Siddhartha Bhattacharyya
Nabendu Chaki · Debanjan Konar
Udit Kr. Chakraborty
Chingtham Tejbanta Singh   *Editors*

Advanced Computational and Communication Paradigms

Proceedings of International Conference on ICACCP 2017, Volume 2

🐎 Springer

Advances in Intelligent Systems and Computing

Volume 706

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

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

Siddhartha Bhattacharyya
Nabendu Chaki · Debanjan Konar
Udit Kr. Chakraborty · Chingtham Tejbanta Singh
Editors

Advanced Computational and Communication Paradigms

Proceedings of International Conference on ICACCP 2017, Volume 2

Editors
Siddhartha Bhattacharyya
Department of Computer Application
RCC Institute of Information Technology
Kolkata, West Bengal
India

Nabendu Chaki
Department of Computer Science
   and Engineering
University of Calcutta
Kolkata, West Bengal
India

Debanjan Konar
Department of Computer Science
   and Engineering
Sikkim Manipal Institute of Technology
Majitar, Rangpo, Sikkim
India

Udit Kr. Chakraborty
Department of Computer Science
   and Engineering
Sikkim Manipal Institute of Technology
Majitar, Rangpo, Sikkim
India

Chingtham Tejbanta Singh
Department of Computer Science
   and Engineering
Sikkim Manipal Institute of Technology
Majitar, Rangpo, Sikkim
India

Siddhartha Bhattacharyya would like to dedicate this volume to his beloved wife Rashni, his cousin sisters Aparna, Archana, Anjana, Mithu, Mita, Piyali, Tanushree, Patralekha, Pamela, Satarupa and Priyanka

Debanjan Konar would like to dedicate this volume to his parents Debidas Konar and Smritikana Konar

Tejbanta Singh Chingtham would like to dedicate this volume to his parents who had shown me hard work and perseverance is the only key in life and to my wife (Lakhi) and my son (Gepel Zayden) for the time I could not spend with you in pursuit of my dreams!

Preface

The Byzantine real-life problems faced by humanity can be catered to by the technological advancements in soft computing approaches. The main aim of such computing methodologies is to provide resilient solutions for these problems. The applications comprise of the different technologies in the fields of computer networking and data communication, cyber security, signal processing, computer vision and image processing, computational perception and cognition, human–computer interaction, adaptive computation and machine learning, and so on. This volume of AISC consists of accepted papers presented at ICACCP 2017, the First International Conference on Advanced Computational and Communication Paradigms 2017. It is aimed at introducing to the aspiring readers the latest trends in advanced computing technologies and communication paradigms. The aforementioned conference was the first of its kind which was hosted by the Department of Computer Science and Engineering, Sikkim Manipal Institute of Technology, Majitar, Rangpo, East Sikkim, Sikkim, and technically collaborated with Computer Society of India (CSI), Kolkata.

The aim of ICACCP 2017 was to address significant areas of research and development in advanced computational and communication paradigms and cyber security, thus providing immense coverage of the advanced computational paradigms and communication techniques which give rise to infallible solutions to the emerging problems faced by humanity and immense potential for future innovations and applications.

ICACCP 2017 received a tremendous response from the academic community. There were about 550 technical paper submissions, and finally after peer-reviewing, 185 high-quality papers were accepted and 142 papers were registered for oral presentation and possible publications in Springer (AISC and LNEE) series. This volume of AISC series comprises 77 papers. In addition, 9 general sessions and 5 special sessions were scheduled in ICACCP 2017.

The organization of the ICACCP 2017 conference was entirely voluntary. The review process required stupendous effort from the members of the International Technical Program Committee, and so they deserve accolades. We would like to express our sincere thanks to the host of ICACCP 2017, Sikkim Manipal Institute of

Technology, and to the publisher, Springer, for their never-ending support in organizing the conference.

Finally, we would like to extend our thanks to all the authors for their superlative contributions. The amicable attitude of conference supporters and contributors has made this event a grand success which will be remembered for long.

Kolkata, India	Siddhartha Bhattacharyya
Kolkata, India	Nabendu Chaki
Majitar, Rangpo, India	Debanjan Konar
Majitar, Rangpo, India	Udit Kr. Chakraborty
Majitar, Rangpo, India	Chingtham Tejbanta Singh
September 2017	

Contents

About the Editors

Dr. Siddhartha Bhattacharyya [FIETE, FIEI, SMIEEE, SMACM, SMIETI, LMCSI, LFOSI, LMISTE, MIET (UK), MIAENG, MIRSS, MIAASSE, MCSTA, MIDES, MISSIP, MSDWIC] did his bachelor's degree in Physics, bachelor's degree in Optics and Optoelectronics, and master's degree in Optics and Optoelectronics from University of Calcutta, India, in 1995, 1998, and 2000, respectively. He completed his Ph.D. in Computer Science and Engineering from Jadavpur University, India, in 2008. He is the recipient of the University Gold Medal from the University of Calcutta for his Masters. He is the recipient of the coveted Adarsh Vidya Saraswati Rashtriya Puraskar in 2016; Distinguished HoD and Distinguished Professor awards conferred by Computer Society of India, Mumbai Chapter, India, in 2017; and Bhartiya Shiksha Ratan Award conferred by Economic Growth Foundation, New Delhi, in 2017. He received the NACF-SCRA, India award for Best Faculty for Research in 2017. He received Honorary Doctorate Award (D. Litt.) from the University of South America in 2017. He is also the recipient of the South East Asia Regional Computing Confederation (SEARCC) International Digital Award ICT Educator of the Year at Colombo, Sri Lanka, in 2017. He is currently the Principal of RCC Institute of Information Technology, Kolkata, India. In addition, he is also serving as the Dean (Research and Development and Academic Affairs) of the institute. He is a co-author of four books and the co-editor of 13 books and has more than 185 research publications in international journals and conference proceedings to his credit. He has got a patent on intelligent colorimeter technology.

Dr. Nabendu Chaki is a Professor in the Department Computer Science and Engineering at University of Calcutta, Kolkata, India. He completed his Ph.D. from Jadavpur University, India, in 2000. He is sharing seven international patents including four US patents. He had been quite active during 2009–2015 toward developing international standards for software engineering and cloud computing as a member of Global Directory (GD) for ISO-IEC. Besides editing nearly 30 book volumes, he has authored seven text and research books and has more than 160 Scopus Indexed research papers in journals and international conferences. His areas

of research interests include distributed systems, image processing, and software engineering. He has served as a Research Faculty in the Ph.D. program in Software Engineering at US Naval Postgraduate School, Monterey, CA. He is a visiting faculty member of many universities in India and abroad. He is the founder Chair of ACM Professional Chapter in Kolkata.

Mr. Debanjan Konar is currently working as Assistant Professor in the Department of Computer Science and Engineering at Sikkim Manipal Institute of Technology. He is also pursuing his Ph.D. from the Department of Electrical Engineering, IIT Delhi. He is currently associated with the quantum-inspired soft computing research community in India and abroad. His important research contributions include quantum computing, rough set theory, natural language processing (NLP). He has published in esteemed international journals and conference. He is recognized reviewer of the *Journal of Applied Soft Computing*. He is a member of IEEE and Computer Society of India.

Mr. Udit Kr. Chakraborty achieved his bachelor's and master's degrees in Computer Science and Engineering from the North Eastern Regional Institute of Science and Technology and Sikkim Manipal University in 2000 and 2009, respectively. At the time of compilation of this volume, he awaits his thesis defense from the University of Calcutta, Kolkata, in the field of intelligent learner computer interaction. He has 14 years of teaching experience prior to which he spent 3 years in the industry working on security systems and telemedicine implementations. His research and teaching interests lie in the areas of soft computing, machine learning, natural language processing, and theory of computation. He has also co-authored a book on soft computing.

Dr. Chingtham Tejbanta Singh, Ph.D. is a Professor in the Department of Computer Science and Engineering. He has 18 years of teaching experience in computer science and engineering, and his areas of interest include robotics, computational intelligence, human–computer interaction, and artificial intelligence. He has a large number of publications both in journals and in conferences; besides, he has also served as a associate editor of many journals and as a technical program committee of many international conferences of high repute. He has also involved in the execution of projects with a grant from AICTE, MHRD, and Microsoft Research India (MSRI). Currently, he is on board of *International Journal of Intelligent Systems and Applications in Robotics*.

A Proposed Artificial Neural Network (ANN) Model Using Geophone Sensors to Detect Elephants Near the Railway Tracks

Rakesh Kumar Mandal and Dechen Doma Bhutia

Abstract Elephant detection system is a subject of interest these days. Expert systems may be designed to enhance the efficiency of these systems. Artificial Neural Networks (ANNs) may be implemented in order to accomplish the task and research may be carried out in this field. In this paper, a method is proposed which detects the presence of elephants near the railway tracks and instantly activates the simulator to drive away elephants from the railway tracks. The simulators may be virtual fire or cracker sound. Elephants are scared of bee sound. So, bee sound may also be used as simulators. The proposed ANN used here is an unsupervised type of ANN, where a weight detection algorithm has been designed to get rid of the ambiguity.

Keywords Artificial neural network · Expert systems · Geophone sensors Simulators

1 Introduction

The human population is increasing at an alarming rate, basic necessities are becoming scarce due to which the habitats of wild animals are being made available for human needs. In India, this is a very common fact that the people have since time immemorial exploited the forests and its produce for food, shelter, and other day-to-day necessities. As urbanization became a necessity, communication could not be left behind and even before aeroplanes or cars, railways became an effective

R. K. Mandal (✉) · D. D. Bhutia
Department of Computer Science & Application, University of North Bengal,
Darjeeling, West Bengal, India
e-mail: rakesh_it2002@yahoo.com

D. D. Bhutia
e-mail: dechendomabhutia2014@gmail.com

form of communication that catered to travel, business, and other human needs. Due to space constraints, the railway tracks had to be laid through forests, so the human and animals encounter became inevitable. The most common were the railway tracks that ran through the elephant corridors and due to less effective infrastructure on the detection of the elephants crossing the railway tracks, the death of these animals has increased at an abnormal rate. Very little work has been done for the effective detection and early warming for these animals. Various methods can be devised using sensors to produce useful data and these sensors in turn would be connected to distant servers which use Artificial Neural Network (ANN) for efficient and early detection of the animals to avoid any fatal accidents.

Some manual work has also been done in the animal detection using various sensors like acoustic, seismic, wireless sensors, etc. [1–3]. Koik and Ibrahim (2012) have done a literature survey on animal detection methods using digital images [4]. Prabhu and Praveen Kumar (2016) have done a review work on recent advances in elephant tracking and detection systems [5]. Sharma and Shah (2013) have done a brief overview on different animal detection methods [6]. Nirmal and Sugumar (2014) have worked on spectral energy magnitude and highest pitch frequency produced by elephants to detect their presence [7]. Sugumar and Jayaparvathy (2013) have worked on a system based on image feature extraction of elephants [8]. Sugumar and Jayaparvathy (2013) have worked on the elephant detection system using migration patterns of the elephants throughout the year [9]. Shaikh et al. (2015) have worked on image mining algorithm to detect animals [10]. Rahayani et al. (2014) have worked on elephant detection system using radio signals [11]. Rangdal and Hanchate (2014) have worked on feature extraction system using Haar of Oriented Gradients. [12].

This paper is divided into three sections. Section 1—discusses about the working of geophone sensors. Section 2—discusses the overall methodology of the proposed system. Section 3—discusses the result analysis.

2 Methodology

In the proposed method, a 20-km-long forest stretch is considered. Four geophone sensors can be installed along the railway track passing through the forest. The distances between the consecutive sensors are kept uniform which is 5 km. The sensors are capable of sensing the presence of elephants within a radius of 15 km. So, the sensing ranges are divided into three regions with radii 15, 10, and 5 km which produce varying ranges of frequencies. The sensors can send signal to a nearby server with the help of a wireless device connected to each sensor. Figure. 1 demonstrates the working of the proposed method.

All the four sensors are connected to two timers installed at the nearby railway stations located at the two ends of the forest. The timer activates the timer when a
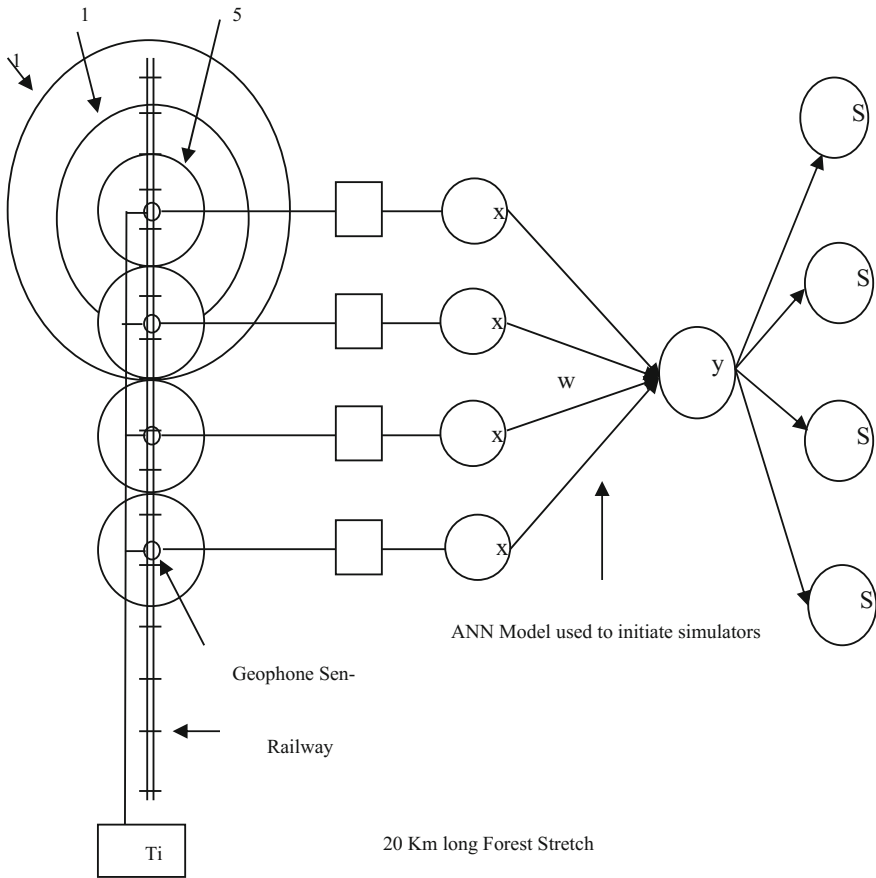
Fig. 1 Proposed elephant detection system

train is about to enter into the stretch from either end of the forest. The railway track may pass through a possible elephant corridor from where herds may cross. Most of the traditional systems are effective on the corridors. Here, the idea is to save single elephant who may try to cross the track from the location which may not be the corridor. The wireless devices connected to the sensors send signals to the server. The server after reading the frequency uses function "f" and produces a binary output vector which is presented to an ANN. The ANN consists of four input neurons "x_1", "x_2", "x_3", and "x_4". The four input neurons are connected to an output neuron "y" using weighted links. The weight vector is "w_i" which is calculated using Algorithm 1.

Algorithm 1: Weight Calculation

```
Step 1: Initialize weight vector wᵢ = {1, 2, 3, 4}
Step 2: Find out the power set of wᵢ.
Step 3: Go to Step 4 if the addition of the
        elements of the subsets of the power set
        generates identical values else go to
        Step7.
Step 4: Set i = 1, j = 3, aᵢ = wᵢ
            c = a[i] + a[i+1]
            a[j] = c + 1
Step 5: Set i = 1, j = 4
            c = a[i] + a[i+1] + a[i+2]
            a[j] = c + 1
Step 6: Set wᵢ = {a[1], a[2], a[3], a[4]}
            Go to Step 2
Step 7: Stop.
```

"y_out" is the net output calculated as follows:

$$y\_out = \sum_i X_i W_i \tag{1}$$

$$y = g(y\_out) \tag{2}$$

The activation "y" is calculated using the function "g", where
$g = \{S_1$ if (y_out == 1) or S_2 if (y_out == 2) or S_3 if (y_out == 4) or S_4 if (y_out == 8) or S_1, S_2 if (y_out == 3) or S_1, S_3 if (y_out == 5) or S_1, S_4 if (y_out == 9) or S_2, S_3 if (y_out == 6) or S_2, S_4 if (y_out == 10) or S_3, S_4 if (y_out == 12) or S_1, S_2, S_3 if (y_out == 7) or S_1, S_2, S_4 if (y_out == 11) or S_2, S_3, S_4 if (y_out == 14) or S_1, S_2, S_3, S_4 if (y_out == 15)\}$.

Using the above method, the movements of the elephants are detected near the geophone sensors. The frequency within the range of 5 km initiates the sensors and sends binary input signals to the input neurons of the ANN. The presence of the elephants produces an input "1" and the absence produces an input "0". The output is calculated and function "g" activates the corresponding simulators, situated at different locations, which either creates virtual fire or humming bee sound to drive away the elephants from the railway tracks.

Table 1 Geophone sensor readings at different radii

S. No.	Radius	Geophone reading (mV)
1	15	6.3–6.6
2	10	7.1–7.7
3	05	8.3–8.8
4	00	10.1–10.15

Table 2 Simulation response due to the presence of elephants

Geophone	Geophone reading (mV)	Active simulator
G-1	8.2	S_1
G-2	7.1	Nil
G-3	6.1	Nil
G-4	8.1	S_4

3   Result Analysis

Some experimental geophone sensor readings are considered to test the efficiency of proposed ANN. The geophone sensor output is measured in millivolts "mV". The output trend for one of the lightest elephant is considered to test the system as given in Table 1.

   Table. 2 demonstrates a case study where two small elephants were found near the railway tracks within 5 km range of geophone sensors G-1 and G-4. The output produced by the other geophone sensors indicates the presence of elephants within 10 and 15 km range. When this data is presented to the proposed ANN, the ANN simulates S_1 and S_4.

4   Discussion

The traditional system emphasizes on presence of elephants near elephant corridors and takes precautionary measures accordingly. The proposed method emphasizes on any possibility of the presence of single Elephant away from the corridors. Most of the traditional methods are manual but the proposed method is automated. The elephants sometimes use dry branches to avoid electric fences and try to cross the railway tracks but in the proposed system, no such possibility is there as electric fences are not used. In traditional methods, manual follow-up of the herds by forest officials is required which is not required in the proposed method.

5   Conclusion

Most of the work done in this field is centered on manual process. The proposed system emphasizes on automation of the system using Artificial Neural Network (ANN). In most of the work, elephant corridors are the main targets to find their presence. The proposed system can find out the presence of elephants at any location. The experiments are successfully done on the ANN developed here. The weight finding algorithm developed in this paper also produces successful results. The future work will be done on the hardware implementation of the system using wireless devices so that it can be used to save the lives of Elephants.

References

1. Kanchan, V.: Survey paper on elephant tracking using acoustic sensor. Int. J. Sci. Dev. Res. (IJSDR) **1**(3), 280–283(2016). ISSN: 2455-2631
2. Prabhu, M.: An efficient surveillance system to detect elephant intrusion into forest borders using seismic sensors. Int. J. Adv. Eng. Technol. **7**(1), 166–171 (2016). E-ISSN 0976-3945
3. Sasikumar, G., Vignesh, R.H., Natheem, M.S.: An analysis on animal tracking system using wireless sensors. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **4**(9), (2014). ISSN: 2277 128X
4. Koik, B.T., Ibrahim.: A literature survey on animal detection methods in digital methods in digital images. Int. J. Future Comput. Commun. **1**(1) (2012)
5. Prabhu, M., Praveen Kumar, G.: Review on recent advances in elephant tracking and detection. Sens. Approach Asian J. Inform. Technol. **15**(6), 1132–1138 (2016)
6. Sharma, S., Shah, D.J.: A brief overview on differential animal detection methods. Int. J. (SIPIJ) **4**(3) (2013)
7. Nirmal, P.J., Sugumar S.J.: Surveillance and tracking of elephants using vocal spectral information. Int. J. Res. Eng. Technol. **3**(7) (2014)
8. Sugumar, S.J., Jayaparvathy, R.: Automated unsupervised elephant image detection system as a solution to human elephant conflict. In: Proceedings of the International Conference on Multimedia Processing. Communication and Information Technology, MPCIT (2013). DOI: 03. AETS
9. Sugumar, S.J., Jayaparvathy R.: An early warning system for elephant intrusion along the forest border areas. Curr. Sci. **104**(11) (2013)
10. Shaikh, S., Jadhav, M., Nehe, N., Verma, U.: Int. J. Adv. Found. Res. Comput. **2** (2015). ISSN 2348-4853
11. Rahayani, R.D. Gunawan, A., Ariwibowo, A.U.: Implementation of radio frequency as elephant presence detector for the human elephant conflict prevention. Innov. Syst. Des. Eng. **5**(5) (2014)
12. Rangdal, M.B., Hanchate, D.B.: Animal detection using histogram oriented gradient. Int. J. Recent Innov. Trends Comput. Commun. **2**(2), 178–183. ISSN: 2321-8169

Local Region with Optimized Boundary Driven Level Set Based Segmentation of Myocardial Ischemic Cardiac MR Images

M. Muthulakshmi and G. Kavitha

Abstract In this work, an attempt is made to segment endocardium and epicardium of left ventricle in normal and myocardial ischemic cardiac magnetic resonance (CMR) images using local region with optimized boundary driven level set. Myocardial ischemia (MI) is a cardiac disorder that results in deprivation of oxygen supply to myocardium and can be analyzed by study of abnormal anatomical changes in CMR. This study is carried out on short-axis view CMR images from Medical Image Computing and Computer-Assisted Intervention (MICCAI) database. The edges are computed by simple Laplacian and Laplacian of Gaussian (LOG) operator. LOG is optimized to obtain enhanced edges of endocardium and epicardium. The quality of edge is validated with edge preservation index (EPI) and gradient magnitude similarity deviation (GMSD) measure. Local region with optimized boundary (LROB) driven level set is utilized for simultaneous segmentation of endocardium and epicardium of left ventricle in CMR images. The results are compared with local region (LR) driven and LR with LOG-driven level set. Further, the efficacy of the segmentation is validated with different similarity measures. The optimized LOG image visually shows better endocardium and epicardium contours. Optimized LOG with a higher EPI and lower GMSD provides better enhanced edges compared to Laplacian and LOG functions. The computed similarity measures for LR with LOG-driven level set are significantly higher compared to LR-based level set for segmentation of endocardium and epicardium. Further, LROB-driven level set shows higher similarity measures than LR with LOG-driven level set. Thus, LROB-driven level set provides better segmentation accuracy for epicardium and endocardium of left ventricle than LR-based level set and LR with LOG-driven level set. The efficiently segmented endocardium and epicardium could aid the diagnosis of myocardial ischemia with their ability to quantify anatomical changes in LV.

M. Muthulakshmi (✉) · G. Kavitha
Department of Electronics Engineering, MIT Campus, Anna University, Chennai, India
e-mail: lakshmingm.2@gmail.com

G. Kavitha
e-mail: kavithag_mit@annauniv.edu

1   Introduction

Worldwide, cardiac disorder is a leading cause of mortality [1]. The cardiac disorders comprise of spectrum of diseases that include myocardial ischemia, myocardial infarction, congenital heart disease, diastolic dysfunction, and contractile dysfunction. The inconsistent oxygen supply to myocardium results in myocardial ischemia (MI). MI patient exhibit impaired ventricular filling, increased myocardial stiffness, delayed relaxation, or irregular contraction [2]. The mortality rate due to MI can be reduced significantly by proper diagnosis and therapeutic interventions at an earlier stage.

Cardiac magnetic resonance (CMR) images are considered to be an effective noninvasive modality for the diagnosis of cardiac disorder. CMR images have lack of ionizing radiations and better soft tissue contrast. Left ventricle (LV) in CMR images is a significant predictor of heart health and progression of disease [3]. Left ventricle anatomy has inter-patient and age variability. Manual delineation of left ventricle is subjective, consumes more time, and prone to interviewer variability [4]. However, automatic segmentation of left ventricle is challenging due to the identical intensities of papillary muscles and myocardium.

In the past decades, image- and model-driven approaches have been used for segmentation of left ventricle [5]. Image-based techniques comprise of pixel intensity [6], region growing [7], prior knowledge [8], and active contour models (ACM) [9]. The efficacy of pixel-based methods may reduce when there are overlapping intensities. Region growing algorithms have a drawback of over segmentation due to inclusion of inappropriate parts. Prior knowledge creation is tedious and requires large set of data samples. Active contour model is most widely used LV segmentation method with its ability to propagate curve by proper design of energy functions with anatomical assumptions as constraints. Model-driven approaches include active shape model [10], Atlas [11], and deep learning methods [12].

Recently, majority of work is based on ACM and level set [13] based ACM. Here, an initial curve moves toward the object boundaries via minimization of energy function. Their characteristics include sub-pixel accuracy in object boundaries, smooth and continuous contours, and robust tool with shape priors. Variational level set is used for segmentation of biological structures in normal brain MR images with intensity inhomogeneity [14]. The epicardium and endocardium of LV are extracted using different variations of level set [15, 16]. The local region based level set method is sensitive to placement of initial contour. The segmentation time is related to location of initial contour. These disadvantages can be overcome by addition of edge information to the energy function of the level set method.

In this work, local region with optimized boundary (LROB) driven level set [17] is used for segmentation of left ventricle in cardiac MR images. The optimized boundary is obtained by optimization of Laplacian of Gaussian (LOG) function that can smooth the uniform regions and enhance the boundary region. The left ventricle and myocardium in CMR have intensity inhomogeneity. Thus, edge enhancement can aid evolve of contours toward the appropriate boundary. The efficiency of this methodology is compared with local region and local region with LOG-driven level set. The enhanced edges and the segmented results are validated with appropriate quantitative measures. The paper is organized as follows: Sect. 2 discusses the database and methodology, Sect. 3 shows the results obtained with different methods and includes the discussion of results.

2 Materials and Methods

2.1 Database

This work utilizes cardiac cine MR images from Medical Image Computing and Computer-Assisted Intervention (MICCAI) left ventricle segmentation database [18]. Cine steady-state free precession (SSFP) MR short-axis (SAX) images are procured from a 1.5T GE Signa MRI. The database encompasses of CMR images of 45 patients. This includes subjects with different conditions such as Normal, Ischemic heart failure, non-Ischemic heart failure, and hypertrophy. Manual ground truth contours are given by experts. The database includes 12 ischemic and 9 normal subjects, among which 4 are females and 17 are male subjects.

2.2 Methodology

Local region and optimized boundary driven level set is used for the simultaneous segmentation of endocardium and epicardium of left ventricle in CMR images. Here, the local region energy is combined with optimized Laplacian of Gaussian energy to evolve the multiphase level set [17].

$$E^{LROB} = E^{LR}(\Phi, f_1, f_2) + E^{OB}(\Phi), \tag{1}$$

where $E^{LR}(\Phi, f_1, f_2)$ is the multiphase local region based energy [14] as given below

$$E^{LR}(\Phi, c, b) = \int \sum_{i=1}^{N} e_i(x) M_i(\Phi(x)) dx \tag{2}$$

In multiphase local region based energy, k-level set functions $\Phi_1, \Phi_2, \ldots, \Phi_k$ are used that are defined by $M_i(\Phi_1(y), \ldots, \Phi_k(y))$ membership functions, b specifies bias field and cluster center is given by c_i. In our case, $k = 2$ and number of region $N = 3$. Here, e_i is the Gaussian weighted energy based intensity descriptor given as

$$e_i = |I - bc_i|^2, \tag{3}$$

where $i = 1$ to N, I denotes the original image. For preset Φ and c, the optimal bias field $\hat{b}$ is given by

$$\hat{b} = \frac{(IJ^{(1)}) * K}{J^{(2)} * K}, \tag{4}$$

where $J^{(1)} = \sum_{i=1}^{N} c_i u_i$ and $J^{(2)} = \sum_{i=1}^{N} c_i^2 u_i$, K is a Gaussian radial basis kernel function. For preset Φ and b, the optimal cluster center $\hat{c}$ is given as

$$\hat{c}_i = \frac{\int (b * K) I M_i(\Phi) dx}{\int (b^2 * K) M_i(\Phi) dx} \text{ for } i = 1, 2, \ldots, N \tag{5}$$

The optimized boundary energy function [20] is given by

$$E^{OB}(\Phi) = \omega L_\Delta(\Phi) + \upsilon L(\Phi) + \mu P(\Phi), \tag{6}$$

where ω, υ, μ are variable coefficients, $L(\Phi)$ is the length term, and $P(\Phi)$ is the distance regularization term. $L_\Delta(\Phi)$ represents the optimized LOG function in LROB-driven level set. The optimized LOG function is obtained by solving the following iterative equation:

$$\frac{\partial L_\Delta(\Phi)}{\partial t} = g(|\nabla I|) \times LG - (1 - g(|\nabla I|)) \times (LG - \beta \times \Delta(G_\sigma * I)), \tag{7}$$

where LG is the LOG term, $g(|\nabla I|)$ is the edge indicator term, β is a positive constant, and G_σ is a Gaussian kernel.

2.3   Validation Measures

The quality of computed edges is analyzed by edge performance index (EPI) [19] and gradient magnitude similarity deviation (GMSD) measures [20]. EPI ranges from 0 to 1. A higher value of EPI indicates better sharpened edges. GMSD reflects the amount of distortion in the processed image compared to the original image. GMSD has a range of 0–1. Lower GMSD signifies good image perceptual quality.

The segmented images are quantitatively validated by similarity measures such as simple matching, Rogers and Tanimoto, Hamann measure, Sokal and Sneath-II and Baroni-Urbani, and Buser-I [21]. All the similarity measures have a range from 0 to 1 except Hamann measure that has range from −1 to +1. In the case of all the similarity measures, a value close to +1 gives a better correlation between segmentation result and ground truth.

3 Results and Discussions

The proposed local region and optimized boundary driven level set is applied for segmentation of endocardium and epicardium of left ventricle (LV) from cardiac magnetic resonance (CMR) images. The images for this study are obtained from MICCAI left ventricle segmentation database. The considered CMR images include 9 healthy and 12 myocardial ischemic subjects. For this analysis, 20 healthy and 20 ischemic images are utilized. The proposed framework is implemented in MATLAB R2012b on an Intel(R) Core (TM) i3 CPU@ 1.7-GHz.

In this work, optimized edges and local region are used to drive the level set evolution. The edge detection is performed with simple Laplacian, LOG, and optimized LOG operators. The optimization of LOG plays a significant role in the proposed method. The optimization of LOG utilizes the following parameters: σ, a, β, Δt, and number of iterations. The parameter σ affects the width and extent of smoothness caused by LOG function, β controls the level of enhancement in edges. After carrying out several trials with the given CMR images, the parameter values for better optimized edges are obtained as $\sigma = 3$, $\beta = 3$, $\Delta t = 0.01$, and number of iterations as 100. The parameter α plays a key role in the optimization process. The computed edge based sharpened images are validated using EPI and GMSD measure. A higher value of EPI and lower value of GMSD indicates better edge quality.

The general range of α is 0.01–1. EPI and GMSD are calculated for different values of α. The parameter α used in optimization of LOG determines the area to be smoothed or sharpened in an image. The effect of smoothness is enhanced by a lower value of α. However, images with weak edges require a higher α value. α is varied from 0.01 to 1 A higher EPI and lower GMSD value for optimized LOG image is obtained with $\alpha = 0.01$ indicating better edge quality. In this proposed method, thus α is chosen as 0.01 for optimized LOG.

Figure 1 shows the edges obtained for end-diastole (ED) and end-systole (ES) images with Laplacian, LOG and optimized LOG for $\alpha = 0.01$. The original ED and ES images for ischemic subjects are shown in Fig. 1a, e. The edges derived from Laplacian, LOG, and optimized LOG function are depicted in Fig. 1b–d, respectively for an ED image. Correspondingly, Fig. 1f–h represents edges obtained from Laplacian, LOG, and optimized LOG function, respectively, for an ES image. Visually, the edges obtained from LOG operator are better than
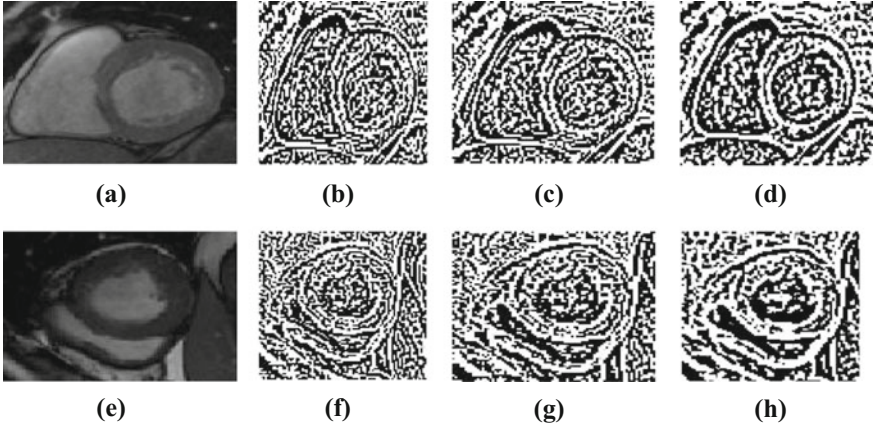
Fig. 1 Comparision of edges obtained from different functions: **a** End-diastole image, **b** laplacian output, **c** LOG output, **d** optimized LOG output, **e** end-systole image, **f** laplacian output. **g** LOG output and **h** optimized LOG output for $\alpha = 0.01$

Laplacian. Further, the edges are still enhanced and boundaries are clearer when the optimized LOG function is applied to the images.

Second-order differential operators are better edge predictors under intensity inhomogeneity conditions. Intensity inhomogeneity causes much zero-crossing detection with LOG. In order to achieve smoothness in homogeneous regions and steepness near object boundaries, optimization of LOG is performed.

The performance of optimized LOG is compared with LOG and Laplacian operator using EPI and GMSD. Figure 2a depicts the EPI obtained for Laplacian, LOG, and optimized LOG. A higher EPI implies a better edge quality. EPI is significantly higher for LOG compared to Laplacian operator for all the images. In majority of cases, EPI is slightly higher for optimized LOG compared to LOG. The GMSD computed for Laplacian, LOG, and optimized LOG is shown in Fig. 2b. A lower GMSD specifies enhanced edges. LOG outperforms Laplacian for all the images with less GMSD. For majority of the considered images, GMSD is slightly lower for optimized LOG compared to LOG. Optimized LOG with higher EPI and lower GMSD indicates a better boundary detection capability compared to other algorithms.

The optimized LOG along with local region is used to evolve the level set for segmentation of LV in CMR images. The experimental parameters used in the level set are as follows: $\sigma = 10$, $\mu = 1$, $\upsilon = 0.002 \times 255^2$, $\omega = 2$, and $\Delta t = 0.1$. The parameter σ governs the weightage and radial distance of the Gaussian kernel, the length of the contour is changed by variation in υ and μ affects the distance regularization term. Parameter ω controls the influence of optimized LOG in the energy function computation and is set as large value for weak edges and vice versa. For positive values of ω, curve evolves along the outer boundary of the object and for negative values, curve evolves along the inner boundary of the object. In
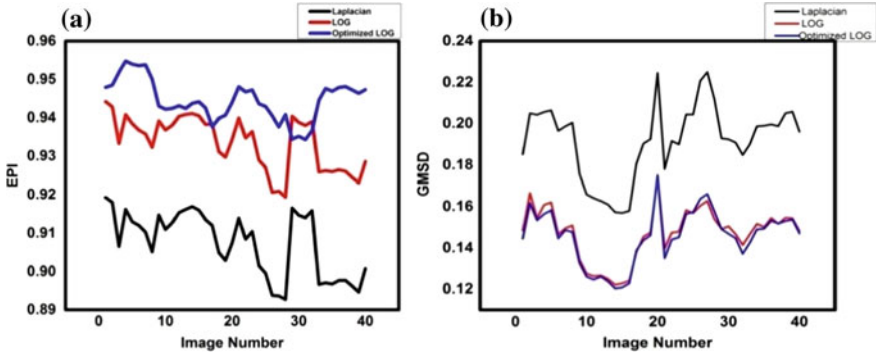
Fig. 2 **a** EPI for Laplacian, LOG and optimized LOG functions. **b** GMSD for Laplacian, LOG and optimized LOG functions

this work, ω is chosen as 2. The segmented boundary of endocardium and epicardium at ED and ES of LV for healthy and ischemic subjects obtained with the proposed local region with optimized boundary driven level set are shown in Fig. 3. The segmentation results of healthy subject at ED, healthy subject at ES, ischemic

(a)   (b)   (c)   (d)   (e)
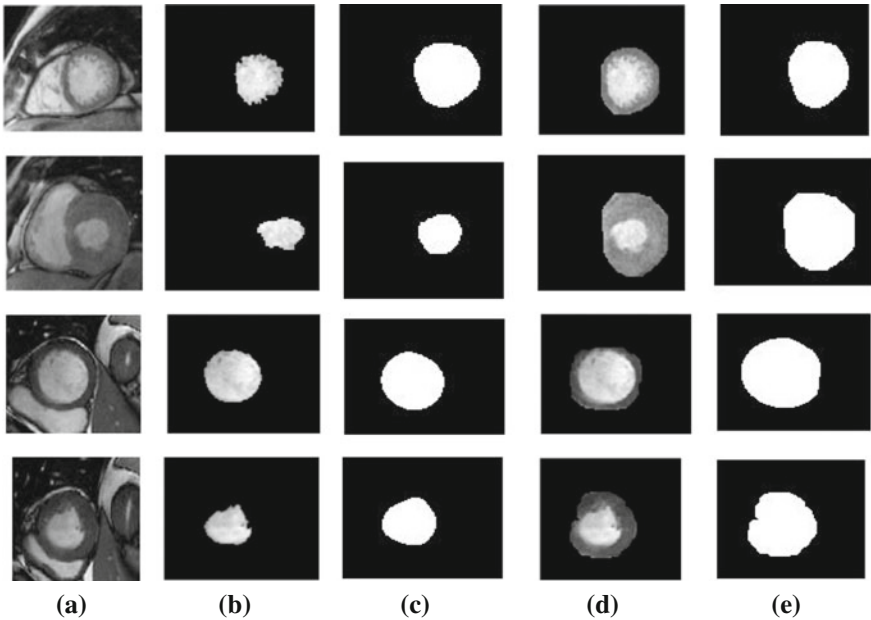
Fig. 3 Segmented endocardium and epicardium of left ventricle. (α = 0.01 at 100 epochs) Row 1: ED images of healthy subject, Row 2: ES images of healthy subject, Row 3: ED images of ischemic subject, Row 4: ES images of ischemic subject. **a** Original image, **b** segmented endocardium, **c** ground truth for endocardium, **d** Segmented epicardium, and **e** ground truth for epicardium

subject at ED, and ischemic subject at ES are shown from the first row to last row in Fig. 3. The images from Fig. 3a–e represent original image, segmented endocardium, ground truth of endocardium, segmented epicardium, and ground truth of epicardium, respectively. Figure 3 qualitatively validates the ability of the proposed method to segment both endocardium and epicardium of LV in CMR images for healthy and ischemic subjects.

Further, quantitative validation of the proposed segmentation results is done with different similarity measures. In this work, similarity measures such as simple matching (SM), Rogers and Tanimoto (RT), Hamann measure (HM), Sokal and Sneath-II (SS-II) and Baroni-Urbani, and Buser-I (BUB-I) are considered [21]. All the similarity measures range from 0 to 1. A value of 1 indicates better similarity between segmented image and ground truth.

Table 1 indicates the mean values of different similarity measures obtained for local region (LR), local region with LOG, and the proposed LROB-driven level set methods. In the case of endocardium segmentation, the mean value of all similarity measures is higher for LR with LOG-driven level set compared to LR-driven level set. However, LROB produces significantly higher mean value for SM, RT, HM, and SS-II measures compared to LR and LR with LOG-driven level sets. Though LR with LOG and LROB-driven level set produces same BUB-I measure, it is higher than LR-driven level set.

All the similarity measures are slightly higher for epicardium segmentation by LROB-driven level set compared to LR and LR with LOG-driven level set. Finally, LROB-driven level set is able to better segment endocardium and epicardium of LV with higher efficacy.

The box plot of various similarity measures obtained from the comparison of endocardium segmentation results with the ground truth is represented in Fig. 4. The similarity measures SS–II, SM, HM, RT and BUB-I ranges from 0.97 to 0.99, 0.95 to 0.99, 0.93 to 0.97, 0.93 to 0.97, and 0.88 to 0.97, respectively.

Figure 5 shows the box plot of various similarity measures obtained for epicardium segmentation. The range of similarity measures for SS–II and SM is from 0.97 to 0.99 and 0.93 to 0.97, respectively. The HM ranges from 0.86 to 0.94,

Table 1 Mean value of various similarity measures for LR, LR with LOG, and LROB-driven level sets for endocardium and epicardium segmentation for normal and abnormal images (LRLS—Local region driven level set, LOGLS—Laplacian of Gaussian driven level set, LROBLS—Local region with optimized boundary driven level set)

Region	Segmentation methods	Mean values of similarity measures				
		SM	RT	HM	SS-II	BUB-I
Endocardium	LRLS	0.9739	0.9495	0.9478	0.9867	0.9271
	LOGLS	0.9740	0.9496	0.9479	0.9868	0.9272
	LROBLS	**0.9756**	**0.9530**	**0.9513**	**0.9876**	**0.9272**
Epicardium	LRLS	0.9481	0.9029	0.8962	0.9731	0.9223
	LOGLS	0.9481	0.9030	0.8963	0.9731	0.9224
	LROBLS	**0.9482**	**0.9031**	**0.8964**	**0.9732**	**0.9225**
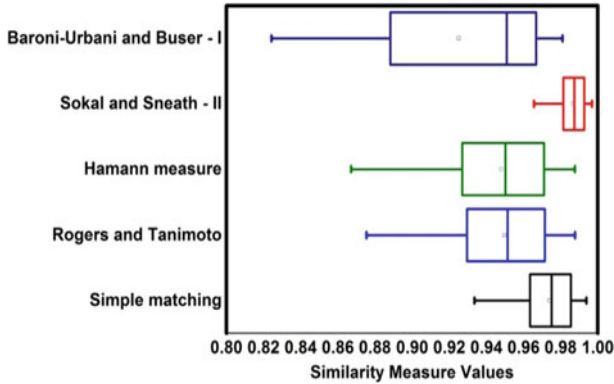
Fig. 4 Box plot representation of various similarity measures for endocardium segmentation
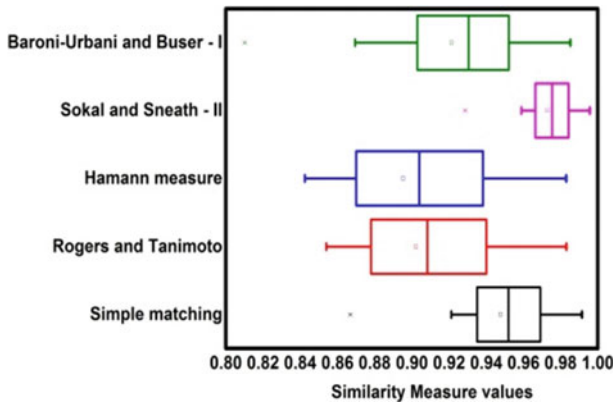
Fig. 5 Box plot representation of various similarity measures for epicardium segmentation

whereas RT measure has a range of 0.87–0.94. BUB-I measure varies from 0.90 to 0.95. A better value for all the similarity measures ensures good quality segmentation by the LROB-driven level set. Some of the reports for segmentation of endocardium and epicardium on the same database are in Table 2. This shows that the LROB-driven level set was able to perform at par with the state of the art techniques.

In the analysis of myocardial ischemia, the clinical indices such as ejection fraction, stroke volume, and myocardial mass of LV play a significant role. These indices are derived from LV volume of ED and ES slices. Thus, only ED and ES slices are considered for this analysis. The segmentation outputs are validated for ED and ES slices of normal and myocardial ischemic subjects. For further analysis of MI tracking of LV volume for the entire cardiac cycle is required. Therefore, the future scope of the work is segmentation of all slices of the normal and abnormal

Table 2 Left ventricle segmentation results for different methods

Segmentation algorithm	Mean dice metric			
	Endocardium		Epicardium	
	Normal subjects	Ischemic subjects	Normal subjects	Ischemic subjects
LROB-driven level set (proposed)	**0.972**	**0.976**	**0.951**	**0.945**
Ngo et al. (2017)	0.910	0.910	0.940	0.940
Avandi et al. (2016)	0.940	0.940	0.940	0.940
Liu et al. (2015)	0.950	0.950	0.920	0.920
Hu et al. (2013)	0.880	0.910	0.930	0.950
Huang et al. (2011)	0.870	0.930	0.920	0.940

subjects. The apex slices have a small blood pool. Hence, challenges will be posed in the segmentation of apex slices compared to basal and mid-cavity slices.

4 Conclusions

In this study, endocardium and epicardium of left ventricle are segmented using local region and optimized boundary driven level set in CMR images. The edges are derived for the images with Laplacian, LOG, and optimized LOG operators. The quality of the enhanced edges is evaluated with the use of EPI and GMSD measures. The simultaneous segmentation of endocardium and epicardium of left ventricle is carried out with LR, LR with LOG, and LROB-driven level set. The segmentation results are validated with different similarity measures. The novelty in the work includes the generation of better boundaries of LV blood pool and epicardium with optimized LOG, segmentation of endocardium and epicardium of LV with local region and optimized boundary driven level set. The segmentation process is robust to initialization of contours and produced better edges.

Optimized LOG produced a higher EPI and lower GMSD values. This proves that optimized LOG gives enhanced edges compared to Laplacian and LOG functions. This could be substantiated by optimized LOG's ability to smooth homogeneous regions and enhance boundaries. The LR with LOG-driven level set resulted in higher similarity measures compared to LR-driven level set for segmentation of endocardium and epicardium. In both the scenarios, LROB-driven level set showed still higher similarity measures than LR with LOG-driven level set. Further, LROB-driven level set has better segmented endocardium compared to epicardium. Thus, LROB-driven level set has shown better segmentation accuracy of epicardium and endocardium of left ventricle than LR-driven level set and LR with optimized LOG-driven level set. The segmented left ventricle can further be used for calculation of cardiac indices which could aid the diagnosis of myocardial

ischemia. This can be a good supplement to radiologists involved in analysis of cardiac pathophysiology.

References

1. http://www.who.int/mediacentre/factsheets/fs317. Accessed Sept 2016)
2. Suinesiaputra, A., McCulloch, A.D., Nash, M.P., Pontre, B., Young, A.A.: Cardiac image modelling: breadth and depth in heart disease. Med. Image Anal. **33**, 38–43 (2016)
3. Dakua, S.P., Sahambi, J.S.: Detection of left ventricular myocardial contours from ischemic cardiac MR images. IETE J. Res. **57**(4), 372–384 (2014)
4. Petitjean, C., Dacher, J.N.: A review of segmentation methods in short axis cardiac MR images. Med. Image Anal. **15**(2), 169–184 (2011)
5. Peng, P., Lekadir, K., Gooya, A., Shao, L., Petersen, S.E., Frangi, A.F.: A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. Magn. Reson. Mater. Phys., Biol. Med. **29**(2), 155–195 (2016)
6. Hu, H., Liu, H., Gao, Z., Huang, L.: Hybrid segmentation of left ventricle in cardiac MRI using Gaussian-mixture model and region restricted dynamic programming. Magn. Reson. Imaging **31**(4), 575–584 (2013)
7. Huang, S., Liu, J., Lee, L.C., Venkatesh, S.K., San Teo, L.L., Au, C., Nowinski, W.L.: An image-based comprehensive approach for automatic segmentation of left ventricle from cardiac short axis cine MR images. J. Digit. Imaging **24**(4), 598–608 (2011)
8. Kaus, M.R., von Berg, J., Weese, J., Niessen, W., Pekar, V.: Automated segmentation of the left ventricle in cardiac MRI. Med. Image Anal. **8**(3), 245–254 (2004)
9. Folkesson, J., Samset, E., Kwong, R.Y., Westin, C.F.: Unifying statistical classification and geodesic active regions for segmentation of cardiac MRI. IEEE Trans. Inf Technol. Biomed. **12**(3), 328–334 (2008)
10. Leiner, B.J., Olveres, J., Ramírez, B.E., Arámbula, F., Vallejo, E.: Segmentation of 4D cardiac computed tomography images using active shape models. In: SPIE 8436 Optics Photonics and Digital Technologies for Multimedia Applications II 84361E (2012)
11. Bai, W., Shi, W., Ledig, C., Rueckert, D.: Multi-atlas segmentation with augmented features for cardiac MR images. Med. Image Anal. **19**(1), 98–109 (2015)
12. Avendi, M.R., Kheradvar, A., Jafarkhani, H.: Combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Med. Image Anal. **30**, 108–119 (2016)
13. Ammar, M., Mahmoudi, S., Chikh, M.A., Abbou, A.: Endocardial border detection in cardiac magnetic resonance images using level set method. J. Digit. Imaging **25**, 294–306 (2012)
14. Li, C., Huang, R., Ding, Z., Gatenby, C., Metaxas, D., Gore, J.: A variational level set approach to segmentation and bias correction of images with intensity inhomogeneity. Med. Image Comput. Comput. Assist. Interv. **11**(Pt 2), 1083–1091 (2008)
15. Liu, Y., Captur, G., Moon, J.C., Guo, S., Yang, X., Zhang, S., Li, C.: Distance regularized two level sets for segmentation of left and right ventricles from cine-MRI. Magn. Reson. Imaging **34**(5), 699–706 (2016)
16. Ngo, T.A., Lu, Z., Carneiro, G.: Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Med. Image Anal. **35**, 159–171 (2017)
17. Ding, K., Xiao, L., Weng, G.: Active contours driven by region-scalable fitting and optimized Laplacian of Gaussian energy for image segmentation. Sig. Process. **134**, 224–233 (2017)
18. Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A.J., Wright, G.A.: Evaluation framework for algorithms segmenting short axis cardiac MRI. MIDAS J. Card. MR Left Ventricle Segmentation Challenge. http://hdl.handle.net/10380/3070

19. Kuppusamy, P.G., Joseph, J., Sivaraman, J.A.: A full reference Morphological Edge Similarity Index to account processing induced edge artefacts in magnetic resonance images. Biocybern. Biomed. Eng. **37**, 159–166 (2017)
20. Xue, W., Zhang, L., Mou, X., Bovik, A.C.: Gradient magnitude similarity deviation: a highly efficient perceptual image quality index. IEEE Trans. Image Process. **23**(2), 684–695 (2014)
21. Raja, N.S.M., Kavitha, G., Ramakrishnan, S.: Analysis of vasculature in human retinal images using particle swarm optimization based tsallis multi-level thresholding and similarity measures. In: International Conference on Swarm Evolutionary and Memetic Computing— SEMCCO: Swarm Evolutionary and Memetic Computing, pp. 380–387 (2012)

An Innovative Approach for Automatic Genre-Based Fine Art Painting Classification

Alexis Paul and C. Malathy

Abstract Recent advances in digital image processing, computer vision, etc., have led to many approaches to the classification of fine art painting. The focus was made to develop an automatic painting classification system based on their genre. The expanding database of digital painting images makes it imperative to develop an automated method to annotate paintings with metadata as painter, genre, painting tool used, style, etc., so that problems like image retrieval, searching, organizing, and artistic recommendations become convenient and efficient. The aim was to classify painting database into five genres. The database consisted of 1229 digital painting images. The method adopted for the task was feature extraction from the images, using each feature individually for classification and then combining the features based on weighted majority voting to form an ensemble classifier. It was observed that Local Binary Pattern (LBP) was the best performing feature. The ensembled model achieved an accuracy of 80.41%. We have included analysis of our work and have discussed the performance of the various features deployed for the painting classification task.

Keywords Art painting genre · Painting classification · Feature extraction
SVM classification · Ensembling · Weighted majority voting

1 Introduction

Recently, art has digitalized itself in the form of virtual galleries like WikiPaintings [1], Caterpillar Labs [2], Olga's gallery [3], Pintura [4] which has made it all the easier and convenient for artists, art lovers, and others who are interested in art to buy, sell, and appreciate the works of artists from all over the world. Fine art

A. Paul (✉) · C. Malathy
SRM University, Chennai 603203, India
e-mail: alexispaul_no@srmuniv.edu.in

C. Malathy
e-mail: malathy.c@ktr.srmuniv.ac.in

paintings can be broadly classified based on genre and style. This paper considers genre-based painting classification as a task for applying machine learning.

Painting classification is a difficult human task and traditionally only human experts were entrusted this task. Some elements of paintings which help in classification are: color mixing, brush stroke, line styles, gradient, objects, etc. Using image processing techniques, features are extracted from images and the extracted features are used for classification. Articles [5–7] have used this approach for the aforementioned task.

In this work, we have followed a similar approach, a set of features were extracted from the painting images and then based on the classification accuracy obtained from each of the individual features, and the features were combined using weighted majority voting to form an ensemble model. The final accuracy reported is that of the ensembled model. Automated art classification can find potential use in content-based image retrieval, automatic artistic recommendations, etc.

In the next section, we have highlighted some of the past related works followed by our approach to the addressed problem. We have also outlined how we have collected our dataset, features chosen for the classification task, and classification algorithm deployed. Finally, we conclude by presenting an analysis of the results obtained and scope for future works. Figure 1 depicts some of the paintings in our database.

(a) Cityscape   (b) Flower   (c) Calligraphy   (d) Portrait

(e) Still Life

Fig. 1 Some paintings used as training examples, one from each genre

2 Related Work

With the evolution of computer vision and machine learning, recently there has been a surge in interest toward using machine learning for fine art related problems. However, research done in classifying paintings based on genre is sparse. Mentioned below are few works that focus on the classification of paintings by genre.

Condorovici et al. [5] describes the classification of six genres: Baroque, Renaissance, Rococo, Romanticism, Impressionism, and Cubism. The features used in this work luminance, edge, and color. The database consisted of 5769 paintings and various classifiers like Linear SVM, Nonlinear SVM, Real AdaBoost, Modest AdaBoost, and Multilayer Perceptron was used and an accuracy of 49% was achieved with Multilayer Perceptron classifier for classifying the paintings into their respective genres.

Agarwal et al. [6] describe the classification of six genres: Abstract Paintings, Interiors, Landscapes, Portraits, Sculpture, Wildlife and ten styles: Abstract Expressionism, Baroque, Cubism, Impressionism, Expressionism, Pop Art, Rococo, Realism, Renaissance, Surrealism. The features used in this work were SIFT, GIST, HOG combined with LBP, GLCM, and color. The genre database consisted of 1800 paintings and the style database consisted of 3000 paintings. Using SVM with the chi-squared kernel, an accuracy of 84.56% was achieved for the genre database and an accuracy of 62.37% was achieved for the style database.

Cetinic et al. [7] describe the classification of seven genres: Portrait, Still Life, Cityscape, Landscape, Nude, Flower, and Animal. The features used in this work were CNN-derived features, SIFT, GIST, HOG, GLCM, and HSV color histograms. The database consisted of 7000 paintings. Using SVM with the RBF kernel an accuracy of 77.57% was achieved.

3 Methodology

The goal was to classify a set of paintings based on their genres. This section first describes how we collected the training and the testing data followed by the features selected and the process of extracting the mentioned features from the paintings. Finally, we describe the deployed classification algorithm and the ensembling method used to obtain the final results.

3.1 Data Collection

We collected the painting images from https://www.wikiart.org/ [1]. The images which we downloaded belonged to these genres: Cityscape (CS), Flower Painting (FP), Calligraphy (CLG), Portrait (POR), and Still Life (SL). The paintings images

Table 1 Number of images in database

Genre	Training set	Testing set
CS	240	60
FP	129	32
CLG	135	33
POR	240	60
SL	240	60
	Total = 984	Total = 245

were of varying sizes which could be an advantage for us as it will add robustness to our built classifier.

Table 1 shows the number of images for each genre in our training and testing set. The ratio of training set size to testing set size was 80:20.

3.2 Feature Selection and Extraction

In this section, we name and describe the features selected for training and testing the classifier. The features that we considered for the classification task were: Dense SIFT [8], Dense HOG2X2 [9, 10], Local Binary Patterns (LBP) [11], Gradient Local Auto-Correlation (GLAC) [12], Color Naming [13, 14], and GIST [11]. We highlight the feature extraction process below [15]:

- **Dense SIFT**: SIFT is known for its object detection capability and has shown impressive performance when applied for the task of painting classification [6, 7]. We extract SIFT [8] descriptor in a dense manner (i.e., on a grid) at multiple patch sizes, and then apply the bag-of-words + spatial pyramid pipeline explained below. With a dictionary size of 400 visual words, the feature vector obtained was 1×8400 dimensional.
- **Dense HOG2X2**: HOG is also known for object detection in images. We extract HOG [10] in a dense manner (i.e., on a grid) [9] and concatenate 2×2 cells to obtain a descriptor at each grid location. Then apply the bag-of-words + spatial pyramid pipeline explained below. With a dictionary size of 400 visual words, the feature vector obtained was 1×8400 dimensional.
- **LBP**: LBP is known for capturing texture in images and has shown impressive performance in [6]. We extract nonuniform Local Binary Pattern [11] descriptor (neighborhood: 8, transitions: 2), and concatenate 3 levels of spatial pyramid to obtain final feature vector. The feature vector obtained was 1×1239 dimensional.
- **GLAC**: GLAC features utilizes 2nd order statistics, i.e., spatial and orientational autocorrelations of local gradients. It enables us to extract richer information from images and to obtain more discriminative power than standard histogram based methods. The image gradients are sparsely described in terms of magnitude and orientation. From a geometrical viewpoint, the method extracts

information about not only the gradients but also the curvatures of the image surface. GLAC features were extracted on spatial bins using MATLAB implementation [16]. The feature vector obtained was 1 × 3996 dimensional.

- **Color Naming**: In order to reduce the effect of photometric variance we convert the image to 12 color names [13, 14] and extract dense overlapping patches of multiple sizes in the form of a histogram of color names. Then apply the bag-of-words + spatial pyramid pipeline explained below. With a dictionary size of 400 visual words, the feature vector obtained was 1 X 8400 dimensional.
- **GIST**: GIST is useful for capturing the spatial character or shape of images. GIST descriptor describing the spatial envelope of the image [17] was computed. The feature vector obtained was 1 × 512 dimensional.

Bag-of-words pipeline: using a random sampling of the extracted features from various patches, learn a dictionary using k-means [18], and apply locality-constrained linear coding (LLC) [19] to soft-encode each patch to some dictionary entries. Then, as shown in [19], we apply max pooling with a spatial pyramid [20] to obtain the final feature vector. We use LLC as it allows the use of a linear classifier for classification instead of using nonlinear kernels. For the features, other than GLAC, MATLAB toolbox [21] was used.

3.3 Classification Algorithm and Ensembling

The classification algorithm deployed was Support Vector Machines (SVM) [22] with the chi-squared kernel implementation [13, 23]. Each feature individually was used for the classification task. Various kernels like RBF, linear, sigmoid and chi-squared were tested and it was found that chi-squared and linear kernel were giving the best results. Specifically, for LBP, GLAC, Color Naming, GIST features, the chi-squared kernel was chosen because of its better performance over linear kernel. Whereas for SIFT and HOG features, the linear kernel was chosen. The reported accuracy for these features was all 10-fold cross-validated.

We obtained six classification accuracies for each of the six features (Fig. 2). Using a weighted majority voting algorithm, these features were combined to form an ensemble model. The weight assigned to each feature was their 10-fold cross-validated classification accuracy. Thus, ensuring an unbiased weight assignment, we obtained the final accuracy for our classifier.

4 Analysis of Results

The overall accuracy achieved using the ensembled model was 80.41%. Figure 2 shows the accuracy obtained using each feature and the ensemble model, Table 2 shows the confusion matrix obtained for the ensemble model.
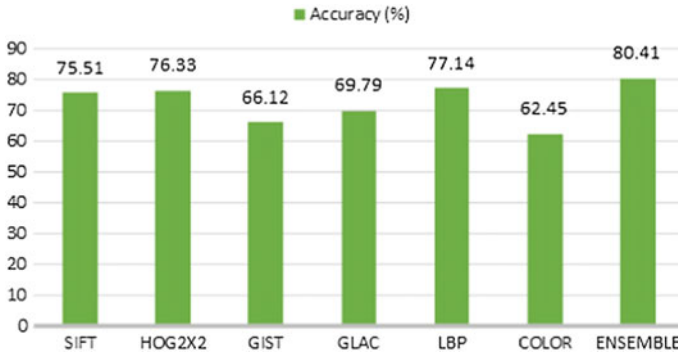
Fig. 2 Overall accuracy on dataset

Table 2 Confusion matrix for the ensemble model

Type of painting	CS	FP	POR	SL	CLG
CS	**0.87**	0.00	0.05	0.08	0.00
FP	0.00	**0.72**	0.03	0.22	0.03
POR	0.03	0.03	**0.82**	0.12	0.00
SL	0.13	0.05	0.07	**0.68**	0.07
CLG	0.00	0.03	0.00	0.00	**0.97**

After a careful analysis of the confusion matrix obtained for each feature, we could conclude about the performance of the each of them. Table 3 shows the best accuracy achieved for each genre and the corresponding feature involved when performance of individual features was considered.

LBP emerged as the best feature and it classified the flower paintings, still life paintings and calligraphy paintings with the highest accuracy. Calligraphy and flower paintings (in some cases) had a well-defined texture which was successfully captured by LBP. Still life paintings in our database consisted of a mixture of objects and defined texture, due to this SIFT which is well known detecting object performed as good as LBP in classifying still life paintings. SIFT was also able to capture the buildings in the Cityscape paintings with good accuracy.

From the confusion matrix in Table 2, it can be inferred that the classifier was most confused between the two genres: Flower Painting (FP) and Still Life Painting (SL). A glimpse into the FP and SL database reveals that there were lots of

Table 3 Genre versus Feature

Genre	Best feature	Accuracy (%)
CS	Dense SIFT	86.67
FP	LBP	68.75
POR	Color naming	90.00
SL	LBP	68.33
CLG	LBP	93.94

Fig. 3 **a** Flower painting misclassified as Still Life and, **b** Still Life painting misclassified as cityscape
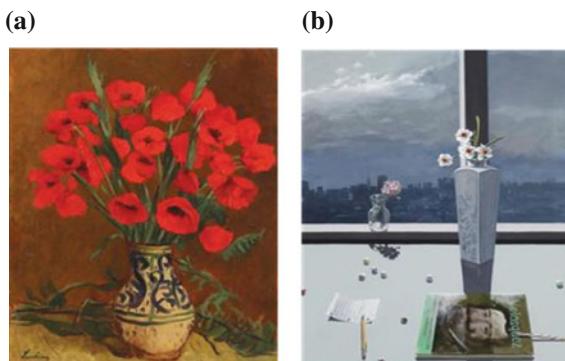
paintings in SL database depicting flowers mixed up with fruits, other objects in some cases. Due to this many FP paintings were misclassified as SL paintings. And strangely enough, few SL paintings were misclassified as CS paintings. Figure 3 shows an instance of misclassification.

Another interesting observation is that the color naming feature was classifying portrait paintings with high accuracy of 90%. This is indicative of the fact that portrait paintings can easily be segmented from the background and thus due to a high degree of similarity between portrait paintings, their classification was yielding. Also, the portrait paintings consisted of only a few number of distinguishable colors and thus converting them into color names consisting of 12 colors was fruitful and the results obtained were promising.

5 Comparison with Previous Work

In this section, we provide an in-depth detailed comparison with the latest work done in this field. Cetinic et al. [7] although have experimented with seven genres, we have four genres in common and few features in common as well. But the feature extraction process differs and this has led to significant difference in results obtained for the individual features. SIFT, HOG features computed in [7] used a very high dimensional feature vector with a K-means vocabulary of 1000 words (for SIFT) but still the classification accuracy achieved was 48.28% for SIFT and 42.42% for HOG, we have demonstrated a different method for SIFT, HOG feature extraction and we report a much better accuracy even with a smaller dictionary size of 400 visual words.

For color features, HSV (Hue, Saturation, Value) histograms were used in [7] and an unacceptable accuracy of 19.07% was achieved. We have demonstrated the use of color naming feature which has been tested for the first time for the task of painting classification. Just by using 12 color names and dictionary size of 400 visual words, we achieved an accuracy of 62.45% which was impressive and

outperformed the color feature used in [6, 7]. We would like to highlight the ensembling technique used in [6] which was boosting and it boosted the accuracy to around 2% more than the accuracy achieved by SIFT which was the best performing feature in [6]. The ensembling approach used by us was weighted majority voting which has shown a better performance by increasing the accuracy by more than 3% and is also computationally very cheap as compared to boosting.

6   Future Works

Although we achieved a good accuracy of 80.41%, there is still a lot of scope of improvement, for example, painting images could be preprocessed using image processing techniques before feature extraction. Different features could be tested on a larger database. Ensembling methods like boosting and stacking could be used for forming the ensemble model. Finally, these set of features could be tested on a database consisting of different painting genres.

7   Conclusion

We have presented a modest scheme for genre-based painting classification. Our classification system comprised of feature extraction from a set of images and then combining the features using weighted majority voting to form an ensemble model. An accuracy of 80.41% was achieved using the ensembled model. By comparing the different features, it was found that LBP was performing better than the rest. The results obtained are better than the method proposed in [7] though in [7] a larger database was experimented with and even Convolutional Neural Network (CNN) based features were used also the number of genre were more than ours. But we had four genres common which made us kind of conclude that our work is better than theirs. The method employed in [6] achieves an accuracy of 84.56% for the task of genre-based painting classification but the genre they had selected could easily be distinguished so that their classification was easier.

References

1. WikiPaintings: https://www.wikiart.org/
2. Caterpillar labs: http://www.caterpillarlabs.com/
3. Olga's gallery: http://www.abcgallery.com/
4. Pintura: http://pintura.aut.org/
5. Condorovici, R.G., Vertan, C., Florea, L.: Artistic genre classification for digitized painting collections. UPB Sci. Bull. Series C **75** (2013)

6. Agarwal, S., et al.: Genre and style based painting classification. In: 2015 IEEE Winter Conference on Applications of Computer Vision. IEEE (2015)
7. Cetinic, E., Grgic, S.: Genre classification of paintings. In: ELMAR, 2016 International Symposium. IEEE (2016)
8. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
9. Russell, B.C., et al.: LabelMe: a database and web-based tool for image annotation. Int. J. Comput. Vis. **77**(1–3), 157–173 (2008)
10. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1. IEEE (2005)
11. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–987 (2002)
12. Kobayashi, T., Otsu, N.: Image feature extraction using gradient local auto-correlations. In: European Conference on Computer Vision. Springer, Berlin, Heidelberg (2008)
13. Van de Weijer, J., Schmid, C., Verbeek, J.: Learning color names from real-world images. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition. IEEE (2007)
14. Khan, R., et al.: Discriminative color descriptors. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2013)
15. https://github.com/adikhosla/feature-extraction/blob/master/FEATURES.md
16. https://staff.aist.go.jp/takumi.kobayashi/code/glac.zip
17. Oliva, A., Torralba, A.: Modeling the shape of the scene: a holistic representation of the spatial envelope. Int. J. Comput. Vis. **42**(3), 145–175 (2001)
18. Elkan, C.: Using the triangle inequality to accelerate k-means. In: ICML, vol. 3 (2003)
19. Wang, J., et al.: Locality-constrained linear coding for image classification. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2010)
20. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 2. IEEE (2006)
21. https://in.mathworks.com/matlabcentral/fileexchange/50448-computer-vision-feature-extraction-toolbox?requestedDomain=www.mathworks.com
22. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) (2)3, 27 (2011)
23. Libsvm chi-square kernel: http://wmii.uwm.edu.pl/˜ksopyla/projects/libsvm-with-chi-squared-kernel/

Asymmetric Cryptosystem Using Affine Transform in Fourier Domain

Savita Anjana, Indu Saini, Phool Singh and A. K. Yadav

Abstract An improved image encryption scheme that uses affine transform and asymmetric keys in Fourier transform domain has been proposed. The scheme is validated for grayscale images through numerical simulation using MATLAB 7.14. In the proposed scheme, the decryption keys are different from the encryption keys and are obtained by phase truncation Fourier transform method. The performance of the scheme is evaluated in terms of the metrics such as correlation coefficient, mean-squared error, and peak signal-to-noise ratio. We have carried out the sensitivity analysis relative to the affine transform parameter, which serves as an additional security feature. The robustness of the scheme is demonstrated by showing its resistance against noise and occlusion attacks. Since this is the first study that uses affine transform in the phase-truncated Fourier transform based asymmetric cryptosystem, it provides a new scheme for image encryption with enhanced security.

Keywords Affine transform · Asymmetric cryptosystem · Fourier transform

1 Introduction

In the recent years, there have been tremendous research activities going on in the field of information security due to fast adoption of digital technology. These efforts are aimed at improving the existing techniques as well as developing new schemes which are efficient and secure. During the past decade, various techniques [1] have been developed for secure communication and transmission of the data. The schemes are broadly classified as purely optical, purely digital, and hybrid techniques, each having its own advantages. Most of the recent optical image

S. Anjana · I. Saini · P. Singh (✉)
Department of Applied Sciences, The NorthCap University, Gurugram 122017, India
e-mail: Phool.singh24@gmail.com

A. K. Yadav
Department of Applied Mathematics, Amity University Haryana, Gurugram 122413, India

encryption studies emerged from the work of Refregier and Javidi [2]. This work was based on double random phase encoding (DRPE). In order to transmit an image over a network securely, two kinds of cryptosystems can be used: symmetric and asymmetric. The classical DRPE scheme [3–6] used two random phase masks as encryption keys and the same is used for decryption thereby rendering the scheme as symmetric.

It is well reported that symmetric cryptosystems based on DRPE are vulnerable to various attacks [7, 8]. In symmetric systems, the same key (the secret key) is used to encrypt and decrypt the data, whereas in asymmetric cryptosystems, a pair of different keys, one for encryption known as the encryption key and the other for decryption known as the decryption key is used [9].

The concept of asymmetric keys based on phase-truncated Fourier transform (PTFT) was used by Qin and Peng [9], followed by many others [10–13] with the aim of enhancing security against the known attacks. Thus, in the present study, a new encryption scheme is proposed in which we have used two random phase masks as encryption keys whereas the decryption keys are obtained by PTFT method. The proposed asymmetric scheme has used affine transform to scramble the pixels of the input image. The use of affine transform adds another parameter for higher security. Thus, it improves the degree of security provided by the earlier asymmetric cryptosystems. The present scheme is validated for grayscale images.

2 The Principle

2.1 Affine Transform

The affine transform [14] (AFT) of an image is a linear map defined from a given set of pixels (x, y) to another set of pixels (x', y'). It is used to move pixels to new locations, thus providing scrambling of pixels. Generally, it is used to change the original position of a pixel by using a pair of transformation equations

$$x' = P_x(x, y); \; y' = P_y(x, y), \tag{1}$$

where $P_x(x, y)$ and $P_y(x, y)$ are polynomials in x and y. In matrix form, it can be represented as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} p + qx \\ r + sy \end{bmatrix} (mod \, N). \tag{2}$$

Here, 'mod' represents the modulo operation. Further, p, q, r and s are any numbers chosen between 1 and N such that $gcd(q, s)$ equals 1. This condition is necessary for the mapping to be one to one. The number of iterations (w) in the affine transform will serve as the encryption parameter.
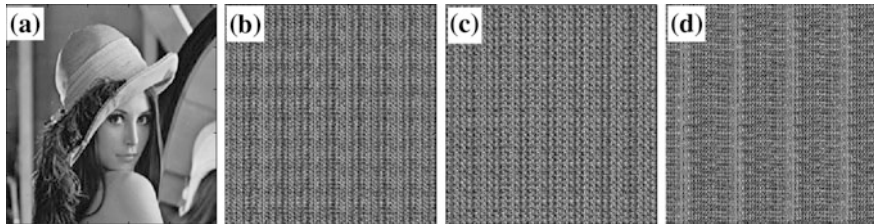
Fig. 1 Affine Transform (*AFT*) of an image, **a** Original Image of Lena; **b–d** *AFT* of image of Lena with $w = 5$, 15, and 30 respectively

The inverse affine transform (IAFT) is given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (x' - p)q^{-1} \\ (y' - r)s^{-1} \end{bmatrix} (mod\,N) \tag{3}$$

Here, q^{-1} and s^{-1} are modulo inverse of q and s, respectively. The effect of affine transform for different values of parameter (w) is shown in Fig. 1 for a grayscale image of Lena of size 256×256 pixels.

2.2 Fourier Transform

The Fourier transform of a function $f(x, y)$ is actually its frequency domain representation. Mathematically, it is represented as

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(ux + vy)} dx\, dy \tag{4}$$

and the inverse Fourier transform is given by

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{i2\pi(ux + vy)} du\, dv. \tag{5}$$

3 The Proposed Scheme

As shown in the flowchart (Fig. 2) of the present scheme, an input image $f(x, y)$ is first subjected to an affine transform. This transformed image is bonded with a random phase mask RPM1 $= \exp\{2\pi im(x, y)\}$, where $m(x, y)$ is randomly distributed in (0, 1). Thereafter, we apply Fourier transform (FT) followed by phase

(a)
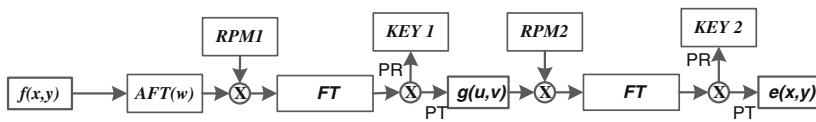
(b)
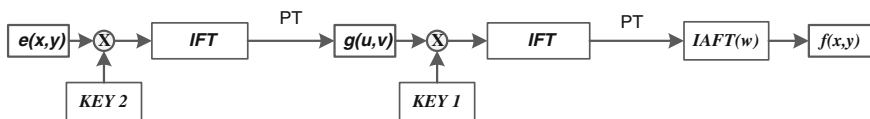
Fig. 2 Flowchart of the proposed scheme, **a** encryption process, **b** decryption process

truncation (PT), which means only the amplitude part will be retrieved. The phase reserved part (PR) will serve as one of the decryption keys (KEY1). The amplitude part bonded with another random phase mask $RPM2 = \exp\{2\pi i n(u, v)\}$ in the frequency domain is Fourier transformed resulting in encrypted image $e(x, y)$ after phase truncation. The phase reserved part (PR) will serve as the second decryption key (KEY2). Mathematically, the encryption process can be shown as

$$g(u, v) = PT[FT[AFT\{f(x, y)\} * RPM1]] \tag{6}$$

$$e(x, y) = PT[FT[g(u, v) * RPM2]] \tag{7}$$

Here, FT represents Fourier transform and PT represents phase truncation.

Similarly, the decryption process can be shown as

$$f(x, y) = IAFT[PT[IFT[PT\{IFT(e(x, y) * KEY2)\} * KEY1]]], \tag{8}$$

where *IFT* represents inverse Fourier transform.

4 Results and Discussion

To demonstrate the validity of the proposed scheme, a grayscale image of Lena (Fig. 3a) of size 256 × 256 pixels is used. The simulation results are obtained using MATLAB 7.14. In the computations, value of the affine parameter is set to $w = 15$. The input image is encrypted using the scheme detailed in Fig. 2a. For decryption, the process used is shown in Fig. 2b. The quality of decryption can be evaluated by the values of correlation coefficient, mean-squared error and peak signal-to-noise ratio for the input-decrypted image pair. The validation of the present scheme is shown in Fig. 3. The encrypted image is provided in Fig. 3b which is completely random and resembles stationary white noise. Figure 3c shows the decrypted image
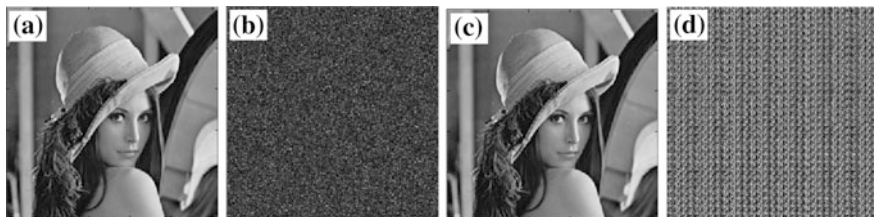
Fig. 3 Results of validation of the proposed scheme. **a** Input grayscale image (Lena) 256 × 256 pixels; **b** encrypted image; **c** recovered image; **d** recovered image with incorrect AFT parameter ($w = 16$, whereas correct value is $w = 15$)
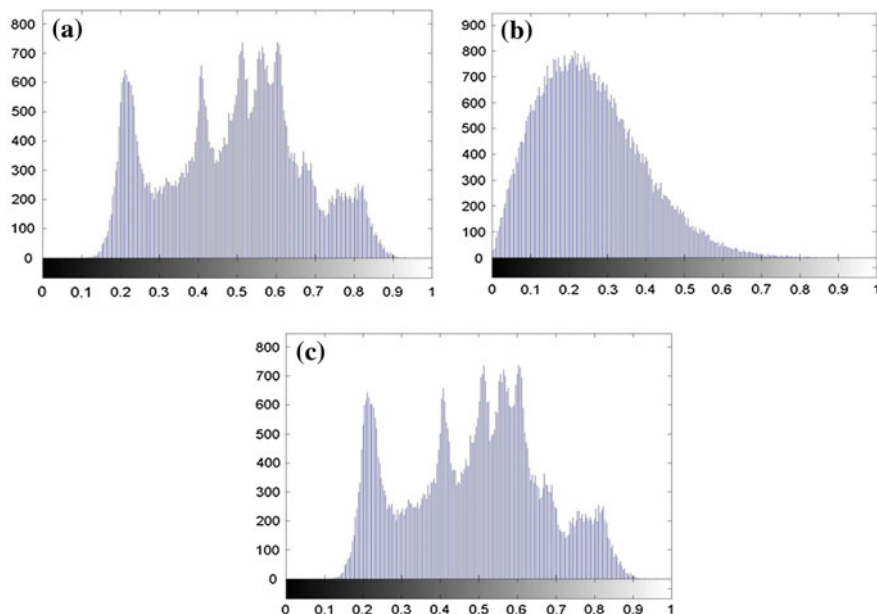
Fig. 4 Histograms of grayscale image of Lena **a** input, **b** encrypted, **c** decrypted

with $CC = 1.00, MSE = 4.5211 \times 10^{-27}$, and $PSNR = 311.5784$. The decryption is quite sensitive to the affine parameter (w), as shown in the decrypted image (Fig. 3d) for wrong parameter.

The efficacy of image encryption by the present scheme is examined by analyzing the histograms (Fig. 4) and 3D plots (Fig. 5) of the input, encrypted and decrypted images. It is observed that the histogram of the encrypted image (Fig. 4b) is uniformly distributed over the normalized range of pixels unlike that of the input and decrypted images which show identical distribution. A similar behavior is seen through the 3D plots in Fig. 5.

For sensitivity analysis with respect to the encryption parameter (w), correlation coefficient (CC), mean-squared-error (MSE), and peak signal-to-noise ratio ($PSNR$) are calculated using the formulae
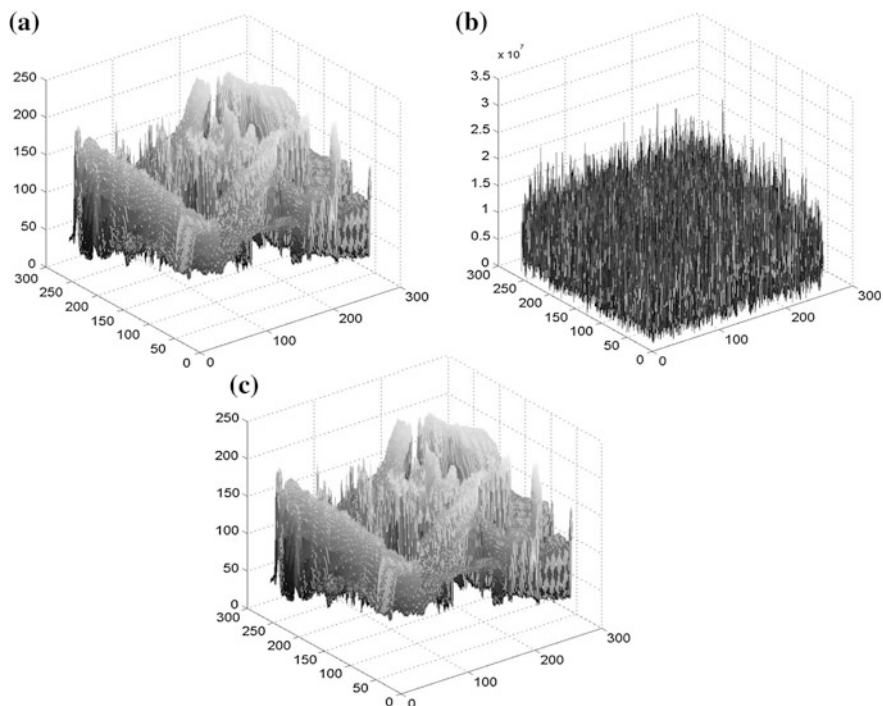
Fig. 5 3D plots of grayscale image of Lena **a** input, **b** encrypted, **c** decrypted

$$CC = cov(I_o, I_r)/\sigma(I_o)\sigma(I_r) \tag{9}$$

$$MSE = \frac{1}{N \times N} \sum_{x=1}^{N} \sum_{y=1}^{N} |I_o(x,y) - I_r(x,y)|^2 \tag{10}$$

$$PSNR = 10 * \log \frac{255^2}{\left[\frac{1}{N \times N} \sum_{x=1}^{N} \sum_{y=1}^{N} |I_o(x,y) - I_r(x,y)|^2\right]}, \tag{11}$$

where $I_o(x,y)$ and $I_r(x,y)$ denote the pixel values of the input image and the recovered/decrypted image, respectively. Here, *cov* is covariance and σ is standard deviation. Figure 6 shows the plots of *MSE* and *PSNR* relative to the affine transform parameter (*w*) values. The curves reveal that the present scheme is highly sensitive to the AFT parameter (*w*).

Most of the symmetric cryptosystems are known to suffer from their vulnerability to several known attacks. The proposed scheme, being asymmetric, is expected to resist such attacks. The resistance to noise attack is tested by adding salt and pepper noise and Gaussian noise of varying densities/standard deviations to the encrypted image and then analyzing the corresponding decrypted image. Figure 7 shows the results of the noise attack. Figure 7g–h shows a gradual decline in CC
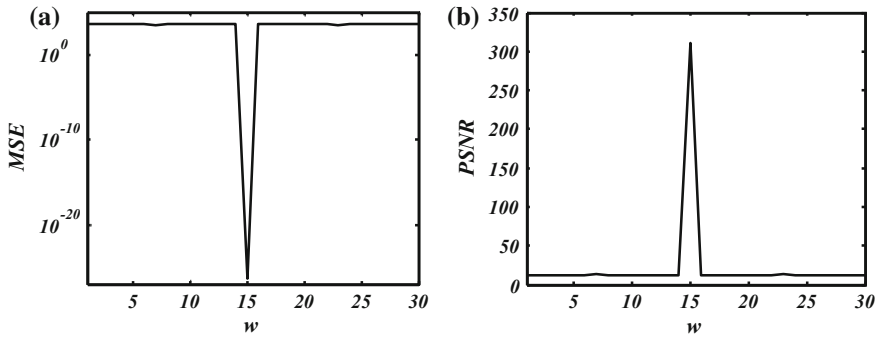
Fig. 6 Sensitivity plots for various values of affine transform parameter (*w*); **a** Mean-squared error (*MSE*). **b** Peak signal-to-noise ratio (*PSNR*)
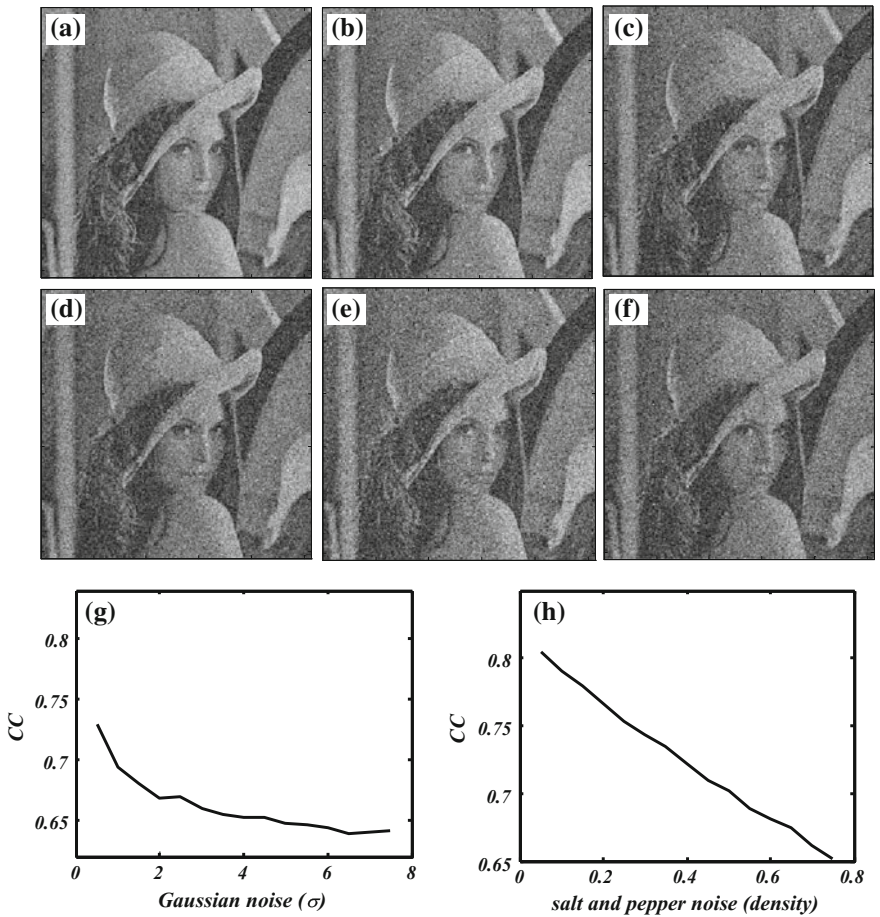
Fig. 7 Results of decryption in presence of salt and pepper noise **a–c** with densities 0.05, 0.25 and 0.55 respectively; and Gaussian noise **d–f** with zero mean and standard deviation 1, 2.5 and 5 respectively; **g**, **h** are plots of correlation coefficient (*CC*) as a function of Gaussian noise, and salt and pepper noise, respectively
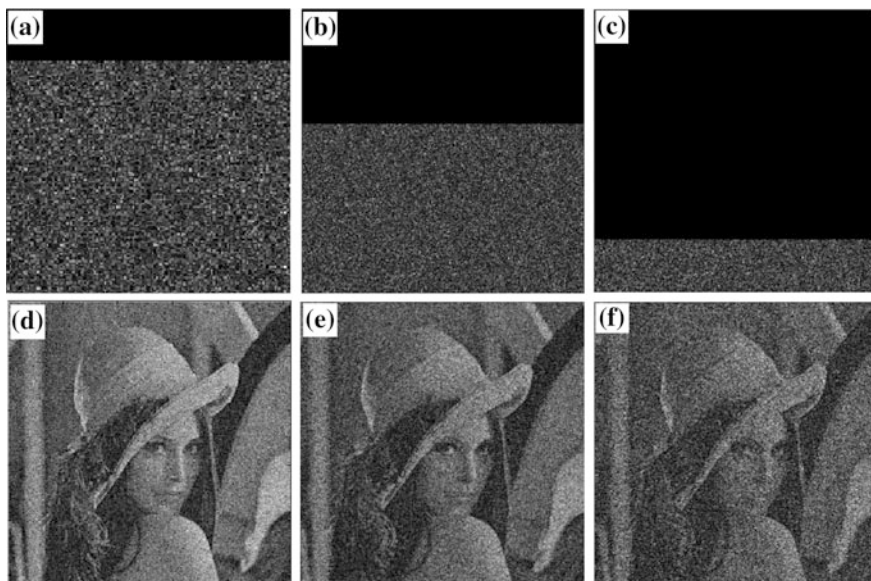
Fig. 8 Occlusion results for the grayscale image for varying degrees of occlusion; **a–c** show encrypted images with 20%, 40%, and 80% occlusion; **d–f** show the corresponding recovered images

curve with the increase in noise strength in case of Gaussian noise as compared to the salt and pepper noise. Clearly, it is observed that the scheme is robust against the noise attacks.

Figure 8 shows the results of occlusion attack. The encrypted image occluded uniformly by 20%, 40% and 80% is shown in Fig. 8a–c, respectively. The second row (Fig. 8d–f) shows the corresponding recovered images. From these recovered images and the CC curve (Fig. 9), it is quite clear that the proposed scheme is robust against the occlusion attack also.

Fig. 9 Plot of correlation coefficient (*CC*) versus percentage occluded area
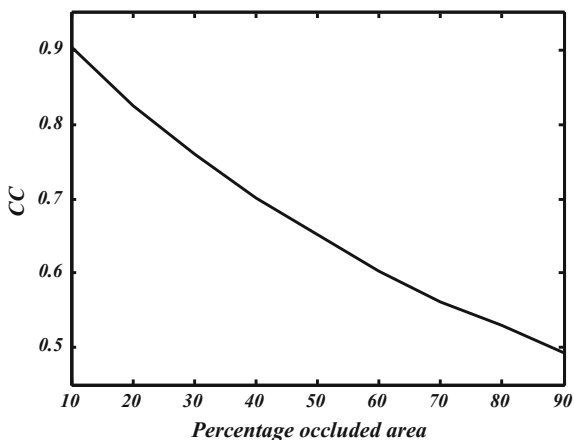
5 Conclusions

In this paper, we have proposed a new scheme for image encryption using affine transform and asymmetric keys. The pair of keys used for encryption and decryption is obtained by phase truncation in Fourier transform domain. The phase reserved parts serve as the decryption keys. We have shown that the scheme is so sensitive to the parameter of the affine transform (w) that even with a slight change in w-value, the original image cannot be recovered. The scheme is analyzed against various noise and occlusion attacks through simulations using MATLAB 7.14. The results indicate that the proposed encryption scheme provides enhanced security and robustness.

References

1. Stallings, W.: Cryptography and Network Security: Principles and Practices. Pearson/Prentice Hall (2006)
2. Refregier, P., Javidi, B.: Optical image encryption based on input plane and Fourier plane random encoding. Opt. Lett. **20**, 767–769 (1995)
3. Situ, G., Zhang, J.: Double random-phase encoding in the Fresnel domain. Opt. Lett. **29**, 1584–1586 (2004)
4. Unnikrishnan, G., Singh, K.: Double random fractional Fourier-domain encoding for optical security. Opt. Eng. **39**, 2853–2859 (2000)
5. Singh, P., Yadav, A.K., Singh, K., Saini, I.: Optical image encryption in the fractional Hartley domain, using Arnold transform and singular value decomposition. AIP Conf. Proc. **1802**, 020017 (2017)
6. Singh, P., Yadav, A.K., Singh, K.: Phase image encryption in the fractional Hartley domain using Arnold transform and singular value decomposition. Opt. Lasers Eng. **91**, 187–195 (2017)
7. Carnicer, A., Montes-Usategui, M., Arcos, S., Juvells, I.: Vulnerability to chosen-cyphertext attacks of optical encryption schemes based on double random phase keys. Opt. Lett. **30**, 1644–1646 (2005)
8. Peng, X., Zhang, P., Wei, H., Yu, B.: Known-plaintext attack on optical encryption based on double random phase keys. Opt. Lett. **31**, 1044–1046 (2006)
9. Qin, W., Peng, X.: Asymmetric cryptosystem based on phase-truncated Fourier transforms. Opt. Lett. **35**, 118–120 (2010)
10. Sui, L., Lu, H., Ning, X., Wang, Y.: Asymmetric double-image encryption method by using iterative phase retrieval algorithm in fractional Fourier transform domain. Opt. Eng. **53**, 026108 (2014)
11. Liu, H., Kadir, A.: Asymmetric color image encryption scheme using 2D discrete-time map. Signal Process. **113**, 104–112 (2015)
12. Wang, Y., Quan, C., Tay, C.J.: Asymmetric optical image encryption based on an improved amplitude–phase retrieval algorithm. Opt. Lasers Eng. **78**, 8–16 (2016)
13. Chen, H., Tanougast, C., Liu, Z., Sieler, L.: Asymmetric optical cryptosystem for color image based on equal modulus decomposition in gyrator transform domains. Opt. Lasers Eng. **93**, 1–8 (2017)
14. Nag, A., Singh, J.P., Khan, S., Ghosh, S., Biswas, S., Sarkar, D., Sarkar, P.P.: Image encryption using affine transform and XOR operation. In: 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies, pp. 309–312 (2011)

A Novel Technique for an Adaptive Feedback Canceller for Hearing Aids

**Ajay Jatav, Ruchi Mehra, Tannu Bala, Gagandeep Singh,
Raman Arora, Gunjan Dogra and Mandeep Kaur Bedi**

Abstract This chapter is focused on the implementation aspects of adaptive feedback canceller algorithms and their computational complexity when reducing misalignment and convergence rates. When an adaptive algorithm filter was used for modeling the acoustic feedback, there was wide misalignment due to a fixed step size. Through the use of the prediction–error method (PEM), the bias in the algorithm for an adaptive filter was reduced. The PEM used a variable step size and a full range of adaptive filters were used as a trade-off between the misalignment and the convergence speed. Various performance measures were considered in order to study these algorithms: misalignment, maximum stable gain, added stable gain, and the algorithm execution time. The disadvantage of misalignment and convergence when changing step size has been addressed using a new algorithm that has automatic step size adjustment. This new algorithm demonstrated effectiveness in controlling misalignment. The findings reveal that the misalignment, maximum added gain, and added stable gain improved with the use of the new adaptive filter algorithm. Despite this, the PEM did not satisfy user requirements and so a new system named AFC-PEM MPVSS is proposed. Furthermore, work has been done to measure the quality of signal.

Keywords Prediction–error · Feedback canceller · AFC-PEM MPVSS

1 Introduction

An acoustic feedback loop is generated when using an acoustic network in conjunction with a loudspeaker and microphone signal, and is a fundamental issue with regards to everyday hearing aid use. Moreover, acoustic feedback reduces what can be heard by the user, regularly presenting itself as "howling" when the hearing aid output becomes unstable. Furthermore, to eliminate the destructive consequences of

A. Jatav (✉) · R. Mehra · T. Bala · G. Singh · R. Arora · G. Dogra · M. K. Bedi
Webtunix Solutions Pvt. Ltd., Mohali, India
e-mail: webtunix@gmail.com; info@webtunix.com

feedback loop, an acoustic feedback cancellation (AFC) method has been designed which makes use of an adaptive filter. In contrast to an acoustic echo cancellation (AEC) machine, the AFC process is a closed-loop system. For that reason, waves and loudspeaker signals in an AFC device are interrelated, as a result the resolution of both waves can be biased [1].

1.1   Statement of the Problem

Acoustic feedback in hearing aids causes undesired coupling between the loudspeaker (also called a receiver) and the microphone. As a result, hearing aids generate an excessive undesired noise instead of outputting the desired signal when the volume is increased. Increasing the distance between the loudspeaker and microphone of the hearing aid increases its hearing capacity and improves its noise-cancellation ability. The problem of undesired coupling is shown in Fig. 1 for a hearing aid with a particular microphone. The following text gives the notation used in the analysis. The so-called ahead path is: $G(q) = g_0 q^0 + \cdots + g L_{G-1} q^{-L_G+1}$, where L_G denotes filter length and $G(q)$ denotes the normal signal-processing route of the device. We expect that the delay d_G of $G(q)$ should be the minimum first sample, which means $g_0 = 0$. $F(q, n)$ is the path of the feedback and $u_1[n]$ and $y[n]$ are the microphone and loudspeaker signals, respectively. The reference signal is expressed via $x[n]$ and the response signal is expressed by $v[n] = F(q, n)u_1[n]$. Because of the undesired coupling, amplified sound $u_1[n]$, sent through the loudspeaker, is again fed into microphone, which leads to closed-loop formation between reference and loudspeaker Signals $u_1[n]$ [2].

$$\frac{G(q)}{1 - G(q)F(q, n)} \tag{1}$$

Instability happens if the loop advantage $|G(e^{i\omega})F(e^{i\omega})|$ crosses at an angular frequency $\omega \in [0, \pi]$ where the loop phase equals more than 2π (Fig. 2).

Fig. 1 Acoustic feedback [2]
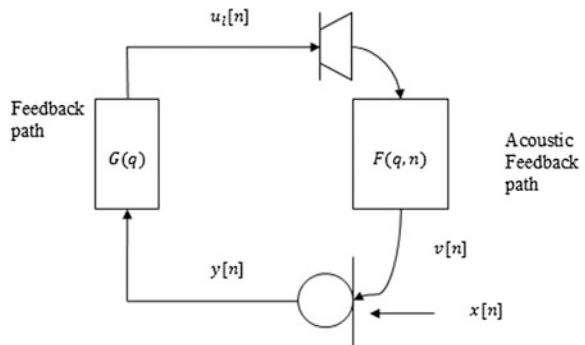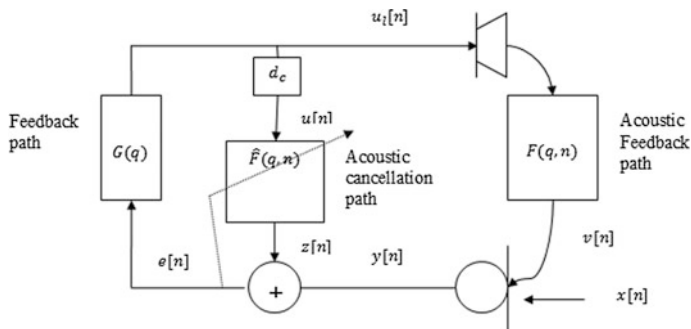
Fig. 2 Adaptive feedback canceller [2]

2 Background

A huge number of studies on the perceptual evaluation of speech quality (PESQ) have been performed. This section gives an overview of previous studies related to this field. When [4] measured the accuracy of PESQ it was shown that PESQ was excessively sensitive to delay, time, processing effects, equipment, and losses. They concluded that PESQ was an ineffective assessment mechanism.

In other research [5], PESQ was presented using test cases in order for training to improve accuracy. It demonstrated an understanding of audio processing, discrete volume levels, packet loss concealment, and jitter buffer adjustment using the PESQ algorithm. They explained the impact on flawed time alignment and calibration errors on estimated subjective opinion.

It was concluded by [6] that PESQ overestimated for community loss but for more extreme conditions, it underestimated.

Pennock [7] offers a fundamental exploration of PESQ accuracy as a predictor of high-quality, speech-based, formal listening tests parallel to PESQ simulations. He concluded that "Outcome point out that PESQ algorithm is a very beneficial in serving to recognize its talents performance issues. However, there are drawbacks of using PESQ for confirmation of performance of voice quality, aggressive evaluation, and procedure optimization," and further "it's included that predominant choices shall be established on results from perceptual experiences and that PESQ just isn't correct adequate to specify exceptional requisites in Service level agreement (SLAs)." In step with writer performance statistics provided in P. 862 are positive. In this work, it is also claimed that regression mapping associated with ITU testing makes fewer errors than basic mapping, and eliminates mistakes that PESQ is sometimes inclined to make.

Morrissey [8] considers the characteristics of PESQ among different algorithms. Testing for an intrusive approach requires observing tools and programs on both sides of a dialog. On the one hand, this would make it hard for PESQ to inspect the result of gateway having PSTN. On the other hand, users have some experience explaining that PESQ is perfect for detecting one-dimensional, fine changes such as

noise degree. In this way, if all the parameters are comparable to codec and other algorithms then it stored unaffected variations and it is acceptable as a result of one parameter may also be detected effectively. Further this consistency no longer applies to alterations in codec, network situations, or every other essential change so as to have an impact on a few great dimensions.

3 Verification and Results

The results are evaluated for speech and music signals for AFC-PEM MPVSS with a step size of 0.001 and AFC-PEM with a step size of 0.01. The measurement is done using 2 acoustic paths and 2 hearing aids [16]. Where the environment considered entirely different, $(F_1(f))$ is calculated in free space, second is evaluated with telephone receiver. The input signal used is a real human voice [20] for both speech and noisy speech signals, furthermore, a real music signal, John Lennon's sound track *Imagine*, was also used. The simulation lasts 80 s. MIS, MSG, and ASG parameters are used to analyze the performance of our proposed method. The description is given below (Figs. 3 and 4).

$$\text{MIS} = 10 \log_{10} \left(\frac{\|f - \hat{f}\|_2^2}{\|f\|_2^2} \right) \tag{2}$$

While ASG is computed according to [7, 8]:

$$\text{ASG} = 10 \log_{10} \frac{1}{\max_\propto \left| F(\propto) - \widehat{F}(\propto) \right|^2} - 10 \log_{10} \frac{1}{\max_\propto |F(\propto)|^2} \tag{3}$$

Figure 5 shows the output from the proposed technique using a speech input signal. From the figure we can interpret that AFC-PEM MPVSS shows significant enhancement in tracking and convergence rate when compared to AFC-PEM with a lower step size $\mu_2 = 0.001$, at the same time obtaining a lower steady state error than AFC-PEM with a higher step size $\mu_1 = 0.01$.

In Fig. 5a–c it is quite easily visible that howling occurs when the feedback path changes. In the case of AFC-PEM $\mu_2 = 0.001$ the howling period is longer than AFC-PEM $\mu_1 = 0.01$. So in the case of a step size of 0.01 the system converges faster, as shown in Fig. 5a, while at the same time it provides a larger steady-state error and increased variation in performance. While increased howling reduces the steady-state error and lessens the variation in performance, it makes the systems converge very slowly as shown in Fig. 5b, correspondingly, in Fig. 5c it shows a trade-off in performance behavior. When we look at the graphs of our proposed system they show an effective reduction in howling period, and also show improvements in steady-state error and tracking rate. The proposed approach is an
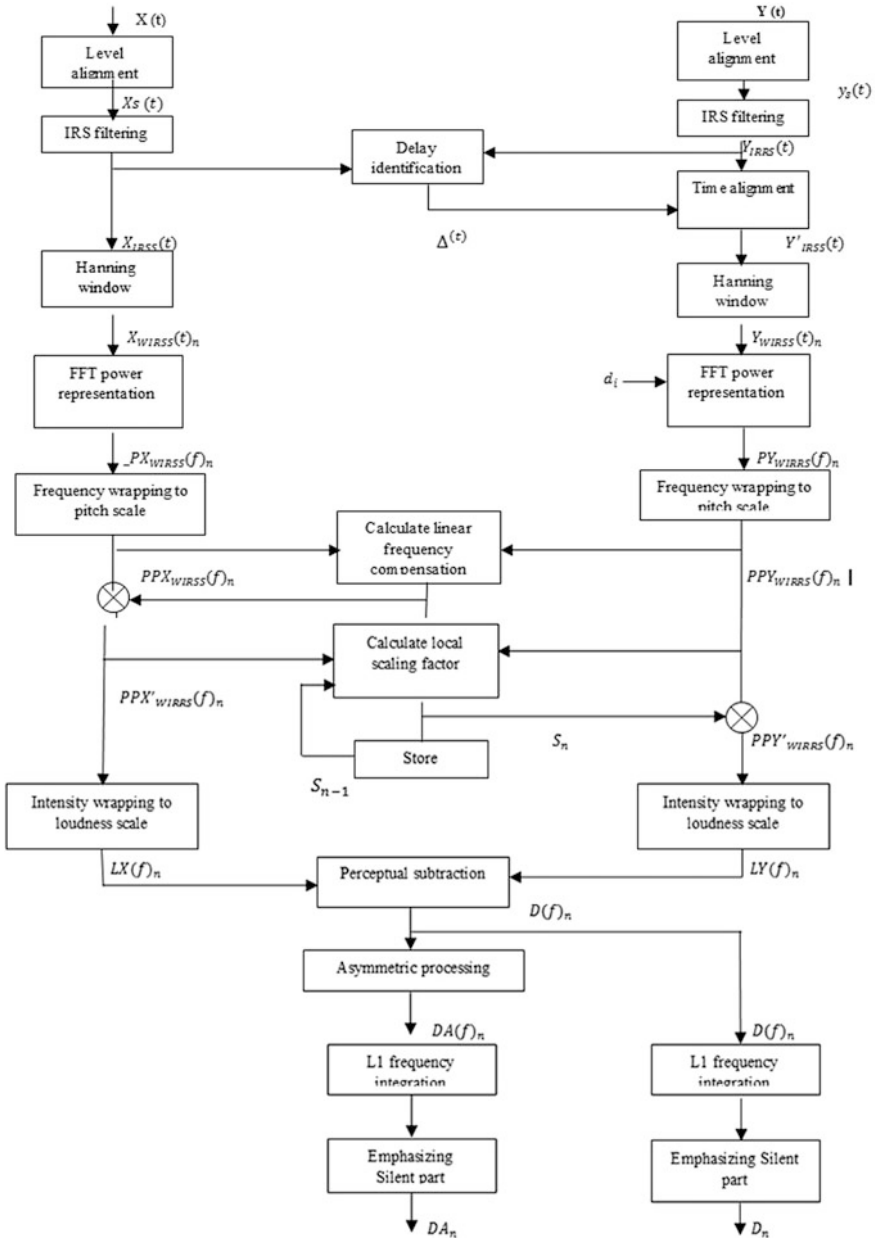
Fig. 3 Flow diagram of the proposed technique
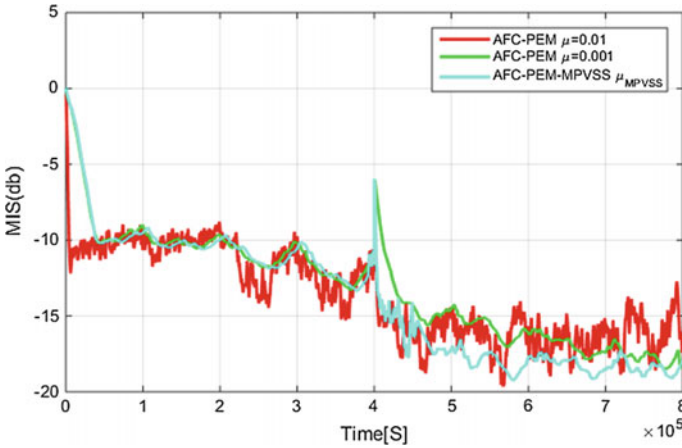
Fig. 4 Performance of the AFC-PEM with step sizes of 0.01–0.001, and AFC-PEM MPVSS for speech signal misalignment

optimum solution. Also, in Fig. 5d we can see the MSG of our proposed system is optimum as it remains stable throughout, signifying a lower steady-state error. All the above-mentioned results are evaluated using speech signals and in the next section we evaluate the results using music signals.

Figure 6 shows the output for the proposed technique for music signals, demonstrating that AFC-PEM MPVSS showed significant improvement in tracking and convergence rate when compared to AFC-PEM with a step size $\mu_2 = 0.001$. At the same time it shows a lower steady-state error than AFC-PEM with a higher step size $\mu_1 = 0.01$.

Figure 7a–c depict the performance in terms of music signal. As with the speech signal, the performance remains the same in this case. Similar to AFC-PEM $\mu_1 = 0.01$, the howling period is significantly reduced compared to AFC-PEM $\mu_2 = 0.001$. Low-level howling causes the system to converge faster, as shown in Fig. 7a, and at the same time it provides a larger steady-state error and high variation in performance, while increased howling reduces the steady-state error and lessens variation in the performance. This makes the system converge very slowly as shown in Fig. 7b, likewise, Fig. 7c shows the trade off in performance. When we look at the graphs of our proposed system, they demonstrate effective reduction in the howling period and also an improvement in steady-state error and tracking rate. The proposed approach is an optimum solution. Figure 7d also shows the MSG of our proposed system is optimum, since it remains stable throughout, which signifies a lower steady-state error.

It can therefore be said that the proposed AFC-PEM MPVSS adapts using variable step sizes. Hence, it utilizes a large step size when the system is not stable and is showing a change in feedback path, and a small step size when the system has converged.
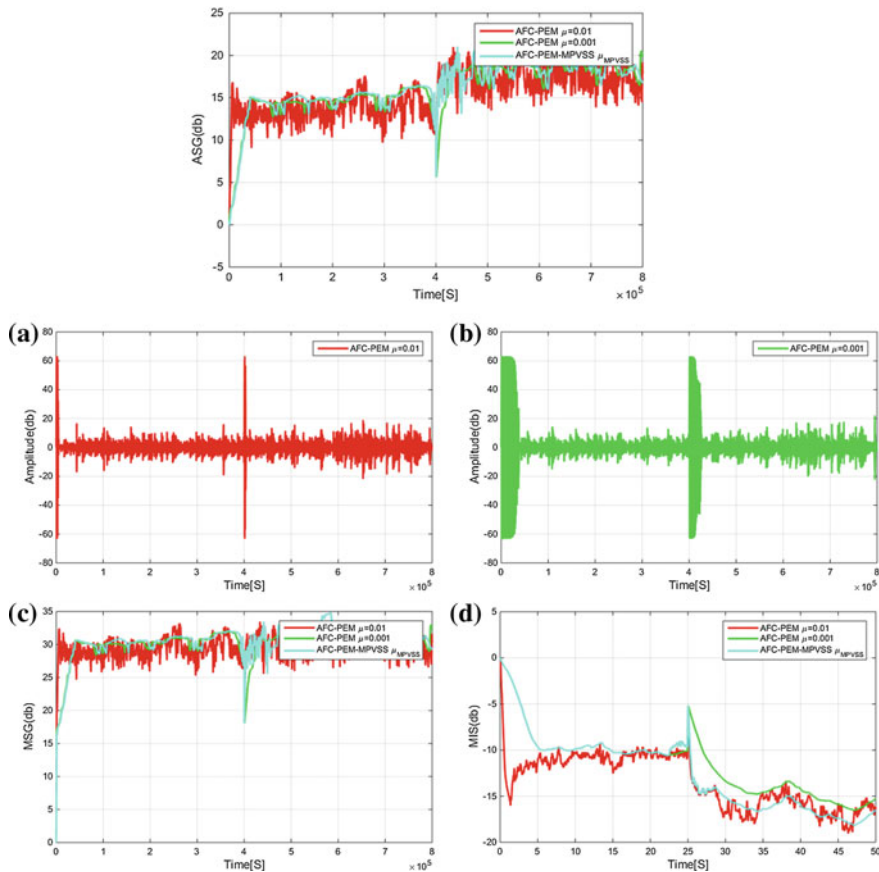
Fig. 5 Performance of AFC-PEM with step sizes of 0.01–0.001, and AFC-PEM MPVSS for speech signal. **a** Loudspeaker signals for incoming speech: AFC-PEM with $\mu_1 = 0.01$. **b** Loudspeaker signals for incoming speech: AFC-PEM with $\mu_2 = 0.001$. **c** Performance of AFC-PEM with step sizes of 0.01–0.001, and AFC-PEM MPVSS with maximum stable gain (MSG). **d** Performance of AFC-PEM with step sizes of 0.01–0.001, and AFC-PEM MPVSS misalignment for music signals

3.1 Perceptual Evaluation of Speech Quality

Results are shown of tests made on different codecs utilizing the PESQ measurement algorithm. For calculating PESQ values, in the case of both music and speech signals, the reference signal and the output signals are compared. The results are shown in Tables 1 and 2.

We utilized PESQ for the perceptual evaluation of speech and music. As shown in the above results for ASG and MIS, evaluation has been made for the same cases with the help of a 60 s input sequence. Results are compared based on three
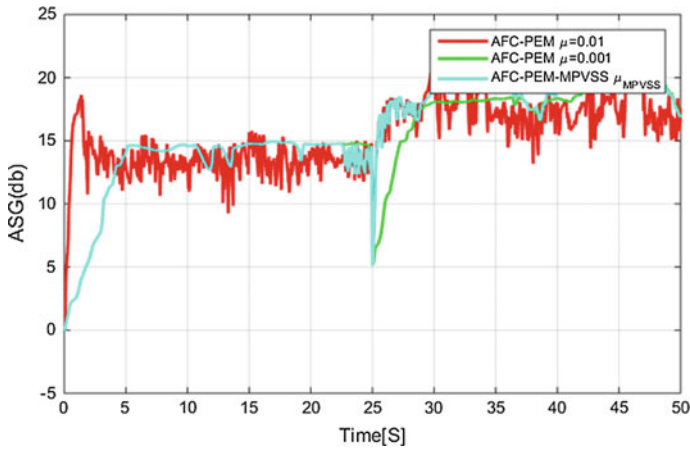
Fig. 6 Performance of AFC-PEM with step sizes of 0.01–0.001, and AFC-PEM MPVSS added stable gain for a music signal
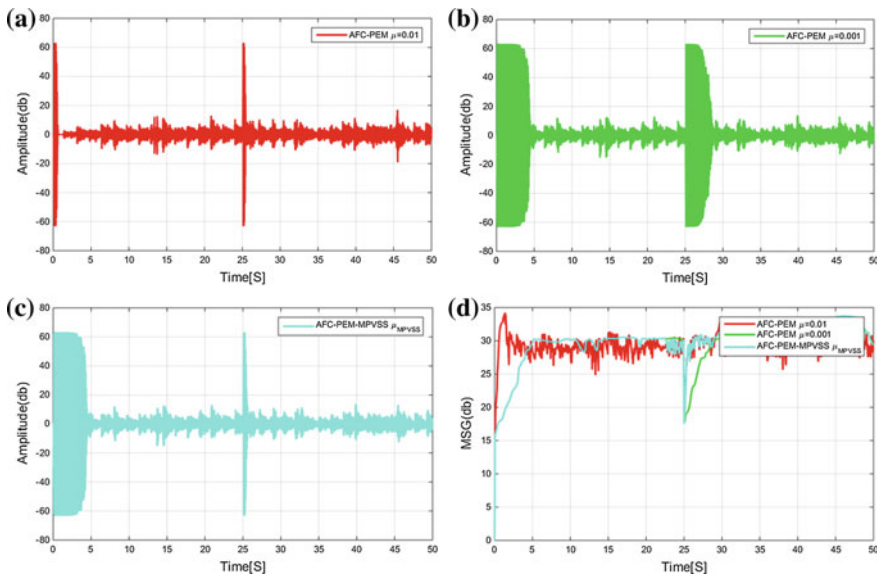
Fig. 7 **a** Loudspeaker signals for incoming music for AFC-PEM with $\mu_1 = 0.01$. **b** Loudspeaker signals for incoming music for AFC-PEM with $\mu_2 = 0.001$. **c** Loudspeaker signals for incoming music for AFC-PEM MPVSS. **d** Performance of AFC-PEM with step sizes of 0.01–0.001, and AFC-PEM MPVSS with maximum stable gain

Table 1 Speech signal

Parameters	PESQ	AVE-MIS	AVE-MSG	AVE-ASG
PEM step size of 0.01	4.2932	–13.4110	29.4763	15.4498
PEM step size of 0.001	4.1135	–13.0845	30.1844	16.1569
Variable step size	4.3958	–13.9193	30.4167	16.3892

Table 2 Music signal

Parameters	PESQ	AVE-MIS	AVE-MSG	AVE-ASG
PEM step size of 0.01	4.2932	–13.4110	29.4763	15.4480
PEM step size of 0.001	4.1457	–11.4678	29.3695	15.3419
Variable step size	4.3094	–13.7714	30.2882	16.2606

different values of PEM, i.e., for PEM step sizes of 0.1, 0.001, and for variable step sizes.

3.2 Perceptual Evaluation of Speech Quality for a Speech Signal

See Table 1.

3.3 Perceptual Evaluation of Speech Quality for a Music Signal

See Table 2.

4 Conclusions

Acoustic feedback is a major problem associated with modern hearing aids. With the advancement of low-power, digital signal–processing devices, demand has increased for a more efficient algorithm which provides improved sound performance, and can deal with the existing acoustic feedback problem. In this chapter we have developed a modified algorithm which can eliminate steady-state error without affecting the rest of the parameters, namely convergence rate. The modified algorithm is applied to the lower and upper bound limit in existing AFC-PEM and we named it MPVSS. We have tested the algorithm for two different signals, speech and music signals, and it shows improved performance in each case. It can be said

that MPVSS is superior to all other algorithms discussed in this chapter. It shows a desirable trade-off between steady-state error and convergence rate. At the beginning of any simulation the algorithm requires a large step size because the system will be unstable at this time—there is also a change in feedback path. However, when the system has converged, it uses a smaller step size. Moreover, it is also clear from the results of PESQ that the quality of the speech signal is better than that of the music signal.

References

1. Schepker, H., Doclo, S., Hai, M., Dam, H., Nordholm, S., Tran, LT.: A Modified Practical Variable Step-Size For Feedback Cancellation Using Pre-whitening Filters
2. Rombouts, G., Moonen, M., Wouters, J., Spriet, A.: Adaptive feedback cancellation in hearing AIDS. J. Frankl. Instit. **343**(6) 545–573 (2006)
3. Pandey, A.: Perceptually Motivated Signal Processing for Digital Hearing AIDS (2011)
4. Beerends, J.G., Hollier, M.P., Hekstra, A.P., Rix, A.W.: Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, p. 862. ITU-T Recommendation (2001)
5. Sun, L., Ifeachor, E., Qiao, Z.: Case study of PESQ performance in live wireless mobile VoIP environment. In: IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1–6 (2008)
6. Marsh, I., Gronvall, B., Varela, M.: A systematic study of PESQ's behavior (from a networking perspective). In: Proceedings of the Measurement of Speech and Audio Quality in Networks, May (2006)
7. Pennock, S.: Accuracy of the perceptual evaluation of speech quality (PESQ) algorithm. In: Proceedings of the Measurement of Speech and Audio Quality in Networks Line Workshop (MESAQIN'02), p. 19 (2002)
8. Morrissey, P.: How to measure call quality. Netw. Comput. (2005)
9. Stemerdink, Jan A., Beerends, John G.: A perceptual speech-quality measure based on a psychoacoustic sound representation. J. Audio Eng. Soc. **42**, 115–123 (1994)
10. Recommendation P ITU-T: 830, Subjective Performance Assessment of Telephone-Band and Wideband Digital Codecs. Int. Telcommun. Union Radiocommun. Assem (1996)
11. ITUT Rec: p. 48: Specification for an intermediate reference system. International Telecommunication Union, CH-Geneva (1988)
12. Beerends, J.G., Hollier, M.P., Hekstra, A.P., Rix, A.W.: Perceptual evaluation of speech quality (PESQ)—a new method for speech quality assessment of telephone networks and codecs. In: 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. Proceedings (ICASSP'01), vol. 2, pp. 749–752 (2001)
13. Feldtkeller, R., Zwicker, E.: Das Ohr als Nachrichtenempfänger (1967)
14. ITU-T Study Group 12: Review of Validation Tests for Objective Speech Quality Measures. Document COM, pp. 12–74, March (1996)
15. Albu, F., Coanda, H., Rotaru, M.: A variable step size modified decorrelated NLMS algorithm for adaptive feedback cancellation in hearing AIDS. In: Proceedings of the ISETC, vol. 2012, pp. 263–266 (2012)
16. Tran, L.: Varible Step-Size Control for PEM (2016)
17. Beerends, J.G., Hollier, M.P., Hekstra, A.P., Rix, A.W., Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, p. 862. ITU-T Recommendation (2001)

18. Manjunath, T.: Limitations of perceptual evaluation of speech quality on VoIP systems. In: 2009 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, pp. 1–6 (2009)
19. ITUT Rec: p. 48: Specification for an Intermediate Reference System. International Telecommunication Union, CH-Geneva (1988)
20. Urban, F., Hellgren, J.: Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification. IEEE Trans. Speech Audio Process. **9**, 906–913 (2001)

Preprocessing of Skin Cancer Using Anisotropic Diffusion and Sigmoid Function

Kartik Sau, Ananjan Maiti and Anay Ghosh

Abstract Skin cancer is one of the atrocious diseases observed in the western part of the world due to exposure to the ultraviolet (UV) rays approaching from the sun- and human-made tanning beds. The survival rate of skin cancer is very high if it is detected at an early phase and treated surgically. To detect it, preprocessing of affected skin lesion images is essential. Here, we are representing a technique for preprocessing of skin lesion via contrast enhancement followed by anisotropic diffusion and sigmoid function. In this method, we critically normalized the skin lesion images followed by removing Gaussian noise and preserving some feature by anisotropic diffusion. For more improvement of it, we applied sigmoid function in the spatial domain of the skin lesion image. Here, we critically consider different parameters of anisotropic diffusion and sigmoid function. This innovative method has been successfully used in various low contrast affected skin lesion images. All most in all the cases, it gives the satisfactory results in terms of MSE PSNR, and SSIM values. This proposed method can be used to improve the quality of low contrast images in medical science, satellite imaging, and different industries. The said technique can be applied successfully in various applications.

Keywords Skin cancer · Contrast enhancement · Normalization
Anisotropic diffusion function · Sigmoid function · PSNR

1 Introduction

In the western part of the world, mainly Europe, Australia, and America, skin cancer is one of the most grievous diseases among many white skinned people. The main reasons behind it are ultraviolet light exposure, either from the sun or tanning beds. The issue is lethal in regions like high elevation or close to the equator where daylight

K. Sau (✉) · A. Maiti · A. Ghosh
Department of Computer Science and Engineering, University of Engineering
and Management, New Town, Kolkata, West Bengal, India
e-mail: kartik_sau2001@yahoo.co.in

exposure is extreme. Fair-toned persons, with brown or blue eyes, and individuals with blond or red hair are particularly susceptible. Persons, affected by skin cancer once, may have a 20% chance of rising a second skin cancer in the following 2 years [1]. Every year, approximately 4% of cancer increases for the white people. As a result, it spreads for all the white people, after some particular time frame. The Skin Cancer Foundation mentions that everybody should examine his or her body on a monthly basis. So identification is crucial for new or changing lesions that might be cancerous. Skin cancers found and detached first are almost always remediable. To screening the skin cancer, preprocessing is essential. Nowadays, certain techniques are available to minimize the costs of screening skin cancer, followed by the development of automated skin cancer detection. Most of these techniques are well appreciated. In response to the demand, almost every year, the more different techniques are introduced by various researchers. Still, these techniques are not sufficient to meet the current requirement. To meet the current requirement, here we critically normalize the skin lesion images followed by removing the Gaussian noise (if any) then preserving the edge details by anisotropic diffusion. For more improvement of it, we apply sigmoid function in the spatial domain.

The current paper is organized as follows: In Sect. 2, we have presented the different existing methods as a literature survey. In Sect. 3, we have presented the newly proposed method to overcome the various difficulties as mentioned in Sect. 2. For more visual improvement of skin cancer, we apply sigmoid function in spatial domain in Sect. 4. The experimental results of our novel proposed method along with conclusions are presented in Sect. 5 and Sect. 6, respectively.

2 Literature Survey

The proper screening of melanoma type skin cancer depends on the quality of the said images. In the current scenario, all the acquired images may not be good in quality, to get the better quality images contrast have to be adjusted for proper selection of images. Contrast is an important property by which said the image could be judged as good or poor. There are different methods available in the literature for screening the skin cancer. Some such recent methods are listed below, for screening the melanoma skin cancer.

In 2015, Jain et al. proposed a method for preprocessing the melanoma skin cancer based on gamma correction in their research work. It is described vividly in [2]. In this paper, authors try to improve the quality of skin images by gamma correction followed by resizing the images and correction of brightness. In 2016, Ihab et al. mentioned a technique to the diagnosis of skin cancer [3]. Here, authors applied a 2D median filter to remove the noises from the corrupted skin images with a mask size 5×5 in each RGB channel of an image. For removal of unwanted hair, two morphological techniques are used on grayscale images. Authors applied a brightness correction technique on each channel R, G, and B separately to achieve

the brightness of the image. B channel is selected for the purpose of segmentation, which is converted to binaries form by using Otsu's method. Finalize preprocessed image represents the area of the skin lesion. In 2015, Sumithra R. et al. proposed a technique for preprocessing the skin lesions, based on Gaussian smoothing techniques of mask size 3 × 3, which will help to detect skin cancer [4]. It reduces the noises from the corrupted skin lesion also. In this article, authors also nicely described the segmentation and classification of skin lesions.

3 Proposed Method

3.1 Normalization

To normalize skin lesion images, apply a normalization technique. This technique converts as parts of skin lesion images have the same intensity. It is presented as follows:

$$I'(x, y) = \begin{cases} \phi_d + \lambda & \text{if } I(x, y) \leq \phi \\ \phi_d - \lambda & \text{if } I(x, y) \geq \phi \end{cases} \tag{1}$$

$$\text{where} \quad \lambda = \sqrt{\frac{\rho_d (I(x, y) - \phi)^2}{\rho}} \tag{2}$$

Here, $I'(x, y)$ is the normalized intensity; ϕ_d is the average gray value of the desired image; ρ_d is the variance of intensity of the desired image; and ϕ is the average gray value of input images. ρ is the difference of skin images.

3.2 Anisotropic Diffusion Function

After normalization, we apply the anisotropic diffusion (AD) to reduce the noises (if exist) and preserve some features (like edges) of original skin images. The basic equation of anisotropic diffusion is presented as follows:

$$\frac{\partial I(x, y, t)}{\partial t} = div(\|(g(\nabla I(x, y, t)))\| \nabla I(x, y, t)), \tag{3}$$

where t is a time parameter, $I(x, y, 0)$ is the original image. $\nabla I(x, y, t)$ indicates the gradient of the input image at time t and $g(.)$ is the so-called monotonically decreasing conduction function. This function is chosen to satisfy $g(x) \to 1$ as $x \to 0$ and $g(x) \to 0$ as $x \to \infty$ so that the diffusion will mainly take place in the interior regions and will not affect the region boundaries where the magnitude of

g(.) is large. The different conduction functions proposed by different researchers (Perona and Malik, Black et al.) are as follows:

$$g_1(x) = \exp\left(-\left(\frac{x}{k}\right)^2\right) \tag{4}$$

$$g_2(x) = \frac{1}{1 + \left(\frac{x}{k}\right)^2} \tag{5}$$

$$g_3(x) = \begin{cases} \frac{1}{2}\left[1 - \left(\frac{x}{k}\right)^2\right]^2 & \text{If } x \leq S \\ 0 & \text{elsewhere} \end{cases} \tag{6}$$

The rate of diffusion is controlled by the gradient threshold parameter (κ), and it serves as a soft threshold between the image gradients that fare attributed to noise and those attributed to edges. The Eqs. (4) and (5) proposed by Perona and Malik respectively whereas Eq. (6) suggested by the Black et al., where, the above anisotropic diffusion equation is discretized [5] as follows

$$I_{t+1}(s) = I_t(s) + \frac{\lambda}{|\eta_s|} \sum_{p \in \eta_s} g_k(|\nabla I_{S,P}|) \nabla I_{S,P} \tag{7}$$

where "I" is a discretely sampled image, s denotes the pixel position in the discrete image plane, t denotes the iteration steps, g is the conduction function, and the value of k indicates that the different conductance functions. Consider $\lambda \in (0, 1)$ determine the rate of diffusion and η_s represents the spatial eight-pixel neighborhood of the pixel "s": $\eta_s = \{N, S, E, W, NE, ES, SW, WN\}$. Where N, S, E, and W are the north, south, east, and west neighbor of the pixels, respectively. Consequently, cardinality of η_s gives us eight (except image border). The symbol ∇ is defined as the difference between adjacent pixels in each direction with pixels. Therefore, we can write $\nabla I_{s,p} = I_t(p) - I_t(s)$.

Where

$$p \in \eta_s = \{N, S, E, W, NE, ES, SW, WN\}. \tag{8}$$

3.3   Selecting the Conduction Function

According to Perona and Malik [5], $g_1(.)$ conductance function favors high contrast edges over low contrast edges, whereas $g_2(.)$ function favors wide region over the smaller one. The $g_3(.)$ function produces the sharper edges as its rate convergence is very high [6]. Now, all the conductance functions can be scaled so that their respective flow function $\phi(x)$ ($= xg(x)$) reaches same maximum value or same

brightness flow at same point x = k. This idea gives the following of modified conductance.

$$g_1(x) = \exp\left(-\left(\frac{x}{k\sqrt{2}}\right)^2\right) \tag{9}$$

$$g_2(x) = \frac{1}{1 + \left(\frac{x}{k}\right)^2} \tag{10}$$

$$g_3(x) = \begin{cases} 0.675\left[1 - \left(\frac{x}{k\sqrt{5}}\right)^2\right]^2 & \text{if } x \le k\sqrt{5} \\ 0 & \text{elsewhere} \end{cases} \tag{11}$$

The conductance function $g_1(.)$ and $g_2(.)$ can smooth the entire image. Whereas $g_3(.)$ proposed by Tukey's weight the brightness flow is much faster than others, but it stops the diffusion near to the edges. As a result, some important image details like edges are preserved. This scale space comparison admires $g_3(.)$ function which produces a better result. The decrement of brightness flow is much faster for $g_3(.)$, thus it prevents a good amount of edges beyond a specific threshold S. The noises and the edges of an image are differentiated by a boundary, here the threshold S is considered as that boundary, which also states that the local gradient value under S is considered as noises and they are smoothed away, and the gradient values beyond S are considered as edges where the diffusion is stopped, and the edges are preserved. Therefore, all the conductance function are scaled in such a way that their corresponding flow function reaches to zero at the same point so that comparison is drawn among all the function. $g_2(.)$ function approaches to zero very slowly. As a result, it smoothed away the noises more efficiently than the preservation of edges. The $g_1(.)$ function never reaches to zero. As a result, it scaled that it reaches to very minimum value at point S where the other two functions reach to zero and both $g_2(.)$ and $g_3(.)$ are arranged in such a way so that their flow function produces maximum diffusion values.

3.4 Calculating the Gradient

If the conduction function is selected properly, then anisotropic diffusion (AD) scheme can reduce the noises [7–9] efficiently. It also preserves the edges of skin images. The said scheme is incapable of reducing noises and preserving the edges for a high level of noises. To solve this problem, replace the term $g(|\nabla I(x,y,t)|)$ with the modified term $g(\nabla|G_\sigma * I(x,y,t)|)$. Here, G_σ is the Gaussian filter and σ indicates the square root of the variance of noises of Gaussian filters [10]. This procedure facilitates us to compute the conductance function's argument which is the local image gradient, with the smoother version of the image with each

every iteration. The automatic estimation of the standard deviation of the Gaussian filter σ is highly dependent on the amount of noise present in the image. As a result, the suitable sliding window of 25×25 or 64×64 is considered so that we can detect the maximum identical chunk of pixels of the image. The standard deviation of the pixels for given image block is measured by the uniformity of the pixels. The standard deviation produced by the most of the significant blocks of the image is considered as the last standard deviation of the Gaussian filter.

3.5   Estimating Gradient Threshold Parameter

The gradient threshold is it detects one of critical parameter since a boundary of an image. Again, it is also responsible for noises and edges for skin lesion image. There are different techniques prescribed in literature to compute the gradient threshold parameter. Some such techniques are listed below.

1. Based on noise estimation, Perona and Malik [5] calculate the gradient threshold parameter using noise estimator described by Canny [11]. In this method first, compute the histogram of absolute values of the gradient of the entire image. The gradient threshold (S) is equal to the 90% values of its cumulative sum of said histogram.
2. Black et al. [5] proposed a method to calculate the gradient threshold parameter based on median by the formula $S = 1.4826\,\mathrm{MAD}(\nabla I)$

$$\text{where} \quad \mathrm{MAD} = \mathrm{meadian}(\|\nabla I - \mathrm{median}(\|\nabla I\|)\|) \tag{12}$$

It is constant throughout the entire image and $\|.\|$ indicates norm function in this situation.
3. Voci et al. suggested a technique [12] for estimating the gradient threshold parameter by the formula

$$S = \frac{\sigma \|I\|_p}{rc}, \tag{13}$$

where σ is a variance of the entire image and it is constant for a particular image, r and c signify the number of rows and columns of image and $\|I\|_p$ is the image p-norm and is defined as

$$\|I\|_p = \left(\sum_{i,j \in I} |I|^p \right)^{\frac{1}{p}} \tag{14}$$

The first two procedures for estimating the gradient threshold parameter are based on absolute values of the gradient as mentioned above, while we are dealing

discrete anisotropic diffusion scheme. The discrete AD scheme deals with eight different AD values as the present methods based on eight-neighborhood. The gradient is calculated by the difference between the current pixel values with directed neighborhood pixel values. The implementation of ∇ in discrete domain indicates the scalar values instead of vector values. So, eight different threshold parameters along eight different directions are estimated with the help of directional difference of intensity values. However, the calculated absolute difference values along north and south are almost equal, it is also the same along east and west directions and so on. As a result, eight threshold parameters reduce to four threshold parameters. Thus, the discrete AD scheme reduces to

$$
\begin{aligned}
I_{t+1}(S) = I_t(S) + \frac{\lambda}{\eta_s} \Bigg[&\sum_{p \in N, S} g(\nabla I_{s,p}) \nabla I_{s,p} + \sum_{p \in E, W} g(\nabla I_{s,p}) \nabla I_{s,p} \\
+ &\sum_{p \in NE, SW} g(\nabla I_{s,p}) \nabla I_{s,p} + \sum_{p \in SE, NW} g(\nabla I_{s,p}) \nabla I_{s,p} \Bigg]
\end{aligned}
\tag{15}
$$

To select the single threshold for a pair of directed threshold, we pursue the following steps. Select a block of size 3×3 of size 3×3 as shown below.
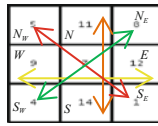
Let μ_N, μ_S be the mean of intensities of north and south block respectively, whereas σ_N, σ_S denote standard deviation of the same. P_N, P_S be the prior probability of selecting the north block threshold and south block threshold respectively such that $P_N + P_S = 1$. The optimal threshold is calculated as follows, suppose gray level in the image is denoted by z and it follows $N(\mu_N, \sigma_N)$ for north region and $N(\mu_S, \sigma_S)$ for south region. The probability density function of the gray level of north region is $p_N(z)$, and that of the south region is $p_S(z)$. Then, we can write

$$
p_N(z) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(\frac{-(z - \mu_N)^2}{2\sigma_N^2} \right)
\tag{16}
$$

$$
p_S(z) = \frac{1}{\sqrt{2\pi}\sigma_S} \exp\left(\frac{-(z - \mu_S)^2}{2\sigma_S^2} \right)
\tag{17}
$$

Hence, pdf of gray level for the north and south region, i.e., normalize gray level histogram is $p(z) = P_N p_N(z) + P_S p_S(z)$. Let us assume that the average gray level of north is less than or equal to the mean gray level of south region. Then, the threshold t that makes a pixel to north region if $(z \leq t)$ or to south region if $(z < t)$, must satisfies as $\mu_1 \leq t \leq \mu_2$ it distinguish between the pixel of north and south the

probability of erroneously classifying a pixel (that belongs to south) to north is $E_1(t) = \int_{-\infty}^{t} p_s(z)dz$ the probability of erroneously classifying a pixel (that belongs to north) to south is $E_2(t) = \int_{-\infty}^{t} p_n(z)dz$ $\mu_1 \leq t \leq \mu_2$. The total error due to said threshold t is $E(t) = P_S E_1(t) + P_N E_2(t)$. To find the minimum E(t), we differentiate on t and equal to result to zero; we get $P_S p_S(t) = P_N p_N(t)$. After calculation, we get.

Using (16) and (17) with respect to that

$$at^2 - bt + c = 0 \quad \text{or} \quad t = \frac{b \pm \sqrt{4 - 4ac}}{2a} \tag{18}$$

Where, $\quad a = \sigma_N^2 - \sigma_s^2 \quad b = 2(\mu_s \sigma_N^2 - \mu_s \sigma_s^2) \quad c = \sigma_N^2 \mu_s^2 - \mu_N^2 \sigma_s^2 - 2\sigma_s^2 \sigma_N^2 \ln\left(\frac{P_S \sigma_N}{P_N \sigma_S}\right)$

$$\tag{19}$$

t = estimated threshold of north–south region, σ_1 = standard deviation of gradient image north 3 × 3 region, σ_2 = standard deviation of gradient image south 3 × 3 region, μ_1 = mean of the gradient image north 3 × 3 region, μ_2 = mean of the gradient image south 3 × 3 region, p_1 = priori probability that a pixel belongs to gradient image north 3 × 3 region, p_2 = priori probability that a pixel belongs to gradient image south 3 × 3 region. As a result, it reduces to a single threshold for a pair of north and south direction. Similarly, we can apply it to another pair of directions like NE–SW, E–W, and SE–NW. As a result, it reduces to four thresholds.

3.6 Stopping Criteria

The AD filtering is highly sensible to the number of iteration. The choice of iteration of AD scheme is crucial, since overestimating it may result in blurring the right edge, while underestimating it may leave unfiltered noise. As number iteration is proportional to the time so optimal iteration can be considered as optimal threshold time (T). Therefore, the optimal threshold time (T) should be estimated based on each filter version of the noisy image. For these, there are different methods available in the literature [13, 14], one such method is listed below. Mrázek and Navara [15] proposed a method to select the optimal threshold based on minimizing the correlation between $I_0 - I_t$ and I_t. It is represented as follows:

$$T \equiv \arg \min_{t} \frac{\text{cov}(I_0 - I_t, I_t)}{\sqrt{\text{var}(I_0 - I)\text{var}(I_t)}} \tag{20}$$

$\bar{I}$ is ideal image, $\bar{N}$ is additive uncorrelated noise of variance σ_n^2, such that $I_0 = \bar{I} + \bar{N}, I_t$ is a modified image at tth iteration.

3.7 Contrast Adjustment Using Sigmoid Function

The sigmoid function is a nonlinear continues function of "S" shape, and its name derives from its shape. There is a wide verity of the sigmoid functions are used in an ANN including logistic and hyperbolic function. The sigmoid function defined for all real values is bounded as 0 to 10, differentiable and has a positive derivative at each point. It is defined for the input x as follows:

$$f(x, a, c) = \frac{1}{1 + e^{-a(x-c)}} \qquad f(x) = \frac{1}{1 + e^{-x}} \tag{21}$$
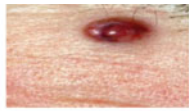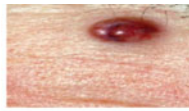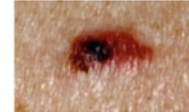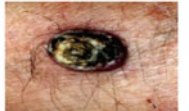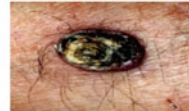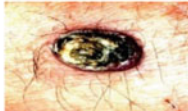
This new image enhancement approach can be regarded as point process where we need direct processing on every single pixel of a particular image apart from other pixels of that image so that the dynamic range can be changed. In this method, a nonlinear activated sigmoid function is convoluted in the form of a mask with the target images and with a factor. Depending on the amount of darkness and brightness in the original image, the factor is used to evaluate the amount of contrast within that image. The mask starts its operation from the upper right corner and goes through pixel by pixel of the entire image. Every single pixel's intensity value is added to the mask value to produce the output intensity value. The following formula defines the output image's intensity values. Simple exponential growth where a dynamic linear limiting control is present is described by the above formula by which the poor contrast of any image is adjusted

$$O(i,j) = I(i,j) + I(i,j) \times C \times 1 / \left(1 + e^{-g(x)}\right), \tag{22}$$

where O(i, j) is the output image. I(i, j) is the unrefined input image. C is a factor. The value of C is enhancement process' objective dependent as the value is not limited within one, two, and three, the user can give own desired value, depending on the contrast you want. For 8-bit images, the pixel value varies in the range of 0–255. After application of the above process the output image also should maintain that particular dynamic range so that it must not cross both the highest and lowest limit.

4 Experimental Results

To evaluate the performance of the proposed method, here we considered three different kinds of images as shown in the Fig. 1. It has a different kind of cancers cells like malignant, Basel cell carcinoma and Squamous cell carcinoma. The said images have variation in lesion sizes, color, and as well as contrast. Also, surrounding skin has some unwanted artifacts (like the presence of hair). The proposed method is applied in a MATLAB 2015 environment, which was installed in processor i5 of the standard machine with 8 GB RAM. It gives the satisfactory results

Original image	Normalize image	Anisotropic image	Sigmoid image
Basal Cell	MSE =0.2105, PSNR=56.3560, SSIM=0.9991		
Malignant	MSE=0.2487,PSNR=57.8907,SSIM=0.9987		
Squamous Cell	MSE=0.1545,PSNR=58.9807,SSIM=0.9993		

Fig. 1 Stating different stages of preprocessing of proposed method

in term of mean square error (MSE), peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) also.

5 Conclusion

The present paper describes enhancing the skin lesion images by applying normalization and AD followed by sigmoid function. The main novelty of this approach is enhancing, reducing the noises and preserving the image details (like edges) also. The proposed method is tested on low contrast gray images as well low contrast color images. The experimental result shows the effectiveness of proposed method. In short, the algorithm is very simple and has the ability to construct better quality images; the addition of conventional approaches to this algorithm makes the algorithm strong, and an excellent contender for the real life implementation; the applications having issues with the image contrast can successfully solve their problem with this algorithm.

Some concluding observations from the study are given below.

- The size of the skin lesion image may be varied, and it may be grayscale and as well as color images, here we consider only color images.
- The proposed method is applicable for low contrast skin lesion images along with bright images also.

- The proposed method automatically selects only lesion portion from the input image.
- The quality of the reconstructed image is good enough. The MSE, PSNR, and SSIM values of reconstructed images are satisfactory for each case.
- The proposed method also removes the Gaussian noises if any and preserves the image details like edges of the skin lesion image via anisotropic diffusion.
- The quality of the output images is good enough, so it helps to identify skin cancer properly via feature extraction.

This proposed method can be used to improve the quality of low contrast images in medical science, satellite imaging, and different industries.

References

1. U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute, Melanoma and Other Skin Cancers
2. Jain, S., Pise, N.: Computer aided melanoma skin cancer detection using image processing. Procedia Comput. Sci. **48**, 735–740 (2015)
3. Zaqout, I.S.: Diagnosis of skin lesions based on dermoscopic images using image processing techniques. Int. J. Signal Process. Image Process. Pattern Recogn. **9**(9), 189–204 (2016)
4. Choudhari, S., Biday, S.: Artificial neural network for skin cancer detection. Int. J. Emerg. Trends Technol. Comput. Sci. **3**(5) (2014). ISSN 2278-6856
5. Black, M.J., Sapiro, G., Marimont, D.H., Heeger, D.: Robust anisotropic diffusion. IEEE Trans. Image Process. **7**(3), 421–432 (1998)
6. Catté, F., Lions, P.L., Morel, J.M., Coll, T.: Image selective smoothing and edge detection by nonlinear diffusion. SIAM J. Numer. Anal. **29**(1), 182–193 (1992)
7. Li, X., Chen, T.: Nonlinear diffusion with multiple edginess thresholds. Pattern Recogn. **27**(8), 1029–1037 (1994)
8. Whitaker, R.T., Pizer, S.M.: A multi-scale approach to non uniform diffusion. CVGIP: Image Underst. **57**(1), 99–110 (1993)
9. Petrou, M., Petrou, C.: Image Processing the Fundamentals. Wiley (2010)
10. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **6**, 679–698 (1986)
11. Voci, F., Eiho, S., Sugimoto, N., Sekiguchi, H.: Estimating the gradient threshold in the Perona-Malik equation. IEEE Signal Process. Mag. **21**(3), 39–46 (2004)
12. Chanda, B., Majumder, D.D.: Digital image processing and analysis. PHI Learning Pvt Ltd. (2004)
13. Gilboa, G., Sochen, N., Zeevi, Y.Y.: Estimation of optimal PDE-based denoising in the SNR sense. IEEE Trans. Image Process. **15**(8), 2269–2280 (2006)
14. Mrázek, P., Navara, M.: Selection of optimal stopping time for nonlinear diffusion filtering. Int. J. Comput. Vis. **52**(2), 189–203 (2003)
15. Hassan, N., Akamatsu, N.: A new approach for contrast enhancement using sigmoid function (2004)

An Incremental Algorithm for Mining Closed Frequent Intervals

Irani Hazarika and Anjana Kakoti Mahanta

Abstract Interval data are found in many real-life situations involving attributes like distance, time, etc. Mining closed frequent intervals from such data may provide useful information. Previous methods for finding closed frequent intervals assume that the data is static. In practice, the data in a dynamic database changes over time, with intervals being added and deleted continuously. In this paper, we propose an incremental method to mine frequent intervals from an interval database with n records, where each record represents one interval. This method assumes that intervals are added one by one into the database and each time an interval is added to the database, our proposed method will mine all the newly generated closed frequent intervals in O(n) time.

1 Introduction

Data mining is an important part of machine learning and pattern recognition. There are many real-life situations in the field of medical, telecommunication, finance, etc., in which intervals are associated with transactions. In general, an interval is the form $[l_1, r_1]$ where l_1 is the lower end and r_1 is the upper end of the interval. Suppose, we consider a telecommunication database in which the information of the calling time of each call is saved. Then, they may keep the starting time and ending time of each call as an interval [start time, end time] (as shown in Table 1). Mining information from such data may give lots of useful information.

I. Hazarika (✉) · A. K. Mahanta
Department of Computer Science, Gauhati University, Guwahati, India
e-mail: queensarathi@gmail.com

A. K. Mahanta
e-mail: anjanagu@yahoo.co.in

Table 1 An interval database

Call no	Call time
1	[10:00 10:30]
2	[11:10 11:20]
3	[10:15 10:45]
4	[10:20 10:40]

If we arrange intervals in 1D space, then 13 possible relationships may occur between two intervals [1] and these are before, after, meets, meet by, overlap, overlap by, during, includes, starts, start by, finishes, finished by, equal. These relationships can be reduced to only 7 relations—before, meet, overlap, during, start, finishes, equality. An interval $X = [l_x, r_x]$ contains an another interval $Y = [l_y, r_y]$ ($X \subseteq Y$) iff $l_x \leq l_y$ and $r_x \geq r_y$. Given a database of intervals, support of an interval X is defined as number of intervals in the database that contain X. An interval whose support is greater than or equal to user-defined minimum support (ρ) is called frequent interval. In Fig. 1, if minimum support is 2 then frequent intervals are all possible intervals within [10:15 10:30] and [10:20 10:40] and other intervals are infrequent. A frequent interval that is not properly contained in any frequent interval is called maximal frequent interval [2, 3]. Here, [10:15 10:30] and [10:20 10:40] are maximal frequent intervals. Again, intervals which are not properly contained in any interval whose support is equal to its support are called closed intervals [4] and closed intervals which are frequent are called closed frequent intervals [5, 6]. In the above example, closed intervals are [10:15 10:30], [10:20 10:30], [10:20 10:40], [11:10 11:20] and closed frequent intervals are [10:15 10:30], [10:20 10:30], [10:20 10:40]. Mining maximal frequent intervals give us dense ranges in which we are interested but by mining closed frequent intervals, we are able to find more dense ranges within the maximal frequent intervals. In the above example, if we are interested to get the ranges in which two calls are active simultaneously then by mining maximal frequent intervals, we get the ranges [10:15 10:30] and [10:20 10:40]. But inside the range [10:15 10:30], there is a range [10:20 10:30] which is more dense than this range, because there are three calls active in this range. To find this range [10:20 10:30], we have to mine the closed frequent intervals. In [5, 6], authors have proposed methods to mine closed frequent intervals from a static interval database. But in dynamic database each time, a new interval is added to the database, the set of closed frequent intervals changes. New closed frequent intervals are to be generated which were infrequent or unseen intervals in the old database. The algorithm proposed here can be used to mine all these newly generated closed frequent intervals. The algorithm assumes that intervals are added to the database one at a time.
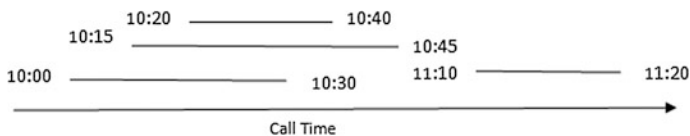
Fig. 1 Arrangement of the call times of Table 1 as a line along X-axis

In Sect. 2, related works are given. In Sect. 3, preliminary definitions and the theoretical background are given. In Sect. 4, the problem definition and in Sect. 5, the proposed algorithm with its explanation and example are given. In Sect. 6, experimental results are shown. In Sect. 7, concluding statements are given.

2 Related Works

In the papers [7–9], the authors have proposed algorithms to cluster interval data and also discussed about different distance measures for interval data. In paper [10], the authors have proposed an indexing and efficient indexing technique for fast retrieval of similar interval time sequences from large databases by addressing the issue of (dis)similarity measures for interval time sequence. The problem of mining maximal frequent intervals was first proposed by Lin [3]. First, it stores the input intervals in a data structure called I-Tree. Next, a preorder traversal algorithm is used to discover all maximal frequent intervals. In [11], authors have proposed a method for mining maximal sparse intervals, which uses maximal frequent intervals to mine maximal sparse intervals. Another algorithm MintMiner [11] uses linear representation of intervals to find all maximal frequent intervals in linear time. In [4], an incremental method to mine closed interval is proposed which uses a Closed Interval Tree (CI-Tree) for mining all closed intervals in the database. In [5], the authors proposed a method for mining closed frequent intervals by using maximal frequent intervals. Another data structure called IS-Tree [3] was also proposed to mine all closed frequent intervals. In [6], authors have proposed another method to mine closed frequent intervals from an interval dataset in O(n2) time. This method assumes that the interval database is static. The method can be divided into two parts—preprocessing and main algorithm. In preprocessing step, the interval database is represented in an array structure as in [3]. In the second part, the main algorithm is applied on the structure to find all closed frequent intervals present in the interval database. To generate closed intervals, the algorithm used a temporary singly linked list to store left ends of intervals so that when a suitable right end is obtained, then these left ends and that right end can be paired to form closed frequent intervals. From above, it is seen that the entire method can mine the closed frequent intervals from a static database. The incremental method proposed in this paper can mine closed frequent intervals from a dynamic interval database without using any extra memory space.

3 Preliminaries and Theoretical Background

In this section, definitions of terms that are used throughout the work are provided. A number of theorems stating certain properties of closed frequent intervals that have been used in designing the algorithm are also proved.

3.1 Preliminaries

Given a totally ordered discrete domain D, a non-empty subset I of D is called an interval iff for all $l, r \in I$ and $c \in D$, $l \leq c \leq r$ implies $c \in I$. If for a given interval I, $l \leq c \leq r$ holds for all $c \in D$ where l and r are two specific elements in I, then I is denoted by $[l, r]$, and l is the left end point and r is called the right end point of I. The interval I is a closed interval because for any $c \in I$, $l \leq c \leq r$. An interval is called *frequent interval* if its support is greater than or equal to user-defined minimum support (ρ). An interval $[l, r]$ is said to be *closed* if the support of any interval properly containing $[l, r]$ is less than the support of $[l, r]$ i.e. if $[l, r] \subset [l', r']$ then sup $[l', r']$ is less than sup$[l, r]$. An interval $[l, r]$ is said to be a *closed frequent interval* (*CFI*), if it is a frequent interval and it is closed.

3.2 Theoretical Background

When an interval $[l\ r]$ is added to the database then each end point present in the database can be marked as "old_end" or "new_end". An end point is called "old_end" if it was present in the database before adding the new interval $[l\ r]$; otherwise it is called "new_end". Each closed frequent interval, which is created after adding $[l, r]$, is mentioned here as newly generated closed frequent interval.

For the sake of completeness, the following two theorems (T1 and T2) that were proved in [4] have been stated below.

T1: If $[l\ r]$ is an interval in a dataset, then all the intervals $[l'\ r]$ are closed intervals, if $l \leq l' \leq r$ and part(l') $\geq r$ where $[l'\ \text{part}(l')]$ is another input interval.
T2: If $[l\ r]$ is an interval in a dataset, then all the intervals $[l\ r']$ are closed intervals, if $l \leq r' \leq r$ and part(r') $\leq l$ where $[l'\ \text{part}(l')]$ is another input interval.

Some properties of the newly generated closed frequent intervals are stated below and proved as theorems.

Theorem 1 *If $[l, r]$ is an interval added to the database and if $[l'\ r']$ is a newly generated closed frequent interval then $[l'\ r'] \subseteq [l\ r]$ i.e. $l \leq l'$ and $r' \leq r$.*

Proof If $[l'\ r']$ is a newly generated closed frequent interval, then either of the following two cases has to be true. The first case is that $[l'\ r']$ was a closed interval earlier but was not frequent. In this case, the newly added interval should increase the support of $[l'\ r']$ and so $[l'\ r'] \subseteq [l\ r]$. In the second case, $[l'\ r']$ was not closed earlier. In this case, either l' or r' should not be present as end points earlier. If l' was not present then $l' = l$ and if r' was not present, then $r' = r$ will have to be true. Now, from properties of closed intervals [4] if $[l' = l, r']$ is closed then $r' \leq r$ and if $[l', r' = r]$ is closed then $l \leq l'$. Therefore, in either case $[l'\ r'] \subseteq [l\ r]$.

Theorem 2 *A closed interval $[l'\,r']$ such that $[l'\,r'] \subseteq [l\,r]$ and l', r' are "old_end", will be a newly generated closed frequent interval if support of $[l'\,r']$ is equal to user-defined minimum support.*

Proof Given, $[l'\,r']$ is a closed interval such that $[l'\,r'] \subseteq [l\,r]$ and l', r' are "old_end". Also $[l'\,r']$ is newly generated closed frequent interval. Suppose support of $[l'\,r']$ is less than user-defined minimum support, then $[l'\,r']$ is not a frequent interval, which contradicts that $[l'\,r']$ is newly generated closed frequent interval. Suppose, support of $[l'\,r']$ is greater than minimum support. Then before adding $[l\,r]$ support of $[l'\,r']$ is at least equal to minimum support and hence $[l'\,r']$ is being frequent or closed frequent before adding $[l\,r]$. If $[l\,r]$ was closed frequent before adding $[l\,r]$, then it is already in closed frequent list, which contradict that $[l'\,r']$ is newly generated closed frequent interval. Otherwise if $[l'\,r']$ is frequent before adding $[l\,r]$ to become closed frequent interval at least one l' or r' must be a "new_end", which contradict that $l'\,r'$ are "old_end" and thus it contradict that $[l'\,r']$ is newly generated closed frequent interval. Thus, support of $[l'\,r']$ is equal to minimum support.

Next suppose, $[l'\,r']$ is an already closed frequent interval. Now, support of $[l'\,r']$ is equal to minimum support. Thus, before adding $[l\,r]$, support of $[l'\,r']$ must be less than minimum support, i.e., $[l'\,r']$ must be infrequent, because $l \leq l' \leq r' \leq r$. This contradicts that $[l'\,r']$ is an already closed frequent interval. Thus, $[l'\,r']$ is newly generated closed frequent interval if support of $[l'\,r']$ is equal to minimum support.

Theorem 3 *If $[l'\,r']$ is a newly generated closed frequent interval such that both l', r' are "old_end", then there is no other newly generated closed frequent interval $[l''r'']$ such that $[l''r''] \subset [l'r']$.*

Proof Suppose $[l''r'']$ is newly generated closed frequent interval. Then support of $[l''r'']$ must be greater than minimum support, because $[l''r''] \subset [l'r']$ and support of $[l'r']$ is equal to minimum support. Thus, support of $[l''r'']$ must be at least equal to minimum support before adding $[l\,r]$, which means $[l''r'']$ is an already closed frequent interval. This contradicts the above assumption. Thus, $[l''r'']$ is not a newly generated closed frequent interval.

Theorem 4 *If $[l'\,r']$ is a newly generated closed frequent interval such that l' is a "new_end" and r' is "old_end", then there will not be any newly generated closed frequent interval $[l''r']$ such that $[l''r'] \subset [l'r']$ where l'' is "old_end".*

Proof Suppose, $[l''r']$ is newly generated close frequent interval. Then support of $[l''r']$ must be equal to minimum support because l'', r' are "old_end" (Theorem 2).

Again $\begin{bmatrix} l' & r' \end{bmatrix}$ is a newly generated closed frequent interval and l' is a "new_end". Thus, support of $\begin{bmatrix} l' & r' \end{bmatrix}$ must be greater than equal to minimum support.

Again, $\begin{bmatrix} l'' & r' \end{bmatrix} \subset \begin{bmatrix} l' & r' \end{bmatrix}$. Thus, support of $\begin{bmatrix} l'' & r' \end{bmatrix}$ must be greater than minimum support. Thus, $\begin{bmatrix} l'' & r' \end{bmatrix}$ is not a newly generated close frequent interval because l'', r' are "old_end" (Theorem 2). This contradicts the assumption. Thus $\begin{bmatrix} l'' & r' \end{bmatrix}$ is not a newly generated closed frequent interval.

Theorem 5 *If $\begin{bmatrix} l' & r' \end{bmatrix}$ is a closed frequent intervals before adding a new interval $[l\ r]$ into the database then $\begin{bmatrix} l' & r' \end{bmatrix}$ remains closed frequent after adding the new interval $[l\ r]$.*

Proof It is obvious.

4   Problem Definition

Suppose ID_{old} denotes an interval database. CFI_{old} denotes set of closed frequent intervals present in ID_{old}. If an interval $I = [l, r]$ is added to ID_{old} then the updated database is denoted as ID. Suppose CFI denotes the set of closed frequent intervals present in the database ID. Then, $CFI = CFI_{old} \cup CFI_{new}$. CFI_{new} denotes all the newly generated closed frequent intervals after adding I to ID_{old}. If there is no such interval then $CFI = CFI_{old}$. The aim of the proposed method is to mine all these newly generated closed frequent intervals after adding an interval I to the interval database ID_{old}. It is to be mentioned that minimum support ρ here is the absolute minimum support and not any percentage value. So if an interval was frequent in ID_{old} it will remain frequent in ID as well.

Example 1 Suppose at some point of time, the interval database holds the following input intervals—[3 10], [4 10], [11 15], [15 17], [3 10]. After adding an interval [4 13] in the database the information in the array structure, *Interval_Info_Array* will be

Index	0	1	2	3	4	5	6	7	8	9
End	3	4	4	10	10	11	13	15	15	17
Link	3	4	6	0	1	7	2	5	9	8
Freq	2	1	1	2	1	1	1	1	1	1

If minimum support is 2 then before adding [4 13], the closed frequent intervals are [3 10], [4 10]. After adding [4 13], closed frequent intervals will be [3 10], [4 10], [11 13]. Thus, [11 13] is newly generated closed frequent intervals.

5 Proposed Method

The proposed method is designed to mine closed frequent intervals incrementally from a collection of intervals which grows over time. When a new interval is added to the collection, then the proposed method will mine all the newly generated closed frequent intervals. From the theorems T1 and T2, it is easy to see that a closed frequent interval in the new dataset will be a newly generated closed frequent interval if either its left end is l or its right end is r or it is properly contained in $[l, r]$ and has support equal to minimum support.

5.1 Array Structure to Store a Set of Intervals

Information of all the intervals present in the interval database are stored in an array structure called Interval_Info_Array as shown in Example 1 (Sect. 5.2). This Interval_Info_Array has four fields: Index—it is array index, End—it keeps end points of all distinct intervals in ascending order, Link—for each end e, in link field the index of the associated end is kept, i.e., if e is a left end of an interval [e r] then Link for e will store the Index of its right end r, Freq—it denotes the number of occurrences or frequency of an interval in the database. If an interval [l r] is present 2 times in the database, then frequency for both ends l and r will be 2. In [3, 11], this array structure is used to store intervals from a static interval database. So for the dynamic interval database, when a new interval is added to the database, then both the ends of the interval is added to the structure by searching for the appropriate position in the structure along the field "End". After finding the appropriate position, the end is stored by shifting the columns to one position right. In this process, the corresponding values in "Link" are also modified if it is necessary. If an interval with same ends is already available, then only the frequencies for the ends are increased by one. Thus, adding an interval to the structure will take $O(n)$ time. In this structure, an end e is a left end, if value in "Link" is greater than its Index; otherwise e is a right end.

5.2 Algorithm

The following algorithm is proposed to mine the newly generated closed frequent intervals from an interval database. When a new interval I = [l, r] is added to the database, first the interval I = [l, r] is added to the structure *Interval_Info_Array* and also find whether l and r are "old_end" or "new_end". Then, the following algorithm is applied on the structure *Interval_Info_Array* and it takes only $O(n)$ time to find all the newly generated closed frequent intervals. Because to find all the newly generated closed frequent intervals, it scans the ends in the structure from left to right only once.

Before going to the steps of the algorithm, here is the description on how the algorithm mines the newly generated closed frequent intervals

1. The "End" field of the structure *Interval_Info_Array* will contain all the left and right ends of the intervals in sorted order.
2. As in Theorem 1, the newly generated closed frequent intervals are contained in $[l\ r]$.
3. So, first we scan the ends in structure from left to right to get the appropriate position for left end l and after that insert l in its position by sifting ends on the right side of it. Note the position of l and update the "Link" of the partners of all subsequent ends. Similarly, the right end r can also insert into the structure.
4. After that, for each left end l', $l \leq l' \leq r$, we will try to find all the right ends r', $l \leq r' \leq r$, for which $[\ l'\ r']$ are newly generated closed frequent intervals using the Theorems 2 and 3.
5. We use a variable "count" [4] to keep number of intervals that contain the left end l' at some point during execution.
6. The searching for right ends for l' will start if

   (a) "count" for l' is greater than equal to minimum support because if an interval is frequent then its ends are also frequent, i.e., number of intervals containing that end is also greater than minimum support
   (b) Current right end is less than r (Theorem 1).
   (c) Maximum right end (right_max) associated with l' is greater than equal to the current right end (from the theorem of closed interval in [4]).

7. Searching for right ends l' will be done sequentially from left to right in the structure. In this process, each time a right end r' is found, the algorithm checks whether $[l'\ r']$ is newly generated closed frequent intervals or not as in Theorems 2 and 3. After that if left end of r' is less than l', then it decrements the frequency value of r' from "count". Such a right end denotes ending of intervals which contain l'. The searching for right ends for l' will continue till the above conditions are true. Thus, till the last right end (e') to which the algorithm proceed, "count" value of l' will be greater than equal to minimum support. If l', e' are "old_end" then $[l'\ e']$ is newly generated closed frequent interval if support is equal to minimum support (Theorem 2), otherwise it is already a closed frequent interval. Again, if l' is a "new_end" then $[l'\ e']$ is a newly generated closed frequent interval (Theorems 1 and 2). We will not get any newly generated closed frequent intervals in $[l'\ e']$ (Theorems 4 and 5). Thus, when searching for right end for l' stops, the algorithm proceeds to find next left end of l' and takes the next left end as l'. Searching for the right ends starts from the last right end (excluding it) or from this next left end if last right end is less than it. If the last right end (e') is equal to r and r is a "new_end" then we need not find the right ends for the next left ends (Theorem 1). We just have to check whether $[l'\ e']$ is newly generated closed frequent interval or not. This process will be continue until l' becomes greater than equal to r.

The steps of the algorithm are

```
1. count=0; right_max=0;
2.  for (x=l_min; x ≤ l; x= next_end)//l_min: minimum left end
 2.1 if ( x is a left end )
  { count = count + x.Freq;
    if (x==l && (x.Link).End > right_max)
       right_max=(x.Link).End;
  } else count = count - x.Freq;
3. e'.End = x  and e'.Index= Index of x;
4. for (l'=x; l' ≤  r;  find_next_left_end(l'))
 4.1 while (count  ≥  min_sup)
  4.1.1 if (e'.End > r ) break;
  else
  {4.1.1.1 if (right_max  >  e'.End)
   a)if((e'.End==r&&e'.End is a left end)||(e'.End<r))
     find right end  r' next to e' (e'.End ≤ r' ≤r );
   b) if [l'  r'] is closed
    b.1 if(l'is'old_end'&&r'is'old_end'&&count==minsup)
         [l',r']is a newly generated CFI
    b.2if((l'is'new_end'||r'is'new_end')&&(count≥minsup))
         [l',r'] is a newly generated CFI
    b.3 if (r' is 'new_end')break;
     else{next_end=r';
        do {e'.End=next_end;
              e'.Index = next_end.Index;
              if ((e'.Link).End ≤ l' )
                 count = count -e'.Freq ;
              next_end = next end of e';
             }
        while(e'.End==next_end&&next_end is right end);
   else
     { count = count_old;
       break ;}
    }
4.2if(e'.End==r&&e'.End is a right end&&r is'old_end')
   break;
find_next_left_end(l')
5. next_end=next left end of the left end l';
6. if (next end ≤ r)
6.1.if there is a right end r''such that e'.End<r''<l'_next
     count = count -  r'.Freq' ;
6.2. count_old=count;
6.3. right_max=0;
6.4. do {  l'=next_end;
       if((l'.Link).End > right_max)
          { right_max=(l'.Link).End; }
       count = count + l'.Freq
       next_end= next end of the left end l'.
        }
     while (( l' == next_end )&& next_end is left end);
 6.5. if (  l' >  e'.End  )
     { e'.End=l'; e'.Index=l'.Index; }
```

Table 2 Newly generated closed frequent intervals for the feature Gills of Fish dataset when minimum support is 2

Sl. no.	Interval added to the database	Newly generated closed frequent intervals
1	[393.71 853.1]	NIL
2	[354.22 1976.38]	[393.71 853.1]
3	[270 823.23]	[354.22 823.23], [393.71 823.23]
4	[0 145.19]	NIL
5	[20 41.3]	[20 41.3]
6	[85.57 199.83]	[85.57 145.19]
7	[38.16 186.22]	[38.16 41.3], [38.16 145.19], [85.57 186.22]
8	[55.12 98.65]	[85.57 98.65], [55.12 98.65]
9	[2.8 147.71]	[2.8 145.19], [85.57 147.71], [38.16 147.71]
10	[0 9.45]	[0 9.45], [2.8 9.45]

6 Experimental Results

For the purpose of experimentation, the proposed method is applied on the fish interval dataset available in https://lhedjazi.jimdo.com/useful-links.

Fish Dataset

This dataset contains 12 species of freshwater fish, each one being described by 13 interval features {Length, Weight, Muscle, Intestine, Stomach, Gills, Liver, Kidneys, Liver/muscle, Kidneys/muscle, Gills/muscle, Intestine/muscle, Stomach/muscle}.

Results: For the purpose of mining closed frequent intervals incrementally, we consider the interval feature Gills and minimum support as 2. Two data with empty intervals (i.e., [0 0]) are removed as it has no meaning for the purpose of closed frequent interval mining. The newly generated closed frequent intervals after input the intervals one by one are shown in Table 2.

7 Conclusion

In this paper, an algorithm has been proposed to mine closed frequent intervals from an interval database. Previous methods [3, 5, 6] on mining closed frequent intervals assume that the interval database is static. However in real life, we hardly come through static databases. If an algorithm is not incremental in nature then whenever new data comes in the whole mining process has to be done afresh. The algorithm proposed here is an incremental algorithm to mine closed frequent intervals from a dynamic interval database. Each time an interval is added to a database with n intervals, the proposed method finds all the newly generated closed

frequent intervals in O(n) time. Thus, it will take O(n2) time to find all the closed frequent intervals from a database with n intervals. Thus, the time complexity of this method is better than the methods proposed in [5] which is $O(n^4)$ and in [3] which is O(n2lg n). Its complexity is same as the method proposed in [6] which was however not incremental in nature.

References

1. Allen, J.F.: Maintaining knowledge about temporal intervals. Commun. ACM **26**(11), 832–843 (1983)
2. Lin, J.L.: Mining maximal frequent intervals. In: Proceedings of ACM Symposium on Applied Computing, New York, USA, pp. 426–431 (2000)
3. Dutta, M.: Development of efficient algorithm for some problems in interval data mining. Ph.D. thesis, Department of Computer Science, Gauhati University, India (2012)
4. Sarmah, N.J., Mahanta, A.K.: An incremental approach for mining all closed intervals from an interval database. In: IEEE International Advance Computing Conference (IACC), pp. 529–532 (2014)
5. Mahanta, A.K., Dutta, M.: Mining closed frequent intervals from interval data. Int. J. Appl. Sci. Adv. Technol. **1**(1), 1–3 (2012)
6. Sarma, N.J.: Study and design of algorithms for certain problems in interval data mining. Ph.D. thesis, Department of Computer Science, Gauhati University, India (2016)
7. De Souza, R.M.C.R., De Carvalho, F.D.A.T.: Clustering of interval data based on city–block distances. Pattern Recogn. Lett. **25**, 353–365 (2004)
8. Gowda, K.C., Diday, E.: Symbolic clustering using a new dissimilarity measure. Pattern Recogn. **24**(6), 567–578 (1991)
9. Ichino, M., Yaguchi, H.: Generalized Minkowski metrics for mixed feature-type data analysis. IEEE Trans. Syst. Man Cybern. **24**, 698–708 (1994)
10. Roh, J.W., Yi, B.K.: Efficient indexing of interval time sequences. Inf. Process. Lett. **109**(1), 1–12 (2008)
11. Sarmah, N.J., Mahanta, A.K.: An efficient algorithm for mining maximal sparse interval from interval dataset. Int. J. Comput. Appl. **107**(16), 28–32 (2014)

Crime Pattern Analysis by Identifying Named Entities and Relation Among Entities

Priyanka Das and Asit Kumar Das

Abstract The present work proposes an unsupervised method for identifying named entities from a corpus of crime reports containing information on crime against women in Indian states and union territories and subsequently discovers substantial relations among the identified named entities. For discovering the relations, different types of entity pairs have been chosen and similarity among them has been measured based on the intermediate context words. Depending on the similarity score, clustering technique has been applied that forms several clusters of named entity pairs. Each cluster consists of a representative entity pair and relation of that representative pair corresponds to the relation of the whole cluster formed, leading to the relational labelling of the clusters. This method does not desire any time consuming richly annotated corpora and the result with high F-measure values depicts the effectiveness of this method.

Keywords Crime analysis · Named entity recognition · Relation extraction

1 Introduction

Recognising named entities from textual datasets and subsequently discovering unrestricted relationships among them is an essential part for any information extraction task. Though most of the blogs, newswire, websites, and governmental documents provide information for crime, human annotation for these colossal

P. Das (✉) · A. K. Das
Indian Institute of Engineering Science and Technology,
Shibpur, Howrah 711103, India
e-mail: priyankadas700@gmail.com

A. K. Das
e-mail: akdas@cs.iiests.ac.in

volumes of heterogeneous data sources is a tedious task. Crime pattern analysis from textual data not only requires Natural Language Processing (NLP) applications [1], but also utilises the text mining techniques with the domain knowledge. Newspapers [2] containing crime information have been considered from which entities (name of certain things) [3, 4] have been detected and these entities helped in the crime pattern analysis. Not only the named entity recognition, but significant research work exists on extracting relevant relation among the entities. Initially, the project called 'ASTREA' [5] developed a relation extraction system for crime investigation. Basili et al. proposed a method called 'REVEAL' [6] that employs variants of support vector machine for automatic relation extraction in case of crime investigation. Yao et al. [7] suggested a series of probabilistic models that creates triples of entity pairs and a syntactic dependency path is generated forming groups of same textual expressions. A hybrid model called 'Powerset Factz' [8] was developed that extracts relationships among different types of entities by linguistic analysis and semantic typing information from databases. The model has proven to give outperforming results when compared to other methods based on entropy label ranking. Unsupervised relation discovery followed by developing particular patterns for the discovered relations was done in Akbik et al. [9].

The present work accentuates on recognition of named entities and determining relation between the entities. Initially, a large corpus of crime reports have been extracted from online newspapers with crime reports against women in 29 states and 4 union territories of India. After data collection, all the necessary preprocessing is done. Then named entities are identified by noun phrase chunking (NP-chunking) that recognises the names of several entities like organisations, places such as states, cities, persons, victims, offenders involved in crime, and many more. Upon identifying the named entities, the goal is to discover the relation among the identified entities by application of clustering technique. Inspired by the approach of relation discovery by Hasegawa et al. [10], the present work extends the relation determination task in the field of crime analysis where, different types of named entity pairs (NE-pairs) have been chosen from the corpus and similarity has been measured among the NE-pairs with the intermediate words as context vectors. A complete graph is generated based on the similarity score and finally clustering technique has been applied that forms different groups of entity pairs which are similar in context. NE-pairs having same context falls in one cluster and vice versa. Each cluster has a representative pair in it and the relation of that representative also defines the relation of the corresponding cluster. The relations are defined as the crime types like rape, abduction, murder, molestation, sexual harassment, abuse, dowry harassment, domestic violence, etc.

The rest of the paper is organised as follows: Sect. 2 describes the proposed work in details. Section 3 shows the results of the proposed method followed by conclusion and future work in Sect. 4.

2 Proposed Framework

This section describes the present methodology in a skilful way. The objective of this work is to identify the meaningful entities from the dataset and find substantial relation among the entities. The detailed techniques are discussed as follows:

2.1 Data Collection and Preparation

Few online classified newspapers like, The Times of India, The Hindu, The Indian Express and websites like www.asianage.com, www.tribuneindia.com, www.dnaindia.com and www.iamnirbhaya.me, [11] providing information of crime against women in India and its union territories, are considered for extracting the data.

(A) Web Crawl: A customised Python program is implemented that discards the unnecessary details present in the webpages and brings out the useful data from the above-mentioned e-papers. The extracted data contains crime reports for 29 states and 4 union territories of India for over a time period of July 2004 to December 2016 and these data are stored for further analysis. Each state has nearly about 850 crime reports constituting news of different types of crime.

(B) Preprocessing: Initially, the data was collected and stored for each state, but it was observed that many reports got extracted only with a name of a district or city. So, again a customised Python program is used that employs the Google Geocoding API which automatically sorts the crime reports of the districts and cities according to the particular states. Other necessary preprocessing like stopword removal and stemming has been done using analytical applications of Natural Language Processing.

2.2 Named Entity Recognition

The present work considers the statewise sorted dataset and with the help of a sentence segmenter, the text of the dataset is split into several sentences. Once the sentences are segmented, word-level tokenisation provides multiple tokens. Then, noun phrase chunking (NP-Chunking) is performed that searches for chunks related to each noun phrase. Often a single noun phrase in the dataset consists of two or more nested noun phrases, but the program for the present framework only deals with the simpler chunks which are mostly smaller than the complete noun phrases. NP-Chunking is mostly done by tagging the sentences with parts of speech. Here, the method is defined with a chunk grammar that sets the rule for chunking with regular expressions present in the dataset. The rule for the NP-Chunker searches for the

presence of any article or determiner in the sentence followed by adjectives and then any noun. Thus, the method detects the named entities relating to name of various person, organisations, states, geopolitical entities like any area, streets, towns/cities or villages, time, money, and many more. The work considers all possible subtypes of the entities as mentioned in Sekine et al. [12] present in the dataset.

2.3 Similarity Among Entity Pairs

Even though several domains of entity pairs exist, only five different domains of NE-pairs namely, ORG-LOC (organisation-location), PER-LOC (person-location), PER-ACT (person-action), ORG-TIME (organisation-time), and PER-PER (person-person) are considered that can provide a better visualisation of the crime. In each case, relation of the first one has been detected with the latter. Figure 1 shows the framework in brief. The accumulation has been done separately for each domain of the pairs and the presence of all stemmed intermediate words is considered as context of the pair of entities. Here, the order of occurrence of the pairs also has been taken into consideration. For example, (E_1, E_2) and (E_2, E_1) where E_1 and E_2 are distinct
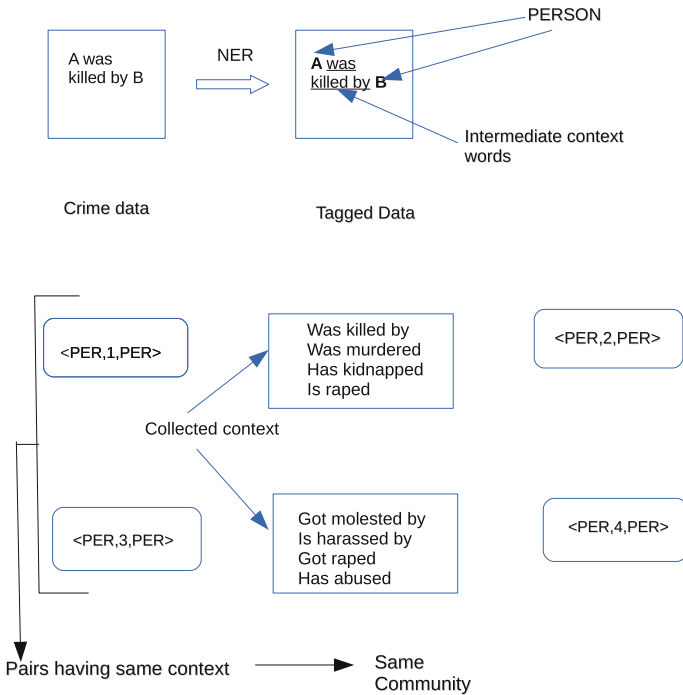
Fig. 1 Proposed framework in brief

named entities with different context present in different crime reports. Now for each NE-pair, a context vector is created whose values are generated by calculating the term frequency inverse document frequency (TF-IDF). Let R be an extracted crime report, a context word c and an individual report $r \in R$, then the value for c in r is calculated using (1).

$$c_r = f_{c,r} * log(|R|/f_{c,r}),$$ (1)

where $f_{c,r}$ is the frequency of c in report R.

Then cosine similarity has been measured for the context vectors of entity pairs using (2). For example, it compares one PER-LOC pair with other PER-LOC and same as in case of other pairs of different domains.

$$Cos(X, Y) = \frac{X \cdot Y}{\|X\| \|Y\|},$$ (2)

where X and Y are two context vectors in a particular domain of NE-pairs.

Based on the similarity score, a complete weighted undirected graph known as *Entity Similarity Graph*, $G = (N, E)$ is formed, where N represents set of nodes of the graph and E, the set of edges connecting the nodes. Here, n_1 and n_2 are two nodes in N, representing two context vectors, say X and Y, whereas, the edge between the nodes n_1 and n_2 has a weight reciprocal of the similarity score between X and Y. The graph is made sparse based on a threshold $t_i(G)$ associated to each node $n_i \in N$ in graph G. For each node, n_i, $t_i(G)$ is set as the average similarity of it with other nodes in the graph and the edge incident with this node is removed if its associated weight is more than the average value. But for node in other terminal of edge, this value may be higher compared to its associated threshold. But it is not allowed with the assumption that relationship should be established from both ends. Otherwise, it will not persist or will not be stable. All the entity pairs have been considered as nodes. Higher the similarity among the pairs, lesser is the weight of the corresponding edges.

2.4 Relation Extraction

Once the sparse weighted *Entity Similarity Graph* is created, clustering technique has been incorporated for the relation discovery task. Infomap clustering algorithm [13] has been applied on the graph that forms different clusters of NE-pairs similar in context. Thus, the pairs which possess similar context words between them are grouped in one cluster and pairs having different context falls in different clusters. This algorithm also provides a representative for each group of entity pairs which leads all the pairs in the cluster. Now, the aim is to label the clusters according to the relation which mainly refers to the crime types like cluster C_1 is labelled as 'Rape', cluster C_2 is labelled as 'Assault' and so on. The present work considers the relation of the representatives of each cluster for labelling the respective clusters. The representative of a cluster is a particular NE-pair and relation can be determined of that

representative pair with the help of prior knowledge of the context words present in the dataset. The relation of representative pair is chosen as the characterisation of the whole cluster from which the representative pair belongs.

3 Experimental Results

As discussed in the previous sections, the proposed framework has been applied on the extracted reports containing information of crime against women for 29 states and 4 union territories of India for over a time period of July 2004 to December 2016. For 33 states, there was a total of 28,050 crime reports based on different crime types. After few essential preprocessing, the dataset provides 3,36,600 number of sentences which are further split into 5,12,533 number of tokens. Then noun phrase chunking is done that gives near about 8,00,210 named entities (including the basic and subtypes). Table 1 shows the different domains of NE-pairs and their corresponding numbers as chosen from the dataset. For the present work, entity pairs with at least four intermediate context words have been chosen and pairs having less than four intervening words have been discarded. Along with more frequent pairs, less frequent entity pairs also have been chosen.

Then, cosine similarity has been measured among the pairs and based on the similarity score, a complete weighted undirected *Entity Similarity Graph* is generated where the entity pairs are represented as nodes and the connecting edges have a weight reciprocal to the similarity score. Infomap algorithm has been applied on this graph and Table 2 shows number of different clusters formed from NE-pairs having similar context words. Now, it would be unwise to provide a graph as an output of the clustering algorithm, as the figure looks clumsy with huge number of pairs. So, for a better visualisation of the Infomap clustering technique, few random pairs have been chosen for the PER-ACT domain and Fig. 2 shows the results where few entity pairs are divided into 12 clusters shown in 12 different colours and each cluster contains the NE-pairs as shown in the numbered nodes. The process has been repeated parallelly for other mentioned domains as well.

Now, each cluster has a representative NE-pair in it and its context words present in the dataset help in determining the relation of the representative pair and that particular relation is chosen as the characterisation of the whole cluster from which

Table 1 Details of the entity pairs

Domain	No. of pairs
ORG-LOC	2783
PER-LOC	5105
PER-ACT	4237
ORG-TIME	3114
PER-PER	5388

Table 2 Result of infomap
clustering technique

Domain	No. of clusters formed
ORG-LOC	52
PER-LOC	64
PER-ACT	41
ORG-TIME	43
PER-PER	67

Fig. 2 Sample result for
clusters of few randomly
chosen PER-ACT pairs
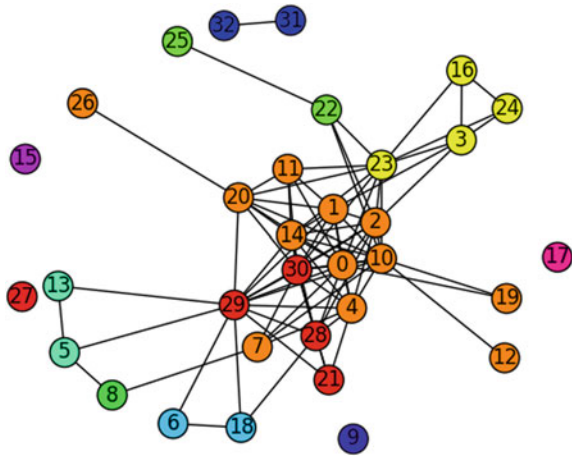
the representative pair belongs. Thus, the relation labelling has been done for all
of the clusters formed for each domain of pairs. The relations are defined as various
crime types like 'rape', 'murder', 'abduction' etc. Table 3 shows in detail the number
of entity pairs with their labelling of crime type.

The accuracy of the proposed framework is evaluated by recall, precision and
F-measure methods. The details of the cluster evaluation techniques for labelling of
NE-pairs are given in (3–5) as follows:

$$Precision\,(P) = \frac{Pairs\_correctly\_labelled}{All\_pairs\_labelled\_automatically} \tag{3}$$

$$Recall\,(R) = \frac{Pairs\_correctly\_labelled}{Pairs\_labelled\_manually} \tag{4}$$

$$F - measure\,(F) = \frac{2PR}{P + R} \tag{5}$$

Table 4 provides the information that this method gives better result for the PER-
LOC, PER-ACT and PER-PER domain than other domains of NE-pairs.

Table 3 Relation labelling for the named entity pairs

Relations (No. of NE-pairs)

Domain	Rape	Abduction	Murder	Sexual harassment	Human trafficking	Dowry harassment	Molestation	Others
ORG-LOC	510	237	383	427	275	341	300	310
PER-LOC	874	620	532	558	542	789	570	620
PER-ACT	630	521	610	587	389	609	514	377
ORG-TIME	427	389	255	413	421	633	424	152
PER-PER	854	745	604	621	653	632	677	602

Table 4 Results for Precision, Recall and F-measure in (%)

Evaluation metrics

Domain	Precision	Recall	F-measure
ORG-LOC	75	82	78
PER-LOC	85	81	82
PER-ACT	84	79	81
ORG-TIME	76	78	76
PER-PER	87	80	83

4   Conclusion and Future Work

The present work emphasised on an unsupervised approach of named entity recognition and relation extraction task from crime data. It used a clustering algorithm that forms clusters of entity pairs similar in context. The method also considered less frequent NE-pairs as few meaningful relations can be discovered from them too. It has been observed that entities have been identified with high recall and precision and relation labelling of the clusters have been done in an efficient manner. As a future work, further inferences can be drawn by realising paraphrases for crime pattern analysis.

References

1. Steven Bird, E.K., Loper, E.: Natural Language Processing in Python (2007)
2. Arulanandam, R., Savarimuthu, B.T.R., Purvis, M.A.: Extracting crime information from online newspaper articles. In: Second Australasian Web Conference (AWC 2014), vol. 155, 31–38, Jan. 2014
3. Shabat, H., Omar, N., Rahem, K.: Named entity recognition in crime using machine learning approach. In: Information Retrieval Technology, pp. 280–288, Dec. 2014
4. Shabat, H.A., Omar, N.: Named entity recognition in crime news documents using classifiers combination. Middle-East J. Sci. Res. **23**(6), 1215–1221 (2015)
5. IRSIG-CNR: Astrea, Information and Communication for Justice
6. Basili, R., Giannone, C., Del Vescovo, C., Moschitti, A., Naggar, P.: Kernel-based relation extraction for crime investigation. In: AI*IA, Citeseer, pp. 161–171 (2009)
7. Yao, L., Haghighi, A., Riedel, S., McCallum, A.: Structured relation discovery using generative models. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. EMNLP '11, Stroudsburg, PA, USA, pp. 1456–1466. Association for Computational Linguistics (2011)
8. Syed, Z., Viegas, E.: A hybrid approach to unsupervised relation discovery based on linguistic analysis and semantic typing. In: Proceedings of the NAACL HLT 2010 First International Workshop on Formalisms and Methodology for Learning by Reading. FAM-LbR '10, Stroudsburg, PA, USA, pp. 105–113. Association for Computational Linguistics (2010)
9. Akbik, A., Visengeriyeva, L., Herger, P., Hemsen, H., Löser, A., et al.: Unsupervised discovery of relations and discriminative extraction patterns. Citeseer

10. Hasegawa, T., Sekine, S., Grishman, R.: Discovering relations among named entities from large corpora. In: Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics. ACL '04, Stroudsburg, PA, USA, p. 415. Association for Computational Linguistics (2004)
11. I AM NIRBHAYA: www.iamnirbhaya.me. Accessed Mar 2014
12. Sekine, S., Sudo, K., Nobata, C.: Extended named entity hierarchy. In: Third International Conference on Language Resources and Evaluation (LREC-2002), pp. 1818–1824, Feb 2002
13. Rosvall, M.: Infomap. http://www.mapequation.org/code.html

A Secure High-Capacity Video Steganography Using Bit Plane Slicing Through (7, 4) Hamming Code

Ananya Banerjee and Biswapati Jana

Abstract Achievement of high-capacity data hiding using a digital media is an important research issue in the field of steganography. In this paper, we have introduced a novel scheme of data hiding directly within the video stream **using** bit plane slicing through (7, 4) Hamming code with the help of shared secret key. In the proposed scheme, a secret logo image is embedded within the cover video stream for authentication and ownership identification through Hamming code based video steganography. Each frame of secret video has been separated into individual three basic color blocks (R, G and B) and then partitioned into (3 × 3) pixel blocks. After that, each color block is sliced up into 4 bit planes starting from LSB plane. The pixels' positions of cover images are randomly selected by Pseudorandom Number Generator (PRNG) using a shared secret seed value and data embedding performed using (7, 4) Hamming code. As a result, 36 bits secret data can be embedded within a (3 × 3) pixel block which is almost eight times greater than Ramadhan and Khaled's scheme (Systems, applications and technology conference (LISAT), 2014 IEEE Long Island, 2014) [1]. Here, we achieve a high payload with good visual quality stego video. **Furthermore, the video compression is lossless so the video file size is strictly preserved for post-data embedding**.

Keywords Video steganography · Hamming code · Least significant bit (LSB)
Bit plane · Data hiding

A. Banerjee (✉) · B. Jana
Department of Computer Science, Vidyasagar University,
Midnapore 721102, West Bengal, India
e-mail: anaanya.2011@gmail.com

B. Jana
e-mail: biswapatijana@gmail.com

1 Introduction

Steganography is the art and science of hidden data communication. Till date, many data hiding schemes [2] are developed but only few of them are considered being more secured and have less distortion. The data hiding schemes are useful in many application areas to solve the problem of ownership identification, copyright protection, authentication, verification, and more. The main aims of data hiding schemes are to ensure extraction of secret data and recovery of original object from stego media. On the other hand, data should stay hidden in stego media even if the eavesdropper tampered the stego or degrading through natural phenomenon like transmission resampling, compression, or filtering, etc. The main drawbacks of data hiding schemes are not to provide a good solution in such cases. The degree of distortion will be high due to increase of data embedding capacity that should be balanced mathematically using spread spectrum. The data embedding in video is considered to be more unsuspicious and secured and less exploration has been done till today in this research area using Hamming code.

2 Related Work

Video hiding inside a video stream using nonuniform rectangular partition is done by Sheng et al. [3]. Then, another video hiding scheme is proposed by Yadav et al. [4] based on LSB technique which replaces the least significant bits of pixels selected to hide the secret information. Video as a collection of numerous frames has greater data hiding capacity as the small color change in the whole video stream is hard to detect in human eyes. Dasgupta et al. [5] proposed hash-based LSB techniques in spatial domain where the bits of the message can be inserted in intensity pixels of the video in LSB positions. Here, we have proposed video steganography using (7, 4) Hamming code for color images. We have divided R, G, B color pixels in bit plane [6] starting from LSB to LSB-3 (up to 4 bit plane) partitioned into (3×3) blocks and then apply Hamming code based data hiding scheme. In this scheme, 36 bits of data are embedded within nine pixels which is more higher than other existing LSB technique [1, 4, 5] and most of the LSB techniques are prone to attack [7, 8]. Also, it maintains high visual quality. Additionally to provide more security, the message is encrypted using symmetric key encryption techniques. Thus, we have achieved secure steganographic system for hiding data in video stream using both cryptography and steganography techniques.

3 Proposed Method

A video stream consists of collection of frames and the secret data is embedded in these frames as payload. The cover video is broken down into frames before embedding. Each frame is now considered as a cover image. Now, the proposed technique has been applied to conceal the secret data in the carrier frames.

Suppose, I is considered as the cover frame image of size ($M \times N$), and I' is the marked image after embedding data $D = \{d_1, \ldots, d_X\}$, where $d_i \in \{0, 1\}$, $1 \leq i \leq X$. Here, H is a parity check matrix of the Hamming code. Let H be

$$H = \begin{vmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{vmatrix},$$ Embedding capacity is an important

metric for data embedding. It is measured by the number of secret bits that can be embedded into a cover image. The embedding capacity is calculated as [2] $ER = L/M \times N$ bpp where L is the length of the secret message. Before embedding the secret data, we take 36 bit secret key k_1 which is known to both the sender and the receiver, to encrypt the secret data bit using symmetric key encryption. We have taken each pixel block of size (3×3) and 4 bit plane of each pixel is used to embed the data, which results in (3×3) $\times 4 = 36$ bits of data (D_1) in one iteration. As an additional security measure, instead of choosing the cover image pixel block serially, we will use Pseudorandom Number Generator (PRNG) function with a secret predefined seed k_2 (which is only known to the sender and the receiver) to determine the next available block for embedding. Since this seed will be known to the sender and receiver only, the generated unique pattern of pixel block selection can be used in embedding and extraction process securely. The data embedding procedure is enlisted in Algorithm 1 and the data extraction procedure is depicted in Algorithm 2 (Fig. 1).

Algorithm 1: Data embedding process

  Input: Cover video, secret data bits D, Hamming matrix H, secret key k_1, and seed value k_2
  Output: A stego video

  Step 1: Extract each frame from video stream as a color image I of size ($M \times N$).
  Step 2: Collect random sequence of pixel blocks of size (3×3) from cover image $I_{M \times N}$ using PRNG (k_2). Say the pixel blocks are $X_1, X_2, \ldots, X_{MN}$.
  Step 3: Convert X_i into three separate RGB color blocks X_{iR}, X_{iG}, and X_{iB}.
  Step 4: Convert each X_i's into binary form.
  Step 5: Perform bit plane slicing of each X_i 's up to 4 bit plane starting from LSB that is $X_{iR(LSB)}$, $X_{iR(LSB-1)}$, $X_{iR(LSB-2)}$, $X_{iR(LSB-3)}$.
  Step 6: Take $c = X_{iR(LSB)}$ and calculate the syndrome $S_1 = (H \times (c)^{T})^{T}$.
  Step 7: Perform $D_1' = (D_1 \oplus k_1)$; $k_1 = 36$ bit length and D_1 is also same length.

Fig. 1 Pictorial diagram of the proposed data hiding scheme

Step 8: Take 3 bits secret data $d_i = \{d_1, d_2, d_3\}$ from D_1' where $d_i \in \{0, 1\}$.

Step 9: Calculate $S_2 = (d_i \oplus S_1)$; if $S_2 = 0$, no change, otherwise flip a bit at the positional value of S_2 and generate H'.

Step 10: Compute $S_3 = (H' \oplus c)$ and store the data.

Step 11: Replace the matrix (c) with S_3 and update $X_{iR(LSB)}$.

Step 12: Repeat Step 4 to 10 using $X_{iR(LSB-1)}$, $X_{iR(LSB-2)}$ and $X_{iR(LSB-3)}$.

Step 13: Repeat Step 5 to 11 to embed secret data on X_{iG} and X_{iB} color blocks.

Step 14: Repeat Step 2 to 12 to embed secret data on each and every random sequence of $(X_i's)$ of pixel blocks.

Step 15: Finally, after combining each stego block, we get stego frame (I') of size $(M \times N)$.

Step 16: Generate stego video stream with encoded frames.

Step 17: End.

Algorithm 2: Data extraction process

Input: Stego video stream, Hamming matrix H, secret key k_1, and seed value k_2

Output: Original Secret Message D.

Step 1: Convert the video stream into frames. Extract each frame as a color cover image I' of size $(M \times N)$.

Step 2: Use PRNG with predetermined seed k_2 to determine the stego pixel of random sequence X'_i of size $[3 \times 3]$ from stego image I'.

Step 3: Separate RGB components into X'_{iR}, X'_{iG}, X'_{iB}.

Step 4: Convert into binary form of each X'_{iR}, X'_{iG}, and X'_{iB}.

Step 5: Perform 4 bit plane slicing of each X'_i's starting from LSB that is $X'_{iR(LSB)}$, $X'_{iR(LSB-1)}$, $X'_{iR(LSB-2)}$, $X'_{iR(LSB-3)}$.

Step 6: Take $c' = X'_{iR(LSB)}$ and calculate the syndrome $S' = (H \times (c')^T)^T$.

Step 7: Concatenate syndrome S' with data unit of D' that is $D' = D' \| (S')$.

Step 8: Repeat Steps 4 to 6 using X'_{iG} and X'_{iB}.

Step 9: Compute $D_i = D' \oplus k_1$.

Step 10: Repeat Steps 2 to 8 using next random sequence of X_i block.

Step 11: Concatenate D_i's, we get original secret message D.

Step 12: End.

3.1 Numerical Illustration

Example 1 Data Embedding

1. Let I is a color pixel block with (3×3) pixel. $D = \{d_1, d_2, …, d_{36}\} = \{0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0\}$. $k_1 = \{0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0\}$ and $ER = 36/(3 \times 3) = 4$ bpp and $D' = D \oplus k_1 = \{0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0\}$

2. Divided into three RGB pixel blocks shown below.

$$R = \begin{vmatrix} 141 & 141 & 141 \\ 141 & 142 & 142 \\ 140 & 143 & 143 \end{vmatrix} \quad G = \begin{vmatrix} 143 & 141 & 142 \\ 141 & 144 & 145 \\ 143 & 146 & 147 \end{vmatrix} \quad B = \begin{vmatrix} 147 & 145 & 145 \\ 145 & 149 & 150 \\ 146 & 150 & 151 \end{vmatrix}$$

3. Take red pixel block and transform into binary number matrix.

$$R = \begin{vmatrix} 10001101 & 10001101 & 10001101 \\ 10001101 & 10001110 & 10001110 \\ 10001100 & 10001111 & 10001111 \end{vmatrix}$$

4. Divide it into 4 bit plane matrices starting from LSB.

$$R_{LSB} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{vmatrix} \quad R_{LSB-1} = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix} \quad R_{LSB-2} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

$$R_{LSB-3} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

5. Read the LSB matrix and form a 1D matrix. c = [1 1 1 1 0 0 0 1 1]
6. Calculate the syndrome

$$S_1 = H \times (c)^T = \begin{vmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{vmatrix}$$

$$\times \begin{vmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{vmatrix}^T = \begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix}$$

7. Transpose the syndrome and XOR with the secret data bit, i.e., [1 0 1] ⊕ [0 1 1] = [1 1 0] which matches with the fifth column of Hamming matrix.
8. Generate the code H' = [0 0 0 0 1 0 0 0 0] and XOR with the original code c.

S_3 = [1 1 1 1 0 0 0 1 1] ⊕ [0 0 0 0 1 0 0 0 0] = [1 1 1 1 1 0 0 1 1].

9. Transform into a new LSB matrix.

$$R'_{LSB} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix}$$

10. Similarly compute the LSB-1, LSB-2, and LSB-3 matrices as follows:

$$R'_{LSB-1} = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} \quad R'_{LSB-2} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix} \quad R'_{LSB-3} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

11. Update all four modified binary matrices to their corresponding position in original RED pixel matrix.

$$R'_{3\times3} = \begin{vmatrix} 10001101 & 10001101 & 10001101 \\ 10000101 & 10001111 & 10001100 \\ 10001100 & 10001111 & 10001011 \end{vmatrix} = \begin{vmatrix} 141 & 141 & 141 \\ 133 & 143 & 140 \\ 140 & 143 & 139 \end{vmatrix}$$

12. Similarly get updated green and blue pixel matrices.

Example 2: Data Extraction

1. The marked frame pixel block is divided into three RGB color pixel blocks.

$$R = \begin{vmatrix} 141 & 141 & 141 \\ 133 & 143 & 140 \\ 140 & 143 & 139 \end{vmatrix} \quad G = \begin{vmatrix} 143 & 141 & 134 \\ 143 & 144 & 149 \\ 143 & 146 & 146 \end{vmatrix} \quad B = \begin{vmatrix} 147 & 145 & 149 \\ 145 & 149 & 150 \\ 147 & 156 & 151 \end{vmatrix}$$

2. Take red image pixel block and transform into binary numbers.

$$\begin{vmatrix} 10001101 & 10001101 & 10001101 \\ 10000101 & 10001111 & 10001100 \\ 10001100 & 10001111 & 10001011 \end{vmatrix}$$

3. Divide it into 4 bit plane matrices starting from LSB.

$$R_{LSB} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} \quad R_{LSB-1} = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} \quad R_{LSB-2} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix}$$

$$R_{LSB-3} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

4. Read LSB matrix and form a 1D matrix. c = [1 1 1 1 1 0 0 1 1]
5. Calculate the syndrome $S_1 = H \times (c)^T =$

$$
\begin{vmatrix}
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1
\end{vmatrix}
\times \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{vmatrix}^T =
\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}
$$

6. Transpose the syndrome to get secret data bits d = [0 1 1]
7. Repeat the above steps until we do not get the secret data bits. Concatenate all the data bits to get the data, that is, D' = {0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0}.
8. XOR the modified secret data with secret key k_1 to get the original secret data bits that is D = {0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0}.

4 Experimental Result and Comparison

The scheme is implemented using NetBeans IDE 8.0 on standard color images to measure the performance. The standard cover video sequences are Audio Video Interleave (AVI) format with the size of (192 × 352). These are collected from the video database of department of computer science at University of Mannheim [9]. The secret message is a binary image logo. Upon extraction, the secret data is retrieved without any loss or noise. The qualities of the stego frame images are measured using Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR).

$$
MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[I(i,j) - I'(i,j) \right]^2 \tag{1}
$$

$$
PSNR = 10 \log_{10} \frac{255^2}{MSE} (dB) \tag{2}
$$

Following tables are representing the PSNR values of stego videos generated by our algorithm with varying payload. The average capacity of data embedding per frame is 16.5 KB (when ER = 2 bpp), 24.68 KB (when ER = 3 bpp), 32.9 KB (when ER = 4 bpp) which is almost eight times that of other existing algorithms [1]. The standard video sequences are 25 fps. So, we can embed maximum 823 KB data per second which is a high payload compare to any other algorithm [5] (Tables 1, 2, 3 and 4).

Table 1 Result of PSNR, SSIM from our proposed method when payload 2 bpp

	Akiyo	News	Cars	Train	Flamingo	Meer kat	Highway	Ice	Football
PSNR	50.17	50.17	50.18	50.18	50.17	50.18	50.17	50.18	50.17
SSIM	0.9936	0.9944	0.9928	0.9969	0.9932	0.9977	0.9966	0.9926	0.9976

Table 2 Result of PSNR, SSIM from our proposed method when payload 3 bpp

	Akiyo	News	Cars	Train	Flamingo	Meer kat	Highway	Ice	Football
PSNR	43.90	43.92	43.94	43.94	43.95	43.95	43.94	43.91	43.93
SSIM	0.9740	0.9776	0.9714	0.9876	0.9727	0.9908	0.9862	0.9705	0.9904

Table 3 Result of PSNR, SSIM from our proposed method when payload 4 bpp

	Akiyo	News	Cars	Train	Flamingo	Meer kat	Highway	Ice	Football
PSNR	37.81	37.82	37.87	37.86	37.93	37.88	37.87	38.04	37.85
SSIM	0.9065	0.9190	0.8985	0.9538	0.9020	0.9652	0.9501	0.8976	0.9636

Table 4 The performance comparison of different schemes under different video sequences

Sequence	Maximum capacity (kbits/s)			PSNR (dB)			SSIM		
	Xu et al.'s [10]	Wang et al.'s [3]	Proposed scheme	Xu et al.'s [10]	Wang et al.'s [3]	Proposed scheme	Xu et al.'s [10]	Wang et al.'s [3]	Proposed scheme
Stefan	17.8017	29.6121	6584	38.33	37.13	37.81	0.9825	0.9811	0.9665
Table	8.2683	12.7662	6590	37.91	37.39	37.82	0.9537	0.9526	0.9537
Mobile	1.8054	5.5506	6586	38.32	37.89	37.87	0.9871	0.9867	0.9567
Hall	0.5991	0.9696	6588	40.26	40.16	40.12	0.9743	0.9742	0.9788
News	0.4956	0.8181	6585	40.75	40.68	37.82	0.9848	0.9847	0.9190

5 Security Analysis

Security analysis is an important key factor of data hiding process. In this paper, we have used two levels of security to enhance our proposed scheme from security perspective. First, we take a 36 bits secret key and encrypt the secret data bits using symmetric key encryption. As it is only known to the sender and receiver, the third party will not be able to decrypt it without knowing the secret key. In the second level, we have taken a secret seed which is also known to the receiver and sender only. Using this seed, we generate a sequence of unique numbers with the help of PRNG function. We have taken the cover image pixel blocks according to the generated numbers. So without knowing this seed, no one will be able to predict the number sequence.

We also verified our algorithm against standard measurement like SSIM. The Structural Similarity (SSIM) index is a method for measuring the similarity between two images. The SSIM index can be viewed as a quality measure of one of the images being compared provided the other image is regarded as of perfect quality. From the tables, it is observed that the SSIM values of all test images are nearer to 1.

6 Conclusion

In this paper, we introduced a novel secure data hiding scheme using Hamming Code for video steganography. Bit plane slicing of the each RGB color cover frame pixel block is also introduced to increase data hiding capacity. So the data embedding rate is raised up to 4 bpp which is greater than other existing schemes [1, 4, 5]. In our algorithm, PSNR is also high compared to existing schemes [4, 5] which means that we generate better visual quality stego videos. From security perspective, we introduced a shared secret key to find suitable bit pattern through XOR operation during data embedding as well as data extraction. The cover video frame block has been chosen in random location through PRNG which enhances security. We have tested our stego image with SSIM and observed that the proposed scheme is preferable for data embedding where visual quality and security constraint needs to be maintained for high payload. In future, the scheme has been extended to enhance security, capacity, and quality in different domains for video-based steganography.

References

1. Mstafa, R.J., Elleithy, K.M.: A highly secure video steganography using Hamming code (7, 4). In: Systems, Applications and Technology Conference (LISAT), 2014 IEEE Long Island. IEEE (2014)
2. Cao, Z., Yin, Z., Hu, H., Gao, X., Wang, L.: High capacity data hiding scheme based on (7, 4) Hamming code. SpringerPlus **5**(1), 175 (2016)
3. Hu, S.D.: A novel video steganography based on non-uniform rectangular partition. In: 2011 IEEE 14th International Conference on Computational Science and Engineering (CSE). IEEE (2011)
4. Yadav, P., Mishra, N., Sharma, S.: A secure video steganography with encryption based on LSB technique. In: 2013 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC). IEEE (2013)
5. Dasgupta, K., Mandal, J.K., Dutta, P.: Hash based least significant bit technique for video steganography (HLSB). Int. J. Secur. Priv. Trust Manage. (IJSPTM) **1**(2), 1–11 (2012)
6. Banik, B.G., Bandyopadhyay, S.K.: Image Steganography using BitPlane complexity segmentation and hessenberg QR method. In: Proceedings of the First International Conference on Intelligent Computing and Communication (pp. 623–633). Springer Singapore
7. Fridrich, J., Du, R., Meng, L. Steganalysis of LSB encoding in color images. In: Proceedings of ICME 2000, Jul.–Aug. 2000, N.Y., USA
8. Westfield, A., Pfitzmann, A.: Attacks on steganographic systems. In: Proceedings of 3rd Info. Hiding Workshop, Dresden, Germany, Sept. 28–Oct. 1, pp. 61–75 (1999)
9. University of Mannheim, Department of Computer science. http://ls.wim.uni-mannheim.de/de/pi4/research/projects/retargeting/test-sequences/
10. Xu, W., Wang, R.D., Shi, Y.Q.: Data hiding in encrypted H.264/AVC video streams by codeword substitution. IEEE Trans. Inform. Forens. Secur. **9**(4), 596–606 (2014)

Conceptual Design of Next Generation Security System Based on Thought Form Image Patterns

Rai Sachindra Prasad

Abstract Research on biophotons imaging techniques has now advanced to a level where it is possible to capture biophotons in true color two-dimensional image form. On the other hand, biophysicists have proved that thought process has definite impact on biophotons emission from human biofield. These two developments are enough to predict the prospects of development of a next generation security system to protect the vital sensitive installations and precious human lives without the need of the present day frisking and scanner aided security system. This paper is devoted to describe a conceptual design for such a security system. To demonstrate its functioning, true color images of thought forms from Theosophical Society literature are used. Most of the images were used in experiments for recognition of their patterns in our three recent papers where it was shown that thought forms had patterns of human behavior which matched with the comments attributed to the images in theosophical texts. These images were captured by theosophists using their power of clairvoyance, and published nearly 100 years ago. The three factors, viz., (i) patterns recognition in our papers, (ii) advances made in biophotons imaging techniques, and (iii) changes in thought process resulting in changes in human biofield, are the essential components on which a conceptual design of the next generation security system is proposed.

1 Introduction

In recent years, research on biophotons has attracted the attention of scientists all over the world. The nature and behavior of human beings is reflected in their biofield, i.e., in the characteristics of the biophotons. Whereas the environment and

R. S. Prasad (✉)
Department of EEE, D S Institute of Technology and Management,
Dr A K Technical University, Ghaziabad, Uttar Pradesh, India
e-mail: drsachpd@yahoo.com

biological organ systems are governed, to a large extent, through programmes inherited (biological system through DNA, and environmental system through matter–energy interaction); thought process is a manifestation of the interaction between environment and biological system. This thought process can be assessed either subjectively through the science of psychology, or through biophotons imaging [1–8] since human biofield is the interaction between thought process and the environment. The interest in this area world over is due primarily to the fact that it is now accepted that biophotons emission is a universal phenomenon linked to all living systems, and provides a new noninvasive powerful tool in medical diagnostics [5–8]. In an experiment, influence of intention, which is a thought process, on photon emission from the hand in eight subjects was investigated [9]. If a person tried to intentionally change his or her photon emission, the mean photon count decreased. Not all subjects were able to influence photon emission with similar success [3, 9–11]. Experiment on the effect of meditation, a state of mind, on photon emission was also performed [11–13]. In one experiment [12], the influence of meditation on photon emission from the hands and forehead was studied. Data indicated that photon emission changed after meditation. In one subject with high pre-meditation values, photon emission decreased during meditation and remained low in the post-meditation phase. Studies on the effect of consciousness on biophotons emission were also carried out in [13, 14] though this study did not give conclusive results. Korotkov [15] defines, consciousness as the ability of a subject to react to environmental information and change the behavior in accordance with this information.

From the foregoing discussions, it is obvious that the nature and behavior of human beings is linked to the thought process. In other words, thought process, captured in an image form and scientifically analyzed using image processing tools, can reveal the mental attitude. That this is going to be a revolutionary step in applications in medical, social and security environments is central to the design of a possible security system. The most important applications can be in social field to discriminate and segregate individuals with violent nature, or those who are "dangerous" elements from other considerations. In medical field, the analysis of thought forms can reveal the nature and behavior for applications in alternative and complementary medicines. In security environment, its application is expected to completely replace the scanner and frisking methods, etc., used presently at sensitive installations, including nuclear installations, airports, and metros. While the first and second applications of human biofield are easily traceable in literature survey, some of which are already shown in [1–15], the third possible application in security environments is a realistic projection in near future, the topic of this paper.

An important characteristic of biophotons is exhibiting of spectrum of colors. The spectral distribution of biophotons emission was analyzed in [16]. Recognizing the great importance of research on biophotons, many top-ranked institutions in the world have established a separate discipline of biophotonic research [17].

It is expected that in the near future, a joint group of scientists and engineers would be able to develop an image capturing system for biophotons emission from the whole human body, and not only from selected anatomic locations.

1.1 Thought Form Images from Theosophical Texts, Why?

A detailed study of theosophical texts including [18, 19] reveals that a great deal of similarity exists between the theosophists' and biophysicists' view of structural form of human beings. This is shown diagrammatically:

Theosophists' view: Physical body➜Astral Body ➜Color
Biophysicists' view: Physical Body➜Biofield➜Biophotons➜Color

The first similarity of views is in respect of astral body and biofield which are both clouds of electromagnetic field. The second similarity is in respect of color. Theosophists claim (without scientific proof though) that astral body has spectrum of colors; the biophysicists also view in exactly similar manner (with scientific proof) that biophotons have spectrum of colors. This similarity generated interest in authors for subjecting the thought form images from theosophical texts on image analysis tools to verify the rationale of comments attributed to the images. Thought form images used for analysis in this paper were taken from theosophical texts [20, 21], authored by C W Leadbeater and Annie Besant, renowned theosophists of the twentieth century. The rationale behind comments of Good or Bad was not explained with scientific proof, obviously due to absence of image processing tools. The theosophists based their comments in [21] on the following three principles: "(i) Quality of thought determines colour, (ii) Nature of thought determines form, and (iii) Definiteness of thought determines clearness of outline". With preset belief of color, shape, and outline being the pivots for qualifying the images in Good and Bad patterns, behavior of individuals in their captured thought forms was interpreted by the authors with comments attributed to the images [21]. For example, all shades of red were considered responsible for rendering the images in Bad category, but shades of blue color were considered responsible to render the images in Good pattern. Green color was given a lower priority over blue color. Similarly, if the shape and outline are regular and smooth then such images were considered as possible candidates in Good category if the color is not red. On the contrary, all images with disjointed or irregular shape were considered in Bad category if not in blue color. For more details, readers are referred to the text in [21]. It is worthwhile to mention that the influence of color is strongly supported in the texts of other religions [22, 23]. From a scientific view point, this preset belief, howsoever convincing it may be, would not be worthy of note unless these are tested on scientific platform.

1.2   Origin and Features of Thought Forms Images

Thought form images of some unidentified individuals were captured using the power of clairvoyance (claimed by the theosophists) in real situations. The hand-drawn and hand-painted images by three friends of the two authors were attributed comments based on the three aspects of color, shape, and outline. An important feature of thought form images is worthy of note since this was relevant to the results of pattern recognition in our recent papers [22–24]. The patterns identification on the aspect of color was reported in [22, 23], while shape-based pattern recognition was reported in [24]. Among the images which were used for pattern recognitions, there were some which involved the thought forms of more than one individual; as a result of this, the comments for these images are a mixture of both Good and Bad patterns. For example, images titled "At a Funeral", "At a Street Accident", and "On a Shipwreck" [21] were all real situations involving two or three unidentified individuals. In images of single person's thought form, the comments, in general, are either in Good or in Bad pattern.

The rest of the paper is organized as follows. In the next section, we highlight the advances made during the last two decades in imaging techniques of biophotons. In Sect. 3, experimental studies of scientists to prove the effect of thought process on biofield are discussed. Section 4 describes the techniques used by the authors for patterns identification in thought form images. Since the proposed security system is designed to identify only Bad and undesirable patterns of thought forms, we define them presenting their images in Sect. 5. These images are carefully selected on observing the comments attributed to them either in Bad or in Very Bad category in respect of the behavior of individuals in real situations. In Sect. 6, we present and discuss salient features of the proposed next generation security system with a flowchart. Finally, in Sect. 7, we conclude with our views of the future scenario of world order and use of the proposed system.

2   Advances in Biophotonics Research

Since a Biophotonic camera worthy of capturing thought form images is central to the long-term research objective of the authors, a brief review of literature is considered relevant. It is expected that in near future, the capturing device would also encompass imaging of biophotons emission of thought forms. Ever since the experiment on biophoton detection and measurement was reported by Popp [5], there has been a spurt in research activities in this area. In the past decades, only one-dimensional detection and measurement of ultra-weak photon emission was achieved using low-noise photomultiplier tube. Current development in both hardware and software in highly sensitive charge coupled device (CCD) based camera system allowed two-dimensional imaging of spontaneous photons emission [8]. Most of the advances in biophotonic research have taken place in developing a

CCD-based system. The motivation for this has been almost exclusively in applications for cancer diagnostics using the powerful tool of nonintrusive optical methods. Cancer detection system has been designed and implemented in [25].
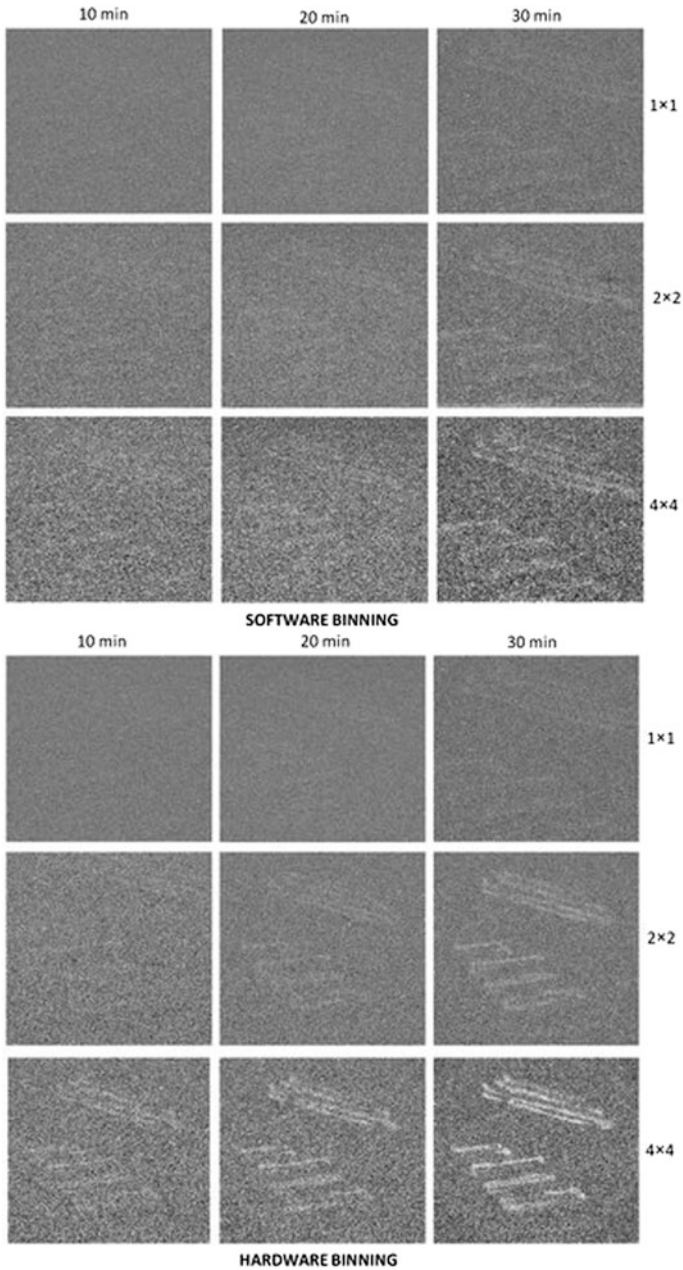
Fig. 1 Two-dimensional imaging (biophotons)

Figure 1 shows the two-dimensional images of biophotons reported in [25]. Schematic diagram of a basic two-dimensional imaging system is shown in Fig. 2 with essential components required in the image capturing device. It shows a biophotonic imaging system consisting of a liquid nitrogen cooled highly sensitive CCD camera system such as ATC 200 C (Photometrics), and a completely light shielded enclosure with no access of external electromagnetic field. The CCD camera system consists of a back illuminated thinned type Si-CCD [25]. Images of photons emission from tumors transplanted in mice, using CCD camera system, are reported in [8], where it was concluded that microscopic finding and biophoton intensity measurement were positively correlated and commented that this noninvasive optical technique has the potential to detect tumor viability.
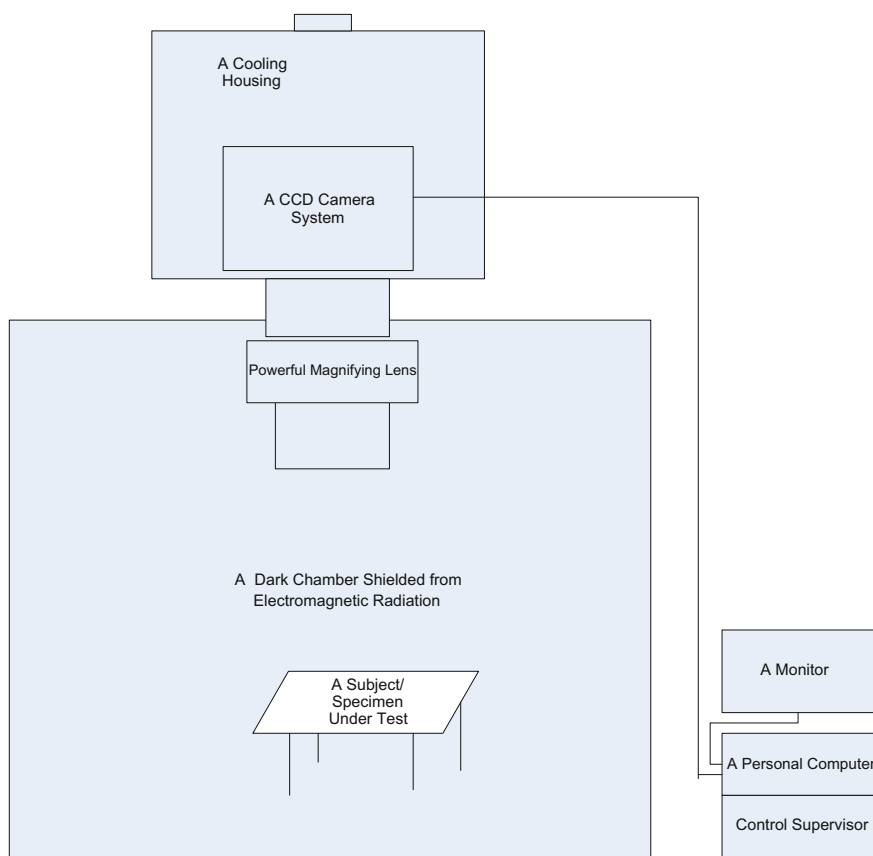
Fig. 2 Schematic of biophotonic camera system for imaging

3 Effect of Thought on Human Biofield

Several scientists and thinkers have elaborated on the significance of the effect of thoughts on biofields of human beings. Prof Hans Peter Duerr, the celebrated German Physicist remarked "understand life in terms of modern physics rather than biochemistry". Human biofield is a cloud of complex, highly dynamic electromagnetic field surrounding the human body, and is affected by the thought process of individuals as a result of looping interface with the environment. Biofield, in general, affects and is affected by the environment. Typically, the biofield may act directly on molecular structures, changing the configuration of molecules in functionally significant ways as well as may transfer bio-information through energy signals [26]. Experiments are reported of brain–computer interface (BCI) controlling a robotic quad-copter in three-dimensional physical space using noninvasive scalp electroencephalogram (EEG) [27]. If brain waves are associated with thoughts, then many phenomena can be explained. That the thought process is interlinked with brain waves is amply demonstrated by scientist all over the world from the late twentieth century to the present day. Rubic [28] states that bio-information may also extend beyond the physical to the ultraphysical realm to include information conveyed by means of the subtlest thought, intent, belief, power of will, and prayer. The concept of subtle energy bodies is also integral to the ancient Eastern philosophical view of human being that arose in India [29]. Indian philosophy and Ayurvedic system of medicine maintain that in addition to the physical body, there is a subtle body possessing various sheaths or bodies [30]. Some of the experiments showing effect of intention and meditation on changing the characteristics of biofield are previously described. The human biofield or energy field has the potential to be corrupted and the energy flow blocked [31]. Experimental data on the study of the effect of conscious intention on the genetic code indicate that conscious intention can influence properties of DNA. The implications of this research are profound since it is possible that we might also be able to consciously change the primary structure of DNA—the genetic code itself [32].

4 Imaging Tools Used for Patterns Recognition

Since the comments attributed to the thought forms images [21] were guided by three factors related to Color, Form (Shape), and Outline jointly, as previously stated, the imaging tools selected were HSV (Hue Saturation and Value) space for color in our papers [22, 23]. For patterns recognition on shape, we used Radon Transform (RT) with Histogram (Hist) of Grayscale images [24]. On patterns recognition using the third aspect of outline in our recent paper [33], it was found that the performance was very poor compared to the first two factors of color and shape. Whereas the comments on the images in theosophical literature [21] were influenced on the joint consideration of all three aspects of color, shape, and outline,

we investigated these three factors separately in order to assess the relative influence of each when compared with the comments attributed to the images [21]. Results showed that use of HSV, and RT with Histogram, identified the patterns correctly in 85% and 95% of the sample size, respectively. We now briefly describe the use of HSV [22, 23] and RT [24].

4.1 HSV Space

The HSV color space is based on the intuitive appeal of attributes such as tint, shade, and tone which enable human beings to differentiate between colors. The set of equations which can be used to transform a point in RGB coordinate system to the appropriate value in HSV space are well documented in several books and papers [34, 35]. However, in our paper [22, 23], the built-in MATLAB function for conversion of RGB space in scaled units (0 1), was used. The concept of Hue, Saturation, and Value can be easily visualized pictorially in Fig. 3. Hue follows a circular scale, increasing in the anticlockwise direction; Saturation increases along the radial lines from the center of the circle of the cone (upside down) to its full value at the periphery of the circle, while Value follows the vertical line touching the center. The rightmost edge of this cone slice shows the greatest amount of the dominant red Hue. Also from right to left along the circular scale in the cone, the Hue becomes less dominant and eventually becomes completely desaturated along the vertical center of the cone where Saturation is zero. These characteristics were

Fig. 3 Basics of Hue, Saturation and Value
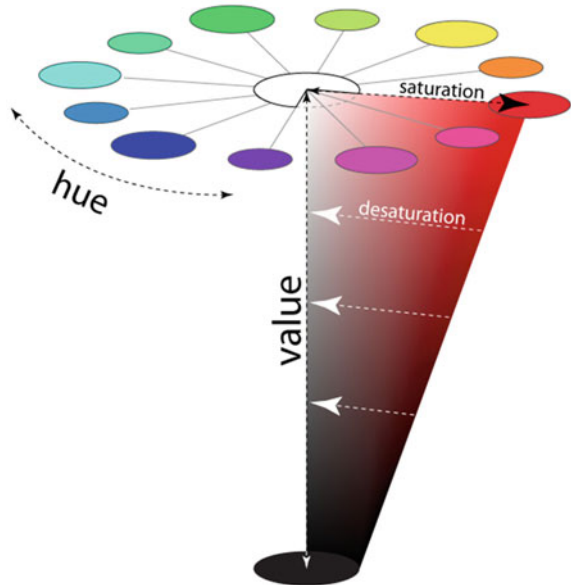
exploited in pattern recognition in [22, 23]. Since Hue increases from a "zero" value in the anticlockwise direction assuming full value as it reaches toward its starting point, the low and high Hue values are significant in rendering the pattern in Bad category. Similarly, Saturation at a low value also contributes in identifying pattern in Bad category. The thresholds of these three values were obtained using scatter diagrams of Hue and Saturation values against selected pixel positions on each image of the sample size of 21 images [22]. Final choice of thresholds was obtained on trials using probable candidates from the low and high zones in Fig. 4, and low zone in Fig. 5. For trial, four possible candidates of threshold from each of low and high zone in Fig. 4, and low zone in Fig. 5 were selected. The trial and error approach is similar to neural network with feedback. The final thresholds selected were: Hue (low) = 0.050, Hue (high) = 0.960, Saturation = 0.250. Using these

Fig. 4 Scatter of Hue [22]

Scatter of Hue (64 Pixel positions) on 21 Images
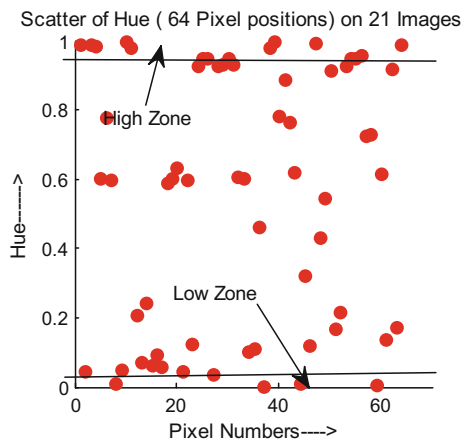
Fig. 5 Scatter of Saturation [23]

Scatter of Saturation (64 Pixel Positions) on 21 Images
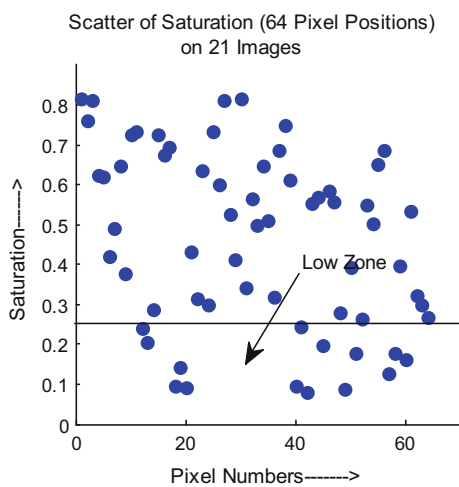
thresholds in a sample size of 32 images, 22 images forming the training group, and the rest forming the test group, patterns were identified either in Good or in Bad category [23].

4.2   Radon Transform (RT) and Histogram

RT has been adequately described in [36, 37]. RT is a projection of the image intensity along radial lines oriented at some specific angles. The projections can be computed over the range [0 180] degrees. Since RT of an intensity image is the sum of the RTs of all pixels in the image, built-in function of RT in MATLAB was used to view each of the images in parameter space. Applying RT on intensity image on specific angles (degrees) returns a vector xp containing the radial coordinates corresponding to each row of R, where R is the RT on intensity image in the matrix [R xp]. The radial coordinates returned in xp are the values along the x-axis, which are oriented at specific angles, counterclockwise from the x-axis. The origin of both axes is the center pixel of the image. In a few cases, where Bad and Good patterns are not clearly defined by the Histogram data (discussed next), the subjective methods of image brightness, and edge detection (binary image) are used. Since RT is closely related to Hough Transform (HT), using the MATLAB codes for RT to implement a form of HT, straight lines were detected in the images. Each of the true color thought form images was imported into the MATLAB workspace for applying RT on image intensity, and RT on binary image (BW).

An image histogram shows the number of pixels for each tonal value [36]. However, in [24], we used the overall area of the histogram as identifier of patterns in images. The quantitative measurement of the area occupied by histogram was estimated graphically, shown in Fig. 6. In classification of the patterns, the spread of y data on either side of a logically assumed pixel cutoff value (threshold) of 100 pixels was used as the marker for low (dark) if the spread is on the left of threshold, else high (bright) when the spread is on the right side. In the former case, it was

Fig. 6 Graphically computed Hist area
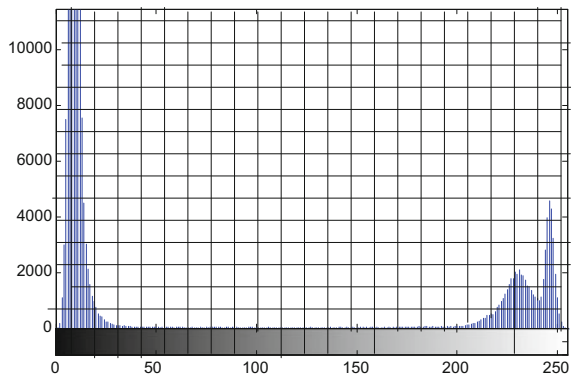
identified in Bad, and in the latter case in Good patterns. In few cases where the spread is found lying on both sides of the threshold, the identification is not defined clearly. For these cases, subjective methods were used to support the objective method.

5 Thought Forms Images of Undesirable Characteristics

Careful selection of a group of eight thought form images from among the images published in theosophical text [21] forms the prototypes of undesirable characteristics for comparison with patterns of thought forms of individuals, when subjected to screening in the proposed security system. It is important to note that out of the three principles guiding the comments in theosophical text [21], only the first two relate to the nature and behavior of individuals; the third principle (outline) only influences the definiteness of thoughts of individuals [21]. In security system, we are concerned only with the nature and behavior of individuals. Images identified unambiguously in [23, 24] with undesirable characteristics, i.e., "dangerous" for public safety, are shown in Table 1 with their assigned numbers and titles as given in text [21].

Table 1 Thought forms (undesirable characteristics)

Sr. no.	Image nos.	Area in % /Spread x-axis		HSV	RT
1	22, 23	98%	0–50	VB	VB
		02%	50–100		
2	24	93%	0–100	VB	VB
		07%	100–150		
3	32	95%	0–100	VB	VB
		05%	0–150		
4	13	100%	0–100	B	VB
5	29	100%	0–100	B	VB
6	27	100%	0–40	B	VB
7	28	95%	0–100	B	VB
		05%	100–130		
8	25	95%	0–50	B	VB
		05%	100–200		

22 Murderous rage, 23 Sustained anger, 27 Sudden fright, 28 Selfish greed, 29 Greed for drink, 32 Gambler, 13 Grasping animal affection, 25 Watchful jealousy; 24 Explosive anger

6   Design of Thought Form Patterns-Based Security System

The flowchart in Fig. 7 explains the functioning of the proposed security system. If (see Fig. 7) the patterns emerging in both subjective and objective methods are in agreement, then further action would be dependent upon the degree of security risks involved (Table 2). To explain it further, if the final outcome of patterns is same as the prototypes shown in Table 1, then the degree of risks corresponding to the pattern would be as shown in Table 2. In case there is no agreement between the subjective and the objective methods, then benefit of doubt about the risk is to be given to the person whose thought form was analyzed. In such cases, the subject's personal details shall be noted before permitting the subject free to resume intended destination.

The flowchart of the proposed design of thought form image pattern-based security system, shown in Fig. 7, has the following salient features:

1. Both subjective and objective methods employed for decision regarding the patterns.
2. Unless both subjective and objective methods agree on a pattern, the final decision to restrain or allow the subjects freedom to resume the journey, is not taken.
3. Image processing tools for identifying patterns in both groups (subjective and objective) have been thoroughly tested in our reports [22–24] over all types of thought forms images available in text [21]. Only those thought form images which proved to have undesirable characteristics for society are used as prototypes in the proposed security system. Both groups showed complete agreement in the eight thought forms images of undesirable characteristics, but with different severity.
4. Since the eight Bad patterns of thought forms images (Table 1) are not unique, it is likely that there may be some which will have divided decisions in both the subjective and objective methods. Therefore, to safeguard that no innocent subject is harassed, the algorithm has incorporated a safety mechanism of complete agreement between the methods used.

Before closing, images of all eight thought forms with undesirable characteristics, used as prototypes in the design of security system, are shown in Figs. 8, 9, 10, 11, 12, 13, 14, and 15.

Fig. 7 Flowchart of a next generation security system

Table 2 Severity of security risk

Sr. nos.	Hist area of Images	Type of risk	Risk level	Remark (A = Area)
1	22, 23, 24, 32 95% A <= 100	Violent	2	1 Area Shown in % for RT
2	13, 25, 27, 28, 29 95% A <= 100	Aggressive	1	2 The Severity is '2' when HSV, and RT agree, else is '1'
3	Others	No risk	0	

22 Murderous rage, 23 Sustained anger, 27 Sudden fright, 28 Selfish greed, 29 Greed for drink, 32 Gambler, 13 Grasping animal affection, 25 Watchful jealousy; 24 Explosive anger

Fig. 8 Image 22, 23

Fig. 9 Image 24

Fig. 10 Image 32

Fig. 11 Image 13
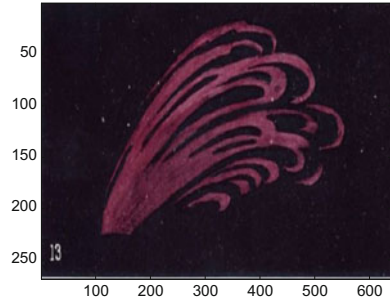
Fig. 12 Image 29

Fig. 13 Image 27

Fig. 14 Image 28
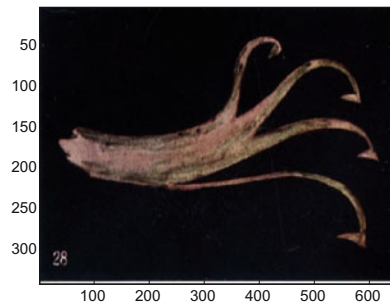
Fig. 15 Image 2

7   Conclusions

A conceptual design for security system has been proposed on two realistic assumptions: (i) thought forms' induced biophotons captured by a sophisticated camera system would be of the same patterns as those selected from the theosophical literature, (ii) it is possible to develop a sophisticated biophotonic camera system to capture thought forms in true color image. The proposed design of the security system is expected to replace the present day security system which is effective only in detecting explosives and drugs but cannot detect the dangerous intentions of persons. We strongly believe that with continued research efforts toward developing the specific biophotonic camera, such a security system is going to be realized sooner than later.

References

1. Ignatov, I., Mosin, O.: Results of bioinfluence of Dimitar Risimanski with biophysical model systems. J. Med. Physiol. Biophys. **24**, 5–17 (2016)
2. Ignatov, I. et al.: Evaluating of possible methods and approaches for registering electromagnetic waves emitted from the human body, **2**, 96–116 (2014)
3. Chiang, L.H., Khong, P.W., Ghista, D.: Bioenergy based medical diagnostic application based on gas discharge visualization. In: Conference on Proceedings IEEE Engineering in Medicine and Biology Society, vol. 2, p. 1533 (2005)
4. Marinov, M., Ignatov, I.: Color Kirlian spectral analysis. In: Color Observation with Visual Analyzer. Euromedica, Hanover, pp. 57–59 (2008)
5. Cohen, S., Popp, F.A.: Biophoton emission of the human body. Indian J. Exp. Biol. **41**, 440–445 (2003)
6. Van Wijk, R., Van Wijk, E.P.A.: An introduction to human biophoton emission. Forsch Komplementärmed Klass Naturheilkd **12**, 77–83 (2005)
7. https://www.ncbi.nlm.nih.gov/pubmed/20637006
8. http://www.nature.com/articles/srep01211/figures/2
9. Vekaria, M.: Biophoton emission and intentionality. Ph. D. dissertation, Californian Institute for Human Science, Encinitas (2003)
10. Korotkov, K.G., et al.: Application of electrophoton capture (EPC) analysis based on gas discharge visualization (GDV) technique in medicine: a systematic review. J. Altern. Complement. Med. **16**(1), 13–25 (2010)
11. Van Wijk, E.P.A.: Anatomic characterization of human ultra-weak photon emission in practitioners of transcendental meditation™ and control subjects. J. Altern. Complementary Med. **12**(1), 31–38 (2006)

12. Van Wijk, E.P.A., Ackerman, J., Van Wijk, R.: Effect of meditation on hand and forehead ultraweak photon emission. Forsch Komplementärmed Klass Naturheilkd. https://doi.org/10.1159/000084028
13. http://www.telegraph.co.uk/news/health/news/8862275/Meditation-improves-the-immune-system-research-shows.html
14. Dobrin, R., et al.: Experimental measurements of the human energy field. In: Krippner, S. (ed.) Psychoenergetic Systems: The Interface of Consciousness, Energy and Matter, pp. 227–230. Gordon and Breach, New York (1979)
15. Korotkov, K.G., Williams, B., Wisneski, L.A.: Assessing biophysical energy transfer mechanisms in living systems: the basis of life processes. J. Altern. Complement Med. **10**(1), 49 (2004)
16. Inaba, H.: Super-high sensitivity systems for detection and spectral analysis of ultraweak photon emission from biological cells and tissues. Experientia **44**, 550–559 (1988)
17. http://www.blt.uni-saarland.de/index_e.htm
18. www.hermetics.org/pdf/theosophy/C.W._Leadbeater_-_An_Outline_of_Theosophy.pdf
19. http://blavatskyarchives.com/theosophypdfs/leadbeater_the_astral_plane_its_scenery_inhabitants_1895.pdf
20. Leadbeater, C.W., Besant, A.: Thought Forms. Project Gutenberg Release #16269
21. http://www.anandgholap.net/Thought_Forms-AB_CWL.htm#THE_MEANING_OF_THE_COLOURS
22. Prasad, R.S., Prasad, S., Prasad, V.: Patterns identification in thought-form images. In: Proceedings of the 2nd IEEE International Conference ICECCT 2017, Coimbatore, Tamil Nadu, India, February 22–24, published in IEEE Xplore 23 November 2017   pp 1–7. https://doi.org/10.1109/ICECCT.2017.8117942
23. Prasad, R.S., Prasad, S., Prasad, V.: Patterns recognition in thought form images using chromaticity parameters. In: Proceedings of the IEEE International Conference WiSP-NET2017, Chennai, Tamilnadu, India, March 22–24, pp. 2707–2714 (2017) (978-1-5090-9/17/$31.00 © 2017IEEE)
24. Prasad, R.S., Prasad, S., Prasad, V.: Pattern recognition in thought-form images using radon transform and histograms. [1] In: Proceedings of the Int conf on Bioinformatics and Biomedical Image Processing, ICBIP '17, August 23–25, Kitakyushu, Japan © 2017 ACM ISBN 978-1-4503-5268- 0/17/08…$15.00. https://doi.org/10.1145/3133793.3133806
25. Takeda, M., et al.: Biophoton detection as a novel technique for cancer imaging. J Cancer Sci. **5**(8), 656–661 (2004)
26. Trivedi, M.K., Tallapragada, R.M.: Biofield energy signals, energy transmission and neutrinos. Am. J. Mod. Phys. **5**(6), 172–176 (2016)
27. LaFleur, K. et al.: Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface. J. Neural Eng. **10**(4) (2013)
28. Rubic, B.: Energy medicine and unifying concept of information. Alter. Ther. Health **1**(1), 34–39 (1995)
29. Rubic, B.: The Human Biofield and a Pilot Study on Quigong, pp. 151–170 (2004)
30. http://bdigital.ufp.pt/bitstream/10284/775/1/151-172Cons-Ciencias%2002-2.pdf
31. http://www.red-spirit-energy-healing.com/human-biofield.html
32. http://www.item-bioenergy.com/infocenter/ConsciousIntionononDNA.pdf
33. Prasad, R.S.: Thought-forms patterns identification using histograms of the derivative of intensity outline. Accepted for presentation at the IEEE 2017 International Conference ICACCI, Manipal University, Karnataks, India, 13–16 Sept 2017
34. 2000-SpringerMonograph ON IMAGE PrOCESSING-1(1).pdf. Available on-line
35. Loesdau, M., Chabrier, S., Gabillon, A.: Hue and saturation in the RGB color space laboratoired'excellence CORAIL, Géopôle du Pacifique Sud EA4238 Université de la Polynésie française, Tahiti

36. Belongie, S., et al.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(24), 509–522 (2002)
37. Tabbone, S., et al.: Histogram of radon transform. A useful descriptor for shape retrieval. In: 19th International Conference on Pattern Recognition—ICPR (2008)

Author Biography

Prof. Rai Sachindra Prasad Born in 1943, Prof. Prasad having graduated in Electrical Engineering from Bihar University, obtained Post Graduation in Instrument Technology, from IISC, Bangalore, and PhD (Systems) from IIT, Kharagpur, India. Has over forty years of experience in academics and research in technical institutions and universities in India as Chief Scientist-cum-University Prof/ Professor/Dean. He has guided PhD research in the area of Chaos theory and Smart grids, besides several MTech projects. He has travelled widely including Germany, France, and USA on academic visits. He is an acknowledged specialist of smart grid and systems engineering and has reviewed several papers in IEEE, Elsevier International journals. He is Fellow of Institution of Engineers and Member of IEEE and other professional bodies, besides serving on Editorial board of some reputed journals. Till date he has over 70 publications in international journals and conferences, some of which are highly cited. His full time current interest is in the area of Bioinformatics and Biomedical Engineering, particularly in image processing of Thought forms and human biofield on which he has already published (2017) six papers in IEEE and ACM sponsored International conferences in India, Japan, and USA. His current assignment is Visiting Professor in DS Institute of Technology and Management, Dr Abdul Kalam Technical University, India.

A New Data Hiding Method Using Block Pixel Intensity Range

Sujit Kumar Das and Bibhas Chandra Dhara

Abstract In this paper, we have proposed a new data hiding method for grayscale image based on the block pixel intensity range. This work is inspired by the Varasaki's data hiding method. Here, the cover image is symmetrically divided into blocks of size 3×3. The intensity range of each block is computed and this intensity range is divided into four zones. For each block, the neighboring pixels are modified to place into a zone according to the message bit so that the message can be reconstructed losslessly. Our proposed method gives better quality stego image than LBP-based data hiding.

Keywords Data hiding · LBP · LBP data hiding · Block intensity range

1 Introduction

In recent times, data hiding method became a popular technique for the secure communication. In data hiding, a secret message is embedded into a cover data, such as audio, video, and image, in such a way that the distortion of cover data is minimum. The cover data with embedded message is called stego data. Most of the data hiding techniques prefer image as cover data. The data hiding in cover image can be carried out in (i) spatial domain and (ii) frequency domain. In spatial domain, the message is embedded by direct manipulation of pixels value of the cover image. The least significant bits (LSB) substitution and pixel value differencing (PVD) are well-known data hiding techniques in spatial domain. In LSB method, r message bits replaced the r least significant bits of a pixel. To make distortion minimum in case of LSB based method, several research works [1–4] have been proposed. The main advantage of LSB method is that it has high embedding capacity, but it lacks security. The LSB

S. K. Das (✉) · B. C. Dhara
Department of Information Technology, Jadavpur University, Kolkata, India
e-mail: sujit.cse.jgec@gmail.com

B. C. Dhara
e-mail: bibhas@it.jusl.ac.in

can be easily detected by RS-steganalysis [5] attack. On the other hand, in PVD techniques, difference of pixel values is used to embed data. The PVD based data hiding method is introduced in [6] where the cover image is divided into nonoverlapping blocks with two pixels and difference of these two pixel values is used for embedding purpose. Later, a number of methods using PVD technique [7–9] have proposed to improve hiding capacity as well as quality of the stego image.

In transform domain, the cover image is first transformed into frequency domain, then secret data is embedded and finally back to spatial domain. For data hiding purpose, people have used DCT [10, 11], DFT [12, 13], and DWT [14, 15] methods.

Texture information is also used for data hiding purpose. An adaptive data hiding method has been proposed in [16], where textural analysis of the cover image is done in wavelet domain. Varasaki et al. [17] have used the local binary pattern (LBP) for data hiding. In [17], the cover image is divided into nonoverlapping blocks of size 3×3. The local binary pattern (LBP) of each block is computed. Then the binary pattern is modified according to the message bits in such a manner so that the binary pattern of the embedded image block gives the message bits.

In this paper, we have proposed a new data hiding method for grayscale images by extending the concept of [17]. In this method, the cover image is divided into 3×3 blocks. For each block, the block pixel intensity range is computed and symmetrically divided into four zones where two zones represent bit 1 and other two indicate 0. Then, 8 neighbor pixels are placed into these four zones according to message bits. The organization of rest of this paper is as follows: Related work is described in Sect. 2. The proposed method is presented in Sect. 3. Experimental result is demonstrated in Sect. 4. Finally, conclusions are drawn in Sect. 5.

2 Related Work

This proposed work is inspired by Varasaki's et al. [17] method. Here, we have extended the method to increase the quality of the stego images. Now, we are going to introduce the Varasaki's method.

2.1 Data Hiding Based on LBP

LBP is first defined in [18]. It is a grayscale invariant texture measure and very useful operator for feature extraction. In LBP, an image is partitioned into blocks (B) of size 3×3. In each B, a neighbor pixel p_i is labeled by $b_i \in \{0, 1\}$ respect to the center pixel p_c. So, from each image block, a binary pattern of length eight is constructed. An example of computing binary pattern for an image block is shown in Fig. 1. In [17], using this binary pattern a data hiding method is proposed wherein each B, a message $m = m_7 m_6 \ldots m_1 m_0$ of length eight is embedded by modifying the

P_0	P_1	P_2
P_7	P_c	P_3
P_6	P_5	P_4

b_0	b_1	b_2
b_7		b_3
b_6	b_5	b_4

50	51	44
53	50	49
47	57	46

1	1	0
1		0
0	1	0

(a)   (b)   (c)   (d)
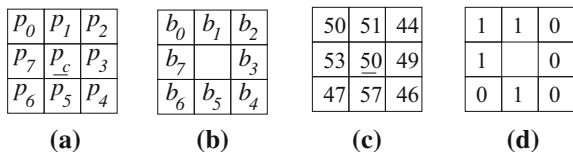
Fig. 1 Illustration of LBP **a** general image block, **b** binary pattern of (**a**), **c** example of image block, and **d** binary pattern of (**c**)

(a) two zones as in [17]   **(b)** four zones in the proposed method
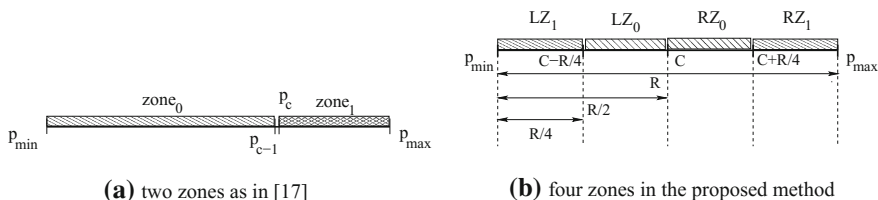
Fig. 2 Partition of block intensity range

B as B'. The modification is done in such a way that the LBP pattern of B' gives the message m. The modification of image block ($B \rightarrow B'$) is done as follows.

For each block B, block intensity range, $R = p_{max} - p_{min}$ (where p_{max} and p_{min} are maximum and minimum intensity of the block, respectively) is computed. In [17], R is divided into two regions $zone_1 \in [p_c, p_{max}]$ and $zone_0 \in [p_{min}, p_c)$ (as given in Fig. 2a) where $zone_1$ and $zone_0$ stand for bit 1 and 0, respectively. So, during the embedding a pixel p_i is placed into $zone_j$ if $m_i = j$ for $i \in \{1, 2, \ldots, 8\}$ and $j \in \{0, 1\}$. If $b_i \neq m_i$, the region associated with pixel p_i has to be changed; otherwise no change is required. A pixel p_i is modified as follows:

$$p'_i = \begin{cases} p_i & \text{if } b_i = m_i \\ p_c & \text{if } b_i = 0 \text{ and } m_i = 1 \\ p_c - 1 & \text{if } b_i = 1 \text{ and } m_i = 0 \end{cases} \tag{1}$$

3 Proposed Method

From the LBP-based data hiding method, we observed that neighbor pixels of a block B are thresholded respect to p_c without considering the intensity profile of the block and a significant amount of error may be added when partition of block is nonsymmetric (p_c is not midpoint of R) or block intensity range is comparatively large. In this proposed method, we have tried to resolve the above problems. In the present work, we consider the block intensity range $R = p_{max} - p_{min}$ and partition R, so that

less error will be added. In the following two subsections, we have described the proposed data hiding and data extraction methods.

3.1 Data Embedding

During data embedding, the cover image is divided into non-overlapping blocks then the pixel intensity range of a block is used for embedding. Let the size of cover image is $M \times N$ (where M and N both are divisible by 3) and it is divided into blocks (B) of size 3×3. The block intensity range $R = p_{max} - p_{min}$ is partitioned it into four zones as LZ_1, LZ_0, RZ_0 and RZ_1 (as shown in Fig. 2b) where C is the midpoint of the range, $p_{min} \leq LZ_1 \leq (C - R/4)$, $(C - R/4) < LZ_0 < C$, $C \leq RZ_0 < (C + R/4)$ and $(C + R/4) \leq RZ_1 \leq p_{max}$. In each block, we can embed 8 bits of message. According the message bit, we place the corresponding pixel into one of these four zones. After embedding, the pixel associated with zone indicates the bit value of message, i.e., LZ_0 and RZ_0 indicate 0 bit and other two represent 1. We have modified p_i in two ways: when $R \leq 1$, we adopt simple LBP based method (i.e., modify p_i using Eq. (1)) and

$$\text{When } R \geq 2, \quad p'_i = \begin{cases} p_i & \text{if } m_i = 1 \text{ and } p_i \in \{LZ_1, RZ_1\} \\ & \text{or } m_i = 0 \text{ and } p_i \in \{LZ_0, RZ_0\} \\ \lfloor C - \frac{R}{4} \rfloor & \text{if } m_i = 1 \text{ and } p_i \in \{LZ_0\} \\ \lceil C + \frac{R}{4} \rceil & \text{if } m_i = 1 \text{ and } p_i \in \{RZ_0\} \\ \lceil C - \frac{R}{4} \rceil & \text{if } m_i = 0 \text{ and } p_i \in \{LZ_1\} \text{ and } R < 4 \\ \lfloor C + \frac{R}{4} \rfloor & \text{if } m_i = 0 \text{ and } p_i \in \{RZ_1\} \text{ and } R < 4 \\ \lfloor C - \frac{R}{4} \rfloor + 1 & \text{if } m_i = 0 \text{ and } p_i \in \{LZ_1\} \text{ and } R \geq 4 \\ \lceil C + \frac{R}{4} \rceil - 1 & \text{if } m_i = 0 \text{ and } p_i \in \{RZ_1\} \text{ and } R \geq 4 \end{cases} \quad (2)$$

After the embedding, it may be possible that p_{min} and p_{max} remain unchanged, either one is changed or both are changed. If either one or both are changed, then R may also be changed, then lossless data extraction is not possible. To get back the original message, we have to execute post-processing for the block to retain the R. Let p'_{min} and p'_{max} be the minimum and maximum value in embedded block. Case I (when one is changed): if $p_{min} \neq p'_{min}$, we have chosen a pixel to updated with p_{min}. It can be chosen from $\{p_c, p'_{min}\}$ with minimum distortion. Similarly, we can keep R as intact when $p_{max} \neq p'_{max}$. Case II (when both are changed): in this condition, we have required two pixels to keep the value of p_{max} and p_{min} so that R remained unchanged. We have chosen two from $\{p_c, p'_{min}, p'_{max}\}$ so that distortion is minimum. It may be noted that while we have to change p'_{min} or p'_{max} to maintain the block range, we can only consider the pixel for which message is '1' (otherwise, we cannot
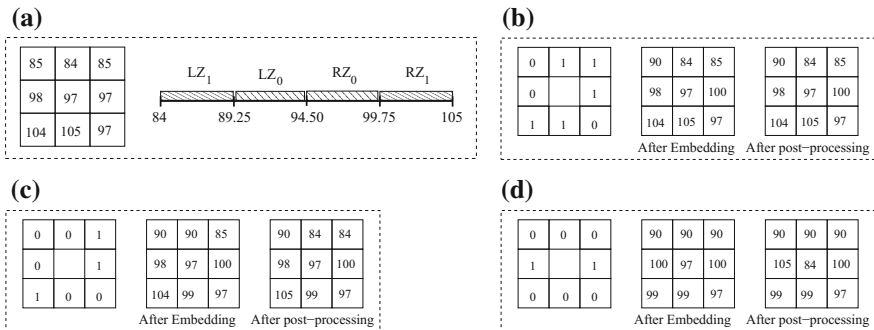
Fig. 3 Data hiding using proposed method: **a** an image block with zones, **b**, **c**, **d** embedded block with different messages

recover the original message). It may be noted that when all message bits are '0', then we cannot preserve p_{min} and p_{max} in the embedded block and in that case, we update all the pixels including p_c with the average pixel intensity of the original block so that during extraction, LBP method can be employed and all 0's message can be extracted. Examples of data embedding are shown in Fig. 3. In Fig. 3a, we have considered an image block with its zones. Figure 3b shows the message embedding followed by post-processing to ensure block range and we note that there is no effect of the post-processing step. In Fig. 3c, we have seen that after embedding $p'_{min} \neq p_{min}$ and $p'_{max} \neq p_{max}$ and applying post-processing step, we ensure the block intensity profile by changing p'_{min} and p'_{max}. In case of Fig. 3d, to ensure the block range p'_{max} is set to p_{max} and p_c is set to p_{min}. In this work, the block obtained after the post-processing is considered as stego block. The algorithm of the data hiding technique is following:

Algorithm 1: Data Hiding

Input: Cover image(C), message(M)
Output: Stego image (C')

---

1   Divide C into blocks of size 3×3
2   For each block B do

   2.1   Find p_{max} and p_{min} of each block and compute $R = p_{max} - p_{min}$.
   2.2   Divide R into four zones LZ_1, LZ_0, RZ_0 and RZ_1
   2.3   Select next 8 bits from M as $m = m_7 m_6 \dots m_0$
   2.4   If all message bits are '0'
      2.4.a   Set all pixels of B' as the average of B
   2.5   Otherwise,
      2.5.a   Modify each pixel p_i using Equn. (1) or (2) and define B'
      2.5.b   if needed, execute post-processing to ensure that $p'_{min} = p_{min}$ and $p'_{max} = p_{max}$ and modify B' accordingly.

3   Return C'

90	84	85
98	97	100
104	105	97

(a)

LZ_1 LZ_0 RZ_0 RZ_1

84   89.25   94.50   99.75   105

(b)

0	1	1
0		1
1	1	0

(c)

0 1 1 1 0 1 1 0

(d)
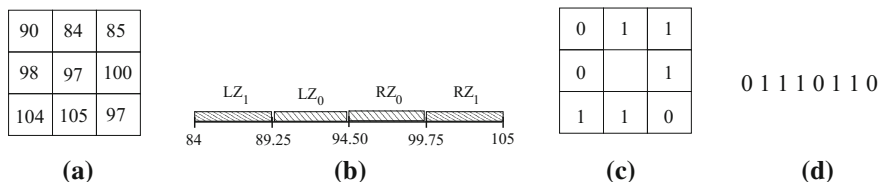
Fig. 4 Data extraction example: **a** stego image block **b** block intensity zones **c** message block **d** extracted message

3.2 Data Extraction

The data extraction process ensures that original message can be reconstructed from the stego image. Data extraction method is simple and just opposite of the embedding method. From an embedded block, we compute the intensity range R and partition the intensity accordingly (when applicable). If $R = 0$, we simple assign $m_i = 0$ for all $i \in \{0, 1, \ldots, 7\}$. When $R = 1$,

$$\text{When } R = 1, \quad m_i = \begin{cases} 0 \text{ if } p'_i = p_{min} \\ 1 \text{ if } p'_i = p_{max} \end{cases} \tag{3}$$

$$\text{If } R \geq 2, \quad m_i = \begin{cases} 0 \text{ if } p'_i \in \{LZ_0, RZ_0\} \\ 1 \text{ if } p'_i \in \{RZ_1, RZ_1\} \end{cases} \tag{4}$$

So, the message extraction process is as follows:

Algorithm 2: Data Extraction

Input: Stego image(C')
Output: Message M

---

1   Divide C' into blocks of size 3×3
2    $M = \phi$
3   For each block B

    3.1   Find p_{max}, p_{min} and R
    3.2   Divide R into four zones LZ_1, LZ_0, RZ_0 and RZ_1
    3.3   If $R \neq 0$
      3.3.a   Find $m = (m_7 m_6 \ldots m_0)$ using Equn. (3) or (4)
    3.4   Otherwise, set $m = (m_7 m_6 \ldots m_0) = 0$
    3.5 $M = M || m$

4   Return M

One example of data extraction for $R > 1$ is shown in Fig. 4.

Table 1 Comparison results between LBP method and our proposed method

Image	PSNR		SSIM	
	LBP method	Proposed method	LBP method	Proposed method
Peppers	28.43	34.44	0.9952	0.9958
Boat	27.64	33.61	0.9951	0.9959
Lake	26.38	32.05	0.9953	0.9959
Airplane	28.15	33.92	0.9953	0.9959
Splash	31.32	37.40	0.9952	0.9958
Zelda	33.25	38.65	0.9951	0.9959
Couple	28.93	35.24	0.9951	0.9958
Average	29.16	35.04	0.9952	0.9958

4 Experimental Results

In this section, we have given experimental results to demonstrate the performance of the proposed method. We have considered seven standard test images (of size 512×512) in this experiment. For experimental purpose, we have resized the images of size 510×510. According to the proposed method, theoretical capacity of each test images is $\frac{510}{3} \times \frac{510}{3} \times 8 = 170 \times 170 \times 8$ bits (or 28,900 bytes). Here, we have considered a random binary string of size 28,900 bytes where randomness of the string is tested using NIST software [19] and the string successfully passes all the tests. In this experiment, entire string is embedded into the cover images, i.e., we have used 100% payload. For comparison purpose, we have compared the result of the proposed method with the LBP-based method [17]. We measure the quality of the stego images in terms of peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) [20]. The comparative result is reported in Table 1. This table shows that SSIM of both the methods are very close and it is note that SSIM ($\in$ [0, 1]) and close to '1' means stego image is visually almost same as original image. Again, from Table 1, we note that PSNR of the proposed method is much better than LBP method, i.e., proposed method is superior than the LBP based method.

5 Conclusion

In this work, we have extended the LBP-based data hiding method to improve the quality of the stego images. To improve the quality block intensity, range is divided into four zones instead of two zones like LBP method. This division of the proposed method is symmetric, but in LBP method, it may not be symmetric (since it depends on p_c). The experimental result shows that the quality (PSNR and SSIM) of the stego images is better than LBP method. This method may be further extended to increase the data hiding capacity.

References

1. Wang, R.Z., Lin, C.F., Lin, J.C.: Image hiding by optimal lsb substitution and genetic algorithm. Pattern Recogn. **34**(3), 671–683 (2001)
2. Yang, C.H.: Inverted pattern approach to improve image quality of information hiding by lsb substitution. Pattern Recogn. **41**(8), 2674–2683 (2008)
3. Luo, W., Huang, F., Huang, J.: Edge adaptive image steganography based on lsb matching revisited. IEEE Trans. Inf. Forensics Secur. **5**(2), 201–214 (2010)
4. Leng, H.S., Tseng, H.W.: High payload data hiding based on just noticeable distortion profile and lsb substitution. In: Advances in Intelligent Information Hiding and Multimedia Signal Processing: Proceeding of the Twelfth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Nov. 21–23 2016, Kaohsiung, Taiwan, vol. 1, pp. 59–66. Springer (2017)
5. Fridrich, J., Goljan, M., Du, R.: Reliable detection of LSB steganography in color and grayscale images. In: Proceedings of the 2001 Workshop on Multimedia and Security: New Challenges, pp. 27–30. ACM (2001)
6. Wu, D.C., Tsai, W.H.: A steganographic method for images by pixel-value differencing. Pattern Recogn. Lett. **24**(9), 1613–1626 (2003)
7. Chen, J.: A pvd-based data hiding method with histogram preserving using pixel pair matching. Signal Process. Image Commun. **29**(3), 375–384 (2014)
8. Shen, S.Y., Huang, L.H.: A data hiding scheme using pixel value differencing and improving exploiting modification directions. Comput. Secur. **48**, 131–141 (2015)
9. Hussain, M., Wahab, A.W.A., Ho, A.T., Javed, N., Jung, K.H.: A data hiding scheme using parity-bit pixel value differencing and improved rightmost digit replacement. Signal Process. Image Commun. **50**, 44–57 (2017)
10. Lin, C.C., Shiu, P.F.: High capacity data hiding scheme for DCT-based images. J. Inf. Hiding Multimed. Signal Process. **1**(3), 220–240 (2010)
11. Lin, Y.K.: A data hiding scheme based upon dct coefficient modification. Comput. Stand. Interf. **36**(5), 855–862 (2014)
12. Liao, X., Li, K., Yin, J.: Separable data hiding in encrypted image based on compressive sensing and discrete fourier transform. Multimed. Tools Appl. 1–15 (2016)
13. Bhattacharya, T., Dey, N., Chaudhuri, S.: A session based multiple image hiding technique using DWT and DCT (2012). arXiv:1208.0950
14. Kumar, M., Agrawal, S., Pant, T.: SVD-based fragile reversible data hiding using DWT. In: Proceedings of Fifth International Conference on Soft Computing for Problem Solving, 743–756. Springer (2016)
15. Baby, D., Thomas, J., Augustine, G., George, E., Michael, N.R.: A novel DWT based image securing method using steganography. Proc. Comput. Sci. **46**, 612–618 (2015)
16. Wang, Y., Yu, Y.: An adaptive data hiding in wavelet domain based on texture analysis of image. In: International Conference on Audio, Language and Image Processing, 2008 (ICALIP 2008), pp. 1369–1372. IEEE (2008)
17. Varsaki, E.E., Fotopoulos, V., Skodras, A.N.: Data hiding based on image texture classification. Signal Image Video Process. **7**(2), 247–253 (2013)
18. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–987 (2002)
19. Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E.: A statistical test suite for random and pseudorandom number generators for cryptographic applications. Technical report, Booz-Allen and Hamilton Inc Mclean Va (2001)
20. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)

Fatigue Detection Based on Eye Tracking

Ashis Pradhan, Jhuma Sunuwar, Sabna Sharma and Kunal Agarwal

Abstract This paper presents the development of a fatigue detection system that would be capable of detecting an individual's level of alertness through live video acquisition. The approach is to build a nonintrusive system that uses computer vision methods to localize face, eyes, and iris positions to measure level of eye closure within an image, which, in turn, can be used to identify visible eye signs associated with fatigue leading to a sleepy state. The aim here is to detect this state early enough and issue a warning or alert in the form of an alarm.

Keywords Eyeblink · Fatigue · Nonintrusive · Alert system
Computer vision

1 Introduction

Fatigue is a state of decreased awareness or alertness associated with a desire or tendency to fall asleep. It impairs judgment, causes longer reaction time, and impairs coordination. Detection of fatigue, leading to drowsiness, involves the observation of eye movements and blink patterns. Fatigue is definitely a natural phenomenon that must be dealt with well, of course by sleeping or taking rest. But fatigue in a situation that demands an attentive state could get disastrous. For

A. Pradhan (✉) · J. Sunuwar · K. Agarwal
Department of CSE, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India
e-mail: ashis.pradhan.2010@gmail.com

J. Sunuwar
e-mail: jhuma.sunuwar@yahoo.com

K. Agarwal
e-mail: mintuthekingrocks@gmail.com

S. Sharma
School of Information Technology, SRM University, Gangtok, Sikkim, India
e-mail: savna27@gmail.com

example, in the transportation sector, the attentiveness of a driver while driving is of utmost importance.

In the USA, a series of studies by the National Transportation Safety Board (NTSB) have proved that sleepiness is one of the major factors in causing accidents. The NTSB came to the conclusion that 52% of 107 single-vehicle accidents involving heavy trucks were fatigue-related. Generally, driver drowsiness is a main factor in 25% of the accidents and 60% of these redound to death [1]. Detection of fatigue can be useful in monitoring the staff and workers at places requiring high level of vigilance. Further, it can be used to detect the fatigue state of security guards. It can also be used in educational institutions to monitor the students in a classroom to detect the drowsiness of students.

2 Methodology

The first category is related to physiological studies. It includes eyelid closure, electrooculogram (EOG) [2], and electroencephalogram (EEG) [3–5]. Although the EOG provides good results, it is not suitable for a real application environment because of the attached electrodes on the users face. This is an intrusive approach and is likely to hinder the natural movement of the user. Among the other methods, eyeblink duration is the most reliable parameter for the detection of the drowsiness level [6]. As eyes are believed to contain vital information of the persons fatigue state, which can be tracked by using digital image processing techniques, and also this system is nonintrusive involving only a camera.

In this paper, we aim to develop a fatigue detection system, by using a method that can constantly monitor the eyes of the person and extract signs of fatigue or drowsiness by tracking eye movements and blink patterns and generating an alert. The system will have a nonintrusive approach and will extract fatigue information about the person by studying the eye movements and blink patterns. Also, the analyzing of the eye is done in real time, and recorded videos are not processed for the same. Our approach to this would be to use a nonintrusive system that would involve the image processing of eyes by monitoring and tracking eye movements and analyzing the blink patterns. Analyzing eyeblink patterns would help to extract the fatigue information of a person and also to develop a system that would be implementable in real-world scenarios.

Levels of fatigue states that we have considered in determining the alert level are [7]:

1. Awake State—Conscious and aware of the surroundings
2. Low Vigilance State—Low level of alertness
3. Drowsy State—A state of being lethargic; half asleep
4. Sleepy State—In a ready state to fall completely asleep

We have considered two modules:

1. **Alert Module**: In this module, based on live video feed, we continuously monitor the eyes of the person and the iris is detected and tracked. Based on the eye tracking, alert is generated on tracing a sleep condition. Also alert is generated on tracking a fixated gaze, left or right iris deflection and also on the facial movements [8].
2. **Fatigue level Module**: In this module, based on live video feed and input from the alert module, we calculate the eyeblink rate, blink amplitude, blink duration, and blink frequency. From the calculated values, fatigue level is detected and also a vacant stare (fixated gaze) condition may also be detected if traced.

Initially, the image is acquired from a live video feed, which is basically a frame. The frame is now used to extract a face image by using the Voila Jones Algorithm. If the face is detected then the face image is considered as the ROI (region of interest), further from the frame, the eye region is extracted and this is set as the new ROI. The eye image thus acquired is now flown into the function for iris detection. Based on this iris detection, we try to conclude whether the eye is opened or closed. In other words, we try to detect a blink. Studying these blink patterns, and tracking the eye gaze and movements thus, we try to conclude whether the person is in awake state, or is fatigued, or is in a sleep state. On the basis of these conclusions, we sound the alarm so that the person becomes vigilant again (Table 1).

The eye open or closed state is determined by calculating the eccentricity and major axis length of the connected component which is supposedly high in the eye closure state. Parameters like the eyeblink rate, eyeblink amplitude, eyeblink duration, and blink frequency are calculated. If the blink duration (the time period taken for a blink) is short and blink interval (time period between two blinks) is long, it is an awake state. Similarly, if the blink duration is short and the blink interval is short, then it is considered as a low vigilance state and if the blink duration is long and blink interval is very long then it is a drowsy state. Sleepy state can be determined with very long blink durations (Fig. 1).

Input image is converted into grayscale image and the grayscale image is complemented in order to highlight the high intensity eye parts. Then median filter is applied in order to filter the noise and connected components are labeled in the image and eccentricity and major axis length is found out of the connected components. If the eccentricity and the major axis length is less than the given threshold value, then proper alert is generated and this process continues. After iris detection

Table 1 Drowsiness stages based on blink behavior [9, 10]

Fatigue stage	Description
Awake	Long blink intervals and short blink durations
Low vigilance	Short blink intervals and short blink durations
Drowsy	Long blink durations
Sleepy	Very long blink durations and/or single sleep events
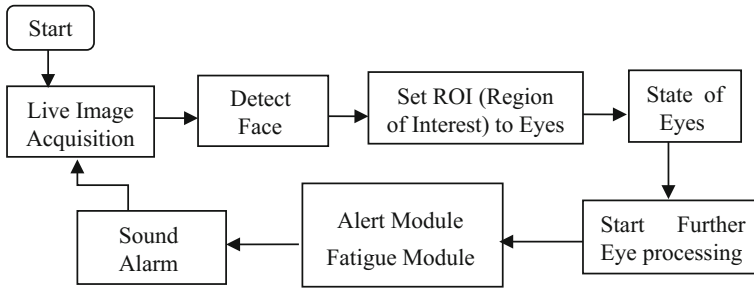
Fig. 1 Basic block diagram

process, alert module is processed. In alert module, we have considered eye opening, eye closure, and face turn detection, iris deflection toward extreme left or right and vacant stare which can be a case of day dreaming. In the Alert Module, alarms and short alerts are issued on detecting a sleep condition and on detecting extreme iris deflections, face turns, and vacant stares, respectively (Figs. 2 and 3).

In the fatigue module, we have considered four parameters:

1. Blink Amplitude—It is the time duration between two consecutive blinks.
2. Blink Rate—It is the time taken for one blink.
3. Blink Frequency—It is the number of blinks in a given time.
4. Constant Gaze—When no blink happens in a certain amount of time, then that can be the case of day dreaming.

Blink frequency and final fatigue level are calculated using the formulae:

$$\text{Blink frequency} = (\text{No. of Closed Frames}/\text{Total No. of Frames}) * 100 \qquad (1)$$

$$\begin{aligned} \text{Fatigue level} = &(0.45 * \text{Blink frequency}) \\ &+ (0.45 * \text{Blink amplitude duration}) + (0.1 * \text{Stare level}) \end{aligned} \qquad (2)$$

After the fatigue level module runs, it gets a final alert level, and then an alarm is issued based on the level of the alertness (Fig. 4).

1. Level 1—(If Final level < 1.6) Alert State—no alarm
2. Level 2—(If Final level > 1.6 & Final level < 2.3) Low Vigilance State—no alarm
3. Level 3—(If Final level > 2.3 & Final level < 3) Drowsy State—mild alarm
4. Level 4—(If Final level > 3) Sleepy State—heavy alarm (Figs. 5, 6, 7, 8 and 9).

Fig. 2 Flowchart for iris detection


```
Start
  ↓
Convert the input image to gray scale
  ↓
Complement the grayscale image
  ↓
Apply Median Filter to Filter out noise
  ↓
Remove unwanted pixels
  ↓
Label all connected components in the image
  ↓
Find eccentricity and Major Axis
length of connected component
  ↓
If eccentricity < TH
& major axis < TH
  Yes ↓        No →  Eye Close Alert
Eye Open
  ↓
Stop
```
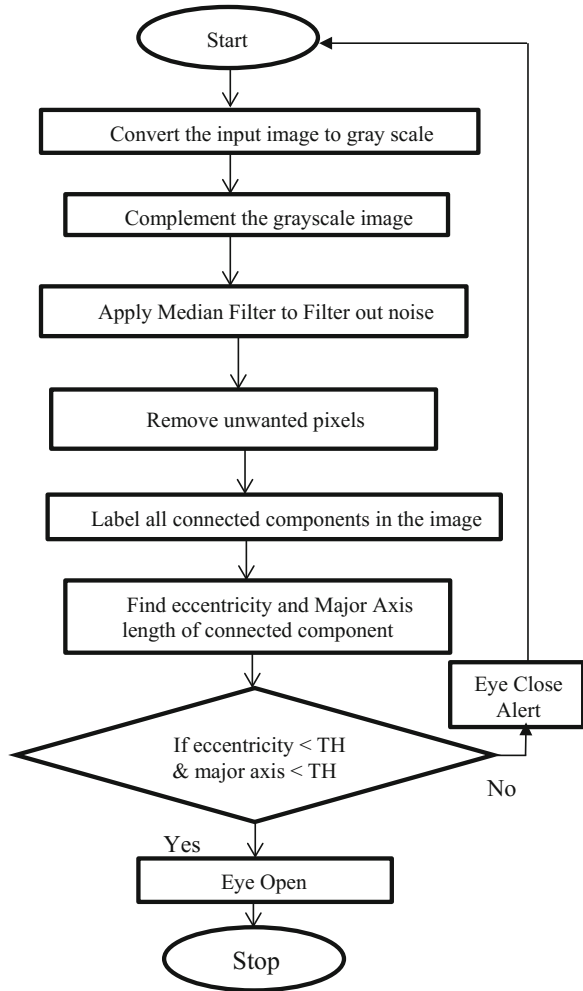
3 Results and Discussions

The system has been tested for its accuracy as well as its usability. In order to measure the accuracy of eyeblink detection, video sequences of each user sitting very close to the camera are taken. It was important to test the system's ability to determine the accuracy and performance speed of the system. So part of objective of these tests is to determine the threshold that can work for multiple users. We tested the detection modules under alert module as well as fatigue detection module. In the alert module, we tested for open eyes, closed eye, face turn, iris deflection as well as a sleep condition. In the fatigue module, we tested for the fatigue level using
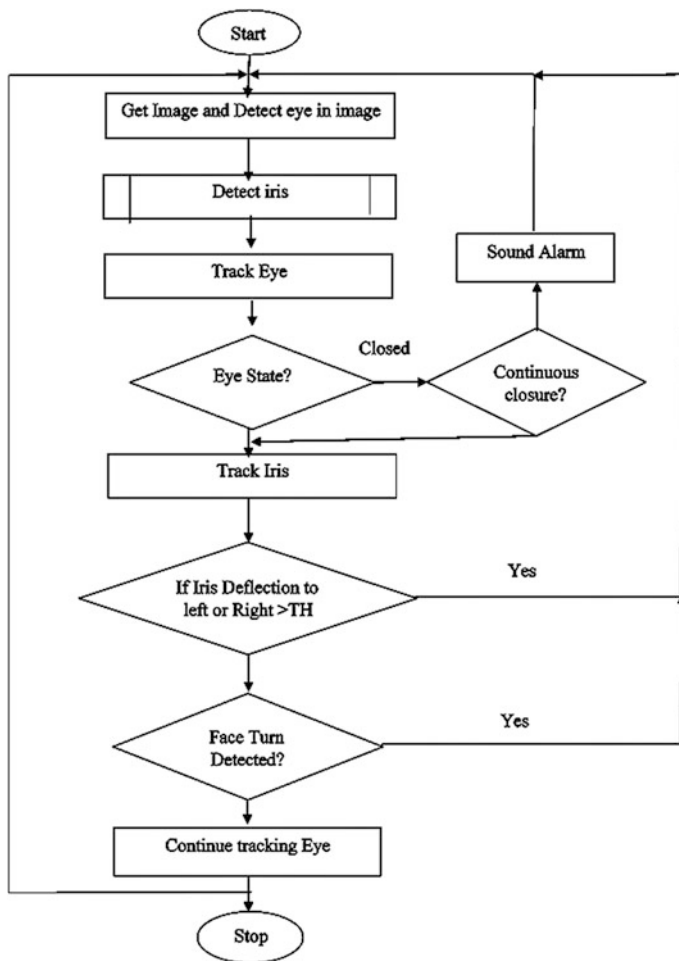
Fig. 3 Flowchart for alert module

the eyeblink rate as well as the eyeblink interval. We also tested for the vacant stare condition of the user (Tables 2, 3 and 4).

The average accuracy of the alert module is 88% and the average accuracy of the Fatigue module is 85%. The detected iris is used to determine the condition of the eye. The centroid of the iris is determined which is compared with iris's centroid of previous frames and change in centroid results in the deflection of the iris. Also, we keep tracking both the eyes and in case only one eye is detected, face turn alert is given.
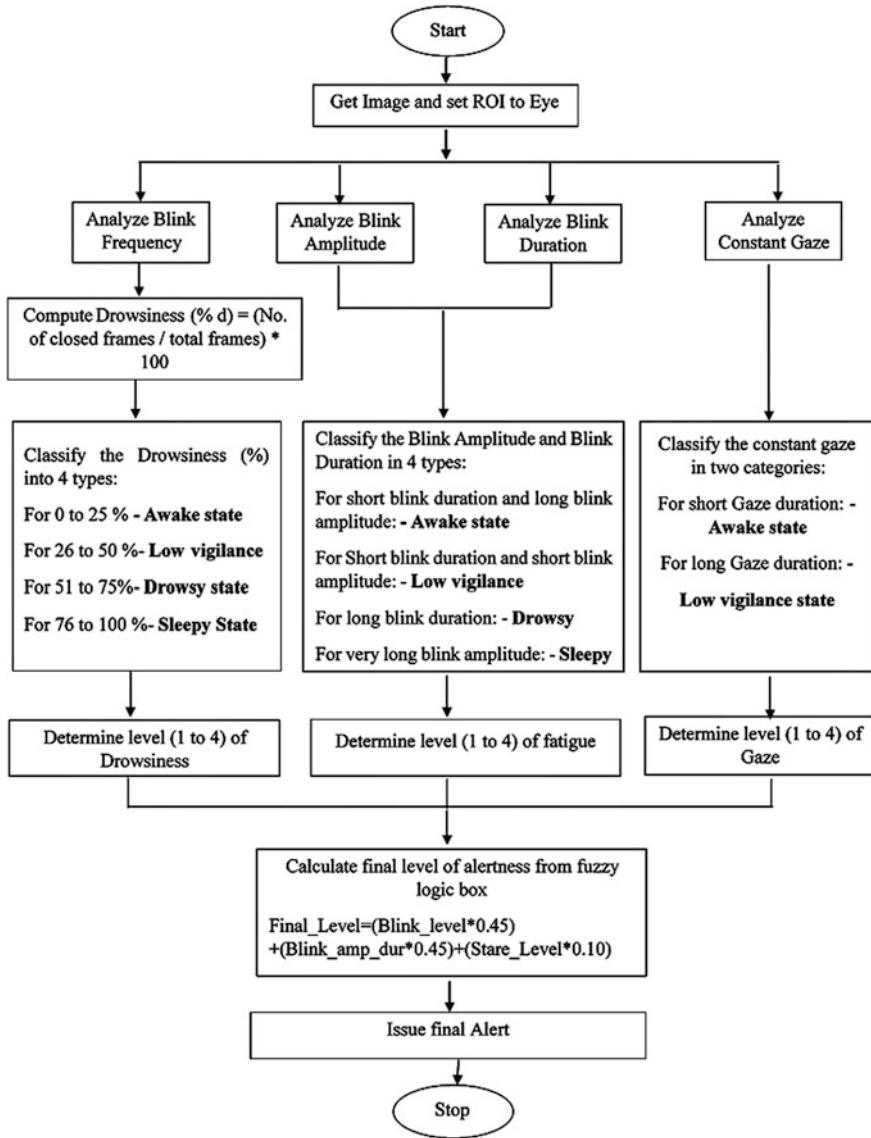
Fig. 4 Flowchart for fatigue module

Fig. 5 Left iris deflection detection

Fig. 6 Right iris deflection detection
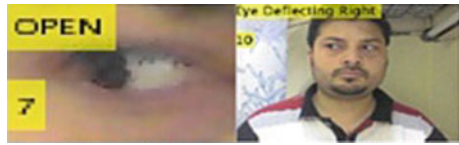
Fig. 7 Face turn toward left detection
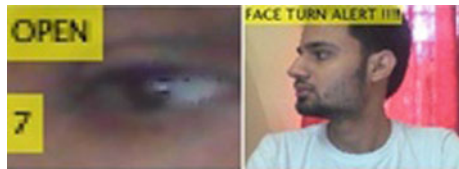
Fig. 8 Face turn toward right detection
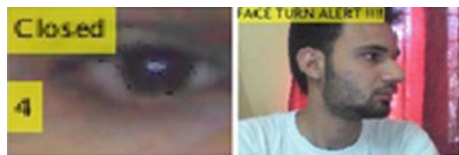
Fig. 9 Vacant stare detected
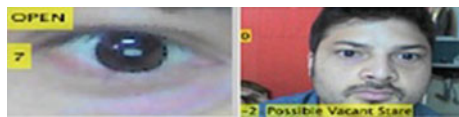
Table 2 Testing for alert module

Test cases	Wrongly identified	Rightly identified	Total no. of samples
Eye open	03	17	20
Eye close	01	19	20
Face turn	00	20	20
Iris deflection	05	15	20
Sleep detection	03	17	20

Table 3 Percentage calculation for accuracy

Test case	Eye open	Eye close	Face turn	Iris deflection	Sleep detection	Accuracy (%)
(%)	85	95	100	75	85	88

Table 4 Accuracy for fatigue module

Test cases	Wrongly identified	Rightly identified	Total no. of samples
Alert level	03	17	20

4 Summary and Conclusions

Lack of sleep, improper diet lacking nutrition, tiredness, and lack of exercise often lead to a fatigue state, which can prove to be hazardous and fatal in case a person is required to be vigilant in certain situations like while driving, air traffic controlling, guard personnel at nuclear reactors, security guard cabins, and so on. Although a proper sleep is the only solution to cure a fatigued person, yet to help the person remain vigilant, necessary alarms and warning systems at the right time can help the person to remain awake. Also, if the iris deflects a lot toward the extreme left or right, then a safety warning alert is issued to prevent distraction. The module was tested and it could detect fatigue and alertness with 88% and 85% accuracy respectively. It can be further implemented for the designing of a proper product after necessary removal of the limitations. The system cannot be used for a person wearing spectacles or sunglasses due to the reflection of light and can be further worked upon to bypass the glasses of the spectacles and thus, read the eye in a proper manner [11].

References

1. Bergasa, L.M., Nuevo, J.U., Sotelo, M.A., Barea, R., Lopez.: Visual monitoring of driver inattention. Studies in Computational Intelligence (SCI) (2008)
2. James, B., Sharabaty, H., Esteve, D.: Automatic EOG analysis: a first step toward automatic drowsiness scoring during wake-sleep transitions. Somnologie, vol. 12, pp. 227–232 (2008)
3. Santamaria, J., Chiappa, K.: The EEG of drowsiness in normal adults. J. Clin. Neurophysiol. **4**, 327–382
4. Vuckovic, A., Radivojevic, V., Chen, A.C.N., Popovic, D.: Automatic recognition of alertness and drowsiness from EEG by an artificial neural network. Med. Eng. Phys. **24**(5), 349 (2002)
5. Nikhil, R.P., Chien–Yao, C., Ko, L.W., Chao, C.F., Jung, T.P., Liang, S.F., Lin, C.T.: EEG–based subject–and session–independent drowsiness detection: an unsupervised approach. EURASIP J. Adv. Signal Process. (2008). ISSN: 1110-8657
6. Dinges, D.: PERCLOS: a valid psycho physiological measure of alertness as assessed by psychomotor vigilance Indianapolis. Federal Highway Administration, Office of motor carriers, Technical report, MCRT-98-006 (1998)
7. Malla, A.M., Davidson, P.R., Bones, P.J., Green, R., Jones, R.D.: Automated video-based measurement of eye closure for detecting behavioral microsleep. In: 32nd Annual International Conference of the IEEE EMBS, Buenos Aires, Argentina, 31 Aug–4 Sept 2010
8. Parmer, S.H., Jajal, M., Brijbham, J.P.: Drowsy driver warning system using image processing. Int. J. Eng. Dev. Res.

9. Hargutt, V., Krüger, H.P.: In: Eyelid Movements and their Predictive Value of Fatigue Stages, Würzburg, Centre for Traffic Sciences (2000)
10. Devi, M.S., Bajaj, P.R.: Driver fatigue detection based on eye tracking. In: 1st International Conference on Emerging Trends in Engineering and Technology, pp. 649–652 (2008)
11. Pradhan, A., Ray, A.: Face detection in night vision images. Int. J. Res. Commer. IT Manage. **4**(6) (2014)
12. Kumtepe, O., Akar, G.B., Yuncu, E.: Driver aggressiveness detection via multisensory data fusion. EURASIP J. Image Video Process. (2016)

Secure Symmetric Key Transmission of Messages Using Random Shuffling of Spiral Matrix and Multiplicative Inverse (RSSMMI)

Sarbajit Manna, Soumya Banerjee, Prantik Panja, Ramkrishna Das and Saurabh Dutta

Abstract A bit level block cipher based symmetric key cryptographic technique is proposed here. Entire plaintext file reads two characters at a time and according to the binary representation of ASCII value of these characters, 16-bit blocks are created and they are represented in a 4 × 4 spiral matrix. Then using random shuffling of spiral matrix and multiplicative inverse, 8-bit ciphertext blocks are generated for each 16-bit plaintext block. So here 50% compression is achieved in terms of size of ciphertext file. Apart from compression, RSSMMI technique has several advantages like formation of symmetric key dynamically and randomly, security, and equal applicability of this technique for large number of files of almost any type.

Keywords Symmetric key · Block cipher · Spiral matrix · Random shuffling Multiplicative inverse

S. Manna (✉) · S. Banerjee · P. Panja
Department of Computer Science, Ramakrishna Mission Vidyamandira,
Belur Math, Howrah 711202, West Bengal, India
e-mail: sarbajitonline@gmail.com

S. Banerjee
e-mail: soumyabanerjee1996@outlook.com

P. Panja
e-mail: prantikpanja@gmail.com

R. Das
Department of Computer Applications, Haldia Institute of Technology, Haldia,
West Bengal, India
e-mail: ramkrishnadas9@gmail.com

S. Dutta
Department of Computer Applications, Dr. B. C. Roy Engineering College,
Durgapur 713206, West Bengal, India
e-mail: saurabhdutta06061973@gmail.com

1 Introduction

Cryptography is the practice and study of techniques for secure communication between sender and receiver in the presence of third parties. Application of cryptography includes ATM cards, computer passwords and e-commerce. In symmetric key cryptography [1, 2], sender and receiver of a message agree upon a common key. The sender uses the common key to encrypt the message which is sent to the receiver over the network and the key is transmitted through a secure channel. The receiver decrypts the encrypted message using the key to get back the original message. To improve efficiency of the proposed RSSMMI encryption technique, a pseudorandom number is [3] issued during symmetric key generation.

RSSMMI technique with example is discussed in Sect. 2. Section 3 contains results on different types of files and comparison of the RSSMMI technique with AES-256 bit technique along with brief description of the technique. Section 4 draws the conclusion.

2 Proposed Technique with Example

2.1 Encryption Technique

The entire plaintext file reads two characters at a time and a ciphertext character for every two plaintext characters is generated at the end of encryption.

2.1.1 Generation of First Byte of the Symmetric Key and Randomly Shuffled Matrix

Let the plaintext be AD. The ASCII values of A and D are 65 and 68 which are 01000001 and 01000100 in binary respectively. The complement of A, A′ is 10111110. Corresponding binary representation of A′ D is 1011111001000100. These bits are taken from LSB to MSB and fit into the square matrix of order 4 × 4 following the concept of spiral matrix [4] along clockwise direction starting from (1, 1) cell (Figs. 1 and 2).

Now, this 4 × 4 matrix is broken into four 2 × 2 matrices with their corresponding location numbers are shown in Fig. 3. Then, a set S consisting of elements 0, 1, 2, 3 is taken to select randomly [5] the elements of S one by one such

Fig. 1 Concept of spiral matrix
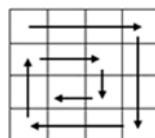
Fig. 2 Spiral matrix representation of the 16-bit block

0	0	1	0
1	1	1	0
1	1	0	0
1	0	0	1

Fig. 3 Spiral matrix before and after randomly shuffling
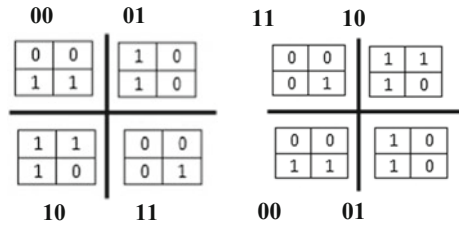
Fig. 4 First byte of symmetric key

H ⟶ | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
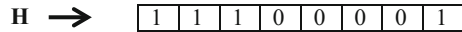
that none element of S is selected twice. Let 3, 2, 0, 1 be the randomly selected elements of S and their corresponding binary representations are 11, 10, 00, and 01, respectively. According to the randomly selected elements of S, a randomly shuffled [6] spiral matrix is generated.

The first byte of symmetric key, H is generated by storing the binary representations of the randomly selected elements (11, 10, 00, 01) of S (Fig. 4).

2.1.2 Generation of Second Byte of the Symmetric Key

Bitwise exclusive OR operations (^) are performed between all the rows of the randomly shuffled matrix starting from top to bottom and in a row from left to right direction to produce a 4-bit block of zeroes and ones. Then, bitwise exclusive OR operations are performed between all the columns of the randomly shuffled matrix starting from leftmost to rightmost column and in a column from bottom to top direction to produce another 4-bit block of zeroes and ones. By combining these 4-bit blocks, second byte of the symmetric key is generated.

Here, 1001 and 1100 are generated after row wise and columnwise exclusive operations respectively. The second byte of the symmetric key, I is shown in Fig. 5.

Fig. 5 Second byte of symmetric key
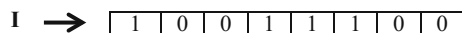
I ⟶ | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |

1	0	1	1	1	1	1	0	0	1	0	0	0	1	0	0
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Fig. 6 16-bit block representation of A′ D with corresponding positions

2.1.3   Generation of Third Byte of the Symmetric Key

To generate the third byte of symmetric key, initial 16-bit block representation of A′ D (1011111001000100) is considered. This 16-bit block along with their corresponding positions is shown in Fig. 6.

Algorithm 1: Third byte of Symmetric Key Generation

Procedure: 16-bit block representation of A′D is stored in array B. 8-bit binary representation of third byte of symmetric key will be contained in P array. ^, | & are bitwise exclusive OR, OR and AND operations respectively.
k ← 0;
for i ← 0, j ← 15; i<= 3, j >= 12; i ← i + 1, j ← j - 1 do
P[k] ← B[i] ^ B[j];
k ← k + 1;
for i ← 4, j ← 11; i<= 6, j >= 9; i ← i + 2, j ← j - 2 do
P[k] ← B[i] | B[j];
k ← k + 1;
for i ← 5, j ← 10; i<= 7, j >= 8; i ← i + 2, j ← j - 2 do
P[k] ← B[i] & B[j];
k ← k + 1;
P1← 4 bit random binary number.
for i ← 0, j ← 0; i<= 3, j <=6; i ← i + 1, j ← j + 2 do
P[j] ← P1[i] ^ P[j];

According to the above algorithm, third byte of the symmetric key, P is shown in Fig. 7.

2.1.4   Extraction and Swapping of Data Bytes from Randomly Shuffled Matrix

Data bits are extracted from shuffled matrix by dividing them into two bytes. First byte is generated by combining the first and fourth row elements of the shuffled matrix from left to right and second byte is generated by combining the third and

Fig. 7 Third byte of symmetric key

P →

0	0	0	1	0	1	1	0

second row elements of shuffled matrix from right to left. Those are stored in E (00111110) and F (01000110) respectively.

A bitwise exclusive OR operation (^) is done between first and second byte of symmetric key, i.e., H and I. If the ASCII value of the result is not even then all even positioned bits of data byte E is swapped with all odd positioned bits of data byte F from right to left. Otherwise, all odd-positioned bits of data byte E are swapped with all even-positioned bits of data byte F from right to left.

Here, as H (11100001) ^ I (10011100) = 01111101 = 125_{10} and it is not even, so all even positioned bits of data byte E (00111110) are swapped with all odd positioned bits of data byte F (01000110) from right to left. So, E and F become 00101011 and 01101100, respectively.

Now, bitwise exclusive OR operations are performed between E (00101011), I (10011100) and F (01101100), P (00010110) and the results are stored in E and F, respectively. So, E and F become 10110111 and 01111010, respectively.

2.1.5 Generation of Fourth and Fifth Byte of the Symmetric Key Using Multiplicative Inverse

Decimal values of E and F are computed. If the decimal value is zero then its multiplicative inverse is also zero. Otherwise, multiplicative inverse [7] is computed with respect to prime number 257. Multiplicative inverse of any integer from 1 to 255 with respect to 257 is between 1 and 255 (both inclusive) and multiplicative inverses of 0 and 256 with respect to 257 are 0 (assumption) and 256, respectively. To ensure getting unique 8-bit multiplicative inverse, it is computed with respect to 257.

Here, E = 183_{10} and F = 122_{10} and the multiplicative inverses of E and F with respect to 257 are 191 and 158, respectively. Binary representations of these multiplicative inverses are stored in E and F respectively. So, now E and F become 10111111 and 10011110, respectively. These are combined as a 16-bit block (1011111110011110) which is 49054 in decimal and is stored in a variable ASC.

Now ASC % 6542 is performed (% is modulus operation) and the result, 3260 is stored in y. From 0 to $2^{16} - 1$, there are total 6542 primes. Let all the primes are stored in an array L in nondecreasing order. Then yth prime (3260th) is selected from L and the result (30139) is stored in z and which is 0111010110111011 in binary. This 16-bit block is broken into two 8-bit blocks M and N such that M and N contains leftmost and rightmost 8-bits, respectively. M (01110101) and N (10111011) are stored as fifth and fourth byte of the symmetric key (Fig. 8).

Fig. 8 Fourth and fifth byte of symmetric key

N → | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 |
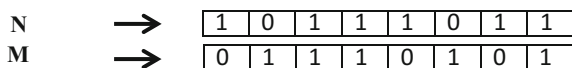
M → | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 |

Fig. 9 Sixth byte of symmetric key
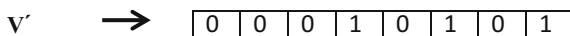
V′ ⟶ | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |

2.1.6 Generation of Sixth Byte of the Symmetric Key

Bitwise exclusive OR operations are performed between E (10111111) and N (10111011) to produce U (00000100) and between F (10011110) and M (01110101) to produce V (11101011). 2's complement of V is computed and the result is stored in V′(00010101) as the sixth byte of symmetric key (Fig. 9).

2.1.7 Generation of Last (Seventh) Byte of the Symmetric Key and the Ciphertext

Bitwise exclusive OR is performed between U and V′ and the result (00010001) is stored in T.

Algorithm 2: Generation of Last Byte (Seventh) of Symmetric Key and the Ciphertext

---

Procedure: Initially n will contain ASCII value of T. y and G will contain the last (seventh) byte of symmetric key and the ciphertext.

def SquareRoot(n)

    $R \leftarrow n$;

    $y \leftarrow func(R)$;

    $x \leftarrow y^2 - R$;

    $G \leftarrow y^2 - (y/2)^2 + x$;

Return G, y;

def func(n)

    $x \leftarrow sqrt(n)$;

    $y \leftarrow floor(x)$;

    if $(x-y) > 0$ then

      return y+1;

    else

      return y;

Here, T (00010001) in decimal is 17. After applying Square Root function on T, ciphertext, C becomes 29 (GS character) and last (seventh) byte of symmetric key D becomes 5 which is 00000101 in binary (Fig. 10).

So, the 56-bit symmetric key from first byte to seventh byte for the plaintext AD becomes H (11100001), I (10011100), P (00010110), N (10111011),

Fig. 10 Last (seventh) byte of symmetric key
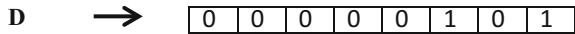
D ➝ | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |

M (01110101), V′ (00010101), D (00000101) and corresponding ciphertext character is Group Separator (GS) character (ASCII value 29).

2.2 Decryption Technique

Algorithm 3: Definition of Reverse Square Root Function

Procedure: C is the ciphertext and D is the last (seventh) byte of symmetric key.
def ReverseSquareRoot(C,D)
    $E \leftarrow 2*D^2 - C - (floor(D/2))^2$;
Return E;

Applying Reverse Square Root function on ciphertext file and last byte of symmetric key and then reverse procedure of encryption, the plaintext file is generated.

3 Result and Comparison

The technique has been implemented using Java [8] on a computer with Intel Core i7 3.6 GHz processor having 8 GB DDR3 RAM. Pearsonian chi-squared test [9] has been performed to test the non-homogeneity between the source and the encrypted files with $(256 - 1) = 255$ degrees of freedom (Table 1).

It is seen from Fig. 11 that the encryption and decryption times are independent of file type. From Fig. 12, it is clear that though chi square values for AES-256 bit technique are higher than RSSMMI technique in most cases, but the RSSMMI technique produces reasonably good chi square values.

4 Conclusion

The symmetric key of RSSMMI technique is based on the plaintext block, random shuffling of spiral matrix and is dynamic in nature. For the same plaintext block, it is very less probable that the same symmetric key will be generated for the next time because of random shuffling. The space requirement of the ciphertext files is

Table 1 Results for different types of files

Source file size (bytes)	File type	Encryption time (ms)	Decryption time (ms)	Values for RSSMMI technique		Values for AES-256 bit technique	
				Chi square value	Degree of freedom	Chi square value	Degree of freedom
5,632	.exe	734	62	14213.12	247	31391.92	255
18,432	.exe	906	203	17378.14	255	37503.92	255
15,360	.exe	859	140	46938.50	254	117656.52	255
8,192	.dll	781	78	22817.55	231	65462.98	255
4,096	.dll	719	47	11515.56	246	25879.49	255
13,312	.dll	875	125	53458.45	254	131359.00	255
69,886	.com	1594	578	92535.88	255	165345.39	255
29,696	.com	1063	250	86859.23	255	189364.90	255
1,131	.com	672	16	1080.03	219	1549.39	254
4,224	.sys	719	47	5997.11	244	16627.13	255
16,168	.sys	875	140	15379.70	255	29337.20	255
22,528	.sys	954	203	40221.70	255	108282.47	255
833	.txt	671	15	630.86	176	527.60	252
32,652	.txt	1094	265	51295.17	209	118768.42	255
23,015	.txt	968	203	35619.36	209	67761.43	255
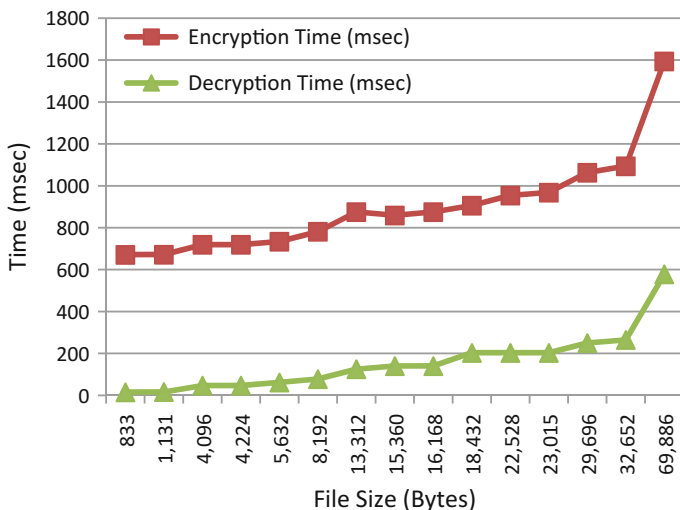
Fig. 11 Graphical representation of encryption time and decryption time versus file size
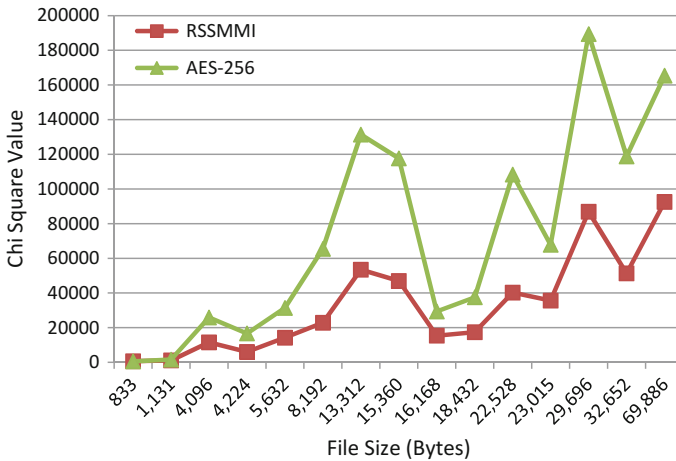
Fig. 12 Graphical comparison of chi square values for RSSMMI and AES-256 technique

half of that of the plaintext files. This technique is very simple to implement using any high level language and is equally applicable to any type of files.

References

1. Forouzan, B.A., Mukhopadhyay, D.: Cryptography and Network Security, 2nd edn. Tata McGraw Hill Education Private Limited
2. Stallings, W.: Cryptography and Network Security: Principles and Practice, 2nd edn. Pearson Education Asia, Sixth Indian Reprint (2002)
3. https://en.m.wikipedia.org/wiki/Pseudorandomness
4. Paul, M., Mandal, J.K.: A novel symmetric key cryptographic technique at bit level based on spiral matrix concept. In: International Conference on Information Technology, Electronics and Communications (ICITEC—2013), Bangalore, India, 30–31 Mar 2013
5. https://en.m.wikipedia.org/wiki/Randomness
6. Gupta, A., Mandal, V.: Enhancing AES algorithm using random shuffle method. Int. J. Innovative Res. Comput. Commun. Eng. **5**(3) (2017). ISSN (Online): 2320–9801, ISSN (Print): 2320-9798
7. Hasan, M.A.: Efficient computation of multiplicative inverses for cryptographic applications, 0-7695-1150-3/01. IEEE (2001)
8. Schildt, H.: Java: The Complete Reference, 9th edn. Tata McGraw Hill
9. Dutta, S.: An approach towards development of efficient encryption technique. Ph.D. thesis, The University of North Bengal (2004)

Odor Source Localization by Concatenating Particle Swarm Optimization and Grey Wolf Optimizer

Upma Jain, Ritu Tiwari and W. Wilfred Godfrey

Abstract A concatenated approach which utilizes the strength of Particle Swarm Optimization (PSO) and Grey Wolf Optimizer (GWO) is proposed for odor source localization by a team of mobile robots. Odor plume is modeled by using the Gaussian distribution. Robots continue random search within the workspace to locate the plume. When one of the robot enters in the vicinity of plume, robot's new positions are calculated by applying concatenation of PSO first then Grey Wolf Optimizer second and vice versa. In order to prevent getting stuck at local minima, concept of search counter is used. Proposed approach is compared with Refined Hybrid PSO and the simulation result shows the validity of the proposed approach over the other.

Keywords Grey wolf optimizer · Particle swarm optimization · Odor source localization · Multi-robot system

1 Introduction

Information of odor is widely used by animals to find food, locate mate and information sharing. Ability of animals to use the odor information has attracted researchers to use robots for finding the odor source in an area. Using robots for detecting the odor source has valuable applications in real life such as toxic gas leak detection, localizing source of fire, etc.

Metaheuristic methods inspired from the natural phenomena such as Particle Swarm Optimization (PSO) [1], Ant Colony Optimization [18], etc., have been applied in many areas including robotics [8, 9, 14] to solve many problems. Among

U. Jain (✉) · R. Tiwari · W. W. Godfrey
ABV-IIITM Gwalior, Gwalior, Madhya Pradesh, India
e-mail: upma.jain88@gmail.com

R. Tiwari
e-mail: tiwariritu2@gmail.com

W. W. Godfrey
e-mail: godfrey@iiitm.ac.in

various nature inspired algorithms PSO [1] has been applied extensively to solve many problems. PSO was proposed by Kennedy and Ebarhart in 1995. PSO is an heuristic intelligence method inspired from bird flocking and fish schooling. Grey Wolf Optimizer (GWO) [15] is a new nature-inspired algorithm proposed by Seyedali Mirjalili et al. in 2014. It is developed on the basis of hunting behavior of grey wolves. In GWO, population is governed by utilizing first three best solutions, which helps to maintain diversity. GWO is preferred because in GWO encircling behavior of prey helps to keep robots within the plume area. To achieve the balance between exploration and exploitation of search space, GWO is used with PSO.

Balancing between exploration and exploitation plays an important role in any control algorithm. Our objective is to utilize the skills of both GWO and PSO to achieve the balancing between exploration and exploitation of the workspace.

Rest of the paper is organized as follows. Section 2 describes the related work. Methodology is given in Sect. 3. Section 4 provides details of simulation setup and results. Conclusion and future work is covered in Sect. 5.

2   Related Work

Research in the field of odor source localization began in the 1990s. At first, single robots were used for odor source localization [5, 10, 16]. With increasing capability of robots and advancement in technology, team of mobile robots was used for odor source localization [2, 9, 18]. Task of odor source localization can be divided into three subtasks namely: plume finding, plume following, and odor source declaration. Random search [6, 9] has been preferred by most of the researchers for plume finding. For odor source declaration, some researchers used the vision-based information [6] while others have considered the odor source to be found when robot reaches in a certain vicinity area [7, 10, 16]. Most of the work is done in the subtask plume following.

In [5], Ishida et al. proposed a new exploratory algorithm for odor source localization. A switchover strategy was used to find and track the plume. Plume tracking was done on the basis of concentration gradient information. Algorithm is able to locate the ethanol from almost anywhere in the room.

In [13], Lino Marques and Anibal T. de Almeida proposed an approach to localize a specific odor source in presence of multiple odor sources. Odor identification was done by using an artificial neural network based pattern recognition algorithm. Odor source localization was performed by three different algorithms in presence of obstacle. Results show that bacteria algorithm converges slowly as compared to moth and gradient-based algorithm.

In [12], Lu et al. give a distributed approach for odor source localization. Position of odor source is predicted by using PSO. Two kinds of information odor concentration and wind speed are used to predict the next move of the robots in the workspace by utilizing two finite time control algorithms.

In [14], Marques et al. presented a new PSO based cooperative algorithm for odor source localization. Algorithm is based on the information exchange among neighboring agents for tracking odor sources. Spreading of agents around the work space is achieved by integrating a global searching behavior with repulsive forces among the agents and crosswind biased motion.

In [9], Jatmiko et al. give a nichie PSO based cooperative approach for odor source localization. Nichie is formed when robot enters in the vicinity of plume area. Performance is improved by dynamically adjusting the size of nichie. Experiments are done to localize multiple odor sources. Results confirm the effectiveness of this approach.

In [11], Li et al. proposed an approach for odor source localization in a ventilated room. PSO based on probability fitness function is used to localize the odor source. Probability fitness function is estimated by Bayesian and fuzzy inference. However, this algorithm is limited to localize only single odor source in the environment.

In [3], Gong et al. proposed an improved PSO-based approach for odor source localization. This algorithm dynamically adjusts the two learning factors and considers wind effect on each particle.

In [17], Yong Zhang et al. proposed a Refined Hybrid PSO by combining Bacterial Foraging Optimization with PSO for odor source localization. Two operations namely chemotaxis and elimination dispersal were integrated with PSO. Chemotaxis helps the robots to follow the plume while elimination dispersal is adopted to prevent stagnation in local optima. Simulation results show the validity of this approach.

3 Methodology

Robots enter from corner of the workspace and start looking for the plume. During plume finding phase, next goal positions of robots are calculated randomly. Once any one of the robot finds the plume, i.e., (concentration > Th1), robots enter into the plume following phase. In this phase, new goal positions of robots are determined by the proposed concatenated method. Robots continue to track the plume until they find maximum intensity area, i.e., (concentration = 255). When any one of the robot finds maximum intensity area, odor source is assumed to be located. Procedure of the proposed method is given in Algorithm 1.

3.1 Proposed Concatenated Method

In plume following phase, next goal position of robots is calculated by both of the possible combinations of PSO and GWO such as, PSO first then GWO second and GWO first and PSO next called as Algorithm 2 (PSO-GWO) and Algorithm 3 (GWO-PSO).

Algorithm 1 Search algorithm

1: Robots initialization;
2: Robots start search from the corner of the workspace; $R_i = (R_1, R_2, ...R_6)$
3: **while** odor source is not found or maximum iteration is not reached **do**
4:     For each robot calculate fitness
5:     Fitness = intensity measured at time t;
6:     **if** fitness>Th1 and search counter < SC **then**
7:         Call Concatenated Method
8:     **else**
9:         Search randomly, assign robots new target positions randomly.
10:    **end if**
11: **end while**

3.1.1 Grey Wolf Optimizer

GWO is inspired from the grey wolves. It works by mimicking the hunting mechanism and leadership hierarchy of grey wolves. Three main steps namely: searching the prey, encircling the prey, and attacking the prey are implemented in hunting mechanism. To simulate the social hierarchy of grey wolves four kinds of grey wolves such as alpha, beta, delta, and omega are defined. Wolves having the highest fitness are considered as alpha. Consequently, next two fittest wolves are considered as beta and delta respectively. Rest of the wolves are considered as omega. Position of omega wolves is determined by using the alpha, beta, and delta wolves position. To mathematically model GWO robot's fitness is considered as the solution. Grey wolf encircling behavior is modeled by the Eqs. 1 and 2.

$$D = | C * X(pr)(t) - X(t) | \tag{1}$$

$$X(t + 1) = | X(pr)(t) - A \odot D | \tag{2}$$

t: current iteration $X(pr)$: prey position vector X(t): grey wolf position vector A and C: Coefficient vectors

$$A = 2 * a * r_1 - a \tag{3}$$

$$C = 2 * r_2 \tag{4}$$

a: linearly decreasing from 2 to 0 over course of iterations r_1, r_2 : random number vectors from 0 to 1.

Positions of alpha, beta and delta are calculated by evaluating formulas from 5 to 10.

$$D_{alpha} = | C_1 * X_{alpha} - X | \tag{5}$$

$$D_{beta} = | C_2 * X_{Beta} - X | \tag{6}$$

$$D_{delta} = | C_3 * X_{delta} - X | \tag{7}$$

$$X_1 = |X_{alpha} - A1 \odot D_{alpha}| \tag{8}$$

$$X_2 = |X_{beta} - A2 \odot D_{beta}| \tag{9}$$

$$X_3 = |X_{delta} - A3 \odot D_{delta}| \tag{10}$$

Position of Remaining robots is calculated by Eq. 11.

$$X(t+1) = (X_1 + X_2 + X_3)/3 \tag{11}$$

3.1.2 Particle Swarm Optimization

PSO works by mimicing the behavior of fish schooling and birds flocking. In PSO, particle's new positions are calculated by global best solution and particle's previous best solution. To update robot's new goal position fitness of robots is evaluated based on the intensity. Then their new velocity and position are calculated using Eqs. 12 and 13.

$$v_m(t+1) = w * v_m(t) + c_1 * rand_1 * (pbest_m - x_m(t)) + c_2 * rand_2 * (gbest - x_m(t)) \tag{12}$$

$$x_m(t+1) = x_m(t) + v_m(t+1), \tag{13}$$

where $v_m(t)$, $v_m(t+1)$: velocity of robot m at time t and $t+1$. w inertia weight. c_1, c_2: positive constants.

$rand_1$, $rand_2$: random numbers varying from 0 and 1. $pbest_m$ is previous best position of robot m, gbest is the position of best performing robot. $x_m(t)$ and $x_m(t+1)$ are the position of robot m at time t and t+1.

Algorithm 2 Algorithm 2 PSO-GWO

1: Evaluate each robot's fitness;
   fitness = intensity measured at time t;
2: Find pbest and gbest based on the fitness.
3: Update velocity and position by equations 12 and 13.
4: Feed these updated position as initial solutions in GWO.
5: Find alpha, beta and delta based on the fitness.
6: Evaluate formulas form 5 to 10.
7: Update final position of robots by equation 11.

Algorithm 3 Algorithm 3 GWO-PSO

1: Evaluate each robot's fitness;
   fitness = intensity measured at time t;
2: Find alpha, beta and delta based on the fitness.
3: Evaluate formulas form 5 to 10.
4: Update position of robots by equation 11.
5: Feed these updated position as initial solutions in PSO.
6: Find pbest and gbest based on the fitness.
7: Update velocity and position by equations 12 and 13.

4 Simulation Setup and Results

Simulation is done in MATLAB environment. An obstacle free workspace is considered for simulation. Environment size is 330 * 430. A single odor source with constant release rate is assumed to be placed in the environment. Wind is considered to be stable with a constant velocity of 0.5 m/s. Distribution of odor plume is modeled as Gaussian distribution [4]. Figure 1 shows the odor model. Intensity of odor across the environment is modeled in an image form. Two thresholds have been used Th1 (>0) and Th2 (=255). When measured intensity is higher than 0 (Th1), robot is considered to be in the vicinity of plume area. Value of parameters used in simulation are given in Table 1. Proposed Algorithms 1 and 2 are compared with Refined Hybrid PSO methods. All three methods were run 50 times to evaluate the average time needed to localize the odor source. Performance evaluation is done on the basis of success rate and time required to finish the search. Table 2 shows the comparison among Algorithms 2 and 3 and Refined Hybrid PSO. It can be seen from Table 2 that time required to locate odor source by both Algorithms 1 and 2 are less as compared to refined hybrid PSO. Also Algorithms 1 and 2 have higher success rates as compared to refined hybrid PSO method. Success rate of proposed method

Fig. 1 Odor source model

Table 1 Odor source model parameters

Parameter	Value
Team size	6
C_1	1.6
C_2	0.6
w	0.5

Table 2 Time required to locate odor sources and success rate

Method	Time required (s)	Success rate (%)
Algorithm 1	28.1773	100
Algorithm 2	17.3251	100
Refined hybrid PSO	67.3202	99

Fig. 2 Performance comparison based on time required to find the maximum intensity area
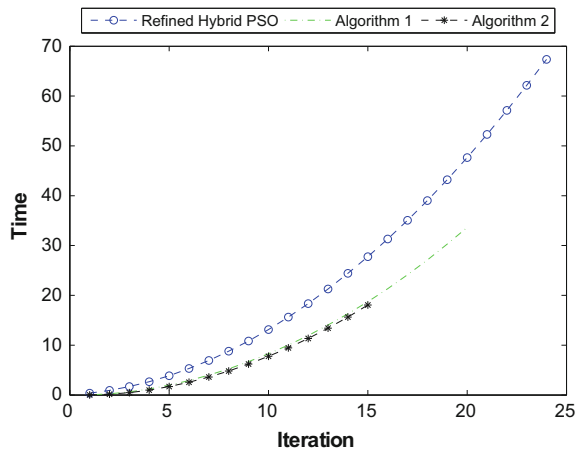
is higher because of the search counter concept. Percentage of success rate may vary in different environments. In this paper, performance evaluation of both approaches is done under assumption like no robot failure, perfect communication, no collision, stable environment, plume follows the Gaussian distribution, and sensors have no errors. Graph of time required for localizing the odor source versus iterations is given in Fig. 2. From graph, it could be observed that proposed Algorithms 1 and 2 converge much faster as compared to refined hybrid PSO method. Because proposed approach achieves better balancing between exploration and exploitation. Algorithm 2 performs best. In this, first PSO has been applied which reduces the search space by moving all the robots toward gbest then on that reduced search space, GWO is applied which helps to maintain diversity within reduced search space.

5   Conclusion and Future Work

Problem of odor source localization is studied and concatenation of PSO first then GWO second and vice versa is used to localize the odor source by a team of mobile robots. Simulation is done to ensure the validity of proposed approach. Results obtained through simulation show that both of the Algorithms 2 and 3 outperform the Refined Hybrid PSO-based approach. Future work involves study of multiple odor source localization problem in time-varying environment by a team of mobile robots.

References

 1. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science. MHS'95, pp. 39–43. IEEE (1995)
 2. Ferri, G., Caselli, E., Mattoli, V., Mondini, A., Mazzolai, B., Dario, P.: Spiral: a novel biologically-inspired algorithm for gas/odor source localization in an indoor environment with no strong airflow. Robot. Auton. Syst. **57**(4), 393–402 (2009)
 3. Gong, D.W., Qi, C.L., Zhang, Y., Li, M.: Modified particle swarm optimization for odor source localization of multi-robot. In: 2011 IEEE Congress on Evolutionary Computation (CEC), pp. 130–136. IEEE (2011)
 4. Hinze, J.: Turbulence, (1975). New York (1959)
 5. Ishida, H., Kagawa, Y., Nakamoto, T., Moriizumi, T.: Odor-source localization in the clean room by an autonomous mobile sensing system. Sens. Actuators B Chem. **33**(1–3), 115–121 (1996)
 6. Ishida, H., Tanaka, H., Taniguchi, H., Moriizumi, T.: Mobile robot navigation using vision and olfaction to search for a gas/odor source. In: Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004), vol. 1, pp. 313–318. IEEE (2004)
 7. Jain, U., Godfrey, W.W., Tiwari, R.: A hybridization of gravitational search algorithm and particle swarm optimization for odor source localization. Int. J. Robot. Appl. Technol. (IJRAT) **5**(1), 20–33 (2017)
 8. Jatmiko, W., Jovan, F., Dhiemas, R., Sakti, A.M., Ivan, F.M., Fukuda, T., Sekiyama, K.: Robots implementation for odor source localization using PSO algorithm. WSEAS Trans. Circ. Syst. **10**(4), 115–125 (2011)
 9. Jatmiko, W., Nugraha, A., Effendi, R., Pambuko, W., Mardian, R., Sekiyama, K., Fukuda, T.: Localizing multiple odor sources in a dynamic environment based on modified niche particle swarm optimization with flow of wind. WSEAS Trans. Syst. **8**(11), 1187–1196 (2009)
10. Kazadi, S., Goodman, R., Tsikata, D., Green, D., Lin, H.: An autonomous water vapor plume tracking robot using passive resistive polymer sensors. Auton. Robot. **9**(2), 175–188 (2000)
11. Li, F., Meng, Q., Li, J., Zeng, M.: P-PSO algorithm based multi-robot odor source search in ventilated indoor environment with obstacles. Acta Automatica Sinica **35**(12), 1573–1579 (2009)
12. Lu, Q., Liu, S., Xie, X., Wang, J.: Decision making and finite-time motion control for a group of robots. IEEE Trans. Cybern. **43**(2), 738–750 (2013)
13. Marques, L., De Almeida, A.T.: Electronic nose-based odour source localization. In: Proceedings of 6th International Workshop on Advanced Motion Control, pp. 36–40. IEEE (2000)
14. Marques, L., Nunes, U., de Almeida, A.T.: Particle swarm-based olfactory guided search. Auton. Robot. **20**(3), 277–287 (2006)
15. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey wolf optimizer. Adv. Eng. Softw. **69**, 46–61 (2014)

16. Willis, M.A.: Biologically-inspired search algorithms for locating unseen odor sources. Technical Report, DTIC Document (2002)
17. Zhang, Y., Zhang, J., Hao, G., Zhang, W.: Localizing odor source with multi-robot based on hybrid particle swarm optimization. In: 2015 11th International Conference on Natural Computation (ICNC), pp. 902–906. IEEE (2015)
18. Zou, Y., Luo, D.: A modified ant colony algorithm used for multi-robot odor source localization. In: International Conference on Intelligent Computing, pp. 502–509. Springer (2008)

Facial Expression Recognition Using Distance Signature Feature

Asit Barman and Paramartha Dutta

Abstract Distance feature has great significance in recognizing facial expressions. Identifying accurate landmarks is a vital as well as challenging issue in the field of affective computing. Appearance model is used to detect the salient landmarks on human faces. These salient landmarks form a grid on the human face. Distances are determined from the one landmark point to another landmark point in grid and normalized. A novel concept of corresponding stability index is introduced which eventually is found to play important role to recognize the facial expressions. Statistical analysis such as range, moment, skewness, kurtosis, and entropy are calculated from normalized distance signature to supplement the feature set. This enhanced feature set is supplied into a Multilayer Perceptron (MLP) to arrive at different expression categories encompassing anger, sadness, fear, disgust, surprise, and happiness. We experimented our proposed system on Cohn-Kanade (CK+), JAFFE, MMI, and MUG databases to training and testing our experiment and establish its superiority performance over the other existing competitors.

Keywords Grid · Distance signature · Stability index · Feature extraction
Feature classification · MLP

1 Introduction

Facial expression analysis is still a rigorous research area in the field of affective computing. Affective Computing [1] is a system that can analyze the facial expression of different categories. Human emotions are varied in human face during verbal

---

A. Barman (✉)
Department of Computer Science & Engineering and Information Technology,
Siliguri Institute of Technology, Darjeeling 734009, India
e-mail: mtechitasit@gmail.com

P. Dutta
Department of Computer & System Sciences, Visva-Bharati University,
Santiniketan 731235, India
e-mail: paramartha.dutta@gmail.com

and nonverbal communication. As such, human face is shown to be most sensitive to emotions. There are many applications in facial expression recognition system such as educational tutoring system, driver safety, video conferencing and surveillance. Ekman and Frisen [2] identified the study of facial expression. They observed six basic facial expressions, viz., fear, surprise, disgust, anger, sadness, and happiness. They used geometric features from eyes, eyebrows, mouth, jaw, etc. These features were encoded into the Facial Action Coding System (FACS). Active Appearance Model (AAM) [3] consists of both shape and texture model. In [4], authors proposed a new framework for formulating the most well-known optimization problems in Active Appearance Models and most importantly for formulating efficient solutions. Artificial neural network is a most well-known popular network to classify the facial expressions. Geometry and Gabor wavelet based facial expression recognition using multilayer perceptron was studied in [5]. Ma and Khorasani [6] used two-dimensional DCT as a feature set over a face image and a constructive one-hidden-layer feedforward neural network was used to classify the expressions. In [7], considered geometric positions of fiducial points offer discriminant information for face action units. These features were used to classify the expressions. Happy [8] proposed salient patch base discriminative features of various expressions and a classifier was used to classify the expressions.

In this article, we introduce features such as distance signature and stability indices in addition to frequently used statistical features, viz., moment, skewness, kurtosis, and entropy. Human face components such as eyes, eyebrows, mouth, jaw, etc., play crucial role in differentiating from one expression to another. Appearance-based method is applied on the human face images to identify the landmarks. Salient landmarks are considered among the region of eyes, eyebrows, nose, and mouth. We consider three points on eyebrow, four points on eye, three points on nose, and four points on mouth region. These salient landmarks form a grid and we calculate the useful distances within the grid. These distances are normalized by distance signature. Subsequently, the stability index is determined from the normalized distance signature which played a vital role to recognize the different expressions. Statistical analysis such as moment, range, skewness, kurtosis, and entropy are calculated from the normalized distance signature. These features are fed as input to the multilayer perceptron for training (testing) and subsequent classification into six basic expressions.

The remaining article is organized as follows. Section 2 presents an overview of proposed methodology, landmark detection, and grid formation. In Sect. 3, we introduce and elaborate the features. Feature classification is provided in Sect. 4. Classification of expressions and result discussions are offered in Sects. 5. Finally, conclusions are drawn in Sect. 6.
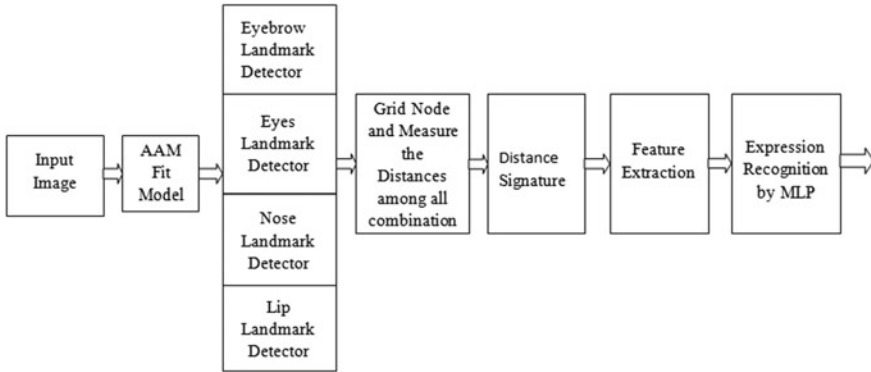
Fig. 1 Block diagram of proposed methodology

2 Proposed System, Landmark Detection and Grid Formation

The proposed methodology of distance signature based facial expression is shown in Fig. 1. Zhu and Ramanan [9] proposed landmark detection of the facial expressions on the face image. In [10], Y. Tie detected landmark points over the displacement and deformations of the facial face during dynamic variations of the expressions. Active Appearance Model (AAM) [4] is used to identify the proper landmarks of face images. Discriminative landmarks are considered in and around the eyes, eyebrows, nose, and mouth region. Extracted landmarks are formed a grid as presented in Fig. 2. These landmarks are connected to each other and stored in an array in a specific order consisting a grid as presented in Fig. 2. Normalized distances are calculated between every pair of landmark points in the grid. A order of such normalized distance measures in a specific order yields a distance signature. These normalized distance signatures are also used to compute the stability index. We also consider statistical analysis such as moment, skewness, entropy, and kurtosis to get prominence of enhanced feature set. These extracted features are used as input of multilayer perceptron to categorize the expressions.

3 Feature Extraction

Selection of appropriate facial features of human faces is a challenging issue in facial expression analysis. Salient landmarks are used to extract the discriminative features which play crucial role in facial expression recognition.

Fig. 2 Landmark detection and grid formation on CK+ and JAFFE databases

```
Grid=
[18 20 22 23 25 27 37 38 40 42 43;
 20 22 23 25 27 37 38 40 42 43 45;
 22 23 25 27 37 38 40 42 43 45 46;
 23 25 27 37 38 40 42 43 45 46 48;
 25 27 37 38 40 42 43 45 46 48 32;
 27 37 38 40 42 43 45 46 48 32 34;
 32 18 20 22 23 25 27 37 38 34 36;
 34 18 20 22 23 25 27 36 37 38 40;
 36 18 20 22 23 25 27 37 38 40 42;
 37 38 40 42 43 45 46 48 49 52 55;
 38 40 42 43 45 46 48 49 52 55 58;
 40 42 43 45 46 48 49 52 55 58 32;
 42 43 45 46 48 49 52 55 58 32 36;
 43 45 46 48 49 52 55 58 32 34 36;
 45 46 48 49 52 55 58 32 34 36 18;
 46 48 49 52 55 58 32 34 36 18 20;
 48 49 52 55 58 32 34 36 18 20 22;
 49 52 55 58 32 34 36 18 20 22 23;
 52 55 58 32 34 36 18 20 22 23 25;
 55 58 32 34 36 18 20 22 23 25 27;
 58 32 34 36 18 20 22 23 25 27 37]
```

3.1 Distance Signature and Stability Index

Detected discriminative landmarks on human faces are induced a grid which is shown in Fig. 2. Distances are computed among pair of landmarks in the grid and distance signature is obtained using Eq. 1. In this paper, m represents the total number of facial images and n indicates the total number of distances. Set of ρ_i^f values in appropriate sequence constitutes the distance signature and various d_i^f are the distances.

Stability index is yet another formidable factor used to recognize the facial expressions due to its sensitivity to six basic expressions. The normalized distance signature is used to compute the higher order signatures using Eq. 2. Here, the rth order distance signature is calculated for each face image. In this experiment, we set the rth order as $r = 1, 2, \dots, 8$. We consider the rth order differential using Eq. 3, where we set a threshold upon the analysis of the differences of higher order factors. This threshold γ is observed as for which value of r satisfies the condition $\Delta s_i^r \leq \gamma$. Such an order r is treated as stability index.

$$\rho_i^f = \frac{d_i^f}{\sum_{i=1}^n d_i^f} \quad f = 1, 2, \dots, m \text{ and } i = 1, 2, \dots, n \tag{1}$$

$$\rho^r = \frac{s_i^r}{\sum s_i^r} \qquad (2)$$

$$\Delta s_i^r = s_i^r - s_i^{r-1} \qquad (3)$$

3.2 Range, Moment, Skewness, Kurtosis, and Entropy

The statistical analyses such as raw moment, skewness, range, kurtosis, and entropy are considered to meaningfully augment the feature set. Range is computed from the difference between maximum and minimum value of the normalized shape signature. The fourth-order moment [11] is identified from normalized distance signature as $l_k = \sum_{i=1}^{n} b(i,k) \times (v(i))$. Here, $v = (v_1, v_2, \ldots, v_n)$ is the normalized distance signature and $b(i,k)$ is i raised to power of k, where $k = 1, 2, 3, 4$. Skewness and kurtosis [11] are computed as $U = \frac{\sum_{i=1}^{n} b(i,3)^2 \times v(i)}{\sum_{i=1}^{n} b(i,2)^3 \times v(i)}$ and $V = \frac{\sum_{i=1}^{n} b(i,4)}{\sum_{i=1}^{n} b(i,2)^2 \times v(i)}$. Here, U and V indicate the skewness and kurtosis, respectively. Entropy [12] is determined using equation $e = -\sum_{i=1}^{n} p_i \log(p_i)$. Where p_i contains the histogram count of the distance signature and $\sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1$.

4 Classification

A multilayer perceptron is a feedforward artificial neural network model that maps set of input data onto set of appropriate outputs [5]. The classification task of extracted features, every face image is identified by vector of nine elements. These features are used as Min-Max process matrices by normalizing the minimum and maximum values of each row feature vector. Extracted features are used as input to multilayer perceptron network and get output responses as per minimum errors. A perceptron consists of a number of nine input nodes, number of eight hidden/intermediate nodes, and number of six output layers. The weights usually excite or inhibit the signal that flows through the edge. The perceptron is found capable of classifying into six basic expressions, viz., anger (A), fear (F), disgust (D), sadness (S), happiness (H), and surprise (G).

5 Experiment and Result

Distance signature based facial expression recognition is validated on Cohn-Kanade (CK+) [13], JAFFE [14], MMI [15], and MUG [16] databases to justify the different facial expressions. Obtained results are compared with different state-of-the-art methods to evaluate the effectiveness of the proposed method.

5.1 Experiment on CK+ and JAFFE Databases

The Cohn-Kanade (CK+) database is a combination of 100 university students aged between 18 and 30 years. To evaluate the system performance, a total of 849 images were selected (566 for training and 283 for testing) and get average rate 98.6% for training. Table 1 shows the performance of different expressions such as A is classified in 49 images properly but 2 images are misclassified with G, D is classified appropriately in 45 images but 2 images are misclassified with A and H. The average performance of the distance signature is 96.3% as evident from Table 5.

The JAFFE [14] database has 213 gray images of seven facial expressions (six basic plus one neutral) of 10 Japanese females. In our discussion, a total of 213 images were considered (116 for training and 97 for testing) and acquires 96.6% for training. In Table 2, the confusion matrix shows the performance of different expressions. From this table, we observe that A, D, S, and G are classified correctly but F is misclassified with A and D. The average recognition rate achieved by the distance signature is 94% as illustrated in Table 5.

Table 1 The confusion matrix on CK+ database using distance signature

	A	D	F	H	S	G
A	49	0	0	0	0	2
D	1	45	0	1	0	0
F	0	0	23	1	0	0
H	1	0	0	63	0	0
S	0	0	0	0	47	2
G	0	1	0	1	0	46

Table 2 The confusion matrix on JAFFE database using distance signature

	A	D	F	H	S	G
A	21	0	0	0	0	0
D	0	20	0	0	0	0
F	2	1	11	0	0	0
H	0	2	0	12	0	0
S	0	0	0	0	14	0
G	0	0	0	0	0	14

Table 3 The confusion matrix on MMI database using distance signature

	A	D	F	H	S	G
A	24	3	2	0	0	0
D	0	17	0	0	3	1
F	2	1	10	0	0	1
H	0	2	0	14	0	0
S	0	0	0	0	18	0
G	1	1	1	0	2	10

Table 4 The confusion matrix on MUG database using distance signature

	A	D	F	H	S	G
A	29	1	0	0	1	0
D	0	30	0	0	0	0
F	0	0	26	0	0	0
H	0	2	0	43	0	0
S	0	0	0	1	28	0
G	1	0	0	0	1	34

5.2 Experiment on MMI and MUG Databases

A total of 325 images were selected (212 for training and 113 for testing) to measure the performance of expressions from MMI database and an average recognition rate of 92.5% could be achieved for training. The proposed procedure uses total 113 images to test the performance of expressions. MMI database offers more challenging issue for facial expression recognition compared to CK+. First, different people give pose of the same expression in various ways. Second, some expressions are found having accessories, such as glasses, headcloth, or moustache. Table 3 shows the promising classification for A, D, H, and S expressions except the F and G. The expressions F and G are misclassified worstly with A, D, S, and G. The average performance of the distance signature reported is 81.5%.

MUG [16] database is also considered to test the system performance of facial expression. Here, the same parameters are used as in CK+ database. A total of 591 images are selected (394 for training and 197 for testing). In Table 4, the confusion matrix reflects the classification of different expressions of distance signature. It is observed that the D and F are correctly recognized but A is misclassified with D and S. The average performance achieved using distance signature is 96.7%.

Table 5 Testing recognition rate on publicly available databases of six basic expressions with different state-of-the-art methods

	Database	A	D	F	H	S	G	Avg.
Distance signature	CK+	96.1	95.7	95.8	98.4	95.9	95.8	96.3
Happy [8]	CK+	87.8	93.3	94.3	94.2	96.4	98.4	94.1
Poursaberi et al. [7]	CK+	87	91.5	90.9	96.9	84.5	91.2	90.3
Zhong et al. [17]	CK+	76.2	94.1	86.1	96.3	88.2	98.7	91.5
Distance signature	JAFFE	100	100	78.6	85.7	100	100	94
Happy [8]	JAFFE	100	86.2	93.7	96.7	77.4	96.6	91.7
Poursaberi el al. [7]	JAFFE	89.3	90.7	91.1	92.6	90.2	92.3	91.1
Distance Signature	MMI	82.8	81	71.4	87.5	100	66.7	81.5
Poursaberi et al. [7]	MMI	80.1	78.2	81.3	83.2	77.1	81	80.1
Zhong et al. [17]	MMI	65.6	72.5	72.5	88.2	71.1	93.8	77.4
Distance signature	MUG	93.5	100	100	95.6	96.6	94.4	96.7

5.3 Compare Analysis with Different State of the Arts

The proposed distance signature method is compared with the results obtained from existing arts in literatures [7, 8, 17, 18]. A comparison task is presented with the other methods. We evaluate our experiments on CK+, JAFFE, MMI, and MUG databases. Table 5 summarizes the comparative results with other methods. It is also noticed that our proposed procedures acquired higher average recognition rate of facial expressions than others reported in [7, 8, 17].

6 Conclusion

This paper proposed a distance signature based facial expression recognition. The performance of the proposed procedure is validated by the recognition rate and comparison with the state-of-the-art reporting. The experimental results also indicate the significant performance improvements of facial expression recognition using distance signature.

References

1. Picard, R.W., Picard, R.: Affective Computing, 252. MIT press, Cambridge (1997)
2. Ekman, P., Frisen, W.V.: Emotion in the Human Face. Prentice Hall, Eagle Woods Cliffs, NJ (1975)
3. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 681–685 (2001)
4. Tzimiropoulos, G., Pantic, M.: Optimization problems for fast aam fitting in-the-wild. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 593–600 (2013)
5. Zhang, Z., Lyons, M., Schuster, M., Akamatsu, S.: Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron. In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998, pp. 454–459. IEEE (1998)
6. Ma, L., Khorasani, K.: Facial expression recognition using constructive feedforward neural networks. Syst. Man Cybern. Part B Cybern. IEEE Trans. **34**(3), 1588–1595 (2004)
7. Poursaberi, A., Noubari, H.A., Gavrilova, M., Yanushkevich, S.N.: Gauss–laguerre wavelet textural feature fusion with geometrical information for facial expression identification. EURASIP J. Image Video Process. **2012**(1), 1–13 (2012)
8. Happy, S.L.: Automatic facial expression recognition using features of salient facial patches. IEEE Trans. Affect. Comput. **6**(1), 1–12 (2015)
9. Zhu, X., Ramanan, D.: Face detection, pose estimation, and landmark localization in the wild. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2879–2886. IEEE (2012)
10. Tie, Y., Guan, L.: Automatic landmark point detection and tracking for human facial expressions. EURASIP J. Image Video Process. **2013**(1), 8 (2013)
11. Young, T.Y.: Handbook of Pattern Recognition and Image Processing (vol. 2): Computer Vision. Academic Press, Inc. (1994)
12. Borda, M.: Fundamentals in Information Theory and Coding. Springer (2011), Technical report. ISBN 978-3-642-20346-6
13. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The extended cohn-kanade dataset (ck+): a complete dataset for action unit and emotion-specified expression. In: Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, pp. 94–101. IEEE (2010)
14. Lyons, M., Akamatsu, S., Kamachi, M., Gyoba, J.: Coding facial expressions with gabor wavelets. In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998, pp. 200–205. IEEE (1998)
15. Valstar, M.F., Pantic, M.: Induced disgust, happiness and surprise: an addition to the mmi facial expression database. In: Proceedings of International Conference on Language Resources and Evaluation, Workshop on EMOTION, pp. 65–70, Malta, May 2010
16. Aifanti, N., Papachristou, C., Delopoulos, A.: The mug facial expression database. In: Proceedings of the 11th International Workshop on Image Analysis for Facial Expression Database, pp. 12–14, Desenzano, Italy, April 2010
17. Zhong, L., Liu, Q., Yang, P., Huang, J., Metaxas, D.N.: Learning multiscale active facial patches for expression analysis. IEEE Trans. Cybern. **45**(8), 1499–1510 (2015)
18. Zhong, L., Liu, Q., Yang, P., Liu, B., Huang, J., Metaxas, D.N.: Learning active facial patches for expression analysis. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2562–2569. IEEE (2012)

A Parallel Interval Type-2 Fuzzy Neural Inference System Using Different Similarity Measures: Comparative Study

Vuppuluri Sumati and C. Patvardhan

Abstract This paper presents the comparison between performance of an Interval Type-2 Subsethood Neural Fuzzy Inference System (IT2SuNFIS) (Sumati et al. Expert Syst. Appl. (Elsevier) 60:156–168, 2016 [27]) using two different similarity measures, implemented on a parallel platform. The inputs to both the system are fuzzified using interval type-2 fuzzy sets (IT2 FS) with Gaussian primary membership function (GPMF) having identical mean but different variance. The signal aggregation of type-2 based activation is performed using product operator. A comparison between subsethood and mutual subsethood has been carried out experimentally, showing better contrast handling capacity of later over former. Consequently, the performance of both the models is tested and compared on a benchmark application of Mackey-Glass time series prediction. It is observed that the performance of mutual subsethood based IT2SuNFIS is better than subsethood-based IT2SuNFIS in terms of result accuracy.

Keywords Interval type-2 fuzzy sets · Neuro-fuzzy systems · Subsethood Mutual subsethood · Differential evolution

1 Introduction

The inferencing capabilities of Interval Type-2 Subsethood Neural Fuzzy Inference System (IT2SuNFIS) are explored on a variety of applications in the paper by Sumati et al. [27]. In IT2SuNFIS, differential evolution (DE) and Artificial Bee Colony–Differential Evolution (ABC-DE) [2] have been used as the basic learning strategies.

V. Sumati (✉)
Department of Physics and Computer Science, Dayalbagh Educational Institute, Agra, India
e-mail: vsumati.29@gmail.com

C. Patvardhan
Department of Electrical Engineering, Dayalbagh Educational Institute, Agra, India
e-mail: cpatvardhan@gmail.com

The signal transmission from the inputs to the rule nodes in IT2SuNFIS takes place through subsethood signal calculated between the input IT2FSs and the antecedent IT2FSs. The subsethood measure for type-n fuzzy sets is derived by Rickard et al. [24] and is used in IT2SuNFIS. On the other hand, another subsethood measure, called *mutual subsethood* measure, has been extensively used in [23, 26, 30] for the case of standard fuzzy sets and it remains undefined in the field of type-2 fuzzy systems.

Type-2 fuzzy sets (T2FS) were introduced by Zadeh [31] and later used by Mendel [12, 18–21] extensively. T2FSs efficiently handle higher order of uncertainty in the data as compared to their standard counterpart, often referred to as type-1 fuzzy sets (T1FS). Type-2 fuzzy logic based systems are being successfully applied to many areas [7] including control [1, 6], time series prediction [3, 5], signal processing [16], healthcare services [4, 9, 14, 15, 22], and recognition [8, 11, 32].

This paper is organized in six sections. Mutual subsethood IT2SuNFIS (MS-IT2SuNFIS) model has been described in Sect. 2. Section 3 compares the performance of subsethood and mutual subsethood methods experimentally. Section 4 describes the performance of MS-IT2SuNFIS on Mackey-Glass time series prediction problem. Section 5 concludes the paper, followed by acknowledgements.

2 IT2SuNFIS and MS-IT2SuNFIS Models

IT2SuNFIS and MS-IT2SuNFIS both are three layered architectures, as shown in Fig. 1. In both the models, the inputs to the network are taken to be IT2 FSs. The antecedent weights to the system are IT2 FSs. The architectural details of the IT2SuNFIS network are explained in the paper by Sumati et al. [27]. MS-IT2SuNFIS is constructed by replacing subsethood method with mutual subsethood method in IT2SuNFIS.

According to the definition of subsethood given by Rickard et al. [24] given in Eq. (1), the subsethood of a set A in B is not equal to the subsethood of B in A. Denominator term in Eq. (1) is the area of membership function of the set whose subsethood value in other set is being determined. Therefore, if the area of set A is bigger than the area of set B, subsethood of A in B is lesser than the subsethood of B in A (refer to Eq. (1)). This makes subsethood measure, a definition, and direction, dependent quantity.

$$S(\tilde{A}, \tilde{B}) = sub(A, B) = \frac{\sum_{k=1}^{N} min(\mu_A(x_k), \mu_B(x_k))}{\sum_{k=1}^{N} \mu_A(x_k)} \tag{1}$$

The mutual subsethood $\xi(F, G)$ between fuzzy sets F and G is then given by

$$\xi(F, G) = \frac{C(F \cap G)}{C(F) + C(G) - C(F \cap G)} \tag{2}$$
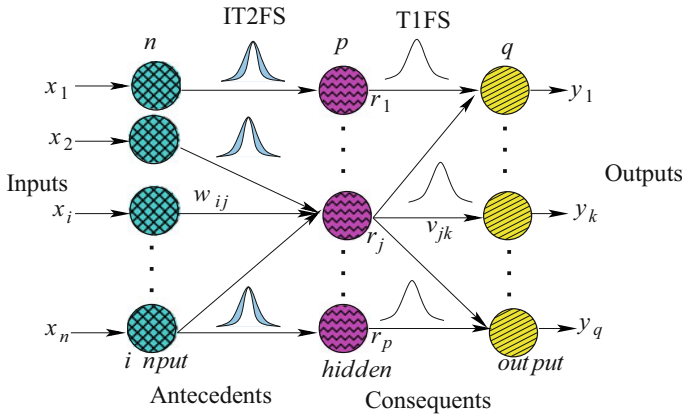
Fig. 1 Architecture of IT2SuNFIS and MS-IT2SuNFIS models

From the definition, mutual subsethood measure will not be 1 even if a particular set is contained within another; it is 1 only if the two sets are identical. In case of IT2FSs, the mutual subsethood is an interval set $[\xi_l, \xi_r]$, where ξ_l and ξ_r represent the leftmost and rightmost mutual subsethood values; it is derived in [28].

3 Subsethood Verses Mutual Subsethood Based Transmission

Type-2 subsethood method (used in IT2SuNFIS) and type-2 mutual subsethood method (used in MS-IT2SuNFIS) appear to have similar names, however, the type-2 mutual subsethood composition has better discriminatory power than the type-2 subsethood composition.

Type-2 mutual subsethood and type-2 subsethood transmission methods are compared for numeric inputs on the basis of the nature of the signals transmitted by composing the inputs and the antecedents. The signals generated by the composition mechanisms are transmitted through a connection weight from an input node to a rule node. The numeric input x_i^c values vary in the range (0, 1) while the antecedent IT2FS weights are kept fixed at particular values of centers w_{ij}^c and lower spreads $w_{ij}^{\sigma^l}$ and upper spreads $w_{ij}^{\sigma^u}$. To facilitate the comparison in experiments, the upper subsethood values of both the type-2 composition transmission methods are plotted.

The IT2FSs are chosen in such a way that their centers, lower spreads, and upper spread values are in set {0.1, 0.5, 0.9}, {0.1, 0.5, 0.9} and {0.12, 0.52, 0.92}, respectively. These values are so chosen as to cover the extreme and the mid-range values in (0, 1). The combination of centers, lower, and upper spread values is shown in Table 1.

Table 1 Combinations of centers and spreads of IT2FSs for the composition mechanism experiment

Exp.	1	2	3	4	5	6	7	8	9
w_{ij}^c	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$w_{ij}^{\sigma^l}$	0.1	0.1	0.1	0.5	0.5	0.5	0.9	0.9	0.9
$w_{ij}^{\sigma^u}$	0.12	0.12	0.12	0.52	0.52	0.52	0.92	0.92	0.92
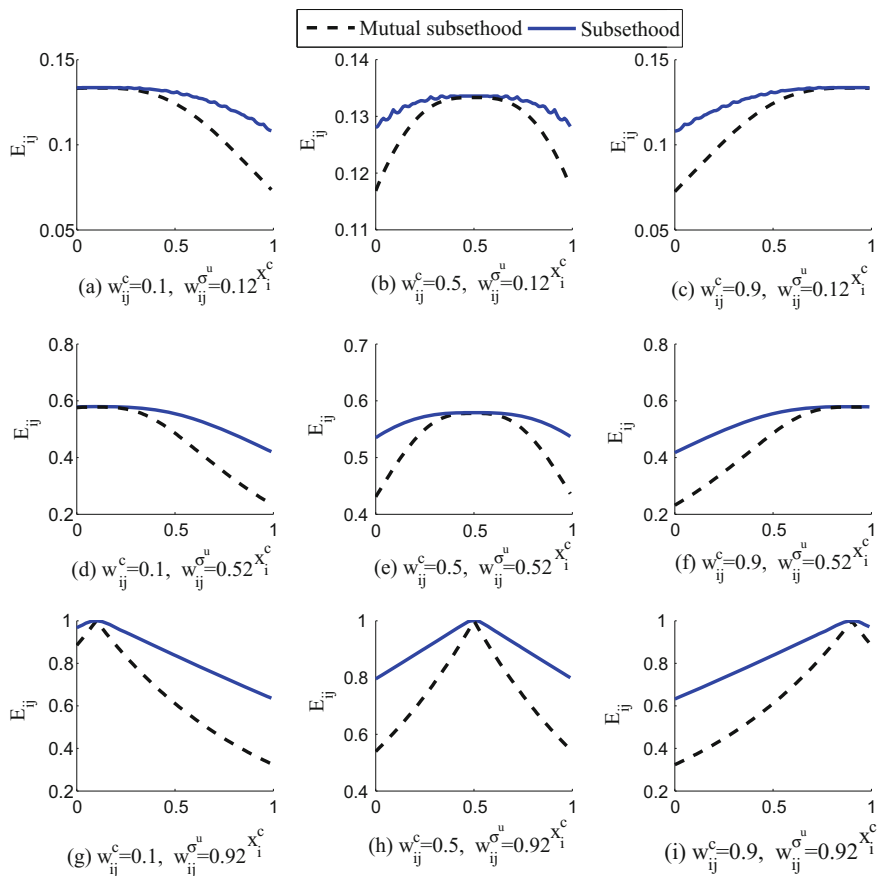
Fig. 2 Signal plots obtained for subsethood and mutual subsethood compositions when the numeric input is fuzzified by a spread of value 0.9

The signal magnitudes generated by both type-2 subsethood and type-2 mutual subsethood based signal transmissions are plotted in Fig. 2 for the input spread values [0.9, 0.92].

It is observed that for all input spreads and weight spreads combinations, the signal magnitudes are high at the points where the centers of both inputs and weights match each other exactly, and also, where input IT2FSs and antecedent IT2FSs are best matched. The major behavioral difference, however, appears in the regions of mismatched input and weight centers. It is observed that over a wide range of spreads (0.1–0.9), between dissimilar input IT2FSs and weight IT2FSs, a large contrast is created by mutual subsethood composition. Mutual subsethood nicely maintains a large difference between minimum and maximum values. It is evident from the sharp decline in the signal magnitudes in the regions of most dissimilar inputs.

Motivated by the better performance of type-2 mutual subsethood over type-2 subsethood method, IT2SuNFIS is integrated with mutual subsethood method in place of subsethood method. The resulting MS-IT2SuNFIS model is investigated on benchmark application Mackey-Glass time series prediction to verify its performance.

In the present study, DE is used to implement learning in both IT2SuNFIS and MS-IT2SuNFIS; DE is based on the evaluation of many population members with total size at least ten times the number of elements in each string. The function evaluation part is the most computation-intensive operation in the present code. This function evaluation part for each population string is sent to each of the available nodes for computation. This process repeats till all the strings in the population are evaluated. Thus, both models are implemented on a parallel platform using Local Area Multi-computer/ Messages Passing Interface (LAM/MPI) library.

4 Mackey-Glass Time Series Prediction

A problem given by Mackey and Glass [17] is considered in this work. This is a widely investigated problem in neural fuzzy and type-2 fuzzy domains [10, 13, 23, 25, 29, 30]. This chaotic Mackey-Glass time series (MGTS) is described by the following delay differential equation:

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \tag{3}$$

As τ in Eq. 3 varies, the system can exhibit either a limit cycle, fixed point, or chaotic behavior. Chaotic behavior is observed For $\tau = 17$ and for this value of τ, the above problem of predicting the time series is attempted.

A $4 - p - 1$ network architecture is used for this prediction problem where input layer consists of four numeric nodes; the output layer consists of a single output node, with p-rule nodes in the hidden layer. Experiments are performed with different rule

Table 2 Performance comparison of different rule MS-IT2SuNFIS networks with IT2SuNFIS

Rules	Network configuration	Free parameters	IT2SuNFIS test NRMSE	MS-IT2SuNFIS NRMSE
3	4-3-1	34	0.018624	0.012838
4	4-4-1	44	0.017101	0.015108
5	4-5-1	54	0.018348	0.013902
6	4-6-1	64	0.015201	0.012667
6	4-7-1	74	0.016413	0.010967

nodes. All the lower spreads are initialized randomly in the range $(0.01, 0.9)$ and centers in the range $(0, 1.5)$. Upper spread values of the antecedent type-2 fuzzy sets are taken to be 0.04 more than their respective lower spread values.

Experiments are carried out with different values of DE parameters and most suitable values found are $F = 0.75$ and $CR = 0.39$. Termination criterion for learning is maximum number of generations. SSE is considered to be the fitness function and normalized root mean square error (NRMSE) is taken as the performance measure. NRMSE is the ratio of RMSE and the standard deviation of the target series [13].

4.1 Simulation Results

The performance of different rule MS-IT2SuNFIS networks is compared with corresponding IT2SuNFIS networks and NRMSE over test samples for both networks is shown in the Table 2. It is seen that MS-IT2SuNFIS outperforms IT2SuNFIS in all the different rule networks in terms of NRMSE over test samples. This experiment also indicates better performance of mutual subsethood based signal transmission over subsethood based transmission.

For a 7-rule IT2SuNFIS and a 7-rule MS-IT2SuNFIS, function approximation capabilities are depicted in Fig. 3. Close observation yields that prediction error in case of mutual MS-IT2SuNFIS is less as compared to IT2SuNFIS for 7-rules network using ABC-DE leaning strategy and same initial conditions. It is also observed that MS-IT2SuNFIS approximated the Mackey-Glass time series better than IT2SuNFIS in 7-rules network. The red dashed line represents predicted series obtained for IT2SuNFIS network and blue solid line represents MS-IT2SuNFIS predicted series obtained for MS-IT2SuNFIS. Black dotted line shows the actual function.
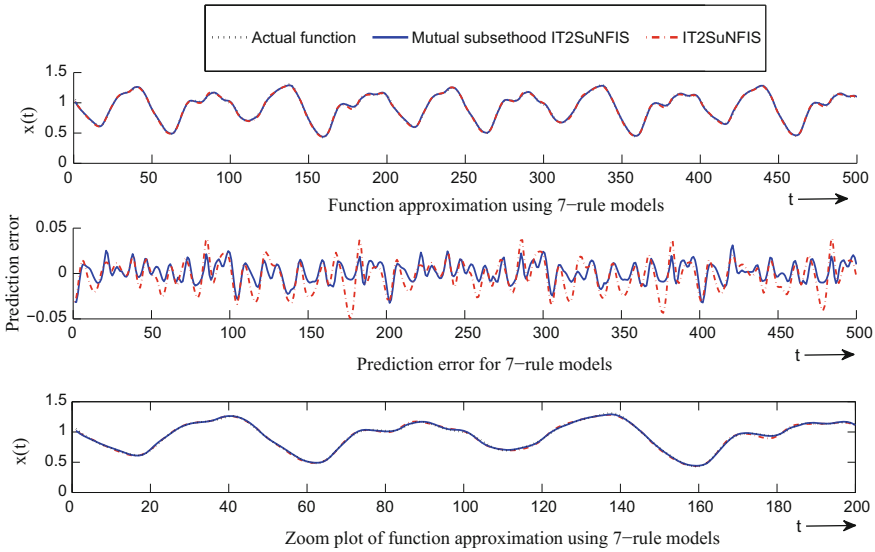
Fig. 3 Comparison between 7-rules IT2SuNFIS and 7-rules MS-IT2SuNFIS

5 Conclusions

This paper compares the performance of subsethood and mutual subsethood based signal transmission methods. It is observed that mutual subsethood based signal distinguishes between two non-correlated values better than subsethood based signal transmission. It is observed that application of mutual subsethood based signal transmission method improves the performance of IT2SuNFIS many folds in Mackey-Glass time series prediction. This experiment provides motivation for further exploration of mutual subsethood based neuro-fuzzy models. Such models are useful in biomedical data where accurate diagnosis and treatment are dependent on the correct classification of the disease type. Better uncertainty handling capacity of T2FSs is useful in problems with partially available data and unseen data.

References

1. Abiyev, R.H., Kaynak, O.: Type 2 fuzzy neural structure for identification and control of time-varying plants. IEEE Trans. Ind. Electron. **57**(12), 4147–4159 (2010)
2. Abraham, A., Jatoth, R.K., Rajasekhar, A.: Hybrid differential artificial bee colony algorithm. J. Comput. Theor. Nanosci. **9**(2), 1–9 (2012)

3. Castillo, O., Castro, J.R., Melin, P., Rodriguez-Diaz, A.: Application of interval type-2 fuzzy neural networks in non-linear identification and time series prediction. Soft Comput. **18**, 1213–1224 (2014)
4. Chumklin, S., Auephanwiriyakul, S., Theera-Umpon, N.: Microcalcification detection in mammograms using interval type-2 fuzzy logic system with automatic membership function generation. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–7. IEEE, Barcelona (2010)
5. Gaxiola, F., Melin, P., Valdez, F., Castillo, O.: Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction. Inform. Sci. **260**, 1–14 (2014)
6. Hagras, H.: A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots. IEEE Trans. Fuzzy Syst. **12**(4), 524–539 (2004)
7. Hagras, H., Wagner, C.: Towards the wide spread use of type-2 fuzzy logic systems in real world applications. IEEE Comput. Intell. Mag. **7**(3), 14–24 (2012)
8. Halder, A., Konar, A., Mandal, R., Chakraborty, A., Bhowmik, P., Pal, N.R., Nagar, A.K.: General and interval type-2 fuzzy face-space approach to emotion recognition. IEEE Trans. Syst. Man Cybern. Syst. **43**(3), 587–605 (2013)
9. John, R.I., Innocent, P.R., Barnes, M.R.: Neuro-fuzzy clustering of radiographic tibia image data using type-2 fuzzy sets. Inform. Sci. **125**, 203–220 (2000)
10. Juang, C.F., Tsao, Y.W.: A self-evolving interval type-2 fuzzy neural network with online structure and parameter learning. IEEE Trans. Fuzzy Syst. **16**(6), 1411–1424 (2008)
11. Juang, C.F., Wang, P.H.: An interval type-2 neural fuzzy classifier learned through soft margin minimization and its human posture classification application. IEEE Trans. Fuzzy Syst. **23**(5), 1474–1487 (2015)
12. Karnik, N.N., Mendel, J.M., Liang, Q.: Type-2 fuzzy logic systems. IEEE Trans. Fuzzy Syst. **7**(6), 643–658 (1999)
13. Kim, J., Kasabov, N.: HyFIS: Adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems. Neural Netw. **12**(9), 1301–1321 (1999)
14. Lee, C.S., Wang, M.H., Hagras, H.: A type-2 fuzzy ontology and its application to personal diabetic-diet recommendation. IEEE Trans. Fuzzy Syst. **18**(2), 374–395 (2010)
15. Lee, C.S., Wang, M.H., Lan, S.T.: Adaptive personalized diet linguistic recommendation mechanism based on type-2 fuzzy sets and genetic fuzzy markup language. IEEE Trans. Fuzzy Syst. **23**(5), 1777–1802 (2015)
16. Liang, Q., Mendel, J.M.: Equalization of nonlinear time-varying channels using type-2 fuzzy adaptive filters. IEEE Trans. Fuzzy Syst. **8**(5), 551–563 (2000)
17. Mackey, M., Glass, L.: Oscillation and chaos in physiological control systems. Science **197**, 287–289 (1977)
18. Mendel, J.: Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions. Prentice-Hall, Upper Saddle River, NJ (2001)
19. Mendel, J.: Type-2 fuzzy sets and systems: an overview. IEEE Comput. Intell. Mag. **2**(1), 20–29 (2007)
20. Mendel, J., Hagras, H., Tan, W.W., Melek, W.W., Ying, H.: Introduction to Type-2 Fuzzy Logic Control: Theory and Applications. Wiley-IEEE Press, Piscataway, NJ (2014)
21. Mendel, J.M., John, R.I.: Type-2 fuzzy sets made simple. IEEE Trans. Fuzzy Syst. **10**(2), 117–127 (2002)
22. Özbay, Y., Ceylan, R., Karlik, B.: Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier. Expert Syst. Appl. **38**(1), 1004–1010 (2011)
23. Paul, S., Kumar, S.: Subsethood-product fuzzy neural inference system (SuPFuNIS). IEEE Trans. Neural Netw. **13**(3), 578–599 (2002)
24. Rickard, J.T., Aisbett, J., Gibbon, G.: Fuzzy subsethood for fuzzy sets of type-2 and generalized type-n. IEEE Trans. Fuzzy Syst. **17**(1), 50–60 (2009)
25. Russo, M.: FuGeNeSys—A fuzzy genetic neural system for fuzzy modeling. IEEE Trans. Fuzzy Syst. **6**(3), 378–388 (1998)

26. Singh, L., Kumar, S., Paul, S.: Automatic simultaneous architecture and parameter search in fuzzy neural network learning using novel variable length crossover differential evolution. In: IEEE World Congress on Computational Intelligence, pp. 1795–1802. Hong Kong (2008)
27. Sumati, V., Chellapilla, P., Paul, S., Singh, L.: Parallel interval type-2 subsethood neural fuzzy inference system. Expert Syst. Appl. (Elsevier) **60**, 156–168 (2016)
28. Sumati, V., Patvardhan, C.: Interval type-2 mutual subsethood fuzzy neural inference system (IT2MSFuNIS). IEEE Trans. Fuzzy Syst. **26**(1), 203–215 (Feb 2018). https://doi.org/10.1109/TFUZZ.2016.2646750
29. Tung, S.W., Quek, C., Guan, C.: eT2FIS: an evolving type-2 neural fuzzy inference system. Inform. Sci. **220**, 124–148 (2013)
30. Velayutham, C.S., Kumar, S.: Asymmetric subsethood product fuzzy neural network (ASuP-FuNIS). IEEE Trans. Neural Netw. **16**(1), 160–174 (2005)
31. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning, part I. Inform. Sci. **8**, 199–249 (1975)
32. Zeng, J., Liu, Z.Q.: Type-2 fuzzy hidden markov models and their application to speech recognition. IEEE Trans. Fuzzy Syst. **14**(3), 454–467 (2006)

A Fuzzy Logic Inspired Approach for Social Media Sentiment Analysis via Deep Neural Network

Anit Chakraborty, Anup Kolya and Sayandip Dutta

Abstract In this paper, we present an efficient method of classification of senti-ment in social media texts, each consisting of single or multiple sentence(s) that most of the time includes pop culture texts. In our experiment, we present an architecture that derives vector representations (i.e., word2vec) of the phrase level sentences. We use some combination of quantitative and qualitative methods for training a recurrent neural network with empirically cross-validating gold-standard array of lexical features, which are precisely synced with sentiment in microblog-like pieces. We leverage a new technique that expands upon previous works on sentence-level lexical sentiment classification, using recurrent fuzzy neural network and use it jointly with a Recursive Neural Network to further improve the classification. We have tested our algorithm against the other state-of-the-art methods on various platforms for better demonstration of our experiment with satisfactory and competitive results.

Keywords Deep learning · Fuzzy logic · Machine learning
Recurrent neural network · Sentiment analysis · Social media analysis
Word2vec

A. Chakraborty (✉) · A. Kolya
RCC Institute of Information Technology, Kolkata, India
e-mail: ianitchakraborty@gmail.com

A. Kolya
e-mail: anup.kolya@gmail.com

S. Dutta
MCKV Institute of Engineering, Howrah, India
e-mail: sayandip199309@gmail.com

1 Introduction

Due to the rapid growth of the Internet, online communication has escalated exponentially and presented newer platforms for the general public to make their feelings and opinions known to the world online. This, in turns, has led to accumulation of enormous amounts of online data, which is rich in user sentiments, emotions, opinions, and evaluations. A broad range of high quality research for fast and accurate sentiment classification has been explored in recent times [1, 2]. Linguistic Inquiry and Word Count or LIWC is an example of such a lexicon [3, 4]. It has been widely used in the social media domain, as its straightforward dictionary and simple word lists are easily inspected, understood, and extended if desired.

Our approach seeks to construct a sentiment analysis engine that is capable in dealing with:

- Versatility and flexibility of artificial neural networks incorporating fuzzy logic system by means of unsupervised learning and planning to attain seamless integration of sentiment classification and approximation with other sensory modalities.
- Considerably faster to deal with live-streaming data and real-time social debates.
- Unaffected from a performance-speed tradeoff while dealing with long corpus of texts.
- Nuanced classification of sentiments over the other extant binary classification (i.e., Positive and Negative) with the help of RFNN.

The organization of the paper constitutes: Review of related works and other state-of-the-art methods in Sect. 2. Section 3 explains our proposed method of fuzzy logic inspired approach for social media sentiment analysis via deep neural network. Experimental results and analysis are presented in Sect. 4. Section 5 concludes the paper and discusses future possibilities for further improvements.

2 Related Works

Previously, linear classification methods, e.g., Supper Vector Machines (SVM) and logistic regression [5] have been used to solve sentiment classification problems. Maximum Entropy and Naïve Based Classifiers have also been studied in this field. In [5], Maas et al. tackle the problem of sentiment classification by learning word vectors, using an unsupervised model to capture nuanced sentimental information. Socher et al. [6] semantic word spaces are ill-equipped to parse meaning of longer phrases. They, instead, propose Recursive Neural Tensor Network for more rigorous classification. "Stanford Sentiment Treebank", also introduced in [6], is a dataset that consists of over 2015,154 phrases with fine-tuned sentiment values which spans over a parse tree containing 11,855 sentences. It shows significant improvement over bag-of-words models. As mentioned in [7], fuzzy logic inspired

artificial neural network has also been explored in the field of sentiment classification and social media analysis.

A large proportion of approaches for sentiment classification greatly rely on intermittent sentiment lexicons. There are some popular lexicons like LIWC1 [4], GI2 [8], Hu-Liu043 [9], etc., where words are divided into binary classes, and on the other hand, lexicons like ANEW4 [10], SentiWordNet5 [11], and SenticNet6 [12] associate words with valence-based scores for sentence-level sentiment intensity.

3   Proposed Method

In this section, initially, we propose the text modeling algorithm to extract efficient and compact vector representation (i.e., Word2vec) according to the word position in the vocabulary. Subsequently, the three-layered deep neural sentiment classification architecture has been introduced via recurrent neural network. Finally, the fuzzy classification based on recurrent fuzzy neural network is introduced for further precise classification of sentiments beyond the classical binary sentiment categorization.

3.1   Text Modeling

In this architecture, each word is matched to a unique vector, which is denoted by a column of the matrix W. It is then indexed according to the word position in the vocabulary. For predicting the next word in a sentence, the sum of the vectors is used.

Formally, if $w_1, w_2, w_3, \ldots, w_T$ is the given sequence of words for training, the vector model tries to maximize the average log probability of w_t at kth word count,

$$\frac{1}{T} \sum_{t=k}^{T-k} \log p(w_t | w_{t-k}, \ldots, w_{t+k}). \tag{1}$$

The prediction is normally done using a multiclass classifier, (e.g., Softmax). We have

$$p(w_t | w_{t-k}, \ldots, w_{t+k}) = \frac{e^{y_{w_t}}}{\sum_i e^{y_i}}. \tag{2}$$

Each of p_{1i} is un-normalized log probability for each output word i, computed as

$$p_1 = b + Uh(w_{t-k}, \ldots, w_{t+k}; W), \tag{3}$$

where b and U are the softmax parameters. h is created by summing or averaging of word vectors obtained from W.

Practically, hierarchical softmax [13] is considered better than softmax for fast training. In our experiment, the structure of the hierarchical softmax is a binary Huffman tree [13].

3.2 Deep Neural Sentiment Classification

Computation of compositional vector representations for sentence-level phrases of variable syntactic type and length are formulated as mentioned earlier. The compositional models parse an n-gram into a binary tree and each leaf node, corresponding to a word, is represented as a vector. Each word is denoted as a d-dimensional vector by initializing all word vectors post random sampling of corresponding weightage from a uniform distribution: U $(-r, r)$, where $r = 0.0001$. In the word embedding matrix $L \epsilon \mathbb{R}^{d \times |V|}$, all the word vectors are arranged, where the size of the vocabulary is denoted by $|V|$.

A softmax classifier takes the optimized word2vectors as feature inputs. The posterior probability is computed over any given label of the word2vector in the following way:

$$h_a = softmax(W_s a), \tag{4}$$

where $W_s \epsilon \mathbb{R}^{5 \times d}$ is the sentiment classification matrix. For the given tri-gram, this is repeated for vectors b and c.

We define the output of a deep neural network $h \epsilon \mathbb{R}^d$ via $V^{[i]} \epsilon \mathbb{R}^{d \times d}$, representing detailed notation for each slice of the following vector equivalent, as:

$$h = \begin{bmatrix} b \\ c \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} b \\ c \end{bmatrix}; h_i = \begin{bmatrix} b \\ c \end{bmatrix}^T V^{[i]} \begin{bmatrix} b \\ c \end{bmatrix}, \tag{5}$$

where $V^{[1:d]} \epsilon \mathbb{R}^{2d \times 2d \times d}$ defines multiple bilinear forms. The RecNN uses this formulation for computing p_1:

$$p_1 = f\left(\begin{bmatrix} b \\ c \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} b \\ c \end{bmatrix} + W \begin{bmatrix} b \\ c \end{bmatrix} \right), \tag{6}$$

where W is defined earlier. p_2, the subsequent parent vector in the tri-gram is computed using the same weights:

$$p_2 = f\left(\begin{bmatrix} a \\ p_1 \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} a \\ p_1 \end{bmatrix} + W \begin{bmatrix} a \\ p_1 \end{bmatrix} \right). \tag{7}$$

3.3 Recurrent Fuzzy Neural Classification

In this section, after obtaining the sentiment scores of all training documents by the sentiment network (RecNN), we are introducing fuzzy logic inspired sentiment approximation.

For accurate sentiment orientation of phrase level sentences of the test data, we define $p_{2i} = |score(h_i)|$ and obtain the fuzzy logic train set (h_1, p_{11}, p_{21}), (h_2, p_{12}, p_{22}), …, (h_n, p_{1n}, p_{2n}). Furthermore, using $\{(h_1, p_{11}, p_{21}), (h_2, p_{12}, p_{22}), …, (h_n, p_{1n}, p_{2n})\}$ we train RecFNN f. Sentiment intensity of h_i of large value indicates robust sentiment of high fuzzy membership degree, making bigger contribution in accurate sentiment distribution.

The prioritization scheme of the RecFNN can be formalized as

$$\max W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j p_{1i} p_{1j} (h_i.h_j), \tag{8}$$

$$subject\ to\ \sum_{i=1}^{n} \alpha_i p_{1i} = 0, \tag{9}$$

$$0 \le \alpha_i \le p_{2i} C, \quad i = 1, 2, …, n. \tag{10}$$

According to Eqs. (8)–(10), we obtain the optimal solution, $\alpha^* = \left(\alpha_1^*, \ \alpha_2^*, …, \alpha_n^*\right)^T$.

In $\alpha^* = \left(\alpha_1^*, \ \alpha_2^*, …, \alpha_n^*\right)^T$, if $\alpha_i^* > 0$, the corresponding h_i is the support vector. If $0 \le \alpha_i^* \le p_{2i} C$, support vector of this type resides in the horizon of the hyper plane.

If $\alpha_i^* = p_{2i} C$, support vector of this type is considered as misclassified sample.

Notably, RecNN and RecFNN models are different, as points with the same α_i^* value may indicate different types of artificial neural network depending on the value of p_{2i}:

In this paper, the Gaussian kernel function $K\left(h_i, h_j\right) = \exp\left(\frac{\|h_i - h_j\|^2}{2\delta^2}\right)$ is used as the kernel for constructing the RecFNN classifier. The corresponding optimal solution is $\alpha^* = \left(\alpha_1^*, \ \alpha_2^*, …, \alpha_n^*\right)^T$. Therefore, the fuzzy optimal classification is

$$f^*(x) = sgn\left\{ \sum_{j=1}^{n} \alpha_j^* p_{1j} K\left(h_i, h_j\right) + b^* \right\}. \tag{11}$$

$$b^* = p_{1i} - \sum_{j=1}^{n} \alpha_j^* p_{1j} K\left(h_i, h_j\right), \quad j\epsilon\left\{j | 0 < \alpha_j^* < p_{2j} C\right\}. \tag{12}$$

To achieve the two-level sentiment classification, we make use of the positive sentiment reviews as the positive class models and negative sentiment reviews as

the negative class models to train the RecFNN. Finally, K classifiers $\{f_1, f_2, \ldots, f_K\}$ are obtained. Thus, each test model h_i consists of K results:

$$\{(f_1(h_i), \ \varepsilon(f_1(h_i))), \ldots, (f_K(h_i), \ \varepsilon(f_K(h_i)))\},$$

where the result of the Kth classifier is $f_K(h_i)$ and confidence of the Kth classifier is $\varepsilon(f_K(h_i))$. Finally, the corresponding label of $max\{\varepsilon(f_1(h_i)), \ \varepsilon(f_2(h_i)), \ldots, \varepsilon(f_K(h_i))\}$ is selected as the label of h_i.

4 Experimental Results

4.1 Experimental Setup

Proposed algorithm is implemented on python 3.6.1 with NLTK package, with an Intel Core 5th Gen i7, 2.5 GHz processor with 6 Gigabytes of RAM.

For fair comparison of our experiment against the broader body of literature, we assess the correlation of calculated sentiment intensity rating to the mean sentiment rating from ten prescreened human raters, along with the 6-class (i.e., very positive, positive, neutral, negative, very negative, compound) classification accuracy parameters of precision, recall, and F1 score. In our analysis, *precision* is defined as the ration between the numbers of true classification, to the total number of elements labeled as belonging to that class (both correct and incorrect labeling count). *Recall* is the ratio of number of correct classification to the total number of classifications that are known to be in the said class. The *F1 score* signifies the overall accuracy, it is defined as the harmonic mean of precision and recall.

4.2 Experimental Analysis

We compared our results to eight state-of-the-art sentiment analysis lexicons: Valence Aware Dictionary for sEntiment Reasoning (VADER) [3], Linguistic Inquiry and Word Count (LIWC), General Inquirer (GI), Affective Norms for English Words (ANEW), SentiWordNet (SWN), SenticNet (SCN), Word-Sense Disambiguation (WSD) using WordNet, and the Hu-Liu04 opinion lexicon (Table 1).

As seen in Tables 2, 3, 4 and 5, in most scenarios, our approach outperforms all the other well-established lexicons for sentiment analysis. In case of social media posts, our approach provides better overall precision, recall, and F1 score than human raters.

Data used for training and validation of RNN model is demonstrated in Table 1. The data used for pre-training phase and data used to create the word embedding do not necessarily have to overlap. The neural network was trained on large number of

Table 1 Data used for training the RNN model

Stages	Tweets	Very Pos.	Positive	Neutral	Negative	Very Neg.
Word embedding	100 M	–	–	–	–	–
Pre-training	30 M	4 M	8 M	5 M	6 M	7 M
Training	20104	3996	7114	2551	3449	2994
Validation	3000	798	617	331	169	1085
Test	21132	4091	3613	3921	3083	6424

Table 2 Comparison of 6-class classification performance on social media posts

	Correlation to ground truth (mean of 10 human raters)	Classification accuracy metrics		
		Overall Precision	Overall Recall	Overall F1 Score
Social Media Posts (5,000 Tweets)				
Ind. Human	**0.909**	0.88	**0.81**	**0.83**
Ours	**0.813**	**0.99**	**0.93**	**0.96**
VADER	**0.799**	**0.99**	**0.92**	**0.96**
Hu-Liu '04	0.713	**0.89**	0.67	**0.74**
SCN	0.542	0.79	0.70	0.72
GI	0.512	0.79	0.51	0.67
SWN	0.441	0.74	0.60	0.61
LIWC	0.606	**0.91**	0.49	0.60
ANEW	0.451	0.79	0.46	0.57
WSD	0.401	0.69	0.44	0.52

Table 3 Comparison of 6-class classification performance on movie reviews

	Correlation to ground truth (mean of 10 human raters)	Classification accuracy metrics		
		Overall Precision	Overall Recall	Overall F1 Score
Metacritic Movie Reviews (8,500 review snippets)				
Ind. Human	**0.898**	**0.96**	**0.92**	**0.90**
Ours	**0.691**	**0.81**	**0.70**	**0.69**
VADER	**0.441**	**0.74**	0.58	0.61
Hu-Liu '04	0.359	0.64	0.48	**0.66**
SCN	0.255	0.61	**0.61**	0.5
GI	0.351	0.69	0.61	0.44
SWN	0.245	0.66	0.59	0.60
LIWC	0.168	0.65	0.31	0.44
ANEW	0.164	0.55	0.40	0.44
WSD	0.339	0.60	0.49	0.59

Table 4 Comparison of 6-class classification performance on product reviews on Amazon.com

	Correlation to ground truth (mean of 10 human raters)	Classification accuracy metrics		
		Overall Precision	Overall Recall	Overall F1 Score
Amazon.com Product reviews (11,000 review snippets)				
Ind. Human	0.925	0.95	0.81	0.89
Ours	0.765	0.92	0.85	0.85
VADER	0.578	0.78	0.61	0.71
Hu-Liu '04	0.555	0.77	0.66	0.66
SCN	0.368	0.59	0.69	0.59
GI	0.421	0.69	0.52	0.56
SWN	0.365	0.62	0.53	0.58
LIWC	0.319	0.74	0.39	0.33
ANEW	0.277	0.68	0.39	0.41
WSD	0.305	0.60	0.59	0.56

Table 5 Comparison of 6-class classification performance on social media posts

	Correlation to ground truth (mean of 10 human raters)	Classification accuracy metrics		
		Overall Precision	Overall Recall	Overall F1 Score
NY Times Editorials (3,500 Article review snippets)				
Ind. Human	0.792	0.82	0.60	0.71
Ours	0.611	0.78	0.61	0.66
VADER	0.492	0.67	0.51	0.55
Hu-Liu '04	0.441	0.74	0.49	0.59
SCN	0.259	0.60	0.50	0.41
GI	0.339	0.64	0.48	0.51
SWN	0.289	0.59	0.51	0.59
LIWC	0.228	0.60	0.21	0.29
ANEW	0.212	0.61	0.41	0.45
WSD	0.235	0.60	0.51	0.52

tweets (30 million) for one epoch, and then it was finally trained on the supervised data (20 k tweets) for about 16 epochs.

The accuracy of the recursive-recurrent approach for training and dev phase are plotted in Fig. 1. We used a learning rate of 0.001 and a regularization strength of 0.0001. The overall accuracy of our algorithm compared to different state-of-the-art methods are graphically depicted in Figs. 2, 3, 4 and 5. In Tables 2, 3, 4 and 5, the best, second best, and third best are represented by the colors *red, yellow*, and *green*, respectively.
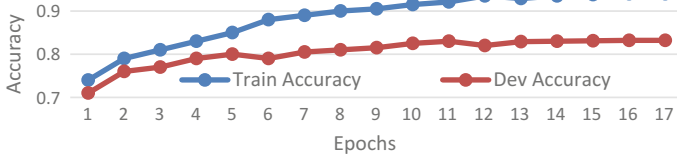
Fig. 1 Transfer learning RNN: training and dev accuracy versus number of epochs

Fig. 2 Performance comparison on sentiment analysis of tweets
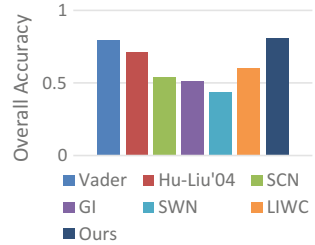
Fig. 3 Performance comparison on sentiment analysis of movie reviews
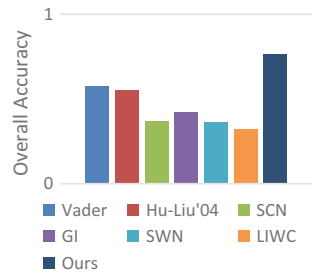
Fig. 4 Performance comparison on sentiment analysis of product reviews
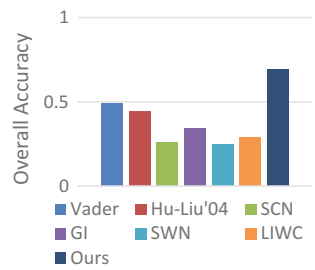
Fig. 5 Performance comparison on sentiment analysis of newspaper editorials
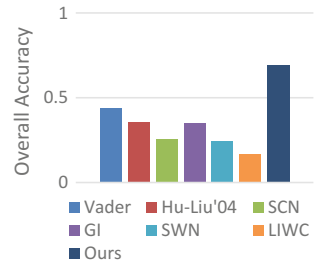
5 Conclusion

In this paper, we extract a vector representation of phrase level sentiments, and use the vector components of a sentence to cross-validate with respect to a gold-standard array of lexical features via recurrent neural network. Furthermore, we segregate the extracted sentimental values in nuanced sentiment classes that make the algorithm more flexible and usable. In future, we aim to explore more robust features via pre-training deep RNNs exploring additional depictive influence available in deeper networks in this architecture and depiction system for other applications, like, personality analysis over social media texts, finding patterns in social media texts, etc.

References

1. Deriu, J., et al.: Leveraging large amounts of weakly supervised data for multi-language sentiment classification. In: Proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee (2017)
2. Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F., Stoyanov, V.: SemEval-2016 task 4: sentiment analysis in Twitter. In: Proceedings of SemEval, pp. 1–18
3. Gilbert, C.J.H.E.: VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Eighth International Conference on Weblogs and Social Media (ICWSM-14). http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf (2014). Accessed 20 Apr 2016
4. Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., Booth, R.J.: The development and psychometric properties of LIWC2007. LIWC.net, Austin, TX (2007)
5. Maas, A.L., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y., Potts, C.: Learning word vectors for sentiment analysis. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 142–150. Association for Computational Linguistics, June 2011
6. Socher, R., Perelygin, A., Wu, J.Y., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C.: Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), vol. 1631, p. 1642, Oct 2013
7. Santos, R.L.S., de Sousa, R.F., Rabelo, R.A.L., Moura, R.S.: An experimental study based on fuzzy systems and artificial neural networks to estimate the importance of reviews about product and services. In: 2016 International Joint Conference on Neural Networks (IJCNN), pp. 647–653. IEEE (2016)

8. Stone, P.J., Dunphy, D.C., Smith, M.S., Ogilvie, D.M.: General Inquirer. MIT Press, Cambridge, MA (1966)
9. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceeding of SIGKDD KDM-04 (2004)
10. Bradley, M.M., Lang, P.J.: Affective norms for English words (ANEW): instruction manual and affective ratings (1999)
11. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA (1998)
12. Cambria, E., Havasi, C., Hussain, A.: SenticNet 2. In: Proceeding of AAAI IFAI RSC-12 (2012)
13. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.: Distributed representations of phrases and their compositionality. In: Advances on Neural Information Processing Systems (2013)

Neighbor Attack Detection in Internet of Things

Arun Thomas, T. Gireesh Kumar and Ashok Kumar Mohan

Abstract Internet of Things (IoT) hegemonies all other technological development the world has seen so far. 6LoWPAN is emerging as the next generation protocol of IoT. 6LoWPAN protocol enables the resource constrained embedded device to connect to the Internet through IPv6. Routing Protocol for Low Power and Lossy Networks (RPL) is specifically designed as a routing protocol for resource constrained device and it is adaptable with the 6LoWPAN. There are many attacks which make RPL insignificant to achieve its desired objectives. Neighbor attack is momentous and is capable of disrupting possible routing path. In this paper, we proposed intrusion detection system which can detect neighbor attack in RPL protocol and a secure root process to prevent the effect of attack on this protocol. The IDS is incorporated by considering location information and received signal strength to identify malicious node. We have also incorporated secure root process which can rectify disruption in routing path after detecting attacks. This method can act as an optimum method for resources constrained environment.

Keywords RPL · IoT · IPv6 · IDS · DODAG and RSSI

1 Introduction

The Internet of Things (IoT) is growing as the part and parcel of our lives. IoT is emerging at fast rate. This new era of technology is expected to be bigger than industrial revolution [1]. Things in IoT can be any device such as temperature sensor, automobile with built in sensor, and farm animal with biochip transponder. We should have zero tolerance toward security loopholes in IoT as they are being implemented in critical sectors like nuclear power plants, hospitals, dams, etc. Routing Protocol for Low Power and Lossy Networks (RPL) is a specially designed

A. Thomas (✉) · T. Gireesh Kumar · A. K. Mohan
TIFAC-CORE in Cyber Security, Amrita School of Engineering,
Amrita Vishwa Vidyapeetham, Amrita University, Coimbatore, India
e-mail: arun2007t@gmail.com

protocol for resources constrained environment. RPL was developed to be used in wireless sensor network. RPL is a routing protocol in network layer to optimize protocol efficiency in IoT like 6LoWPAN [2]. There are many attacks in RPL such as sinkhole, selective forwarding, wormhole, sybil, clone id, neighbor attack, rank attack, and DIS attack [3]. Neighbor attack is a severe attack which can cause alteration in the network, disrupt routing topology in the network, and make the routing algorithm meaningless [4]. No effective techniques exist to counter neighbor attack in RPL. We have designed an efficient intrusion detection system to detect neighbor attack in RPL by keeping an account of resource constrained environment. We have also introduced secure root process which can mitigate the effect of attack on this protocol after detecting the attack. RPL starts the formation of Destination Oriented Directed Acyclic Graph (DODAG) from root node or sink node. RPL and DODAG are maintained using control messages. Root node initially broadcasts DIO messages to build roots in downward direction. Nodes upon receiving DODAG Information Object (DIO) message select the parent of the sender by replying with Destination Advertisement message (DAO). Parent node gives the authorization to join by replying with DIO ACK message [5]. Every node periodically sends DIO message to maintain DODAG and to avoid closed loops. Every node in DODAG has a rank value. The rank is an integer value which is dependent on distance from sink node, delay and other parameters. Nodes nearer to root node will have smaller rank value and nodes far from root node will have larger rank value. The rank value eventually decreases in downstream direction. Ranks are used to avoid loops and to differentiate between parents and siblings in DODAG [6]. In the process of network topology, each router is responsible for identifying stable set of parents on the path toward the DODDAG root. The stable set of parents is referred as preferred parents. The objective function is responsible for deciding ranks in RPL [7]. The objective function uses routing metrics like delay, link, quality, connectivity, etc., and optimization objective for the computation of rank value. The design of optimized objective function is yet an open research issue [8].

The rest of the paper is organized as follows. In Sect. 2, we discuss all the possible attacks in RPL and analyze the impact on the network. We also give brief outline about IDS system to counter each attack in RPL. Section 3, we discuss the architecture of intrusion detection system for detecting neighbor attack. Section 4, we demonstrate how neighbor attack is simulated in cooja environment. We also demonstrated how the proposed intrusion detection system can detect this attack along with secure root process. Section 5, we analyze the future work.

2 Related Works

2.1 Sinkhole Attack

In this attack, compromised node announces optimized path to attract nearby nodes to route tunneled traffic through it. A compromised node captures traffic with the intention to drop the messages and thus reducing the number of messages successfully received to the destination node. IDS system such as SVELTE [4] can detect sinkhole attack. The parent failover and rank authentication are the two techniques to defend again this attack. This attack can cause severe impact in the network when combined with other attack [6].

2.2 Selective Forward Attack

Here, the malicious node works with intent to drop packets selectively, when forwarding messages to other genuine node in the network. In some cases, malicious node could forward all control message and skip the rest of the traffic [9]. The attacker's main intention is to disrupt the routing path. Creating disjoint path between source and destination node is one of the ways to defend against selective forwarding attack. However, we can guard against selective forwarding attack to some extent by encryption and analysis of application level traffic. Heartbeat protocol can also detect the existence of this attack [10].

2.3 Black Hole Attack

Black hole is a special case of selective forwarding. Here, attacker instead of selective forwarding packet, the attacker node silently drops all the packets. Packet delay and control overhead are the effect of attack on the network. However, no techniques have been evaluated to counter this attack [4].

2.4 Sybil Attack

In sybil attack, a malicious node copies several logical identities on the same node. This attack is capable of controlling large parts of network without deploying physical nodes. Techniques which can detect clone id attacks are using distributed hash table, geographical location information of nodes, etc. However no techniques are evaluated to counter this attack [4].

2.5   Neighbor Attack

Neighbor attack creates a drastic impact on RPL protocol. Here, malicious node replicates the DIO message that it received without adding its identity and broadcast them again. Malicious node deceit the victim node pretends to have a neighbor which is not in range. Moreover if victim node observes that the neighbor node has better rank than it, then victim node may request it as the preferred parent and changes the route to out this neighbor. This attack can have catastrophic effect on the network like end to end delay and changes in network topology. However, no techniques are evaluated yet to counter this attack. This attack resembles the wormhole attack but it is limited to selective forwarding of DIO messages [4].

2.6   Local Repair Message

In this attack, malicious node broadcast local repair messages periodically without any problem with link quality. Other nodes upon receiving the local repair message will need to validate the route with the malicious node. This attack can have impact on the network with increased control overhead messages and packet dropping due to temporarily unavailable route. Finite state machine based IDS system can detect this attack [11].

3   Problem Formulation

The neighbor attack creates a drastic impact on RPL protocol. Unfortunately, no techniques have been proven to counter this attack. In this attack, the attacker is only responsible for broadcasting DIO message to victim node which is in its range. Attacker does not include its identity in DIO message and follows normal network behavior. This attack can have acute impact on the network like unoptimized routing path, loops in network topology, and increase in control overhead. The architecture of proposed IDS is shown in Fig. 1. The prime concept of neighbor attack detection works on location verification and RSSI validation. The location information is attached by sensor nodes in the DIO message. The location information is accessed by each sensor node accurately by using X and Y coordinates in the stimulator. When victim receives the broadcasted DIO message from the attacker, it will access location information and RSSI from the received packet. Whenever sensor nodes find new neighbor, distance verification is done using location and RSSI value. First, distance is calculated through location information. Second, we can find distance using RSSI value. Received signal strength indicator can be converted to distance by analyzing simulation environment. Here, sensor nodes validate the distance from location information and RSSI and moreover it
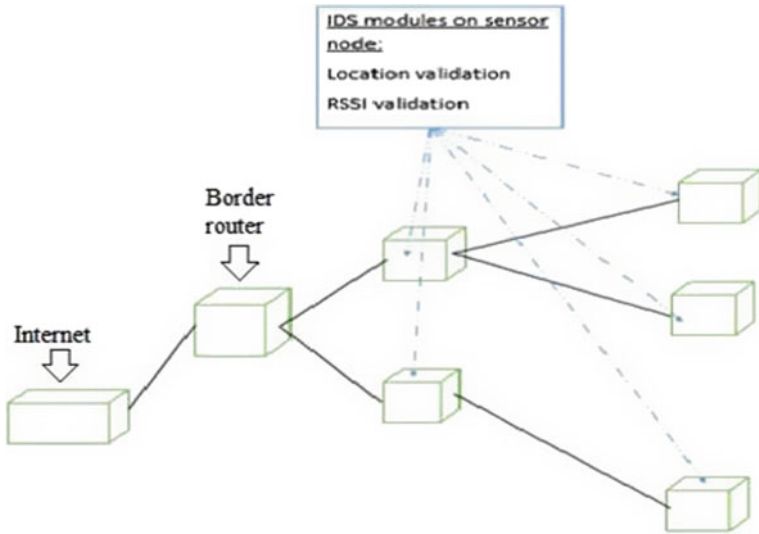
Fig. 1 Architecture of proposed system

should be within fixed communication range. If there is any mismatch beyond the threshold value then the attacker is found as shown in Eq. (1). In attacker case, the rebroadcasted message does not change the location information. Hence, from location information when the distance is calculated, it will be more from the fixed range and also they will mismatch with the distance found using RSSI. We have also incorporated secure root process to recoup from the impact of the attack. After neighbor attack detection, out of range node is not added to neighbor cache of the source.

$$\sqrt{\left((y2-y1)^2+(x2-x1)^2\right)} \,+\, \text{Distance from RSSI} < \text{Threshold value.} \quad (1)$$

y1 and y2 refers to y-axis coordinate position of nodes. x1 and x2 refers to x-axis coordinate position of nodes.

4 Implementation and Result

We run our stimulation setup in cooja [12] environment which produces accurate result. In our stimulation, we use sky mode. We employ RPL as the routing protocol. The placement of nodes is at random as shown in Fig. 2. The node 1 is the border router and other nodes are client node which sends data to node 1. Nodes are placed in such a way that it should be in the range of at least one node. All nodes transmit with 23 dB. The total number of nodes in our simulation environment is
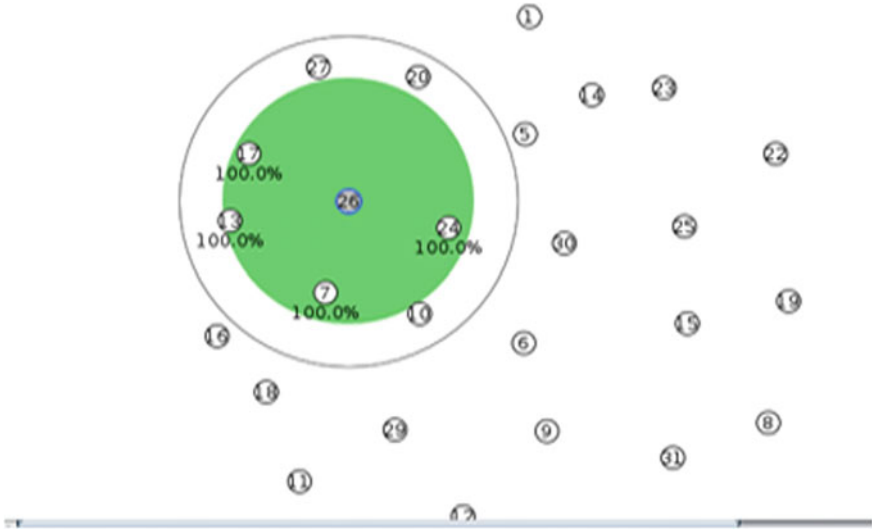
Fig. 2 Neigbors of node 26

31. Node 24 is the attacker node. Node 24 has neighbors namely 26, 5, 30, and 10 as shown in Fig. 2.

Node 26 broadcasts DIO message with rank. Node 17, node 13, and node 24 receives DIO message from node 26. Attacker node 24 broadcast DIO message with same source IP address which is the IP address of node 26. The node 10, node 26, and node 30 receive DIO message from node 24 with the same IP address of node 26. The buffer 100 in Fig. 3 represents the DIO message of node 26. Node 26 is added as neighbor of node 10 and 30. Node 30 is not the neighbor of node 26. Node 30 is not in the transmission range of node 26 and it is not the actual neighbor of node 26.

Now, let us look at the detection mechanism and secure root process for this attack. Our detection works on location information and RSSI value. Initially, we store all 31 nodes positions in an array. For understanding the relation between RSSI value and the distance in our simulation environment, we have broadcasted DIO messages within all 31 nodes by moving nodes at random positions. In our analysis, we have got RSSI value from 13 to 83 dB. We have stored the corresponding distance for each RSSI value from 13 to 83 dB. Here in Fig. 4, we can see node 30's X and Y position. We can also observe the position of node for which DIO message is associated. Whenever a node receives DIO message, it will access the location information and RSSI value from the received packet. We can observe that node 30 computes distance by accessing RSSI value of its received packet. First, we compute difference between coordinate positions of its location and location from that associated DIO. The distance is also computed by analyzing RSSI value from the received packet. If there is a mismatch between distance in location information and RSSI value exceeds threshold value, then the attacker is detected. We have kept the

```
752 00:21.591      ID:30   RPL: Received a DIO from 24·····100
753 00:21.593      ID:10   fe80::212:741a:1a:1a1a
754 00:21.594      ID:30   fe80::212:7418:18:1818
755 00:21.596      ID:30   Change Source IP address 24
756 00:21.599      ID:30   fe80::212:7418:18:1818
757 00:21.602      ID:30   fe80::212:741a:1a:1a1a
758 00:21.604      ID:10   RPL: Neighbor added to neighbor cache fe80::212:741a:1a:1a1a,
     00:12:74:18:00:18:18:18
759 00:21.607      ID:10   RPL: receiving node 10 rank 0 ········buffer 100
760 00:21.610      ID:10   fe80::212:741a:1a:1a1a
761 00:21.613      ID:30   RPL: Neighbor added to neighbor cache fe80::212:741a:1a:1a1a,
     00:12:74:18:00:18:18:18
762 00:21.613      ID:26   RPL: Received a DIO from 24·····100
763 00:21.614      ID:10    RPL: Incoming DIO (id, ver, rank) = (254,128,0)
764 00:21.616      ID:26   fe80::212:7418:18:1818
765 00:21.617      ID:30   RPL: receiving node 30 rank 0 ········buffer 100
766 00:21.618      ID:26   Change Source IP address 24
767 00:21.620      ID:30   fe80::212:741a:1a:1a1a
768 00:21.621      ID:26   fe80::212:7418:18:1818
769 00:21.624      ID:30    RPL: Incoming DIO (id, ver, rank) = (254,128,0)
770 00:21.625      ID:26   fe80::212:741a:1a:1a1a
771 00:21.634      ID:26   RPL: Neighbor added to neighbor cache fe80::212:741a:1a:1a1a,
     00:12:74:18:00:18:18:18
772 00:21.638      ID:26   RPL: receiving node 26 rank 0 ········buffer 100
773 00:21.641      ID:26   fe80::212:741a:1a:1a1a
774 00:21.645      ID:26    RPL: Incoming DIO (id, ver, rank) = (254,128,0)
775 00:21.763      ID:7    aaaa::ff:fe00:1RPL: rpl_process_parent_event recalculate_ranks
776 00:21.790      ID:14   RPL: rpl_process_parent_event recalculate_ranks
```

Fig. 3 Victim nodes receiving DIO message

```
799 00:21.856      ID:30   RPL: Received a DIO from 24·····100·······4·····last_rssi ·88
800 00:21.857      ID:10   Current Source to neighbours Distance ····· 640 64 576
801 00:21.859      ID:30   fe80::212:7418:18:1818
802 00:21.862      ID:10   Change Source IP address : src 24 X_pos ·4 ····· Y_pos ·9
803 00:21.863      ID:30   RRRRRRR : tx_power ····· 23 ······rx_power ·88
804 00:21.865      ID:10   fe80::212:7418:18:1818
805 00:21.866      ID:30   RSSI : rss_ ····· ·88 ······distance_ 1160
806 00:21.868      ID:10   fe80::212:741a:1a:1a1a
807 00:21.871      ID:30   Current Source to neighbours Distance ····· 1165 1156 9
808 00:21.871      ID:10    Current Node 10 X_pos 16 ····· Y_pos 23
809 00:21.874      ID:10    Source Node 26 X_pos ·4 ····· Y_pos ·9
810 00:21.875      ID:30   Change Source IP address : src 24 X_pos ·4 ····· Y_pos ·9
811 00:21.878      ID:30   fe80::212:7418:18:1818
812 00:21.880      ID:10   Original Source to neighbours Distance ····· 1424 rss_dist
     738 dist_diff 686
813 00:21.882      ID:30   fe80::212:741a:1a:1a1a
814 00:21.885      ID:30    Current Node 30 X_pos 58 ····· Y_pos 2
815 00:21.888      ID:30    Source Node 26 X_pos ·4 ····· Y_pos ·9
816 00:21.891      ID:10   RPL: Neighbor added to neighbor cache fe80::212:741a:1a:1a1a,
     00:12:74:18:00:18:18:18
817 00:21.894      ID:30   Original Source to neighbours Distance ····· 3965 rss_dist
     1160 dist_diff 2805
818 00:21.894      ID:10   RPL: receiving node 10 rank 0 ········buffer 100
819 00:21.897      ID:30   Attacker is detected ····· 24 ···· 1
820 00:21.898      ID:10   fe80::212:741a:1a:1a1a
821 00:21.901      ID:10    RPL: Incoming DIO (id, ver, rank) = (254,128,0)
822 00:21.908      ID:30   RPL: Neighbor added to neighbor cache fe80::212:741a:1a:1a1a,
     00:12:74:18:00:18:18:18
```

Fig. 4 Attack is detected

threshold for mismatch as 44.721 m. Here, distance differs by 52.96 (i.e., sqrt of 2805). Since this is more than 44.721, the attacker is detected. We have also introduced secure root process. After attacker detection out of range node is not
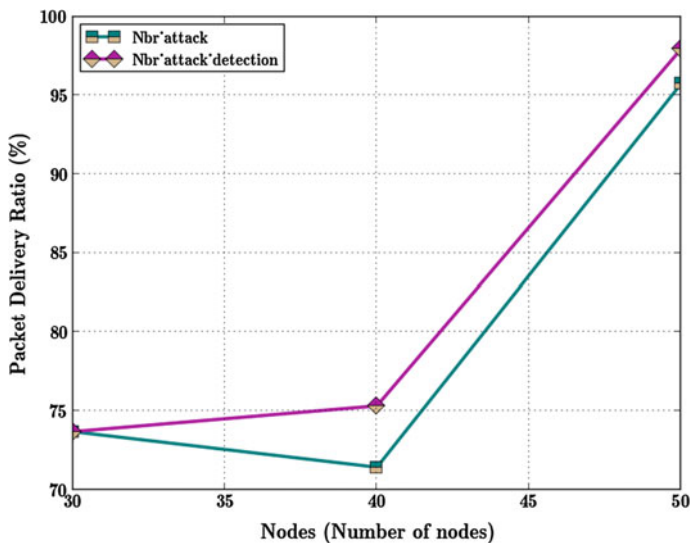
Fig. 5 Packet delivery ratio

added to neighbor cache source. Here in Fig. 4, the attack is detected by node 30 and node 30 does not add 24 as a neighbor.

For better graphical representation, we run our simulation for 31, 41, ad 51 nodes by keeping the same attack scenario. We have considered nodes from 30 to 50 for clear understanding the impact of this attack (Fig. 5).

X coordinates indicate the number of nodes is from 30 to 50. Y coordinates indicate the packet delivery ratio is from 70 to 100. Green line indicates the effect of packet delivery ratio with respect to number of nodes before the attack detection. Pink line indicates the effect of packet delivery ratio with respect to number of nodes after the attack detection. When we analyze the graph, we can conclude that packet delivery ratio increased slightly after detecting this attack when compared to before detecting this attack (Fig. 6).

X coordinates indicate that the number of nodes is from 30 to 50. Y coordinates indicate the throughput and it is measured in Mbps. Green line indicates the change in throughput with respect to number of nodes before the attack detection. Pink line indicates the change in throughput with respect to number of nodes after the attack detection. We can conclude from graph that throughput increased slightly after detecting this attack, when compared before detection (Fig. 7).

X coordinates indicate that the number of nodes is from 30 to 50. Y coordinates indicate that the total delay is from 0 to 12. The total delay is measured in seconds. Green line indicates the alteration in total delay with respect to number of nodes before the attack detection. Pink line indicates the alteration in total delay with respect to number of nodes after the attack detection. We can infer from the graph
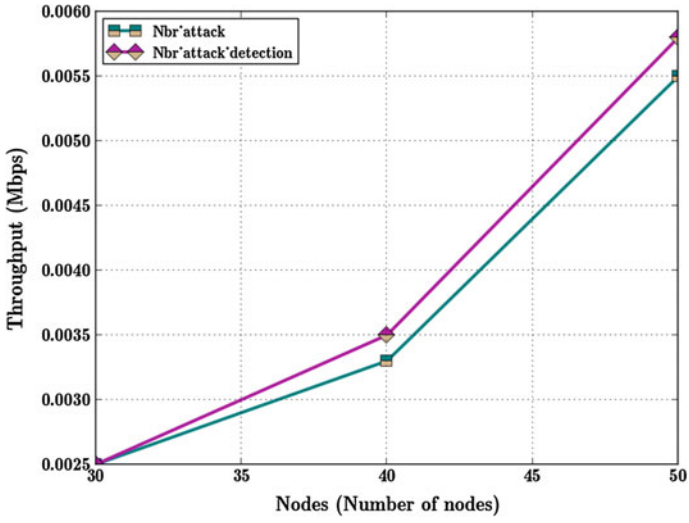
Fig. 6 Throughput
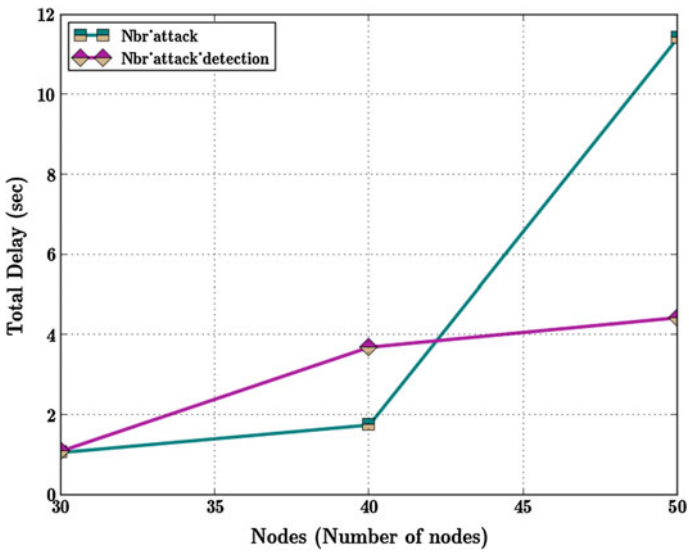
Fig. 7 Total delay

that total delay drastically reduced after detection from node 40 to 50 when compared before detection. The total delay from node 40 to 50 remains to be stable after detection.

5   Conclusion and Future Works

The proposed method is the first work on neighbor attack detection. The proposed concept can detect neighbor attack in cooja stimulator with very high accuracy and is designed according to resource constrained environment. The proposed IDS can detect clone id and sybil attack but not evaluated yet. Research community should descent from traditional IDS system which is based on signature-based detection. There are many potential attacks in RPL and the number of attacks is increasing. The proposed method can be used in wormhole to enhance intrusion detection capability. The location information of node will help research community to mitigate future attacks in RPL.

References

1. http://www.businessinsider.com/iot-ecosystem-internet-of-things-predictions-and-business-potential-2016-7?IR=T1
2. Le, A., Loo, J., Chai, K.K., Aiash, M.: A specification-based IDS for detecting attacks on RPL-based network topology. Information **7**(2), 25 (2016)
3. Rghioui, A., Khannous, A., Bouhorma, M.: Denial-of-Service attacks on 6LoWPAN-RPL networks: threats and an intrusion detection system proposition. J. Adv. Comput. Sci. Technol. **3**(2), 143 (2014)
4. Pongle, P., Chavan, G.: A survey: attacks on RPL and 6LoWPAN in IoT. In: 2015 International Conference on Pervasive Computing (ICPC), pp. 1–6. IEEE, Jan 2015
5. Le, A., Loo, J., Luo, Y., Lasebae, A.: The impacts of internal threats towards routing protocol for low power and lossy network performance. In: 2013 IEEE Symposium on Computers and Communications (ISCC), pp. 000789–000794. IEEE, July 2013
6. Weekly, K., Pister, K.: Evaluating sinkhole defense techniques in RPL networks. In: 2012 20th IEEE International Conference on Network Protocols (ICNP), pp. 1–6. IEEE, Oct 2012
7. IETF, RPL. Routing Over Low Power and Lossy Networks
8. Kasinathan, P., Pastrone, C., Spirito, M.A., Vinkovits, M.: Denial-of-Service detection in 6LoWPAN based Internet of Things. In: 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 600–607. IEEE, Oct 2013
9. Airehrour, D., Gutierrez, J., Ray, S.K.: Secure routing for internet of things: a survey. J. Netw. Comput. Appl. **66**, 198–213 (2016)
10. Gaddour, O., Kouba, A.: RPL in a nutshell: a survey. Comput. Netw. **56**(14), 3163–3178 (2012)
11. Jose, E., Gireesh, K.T.: Gigabit network intrusion detection system using extended bloom filter in reconfigurable hardware: In: Proceedings of the Second International Conference on Computer and Communication Technologies: IC3T 2015, vol. 1, pp. 11–19 (2016)
12. https://sourceforge.net/projects/contiki/files/Instant%20Contiki/

Human Opinion Inspired Feature Selection Strategy for Predicting the Pleasantness of a Molecule

Ritesh Kumar, Rishemjit Kaur, Amol P. Bhondekar and Gajendra P. S. Raghava

Abstract The identification of features responsible for smell of a molecule has been a long-standing challenge. We use cheminformatics and opinion dynamics based optimization algorithm to identify feature subsets of a molecule, which can predict how pleasant a molecule will smell. We have also compared it to standard feature selection techniques. The features identified reveal that three classes of features are primarily responsible for pleasantness. The work may open up some innovative inroads into feature identification and their physical understanding into the olfactory stimulus-percept problem.

1 Introduction

The sense of smell has been the least understood of all our senses and it has been a challenge to predict the smell of a novel molecule by its physicochemical structure, or the physicochemical structure of a novel smell. There has been a considerable amount of study identifying molecular signatures or "rule-based fingerprints" to predict the smell of a group of molecule. For example, the ester group is known to have fruity smell and musky odor is related to three specific conditions: measured

R. Kumar (✉) · R. Kaur (✉) · A. P. Bhondekar
CSIR-Central Scientific Instruments Organisation, Chandigarh, India
e-mail: riteshkr@csio.res.in

R. Kaur
e-mail: rishemjit.kaur@csio.res.in

R. Kumar · R. Kaur · A. P. Bhondekar · G. P. S. Raghava
Academy of Scientific & Innovative Research, New Delhi, India

G. P. S. Raghava
CSIR-Institute of Microbial Technology, Chandigarh, India

standard desorption rates in the range 0.4–1.7, molecular cross sections of 40–57 square A^0, and ratios of 2.8–3.3 of length to breadth of the molecule [1]. More recently, Khan et al. [2] identified principal components to be directly related to the pleasantness of the molecule, Snitz et al. [3] and Kumar et al. [4] tried to find molecular signatures related to the perceptual descriptors and Boyle et al. [5] designed a machine learning pipeline to relate the physicochemical properties to the olfactory receptors. All these works have used machine learning approaches but the data set used is very small and they have not identified some important features which could be directly linked for the prediction of smell. There is also a motivation to find optimal feature subsets in these studies in order to derive physical meaning and virtual screening of large number of molecules.

Evolutionary algorithms and swarm intelligence have been employed for feature subset selection in different areas such as face recognition [6], image analysis [7], sensor selection for electronic nose and electronic tongue [8, 9], gene analysis, and so on. Generally, feature selection techniques are classified in two main categories, filter and wrapper. For filter approaches, different measures such as entropy, distance, rough set theory, fuzzy set theory, etc., have been applied to evaluate the goodness of selected features. In wrapper approaches, different classifiers such as SVM and KNN have been used to evaluate the quality of features selected by search algorithm. Recently, it has been shown that human opinion dynamics can be used to solve complex mathematical problems [10] and referred to as Continuous Opinion Dynamics Optimizer (CODO). It has also been shown to be performing better than other optimization algorithms.

This work aims at identifying important physicochemical features of molecules responsible for predicting the pleasantness of molecules. We performed a wrapper-based feature selection employing CODO and support vector machines for maximizing the accuracy of pleasantness prediction. We have compared it against standard techniques. We believe that by employing and using these features one can design a pleasant or unpleasant molecule. By employing opinion dynamics optimizer, we show their usefulness in this field. It will also contribute new sources and fresh approaches to the fields of modern data analysis and olfaction science.

2 Methodology

2.1 Data Set

We worked on the data published by Keller and Vosshall [11] in which 476 odorants were rated by human subjects for pleasantness on the scale of 0–100, with 0, meaning least pleasant and 100 being the highest. We divide the values less than 50 to unpleasant and more than 50 to pleasant. Further, we obtained the molecular descriptors using Dragon (a commercial software package to generate descriptors). Hence, the dataset has 4870 features (such as number and type of functional groups,

topological and geometrical descriptors etc.) to relate to the perceptual quality of a molecule.

2.2 Continuous Opinion Dynamics Optimizer (CODO)

It is inspired from Durkheim theory of social integration [10] according to which opinion formation is influenced by the two opposing forces: integrative, stemming from people's desire to be a part of the society and disintegrative, coming from individual's desire to seek uniqueness. It is the presence of these two opposing forces that imparts exploitation and exploration capabilities to the individuals, which is the central tenet of optimization algorithms. Let us assume that the society has M number of individuals representing candidate solutions to the problem in D-dimensional space. Each individual i is characterized by its opinion vector $o_i = [o_i^1, o_i^2, \ldots o_i^D] \in R^D$ and social rank SR_i at the particular time or iteration t. The opinions are real valued and initialized randomly using a uniform distribution. The detailed description of the important components of CODO is provided below.

Social Structure/Topology: It defines the connections between the individuals. It can be regular, scale free, random, or fully connected in nature.

Social Influence: In a society, individuals constantly interact with each other and during these social interactions, they influence others and also get influenced by others. This effect is often termed as "social influence" (w_{ij}), defined by $w_{ij} = SR_j(t)/d_{ij}(t)$. It depends on two factors:

1. *Social Ranking (SR)*: For every individual i, the fitness (or function) value f_i, which is the output of the function to be minimized (maximized), is used for calculation of the Social Rank (*SR*). The individual with the minimum (maximum) function value is assigned the highest *SR*.

2. *Distance between individuals (d_{ij})*: It is the distance between individuals i and j in the topological space.

   The better-fit and nearby individuals will have more social influence on others.

Updating Dynamics: The updating dynamics has two components, integrating and disintegrating forces. The mathematical formulation of this dynamics is given by (1).

$$\Delta o_i = \frac{\sum_{j=1}^{N} (o_j(t) - o_i(t)) w_{ij}(t)}{\sum_{j=1}^{N} w_{ij}(t)} + \xi_i(t), j \neq i \tag{1}$$

where t represents the iteration, N is the number of neighbors of individual i, $o_j(t)$ is the opinion of the neighbor j, $w_{ij}(t)$ is the social influence exerted by neighbor j on individual i, and $\xi_i(t)$ represents the "*adaptive noise*".

The first component of this equation represents integrative forces present in the society. It is the weighted average of the opinion differences of an individual from its neighbors, and the weights represent the social influence. The second component of Eq. (1) stands for the disintegrative forces in a society. It is a normally distributive random noise with zero mean and a standard deviation of $\sigma_i(t) = S \sum_{j=1}^{N} e^{-\left|f_{ij}(t)\right|}$, where S is the strength of the disintegrating forces and $|f_{ij}(t)|$ denotes the difference in fitness values of an individual i and its neighbor j. As difference in the fitness values of individual i and its neighbors j decreases, $\sigma_i(t)$ increases and leads to more exploration in search space. The pseudocode for CODO is as follows:

- *society.opinion* = GenerateInitialSociety(Xmin, Xmax);
- *iter* = 0;
- **while** (*iter* < max_iter && error > = min_error) **do**
- *society.fitnessVal* = EvaluateFitnessFcn(*society.opinion*);
- *society.ranking* = CalcRank(*society.fitnessVal*); //It ranks the individuals based on society fitness values
- *iter* = *iter* + 1;
- **for** each individual i **do**
- **for** each dimension d **do**
- Calculate w_{ij} of neighbors j of individual i with respect to dimension d
- Update opinion of individual i as defined in Eq. 1
- **end(for)**
- **end (for)**
- **end (while)**

2.3 Implementation of Wrapper Approach Combining CODO and SVM

Generally, each individual/candidate solution in the society/population is encoded using a vector of n real numbers, where n is equal to total number of features present in data and real numbers represent the weights of features. In our case, we have fixed the length of individual to the desired or reduced number of features and each opinion represents the feature number to be selected. For example, in Fig. 1, the solution vector [2,3,5] indicates that feature number 2, 3, and 5 will be selected. Please note that algorithm works on real numbers, which we round off to the nearest integer representing the feature number. This does not imply that the adjacent features are necessarily co-located in search space.

Each individual represents a features subset and we evaluate their fitness using SVM and 10-fold cross-validation, i.e., in every iteration and for every individual,
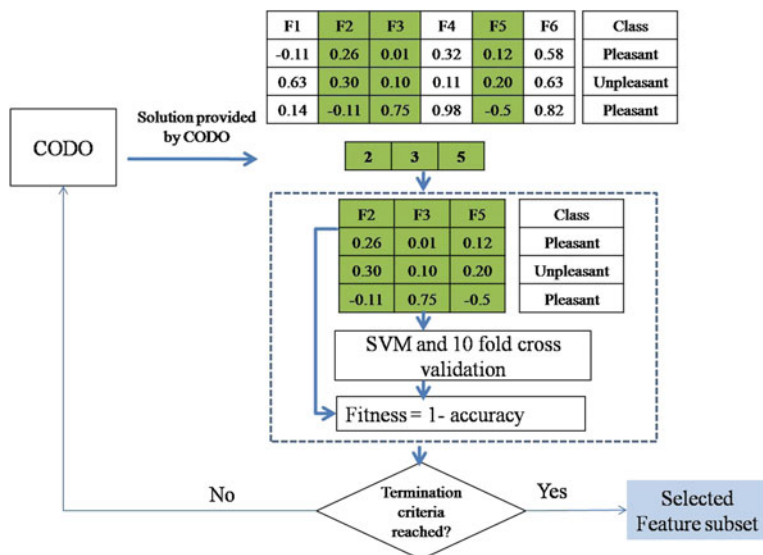
Fig. 1 Methodology for feature subset selection

we split the data into 10 sets and train the SVM classifier on 9 sets and test it on 1 remaining set. The accuracy on the test set represents the fitness of individual, i.e., f = 1−accuracy.

The fitness function aims to minimize the misclassification error on the dataset. The algorithm has been implemented in MATLAB® using standard library for SVM (LIBSVM) [6] and CODO (developed in-house and an open source library has been hosted on https://github.com/rishemjit/CODO).

3 Results

At first, we removed constant columns from the data by removing the columns whose standard deviations were zero. This left us with 3031 columns of physico-chemical features. All these features represent some or other characteristic of a molecule and are highly redundant. For all the analyses, we divided the data into training and test sets (80 and 20%). For feature selection, we performed 10-fold cross-validation on training set, i.e., the accuracy used in the fitness function is based on 10-fold cross-validation on the training data. The results shown here are on the test data.

We employed naive Bayes, SVM with linear kernel, and linear discriminant analysis (LDA) for classifying the full data (without feature selection). Table 1 shows the results obtained. We can see that SVM with linear kernel performs the best on all the features. Thus, we continued with using SVM for further study.

Table 1 Performance results for complete data (without performing feature selection)

	Accuracy (%)	Precision	Recall
Naïve Bayes	63.31	0.604	0.758
LDA	59.73	0.588	0.633
SVM	62.01	0.624	0.673

Table 2 Performance results with (a) f-test based feature selection and (b) single feature based wrapper feature selection

	(a) F-test			(b) Single feature based wrapper feature selection		
	Accuracy (%)	Precision	Recall	Accuracy (%)	Precision	Recall
N = 10	63.33	0.623	0.676	64.53	0.627	0.711
N = 20	64.81	0.625	0.717	62.12	0.617	0.634
N = 30	66.52	0.673	0.67	63.96	0.636	0.675
N = 50	68.97	0.686	0.682	62.10	0.632	0.610
N = 100	66.04	0.655	0.664	60.37	0.595	0.669

The fact that the accuracy is at most 63% on all the features shows the difficulty of the problem.

We further ranked the features according to the f-score or the prediction power of the features. The high ranking features should have higher interclass to intraclass variance. We extracted the top 10, 20, 30, 50, and 100 features. We subjected these features to classification using SVM with linear kernel and results are shown in Table 2a. Clearly, there is an improvement in accuracy and the highest accuracy is at 50 number of features. Further, we used another feature ranking technique where the relative importance of each feature is measured by its classification accuracy. We employed SVM as the classifier. We repeated the procedure of employing 10, 20, 30, 50, and 100 top features for classification. Table 2b shows the results obtained. There is no marked improvement as such in the results.

We employed wrapper-based feature selection using CODO which used SVM as the classifier. The schema of process involved has seen described in the methodology section. The algorithmic parameters such as society size, neighborhood size, topology, strength of disintegrating forces (S), and max number of function evaluations were set to 20, 2, ring, 0.1 and 10000, respectively. We performed the experiments with 10, 20, 30, 50, and 100 features. 30 runs were performed for each experiment with random initialization of the population. Table 3 shows the mean performance along with standard deviation over 30 runs. It can be observed that the maximum mean accuracy of 73.02% has been achieved with 50 features.

Clearly, the best performance has been observed using CODO and SVM. Further, we have used one of the best performing runs with 20 features for visualization with Principal Component Analysis (PCA). This was undertaken to better understand the features and for the ease of visualization. Figure 2 shows the biplot

Table 3 Average performance results (over 25 runs) for wrapper-based feature selection employing CODO and SVM

30 runs, mean performance	Accuracy (%)	Precision	Recall
CODO, N = 10	71.75 ± 1.094	0.64 ± 0.028	0.64 ± 0.061
CODO, N = 20	72.19 ± 1.323	0.66 ± 0.032	0.58 ± 0.043
CODO, N = 30	72.49 ± 0.987	0.64 ± 0.037	0.59 ± 0.068
CODO, N = 50	**73.02 ± 1.335**	0.62 ± 0.040	0.60 ± 0.048
CODO, N = 100	72.80 ± 0.883	0.61 ± 0.027	0.60 ± 0.039

Fig. 2 PCA biplot representation (PC1-PC3) of the 20 features selected using wrapper approach based on CODO and SVM
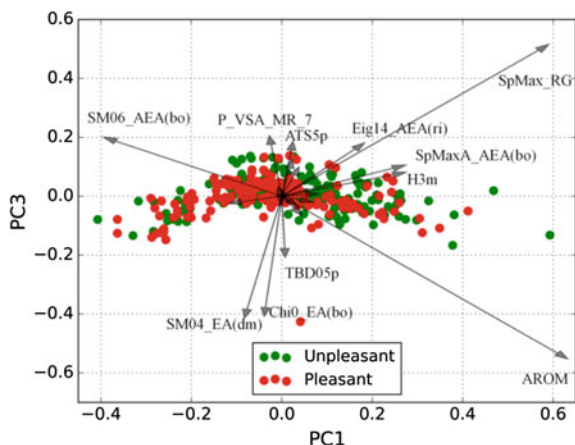
representation (PC1-PC3) of the 20 features. PC1-PC2 does not seem to differentiate between the pleasantness of the molecules. It can be observed that even though there is a huge overlap between pleasant and unpleasant molecules due to the complexity of the dataset, the pleasant molecules are dominant on the left side of the PC1 whereas unpleasant molecules are on the right side. The dominant features seem to stem from three major type of indices, i.e., edge adjacency (e.g., SM06_AEA(bo), SM04_EA(dm)), 3-D features (e.g. SpMax_RG) and aromaticity index (AROM). The presence of these features suggests a long-held belief that the graphical or structural representation along with aromaticity which is also related to molecular weight is responsible for the pleasant smell of a molecule. Further, we also calculated the stability of each feature in terms of its occurrence probability over 30 runs. It was done with 10, 20, 30, 50, and 100 feature subsets separately. Table 4 lists the features with the highest occurrence probability for each feature set size. The specific reason for the presence of these features is a matter of further investigation.

Table 4 Features with highest occurrence probabilities

Total number of features	Feature with the highest occurrence probability	Highest occurrence probability
10	Eig14_AEA(dm)	0.13
20	RDF050m	0.27
30	RDF035u	0.17
50	RDF050m	0.40
100	P_VSA_m_5	0.43

4 Conclusion

In this work, we have identified some important physicochemical features of molecules responsible for predicting the pleasantness of molecules. We have used CODO for selecting important feature subset for predicting pleasantness of molecule. We have also compared it against standard techniques. Clearly, CODO stands out in better accuracy and finding minimum number of features for the task. The selected features have been analyzed and found to be mainly belonging to three subclasses of features of a molecule. The physical meaning of these subclasses need deeper analysis. This technique can also be used for finding the important features responsible for other descriptors of molecules such as musky, fruit, fish, etc.

References

1. Davies, J.T.: Olfactory theories. In: Olfaction, pp. 322–350. Springer (1971)
2. Khan, R.M., Luk, C.-H., Flinker, A., Aggarwal, A., Lapid, H., Haddad, R., Sobel, N.: Predicting odor pleasantness from odorant structure: pleasantness as a reflection of the physical world. J. Neurosci. **27**, 10015–10023 (2007)
3. Snitz, K., Yablonka, A., Weiss, T., Frumin, I., Khan, R.M., Sobel, N.: Predicting odor perceptual similarity from odor structure. PLoS Comput. Biol. **9**, e1003184 (2013)
4. Kumar, R., Kaur, R., Auffarth, B., Bhondekar, A.P.: Understanding the odour spaces: a step towards solving olfactory stimulus-percept problem. PLoS ONE **10**, e0141263 (2015)
5. Boyle, S.M., McInally, S., Ray, A.: Expanding the olfactory code by in silico decoding of odor-receptor chemical space. Elife **2**, e01120 (2013)
6. Ajit Krisshna, N.L., Deepak, V.K., Manikantan, K., Ramachandran, S.: Face recognition using transform domain feature extraction and PSO-based feature selection. Appl. Soft Comput. **22**, 141–161 (2014)
7. Ghamisi, P., Couceiro, M.S., Martins, F.M.L., Benediktsson, J.A.: Multilevel image segmentation based on fractional-order Darwinian particle swarm optimization. IEEE Trans. Geosci. Remote Sens. **52**, 2382–2394 (2014)
8. Kaur, R., Kumar, R., Gulati, A., Ghanshyam, C., Kapur, P., Bhondekar, A.P.: Enhancing electronic nose performance: a novel feature selection approach using dynamic social impact theory and moving window time slicing for classification of Kangra orthodox black tea (Camellia sinensis (L.) O. Kuntze). Sens. Actuators B Chem. **166–167**, 309–319 (2012)

9. Bhondekar, A.P., Kaur, R., Kumar, R., Vig, R., Kapur, P.: A novel approach using dynamic social impact theory for optimization of impedance-tongue (iTongue). Chemom. Intell. Lab. Syst. **109**, 65–76 (2011)
10. Kaur, R., Kumar, R., Bhondekar, A.P., Kapur, P.: Human opinion dynamics: an inspiration to solve complex optimization problems. Sci. Rep. 3 (2013)
11. Keller, A., Vosshall, L.B.: Olfactory perception of chemically diverse molecules. BMC Neurosci. **17**, 55 (2016)

An Ensemble Learning Based Bangla Phoneme Identification System Using LSF-G Features

Himadri Mukherjee, Sourav Ganguly, Santanu Phadikar and Kaushik Roy

Abstract Technology has evolved a lot in the last decade, and various devices have come up for assisting us in our day-to-day life. There has always been a need for simplifying the User Interfaces (UI) of such devices so that they can be easily interacted with, and a speech based UI can be a potential solution. Speech recognition is the task of identification of words from voice signals. Every language consists of a set of atomic sounds called Phonemes which builds up the entire vocabulary of that language. Speech recognition in Bangla is rather a complicated task due to the complex nature of the language like the presence of compound characters. In this paper, a Bangla Phoneme recognition system is proposed towards the development of a Bangla Speech recognition system based on Line Spectral Frequency-Grade (LSF-G) features derived from standard line spectral frequency values. The system has been tested on a Bangla Swarabarna Phoneme dataset of 3290 clips and an accuracy of 94.01% has been obtained with an Ensemble learning based approach.

Keywords Phoneme · Grading · Ensemble learning · Linear spectral frequency
LSF-G

H. Mukherjee (✉) · S. Ganguly · K. Roy
Department of Computer Science, West Bengal State University,
Kolkata 700126, India
e-mail: himadrim027@gmail.com

S. Ganguly
e-mail: sourav.ganguly.iam@gmail.com

K. Roy
e-mail: kaushik.mrg@gmail.com

S. Phadikar
Department of Computer Science and Engineering, Maulana Abul Kalam
Azad University of Technology, Kolkata 700064, India
e-mail: sphadikar@yahoo.com

Fig. 1 Graphical illustration
of the proposed system
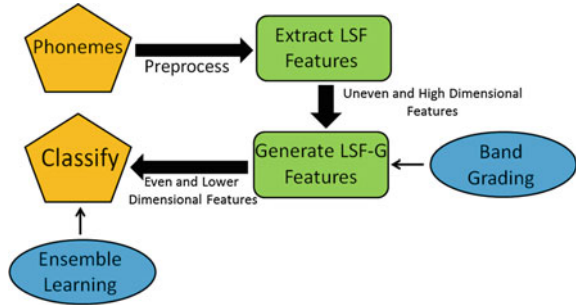
1 Introduction

To make the available devices more user-friendly in order to ease out our day-to-day life, the UI of those devices needs to be brought within our grasp. The UI needs to be developed in such a manner so that we get the feel of interacting with our peers and not a device. Speech recognition based UIs can be a solution to this as speech is the most natural and spontaneous mode of communication. Though Automatic Speech Recognizers (ASRs) in various languages are commercially available but a fully functional Bangla ASR is yet to be developed. Bangla is the 6th most popular language in the world amounting to approximately 874 million speakers worldwide [1], and thus there is a need for the aforesaid. Speech recognition is the technique of identification of words and phrases from voice signals. Every language consists of an inventory of atomic sounds called Phonemes which constitute its Phoneme set. All the words of a particular language are formed by restricted permutation of elements from its Phoneme set. Among them, the vowel (Swarabarna in Bangla) Phonemes are extremely important as because there are only a very few meaningful words which do not have a vowel Phoneme thereby making their identification an important task. In this paper, a Line Spectral Frequency (LSF) based feature namely Line Spectral Frequency-Grade (LSF-G) coupled with an Ensemble learning based classifier has been used to distinguish Bangla Swarabarna (vowel) Phonemes. The proposed system is graphically illustrated in Fig. 1.

In the rest of the paper, the related work is presented in Sect. 2 followed by the details of dataset in Sect. 3. Sections 4 and 5 cast's light on feature extraction and classification, respectively. The result and discussion is presented in Sect. 6, and finally the conclusion is drawn in Sect. 7.

2 Related Works

Scientists have adopted various methodologies for building automatic speech recognizers ever since Dudley's [2, 3] proposal of a speech analysis and synthesis system in 1930. It was, however, Forgie et al. [4] who first tried to recognize speech

in the form of English vowels with the aid of a computer in 1959. Speech recognition systems in English have developed considerably since then. Desai et al. have presented some of the commonly used feature extraction and speech recognition techniques in [5]. An array of highly accurate speech recognizers is now commercially available for English and a few other languages [6–8]. Besacier et al. [9] have presented a survey of Speech recognition systems for various under resourced languages. Though research in Bangla Speech Recognition started as early as in 1975 [10], it is yet to reach considerable maturity. Hasnat et al. [11] proposed a system for the recognition of isolated as well as continuous speech in Bangla with the aid of Hidden Markov Model (HMM) for 100 words recorded by only five speakers. They used a 39 dimensional feature set consisting of 12 MFCC features, a single energy coefficient, and 13 first-order and second-order derivatives each. In the case of isolated speech, accuracies of 90% and 70% were obtained for speaker dependent and independent modes, respectively. In the continuous speech recognition scenario, accuracies of 80% and 60% were obtained for speaker dependent and independent modes respectively. Hasanat et al. [12] classified Bangla Phonemes using 13 reflection coefficients with 13 autocorrelations. They obtained an accuracy of 80% using a Euclidean distance based approach for classifying unknown Phonemes. Ali et al. [13] presented four different models for the recognition of Bangla words. The first one used MFCC features and Dynamic Time Warping (DTW) which produced an accuracy of 78%. The second model was based on linear predictive coding features and DTW which produced an accuracy of 60%. The next was based on Gaussian mixture model and MFCC along with posterior probability function. An accuracy of 84% was obtained for this model. The final model used Linear Predictive Coding (LPC) based MFCC features and DTW which produced an accuracy of 50%. An isolated Bangla word recognizer was proposed by Firoze et al. [14] based on spectral features and fuzzy logic based classifier. An accuracy of 80% was obtained for this system on a dataset of only 50 words. Kotwal et al. [15] used hybrid features (MFCC and Phoneme probability derived from the MFCCs and acoustic features) to classify Bangla Phonemes. A speech corpus of 3000 sentences uttered by 30 male speakers was used to train an HMM classifier which produced an accuracy of 58.53% when tested on a database of 1000 sentences uttered by 10 speakers. Hossain et al. [16] applied various classification techniques for the recognition of six vowels and four consonant Phonemes in Bangla with MFCC features. The system was tested on a small dataset of 300 Phonemes (30 each). Accuracies of 93.66%, 93.33%, and 92% were obtained for Euclidean distance measure, Hamming distance, and artificial neural network respectively.

3 Dataset Development

Data is an important aspect of any experiment whose quality has vital leverage on the final outcome. Care needs to be taken during the course of data collection so that noise and other sorts of error cannot contaminate the collected data. To

Table 1 Bangla vowel phonemes with their IPA symbol, alphabetic representation, pronunciation as in a Bangla word and equivalent English pronunciation

IPA Symbol	ɔ	a	i	u	e	o	æ
Bangla Representation	অ	আ	ই, ঈ	উ, ঊ	এ	ও	অ্যা
Bangla Pronunciation as in	অবাক	আলো	বিশাল	পটু	তেল	ওম	ভেরা
Equivalent English Pronunciation	Pot	Theta	Rig	Pool	Fade	Coat	Lab
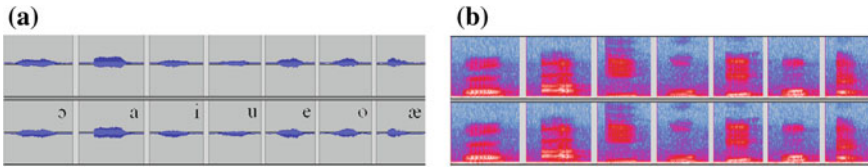
Fig. 2 **a** Amplitude-based representation of the Phonemes. **b** Spectrum based representation of the Phonemes

the best of our knowledge, there is no standard freely available Bangla Phoneme database. Thus, we developed a dataset of our own with the help of 47 volunteers of whom 28 were male and 19 were female aged between 19 to 77. The seven Bangla Swarabarna Phonemes along with their phonetic symbols and their pronunciation in Bangla as well their equivalent English pronunciation are presented in Table 1. The Phonemes were recorded using various locally made headphones including Frontech JIL-3442 with Audacity [17] which were stored in .wav stereo format at a bit rate of 1411 kbps. Each of the volunteers uttered the seven Phonemes one after the other which was repeated 10 times. The Phonemes were then separated using an amplitude based semi-supervised method which ultimately produced a dataset of 3290 ($7 \times 10 \times 47$) Phonemes. The amplitude and spectral content based representations of the Phonemes are presented in Fig. 2.

4 Feature Extraction

4.1 Framing and Windowing

The spectral properties of a speech signal tend to deviate much throughout the entire span of a signal. Thus, in order to facilitate analysis, the signals are partitioned into small parts called frames within which the spectral properties appear to be pseudo-stationary. In order to ensure smoother transition between two consecutive frames, the signals are framed in overlapping manner where a certain number of points from the end of a frame overlap with the starting points of the next frame. In our experiment, the signals were framed into 256 sample point frames with an overlap of 100

points. A signal consisting of n sample points can be partitioned into N overlapping frames of size F with an overlap factor of O as shown in Eq. 1.

Jitters might be observed in frames which produces spectral leakage at the time of Fourier transformation (required for frequency content analysis). In order to remove such components, the frames are multiplied by a windowing function. In our experiment, Hamming window [18] was chosen for this purpose based on trial runs.

$$N = \left\lceil \frac{n - F}{O} + 1 \right\rceil \tag{1}$$

4.2 Line Spectral Frequency Extraction

Line Spectral Frequency [19] is a unique technique of representing linear predictive coefficients, which ensures higher interpolation properties. One of its primary characteristics is its better and effective quantization capability [19]. In this technique, a signal is considered as the output of an all-pole filter H(z). The inverse filter is represented by A(z). A(z) is detailed in Eq. 2 where $a_{1...M}$ represents the predictive coefficients up to the order M.

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + \cdots + a_M z^{-M} \tag{2}$$

The polynomial A(z) is decomposed into two polynomials P(z) and Q(z) presented in Eqs. 3 and 4, respectively, whose roots constitute the LSF representation. The zeros of these polynomials lie on the unit circle and are interlaced with one another thereby facilitating in computation.

$$P(z) = A(z) + z^{-(M+1)} A(z^{-1}) \tag{3}$$

$$Q(z) = A(z) - z^{-(M+1)} A(z^{-1}) \tag{4}$$

4.3 LSF-G Computation

Since LSF feature values were computed for each frame, thus disparate number of feature values were obtained for each clip because of disparity in the length of clips which is one of the primary characteristics of real-world scenario. Moreover, a clip of only 1 second (44.1 KHz) produces 440 frames according to Eq. 1. If only 10 LSFs are extracted for each frame, then a feature of 4400 (440×10) dimension is generated. In order to even the dimensionality and at the same time reduce the same, the LSF-G feature was engendered by grading the bands of the LSFs based on total energy content in ascending manner. This produced a feature of dimension equal to that of the order of the generated coefficients. In the aforesaid

case, the feature dimension is reduced from 4400 to only 10 by applying this technique, thereby reducing the computational burden. In the current experiment, 10, 12, and 15 dimensional LSF-G features were extracted for the clips and the highest accuracy was obtained for the 12 dimensional features. A detailed account of the results is presented Sect. 6.

5 Ensemble Learning Based Classification

Ensemble learning [20] is a technique in which multiple learners (base learners) are trained and combined to solve a problem. Here, the learners construct a set of hypothesis during the training phase and take a decision by combining them in contrast to the single hypothesis learning of various machine learning algorithms. The generalization capability of an Ensemble is much more than that of a base learner.

A Random Forest [21] based classifier was used in the current experiment which is an Ensemble learner composed of an array of decision trees. Random forests efficiently estimate missing data and maintain the accuracy when a large portion of data is missing as often observed in real-world cases. Moreover, Random Forests are fully parallelizable thereby making them an ideal choice for systems with parallel processing capability. Random Forest works by generating a set of decision trees ($T_{1...n}$) and for every such tree a random vector ($\Theta_{1...n}$) is generated which have the same distribution but Θ_k is independent of Θ_1 to Θ_{k-1}. The training set along with Θ_k aids in growing the kth tree thereby producing the classifier h (x, Θ_k), where x represents the input vector. The margin function for such an Ensemble of k classifiers with a randomly drawn training set from the random vector (Y) distribution is presented in Eq. 5.

$$mg(X, Y) = av_k I(h_k(X) = Y) - max_{j \neq Y} av_k I(h_k(X) = j), \tag{5}$$

where $h_k(X) =$ h (x, Θ_k) and I represents the indicator function. The margin measures the extent to which the average number of votes for the correct class exceeds the average of any other class. Greater value of this margin ensures higher confidence in classification. The generalization error is presented in Eq. 6 whose convergence is presented in Eq. 7 for almost all random vector sequences with an increase in the number of trees.

$$PE^* = P_{X,Y}(mg(X, Y) < 0) \tag{6}$$

$$P_{X,Y}(P_\Theta(h(X, \Theta) = Y) - max_{j \neq Y} P_\Theta(h(X, \Theta) = j) < 0) \tag{7}$$

6 Result and Discussion

A five fold cross validation technique was adopted in the experiment for evaluation. Each of the three feature sets was evaluated with the random forest based classifier

Table 2 Accuracy of various feature set dimensions for various number of iterations

Number of iterations	Accuracy (%) for feature set dimensions		
	15	12	10
1000	88.05	93.89	91.37
1500	88.12	**94.01**	91.22
2000	88.24	93.89	91.31
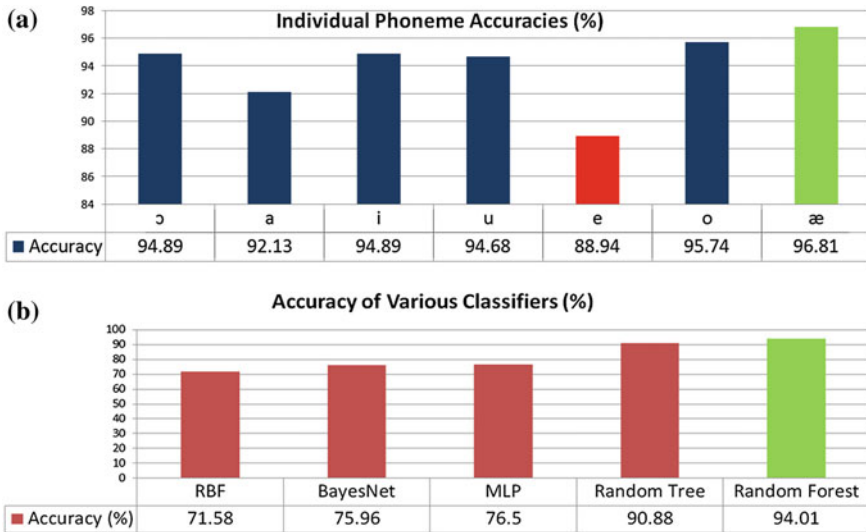
Fig. 3 **a** Individual Phoneme accuracies for ensemble learning. **b** Performance of other classification techniques

for 1000, 1500, and 2000 iterations, respectively, whose accuracies are presented in Table 2 with the highest accuracy denoted in Bold.

It can be observed from the Table that the highest accuracy was obtained for the 12 dimensional feature set at 1500 iterations and thus the result for this set is analyzed in detail. The individual accuracies of the Phonemes are presented in Fig. 3a, where the highest accuracy is highlighted in green and the lowest in red. The interference and similarity of spectral components of the various ambient noises with those of the clips could be a reason for the relatively lower accuracy of the Phoneme "e". The percentages of confusions among all the Phoneme pairs are presented in Table 3 with the highest confusions highlighted in green. It can be seen from the Table that the Phonemes (ɔ and a) constituted the most confused pair. These two Phonemes sound extremely close when pronounced at pace which may be a reason for such confusion because no restrictions were imposed on the volunteers at the time of data collection. A few other classifiers were also applied on this set whose results are presented in Fig. 3b.

Table 3 Percentages of confusions among the Phoneme pairs

	ɔ	a	i	u	e	o	æ
ɔ	-	5.11	0	0	0	0	0
a	5.11	-	2.77	0	0	0	0
i	0.43	3.19	-	1.06	0	0	0.43
u	0	0.21	1.49	-	3.62	0	0
e	0	1.91	0	5.11	-	4.04	0
o	0	0	0	0	2.55	-	1.7
æ	0	0	0	0	0.43	2.77	-

7 Conclusion

In this paper, a Bangla Swarabarna Phoneme recognition system has been presented using a newly proposed LSF-G feature coupled with Ensemble Learning. The system has been tested on an engendered database of 47 volunteers, and an encouraging accuracy has been obtained with a precision of 0.94. In future, we plan to experiment on a larger dataset of consonant Phonemes and also employ various pre processing and pre emphasis techniques to reduce the effect of environmental noise. The silent sections at the clip boundaries will also be identified and removed for further improvement of accuracy of the system. We also plan to experiment with other machine learning techniques and features in the future.

References

1. Ethnologue. http://www.ethnologue.com. Accessed 1 Apr 2017
2. Dudley, H.: The Vocoder. Bell Labs Rec. **17**, 122–126 (1939)
3. Dudley, H., Riesz, R.R., Watkins, S.A.: A synthetic speaker. J. Franklin Institute **227**, 739–764 (1939)
4. Forgie, J.W., Forgie, C.D.: Results obtained from a vowel recognition computer program. J. Acoust. Soc. Am. **31**, 1480–1489 (1959)
5. Desai, N., Dhameliya, K., Desai, V.: Feature extraction and classification techniques for speech recognition: a review. Int. J. Emerg. Technol. Adv. Eng. **3**(12), 367–371 (2013)
6. Microsoft Speech Recognizer. http://windows.microsoft.com/en-in/windows/set-speech-recognition#1TC=windows-7. Accessed 25 Apr 2017
7. Dragon Speech Recognizer. http://www.nuance.com/dragon/index.htm. Accessed 25 Apr 2017
8. Google Speech Recognizer. https://cloud.google.com/speech/. Accessed 25 Apr 2017
9. Besacier, L., Barnard, E., Karpov, A., Schultz, T.: Automatic speech recognition for under-resourced languages: a survey. Speech Commun. (2013)
10. Pramanik, M., Kido, K.: Bengali speech: formant structures of single vowels and initial vowels of words. In: Proceedings of ICASSP, vol. 1, pp. 178–181 (1976)
11. Hasnat, M.A., Mowla, J., Khan M.: Isolated and continuous Bangla speech recognition: implementation performance and application perspective. In: Proceedings of SNLP (2007)

12. Hasanat, A., Karim, M.R., Rahman, M.S., Iqbal, M.Z.: Recognition of spoken letters in Bangla. In: Proceedings of ICCIT (2002)
13. Ali, MdA, Hossain, M., Bhuiyan, M.N.: Automatic speech recognition technique for Bangla words. Int. J. Adv. Sci. Technol. **50**, 51–59 (2013)
14. Firoze, A., Arifin, M.S., Quadir, R., Rahman, R.M.: Bangla isolated word speech recognition. Proceedings of ICEIS **2**, 73–82 (2011)
15. Kotwal, M.R.A., Hossain, Md.S., Hassan, F., Muhammad, G., Huda, M.N., Rahman, C.M.: Bangla Phoneme recognition using hybrid features. In: Proceedings of ICECE (2010)
16. Hossain, K.K., Hossain, MdJ, Ferdousi, A., Khan, MdF: Comparative study of recognition tools as back-ends for Bangla Phoneme recognition. IJRCAR **2**(12), 36–40 (2014)
17. Audacity. http://www.audacityteam.org/. Accessed 25 Apr 2017
18. Paliwal, K.K., Alsteris., L.: Usefulness of phase spectrum in human speech perception. In: Interspeech, pp. 2117–2120 (2003)
19. Paliwal, K.K.: On the use of line spectral frequency parameters for speech recognition. Digit. Signal Process. **2**, 80–87 (1992)
20. Dietterich, T.G.: Ensemble Learning. The handbook of brain theory and neural networks, vol. 2, pp. 110–125 (2002)
21. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)

An Efficient Approach for Detecting Wormhole Attacks in AODV Routing Protocol

**Parag Kumar Guha Thakurta, Rajeswar Guin
and Subhansu Bandyopadhyay**

Abstract In MANET, wormhole link creates an illusion in such a way that two remote regions are directly connected through nodes which seems to be neighbors; however, these are actually distant from one another. The attackers using wormhole can easily manipulate the routing priority in AODV to perform eavesdropping, packet modification, or packet drop. Hence, a two-phase wormhole link detection procedure in AODV routing protocol is proposed to identify the malicious link for avoiding such erroneous transmission. Initially, the round trip time (RTT) and corresponding round trip bit transfer (RTBT) of each link are determined. Here, if RTBT of any link is greater than a dynamic threshold value ($RTBT_{TH}$), then such victim link is marked as a suspicious link. Next, the amount of power required to send a packet of certain size by each node is obtained to verify whether the transmission power of suspected links is reasonably high compared to other links. Various experimental results are carried out to validate the proposed work as well as to show an improvement obtained by the proposed approach in terms of several performance metrics.

1 Introduction

Mobile Ad Hoc Network (MANET) is an emerging wireless networking system in which mobile nodes are associated on an extemporaneous basis. It is self-forming and self-healing, enabling peer-level communications between the nodes without reliance on centralized resources or fixed infrastructure. The absence of both

P. K. G. Thakurta (✉) · R. Guin
NIT Durgapur, Durgapur 713209, India
e-mail: parag.nitdgp@gmail.com

R. Guin
e-mail: guinrajeswar@gmail.com

S. Bandyopadhyay
Brainware University, Kolkata, India
e-mail: subhansu@gmail.com

217

infrastructure and authorization establishes a key concern for the trusted and the non-trusted network [1]. In such adverse scenario, colluding nodes create an illusion that two remote regions of a MANET are directly connected through nodes which appear to be neighbors; however, these are actually distant from one another. The link between those malicious nodes is known as wormhole link, and corresponding nodes are termed as wormhole nodes [2]. Attackers using wormhole can easily manipulate the routing priority in MANET to perform eavesdropping, packet modification, or packet drop. Wormhole attack is normally launched in AODV (Ad hoc On-demand Distance Vector) [3] during the route discovery phase by creating the illusion of one-hop neighbors by wormhole peers. Therefore, the Route Request (RREQ) packets are routed through these wormhole tunnels to reach the destination in terms of low hop count as compared to usual normal path [4]. So, a wormhole attack is immensely harmful for a MANET. Once the wormhole is established, the adversary is able to conduct further attacks and do so with a low possibility of detection. This phenomenon motivates to propose an approach by which wormhole detection and its subsequent troublesome consequences can be avoided.

A procedure to detect the wormhole link is proposed in this paper to identify the malicious link for avoiding erroneous data transmission. The proposed work is executed in two consecutive phases. Initially, during route establishment state, the round trip time (RTT) of each link according to AODV routing protocol is determined. Next, the round trip bit transfer (RTBT) of those links depending on RTT value is computed. If RTBT of any link is greater than a dynamic threshold value $RTBT_{TH}$, then such victim link is marked as a suspicious link. In the later phase, the amount of power required to send a packet of certain size by each node is obtained using a standard power model. As it is known that the wormhole nodes are actually far apart from each other and transmission power (P) is proportional to the distance (d) between two nodes, i.e., $P \propto d$, so obviously transmission power for wormhole nodes is high. Hence, a detection procedure in this phase is executed to verify whether the transmission power of suspected link's obtained in the first phase is reasonably higher compared to other links. The experimental results are carried out to validate the proposed work. It shows the legitimate packet delivery obtained by the proposed method within valid end-to-end delay.

This paper is organized as follows: Sect. 2 discusses related studies for completeness of the proposed work. Section 3 presents the system model. The problem addressed in this paper is described in Sect. 4. The proposed approach is discussed in Sect. 5. Next, various experimental results, in Sect. 6, are shown in connection with the proposed work. Section 7 concludes the work.

2   Related Works

Various strategies for detecting wormhole attacks in mobile networks have been studied till date, for ad hoc topologies. Z Tun et al. present an approach to detect wormhole attack [5]. This paper analyzes the nature of wormhole attack and

proposes RTT-based detection mechanism. Another approach in [6], introduced by Virendra Dani et al., locates the wormhole link in network and tries to recover the performance during the attack conditions. However, Bharti Patidar et al. algorithm [7] needs one extra step to bring the work overload for increasing network lifetime. In [8], an effective method called Wormhole Attack Prevention (WAP) is developed not only for detecting the fake route but also for adopting preventive measures against wormhole nodes during the route discovery phase. The Delay per Hop Indicator (DelPHI) [9] can detect both hidden and exposed wormhole attacks. A new mechanism for detecting wormhole attack allows each device to monitor its neighbors' behavior [10]. Another algorithm called Neighbor-Probe-Acknowledge (NPA) [4] is proposed to detect wormhole attacks on a real wireless mesh network using standard deviation of RTT.

In short, various procedures on wormhole attack detection in routing have been identified for wireless ad hoc networks and the corresponding countermeasures have been addressed in the literature studies. Most of these works are based on RTT. In such scenario, if RTT is high, the wormhole is suspected. However, those methods can lead to false detection of wormhole attacks as RTT can be higher when the route is congested. Therefore, a new metric, known as round trip bit transfer (RTBT), is proposed herein such that false suspect due to congestion can be avoided. A two-phase wormhole link detection procedure in AODV routing protocol is proposed therefore in next to identify the malicious link for avoiding such inaccurate transmission.

3 System Model

This section presents different models and various performance metrics related to the work proposed in this paper. These are as follows:

3.1 Network Model

Let us assume a multi-hop ad hoc network where nodes collaboratively relay packets according to an underlying AODV routing protocol. The network consists of a set of N nodes. Any path P_{SD} is used to route the packet from a source S to a destination D. It is assumed that the S and D collaboratively monitor the performance of P_{SD}. The destination periodically reports to the source about the critical metrics related to routing. If a misbehaving node drops the periodic updates as part of its misbehavior pattern, the source interprets the lack of updates as an occurrence of misbehavior. The presence of such a misbehaving node within the network can be known by checking some performance measurement.

3.2 Adversial Model

In perspective of maliciousness, one or more than one pair of nodes deployed in the network are assumed to be misbehaving. However, wormhole establishment is possible through a single long-range wireless link. The attacker may create tunnel even for the packets not addressed itself as of broadcasting nature of the radio channel. In order to detect such malicious links, the following metrics are defined.

- **RTT**: It is defined as the time difference of route request (RREQ) and route reply (RREP) packets arrival for a node. Suppose RREQ arrives at a node N_i at time X, and the receiving time of RREP packet for that node is Y, then RTT is calculated as follows:

$$RTT_{N_i} = Y - X - RTT_{previous}, \tag{1}$$

where $RTT_{previous}$ denotes the RTT value of the previous node of N_i within P_{SD}.

- **RTBT**: During the time of detection for wormhole link, the parameter RTBT for a node is defined as

$$RTBT \text{ of a node} = \frac{RREQ \text{ packetbytes} + RREP \text{ packetbytes}}{RTT \text{ of that node}} \tag{2}$$

In (2), $RREQ_{packetbytes}$ and $RREQ_{packetbytes}$ denote the size of those packets, respectively. Further, the threshold value of RTBT ($RTBT_{TH}$) is obtained as maximum RTBT value for the nodes in the half of P_{SD}. Experimentally, it is observed that the value of $RTBT_{TH}$ provides more correct results for dividing maximum RTBT value from a set of nodes in P_{SD} by 2 rather than any other possible values.

3.3 Power Consumption Model

A power model is introduced to calculate the amount of power required to transmit a packet of size 'B' bytes from a node to another which is d meter away from it. In order to obtain such power consumption, it is required to find minimum energy [2] for a node to transmit data to others. It is obtained by the following.

$$E_{min} = K_1(D)d^4 + K_2 \qquad [K_1, K_2 \text{ are constants}] \tag{3}$$

Using E_{min}, the transmitted power [2] can be obtained as follows:

$$P_{TX}(m^{'}, d) = \frac{(m^{'} E_{min} + \varepsilon d^n)}{1000} \qquad (4)$$

where $m^{'}$, ε denote data rate of the node and permittivity of the air as medium, respectively. In addition, it is important to mention here that P_{TX} is estimated in WATT. Further, the threshold value of P_{TX} denoted by P_{TH} is obtained depending on the normal transmission range for a node within the network.

4 Problem Description

Usually, AODV uses the hop count to determine the shortest path between sender and receiver. A malicious node can set the false hop count as well as the value of route sequence number. An attacker can tunnel a request packet RREQ directly to the destination node without increasing the hop count value. Thus, it prevents any other routes from being discovered. As a result, the destination node discards all subsequent RREQ packets received and selects the false wormhole tunnel infected route to send the RREP. It leads to a wormhole attack in AODV during data transfer phase. Under such scenario, the work proposed in this paper detects such attack by calculating RTBT of each node in the first phase and subsequently the transmission power of the corresponding node in next. Therefore, such problem can be realized as follows:

Objectives:

$$\begin{cases} \min(\text{RTBT}) \, \forall N \ \in P_{SD} \\ \max(P_{TX}) \, \forall N \ \in P_{SD} \end{cases} \qquad (5)$$

In (5), the node with minimum RTBT value is obtained as a suspected wormhole node, and subsequently the same node is validated as wormhole by the requirement of maximum power consumption for data transmission for S to D in P_{SD}.

5 Proposed Approach

The approach proposed in this work mainly involves two phases for wormhole node detection. Along with these phases, a network deployment is used at the beginning to generate neighbors list of each node according to AODV routing protocol. Next, the malicious node detection phase begins with the determination of the value of RTBT between a pair of nodes N_i and N_j in P_{SD}. Here, the minimum of such RTBT value is selected as wormhole node, because the RREQ and RREP take more time for traveling as compared to others. However, a common link may be shared by

multiple nodes in wireless networks. So, due to congestion, more time is needed to travel a packet using that common link. In such scenario, RTBT would become also low for a legitimate link. To avoid such conflictness, a new phase is introduced depending on the computation of transmission power consumption $P_{TX_{N_i, N_j}} \forall N_i, N_j \in P_{SD}$ for sending a certain size packet from N_i to N_j. By sensing the power value, it can determine the actual distance between a pair of nodes within P_{SD} as it is known that $P_{TX} \propto d^{\alpha}$, where "$\alpha$" denotes path loss factor. So min (RTBT) and max (P_{TX}) are used to make a decision about wormhole link. This procedure to detect wormhole attack is described by the following algorithms.

Algorithm: Phase—1

1. AODV find a route P_{SD} from source node S to destination node D
2. Store the size of both RREQ and RREP packets in one variable Tot_size (Tot_size = size of RREQ + size of RREP).
3. Initialize variable Pre_RTT = 0.
4. Count the number of nodes in P_{SD}.
5. Starting from immediate previous node of D and end with S where N_i \in P_{SD}

   5.1 RTT_{Ni} = The time when N_i get RREP packet – The time when N_i getting RREQ packet – RTT value $(RTT_{N(i+1)})$
   5.2 $RTBT_{Ni}$ = Tot_size / RTT_{Ni}
   5.3 Pre_RTT = RTT_{Ni}

6. Now calculate the threshold value $RTBT_{TH}$
7. If $RTBT_{Ni} < RTBT_{TH}$

   Then, link in between N_i and $N_{(i+1)}$ marked as a suspicious link.
   Else
   The link between N_i and $N_{(i+1)}$ is valid

Algorithm: Phase—2
Calculate a threshold value of P_{TH} which is the maximum power limit for a node required to transmit a packet to its next node.

1. Starting from sender node S to the immediate previous node of D

   1.1 Calculate amount of power needed to send a B bytes packet from N_i to $N_{(i+1)}$.
   1.2 Check if link in between N_i and $N_{(i+1)}$ was previously suspected

      1.2.1 Now check if N_i takes transmission power which is more than P_{TH}.
             //The link $N_i \rightarrow N_{(i+1)}$ marked as wormhole link.
      1.2.2 Else //The link is legitimate link.

   1.3 Else
   // the link is legitimate link.

Time Complexity: The overall time complexity of the proposed work is the sum of the time complexities of AODV routing protocol, the first phase, and the second phase of the proposed algorithms. It is known that AODV routing is executed in O

(2d) time, where "d" denotes the network diameter. In the first phase of proposed algorithm, RTBT for each node of P_{SD} is computed in O(n), where "n" is the number of nodes present in the network. The second phase of proposed algorithm is executed in O(n) time. So, the overall time complexity of the proposed work is O (d + n).

6 Experimental Results

Various experiments are carried out to show the effectiveness of the proposed work. The simulation set up is given next.

6.1 Simulation Setup

The proposed wormhole detection approach in AODV routing protocol is simulated using network simulator ns2 [3]. A network topology of 16 nodes with constant bit rate (CBR) traffic pattern is adopted. Up to four wormhole tunnels (eight wormhole peers) are used for experiments. Various simulation parameters used in different experiments are shown in Table 1.

Table 1 Simulation parameters

Parameter	Value
Area	1440 m × 1200 m
Simulation time	10 s
Number of nodes	16
Traffic model	CBR
Number of wormhole tunnels	1/2/3/4 (up to 8 wormhole peers maximum)
Mac protocol	MAC/802.11
Data rate	256 Kbps
Data packets	1024 Bytes/packet
Routing protocol	AODV
Source type	UDP
Channel type	Channel/wireless channel
Radio propagation model	Propagation/two-ray round wave
Interface queue type	Queue/drop tail
Link layer type	LL
Antenna	Antenna/omni antenna
Maximum packet in ifq	150

6.2 Simulation Studies

The following metrics are defined next for completeness of the subsequent discussions.

- **Packet delivery ratio (PDR)**: It is defined as percentage of the amount of received packets by D to the number of packets sent by S. It is expressed as follows:

$$PDR(\%) = (\text{Number of packet received by} D \times 100)/\text{Number of packet send by } S \tag{6}$$

- **Drop Rate**: It is defined as the number of packets dropped by nodes.

$$DropRate(\%) = ((\text{Number of Packet send } - \text{ Number of Packet Received}) \times 100)/ \\ (\text{Number of Packet send}) \tag{7}$$

- **End-to-end delay**: The average time taken by a data packet to arrive at the destination. It also includes the delay caused by route discovery process and the queue in data packet transmission. Only the data packets that successfully delivered to destinations are counted. Therefore,

$$End-to-end\ Delay = \left(\sum(\text{arrive time} - \text{send time})\right)/\left(\sum \text{Number of connections}\right) \tag{8}$$

- **Throughput**: Throughput is defined as the number of packets received at the destination over a period of time and is measured in kbps unit.

$$Throughput = (\text{Total bytes received by destination node})/(\text{Total simulation time}) \tag{9}$$

Now, it is shown in Fig. 1 that PDR decreases drastically with increase in wormhole links as more wormhole peers perform selective packet dropping. It shows an improvement in PDR obtained by the proposed approach compared to

Fig. 1 Packet delivery ratio versus number of wormhole links
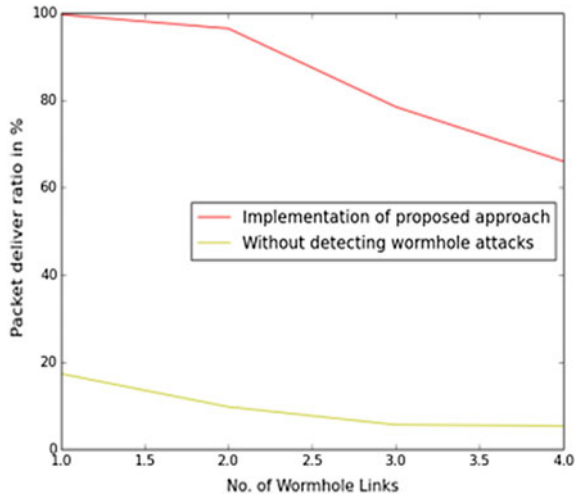
Fig. 2 Packet drop ratio versus number of wormhole links
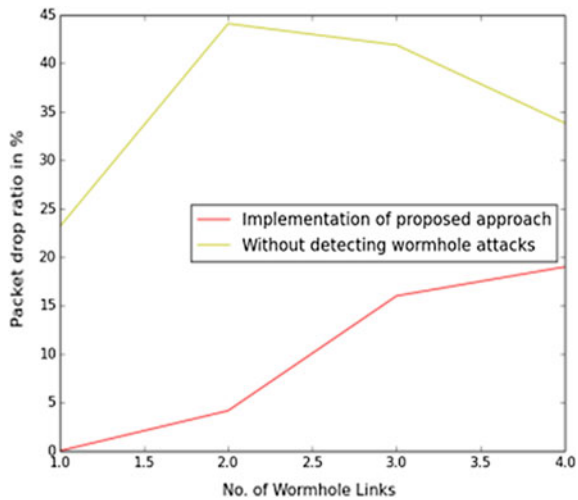
wormhole-infected AODV. The lower value of the packet loss means the better performance of the protocol. Drop rate increases steadily with increasing wormhole links in AODV. As observed in Fig. 2, packet drop rate is reduced by applying proposed approach compared to wormhole-infected AODV.

As shown in Fig. 3, average end-to-end delay increases drastically when number of wormhole links are increased as the link latency is higher for wormhole tunnels leading to more time consumption. In Fig. 4, it is shown that the throughput is generally kept low when wormhole nodes are active. The throughput is increased when proposed algorithm is applied.

Fig. 3 End-to-end delay versus number of wormhole links
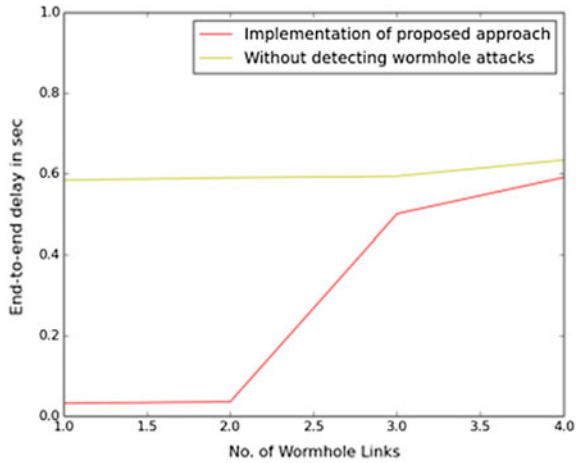
Fig. 4 Throughput versus number of wormhole links
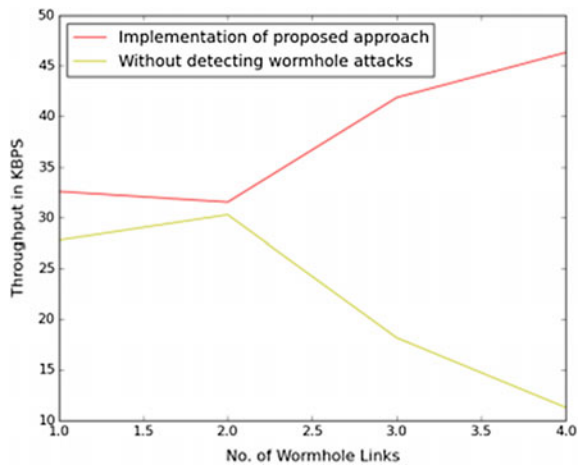
7 Conclusions

An approach toward identifying wormhole links in AODV routing is proposed in this paper to avoid errorneous data transfer. The wormhole link is suspected using the round trip bit transfer method. Transmission power is used to obtain the distance between two immediate nodes, which helps us to confirm the existence of wormhole links along the route. Our proposed approach is not susceptible to false detection of wormhole links. The simulation studies show the effectiveness of the proposed work in terms of several network parameters. A possible future scope of this work is to apply the proposed method into other available routing protocols in MANET to make a concrete comparative study.

References

1. Touzene, A., Al-Yahyai, I.: Performance analysis of grid based AODV routing algorithm for ad hoc wireless networks. Int. J. Commun. Netw. Syst. Sci. **8**, 523–532 (2015)
2. Zhou, J., Cao, J., Zhang, J., Zhang, C., Yu, Y.: Analysis and countermeasure for wormhole attacks in wireless mesh networks on a real testbed. In: 26th IEEE International Conference on Advanced Information Networking and Applications, Japan (2012)
3. Choi, S., Kim, D.Y., Lee, D.H., Jung, J.I.: WAP: wormhole attack prevention algorithm in mobile ad hoc networks. In: IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, (SUTC'08), pp. 343–348, Taiwan (2008)
4. Tun, Z., Thein, N.L.: Round trip time based wormhole attacks detection. In: IEEE Wireless Communications and Networking Conference, pp. 1–6, Las Vegas, USA (2008)
5. Dani, V., Birchha, V.: An improved wormhole attack detection and prevention method for wireless mesh networks. Int. J. Adv. Res. Comput. Commun. Eng. **4**(2), 480–485 (2015)
6. Kodali, R.K., Sarma, N.N.: Low energy routing for WSN's. ACEEE Int. J. Netw. Secur. **5**(4), 4–10 (2014)
7. Misra, R., Manda, C.R.: Performance comparison of AODV/DSR on-demand routing protocols for ad hoc networks in constrained situation. In: International Conference on Personal Wireless Communications, New Delhi, India (2005)
8. Patidar, B., Shivhare, R.R.: Detection of wormhole attack using round trip time and node workload. Int. J. Comput. Appl. **112**(11), 1–4 (2015)
9. Jhaveri, R.H., Patel, A.D., Parmar, J.D., Bhavin, I.: MANET routing protocols and wormhole attack against AODV. Int. J. Comput. Sci. Netw. Secur. **10**(4), 12–18 (2010)
10. Sharma, G., Fatima, M.: An energy efficient approach for wormhole detection and prevention. Int. J. Comput. Appl. **76**(17), 38–42 (2013)

Efficient Contrast Enhancement Based on Local–Global Image Statistics and Multiscale Morphological Filtering

Gunjan Gautam and Susanta Mukhopadhyay

Abstract In this paper, image contrast enhancement is achieved by combining together the local–global image statistics and multiscale morphological filtering (MMF). The proposed method has been executed on two different sets of images, and the result has been compared with that of some existing standard methods, namely histogram equalization (HE), contrast limited adaptive histogram equalization (CLAHE), and multiscale morphology in order to have an outlook on the relative performances. The experimental results manifest that the proposed method produced results superior to the methods compared.

Keywords Contrast enhancement · Multiscale mathematical morphology
Histogram equalization · Adaptive histogram equalization

1 Introduction

Contrast for a digital image is the difference between the higher and lower valued pixel that makes an object present in the image distinguishable. The bigger this difference is, the higher is the image contrast. Contrast enhancement techniques have been employed extensively in image processing and among them, histogram equalization (HE) and adaptive histogram equalization (AHE) [1] are two frequently used methods. As the intensities and other characteristics vary across the entire image, HE being a global natured technique turns out to be less efficacious. Other improved variants of HE [2, 3] proposed in literature are based on division of histograms and hence fail sometimes, specifically when the division is not accurate. To the contrary,

G. Gautam (✉) · S. Mukhopadhyay
Department of Computer Science and Engineering, Indian Institute of Technology (ISM),
Dhanbad, India
e-mail: gautamgunjan29@gmail.com

AHE and its variants (e.g., interpolated AHE, weighted AHE, clipped AHE, etc.) are devised to get over the issues of HE by dealing with the local features of the image and have demonstrated success. Basic form of AHE is about mapping each pixel to an intensity proportional to its rank in the neighborhood pixels. Two major issues associated with AHE are over-enhancement of noise and slower speed [4]. Subsequently, several improvements in basic AHE have been reported till now [4] and CLAHE is one of its variants which limits the noise amplification.

Mathematical morphology is a suitable tool to handle spatial features and the concept of multiscaling enhances its potential to extract the geometry- and size-based features [5]. In this paper, a linear mapping (relying on local and trimmed global mean) has been applied prior to applying multiscale morphological techniques, which results in amplification of graylevel $I(x, y)$ (see Eq. 6). Final results are compared concerning the visual appearance and graylevel co-occurrence matrices (GLCM) [6] based contrast estimation.

The remainder of this paper is organized as follows: Sect. 2 briefly describes some standard techniques for contrast enhancement that we have used for comparative analysis. In Sect. 3, we present the proposed approach in detail. Sections 4 and 5 show the experimental results and conclusions, respectively.

2 Background

Histogram Equalization modifies the dynamic range and contrast of an image by altering the intensity histogram in a desired shape. One important thing to be noted here is that during histogram equalization the overall shape of the histogram should not change. The key idea of HE is to remap the intensity values of the input image into new intensity levels using a transform function created from cumulative density function (cdf) [7]. HE does not ensure that contrast will always be increased. Furthermore, there may be some cases where the result of employing HE can be inadequate.

Contrast Limited Adaptive Histogram Equalization fortifies visibility of the hidden features in an image by functioning on small contextual regions. Each region is then enhanced individually using HE and while combining these neighboring enhanced regions, bilinear interpolation is used to reduce the blocking artifacts [1]. CLAHE clips a portion of the histogram above a value called as clip limit and redistributes them to each histogram bin. HE operates uniformly across the images and therefore is unable to deal with the local contrast. CLAHE, being an upgraded and refined form of AHE [1, 8], surmounts this issue. In this work, CLAHE has been used with three parameters: block size, histogram bins, and clip limit with the values 8×8, 256, and 0.01, respectively.

Mathematical Morphology [9, 10] is a set theory based powerful technique in image processing to identify objects and their features depending upon their size and geometry. *Dilation* ($\oplus$) and *Erosion* ($\ominus$) are two fundamental operations which function by probing an image at each pixel with a proper structuring element (SE).

Multiscale morphology [9, 11–13] is about employing SEs of different scales to extract the scale-specific features. Other two essential operations derived from the fundamental ones are Opening (∘) and Closing (•). In multiscale scenario, both are defined as

$$(A \circ nB)(x, y) = ((A \ominus nB) \oplus nB)(x, y) \tag{1}$$

$$(A \bullet nB)(x, y) = ((A \oplus nB) \ominus nB)(x, y), \tag{2}$$

where A is the grayscale image, B is the SE, and n is an integer, representing the scale of the SE. Tophat transform is an outstanding morphological tool to extract small details (which are smaller than the SE) from an image. The twin classes of tophat transform, *white* tophat and *black* tophat, facilitate feature extraction through acquiring the brighter and the darker features, respectively. Mathematically, *white* tophat transform (T_{TH}) is the difference between the image and its opening by an appropriate SE and its dual, i.e., *black* tophat transform (T_{BH}) is the difference between the closing and the image (see Eq. 3).

$$T_{TH}(A) = A - (A \circ B) \quad \text{and} \quad T_{BH}(A) = (A \bullet B) - A \tag{3}$$

3 Proposed Method

This section describes the proposed approach in detail. Some popular contrast enhancement methods along with their limitations have been discussed in the previous section; however, the scheme presented here conquers these issues and provides a substantially better solution. Figure 1 demonstrates the schematic representation of the proposed method consisting of the following operational stages:

- *Remapping each pixel of input image* based on local–global information
- *Multiscale morphological filtering* using white and black tophat transform.

First of all, the input RGB image is converted into grayscale based on a weighted sum of the R, G, and B components (belonging to each channel, i.e., red, green, and blue) as shown in Eq. 4.

$$0.2989 * R + 0.5870 * G + 0.1140 * B \tag{4}$$

In order to gather the global information, trimmed global mean of the entire image is computed. Trimmed mean helps in reducing the effects of statistical outlier (e.g., noise) bias and so is achieved by removing a certain percentage (8% in our experiments) of the largest and smallest values from the set before evaluating its standard mean. Usually, it is the better representation of the center of the data than the mean, when there are outliers present. If S is a set with x number of total elements, from
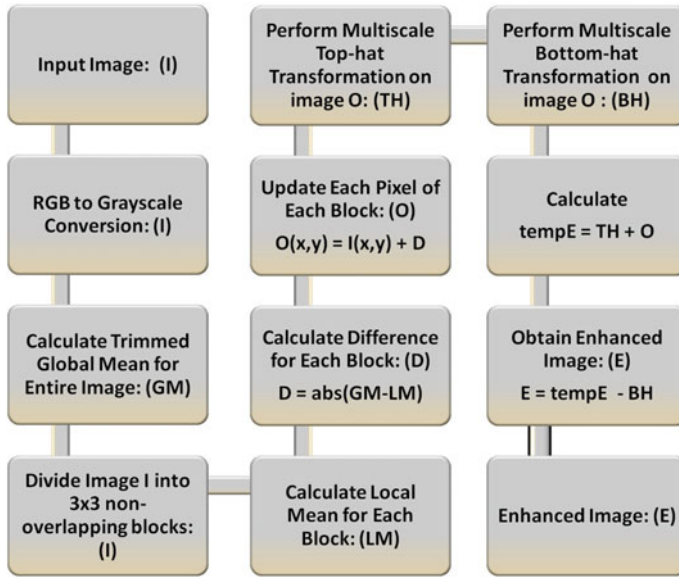
Fig. 1 Block diagram of proposed method

which $p\%$ is to be trimmed out, then the trimmed mean is the mean of S after eliminating n elements from the lower and upper bounds. n can be calculated from Eq. 5

$$n = (x \times (p/100)/2) \qquad (5)$$

Subsequently, the input image undergoes a division of 3×3 nonoverlapping blocks and local mean for each one is calculated. Afterward, the difference between both the means is added to the corresponding pixel's intensity value of the input image and a comparatively better image $O(x, y)$ as an intermediate result of proposed pipeline is achieved (see Eq. 6 and Fig. 1).

$$O(x, y) = I(x, y) + |GM - LM|, \qquad (6)$$

where GM and LM denote the trimmed global mean and local mean of a predefined neighborhood, respectively. Lastly, this better version $O(x, y)$ goes through the multiscale morphological filtering to obtain the final enhanced image. In multiscale morphology, the geometry and size of the structuring element play a vital role to accomplish appropriate results. Typically, both parameters are selected empirically depending on the application. Being rotation invariant, disk-shaped structuring elements are used frequently and therefore the same with increasing radii 1, 2, 3, and 4 are utilized here to perform the multiscale tophat transform. Multiscale white tophat transformation (T_{TH}) is carried out using Eq. 3 in which all the resultant images that have been processed separately through an ordered sequence of T_{TH} based on each

SE are summed up together. Likewise, multiscale black tophat transformation (T_{BH}) has also been executed using Eq. 3. After adding the outcome of multiscale white tophat transform $TH(x, y)$ to $O(x, y)$, final enhanced image $E(x, y)$ is obtained by subtracting the multiscale black tophat transformation outcome $BH(x, y)$ from this result (see Eq. 7).

$$E(x, y) = (O(x, y) + TH(x, y)) - BH(x, y) \qquad (7)$$

4 Experimental Results and Analysis

The proposed scheme has been tested on two sets of biometric images taken from CASIA iris image database [14] and NIST fingerprint database [15], and the results have been compared (see Figs. 3, 4, 5 and 6) with a few standard techniques as described in Sect. 2. 110 images of visually poor contrast from each database have been selected for the experimentation. All the experiments have been performed using MATLAB 9.0 R2016a. We have made the subjective as well as objective evaluation to validate the strength and superiority of the proposed method. Subjective evaluation based on visual quality is depicted in Figs. 5 and 6. Additionally, outcome of the objective evaluation in terms of graphical relationship is presented in Figs. 3 and 4 for fingerprint and iris images, respectively. Here, objective evaluation is grounded on estimating contrast derived from the graylevel co-occurrence matrix (GLCM) [6]. GLCM is created from a scaled version of an image depending upon the number of graylevels.

GLCM describes how often pairs of pixels with particular values and in a specified spatial relationship (or offset) occur in an image. In our GLCM, the final occurrence for a pixel pair (i, j) in the scaled image (with eight graylevels) is the average occurrence of this pair at 1–5 pixel distances in four different directions (i.e., $0°, 45°, 90°$, and $135°$). Also, (i, j) and (j, i) are considered to be two different combinations. Hence, the GLCM size is $8 \times 8 \times 4$ for our test images. Figure 2 is

Fig. 2 Illustration of GLCM offset at the distance of 3 pixel
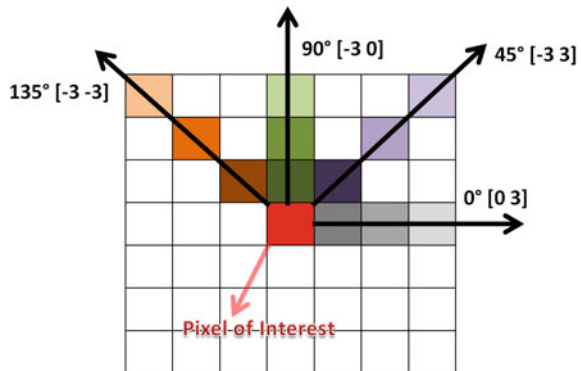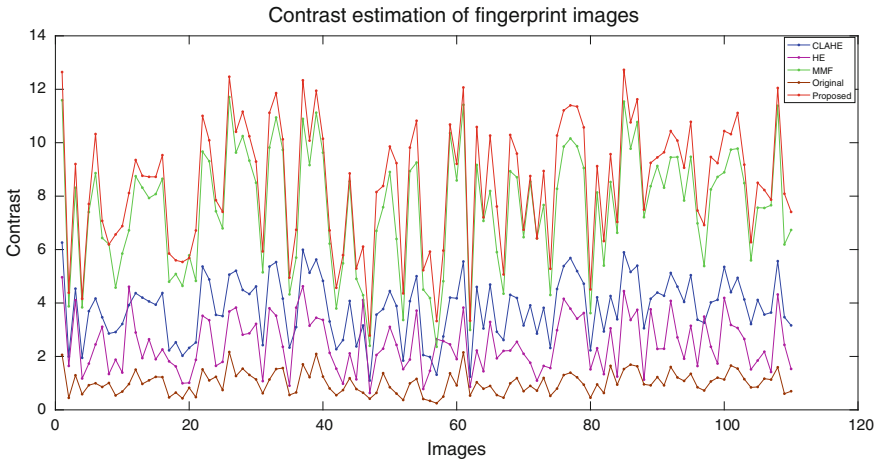
Fig. 3 Quantitative results for 110 fingerprint images based on GLCM
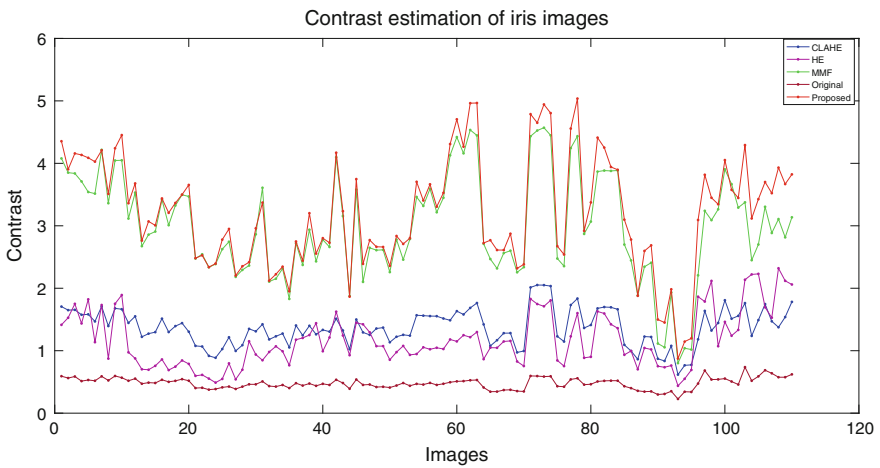
Fig. 4 Quantitative results for 110 iris images based on GLCM

a vivid depiction of how offset of GLCM has been determined. Consequently, the final statistics of comparative results are shown in Table 1 and it is found that the contrast of the images processed by proposed method is higher than that of images
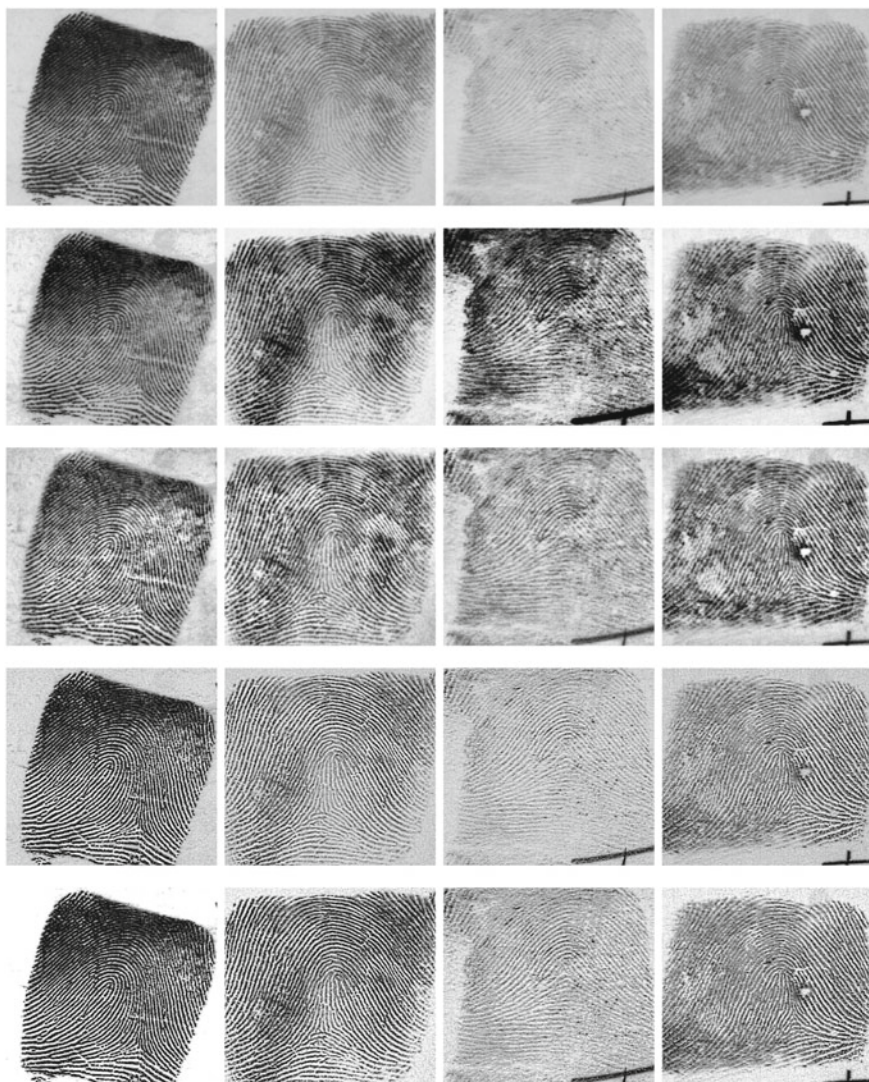
Fig. 5 Enhancement results: First row: Original images. Second row: Enhanced using HE. Third row: Enhanced using CLAHE. Fourth row: Enhanced using multiscale morphological filtering. Fifth row: Enhanced using proposed method
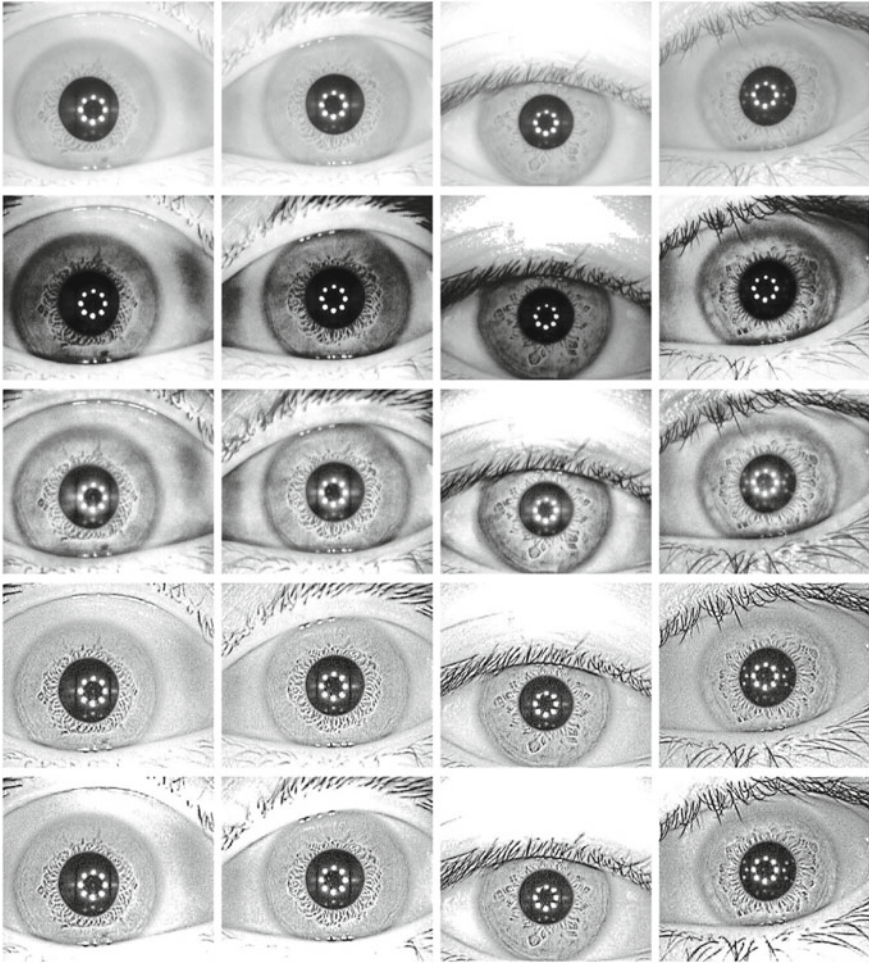
Fig. 6 Enhancement results: First row: Original images. Second row: Enhanced using HE. Third row: Enhanced using CLAHE. Fourth row: Enhanced using multiscale morphological filtering. Fifth row: Enhanced using proposed method

processed by other three methods. Texture region of iris and ridge patterns of finger-print images has also been enhanced significantly. Besides being straightforward, our approach has the potential to effectively handle the poor quality images in an acceptable computation time.

Table 1 Average contrast derived from GLCM for 110 fingerprint and 110 iris images

Methods	Fingerprint/Iris					
	1 Pixel	2 Pixel	3 Pixel	4 Pixel	5 Pixel	Mean contrast
Original	0.3476/0.20817	0.74208/0.36036	1.1056/0.49246	1.3663/0.60939	1.5178/0.7165	1.0159/0.4774
HE	0.76243/0.5228	1.8128/0.93407	2.7425/1.2384	3.3676/1.482	3.6889/1.6971	2.4749/1.1749
CLAHE	1.1242/0.62756	2.7989/1.1029	4.2847/1.4463	5.26/1.7231	5.7277/1.9756	3.8391/1.3751
MMF	2.9356/1.927	6.2968/2.867	8.621/3.2648	9.8094/3.4049	10.1167/3.5277	7.5559/2.9983
Proposed	**3.3318/1.9593**	**6.9912/3.0023**	**9.5866/3.5085**	**10.9959/3.7199**	**11.4291/3.8812**	**8.4669/3.2142**

5 Concluding Remarks

This paper presented an efficient algorithm for contrast enhancement which suitably enhances the image along with the persistence of major information and natural look. The results have been compared with few standard methods and according to the quantitative results, the proposed one is found to be much satisfactory. Its global and simple behavior makes it effective. Though this method is not the fastest, its performance surpasses the performance of the other methods in comparison.

References

1. Zuiderveld, K.: Contrast limited adaptive histogram equalization. In: Graphics Gems IV, pp. 474–485. Academic Press Professional, Inc. (1994)
2. Wang, Y., Chen, Q., Zhang, B.: Image enhancement based on equal area dualistic sub-image histogram equalization method. IEEE Trans. Consum. Electron. **45**(1), 68–75 (1999)
3. Kim, M., Chung, M.G.: Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement. IEEE Trans. Consum. Electron. **54**(3) (2008)
4. Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J.B., Zuiderveld, K.: Adaptive histogram equalization and its variations. Comput. Vis. Graph. Image Process. **39**(3), 355–368 (1987)
5. Maragos, P.: Pattern spectrum and multiscale shape representation. IEEE Trans. Pattern Anal. Mach. Intell. **11**(7), 701–716 (1989)
6. Haralick, R.M., Shanmugam, K., et al.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
7. Ooi, C.H., Isa, N.A.M.: Quadrants dynamic histogram equalization for contrast enhancement. IEEE Trans. Consum. Electron. **56**(4) (2010)
8. Reza, A.M.: Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. J. VLSI Signal Process. **38**(1), 35–44 (2004)
9. Serra, J.: Image Analysis and Mathematical Morphology, vol. 1. Academic press (1982)
10. Lu, H., Li, Y., Zhang, L., Serikawa, S.: Contrast enhancement for images in turbid water. JOSA A **32**(5), 886–893 (2015)
11. Mukhopadhyay, S., Chanda, B.: A multiscale morphological approach to local contrast enhancement. Signal Process. **80**(4), 685–696 (2000)
12. Bai, X., Zhou, F.: A unified form of multi-scale top-hat transform based algorithms for image processing. Opt.-Int. J. Light Electron Opt. **124**(13), 1614–1619 (2013)
13. Das, D., Mukhopadhyay, S., Praveen, S.S.: Multi-scale contrast enhancement of oriented features in 2d images using directional morphology. Opt. Laser Technol. **87**, 51–63 (2017)
14. CASIA: Biometrics ideal test, casia.v4 database. http://www.idealtest.org/dbDetailForUser.do?id=4
15. NIST: Standard Reference Data, Fingerprint database 4. https://srdata.nist.gov/gateway/gateway?keyword=fingerprint

Bag-of-Tasks Intelligent Scheduling Agent (BISA) in Cloud Computing

Preethi S. H. Darius and E. Grace Mary Kanaga

Abstract Cloud computing offers to its users, in theory, infinite computing through utility computing. Scheduling tasks in heterogeneous resources poses a formidable challenge albeit an increase in available computing capacity. A class of tasks termed as Bag-of-Tasks (BoT) is a predominant workload in any large-scale distributed system. We propose an agent-based approach, BISA (BoT Intelligent Scheduling agent), which chooses an appropriate scheduling heuristic and in time learns which policy will provide the most optimal schedule. The agent-based framework is presented and the cloud environment is simulated in CloudSim using threads. The working of the BIS agent is presented, and results from the training phase are divulged and discussed. The results obtained show the framework presented could provide near-optimal solution in minimizing the makespan of a BoT using the most appropriate scheduling heuristic in a given scenario.

Keywords Bag-of-Tasks application · Cloud computing · Intelligent agents

1 Introduction

As the magnitude of computing resources increases, so is the need to efficiently provision them without which the resources will either be underutilized or over-provisioned. One of the challenges in cloud computing is "performance unpredictability" [1]. A major contributing obstacle is the unpredictability of the scheduling algorithms used on some classes of batch processing programs where many of the tasks with parallelism are run in small clusters, which are poorly utilized. This type of task has been categorized as Bag-of-Tasks (BoT).

P. S. H. Darius (✉)
MVJ College of Engineering, Bangalore, India
e-mail: preethi.hepsiba@gmail.com

P. S. H. Darius · E. Grace Mary Kanaga
Karunya Institute of Technology and Sciences, Coimbatore, India
e-mail: grace@karunya.edu

From the analysis of grid workloads in the period of 2003–2010, Iosup and Epema stated that BoT submissions account for over 75% of the tasks and 90% of the overall CPU consumption in grid workload [5]. Both grids and clouds are essentially large-scale distributed systems, and the type of workload that dominates the cloud is also BoTs. The scheduling of independent tasks (bag-of-tasks) that have to be scheduled in heterogeneous systems is known to be an NP-complete problem [4].

Meta-heuristics may sometimes provide near-optimal schedules but the time complexity for producing the schedule may be extremely high when the magnitude of the system increases which is the case in cloud.

One of the most important observations made by Garcia and Sim [3] is this: Due to the *NP-complete* nature of the scheduling problem, there was not a dominant scheduling heuristic for all the BoTs. We hypothesize there is no dominant scheduling heuristic that performs the best in all situations due to the NP-complete nature of the problem.

Our contributions for scheduling bag-of-tasks using agents are as follows:

- To outline a framework of an intelligent agent (BIS) that uses concepts of learning and reinforcement which eventually will use the most efficient type of scheduling heuristic/meta-heuristic to generate a schedule that has the shortest makespan for each BoT.
- To demonstrate the experimental setup in CloudSim to test our hypothesis and study the outcome.
- To present the results of the training phase of the agents that provides a rationale for continuing research in this direction.

2 BIS Agent Framework

A Cloud Controller Agent (CCA) oversees all the agents and initializes the Broker Agent (BA), the Service Provider Agent (SPA), Resource Agent (RA), User Agent (UA), and BoT Intelligent Scheduler Agent (BISA). The agent-based framework adapted from Sim [10] is presented in Fig. 1.

Upon initialization, the SPA creates a datacenter for the private/public cloud in CloudSim and interacts with the RA which is assigned to each host in the datacenter and it is responsible for choosing the allocation policy (timeshared/spaceshared) for the virtual machines (VMs) and the instantiation and allocation of VMs to hosts. UAs are created for each user who wants to submit tasks.

The model for BISA

The BISA (BoT Intelligent Scheduler Agent) is an intelligent agent that senses the environment and based on the current state chooses a scheduling policy and produces a schedule. The intelligent agent we have modeled is adapted from structure of intelligent agent proposed by Russel and Norvig [9].
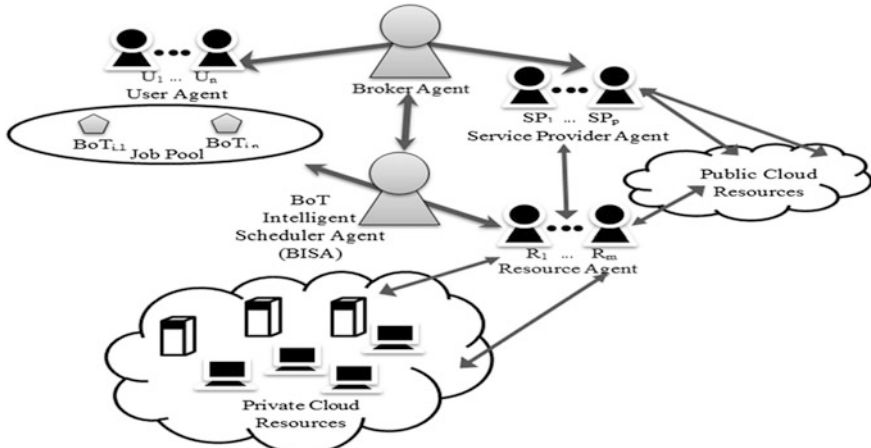
Fig. 1 Agent-based framework

The working of the BIS agent in each stage is outlined below:

Let S be the set of states that the agent has come across where $S = \{s_1, s_2, \dots s_l\}$.

Let BoT_t be a BoT submitted at time t.

Let N_t denote the number of tasks in a BoT_t submitted at time t.

Let I_t be the ideal makespan a BoT submitted at time t.

Let $M(BoT_t, a_j)$ be the makespan of BoT submitted at time t after scheduling with action j.

Let st_t be the submission time of BoT_t.

Let $ft_{i,t}$ be the finish time of task i in BoT_t where $i = 1, \dots N_t$.

Let $size\_mi_{i,t}$ be the size of task i in BoT_t where $i = 1, \dots N_t$ in Million Instructions (MI).

Let H be the set of hosts in a cloud where $H = \{h_1, h_2, \dots h_n\}$.

Let VM be the set of virtual machines in a cloud where $VM = \{vm_{1,1}, vm_{1,2}, \dots vm_{r,p}\}$ where $vm_{r,p}$ denotes a VMp assigned to host r.

Let HN_r be the number of VMs assigned to host r.

Let $mips_{r,p}$ denote the MIPS rating of a VMp assigned to host r.

Let $C = \{LG, ST, SM\}$ denote the classifications of VMs. Large VMs (LG) having mips >80,000, Standard VMs (ST) having mips between 40,000 and 80,000, and Small VMs (SM) having mips below 40,000.

Let $cpu\_util_{c,t} = \{H, M, L\}$ denote the average CPU utilization of all the cloudlets/tasks running on the VM at time t where $c = \{LG, ST, SM\}$. H denotes high CPU utilization (>80%). M denotes medium CPU utilization (40–80%). L denotes low CPU utilization (<40%).

Let A be the set of actions that are available to the agent. $A = \{a_1, a_2, \dots a_m\}$.

Let $r_{l,j,k}$ be the reward for the kth time an action j applied on state l.

Let $Q_{l,j,k}$ be the average of the first k rewards on action j in state l.

The working of the BIS intelligent agent is based on reinforcement learning adapted from Sutton and Barto [11] and can be summarized in the four stages outlined below:

Recognize the state: The BIS agent will calculate the size ratio of a BoT, find out the available number of VMs, and the MIPS share is available at that point in time and also the number of tasks in a BoT.

The size ratio is calculated as number of small tasks divided by the number of large tasks. A higher size ratio denotes a large number of small tasks present as compared to larger tasks.

The average CPU utilization of all cloudlets running on VMs at time t, $cpu\_util_{c,t}$ is calculated for all VMs classified into $C = \{LG, ST, SM\}$.

The size ratio of a BoT and the $cpu\_util_{c,t}$ for every classification of VMs $C = \{LG, ST, SM\}$ form a state.

Assess the effect of possible actions: After the state is recognized, the ideal makespan is calculated for the given BoT. We calculate the reward by how well a policy performs w.r.t. the ideal makespan. The ideal makespan is calculated using Eq. (1) which is the time taken for the largest task, i.e., the task with the longest instruction length (MI) to execute in the fastest machine, i.e., the VM with the largest MIPS rating.

$$I_t = \frac{max \sum_{i=1}^{N_t} (size\_MI_{i,t})}{max \sum_{r=1}^{H} \sum_{p=1}^{HN_r} mips_{r,p}} \tag{1}$$

The makespan of a BoT_t on action a_j is calculated using Eq. (2) which is the difference between the maximum finish time of a task in BoT_t and the submission time of BoT_t.

$$M(BoT_t, a_j) = max \sum_{i=0}^{N_t} (ft_{i,t}) - st_t. \tag{2}$$

This value of the makespan divided by the ideal makespan will be the reward, $r_{l,j,k}$ at the kth time the policy a_j was used on a given state l. Initially, the reward, $r_{l,j,0}$ is 0.

$$r_{l,j,k} = I_t / M(BoT_t, a_j). \tag{3}$$

An incremental implementation is used to calculate the cumulative reward at each step. In order to decide which action to choose at kth time to schedule BoT_t in state l, probability of choosing each action or policy is calculated.

Let $\pi_{l,\tau}(a_j)$ be the probability of choosing action a_j at play τ for state l, i.e., the tth time of the state l is encountered.

Let $P_{l,\tau}(a_j)$ be the preference of an action selected at play τ of the state l.

We set the initial action preferences to 0. The initial reference reward, $\bar{r}_{l,0}$, for every state l where play τ is 0 is set to 0.1 meaning if the makespan of the

scheduling policy is more than 10% as efficient as the ideal makespan, the reward is incremented and the preference for that action is set to a positive value. The preference is the difference between the reward $r_{l,\tau}$ and the reference reward $\bar{r}_t$.

The probability $\pi_{l,\tau}(a_j)$ of selecting an action j at play τ in state l is calculated using Eq. 4.

$$\pi_{l,\tau}(a_j) = \frac{e^{P_{l,\tau}(a_j)}}{\sum_{b=1}^{A} e^{P_{l,\tau}(a_b)}} \tag{4}$$

Decide on action that is matching with the goal: In this stage, the agent will choose the appropriate scheduling heuristic that minimizes the makespan of the BoT. The policy is chosen using *Roulette Wheel Selection* where all the probability for a set of actions in a given state is in the scale of one.

Execute action: Schedule the BoT according to the scheduling policy decided upon in Step 3 by assigning the cloudlets to the VMs.

3 Experimental Setup

The experimental setup of the private cloud in CloudSim is shown in Fig. 2. The setup for this simulation is based on the actual systems in a private cloud setup found in [6]. It consists of seven hosts and 10 VMs allotted to each PE (Processing Element).

The workload for the trial run is taken from LCG (Large Hadron Collider Computing Grid) archive [12]. 100 tasks for 10 BoTs were sampled from the log. The following heuristics were used based on ordering of the BoT and the mapping
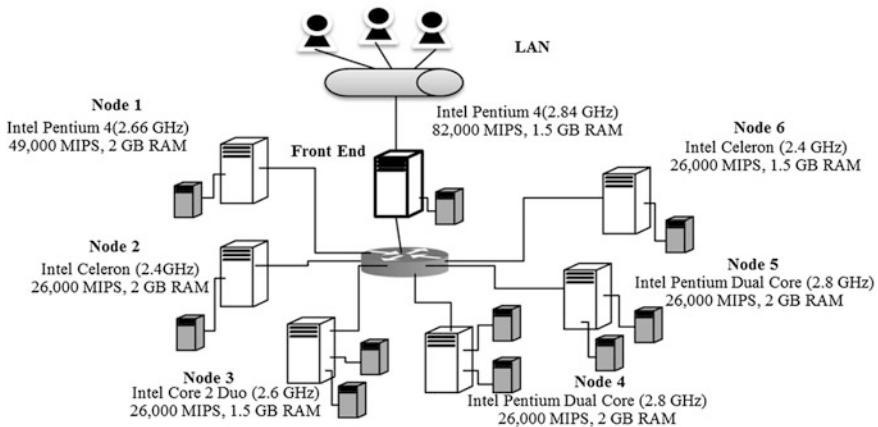
Fig. 2 Private cloud setup in CloudSim

policies adapted from Garcia and Sim [3]. MinET, MinCT, MaxET, and MaxCT are implemented as defined in Maheshwaran et al [7]. These allocation policies are based on three forms of ordering, *Uniform* (*U*), *LtoS* (*Large to Small*), *and StoL* (*Small to Large*). The mapping policies used are as follows:

Random (*R*): In this policy, there is no particular rule that is used and *{U, R}* mapping policy corresponds to FCFS.

Minimum Completion Time (*MinCT*) *and Maximum Completion Time* (*MaxCT*): It assigns each task to the machine that results in the task's earliest completion time in case of MinCT and vice versa for MaxCT.

Minimum Execution Time (*MinET*) *and Maximum Execution Time* (*MaxET*): It assigns each task to the machine that performs that task's computation in the least amount of execution time and vice versa for MaxET.

The agent can thus schedule each BoT for each of these scheduling policies: *{U, R}, {StoL, R}, {LtoS, R}, {U, MinET}, {StoL, MinET}, {LtoS, MinET}, {U, MinCT}, {StoL, MinCT}, and {LtoS, MinCT}.*

4   Results and Discussion

The cloud environment was scheduled in CloudSim [2], and the agents were incorporated as threads. The BoTs were scheduled on all possible scheduling policies for control purposes, and we observed that *{LtoS, MinCT}* performs better than other policies because we have one system which is almost three times and assigning longer tasks to this faster machine first seems to reduce the overall makespan of the BoT.

We feed the BIS agent with the same 1000 tasks for each run and observe the makespan and the scheduling policy chosen by BISA agent. We ran the simulation 10 times. During the start of each BoT, all the machines are free, so we have three states depending only on the size ratio of the BoTs. The CPU utilizations for all three states are $cpu\_util_{LG,t} = L$, $cpu\_util_{ST,t} = L$, and $cpu\_util_{SM,t} = L$.
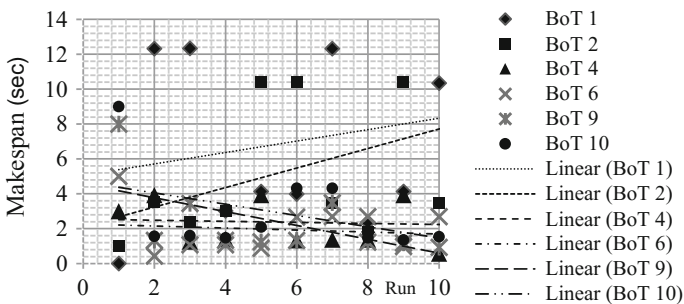
Fig. 3 Scatterplot of BoT in state 1

The scatterplots for BoTs in state 1 are presented in Fig. 3. BoTs 1, 4, 6, 9, and 10 fall into state 1 which corresponds to BoT size ratio 2.

We observe that, except for BoTs 1 and 2, there is a steady decline in makespan for all the other BoTs. The geometric mean of BoTs 1 and 2 is around 500 *MI,* whereas for the other BoTs with the same size ratio, the geometric mean is less than 270 *MI*. So, when the agent is choosing a scheduling algorithm that is less efficient, the loss in efficiency incurred due to that it is also proportionally higher due to the large geometric mean of the tasks.

Figure 4a, b shows the makespan of BoTs in state 2 and state 3, which corresponds to BoT ratio of 3 and 1, respectively. The number of BoTs in these states is comparatively lesser than for state 1. Even though the preference for *{LtoS, MinCT}* which will result in the best schedule is higher, the agent is still at the exploration phase and the number of plays is not sufficient in this state to observe a convergence.

Upon observing the preferences in the table of preferences at the end of the simulation, we observed that *{LtoS, MinCT}* has the highest preference. The agent's efficiency can be increased if we eliminate an action that has been applied k number of times in a given state if the preference is very low or in negative. With some slight modifications in the working of the BIS agent, this study does show promising results for scheduling tasks in heterogeneous large-scale systems. The major findings of our work are as follows:

- *Reinforcement learning* can be used to choose an optimal policy in a state.
- The *size ratio* parameter of a BoT alone is insufficient to distinguish between different types of BoTs.
- *{LtoS, MinCT}* performs best for the underlying heterogeneous hardware configuration presented in this paper.
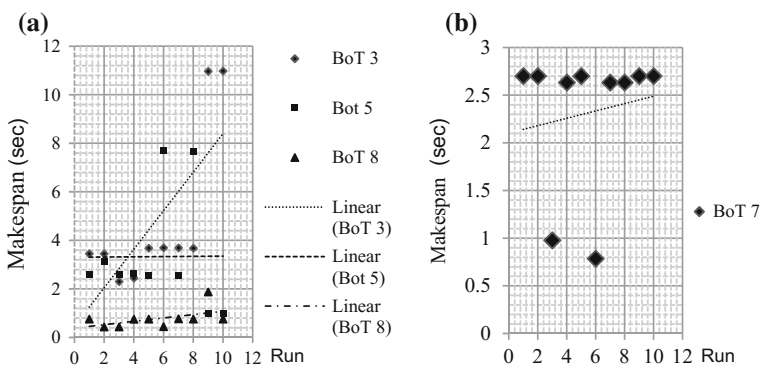
Fig. 4 **a** Scatterplot of BoT in state 2, **b** scatterplot of BoT in state 3

5 Conclusion and Future Work

A framework is presented where agents are used to learn the most optimal scheduling policy for a given state using reinforcement learning. The present study shows that over time, the probability of an agent choosing the optimal policy increases while the agent still learns by exploration. In an unpredictable large-scale distributed system like the cloud, even when the environment changes, the agent is still able to learn and adapt to the new environment because it still chooses a suboptimal solution, which may work better in the new environment.

Our future work consists of developing an agent for cloud bursting that takes into account the cost and network latency factors associated with the public clouds when allocating a task to a resource in the public cloud. We will also implement meta-heuristics that agents can choose from.

Another dimension we want to pursue is to use multi-agents which communicate with each other and "gossip" [8] its learned optimal policy for a given state with other agents. This will significantly reduce the time take for an agent to "learn" the most optimal policy for a given state.

References

1. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M.: A view of cloud computing. Commun. ACM **53**(4) (2010)
2. Calheiros, R.N., Ranjan, R., Beloglazov, A., De Rose C.A.F., Buyya, R.: CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Softw. Pract. Experience **41**(1), 23–50 (2011)
3. Garcia, J., Sim, K.M.: A family of heuristics for agent-based elastic Cloud bag-of-tasks concurrent scheduling. Future Gener. Comput. Syst. **29**(7), 1682–1699 (2013)
4. Ibarra, O.H., Kim, C.H.: Heuristic algorithms for scheduling independent tasks on nonidentical processors. J. ACM (JACM) **24**(2), 280–289 (1977)
5. Iosup, A., Epema, D.: Grid computing workloads. IEEE Internet Comput. **15**(2), 19–26 (2011)
6. Iqbal, W., Dailey, M.N., Carrera, D., Janecek, P.: Adaptive resource provisioning for read intensive multi-tier applications in the cloud. Future Gener. Comput. Syst. **27**(6), 871–879 (2011)
7. Maheswaran, M., Ali, S., Siegal, H.J., Hengsen, D., Freund, R.F.: Dynamic matching and scheduling of a class of independent tasks onto heterogeneous computing systems. In: Eighth Proceedings of Heterogeneous Computing Workshop, pp. 30–44 (1999)
8. Palmieri, G., Luigi, B., Salvatore, V., Aversa, R., Di Martino, B.: A distributed scheduling framework based on selfish autonomous agents for federated cloud environments. Future Gener. Comput. Syst. 1461–1472 (2013)
9. Russell, S.J., Norvig, P.: Artificial intelligence a modern approach. N. J. Prentice Hall **41**, 42 (1995)
10. Sim, K.M.: Agent based cloud computing. IEEE Trans. Serv. Comput. **5**(4), 564–577 (2011)
11. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction, pp. 43–74, MIT Press (2012)
12. The LCG Grid Log. http://www.cs.huji.ac.il/labs/parallel/workload/l_lcg/. Accessed 10 Jan 2016

Modeling a Bioinspired Neuron: An Extension to the H-H Model

Plabita Gogoi, Satyabrat Malla Bujarbaruah and Soumik Roy

Abstract The Hodgkin–Huxley (H-H) model of axonal membrane is one of the most inspiring and popular neuron models. Many variants of this model are present in current literature. However, this model does not encompass the post-synaptic membrane response to synaptic functions that are responsible for the dynamic membrane behavior. The paper thus presents an extended version of the H-H model that incorporates a "conjoint" synaptic model representing the complex synaptic activities and describes how the post-synaptic membrane behaves in presence of a synaptic entity as an input. The simulation is performed in MATLAB environment and the results are presented in the form of graphs.

Keywords Neuron · Synapse · Neurotransmitters · H-H model

1 Introduction

The nervous system is indeed one of the most complex yet mesmerizing systems of the human body. The discrete units of the nervous system are called *neurons* which carry out the responsibility of mediating information transfer between different points in the human body using complex electrophysiological phenomenon. The neuron in turn loosely comprises of a number of discrete structural and functional components out of which, the *synapse* can be considered as a "System-within-a-System" with all its integral electrochemical machinery. The modeling of neuron has therefore been one of the most spectacular breakthroughs of this era [4]. The extensive and exhaustive study of the giant squid membranes led to the pioneering work of Hodgkin–Huxley (H-H)

P. Gogoi (✉) · S. M. Bujarbaruah · S. Roy
Tezpur University, Napaam, Tezpur, Assam, India
e-mail: plabita.gogoi6@gmail.com

S. M. Bujarbaruah
e-mail: baruah.satyabrat@gmail.com

S. Roy
e-mail: xoumik@tezu.ernet.in

model of axonal membrane which is the core of this paper along with the synapse models that describe the complex phenomena of neural transmissions.

The classical H-H model, however, was described only for an axonal membrane patch and did not include the effect of pre-synaptic signals on post-synaptic membrane at the synaptic cleft [5, 6]. Also, the external stimulus to the H-H model in most cases is seen to be a constant current signal or an impulse current. Yet, another possibility can be a current form due to in vivo synaptic interactions [9]. In the proposed model, the latter concept is used to obtain an extended version of the classical H-H model to describe the post-synaptic function's dependence on the in vivo pre-synaptic current using the amalgamation of two models: the classical H-H model and the synapse model; and the effect is observed in the form of variation in membrane voltage of the post-synaptic neuron.

1.1 The Hodgkin–Huxley Model of Nerve Axon

The nerve membrane consists of a phospholipid bilayer that is selectively permeable to various ions present in the fluidic environment of the body. This characteristic nature of the membrane leads to both electrical and chemical gradients across it and aids in generation and propagation of action potentials in suitable conditions. This very electrophysiological system was modeled by Hodgkin and Huxley using a versatile mathematical equivalent and some elegant equations [2, 3, 5]. The H-H model is presented in Fig. 1.

The modeling is done using a *Membrane Capacitance,* C_m, which has a constant value and leads to a current called the *Capacitive Current* depicted by I_C. The "selective ion conductance" behavior itself was modeled using a conductance in series with a battery and together it leads to an *Ionic Current*, I_{ion}, due to movement of ions across the membranes and are further classified as sodium current, I_{Na}, potassium current, I_K, and leakage current, I_L (in case of chloride and calcium).

Fig. 1 The H-H model representing the electrical equivalent of the nerve membrane
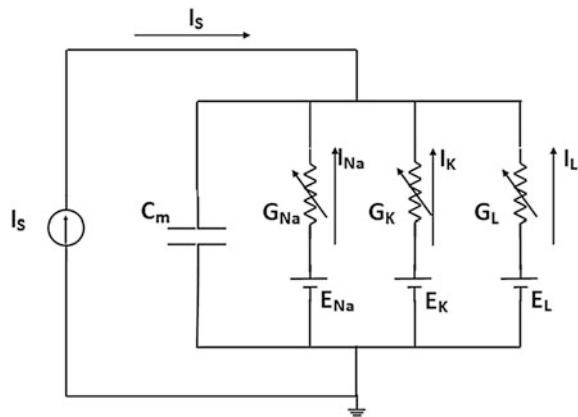
The H-H equations are described by the following mathematical relations that explain the dynamic nature of the membrane [6]:

$$I_S = I_C + I_{ion} \tag{1}$$

$$I_S = C_m \frac{dV_m}{dt} + I_{Na} + I_K + I_L \tag{2}$$

$$I_S = C_m \frac{dV_m}{dt} + G_{Na}(V_m - E_{Na}) + G_K(V_m - E_K) + G_L(V_m - E_L) \tag{3}$$

Here, t is time and V_m is the membrane potential at a specific point of time.

1.2 The Ion Channels

The nerve membrane at specific locations has channels that allow a constrained movement of ions. These channels are of various types, the most important being the *voltage-gated* channels (VGCs) and the *ligand-gated* or *ionotropic* channels (LGCs) [10]. The permeability of voltage- gated channels depends upon the membrane potential value at that point of time. Similarly, permeability of ionotropic channels depends upon the binding activity of neurotransmitters with specific receptors and is specific to the synaptic region of the neuron only. The membrane channels follow a dynamic pattern of opening and closing function that characterises the electrical behaviour of the membrane [9]. At a point of time if p are the fraction of gates which are open, then from probabilistic model we know that (1−p) are the fraction of gates which are in a closed state. The simplest model that represents this theory is depicted by Fig. 2 [10].

Here, α_p and β_p are the forward and backward rate constants, respectively, representing a specific number of times per second that a closed-state gate gets opened and vice versa. The resulting differential equation is as follows [10]:

$$\frac{dp}{dt} = \alpha_p(1 - p) - \beta_p p \tag{4}$$
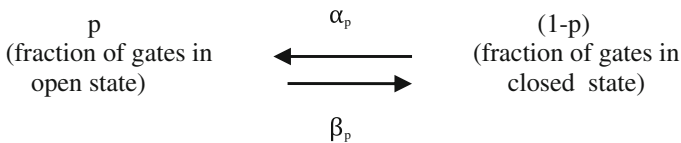
Fig. 2 A simple diagrammatic representation of the channel dynamics

Considering the two major ions, sodium and potassium, the probabilities are denoted as "m" which represents the sodium activation constant, "n" which represents potassium activation constant, and "h" which represents sodium inactivation constant. Consequently, (α_m, β_m), (α_h, β_h), and (α_n, β_n) are the sets of rate constants for the three probabilities. The generalized equation (4) thereby yields the relation of the individual ionic conductances, which are represented as [6].

$$G_{Na} = \bar{g}_{Na} m^3 h \tag{5}$$

$$G_K = \bar{g}_K n^4 \tag{6}$$

Thus, applying Eqs. (5) and (6) in Eq. (3) gives the ultimate current equation [6] and solved using a set of ordinary differential equations [1, 7]:

$$I_S = C_m \frac{dV_m}{dt} + \bar{g}_{Na} m^3 h (V_m - E_{Na}) + \bar{g}_K n^4 (V_m - E_K) + G_L (V_m - E_L) \tag{7}$$

1.3   Modeling the Synapse

The synaptic cleft is a complex region that acts as a communication junction between two neurons. Pre-synaptic potentials of requisite "strength" on reaching the synaptic cleft depolarize the membrane causing influx of calcium ions which in turn triggers neurotransmitter release that binds with specific receptors (LGCs) present in the post-synaptic neuron causing certain morphological changes which may culminate into action potentials. The above complex synaptic transmission phenomena have been modeled in various ways, a great detail of which has been discussed by Roth and van Rossum [8]. Among these, two simple yet effective models are discussed below, namely "Difference of two exponentials" and "Current based" models of synapses.

In "Difference of two exponentials" model of synapse, the synaptic conductance, G_{syn}, is modeled as a sum of two exponential functions which represents the conductance profile due to rapid binding and slow unbinding of neurotransmitters. The equations which define this model are as follows [8]:

$$G_{syn}(t) = \bar{g}_{syn} f (e^{-(t-t_o)/\tau decay} - e^{-(t-t_o)/\tau rise}) \tag{8}$$

$$f = \frac{1}{-e^{-(t_{peak} - t_o)/\tau rise} + e^{-(t_{peak} - t_o)/\tau decay}} \tag{9}$$

$$t_{peak} = t_o + \frac{\tau_{decay} \tau_{rise}}{\tau_{decay} - \tau_{rise}} \ln\left(\frac{\tau_{decay}}{\tau_{rise}}\right) \tag{10}$$

Here, $\tau_{decay} \neq \tau_{rise}$ and G_{syn} has a non-zero value for $t \geq t_o$.

The "Current based" model of synapse accounts for the "linear current–voltage relationship" behavior of most ligand-gated channels in an open state, where G_{syn} can be modeled as an ohmic conductance which when multiplied with a potential difference of $(V_m - E_{syn})$ gives a synaptic current equal to I_{syn} represented as [8]:

$$I_{syn} = G_{syn}(t)(V_m - E_{syn}) \tag{11}$$

2 H-H Model for a Synaptic Input—The Proposed Model

To express the post-synaptic membrane behavior to an incoming pre-synaptic signal at the synaptic junction, the classical H-H model is augmented with the synapse model. The synapse model here is obtained as a product of two synaptic models described in Sect. 1.3 (by putting Eq. (8) in (11)) resulting in the following equation:

$$I_{syn} = \overline{g}_{syn} f(e^{-(t-t_o)/\tau decay} - e^{-(t-t_o)/\tau rise})(V_m - E_{syn}) \tag{12}$$

The post-synaptic membrane is initially considered to be at rest. Assuming occurrence of neurotransmitter binding due to pre-synaptic spikes at synapse, ligand-gated channels will be activated resulting in post-synaptic current as depicted by Eq. (12). This "in vivo" current will now act as a stimulus to the axonal membrane causing its depolarization by activation of voltage-gated sodium and potassium channels in sequence leading to the generation of an action potential. The resulting modified current equation is thus obtained by incorporating Eq. (12) in (7), where G_{syn} is the conductance of synaptic channel at synaptic cleft and I_{syn} is the resultant ionic current due to permeability of this channel with respect to time.

$$\begin{aligned}
&\overline{g}_{syn} f(e^{-(t-t_o)/\tau decay} - e^{-(t-t_o)/\tau rise})(V_m - E_{syn}) \\
&= C_m \frac{dV_m}{dt} + \overline{g}_{Na} m^3 h(V_m - E_{Na}) + \overline{g}_K n^4(V_m - E_K) \\
&\quad + G_L(V_m - E_L)
\end{aligned}$$

The modified version of the H-H model thus obtained is represented in Fig. 3.

2.1 Simulation

The modeling is done considering that the axon hillock is close to the synaptic junction accounting for negligible signal loss due to propagation, and the synapse considered is excitatory in nature ensuring generation of an action potential when triggered by a threshold synaptic current. The simulation is performed in MATLAB environment for the component values as tabulated in Tables 1 and 2 respectively [5, 6, 8, 10].
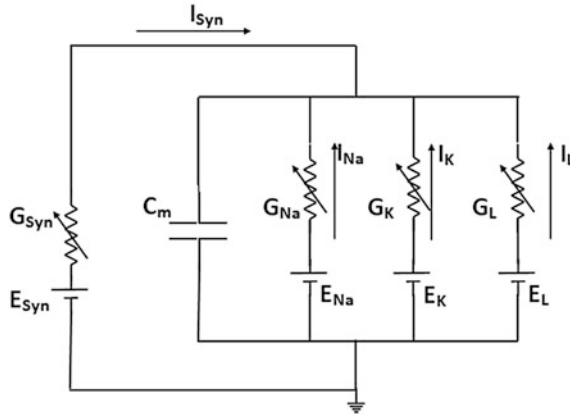
Fig. 3 The H-H model augmented by a synaptic model

Table 1 List of values assigned to the various components of the synaptic model

Parameters	Parameter details	Unit	Value
τ_{rise}	Rising time constant	Second	1 ms
τ_{delay}	Decaying time constant	Second	3 ms
t_o	Transmission delay	Second	0.05 ms
E_{syn}	Synaptic potential	Volt	70 mV
$\bar{g}_{syn}$	Maximum synaptic conductance	Mho	1 mS per cm^2

Table 2 List of values assigned to the various components of the H-H model

Parameters	Parameter details	Unit	Value
C_m	Membrane capacitance per unit area	Farad	1μF per cm^2
g_{Namax}	Maximum sodium conductance	Mho	120 mS per cm^2
g_{Kmax}	Maximum potassium conductance	Mho	36 mS per cm^2
g_{lmax}	Maximum leakage conductance	Mho	0.3 mS per cm^2
E_{Na}	Sodium reversal potential	Volt	–55.17 mV
E_K	Potassium reversal potential reversal	Volt	–72 mV
E_l	Potential for leakage ions	Volt	–49.378 mV
I	Total membrane current	Ampere	0 A
V_m	Initial membrane voltage	Volt	–60 mV

3   Results

The MATLAB simulation results are shown in the Figs. 4 and 5. Figure 4 describes how synaptic current I_{syn} obtained in response to a combined synaptic model varies with time. The magnitude of the synaptic current is positive for the direction of I_s

Fig. 4 Simulation results showing the variation of synaptic current with time
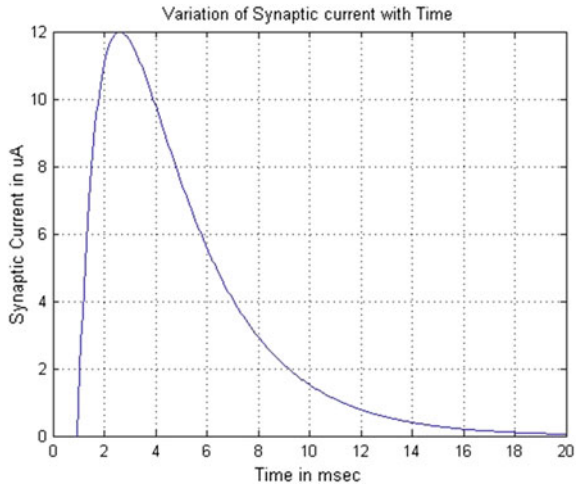
Fig. 5 Simulation results showing the generation of an action potential due to application of a synaptic current input
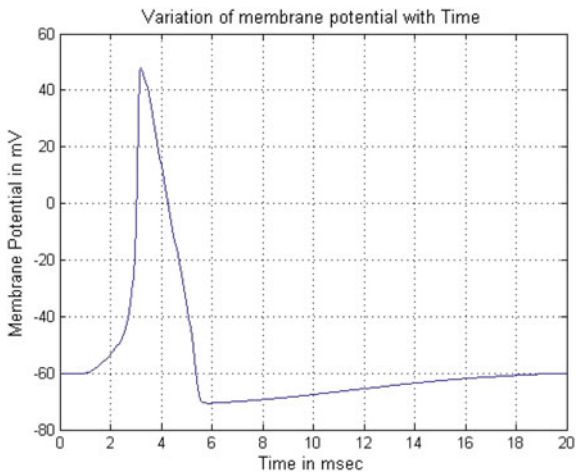
shown in Fig. 3 and hence represents an excitatory synapse response. Figure 5 represents the situation when a suitable synaptic input is applied to the H-H model of neuron causing the membrane potential to depolarize and upon attaining the necessary threshold limit, an action potential is generated with a peak value of 47.759 mV occurring at 3.25 ms. The observations are very similar to the ones obtained for the classical H-H model reacting to a constant external current input [1, 9].

4   Conclusion

The H-H model is implemented in MATLAB for an input current that arises from within the system in the form of a synaptic current due to pre-synaptic activities. The observation consolidates the fact that a synaptic model can be augmented with the basic H-H model to obtain an extended H-H model that is capable of representing the post-synaptic function dependence on the pre-synaptic variations at the synaptic cleft.

References

1. Börgers, C.: The classical Hodgkin-Huxley ODEs. In: An Introduction to Modeling Neuronal Dynamics. Springer International Publishing, pp. 15–21 (2017)
2. Dutta, J.C., Roy, S.: Biologically motivated circuit model for simulation of excitatory and inhibitory synapses. Can. J. Biomed. Eng. Technol. **1**(2), 52–55 (2010)
3. Fitzhugh, R.: Threshold and plateaus in the Hodgkin-Huxley nerve equations. J. Gen. Physiol. **43**, 867–896 (1960)
4. Harmon, L.D., Lewis, E.R.: Neural modeling. Physiol. Rev. **48**, 513–591 (1966)
5. Hodgkin, A.L.: Ionic movements and electrical activity in giant nerve fibers. In: Proceedings of the Royal Society of London. Series B, Biological Sciences, pp. 1–3 (1957)
6. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and application to nerve. J. Physiol. **117**, 500–544 (1952)
7. Nagy, A.M., Sweilam, N.H.: An efficient method for solving fractional Hodgkin-Huxley model. Phys. Lett. A **378**(30), 1980–1984 (2014)
8. Roth, A., van Rossum M.C.W.: Computational Modeling Methods. The MIT Press (2009)
9. Shepherd, G.M.: Creating Modern Neuroscience: The Revolutionary 1950s, Oxford (2010)
10. Wells, R.B.: Introduction to Biological Signal Processing and Computational Neuroscience, Moscow (2010)

A Nonnegative Matrix Factorization Based Approach to Extract Aspects from Product Reviews

Debaditya Barman and Nirmalya Chowdhury

Abstract Due to the unstructured nature of review text, it is very hard to develop an automated opinion mining application to compare various product models based on their various aspects to make a purchase decision. Over the year, various data mining techniques have been proposed to extract aspects of the products. In this paper, we have proposed a technique based on the nonnegative matrix factorization to extract aspects of a product category. Performance of our proposed method has been compared with a very popular aspect extraction technique based on probabilistic latent semantic analysis. We have also given a comparison between common aspects of a particular model under a specific product category from various manufacturers. These comparisons are based on the sentiments expressed by the users on these aspects. These sentiments expressed in various aspects have been extracted using an unsupervised technique.

Keywords Aspect extraction · Opinion mining · Nonnegative matrix factorization · E-commerce

1 Introduction

Online retailers often asked their customers to evaluate their product by providing reviews and ratings to help a potential customer to make a purchase decision. Manual analysis of these reviews is impossible because of the sheer amount of reviews. On the other hand, traditional rating system has a couple of problems. First, there is no mechanism to detect spam ratings. Second, most of the e-commerce sites do not have any feature-wise rating system, so it is very hard for

D. Barman
Department of Computer and System Sciences, Visva-Bharati, Bolpur, India
e-mail: debadityabarman@gmail.com

N. Chowdhury (✉)
Department of Computer Science and Engineering, Jadavpur University, Kolkata, India
e-mail: nirmalya_chowdhury@yahoo.com

the consumers to make an informed decision if he or she wants to buy a product that is more feature-wise rich than other products. In this paper, we have proposed a framework to address above-mentioned problems. Our proposed method analyzes all the reviews expressed on a particular model under a specific product category from various manufacturers to provide aspect-wise comparison to customers.

Topic modeling techniques can be used to extract aspects of these products. These techniques have been used to discover topics from a large set of text documents. We can assume that a text document can contain multiple topics, and each topic is a group of some words. Intuitively, these topics can be treated as aspects. Latent Dirichlet Allocation (LDA) and probabilistic Latent Semantic Analysis (pLSA) are two frequently used models in topic modeling. In 2007, Mei et al. [1] proposed Topic Sentiment Mixture (TSM), which is a probabilistic mixture model based on pLSA. TSM is used to model and extract multiple topics (aspects) along with the associated sentiments (positive and negative) from a set of blog articles. Lin and He (2009) [2] proposed an unsupervised technique called joint sentiment/topic model (JST), which is an extension of LDA to detect sentiment and aspects simultaneously from text. In 2010, Li et al. [3] proposed joint sentiment and topic modeling using sentiment LDA and dependency-sentiment LDA. Their approach is capable of extracting aspects with positive or negative sentiments using inter-dependency of sentiments. Zhao et al. (2010) [4] integrated the concept of maximum entropy in LDA and created MaxEnt-LDA. This hybrid model can extract both aspect and aspect-specific opinion words. In 2012, Mukherjee and Liu [5] proposed Seeded Aspect and Sentiment model (SAS). SAS is a semi-supervised model. It takes seed words for aspect categories as input from the user and produces clusters of extracted aspect terms. Nonnegative Matrix Factorization techniques (NMF) have been used to uncover latent low-dimensional structures present in high-dimensional data and provide a nonnegative, part-based, representation of data. In 2004, Peng and Park [6] proposed Constrained Symmetric Nonnegative Matrix Factorization (CSNMF) algorithm to generate a sentiment dictionary. In 2004, Pauca et al. [7] proposed a hybrid NMF algorithm using gradient descent with constrained least squares method to extract topics and document clusters from a collection of text documents. In 2006, Shahnaz et al. [8] used NMF to identify topics and clusters in a collection of text documents. In 2015, Kuang et al. [9] developed UTOPIAN (User-driven Topic modeling based on Interactive NMF) to cluster and extract topics from text documents.

In this paper, we have proposed an aspect extraction technique based on NMF method. This method decomposes a term-document matrix into two nonnegative factors: term-topic matrix and topic-document matrix. The term-document matrix can be created from a set of reviews on a particular model by considering each review as a document and frequent nouns as terms. Since all the reviews are expressed on a particular product model, number of topics is one here. Terms with a higher value of the term-topic matrix can be treated as potential aspects. By analyzing reviews on several product models under a product category, we can create a pool of potential aspects of a product category. The aspect with the frequency greater than a predefined threshold can be treated as the aspect of a particular

product category. We have also analyzed the sentiments expressed on these extracted aspects to provide better insights about the product to a potential customer.

The rest of this paper is organized as follows: statement of the problem can be found in Sect. 2. Section 3 describes each of the components of our proposed system's framework. We have reported our experimental results in Sect. 4. Concluding remarks can be found in Sect. 5.

2 Statement of the Problem

This paper presents a method that can extract various aspects of a product from a set of reviews by users of a product from various manufacturers and then provide a comparative analysis of the sentiments expressed by the users on these aspects. We have used an unsupervised technique (i.e., Semantic Orientation (SO) [10]) to find the polarities of the sentiments. A review of a product generally consists of some opinions about the products. We can define these opinions [11] as follows:

Definition (Opinion): An opinion can consist of four parts—(ot, oh, s, t)—where t is the timestamp when an opinion holder oh expressed his or her sentiment s about ot (i.e., opinion's target).

Example: A Flipkart user wrote following review of iPhone 6s on 23.04.2017.

"(1) Excellent phone. (2) Excellent service by seller and Flipkart. (3) Camera, battery too good besides other well known features of the phone…"

Here, opinion holder is the Filpkart user and time is April 23, 2017. In sentence 1, "*excellent*" is the sentiment and "*phone*" is the target of the opinion. In sentence 2, "*excellent*" is the sentiment and we have two targets: "*seller*" and "*Flipkart*". Similarly, in sentence 3, we have two opinion targets: "*Camera*", "*battery*", and only one sentiment: "*good*".

Let *entity* denote an opinion's target.

Definition (Entity): An entity e can be represented by the tuple (T, W), where T is a hierarchical representation of different parts or subparts of e and W is the set of attributes of these parts, subparts, or entity. An entity can represent a product, service, topic, company, or event.

Example: A specific brand of a laptop can be an entity, e.g., Apple MacBook Air. A set of attributes of this laptop can be *dimension*, *weight*, *color*, and *price*. Some parts of these laptops are *processor, storage, display,* and *battery*. These parts can have own set of attributes. *Speed (e.g., 1.8 GHz or 2.5 GHz), type* of the *processor (e.g., i5 or i7)* can be the attributes of the processor. These parts can be divided into multiple subparts. For instance, subparts of the storage can be *primary storage* and *secondary storage*.

The hierarchical representation of the entity is very difficult to extract from a set of reviews. Since processing of unstructured natural language is very hard,

recognizing various parts and subparts, establishing relationship between parts and subparts, and extracting attributes of these parts and subparts are very tough tasks. So to simplify the situation, we have adopted a two-level representation where entity is the root and different parts, subparts or attributes are the children. These parts, subparts, and attributes have been represented by an element called *aspect*.

Definition (Aspect): An aspect (a_{ij}) represents parts, subparts, or attributes of an entity (e_i).

We can model the entity (e_i) as a finite set of aspects $A_i = \{a_{i1}, a_{i2}, \ldots, a_{im}\}$. An entity (e_i) can represent a product category (e.g., Mobile, Tablet, TV, etc.) but in the e-commerce websites, users express their opinion on a particular model (e.g., iPhone 5s, MI 4, Moto G Turbo, etc.) under a specific product category from various manufacturers. To be specific, users expressed their sentiments on the aspects of these particular models. So, we have to redefine the opinion by introducing the aspect and a particular model under a product category.

Definition (Opinion): An opinion can consist of five parts—$(\acute{e}_i, a_{ij}, s_{ijkl}, h_k, t_l)$, where at time t_l, an user h_k expressed his or her sentiment s_{ijkl} on an aspect a_{ij} of a product model $\acute{e}_i$.

An opinion document (d_i) consists of opinions expressed on a set of product models $\{\acute{e}_{i1}, \acute{e}_{i1}, \ldots, \acute{e}_{in}\}$ and subset of their different aspects by a set of users $\{h_1, h_2, .., h_k\}$ at some time. Essentially, $\bigcup_{i=1}^{n} \acute{d}_i \equiv d_i, where\, d_i$ contain reviews on $\acute{e}_{ij}$. It is assumed that, in set of opinion documents $\left(D = \{d_1, d_2, \ldots, d_p\}\right)$, each document (d_i) is related to a particular entity (e_i) and $\exists d_i \in D: (P(d_i) \bigwedge \forall d_j \in D: P(d_j) \rightarrow d_i = d_j)$, where the property $P(d_i)$ denotes d_i that represents one and only one entity e_i.

For example, we have two entities: *mobile* and *tablet*. Let $e_1 = mobile$ and $\acute{e}_{11} = Apple\,iphone, \acute{e}_{12} = Samsung\,Galaxy\,note.$ Let $e_2 = tablet$ and $\acute{e}_{21} = Apple\,ipad,$ $\acute{e}_{22} = Lenovo\,Yoga\,tab.$ So d_1 is related to the entity $e_1(mobile)$ consists of two set of reviews $\acute{d}_{11}$(reviews of *Appleiphone*) and $\acute{d}_{12}$(reviews of *Samsung Galaxy note*). Similarly, d_2 is related to the entity $e_2(tablet)$ consists of two sets of reviews $\acute{d}_{21}$(reviews of *Apple ipad*) and $\acute{d}_{22}$(reviews of *Lenovo Yoga tab*). So, the opinion dataset $D = \{d_1, d_2\}$.

Given a set of opinion documents, our objectives are as follows:

Task 1 (Aspect Extraction):

Extract a set of aspects (A_i) of a particular entity e_i.

Task 2 (Aspect Sentiment Classification):

Assign a class label (*positive*, *negative*, or *neutral*) to the extracted aspects depending on the sentiments expressed by the users on these aspects.

3 The Proposed Methodology

In this paper, we have proposed a method based on nonnegative matrix factorization (NMF) [12] technique to extract various aspects of a product based on multiple reviews or opinions expressed on the product by the users. An unsupervised algorithm based on semantic orientation technique [13, 14] has been used to identify the sentiments expressed in these aspects. Figure 1 presents the framework of our aspect extraction system, and afterward, we have briefly described each component of our system.
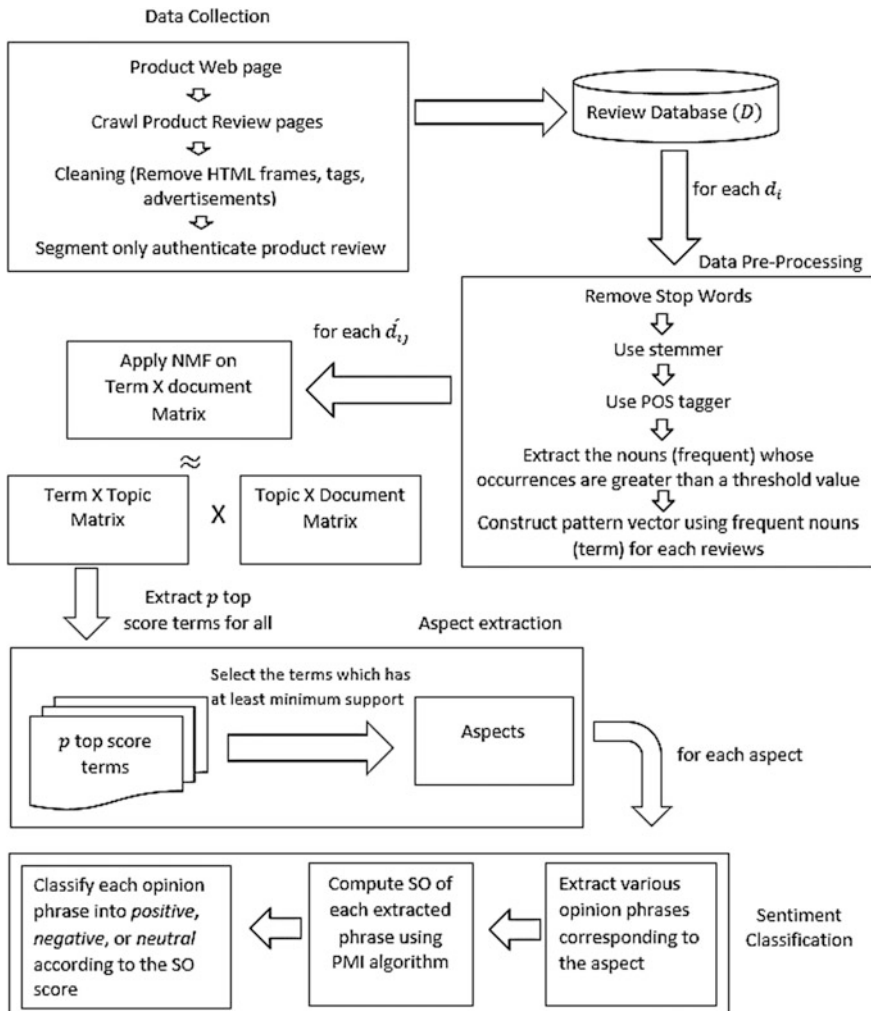
Fig. 1 The system framework

3.1 Data Collection

Given a product's name, our system can automatically locate the product review pages (in Flipkart) and crawl through them. An HTML parser has been developed to extract product review, author, time stamp, ratings, and certified buyer tag by removing various extraneous information (e.g., HTML tag, frame, product description, advertisements, etc.) from these review pages. We have only considered product reviews by certified users to avoid spam reviews. These reviews have been stored in a database to process them offline.

3.2 Data Preprocessing

We can represent a text document by a pattern vector of the frequency of the words presented in the document. Similarly, multiple text documents can be represented by a matrix. Unfortunately, size of this matrix will be very large. Following steps have been taken to reduce the size of this matrix:

Step 1. We have removed stop words (e.g., determiners, conjunctions, prepositions, and some adverbs) and applied *Porter stemmer* to remove morphological affixes from the words to extract root words.
Step 2. Part-of-Speech (POS) tagger has been used to extract only nouns.
Step 3. If number of occurrence of a noun is greater than a predefined threshold, then it is a frequent noun. The threshold is $c * \sqrt{N}$, where N is the number of review corresponding to a particular product model, and c is a constant and $0 < c < 1$. In our experiment, $c = 0.15$, it is obtained experimentally.
Step 4. These frequent nouns are candidates for the aspect selection procedure.

We have constructed a *term-document* matrix using these frequent nouns. Each frequent noun is treated as *term*, and each review is treated as a *document*. This term-document matrix is created for each product model.

3.3 Nonnegative Matrix Factorization

Suppose we have a nonnegative matrix $V \in \mathbb{R}^{m \times n}$ and a rank $k \ll \min(m, n)$. Nonnegative matrix factorization (NMF) [12] tries to decompose V into two nonnegative factors W and H by solving the following optimization problem.

$$\min V_{m \times n} - W_{m \times k} \cdot H_{k \times nF}^2, \quad \text{such that } W \geq 0 \text{ and } H \geq 0 \tag{1}$$

The rank of the approximation k can vary across problem domains, and domain expert can set the value. Frobenius norm is a frequently used parameter to compute

the error between the original matrix V and its low-rank approximation WH. So NMF decomposes the term-document matrix (i.e., V) into two matrices: term-topic matrix (i.e., W) and topic-document matrix (i.e., H). Three frequently used algorithms to solve the optimization function stated in Eq. 1 are multiplicative update rule, Alternating Least Squares (ALS), and Alternating Constrained Least Squares (ACLS). ACLS algorithm has certain advantages over the other two algorithms. ACLS is fastest among aforementioned three algorithms. Unlike multiplicative update rule, convergence time of ACLS is not dependent on how W and H have been initialized. ACLS is not lacking sparsity in the factors W and H, which was major disadvantage of ALS algorithm. Algorithm of the ACLS can be found in the following section.

Algorithm: Alternating constrained least square
Input: A term-document matrix $V_{m \times n}$, k number of topics, $\lambda_W \geq 0$, and $\lambda_H \geq 0$ are the parameters to control sparsity of W and H.
Output: Term-topic matrix $W_{m \times k}$ and topic-document matrix $H_{k \times n}$.
Steps:

Step 1. $W = abs(rand(m, k))$
Step 2. $for\ i = 1: maxiteration$
Step 3. Find solution for H in the equation $(W^T W + \lambda_H I)H = W^T V$
Step 4. $\forall h_{kn} \in H, h_{kn} < 0 \rightarrow h_{kn} = 0 \bigwedge h_{kn} \rightarrow h_{kn}$
Step 5. Find solution for W in the equation $(HH^T + \lambda_W I)W^T = HV^T$
Step 6. $\forall w_{mk} \in W, w_{mk} < 0 \rightarrow w_{mk} = 0 \bigwedge w_{mk} \rightarrow w_{mk}$
Step 7. $end\ for$

3.4 Aspect Extraction

We have selected the terms with top p score from the term-topic matrix $W_{m \times k}$. It may be noted that, for our experiment, $p = 20$. This procedure has been repeated for all the products $(\acute{d}_{ij})$ that belong to a particular entity (d_i). All the terms with frequency greater than a predefined threshold (minSupport $= 60\%$) have been considered as aspects of a particular entity (d_i).

3.5 Sentiment Classification

We have used extracted aspects to mine nearest opinion words with adjective or adverb. Pattern knowledge presented in Table 1 has been used to extract only sentiment part.

A set of positive words (i.e., *Pwords*) and a set of negative words (i.e., *Nwords*) have been defined to calculate semantic orientation of the extracted phrases using

Table 1 Pattern of POS tags for extracting sentiment

	First word	Second word[a]	Third word[a]
1	JJ	NN/NNS	–
2	RB/RBR/RBS	JJ/RB/RBR/RBS	NN/NNS
3	RB/RBR/RBS	VBN/VBD	NN/NNS
4	VBN/VBD	RB/RBR/RBS	NN/NNS
5	VBN/VBD	NN/NNS	–

[a]NN/NNS are not extracted

Point-wise Mutual Information (PMI). Words included in *Pwords* and *Nwords* always represent positive sentiment and negative sentiment, respectively, irrespective of the context. Equation 3 represents PMI between $term_1$ and $term_2$. Using Eq. 6, we can calculate *SO* score of the phrase. Sentiment of the phrase can be extracted through Eq. 7.

$$Pwords = \{good, nice, excellent, positive, fortunate, correct, superior\} \quad (2)$$

$$Nwords = \{bad, nasty, poor, negative, unfortunate, wrong, inferior\} \quad (3)$$

$$
\begin{aligned}
PMI(term_1, term_2) &= \log_2\left(\frac{\Pr(term_1 \wedge term_2)}{\Pr(term_1) \cdot \Pr(term_2)}\right) \\
&= \log_2\left(\frac{\frac{1}{N} * hits(term1\ AROUND\ term2)}{\frac{1}{N} * htts(term1) * \frac{1}{N} * htts(term2)}\right),
\end{aligned}
\quad (4)
$$

where N is the total number of documents in the search engine.

Details of the *Around* operator can be found in [14].

$$SO(phrase) = \sum_{pword \in Pwords} PMI(phrase, pword) - \sum_{nword \in Nwords} PMI(phrase, nword) \quad (5)$$

After placing the expression of *PMI* from Eq. 3 into Eq. 4, we get

$$SO(phrase) = \log_2\left(\frac{\prod_{pword \in Pwords} hits(phrase\ AROUND\ pword) * \prod_{nword \in Nwords} hits(nword)}{\prod_{pword \in Pwords} hits(pword) * \prod_{nword \in Nwords} hits(phrase\ AROUND\ nword)}\right) \quad (6)$$

$$Sentiment(phrase) = \begin{cases} negative, & SO(phrase) < 0 \\ positive, & SO(phrase) > 0 \\ neutral, & SO(phrase) = 0 \end{cases} \quad (7)$$

4 Experimental Results

We have conducted our experiment with three entities (i.e., mobile, laptop, and TV). Reviews provided by certified users on the different products under these entities have been collected from Flipkart. More details regarding the experimental data can be found in Table 2. We have extracted different aspects of these entities with our proposed methodology and compared the results with a very popular aspect extraction technique called pLSA. Due to limited space, we have shown frequent terms and aspects extracted for "*mobile*" entity. Frequent terms (terms with frequency greater than the minSupport = 3) along with their frequency for our NMF-based method and pLSA method can be found in Table 3. These frequent terms have been considered as aspects of an entity. Correctness of our proposed method has been checked by comparing aspects extracted by our method with the ground truth. The ground truth for different entities has been created manually by observing product specification given in their respective product web pages. Details of the ground truth of different entities used for our experimentation can be found in Table 4. We have used two metrics—accuracy (in percentage) and Jaccard index—to measure up the performance of our system. Accuracy measurement indicates how many aspect has been discovered by our method and Jaccard index shows the similarity and diversity of the proposed aspect set and ground truth aspect set. Jaccard index is very useful metric in this kind of scenario because it compares members of the proposed aspect set and ground truth aspect set to see which members are shared and which are distinct. High value of Jaccard index represents higher similarity between these two sets. Perfect Jaccard index is 1. We have

Table 2 Details of the experimental data

Product category	Sl. no.	Product name	Number of reviews
Mobile	1	Apple iPhone 5S	924
	2	Lenovo VIBE P1	1950
	3	Moto G Turbo	1216
	4	Samsung Galaxy J7	374
	5	Microsoft Lumia 640 XL	338
Laptop	1	Apple MacBook Pro	106
	2	Asus EeeBook X205TA	419
	3	Dell Vostro 14 V344	131
	4	HP 15 af114AU	428
	5	Lenovo G Series G50	98
TV (32″)	1	LG	306
	2	Micromax	448
	3	Panasonic	50
	4	Samsung	185
	5	Sony	50
	6	VU	1837

Table 3 Frequent terms along with the frequency

NMF		pLSA	
Terms	Frequency	Terms	Frequency
Battery	5	Battery	5
Camera	5	Camera	5
Service	5	Flipkart	5
Processor	5	Phone	5
Price	5	Price	5
Ram	4	Product	5
Screen	4	Quality	5
Charge	4	Buy	4
Quality	4	Mobile	4
Call	4	Day	3
Android	3	Screen	3
Warranty	3	Time	3
Mobile	3	Processor	3
		Ram	3

Table 4 Ground truth for different entities

Mobile	Laptop	TV
Battery	Battery	Delivery
Camera	Delivery days	Display
Display	Display quality	HDMI
Network type	Keyboard	Installation
Price	Operating system	LED/LCD
Processor	Price	Picture quality
RAM	Processor	Price
Screen size	RAM	Service
Service	Release year	Speaker
Warranty	Screen	USB port
	Service	HD/FHD
	Support	Release year
	Warranty	Remote
	Weight	Warranty

provided performance of our proposed NMF-based approach and pLSA-based approach for three entities in Table 5. We have extracted sentiments expressed by the users for all the aspects of an entity or product category. A potential buyer can study these sentiments expressed on different aspects of a particular product model under a product category to identify its strong and week aspects. It is also possible to compare different product models under a product category from different manufacturers based on a specific aspect. These applications are very useful to make a purchase decision. Due to limited space, aspects present in all the product models (i.e., *support* = 100%) belong to the mobile entity have been used to provide

Table 5 Performance of NMF versus pLSA

Entity name	NMF		pLSA	
	Accuracy (in percentage)	Jaccard index	Accuracy (in percentage)	Jaccard index
Mobile	80.00	0.533	70.00	0.389
Laptop	71.43	0.588	64.29	0.450
TV	76.92	0.556	69.23	0.529
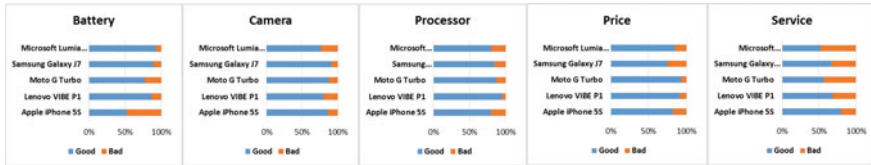
Fig. 2 Comparison of different mobile models based on extracted aspects

a comparison among different product models. This comparison can be found in Fig. 2.

5 Conclusion and Scope for the Further Work

Aspect extraction technique is a crucial part of any opinion mining application. In this paper, we have proposed an aspect extraction technique based on NMF. Three different real-life datasets have been used to measure the performance of our proposed method. Results have been validated against the ground truth and compared with another very popular aspect extraction technique namely pLSA. As you can see from Table 5, our proposed method constantly achieved more than 71% accuracy compared to the accuracy of 64% achieved by pLSA. Highest accuracy achieved by our proposed method is 80%, whereas that of pLSA is only 70%. Using Jaccard index measure, one can conclude that our NMF-based method can identify correct aspects along with very less number of incorrect aspects compared to pLSA-based method.

There are several ways in which we can extend our research works. Since NMF performs very well in the multiple topic scenario, one can use NMF to extract aspects of different products from multi-domain product reviews. Multi-domain product review dataset contains reviews on different products from product category. Each product category has been considered as a domain. It may be noted that these reviews are not labeled. This means the category of a product on which a review has been written is unknown. The limitation of our proposed system is that it can identify only *explicit aspects* (i.e., noun or noun phrases). We can continue our research work to identify *implicit aspects* (i.e., not noun or noun phrases).

References

1. Mei, Q., et al.: Topic sentiment mixture: modeling facets and opinions in weblogs. In: Proceedings of the 16th International Conference on World Wide Web. ACM (2007)
2. Lin, C., He, Y.: Joint sentiment/topic model for sentiment analysis. In: Proceedings of the 18th ACM Conference on Information and Knowledge Management. ACM (2009)
3. Li, F., Huang, M., Zhu, X.: Sentiment Analysis with Global Topics and Local Dependency, vol. 10. AAAI (2010)
4. Zhao, W.X., et al.: Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. In: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics (2010)
5. Mukherjee, A., Liu, B.: Aspect extraction through semi-supervised modeling. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers, vol. 1. Association for Computational Linguistics (2012)
6. Peng, W., Park, D.H.: Generate adjective sentiment dictionary for social media sentiment analysis using constrained nonnegative matrix factorization. Urbana 51, 61801 (2004)
7. Pauca, V.P., et al.: Text mining using non-negative matrix factorizations. In: Proceedings of the 2004 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics (2004)
8. Shahnaz, F., et al.: Document clustering using nonnegative matrix factorization. Inf. Process. Manag. 42(2), 373–386 (2006)
9. Kuang, D., Choo, J., Park, H.: Nonnegative matrix factorization for interactive topic modeling and document clustering. In: Partitional Clustering Algorithms, pp. 215–243. Springer International Publishing (2015)
10. Turney, P.D., Littman, M.L.: Measuring praise and criticism: inference of semantic orientation from association. ACM Trans. Inf. Syst. (TOIS) 21(4), 315–346 (2003)
11. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. 5(1), 1–167 (2012)
12. Lee, D.D., Sebastian Seung, H.: Learning the parts of objects by non-negative matrix factorization. Nature 401(6755), 788 (1999)
13. Turney, P.D.: Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics (2002)
14. Barman, D., Tudu, A., Chowdhury, N.: Opinion classification based on product reviews from an Indian E-Commerce website. In: Proceedings of the Second International Conference on Computer and Communication Technologies. Springer, India (2016)

Extraction of Geometric and Prosodic Features from Human-Gait-Speech Data for Behavioural Pattern Detection: Part II

Raj Kumar Patra, Rohit Raja and Tilendra Shishir Sinha

Abstract This part of the research paper emphasizes on the detection of behavioural pattern from the extracted prosodic and geometrical features using human-gait-speech data. The clusters of these above-extracted features are mapped for the detection of behavioural pattern using soft-computing technique and c-means clustering method. Here, only four features of human-gait and four features of human-speech pattern are used for the formation of clusters. These clusters are mapped to close vicinity with minimum distance measurement, which in return is helpful for the proper classification and decision process, with a positive outcome for the detection of behavioural pattern. The mapping has been done with proper mathematical analysis over each feature of human-gait-speech pattern. The four prosodic features (extracted from human-speech pattern) are speech duration, speech rate, pitch and speech momentum, whereas the four geometrical features (extracted from human-gait pattern) are step length, walking speed, energy or effort and gait momentum, which are clustered. Here, five different natural languages (Hindi, Bengali, Oriya, Chhattisgarhi and English) have been employed for the completion of this part of research, when the subject is talking while walking. The classification process is being carried out with the help of a human-gait-speech-model (HGSM), using Baye's theorem and support vector machine of artificial neural network. The mapping process has been carried out using adaptive-unidirectional-associative-memory (AUTAM) technique with an acceptable limit. The decision process for the detection of behavioural pattern has been done using revolutionary algorithm called

R. K. Patra (✉)
Department of CSE, CMR Technical Campus,
Kandlakoya(V) Medchal Road, Hyderabad, India
e-mail: patra.rajkumar@gmail.com

R. Raja
Department of CSE, Sreyas Institute of Engineering and Technology,
Nagole, Hyderabad, India
e-mail: drrohitraja1982@gmail.com

T. S. Sinha
Camellia Institute of Technology, Digberia, Madhyamgram, Kolkata, India
e-mail: tssinha1968@gmail.com

267

genetic algorithm. Three behavioural patterns have been detected with three class-based moments: happy moments, normal moments and sad moments. An algorithm, called behavioural pattern detection algorithm using human-gait-speech pattern (BPDAHGSP), has been proposed. The complexity measures have been done, and the performance of the overall authentication system has been found very helpful for promoting global biometrical security system using nominal number of features.

1 Introduction

Most of the behavioural detection based research work emphasizes on the features extraction and their matching with the template either in one-dimensional or in two-dimensional domains. Such research work has been carried out either using speech data or image data, but very few researchers have considered both one-dimensional and two-dimensional domains [1]. From the literature, it is not clear about the minimum number of features required for obtaining better performance [2–5]. In this part of the paper, human-gait-speech data has been considered for the extraction of geometrical and prosodic features. The geometrical features extracted are step length, walking speed, energy or effort and walking momentum, whereas prosodic features extracted are speech duration, speech rate, pitch and speech momentum. Since last three decades, the literature has been reviewed related to the present research work and found that image pattern based on human-gait has been used for the separation of foreground and background segments [6] with negligible data loss. Also, the work has been further progressed for the identification of multimodal physiological and behavioural biometrical traits [7] from human-gait image. Similar work has been done for the detection of behavioural trait pattern [8] using soft-computing and high-end computing technique over human-gait image frame. Similar mechanism of research has been applied, for the detection of Parkinson's disease [9] through human-speech and human-gait data. The mechanism of artificial neural networks has been applied for the recognition of emotional behaviour using human-gait patterns [10]. With reference to minimum number of features required for the detection of emotional behaviour through human-gait pattern, critical solution has been proposed [11]. The four geometrical features and four prosodic features, which were extracted from human-gait-speech pattern, have been clustered for classification process. The classification has been done after proper mapping with the corpus or knowledge-based model called human-gait-speech-model (HGSM) [12]. With the relevant statistical techniques, combined auto-cross-correlation and auto-cross-regression, the mapping process has been done for the detection of three behavioural patterns: happy moments,

normal moments and sad moments. The feature vectors of human-gait-speech pattern are clustered, and hence slope or gradient of each of the feature vectors of human-gait and human-speech data are computed. From these gradient or slope values, the behavioural classification has been done and found that linearity index property characteristics are dispersed with three different values: positive, zero and negative. The decision for behavioural pattern detection has been done with three class problems: 'positive for happy moments', 'zero for normal moments' and 'negative for sad moments'. An algorithm has been proposed called behavioural pattern detection algorithm using human-gait-speech pattern (BPDAHGSP), and the complexity has been computed under worst-case and best-case conditions. In worst case, the complexity of BPDAHGSP is '$(p + 4)/p$' and in best case, it is '$(p + 1)/p$', for 'p' total number of occurrences. In the following manner, the present paper has been organized: Sect. 2 proposes behavioural pattern detection, Sect. 3 describes the simulated outcomes and discussions and Sect. 4 gives the concluding remarks and further scope of the work and final part of the paper incorporates all the references.

2 Behavioural Pattern Detection

Let human-speech pattern and human-gait pattern be collected under the conditions stated below:

- The subject is talking while walking on a plain surface;
- Five different natural languages have been used by the subjects.

Let '${}^{Zn}S_{L1}$', '${}^{Zn}S_{L2}$', '${}^{Zn}S_{L3}$', '${}^{Zn}S_{L4}$' and '${}^{Zn}S_{L5}$' be the human-speech pattern with languages spoken in Hindi, Bengali, Oriya, Chhattisgarhi and English, respectively, by the subject 'Zn' while walking. Let '${}^{Zn}G_{L1}$', '${}^{Zn}G_{L2}$', '${}^{Zn}G_{L3}$', '${}^{Zn}G_{L4}$' and '${}^{Zn}G_{L5}$' be the human-gait pattern with above languages spoken, respectively, by the subject 'Zn' while walking, where 'n' is the subject number, which lies in between $1 \leq n \leq$ infinity.

In general, let human-speech pattern is '${}^{Zn}S_{Lm}$' and human-gait pattern is '${}^{Zn}G_{Wm}$'. So the human-gait-speech pattern '${}^{Zn}V_{WLm}$' will yields to

$$^{Zn}V_{WLm} = {}^{Zn}S_{Lm} + {}^{Zn}G_{Wm} \tag{1}$$

2.1 Clustering of Prosodic Features from Human-Speech Data

To obtain the clusters of human-speech pattern, ${}^{Zn}S_{Lm}$, into even and odd components, consider '${}^{S}Z_{T}$' known samples, where superscript 'S' represents

human-speech pattern and the subscript 'T' represents total samples. Let '$^{S}Z_E$' and '$^{S}Z_O$' be the even and odd components or clusters, respectively, for a language spoken 'L_1' and walk 'W_1'. Hence, it yields to

$$^{S}Z_T = {}^{S}Z_E + {}^{S}Z_O \tag{2}$$

So the total sample 'T' yields to summation of even sample 'E' and odd sample 'O'.

So the linearity combination of Eq (2) yields to

$$^{Z1}S_{L1} = {}^{S}Z_T = {}^{S}Z_E + {}^{S}Z_O \tag{3}$$

So, the relation for voiced and unvoiced human-speech pattern yields to the linear combination of even and odd samples for the pattern, $^{Zn}S_{Lm}$. Thus, it becomes

$$^{Zn}S_{Lm} = \rho^T \otimes {}^{S}Z_T, \tag{4}$$

where the linearity factor for total samples, ρ^T, and $\otimes$ is the convolution operator.

Now the mean μ_e for even sample N_e within the even cluster W_e and the mean μ_o for odd sample N_o within the odd cluster W_o yields to total sample mean value, μ_T,

$$\mu_T = \frac{1}{N_e} \sum X_e + \frac{1}{N_o} \sum X_o \tag{5}$$

So the projected mean points yields to the projection of μ_T,

$$\overline{\mu_T} = \frac{1}{N_T} \sum \rho^T \mu_T = \rho^T \mu_T \tag{6}$$

The diversion of the projected means on even and odd samples yields to

$$\left| \overline{\mu_e} - \overline{\mu_o} \right| = \left| \rho^T (\mu_e - \mu_o) \right| \tag{7}$$

Let $L_{in} = \{L_1, L_2, \ldots, L_n\}$ and $D_{out} = \{D_1, D_2, \ldots, D_m\}$ consisting of 'n' and 'm' input language and output codeword, respectively, which are of maximum size. For matching process, a stream of feature vectors, $B_{test\_data\_set} = \{B_1, B_2, \ldots, B_u\}$, have been compared with the values stored in human-gait-speech-model, $A_{trained\_data\_set} = \{A_1, A_2, \ldots, A_q\}$. Hence, to satisfy the linearity index condition, that is, $D_{out} = D_{in}$, adaptive-unidirectional-associative-memory, c-means clustering and genetic algorithm have been applied for the positive matching. Mathematically, the relation is

$$C_{matching} = \arg_{1 \leq q \leq n} \min\{S(D_e, D_o)\} \tag{8}$$

The above measures result in the nearest pattern of language spoken. Thus, each feature vector in the data input stream, $C_{matching}$, has been compared with the codeword in the knowledge-based corpus of the system, and the minimum average distance has been chosen as the best-fit codeword. If the unknown vector is unreachable to known vector, such situation is treated as out-of-limit (OOL) problem. Assigning weights to all the codewords in the database has minimized the OOL problem. So maximum values of the feature vector yield to

$$C_{OOL} = \arg_{1 \leq q \leq n} \max\{S(D_e, D_o)\} \tag{9}$$

The absolute difference, C_{DIFF}, is the cropped pattern and yields to

$$C_{DIFF} = |C_{OOL} - C_{matching})| \tag{10}$$

Dividing Eq. (8) by Eq. (10), it yields to C_{CMR}, where CMR means cluster matching rate is

$$C_{CMR} = \frac{C_{matching}}{|C_{DIFF}|} \tag{11}$$

Similarly, speaking rate has been also computed. Speech duration has been computed using Euclidean distance measure of voiced and unvoiced part of the human-speech pattern. The speech momentum is calculated and the mathematical relation is

$$\alpha_{speech\_momentum} = \prod_{i=0}^{i=n} \alpha_i \quad S_{dur}, \tag{12}$$

where S_{dur} is the speech duration and α_i is the speech learning rate ranging from 0 to 1.

2.2 Clustering of Geometrical Features from Human-Gait Data

To obtain the clusters of human-gait pattern, $^{Zn}G_{Wm}$, into even and odd components, consider '$^{G}Z_T$' known samples, where superscript 'G' represents human-gait pattern and the subscript 'T' represents total samples. Let '$^{G}Z_E$' and '$^{G}Z_O$' be the even and odd components or clusters, respectively, for a human walk 'W_1'. Hence, it yields to

$$^{G}Z_T = {}^{G}Z_E + {}^{G}Z_O, \tag{13}$$

where the total sample 'T' is the summation of even sample 'E' and odd sample 'O'.

So the linearity combination of above equation yields to

$$^{Z1}G_{W1} = {}^{G}Z_T = {}^{G}Z_E + {}^{G}Z_O \tag{14}$$

So, the relation for energy and non-energy value of human-gait pattern yields to the linear combination of even and odd samples for the pattern, $^{Zn}G_{Wm.}$ Thus, it becomes

$$^{Zn}G_{Wm} = \rho^T \otimes {}^{G}Z_o, \tag{15}$$

where the linearity factor for total samples, ρ^T, and $\otimes$ is the convolution operator.

Equations (5) through (11) are used for the computation of the mean, and the walking rate of the human-gait pattern.

Step length has been computed using Euclidean distance measure of all odd frames and all even frames of the human-gait pattern. Now, the gait momentum has been calculated, and the mathematical relation is

$$\beta_{\text{gait\_momentum}} = \prod_{i=0}^{i=n} \beta_i \quad G_{step\_length}, \tag{16}$$

where G_{step_length} is the step length and β_i is the gait learning rate ranging from 0 to 1.

2.3 Gradient or Slope of Human-Gait-Speech Data

Let $^{m}\nabla_{PGF}$ be the slope or gradient of the human-gait-speech pattern, where the subscript PGF means prosodic geometrical features: 'duration', 'rate', 'effort' and 'momentum'. The superscript 'm' is the number of prosodic geometrical feature. Let $^{m}\nabla_{PF}$ and $^{m}\nabla_{GF}$ be the slope or gradient of human-speech and human-gait data, respectively. So mathematically, the relation becomes

$$^{m}\nabla_{PGF} = {}^{m}\nabla_{PF}/{}^{m}\nabla_{GF} \tag{17}$$

If $^{m}\nabla_{PGF}$ is the slope for 'duration' count, and $^{m}\nabla_{PGF}$ is the slope for 'rate' count, and $^{m}\nabla_{PGF}$ is the slope for 'effort' count, and $^{m}\nabla_{PGF}$ is the slope for 'momentum' count, then Eq. (17) becomes

$$^1\nabla_{\text{duration}} = {}^1\nabla_{\text{step\_length}} / {}^1\nabla_{\text{speech\_duration}} \tag{18}$$

$$^2\nabla_{\text{rate}} = {}^2\nabla_{\text{walking\_speed}} / {}^2\nabla_{\text{speech\_rate}} \tag{19}$$

$$^3\nabla_{\text{effort}} = {}^3\nabla_{\text{energy}} / {}^3\nabla_{\text{pitch}} \tag{20}$$

$$^4\nabla_{\text{momentum}} = {}^4\nabla_{\text{gait\_momentum}} / {}^4\nabla_{\text{speech\_momentum}} \tag{21}$$

The values for $^4\nabla_{\text{gait\_momentum}}$ and $^4\nabla_{\text{speech\_momentum}}$ are calculated from Eqs. (16) and (12), respectively. Similarly, the values for $^3\nabla_{\text{energy}}$ and $^3\nabla_{\text{pitch}}$ are calculated from Eqs. (15) and (4), respectively. Considering odd and even frames or samples of data, the differences are measured and hence employed in Eqs. (15) and (4), respectively, results in $^2\nabla_{\text{walking\_speed}}$ and $^2\nabla_{\text{speech\_rate}}$, values.

The algorithm for detecting the behavioural pattern has been depicted below:

Algorithm BPDAHGSP (human-gait-speech data)

Step 1.   GEN: Initial data stream $X(o) = (X_1, X_2, …, X_M)$ and $Y(o,p) = (Y_1, Y_2, …, Y_N)$
Step 2.   READ: Corpus(Size Q) and set the counter, say, $q = 0$
Step 3.   DO WHILE $(q \le Q)$
          GEN: An intermediate data stream $X_r(t)$ & $Y_r(t)$
          GENNEXT: State data stream $X(t+1)$ and $Y(t+1)$ for $X_r(t)$ & $Y_r(t)$
          COMPUTE: WEIGHT for Linearity Index condition, INPUT ==
     OUTPUT COMPUTE: Gradient or Slope for {prosodic geometrical feature: HGSM}
     INR: Increment the duration or length, $t = t + 1$ and the counter, $q = q + 1$
     FITNESS_TEST: $f(X_i)$ & $f(Y_i)$ of each data stream X_i & Y_i
     MAPPING: AUTAM (spch_gait_val, spch_gait_code)   else
     IF (TRUE) THEN
       {BEHAVIOUR: Display 'Acceptable placed in category NOT BOL'}
     ELSE {BEHAVIOUR: Display 'Not Acceptable placed in category
       BOL'} and Display 'Data is OOL (Out-Of-Limit)'}
     ENDDO

Complexity measures: In worst case, let the total number of samples be represented by 'p'. As the mapping process is done for four features, so the execution time is being added up. Thus, the complexity is (loop execution/ total number of occurrences). In worst case, loop execution is 'p + 4'. So the complexity measure in worst case is '(p + 4)/p'. Similarly, in best case, the minimum feature required for mapping process is one, so the execution time is being added up. Thus, the complexity measure in best case is '(p + 1)/p'.

3 Simulated Outcomes and Discussions

The simulations of the human-gait-speech patterns have been done and are clustered using c-means clustering method and also the auto-cross-correlation and regression are graphically plotted, which are shown in Figs. 1 and 2.

Correlation is performed for the similarity and difference pattern matching, whereas regression is performed for the computation of weights with feedback mechanism and hence resulting in a value with optimal outcome and also in the improvement of the performance of the overall detection system.
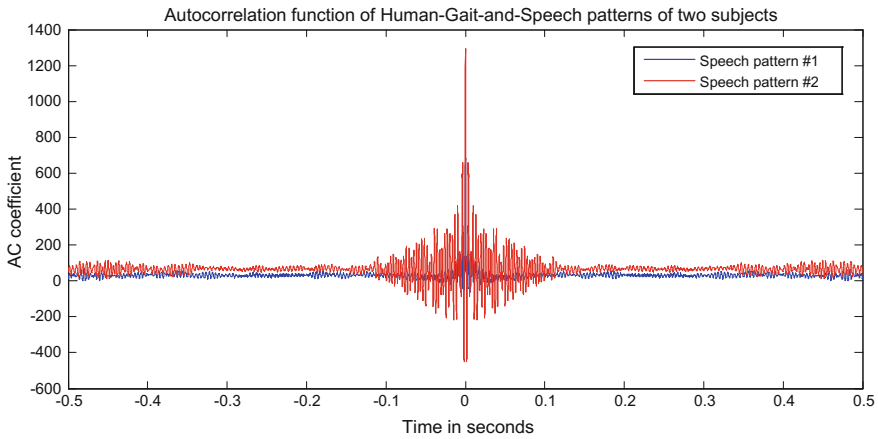
Fig. 1 Auto-cross-correlation and regression of human-gait-speech pattern
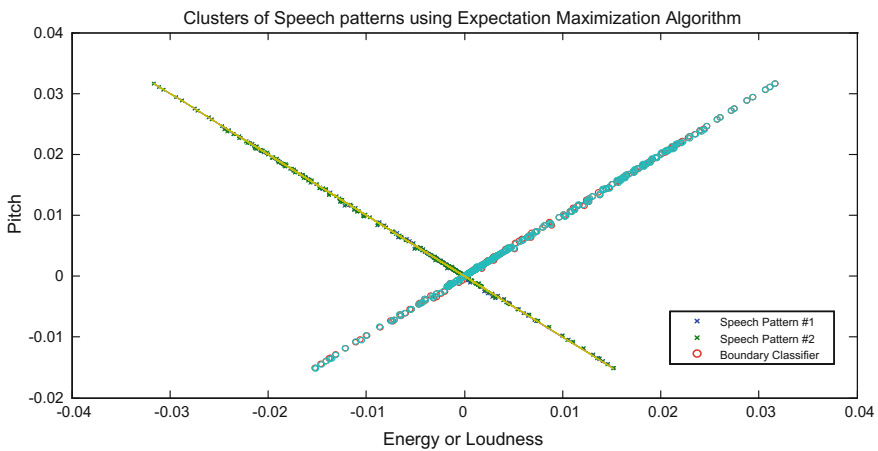
Fig. 2 Clusters of human-speech pattern
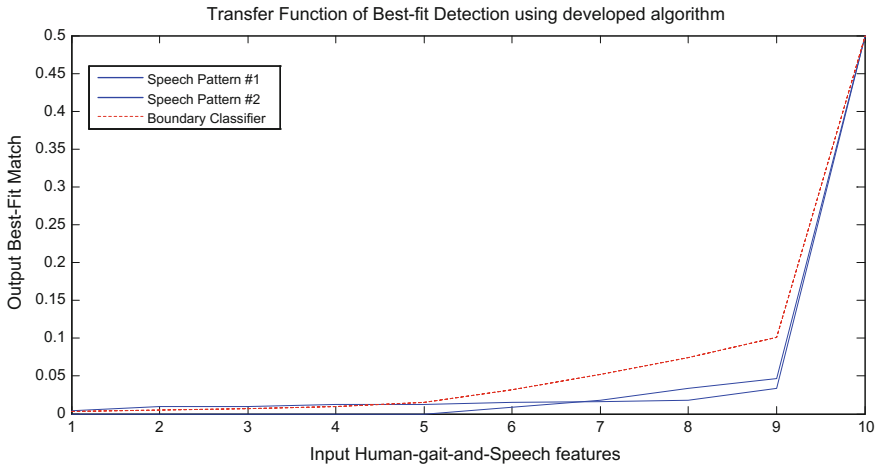
Fig. 3 Transfer function of best-fit detection using BPDAHGSP
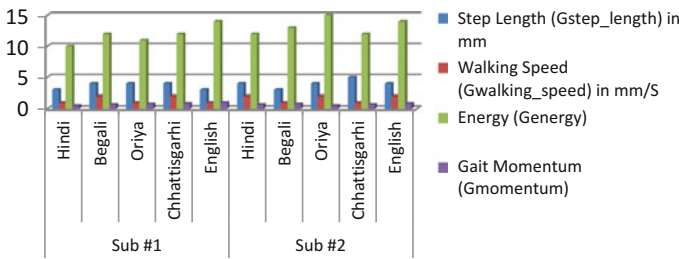
Fig. 4 Graphical representation of geometrical features of human-gait pattern

The relevant transfer function for the best-fit detection using the proposed algorithm BPDAHGSP is shown in Fig. 3.

Figure 4 shows the graphical representation of the geometrical features extracted from human-gait pattern.

Similarly, Figs. 5, 6 and 7 show the graphical representation of different prosodic and geometrical features from human-gait-speech pattern.

Tables 1 and 2 shows the relevant prosodic and geometrical features extracted from human-gait-speech pattern.

Table 2 shows the data in millimeter, which are actually extracted in pixels. The conversion formula that has been used is

$$mm\_value = (pixel\_value * 25.4)/dpi, \tag{22}$$

where dpi $= 96$
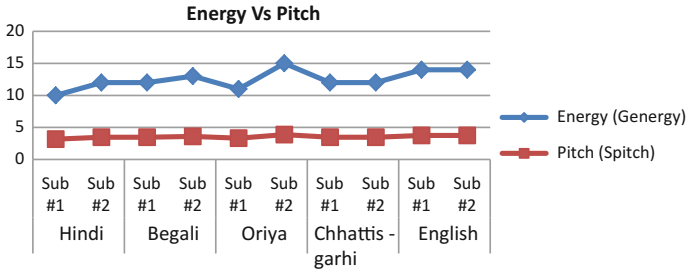
Fig. 5 Graphical representation of energy and pitch of human-gait-speech pattern
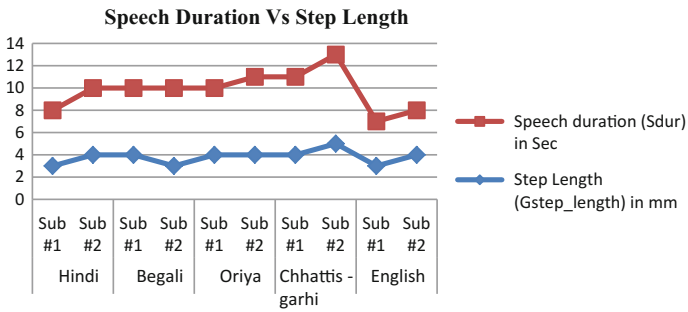
Fig. 6 Graphical representation of speech duration and step length of HGS
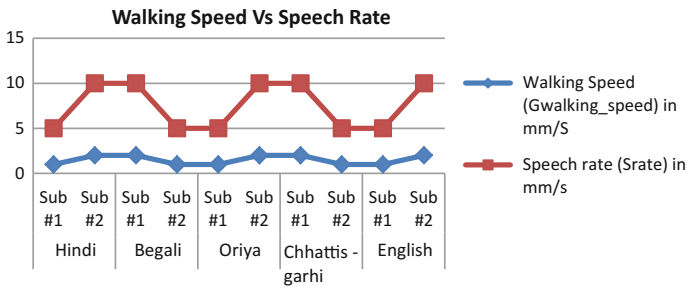
Fig. 7 Graphical representation of walking speed and speech rate of HGS

Table 1 Prosodic features extracted from human-speech data

Subjects	Gender and age	Speaking language pattern	Speech duration (S_{dur}) (s)	Speech rate (S_{rate}) (mm/s)	Pitch (S_{pitch})	Speech momentum ($S_{momentum}$)
Sub #1	Male 30 years	Hindi	5	5	3.16	0.5
		Bengali	6	10	3.46	0.7
		Oriya	6	5	3.31	0.8
		Chhattisgarhi	7	10	3.46	0.9
		English	4	5	3.74	1.0
Sub #2	Male 31 years	Hindi	6	10	3.46	0.7
		Bengali	7	5	3.6	0.8
		Oriya	7	10	3.87	0.5
		Chhattisgarhi	8	5	3.46	0.7
		English	4	10	3.74	0.9

Table 2 Geometrical features extracted from human-gait data

Subjects	Gender and age	Speaking language pattern	Step length (G_{step_length}) (mm)	Walking speed ($G_{walking\_speed}$) (mm/s)	Energy (G_{energy})	Gait momentum ($G_{momentum}$)
Sub #1	Male 30 years	Hindi	3	1	10	0.5
		Bengali	4	2	12	0.7
		Oriya	4	1	11	0.8
		Chhattisgarhi	4	2	12	0.9
		English	3	1	14	1.0
Sub #2	Male 31 years	Hindi	4	2	12	0.7
		Bengali	3	1	13	0.8
		Oriya	4	2	15	0.5
		Chhattisgarhi	5	1	12	0.7
		English	4	2	14	0.9

4   Concluding Remarks and Further Scope of the Work

In the present research work, three behavioural patterns happy moments, normal moments and sad moments have been classified. The above classification has been done based on four prosodic features and four geometrical features, which were extracted from human-gait-speech pattern. These features were clustered and a proper mapping process has been applied for the decision-making regarding the detection of behavioural patterns. The mapping process has been done using auto-cross-correlation and auto-cross-regression techniques. The slope or gradient of each of the feature vectors of human-gait and human-speech data has been computed. This slope or gradient values when plotted give linearity index property characteristics. The dispersion of the gradient

or slope has been calculated which results in +1, 0 and −1 values. The decision for behavioural pattern is +1 for happy moments, 0 for normal moments and −1 for sad moments. An algorithm has been proposed called behavioural pattern detection algorithm (BPDA). In the worst case, the complexity of the proposed algorithm has been computed as '(p + 4)/p', in the best case, the complexity of the proposed algorithm has been computed as '(p + 1)/p', for 'p' total number of occurrences.

The present work may be further applied in a human–computer interaction environment for the detection of human–brain signals and their analysis, along with human-gait-speech pattern, which will be a tri-modal biometric security system. The work may be further carried out for the detection of various health problems related to breathing, walking, speaking, neurological and heart. Also, it can be applied for promoting global multimodal biometrical-based security system.

References

1. Pantic, M., Vinciarelli, A.: Implicit human-centered tagging. IEEE Sig. Process. Mag. **26**(6), 173–180 (2009)
2. Boersma, P., Praat A System for Doing Phonetics by Computer. Glot Int. **5**(9/10), 341–345 (2001)
3. Bezooijen, V.R., Otto, S.A., Heenan, T.A.: Recognition of vocal dimensions of emotion: a three nation study to identify universal characteristics. J. Cross Cult. Psychol. **14**, 387–406 (1983)
4. Juslin, P.N., Laukka, P.: Communication of emotions in vocal expression and music performance: different channels, same code? Psychol. Bull. **129**(5), 770–814 (2003)
5. Ng, A.K., Koh, T.S., Thng, C.H.: A level–wavelet dependent scheme for Image denoising via undecimated wavelet transform. In: Signal and image processing, pp. 576–586 (2007)
6. Cheng, L., Gong, M., Schurmans, D., Calli, T.: Real–time discriminative background subtraction. IEEE Trans. Image process. **20**(5), 1401–1414 (2011)
7. Hossain, S.M.E., Chetty, G.: Human identity verification by using physiology and behavioral biometric traits. Int. J. Biosci. Bioinform. **1**(3), 199–205 (2011)
8. Sinha, T.S., Chakravarty, D., Patra, R.K., Raja, R.: Modelling and simulation for the recognition of physiological and behavioral traits through human gait and face images. In: Awad Kh. Al-Asmari (ed) Discrete Wavelet Transforms—A Compendium of New Approaches and Recent Applications. InTech Open, pp. 95–125 (2013) ISBN 978-953-51-0940-2. https://doi.org/10.5772/52565
9. Cantiniaux, S., Vaugoyeau, M., Robert, D., Horrelou-Pitek, C., Mancini, J., Witjas, T., Azulay, J.P.: Comparative analysis of gait and speech in Parkinson's disease: hypokinetic or dysrhythmic disorders? J. Neurol. Neurosurg. Psychiatry **81**(2), 177–184 (2009)
10. Janssen, D.: Recognition of emotions in gait patterns by means of artificial neural nets. J. Nonverbal Behav. **32**, 79–82 (2008)
11. Roether, C.L., Omlor, L., Christensen, A., Giese, M.A.: Critical features for the perception of emotion from gait. J. Vis. **9**(6), 1–32 (2009)
12. Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviours. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **34**, 334–352 (2004)

Mutation in Path for the Packets in the Network During Journey from Source to Destination

Tarak Nath Paul and Abhoy Chand Mondal

Abstract This paper formulates a genetic algorithm to judge the mutation of path for packets in the network. The predefine routes for network lines to transfer packets, transportation links in liquids and bearer links are static phenomenon. The algorithm discovers mutation in prescribed path due to some breakage in the communication link. The breakage occurs due to natural calamity or for any technical issues. At this juncture in the mid-way dumping or returning, the packet/ package is not wise. Here, a new path has to be redefined from the current node and the packet has to be delivered to the destination. It is good to have delay in delivery instead of been dumping. Because resending the packet is far more time consuming, this makes the phenomenon more dynamic and less time-consuming. The concept is also useful in robotics for path selection. Simulation for path is performed using MATLAB. Results are similar to Dijkstra's algorithm. Dijkstra's algorithm is a static algorithm, whereas the proposed algorithm is dynamic one.

Keywords Selection · Chromosome · Mutation · Crossover
Routing and shorter paths

1 Introduction

There are many possible paths from source to destination before the journey begins. Each possible path is treated as chromosome. After commencement of the journey, it was found that the path is not valid due to some unavoidable disturbance. It can

T. N. Paul (✉) · A. C. Mondal
Department of Information Technology, ICFAI University Jharkhand, Ranchi,
Jharkhand, India
e-mail: paultarak@gmail.com

A. C. Mondal
e-mail: abhoy_mondal@yahoo.co.in

T. N. Paul · A. C. Mondal
Department of Computer Application, The University of Burdwan, Burdwan, India

be natural or technical disaster. At that point of time, it is wise to deliver the packet instead of dumping the packet. A new modification is required in the path to reach the destination and the required modification is termed as "Mutation". Mutation brings dynamic properties in searching shorter path from source to destination. Genetic algorithm is successfully implemented to find the mutated path in the given network [1–3].

In genetic algorithm, the chromosomes encode individuals to an optimization problem that evolves towards more suitable solution. The evolution basically starts from randomly selected fit individuals into a population. Then, they are fused to reproduce a new generation. Fitted individuals are retained, and unfitted individuals are discarded from the new generation. Then, the new generation has been used for the next iteration. It stops when maximum of the generation is reached or a desired fitted population is achieved [4, 5].

Dijkstra's algorithm finds the shortest path between the two desired nodes in a graph. It also helps in finding the shortest path from a given node to all other nodes in the graph. Originally, Dijkstra's algorithm does not use a min-priority queue and executed in the time bound of $O(|V|2)$. The idea of this algorithm was given by Leyzorek in 1957. The implementation is done by min-priority queue [6].

2 Genetic Algorithm

Genetic algorithm consists of basic four steps, namely initialization, selection, reproduction, and termination. In initialization, individual solutions are randomly generated with respect to the problem domain. During selection, a portion of the existing population is selected with the help of fitness function to breed a new generation. In the process of reproduction, crossover and mutation are implemented with the help of chromosomes. For termination, a terminating condition must be reached. The terminating condition may be number of generation allowed or a targeted population is achieved [7, 8].

3 Dijkstra's Algorithm

Dijkstra's algorithm finds the shortest by weight precedence. It solves the shortest path problem with nonnegative values. This algorithm is a greedy algorithm similar to Prim's algorithm. The algorithm begins at the source node and grows a tree structure that shortest path for all nodes [9, 10]. Dijkstra's algorithm keeps two sets of vertices: S → the set of vertices whose shortest paths from the source have already been determined and V–S → the remaining vertices. The other data structures needed are d—array of best estimates of shortest path to each vertex. pi—an array of predecessors of each vertex. The basic mode of operation is as follows: Initialize d and pi, and set S to empty, while there are still vertices in V–S,

i. Sort the vertices in V–S according to the current best estimate of their distance from the source,

ii. Add u, the closest vertex in V–S, to S,

Relax all the vertices still in V–S connected to u.

4 Proposed Algorithm

A new hybrid algorithm for routing problem with genetic algorithm approach is presented when path gets disrupted. The predefined and guided route has decided on the basis of leveling of graph (modify the existing graph into leveled graph). Checking and testing are based upon genetic evolution of the chromosomes. Chromosomes represent the shorter paths. Paths are of variable length. Chromosomes and their genes are used to encode all the possible paths. Fusing chromosomes produces the new one which generates the new shorter path as a feasible solution. During crossover partial exchange of chromosomes occurs at a point which is common, with the help of Geospatial Information System, a graph is considered of 20 cities shown in Fig. 1a. A new reduced graph is created with less number of nodes. To traverse lesser nodes, less time is required. Less time symbolizes high performance.
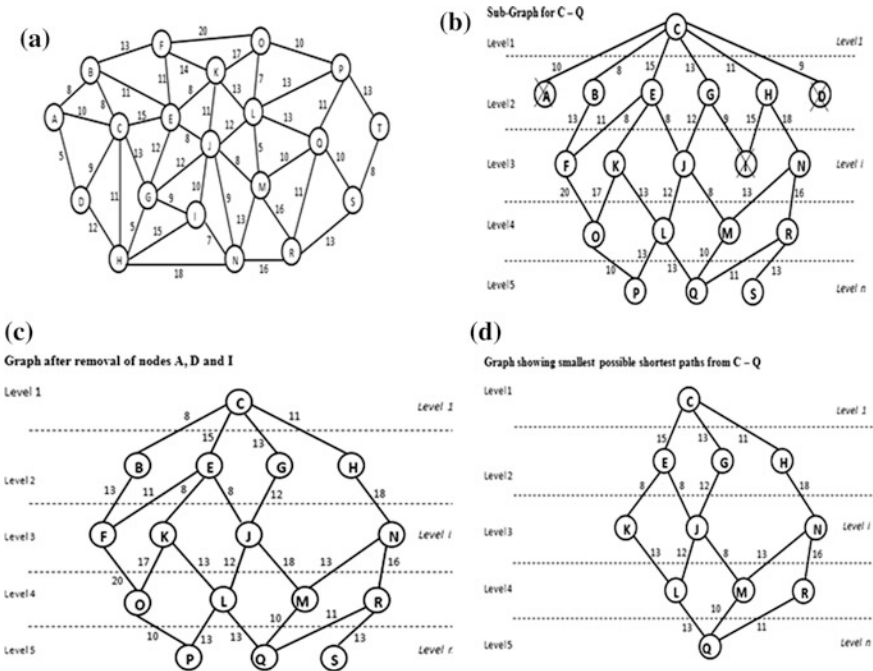
Fig. 1 Graphical representation of nodes **a** Main-graph (Source = C and Destination = Q) **b** after implementing the sub-graph algorithm **c** after deleting the redundant nodes **d** Sub-graph representing minimum nodes

```
  Algorithm to create of Sub-Graph from the Main-Graph
1  Choose Source Node. TSource = Source…….[TSource = Tem-
porary Source]
2  Create First Level
3  Place TSource in the Level
4  Find connecting nodes to TSource and Add it to Queue
5  If all Nodes of the Same Level Traversed: TSource =
Next un - Traversed node in the Same Level, Goto Step 6:
else Goto Step 3
6  Create the next Level (Level++)
7  Add the Queue to the Level
8  If Destination Found: Goto Step 9: else   TSource =
Next un - Traversed node in the Same Level
9  End
```

In the above algorithm, TSource is considered as temporary source. First, choose source from where the journey begins and the destination it ends. A level is created and the source is placed in the first level and the connected nodes are placed in the next level. Then, the next level performs the same operation with every node. Carry on the process until the destination does not reach by all nodes in the last level. Algorithm to create sub-graph from the main-graph as shown in Fig. 1b is evaluated. Dead nodes named as A, D and I are present in the graph. Nodes in dead end are called dead nodes. So, dead nodes are not required in the present graph. After discarding the dead nodes, Fig. 1c is evaluated.

```
  Back Track Algorithm (Elimination Algorithm)
1  Select the nodes of the Queue of the nth Level
2  TDestination = First node of the Queue…[TDestinaiton =
Temporary Destination]
3  If TDestinaiton = Destination: Goto Step 10:else  Goto
Step 4
4  Get the predecessor node of TDestination
5  Delete TDestination
6  TDestination = Predecessor Node
7  If TDestination = Path Node…[Path Node = node which
lies between Source & Destination of a defined Path]:Goto
Step 8: else Goto Step 4
8  TDestination = Next Node in the Queue
9  Goto Step 3
10 If Queue = Empty: Goto Step 11: else Goto Step 8
11 End
```

After implementation of the above algorithms, the graph reduces into levels with lesser number of nodes called leveled graph. With the help of backtrack algorithm, all the undesired nodes are in the nth level and other connected nodes which are not responsible for the formation of path are deleted. First, select all the nodes in the nth level and store them in a queue. And then check every node and its predecessor node for survival. Node other than destination cannot survive in the nth level. The node which is a predecessor of a non-survival node at nth level and a non-participate in the formation of path will not survive. This phenomenon will go on until the non-participating gets eradicated. The resultant is a reduced leveled graph shown in Fig. 1d. This sub-graph is responsible to provide possible path from source to destination. The set will also contain the shortest path. If the shortest path is damaged, then the path will mutate during execution. This will increase the efficiency of the proposed work.

In "possible paths from source to destination algorithm", few data structure and variables are used—Queue, Stack, Source, TSource (Temporary Source), Destination, and Count. Queue is used to track, traverse, and store the path. When any node is visited, the respective edge is locked. All possible paths from source node C to destination node D are shown in Fig. 2a. Considering two chromosomes (3 and 4) in Fig. 2b as parent and fusion is implemented on it and two chromosomes (2 and 5) in Fig. 2c received as daughter chromosomes.
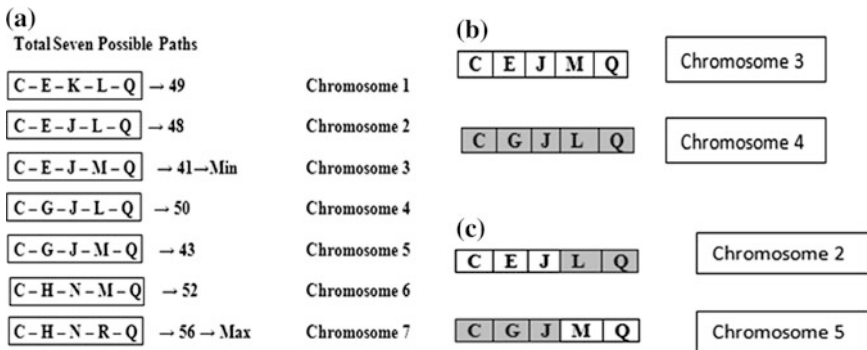
Fig. 2 **a** Showing parent chromosomes (possible path algorithm), and **b** showing the daughter chromosomes, **c** offspring chromosomes

```
  Algorithm to search every possible paths form Source to
Destination.
1  Take Stack, Source, Destination, Count
2  TSource = Source
3  Count = Number of unlocked Connecting Nodes
4  If Count > 1: Goto Step 5: else  Goto Step 8
5  Push TSource to Stack
6  Count--
7  Goto Step 4
8  If TSource has unlock edge: Goto Step 9: else Goto
   Step 11
9  Fetch the Node (Traverse the edge)
10 If Destination Found: Goto Step 11: else Goto Step 13
11 Print the path (Print the Queue)
12 If Stack Empty
13 TSource = TSource → Next
14 Goto Step 3
15 Pop the Stack
16 If Source: Goto Step 17: else Goto Step 20
17 New Queue
18 TSource = Source
19 Goto Step 3
20 New Queue
21 TSource = Popped Node
22 Copy Source to Popped Node to New Queue
23 If all node traversed: Goto Step 24 else Goto Step 3
24 End
```

5   Mutation

Mutation occurs when the journey of an object has started from source to desti-
nation. Then, it has found that the path is ruptured which is predefined by the
proposed algorithm. Then, the question arise, how to reach the destination? Because
the chromosome which defines a healthy individual is no healthier, it is an unfit
body in the pool. A new modification is required in the path to reach the destination
and that modification is termed as "Mutation". Mutation means extrachromosomal
genetic change of element, in the code of the chromosome. Mutation brings
dynamic properties to the proposed algorithm. In the proposed work, a possible
path is represented by the chromosome and the code is which node or vertex comes
after the other. When the transmission is in process, it is a costlier process to return
back from the node, after which the edge is broken. At this point of time, we cannot
stop the process; we need to deliver the packet to the destination. A disrupted edge
means a bifurcated way need to be chosen that will increase the processing time,
execution time, and delivery time. So mutation is a costlier process.

```
  Mutation on the path
1  Get the base node of the disrupted path
2  Search for a connection that reaches to the next level
3  If (find connection) Goto 8 else Goto Step 4
4  Search for a connection that lies in the same level
5  If (find connection) Goto 8 else Goto Step 6
6  Get back to the previous
7  If (it is source node): Communication Fail, If Commu-
   nication Failed: Goto 27 else Goto Step 9: else Goto
   Step 2
8  Take Stack, Source, Destination, Count
9  TSource = Source
10 Count = Number of unlocked Connecting Nodes
11 If Count > 1: Goto Step 5 else Goto Step 8
12 Push TSource to Stack
13 Count--
14 If Count = 0: Goto Step 4 else Goto Step 16
15 If TSource has unlock edge: Goto Step 9 else: Goto
   Step 11
16 Fetch the Node (Traverse the edge)
17 If Destination Found: Goto Step 11 else: if count = 0:
   Goto 18 or else Goto Step 13
18 Print the path (Print the Queue)
19 If Stack Empty: TSource = TSource → Next, Goto Step 3
20 Pop the Stack
21 If Source: Goto Step 17 else if count = 0: Goto 22 or
   else Goto Step 19
22 If TSource = Source: Goto Step 3 else Step 23
23 New Queue
24 TSource = Popped Node
25 Copy Source to Popped Node to New Queue
26 Goto Step 3
27 End
```

6 Results of Mutation

Mutation is purely dependent on the availability of path from the base node. Instead of dumping the packet, it can be delivered to the destination. The plus sign represents the extra distance covered by the packet during mutation. This depends on the weight of the edge. It is never more than the weight of the path. Keeping this in mind, there are possibly three cases available in the proposed work shown in Table 1.

Table 1 Different cases of mutation

Cases	Possibility	Defective edge	Mutated path	Values
Case I	Base node got two descendent connections and n super node connection	J–M	J–L–Q	48+
Case II	Base node got only one descendent connection and one super node connection	K–L	E–J–M–Q or E–J–L–Q	41+ or 48+
Case III	Base node got only one descendent connection and two super node connections	L–Q	J–M–Q or K–E–J–M–Q	41+ or 43+ or 65+
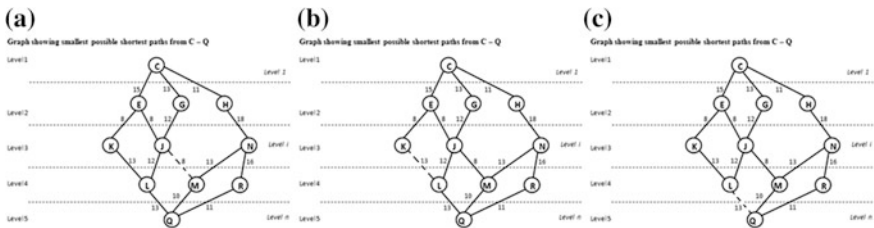
Fig. 3 Different cases of disruption of path. **a** J–M is disrupted, **b** K–L is disrupted, **c** L–Q is disrupted

Case I: In this case, edge J–M is disrupted, as J is the base node and having another connection to the next level (i.e., J–K) which will be used for the transfer of the packet. Using this path J–L, the packet can be delivered to the destination note successfully. The mutated path will be J–L–Q as shown in Fig. 3a.

Case II: In this case, edge K–L is disrupted, as K is the base node and having no connection to the next level. So, it will go back to the previous level (i.e., node E) and find shorter path to the destination. The mutated path will be E–J–M–Q or E–J–L–Q as shown in Fig. 3b.

Case III: In this case, edge L–Q is disrupted, as L is the base node and having no connection to the next level. So, it finds two nodes in its previous level (i.e., K and J). J got connection in the next level; mutation takes place here and the possible path will be J–M–Q. K does not got a connection to the next level so it will go back to its previous level (i.e., E). E got a connection to the next level; again mutation will take place here and the possible path will be K–E–J–M–Q as shown in Fig. 3c.

7 Complexity and Comparative Study

In the most worst case, it can reach the destination vertex as many as n levels, and n is the maximum number of vertex in the graph, if and only if all vertexes lies in a straight line or in a series. The complexity runs up to O (n log n). If the solution space

Fig. 4 Implementation of
Dijkstra's algorithm

$$[\text{C} - \text{E} - \text{J} - \text{M} - \text{Q}] \rightarrow 41$$

need only a set of m nodes, then complexity runs up to O (m log m), where m < n. If Dijkstra's algorithm is implemented in the graph provided, equivalent results have been found. But dynamicity is not feasible in Dijkstra's algorithm (Fig. 4).

The resultant path of Dijkstra's algorithm is being disclosed by the proposed algorithm, that is, chromosome number 3 (Three). Dijkstra's algorithm discusses only the shortest path from source to destination but the proposed algorithms not only discuss the shortest and shorter path but also discuss the dynamic path. If the shortest path is congested, then solutions are provided with other shorter paths, those paths can be availed. Even in crisis situation, mutated paths can be availed.

8   Conclusions

The proposed work is eligible to take runtime decision on finding possible path from source to destination. During travel due to some natural calamity or technical error occurrence, the predefined course is redefined, termed as mutation in the path. This discovers the mutated path from source to destination and also encounters the alternative shorter paths for any adverse condition. If unfortunately the path gets disrupted or disconnected, then it is necessary to pass the traffic by other paths to the destination without taking much time. With the help of the proposed algorithm, a disrupted not defined destination is being achieved or reached with the same complexity.

References

1. Paul, T.N., Mondal, A.C.: Amalgamation of graphs do not affect the search of the set of shorter paths algorithm. Int. J. Adv. Comput. Res. **4**(14), 184–192 (2014)
2. Paul, T.N., Mondal, A.C.: Search the set of shorter paths using graph reduction technique. Int. J. Adv. Comput. Res. **3**(13), 278–287 (2013)
3. Paul, T.N., Mondal, A.C.: Set of shorter paths and network delay deduction using genetic algorithm. In: IEEE ISACC-2015, Assam University
4. Paul, T.N., Mondal, A.C.: Constitute a sub-graph with n-levels to search a set of shorter paths using genetic algorithm. In: IEEE International Conference on Intelligent Computing and Intelligent System, 3-0096-10943, vol. 3, pp. 452–456, Nov 2011
5. Paul, T.N., Mondal, A.C.: Intelligent traversing method using genetic algorithm for long distance. In: International Conference on Computing and System, pp. 69–74, Nov 2010
6. Paul, T.N., Mondal, A.C.: Faster genetic algorithm for traverse large distance. In: National Conference on Computing and System, pp. 56–59, Jan 2010

7. Paul, T.N., Mondal, A.C.: Shorter paths from source to destination. In: COGNITION, National Conference on Innovation in Computing and Information Technology, 25–26 Feb 2015, Amity University Gurgaon

8. Bramlette, M.F., Cusin, R.: A comparative evaluation of search methods applied to parametric design of aircraft. In: Proceeding of ICGA 3, pp. 213–218 (1989)

9. Paul, T.N., Mondal, A.C.: Combination of graph does not affect the execution of finding group of shorter path algorithm. In: Second International Conference on Computing and Systems, pp. 348–356, Sept 2013

10. Xu, M.H., Liu, Y.Q., Huang, Q.L., Zhang, Y.X., Luan, G.F.: An improved Dijkstra's shortest path algorithm for sparse network. Appl. Math. Comput. **185,** 247–254 (2007)

Facial Expression Recognition Using 2DPCA on Segmented Images

Dewan Imdadul Islam, S. R. Ngamwal Anal and Aloke Datta

Abstract In this article, a facial expression recognition technique has been proposed. In this method, initially, a face image is segmented into different sub-images like left eye, right eye, mouth, nose, etc. Then, two-dimensional principal component analysis (2DPCA)-based transformation is performed on each sub-image separately to extract the features. Lastly, classification operation is made to categorize the expression of faces. To demonstrate the effectiveness of the proposed method, results are compared with other existing PCA- and 2DPCA-based methods in terms of overall classification accuracy. Results for the proposed method are found to be encouraging.

Keywords Facial expression recognition · Two-dimensional principal component analysis (2DPCA) · Image cropping · Minimum distance classifier

1 Introduction

Facial expression recognition has been active research field since many years. Facial expression recognition system is a computer application to recognize the expression, i.e. emotion of a person from face images. There are seven universal facial expressions, i.e. surprise, happy, fear, anger, sad, disgust and neutral. The applications of facial expression recognition technique are in the field of security, medical purpose, computer entertainment, human–computer interaction, etc.

There are three main modules for facial expression recognition system and these are image preprocessing, features extraction and classification. In image preprocessing step, acquisition of face images and different types of preprocessing task are performed to make these face images into good quality for further analysis. In features extraction module, extraction of some valuable features from the images is made, so

D. I. Islam · S. R. N. Anal · A. Datta (✉)
Department of Computer Science and Engineering, National Institute
of Technology Meghalaya, Shillong, India
e-mail: alokedatta@nitm.ac.in

that recognition task is performed depending on those features. In the last module, the features are analysed to recognize the facial expression.

There are various methods to recognize the facial expression like Gabor filter selection based [1, 2], linear discriminant analysis based [3–5], principal component analysis (PCA) based [6–9], two-dimensional principal component analysis (2DPCA) based [10, 11], etc. Principal component analysis (PCA) and 2DPCA are most popular methods in this approach. PCA is based on one-dimensional matrix. So in PCA, image matrix needs to transform into 1D vector before calculating the covariance matrix. 2DPCA is based on 2D image matrix. So in 2DPCA, covariance matrix can be calculated directly from image matrix. Eigenvector and eigenvalues are calculated from covariance matrix. PCA- and 2DPCA-based methods find the covariance of overall dataset based on whole image. 2DPCA has some advantage over PCA [10], and these advantages are easy to calculate the covariance matrix accurately, identify the particular eigenvector quickly, low computational cost and good recognition rate.

In the proposed method, main concentration is on partitioning the whole image into small sub-images and then perform 2DPCA over small sub-images, because it gives more significant information as it focuses on smaller units, i.e. local details. Classification is used for finding the expression of an image. 2DPCA has more recognition rate than PCA. Japanese female facial expression (JAFFE) dataset has been used in experiment. Experimental result shows that 2DPCA gives more recognition rate when applied on face parts instead of entire face.

The remainder of this article is organized as follows. The proposed scheme is outlined in Sect. 2. Section 3 contains experimental details and result analysis. Conclusive remarks are given in Sect. 4.

2 The Proposed Method

In the proposed method, 2DPCA-based transformation for extracting features is performed and then minimum distance classifier is used to recognize the face expressions. 2DPCA is not applied directly on face images but over segmented sub-images. The block diagram for proposed system is shown in Fig. 1.

The main three steps are image preprocessing, feature extraction and classification. Feature extraction contains two sub-steps segmentation into sub-images and 2DPCA-based transformation. The details of each step are discussed as follows.
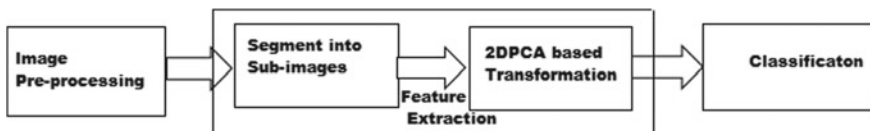
Fig. 1 Structure of proposed facial expression recognition system

2.1 Image Preprocessing

All input images are greyscale image. All image sizes are same. Seven different expressions of images are taken. Faces are detected from the images using Viola–Jones face detection algorithm [12]. The tasks involved in image preprocessing are acquisition of face images, removal of different types of noises, enhancement of image and registration of images. It is assumed that face images are properly preprocessed in the proposed method.

2.2 Feature Extraction

Features are the very important information of images. There are huge number of features in images. There are also some unwanted features in images. So features extraction is required to extract important features from the images [13]. In this step, features are extracted from each face image and perform comparison between faces based on these extracted features.

In the proposed method, the main aim is to extract the meaningful features from faces. For this purpose, each face image is divided into sub-images and then 2DPCA-based transformation is performed to extract features. Here, main concentration is on partitioning the whole image into small sub-images and then perform 2DPCA over small sub-images, because it gives more significant information as it focuses on smaller units, i.e. local details. For example, partition a whole face image into sub-images like mouth, nose, left eye, right eye, etc. Then, 2DPCA is performed over all mouth images, nose images, left eye images and right eye images separately and consider the variance in subgroup basis. The two steps of proposed feature extraction method are discussed as follows.

2.2.1 Partition of Face Image into Sub-image

Some parts of faces like mouth, nose and eyes have more information about face. These parts play important rule in face. So it is better to operate on those portions of the face instead of entire face. Mouth, nose and eyes are cropped from the detected face using computer vision system tool, i.e. CascadeObjectDetector [12] and these portions of faces are called as feature images. All cropped images have different sizes but to perform some operation we need to resize the images. Mouth image is resized as 50 by 50; nose image is resized as 50 by 50; left eye is resized as 40 by 40; and right eye is resized as 40 by 40. Four cropped portions of face image are represented in Fig. 2.
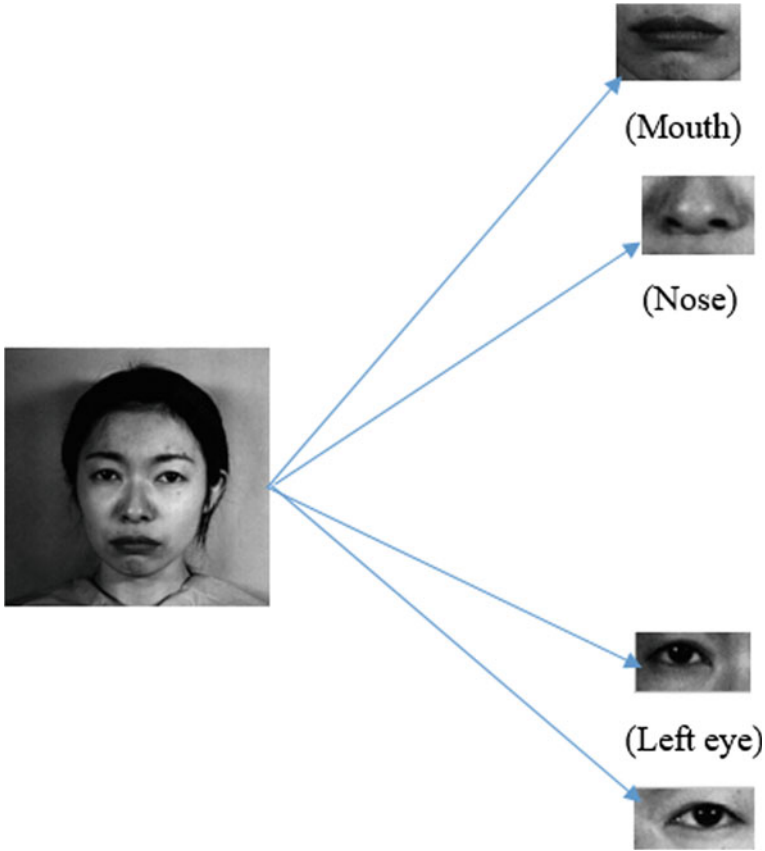
Fig. 2 Segmented sub-images

2.2.2 2DPCA-Based Features Extraction from Sub-images

Two-dimensional principal component analysis (2DPCA) algorithm [10] is used to extract the features from the feature images (mouth, nose, left eye and right eye). All seven expressions are set into seven different classes for training. 2DPCA is performed over mouth, nose, left eye and right eye images, separately, to extract the features. Eigenvectors and eigenvalues are calculated from these images, separately. The steps to extract the features using 2DPCA [14] are as follows.

Let the total number of images (represent by A_i) be M. The mean image μ_A of all images is calculated by the following equation.

$$\mu_A = \frac{1}{M} \sum_{i=1}^{M} A_i \tag{1}$$

The image scatter (covariance) matrix, G_t, is calculated by

$$G_t = \frac{1}{M} \sum_{i=1}^{M} (A_i - \mu_A)^T (A_i - \mu_A) \tag{2}$$

Now, the eigenvectors corresponding to largest eigenvalues of the covariance matrix G_t are calculated. Generally, a set of eigenvalues and corresponding eigenvectors are selected, to make the eigenimages when multiplied it with face images.

Let d be the number of eigenvector corresponding to d largest eigenvalues as the projection axes $X_1, X_2, X_3, \ldots, X_d$.

Then, the projected feature vector is evaluated by multiplying the eigenvector and images. Projected feature vectors are called principal component of images. Principal component contains the significant information of the images.

$$Y_i = A_i X_k, \tag{3}$$

where $i = 1, \ldots, M$ and $k = 1, \ldots, d$. Then, projected feature vectors are stored as

$$Y_i = [Y_1^i, Y_2^i, \ldots, Y_d^i] \tag{4}$$

2.3 Classification

The minimum distance classifier (MDC) is used for classification [10]. In test image, there will be four eigenvectors of cropped images mouth, nose, left eye and right eye, respectively. The distance between the features vector of each sub-image of test image is calculated corresponding to each sub-image of training set. Basically, Euclidean distance is considered here to measure the distance. Feature vector of each sub-image of test image and each sub-image of training sets is represented as

$$(Y_1, Y_2, \ldots, Y_d) \tag{5}$$

and

$$(Y_1^i, Y_2^i, \ldots, Y_d^i), \tag{6}$$

(where $i = 1, \ldots, M$ and $k = 1, \ldots, d$). The mathematical formula to finding the distance is

$$d_{x,y} = \sqrt{[\sum_{i=1}^{m} (x_i - y_i)^2]}. \tag{7}$$

Then, we need to normalize the distances and bring all values into the range [0, 1], after finding all the distance of mouth, nose, left eye and right eye. The mathematical formula to find the normalized distance is

$$X^i = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (8)$$

There will be four normalized values for mouth, nose, left eye and right eye. Then, we have to add all four normalized values and find the minimum normalized value among all the normalized distances. Minimum normalized value will classify the expression of a test image [15].

3 Experiment and Evaluation

In this section, experimental details are discussed for the facial expression recognition using 2DPCA. JAFFE facial expression database is used in this experiment. We have experimented with MATLAB 2013a version. Two major stages in this experiment are extraction of the features images and classify the facial expression by minimum distance classifier.

3.1 Description of the Dataset

JAFFE dataset contains frontal face images of size 256×256 of different facial expressions. All images are in greyscale image. This JAFFE dataset contains seven universal facial expression images. We rearranged the database in seven classes, and each class contains 25 images. In this experiment, for training, 15 number of images are taken for each expression. In training dataset, 105 images are taken and it contains all seven expressions. For testing, 10 images are taken per expression and there will be 70 images for testing. Facial expression is shown in Fig. 3.

3.2 Experimental Details

In this experiment, 25 number of images are taken for each expression. In training dataset, we have divided into seven classes and each class contains 15 number of

Fig. 3 Different facial expressions of JAFFE dataset: 1, surprise; 2, sadness; 3, neutral; 4, joy; 5, fear; 6, disgust; 7, anger

images for training. In testing, we have taken 10 images per class to recognize the expression. There are four parts of each image, and these face parts are mouth, nose, left eye and right eye. We have calculated the feature vectors of each part. We have compared all four feature vector of test image with four features vectors of all trained images. In case of test image, there will be one feature vector for mouth, one feature vector for nose, one feature vector for left eye and one feature vector for right. For training dataset, there will be 105 features vectors for mouth of 105 images, there will be 105 features vectors for nose of 105 images, there will be 105 features vectors for left eye of 105 images and there will be 105 features vectors for right eye of 105 images. We have calculated the Euclidean distance between feature vectors of each part of images. After that, we have calculated the normalized distance. There will be four normalized distances of one test image with each training image. Then, we have added all the four normalized distances. We have to find the minimum normalized distance among all the normalized distance, and that minimum normalized distance will classify the facial expression.

3.3 Result Analysis

15 images per class from the dataset are selected for training and 10 images per class for testing. There is no overlap of any images between training data and test data. We repeated this experiment 10 times by changing the training images and test images from dataset. Class-wise accuracy rate is given in table below. The overall classification accuracy for the proposed method is recognition rate that is 89.86%.

The performance of the proposed method is compared with the result of other two existing method like facial expression recognition using principal component analysis (PCA) [9] and facial expression representation and classification using 2DPCA [10]. The proposed method has higher recognition rate than other two methods. The performance of these three methods is shown in Table 1.

Table 1 Comparison of class-wise and overall classification accuracy (in %)

Method	Happy (%)	Sad (%)	Disgust (%)	Angry (%)	Surprise (%)	Fear (%)	Neutral (%)	Overall classification accuracy (%)
PCA	42.85	68.42	70	71.43	75	50	70	63.95
2DPCA	85	90	85	80	90	85	85	85.71
Partition based 2DPCA	93	90	92	90	90	80	94	89.86

4   Conclusions

This paper presents a new method to recognize the facial expression using 2DPCA by taking some important parts of face (mouth, nose, left eye and right eye) instead of entire face and extracting some significant feature from the feature images by 2DPCA. Eigenvectors and eigenvalues are calculated from the covariance matrix. Minimum distance classifier is used to classify the facial expression. It gives more recognition rate while performing 2DPCA on some significant face parts instead of entire face. The performance of 2DPCA is more efficient than PCA. The proposed method has far better result than other two existing methods like PCA and 2DPCA on entire face. The average facial expression recognition rate is very high in the proposed method.

References

1. Deng, H.-B., Jin, L.-W., Zhen, L.-X., Huang, J.-C.: A new facial expression recognition method based on local gabor filter bank and pca plus lda. Int. J. Inf. Technol. **11**(11), 86–96 (2005)
2. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans. Pattern Anal. Mach. Intell. **27**(8), 1226–1238 (2005)
3. Zhang, L., Tjondronegoro, D.: Facial expression recognition using facial movement features. IEEE Trans. Affect. Comput. **2**(4), 219–229 (2011)
4. Li, M., Yuan, B.: A novel statistical linear discriminant analysis for image matrix: two-dimensional fisherfaces. In: Proceedings of the International Conference on Signal Processing, vol. 2, pp. 1419–1422 (2004)
5. Tsai, C.-C., Chen, Y.-Z., Liao, C.-W.: Interactive emotion recognition using support vector machine for human-robot interaction. In: Proceedings of the International Conference on Systems, Man, and Cybernetics, pp. 407–412 (2009)
6. Pentland, A.: Looking at people: sensing for ubiquitous and wearable computing. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1), 107–119 (2000)
7. Cottrell, G.W., Fleming, M.: Face recognition using unsupervised feature extraction. In: Proceedings of the International Conference on Neural Network, pp. 322–325 (1990)
8. Valentin, D., Abdi, H., O'Toole, A.J., Cottrell, G.W.: Connectionist models of face processing: a survey. Pattern Recognit. **27**(9), 1209–1230 (1994)
9. Gosavi, A.P., Khot, S.R.: Facial expression recognition using principal component analysis. Int. J. Soft Comput. Eng. 2231–2307 (2013)
10. Thai, L.H., Ha, D.T.T.: Facial expression representation and classification using 2DPCA. In: Proceedings of the International Conference on Control, Automation and Information Sciences, pp. 42–47 (2012)
11. Das, D.K.: Comparative analysis of PCA and 2DPCA in face recognition. Int. J. Emerg. Technol. Adv. Eng. **2**(1), 330–336 (2012)
12. Wang, Y.-Q.: An analysis of the Viola-Jones face detection algorithm. Image Process. Line **4**, 128–148 (2014)
13. Oliveira, L., Mansano, M., Koerich, A., de Souza Britto Jr., A.: 2d principal component analysis for face and facial-expression recognition. Comput. Sci. Eng. **13**(3), 9–13 (2011)

14. Yang, J., Zhang, D., Frangi, A.F., Yang, J.-Y.: Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans. Pattern Anal. Mach. Intell. **26**(1), 131–137 (2004)
15. Sun, W., Ruan, Q.: Two-dimension PCA for facial expression recognition. In: Proceedings of the International Conference on Signal Processing, vol. 3 (2006)

Stable and Consistent Object Tracking: An Active Vision Approach

Dibyendu Kumar Das, Mouli Laha, Somajyoti Majumder and Dipnarayan Ray

Abstract Active vision plays an important role for many applications using better understanding of our surrounding environments. Using active vision, it is possible to gather new knowledge or changes in the environment by autonomously moving of eyes, and nearly all animals have this routine biological capability including human. In active vision, it is needed to repeatedly move the gaze from one fixation point to another fixation point to obtain high-resolution image and track the object robustly. The main reason to keep gaze still during fixations is the need to avoid the blur that results from the long response time of the photoreceptors. In this paper, we propose an active vision system which is constantly tracking an object by moving its gaze. In our proposed approach, human–eye movement technique, i.e., saccade and smooth pursuit mechanism, has been used to obtain high-resolution image on the sensor and track that particular object in real time.

Keywords Active vision · Saccade · Smooth pursuit · Gaze
Velocity estimation · Feature matching · Image moment

1 Introduction

In all vertebrates, vision is an ample source of information, they can sense and act accordingly by directing their gaze toward the object and scaning the surrounding. Like all other animal, humans can very efficiently direct and redirect their gaze toward any object, even though its relative location or orientation has been changed and altered. This capability considered as a classical example of active vision system. In fact, eye movements are necessary in order to focus the most important and high-resolution part of the object in our fovea. The movements are saccades,

D. K. Das (✉) · S. Majumder · D. Ray
CSIR-CMERI, Durgaur 713209, West Bengal, India
e-mail: dibs06.it@gmail.com

M. Laha
Indian Institute of Technology (Indian School of Mines), Dhanbad, India

smooth pursuit, Vestibular Ocular Reflex (VOR), and Optokinetic Ocular Reflex (OKR). The fovea sees only the central two degrees of the visual field, if any object is larger than the acuity of fovea, then eye repeatedly shift its gaze to acquire the different portions of the image into the fovea.

Human eye can move and shift their gaze attentively using two type of visual attention, namely overt and covert visual attention. Overt visual attention assimilates to the visual attention which involves the body movement, whereas covert visual attention mentally focuses on the scene without moving one's eye [1].

Additionally, Superior Colliculus (SC) is another major part of the visual system which is a layered multisensory structure, located in the midbrain. Visual inputs from the retina and the visual cortex reach the superficial layers of this laminated structure, while the lower layer processes the visual signal. The area projects its output to the ocular motor pathway in the midbrain and brainstem, where the retinotopic representation is transformed into motor commands [2].

There are six extraocular muscles that control the movement of human eye by which human can shift their gaze constantly. Our active vision research presented here is essentially inspired by the human vision system and specially focused on saccade and smooth pursuit movement. Here, we apply overt visual attention approach to control the camera angle and achieve the similar eye movement of human eye. Two cameras can operate interactively with the surrounding environment using multiple stepper controllers. The task is divided into two parts. Primarily, system has been deployed to detect an object of interest and then segmenting that region of interest point and match those points on successive frame over the time. An optical flow algorithm is applied to find the velocity of object from image data, according to the movement of the object cameras that are moved till their pan/tilt angle smoothly. This mechanism is similar to the smooth pursuit movement of human eye (Figs. 1 and 2).

We demonstrate our system using a hanging ball which is freely movable on its horizontal axis. Camera module persuades that ball with external command is based on an object tracking and velocity estimation algorithm.

Fig. 1 Human eye with its anatomical structure [11]

Fig. 2 Human eye with
extraocular muscle [12]
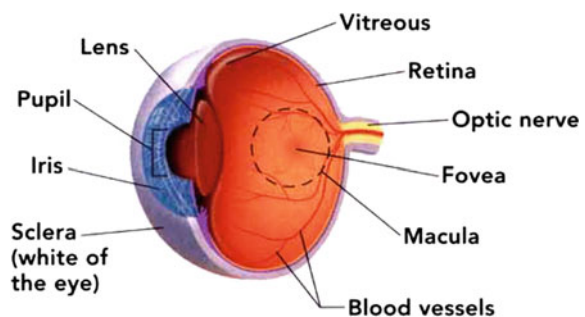
Fig. 2 Human eye with extraocular muscle [12]
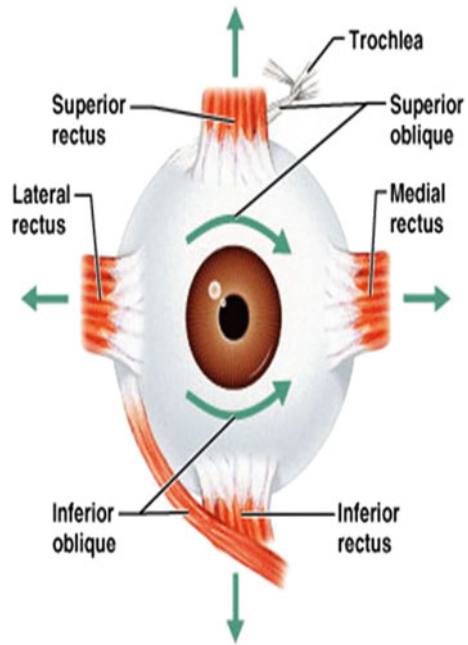
2 Related Works

Wang et al. [3] and Macesanu et al. [4] Research works on active vision system which imitates human eye and controls the camera using open- and closed-loop servo controller to achieve saccade and smooth pursuit movement respectively. In order to control the pan and tilt of the camera, two control approaches have been established [5]. These are velocity control for smooth pursuit and position control for saccadic movement. For attentively facilitating high-resolution images, cameras need to reorient their position in a controlled manner by adjusting the parameter (pan tilt zoom focus aperture) [6]. Research work on the task selection based controller which is used to track the object and take relevant perceptibility measure for active vision system using multiple cameras is done in [7]. Two relevant measures are taken in active vision to manipulate the environment, i.e., top-down and bottom-up approach [8]. Bottom-up is from the stimulus, and top-down is from the interests, knowledge, and memory of the agent-based system. The gaze control is main substance to perform during active vision, memorize the gaze space, and do task-based performance like manipulation was done by Hülse et al. [9]. Object detection based on saccadic movement is described in [10].

3   System Architecture and Methodology

We rigorously study the human eye movement for designing of our system. The system mainly consists of two cameras and three stepper motors with one controller for camera movement, which emulates the human oculomotor reflex. As per the physiology of human eye, there are six types of eye muscle which incorporates the movement of the eye in different directions. Our systems also have the same anthropomorphic properties of human eye movement. Unlike human eye, we restricted our system's DOF in 2 instead of 3. Each camera has one stepper motor for panning its axis independently in horizontal axis, while another one stepper is connected with the total system in tilt motion in vertical axis. The system has the capability of mimicking the saccadic and smooth pursuit movement of human eye by above-mentioned mechanism.

In this section, we also described the preprocessing of real-time video data and how that data can be recalibrated to a control command for further movement of the camera.

3.1   *Image Acquisition and Object Tracking*

Our first step is to capture real-time video by a high-resolution Watec camera, by which video is captured and stored in a particular location for further processing. Videos are fragmented into several frames or images. After capturing all images for a particular scene, we go through the next step with image registration. Using feature-based registration, we are extracting most unique features from an image like point, corner, and border lines from an image to make registration process most robust and reliable.

Feature is nothing but an interest point in an image, where the signal of that image changes frequently. In our method, we use Speeded Up Robust Features (SURF) feature detection for identifying interest points in an image due to its faster detection capability. Feature matching based on local feature is another crucial step for object tracking in successive frame. Object tracking is the process of locating and following the moving object in the sequence of video frames. In the very first step, the whole video has been fragmented into the corresponding image sequence. Feature matching algorithm has been applied to match local features among successive frame and detect the same object in the all subsequent frames. After successfully detecting an object on the next frame, a background modeling algorithm by Gaussian mixture model has been used to segregate the foreground from the background. A morphological operation has to be done for masking and excluding the noise from the binary image followed by the background segregation. Next step is to calculate the image moment or centroid and mark it as red color. Superimpose the centroid onto the frame for tracking the path of the ball on subsequent frame.

3.2 Motion Detection and Velocity Estimation

There are various types of application that relies on motion estimation. In our case, velocity estimation of moving object solely depends on analyzing the motion in image sequence. Image velocities can be measured by different algorithms, i.e., block matching algorithm, gradient-based approach using the optical flow field, spatiotemporal filtering, etc. In our proposed approach, a template matching algorithm with centroid detection technique has been used to track the motion of 2D object on successive frame and estimate the velocity of that object from the voxel (pixel with volume) data.

The camera was fixed at some distance from the ball, and the ball was moving along its x-y axis, so we will not consider the depth parameter for the time being. After finding the centroid of particular object, we had evaluated the corresponding pixel value of that centroid point, which we have to consider for calculating the speed of an object using intrinsic parameter.

Let us assume that I(x, y, t) is the centroid pixel at time t in a given frame that moves by (dx, dy) distance in the next frame taken after dt time. Then

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \tag{1}$$

on subsequent frames.

Following steps are necessary for velocity estimation of given object.

1. Find the centroid and corresponding 2-D pixel coordinate in successive frame.
2. Find the Euclidean distance of two centroids in each successive frame.

   $\text{Dist} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}$

   $i = 1$ to $n-1$

   $j = i+1$ to n

   n = no of frames captured for experiment.
3. Convert the 2-D pixel coordinate into 3-D world coordinate using camera intrinsic parameter for fixed distance.
4. Actual Dist (mm) = Dist * param (calculated earlier in mm).
5. Motion time (s) = n/fps value.
6. Velocity (cm/s) = (Actual Dist/Motion time)/100.
7. Motion and velocity estimate is the necessary step for accomplishing the smooth pursuit mechanism.

(a) **(b)** **(c)** **(d)**
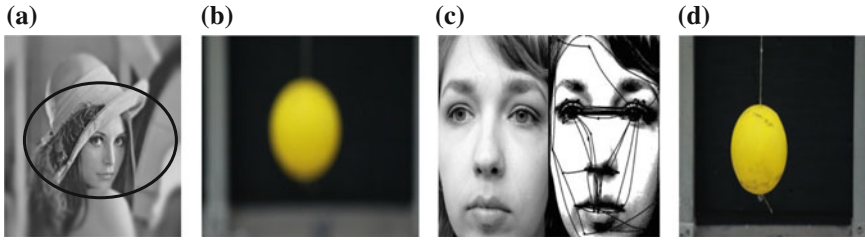
Fig. 3 **a** Foveated portion of an image in human eye visual system [13]. **b** Unfoveated image of the ball during experiment. **c** Tracing of saccade movement of human eye [14]. **d** Foveated image of the ball captured by the camera during experiment

3.3 Saccade Control

Saccade is a fast-gazing, ballistic, and conjugate eye movement. Saccade can be classified as memory-guided saccade, visually guided saccade, anti-saccade, and others. Because all of these saccade movements are consciously initiated, saccade is clearly a type of voluntary eye movement.

We develop our system in such a manner, where two cameras are identifying fixed object by slightly adjusts their lenses and gets a foveated image on the sensor. Camera take an image using its varifocal motorized lens, which was driven by command control using image intensity, based on contrast. As soon as the highest contrast is achieved for a particular object, lens has been stabilized itself and fixates that object in the sensor. If the object is not stationary, then the systems have to track the object by activating smooth pursuit control (Fig. 3).

3.4 Smooth Pursuit Control

Smooth pursuit is another eye movement by which human can smoothly track a moving object continuously; smooth pursuit helps to lock the target in the parafoveal area of the eyes toward moving object. This movement is implemented unconsciously by visual feedback. In addition, the tracking speed of smooth pursuit is determined by the object's speed. Smooth pursuit eye movements are conjugated and under feedback control system which is capable of rapid change in motion with respect to the motion of input image of the object.

In this paper, when the object is moving, high-resolution foveated image of the ball was captured by the saccade movement in each frame; after the successful saccade movement, smooth pursuit loop has begun to track the object by moving its pan and tilt angle accordingly. As mentioned earlier, velocity estimation of the moving object is the foremost step to begin smooth pursuit control. We map and calibrate the velocity value into the PWM voltage for controlling the stepper motor.

4 Experiments and Result

In this section, we illustrate the system setup and how the system works efficiently to track an object. Real-time video has been captured by Watec analog camera with the Matrox frame grabber; then, it has been imported to the MATLAB vision toolbox for further processing and a velocity estimation algorithm has purposefully selected for estimating the velocity of corresponding object in that video. Two cameras and the system are connected to the NEMA-17 stepper motor, by which each camera can separately pan in their horizontal axis and the adjoin system can be tilted in vertical axis driven by that motor.

All the motors are connected to and controlled by the Arduino Mega controller. In the following experiment, we use a pendulum-like structure with yellow ball (Fig. 5) and maintain the distance between pendulum and camera of approximately 2 m. Video has been captured in 25 FPS, and a velocity estimation algorithm finds the velocity and direction of the object and transfers that data to the Arduino controller for generating corresponding PWM signal to operate the stepper motor.

4.1 Smooth Pursuit Mechanism

Smooth pursuit mechanism starts if the object is in motion; otherwise, saccade movement is sufficient to detect high-resolution image of the object. If the object is not fixed in one position, then smooth pursuit takes place followed by saccadic control of the system and track the object in real time (Figs. 4 and 5).

Above figure demonstrated five successive frames of captured video data during experiment, where red dots implicate the centroid with pixel coordinate value. Furthermore, we have already discussed about the velocity estimation in Sect. 3, which will be applied in next step to find the velocity of the ball in the world coordinate system.

4.1.1 Velocity and Direction Estimation

When object left camera's deep fovea region and appears in its parafoveal area, feature matching algorithm has been started to match the template from adjacent
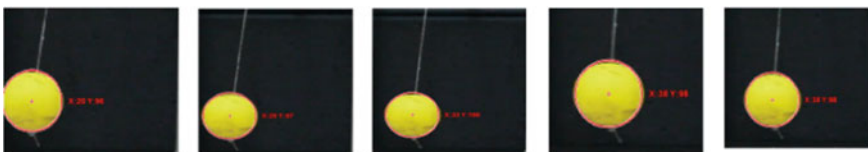
Fig. 4 Detection of centroid
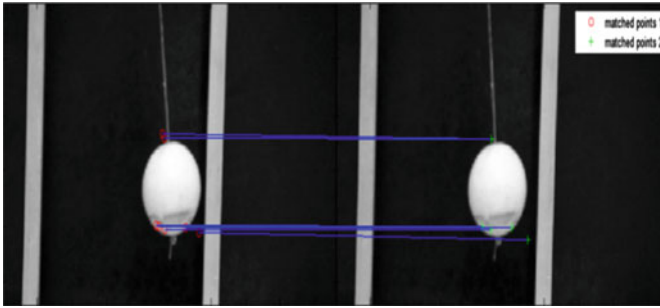
| Object in camera' deep fovea region | Object in camera' parafovea region in next frame |

Fig. 5 Matching points using SURF algorithm

Table 1 Object velocity and corresponding motor current

Frame	1st	2nd	3rd	4th	5th	6th	7th
Velocity cm/s	2.5	2.2	1.96	1.51	1.5	1.33	1.2
Motor current(mA)	50.4	44.35	38.30	30.40	30.24	26.71	24.19

frame or region. As soon as template has been caught, camera will redirect its axis to the moving object using its corresponding world coordinate. The video was captured in 25 fps so motor is controlled in 25 Hz; otherwise, motion blur will happen during processing.

We consider few frames where ball starting from the center position propagates through the right upward direction. Object velocity has been sent to the motor controller as a feedback to rotate the motor with significant stepping value (Table 1).

Fig. 6 Smooth pursuit velocity with seven successive frames
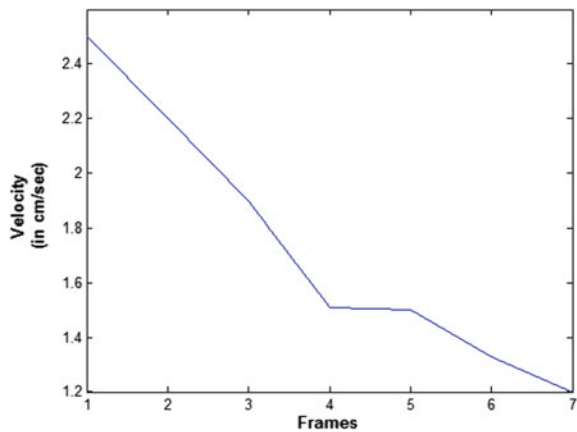
Fig. 7 Supply current in motor according to the object movement
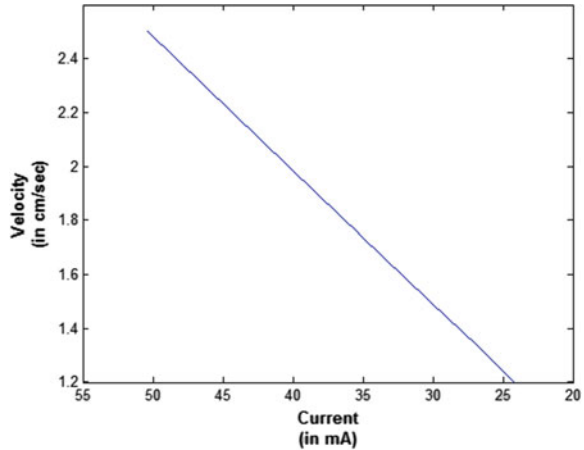
Velocity of the ball gradually decreased while going to the upward direction in all subsequent frames (Figs. 6 and 7).

5   Conclusion and Future Work

In this paper, we proposed an active vision system which imitates the human eye movement; basically, it implements two major eye movements such as saccadic and smooth pursuit movement. We first summarize the human eye anatomy and physiology of its movement, and generate the same control model using various apparatuses like stepper motor, controller, and camera. Our aim in this research paper is to detect the high-resolution image of object and track that object smoothly in more robust way. Next, we plan to evaluate VOR and OKR movement with the help of IMU sensor and implement that active vision module on the mobile robot for wide-area environment monitoring and surveillance.

References

1. Kietzmann, T., Geuter, S., König, P.: Overt visual attention as a causal factor of perceptual awareness. PLoS ONE **6**(7), e22614 (2011)
2. Findlay, J.M., Gilchrist, I.D.: Visual attention: the active vision perspective. In: Vision and Attention, pp. 83–103. Springer, New York (2001)
3. Wang, X., van de Weem, J., Jonker, P.: An advanced active vision system imitating human eye movements. In: 2013 16th International Conference on Advanced Robotics (ICAR). IEEE (2013)
4. Macesanu, G., Grigorescu, S.M., Comnac, V.: Time-delay analysis of a robotic stereo active vision system. In: 2011 15th International Conference on System Theory, Control, and Computing (ICSTCC). IEEE (2011)

5. Land, M.: Motion and vision: why animals move their eyes. J. Comp. Physiol. A. **185**(4), 341–352 (1999)
6. Rotstein, H., Rivlin, E.: Control of a camera for active vision: foveated vision, smooth tracking and saccade. In: Proceedings of the 1996 IEEE International Conference on Control Applications, 1996. IEEE (1996)
7. Ognibene, D., Baldassare, G.: Ecological active vision: four bioinspired principles to integrate bottom–up and adaptive top–down attention tested with a simple camera-arm robot. In: IEEE Transactions on Autonomous Mental Development, vol. 7(1), pp. 3–25 (2015)
8. Marraffa, R., et al.: A bio-inspired attention model of anticipation in gaze-contingency experiments with infants. In: 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL). IEEE (2012)
9. Hülse, M., et al.: Integration of active vision and reaching from a developmental robotics perspective. In: IEEE Transactions on Autonomous Mental Development, vol. 2(4), pp. 355–367 (2010)
10. Yabuta, Y.: Object detection by following saccadic movement. In: MECATRONICS—REM 2016, 15–17 June (2016)
11. http://www.urbaneyemd.com/images/glossary
12. https://s-media-cache-ak0.pinimg.com/originals
13. https://ece.uwaterloo.ca/~z70wang/research/fovea/fovea_model.html
14. http://www.vertigoexpert.org/wp-content/uploads/Vertigo-eye-exercises-440x500.jpg

Vision-Based Forward Kinematics Using ANN for Weld Line Detection with a 5-DOF Robot Manipulator

Don Joe Martin, Aaditya Saraiya, V. Kalaichelvi and R. Karthikeyan

Abstract While robotic manipulators are becoming a common sight in today's industries and fast paced production lines, it is becoming difficult to develop foolproof methods for automation of these manipulators, owing to their geometric and structural variety. Creating a common algorithm for these manipulators would help in setting a base standard for their automation. Trio Motion coordinators are most widely used for robotic manipulators in recent times. The objective of this paper is to create a simple interface based on Visual Basic programming language to coordinate directly with the robot's motion coordinator by bypassing all other programming methods which are otherwise needed for sending commands to the robot. This interface can be easily adapted for further tuning methods and also for more or lesser degrees-of-freedom robotic manipulators. MATLAB has been used for detecting the weld line in the image using image processing techniques. A suitable artificial neural network has been used to give forward kinematic solutions with image coordinates as the input.

Keywords Robotic manipulator · MATLAB · 5-DOF robot
Weld line detection · Artificial neural networks

1 Introduction

High productivity requires manipulators that are capable of performing complex tasks in minimum time. Current industrial manipulators, however, are not controlled to achieve minimum-time motions or are their work places structured so as

D. J. Martin (✉) · A. Saraiya · V. Kalaichelvi
Department of Electrical and Electronics Engineering,
BITS Pilani, Dubai Campus, Dubai, United Arab Emirates
e-mail: f2014017@dubai.bits-pilani.ac.in

R. Karthikeyan
Department of Mechanical Engineering, BITS Pilani,
Dubai Campus, Dubai, United Arab Emirates

to decrease the required task time [1]. This paper aims to create a Visual Basic based GUI which can be utilized to bypass the Trio Motion controllers, which have been used in the TAL (Tata Automation Limited, Pune) Brabo robot. Dung L.T et al. have described methods of utilizing MATLAB/Simulink to model the robot and provide the required control checks [2]. The image that is obtained from the robot is passed through image processing algorithms that help in filtering out unwanted features, thereby producing the weld line to be traced. A few image processing algorithms have been discussed henceforth in the text, with the results relevant to weld line detection. However, even though Vijayarani and Vinupriya [3] found that Canny edge detection is better than Sobel detection, we experimented with both these popular edge detection algorithms. For the case of weld line detection, Canny edge detection was not effective due to the detection of unnecessary image features. The image processing results with Sobel edge detector were satisfactory and have been enumerated in Sect. 5. An artificial neural network model is trained in order to predict the output and to solve for the forward kinematics solutions (X, Y, Z) when the image coordinates (x, y) have been provided to the neural network as input. Dehghani et al. in [4] have compared the results of MLP neural network with a wavelet-based neural network for providing forward kinematic solutions of a HEXA parallel manipulator robot.

2   Modeling of 5-DOF Robot Manipulator

Trio software is used to create the interface between the robot's controller and the operating computer system which is replaced. This is used to run the robot until an appropriate controller is created that can override all of these software's contributions and automate the entire robotic process. Figure 1 shows the 5-DOF TAL Brabo robot utilized for the analysis. A graphical user interface for the robot was built up using ActiveX and Visual Basic. ActiveX is a tunnel mechanism that
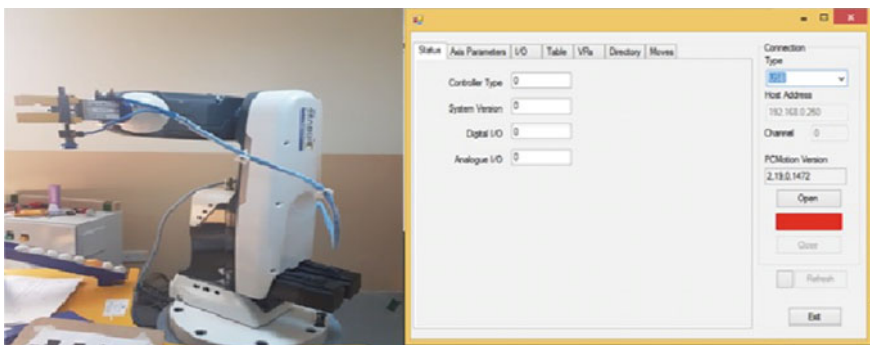
Fig. 1 The TAL robot's response to the Visual Basic commands
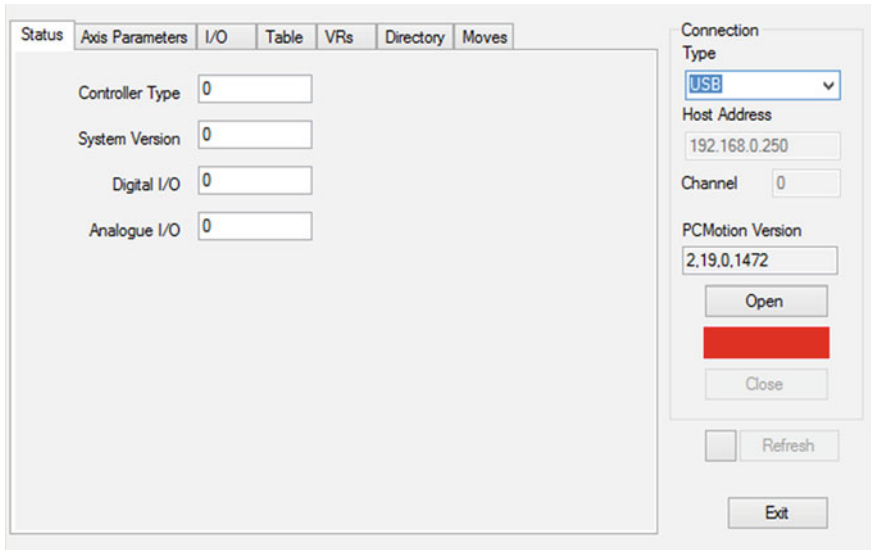
Fig. 2 ActiveX–Visual Basic interfacing for TAL robot

creates a direct connection between the robot's motion controller and the controlling PC, and helps to control the robot motion drivers in an easy way. The Visual Basic program that was created enables the user to send the required coordinates directly to the robot. Figure 2 shows the GUI created for ActiveX–Visual Basic interfacing for TAL robot.

3 Denavit–Hartenberg Parameters

Four DH parameters have been used to describe each of the five single DOF revolute joints used in the serial robotic manipulator [5]. The four DH parameters include two link parameters (a_i, α_i) and two joint parameters (d_i, θ_i) that are shown in Table 1. The DH parameters help to showcase the dimensions of the robot in a succinct manner and provide a standard method to describe the forward kinematics

Table 1 The Denavit–Hartenberg parameters for 5-DOF TAL robot (in cm)

Link no. i	a_i	α_i	d_i	θ_i	Joint variable q_i
1	0	−90	$40(L_1)$	θ_1	θ_1
2	$40(L_2)$	0	0	θ_2	θ_2
3	$40(L_3)$	0	0	θ_3	θ_3
4	0	−90	0	$\theta_4 - 90$	θ_4
5	0	0	$40(L_5)$	θ_5	θ_5

equations for a robotic manipulator. The DH parameters can be explained as given below.

a. Link length (a_i)—Distance between Z_i and Z_{i+1} along axis X_{i+1}
b. Link twist angle (α_i)—Angle between axes Z_i and Z_{i+1} along axis X_{i+1}
c. Joint distance (d_i)—Distance between X_i and X_{i+1} along axis Z_i
d. Joint angle (θ_i)—Angle between X_i and X_{i+1} along axis Z_i

4 Forward Kinematics

Forward kinematic model of the manipulator helps to analyze the transformation from joint space to Cartesian space [5]. The overall transformation from the base to the end effector frame can be modeled by successive transformations from one joint to another as shown in Eq. (1).

$$
{}^0_5T = {}^0_1T \times {}^1_2T \times {}^2_3T \times {}^3_4T \times {}^4_5T \tag{1}
$$

The overall transformation can be given by the following matrix equation given in Eq. (2).

$$
{}^0_5T = \begin{bmatrix}
C_1S_{234}C_5 + S_1S_5 & -C_1S_{234}S_5 + S_1C_5 & C_1C_{234} & C_1(L_2C_2 + L_3C_{23} + L_5C_{234}) \\
C_1S_{234}C_5 - C_1S_5 & -S_1S_{234}S_5 - C_1C_5 & S_1C_{234} & S_1(L_2C_2 + L_3C_{23} + L_5C_{234}) \\
-C_{234}C_5 & C_{234}S_5 & -S_{234} & L_1 - L_2S_2 - L_3S_{23} - L_5S_{234} \\
0 & 0 & 0 & 1
\end{bmatrix} \tag{2}
$$

From the above matrix, the position of end effector (X, Y, Z) can be deduced for the 5-DOF robotic manipulator. S refers to sin θ, C refers to cos θ, and L is length of links. Equation (3) shows the matrix equation which can be used to given the position of end effector in the initial conditions.

$$
T_{end-effector} = \begin{bmatrix}
0 & 0 & 1 & L_2 + L_3 + L_5 \\
0 & -1 & 0 & 0 \\
1 & 0 & 1 & L_1 \\
0 & 0 & 0 & 1
\end{bmatrix} \tag{3}
$$

5 Image Processing for Weld Line Detection

Image processing has been utilized for the process of automatic arc welding. For the process of arc welding, the endpoints of the center weld line are considered as features and the image processing is done in real time to ensure the most accurate

and adaptable results for weld line detection. Certain image preprocessing techniques are applied in order to segment the weld line and detect the endpoints of the weld line. The preprocessing techniques used are explained below.

5.1 Conversion of Image to Grayscale

The original image is captured by the camera at fixed height. The images are of 640×480 resolution and consist of RGB images which have three channels (Red, Green, and Blue). These images are converted into grayscale to make the data easier to handle and to apply segmentation algorithms.

5.2 Sobel Edge Detection

For the process of detection of the center weld line, a Sobel edge detector has been utilized. The Sobel edge detector highlights the edges in the image by finding the absolute gradient at each point. The algorithm uses two 3×3 convolution kernels in order to calculate the derivatives of pixel intensities in both vertical and horizontal directions [3]. For the weld images, the Sobel edge detector has been used in the vertical direction. This allows detection of all vertical edges which was beneficial for the particular grayscale images considered.
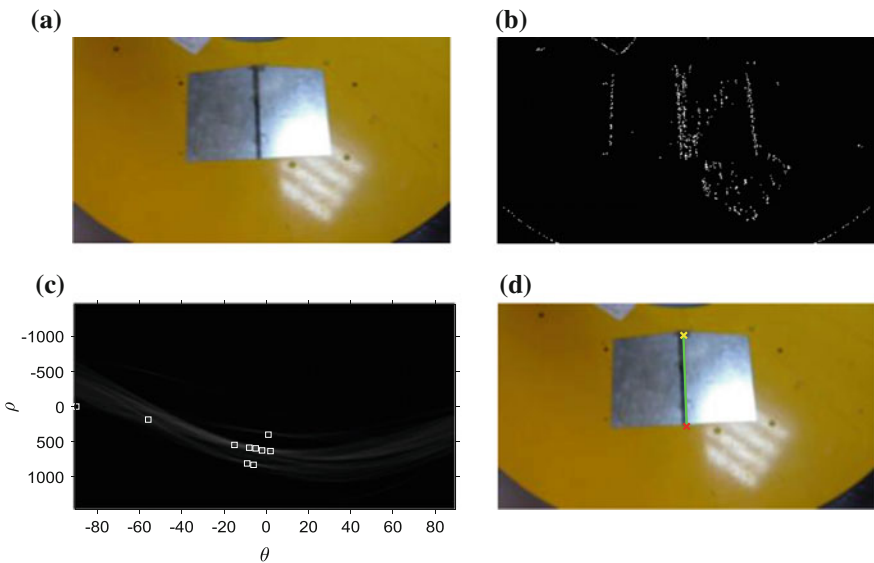
Fig. 3 **a** Original image of weldpiece. **b** Edges detected in weldpiece. **c** Coordinates in the Hough space. **d** Original image with weld line detected
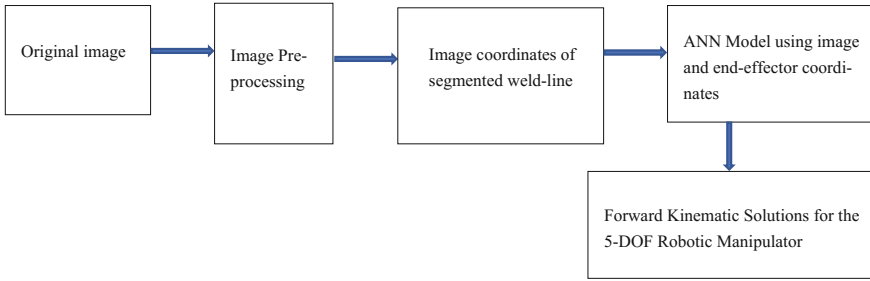
Fig. 4 Workflow of the process of weld line detection and forward kinematics using ANN

5.3　*Hough Line Transform*

After the process of edge detection, Hough line transform is used to detect even imperfect lines in the image. First, the detected edge pixels are transformed into Hough space and stored in an accumulator. These accumulators are used to detect infinite lines which are later reduced to finite lines. Using this, the center weld line, being most prominent in the image, is detected and the endpoints are obtained [6]. Figure 3a–d shows the results which are obtained after completion of image pre-processing. Figure 4 summarizes the entire workflow including the image processing and the neural networks part. The experimental data obtained is used for post-processing and is used as training data for the artificial neural network which is explained in Sect. 6.

6　ANN Modeling for Forward Kinematics

6.1　*Overview*

Neural networks are computational structures which consist of fundamental computational units called neurons. These networks have an ability of classification of the data according to the weights and biases of the neurons. The weights of the hidden layer are varied as per the backpropagation algorithm over numerous epochs in order to introduce various features which can successfully fit the data and aid in reduction of the error metric [4]. The scaled conjugate training method has a better prediction accuracy and is utilized for the proposed work. This method uses information of the second order from the neural network and also takes step-size scaling into consideration. The detailed explanation is given in reference number [7].

6.2 Dataset and Network Architecture

The real-time data collected is used for training the neural network for forward kinematics. Due to the limitation of the intersection between the workspace of the robot and view of the camera, a total of 16 images were captured which were utilized for training the neural network. The testing data was separately captured and involved four images at different orientations from the training data. The prediction accuracy was $\pm 2\%$ for the testing data. A neural network architecture with 2 input neurons, 25 neurons in the hidden layer, and 3 neurons in the output layer is utilized. The target data was prepared by taking the ideal values of the weld lines without any error. Mean squared error (MSE) was chosen as the error metric for this process.

6.3 Methodology

The input values for the neural network include the (x, y) coordinates obtained as endpoints of the weld line from the image after the image processing steps. The neural network maps these input values to the output values of the neural network, which are the (X, Y, Z) coordinates of the end effector of the robot in the linear mode. With these provided coordinates, the end effector can move to the end points of the weld workpiece. These forward kinematic solutions were experimentally verified in real time by moving the Tata TAL Brabo robot to the various positions obtained as output from the neural network.

7 Results and Discussion

The neural network model has many parameters which can be tuned in order to improve the network performance. The following analysis has been carried out in order to validate the performance of neural modeling for 5-DOF robotic manipulator.

7.1 Training the Network with Different Training Algorithms

The neural network was trained with three algorithms which include Levenberg–Marquardt (LM), Scaled Conjugate Gradient (SCG), and Gradient Descent (GD). The scaled conjugate gradient produced the lowest MSE of 46.2962 in 2958 epochs as shown in Fig. 5. The Levenberg–Marquardt algorithm produced the same error,

Best Training Performance is 46.2962 at epoch 2179
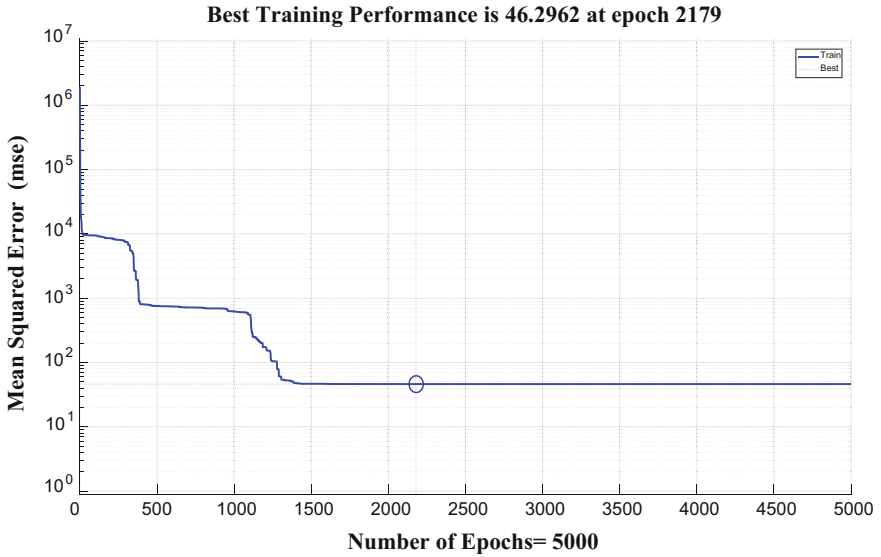
Fig. 5 Mean square error versus no. of epochs

and however did not run for the full number of epochs while the gradient descent algorithm does not converge. Figure 6 gives the prediction accuracy between actual output and neural model output.

Training: R=0.99959
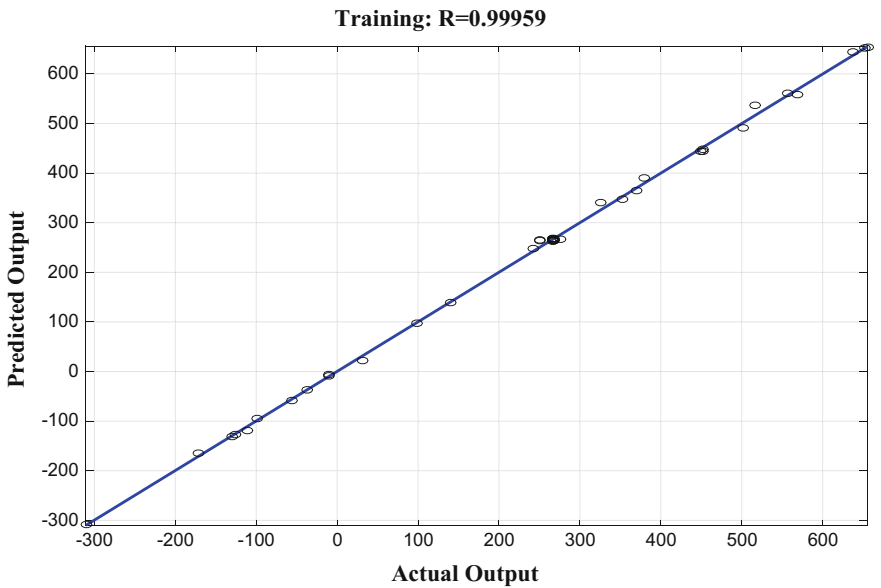
Fig. 6 Regression analysis for actual output and neural model output

Table 2 Performance of error with change in the number of epochs for SCG

No. of epochs	1000	2000	4000	5000	10,000
MSE	601.8986	465.1309	46.2962	46.2962	46.2962

7.2 Training the Network with Different Number of Epochs

From the previous analysis, Scaled Conjugate gradient (SCG) was taken as the best training algorithm hence further analysis was done using that. The performance of Scaled Conjugate gradient was observed over different number of epochs which has been summarized in Table 2. Beyond 4000 epochs, the error remains almost same.

7.3 Training the Network with Different Activation Functions

For the earlier analysis, purelin activation function was used. However, in this analysis, the neural network was trained using scaled conjugate gradient with 5000 epochs using various activation functions. For the log sigmoid activation function, the MSE was 113771.309 at epoch 1664 and for tan sigmoid, the MSE was 113752.175 with epoch 64 with the training not occurring for all epochs. Hence, purelin linear activation function was considered ideal for the analysis.

7.4 Training the Network with Different Number of Neurons in the Hidden Layer

The performance analysis has been extended by changing the number of neurons in the hidden layer. The MSE was observed to be same when using 5, 10, 15, 20, and 25 neurons. However, the convergence was fastest with 25 neurons in the hidden layer. Hence, choosing 25 neurons is most effective if the computation resources are available for the same.

8 Conclusion

A GUI has been successfully created for controlling TRIO controllers using the ActiveX interface through Visual Basic. A neural network based approach has been used to find the forward kinematics solution. This is done by creating a mapping between the detected endpoints in the weld line and the required orientation of the robot to reach to those particular points in Cartesian space. The neural network

performance has been analyzed by tuning the network with the efficient performance parameters. The performance of the neural network has also been verified in real time by testing the output values using a 5-DOF TAL Brabo robot. It can be concluded that the ANN modeling has given satisfactory results and can be used for real-time applications.

References

1. Luh, J.Y.S., Lin, C.S.: Optimum path planning for mechanical manipulators. J. Dyn. Syst. Meas. Control. **102**(2), 142–151 (1982)
2. Dung, L.T., Kang, H.J., Ro, Y.S.: Robot manipulator modeling in Matlab-Simmechanics with PD control and online Gravity compensation. In: IEEE Proceedings International Forum on Strategic Technology, pp. 446–449. IEEE (2010)
3. Vijayarani, S., Vinupriya, M.: Performance analysis of canny and sobel edge detection algorithms in image mining. Int. J. Innov. Res. Comput. Commun. Eng. **1**(8), 1760–1767 (2013)
4. Dehgani, M., et al.: Neural network solution for forward kinematics problem of HEXA parallel robot. In: 2008 American Control Conference, University of Illinois, Illinois, pp. 4214–4219. IEEE (2008)
5. Nugroho, A., Prihatmanto, S., Rohman, S.: Design and implementation of Kinematics model and trajectory planning for NAO humanoid robot in a tic-tac-toe board game. In: Proceedings of IEEE 4th International Conference on System Engineering and Technology, vol. 4, pp. 1–7. IEEE (2014)
6. Hassanein, A. Shehata, et al.: A survey on Hough transform, theory, techniques and applications. Int. J. Comput. Sci. Iss. **12**(1) (2015)
7. Mishra, S., Prusty, R., Hota, P.K.: Analysis of Levenberg-Marquardt and Scaled Conjugate gradient training algorithms for artificial neural network based LS and MMSE estimated channel equalizers. In: 2015 International Conference on Man and Machine Interfacing (MAMI), Bhubaneswar, Odisha, India, pp. 1–7. IEEE (2015)

Performance Measurement and Evaluation of Pluggable to Scheduler Dynamic Load Balancing Algorithm (P2S_DLB) in Distributed Computing Environment

Devendra Thakor and Bankim Patel

Abstract The imbalanced load between clusters is a key issue in distributed computing environment. All existing dynamic load balancing algorithms are post-active, as balancing activities start after system turn into imbalanced state. The better approach is to design pro-active load balancing algorithm which starts working with scheduling algorithms. It helps scheduling algorithms to schedule incoming jobs in such a way that system remains in balanced state. The pluggable to scheduler dynamic load balancing algorithm (P2S_DLB) is designed and evaluated over priority scheduling algorithm in our previous research work. The P2S_DLB is pro-active dynamic load balancing algorithm. In this paper, we have measured and evaluated the performance of P2S_DLB over First Come First Serve (FCFS), Shortest Job First (SJF), and Earliest Deadline First (EDF) scheduling algorithms. The experimental result shows that algorithm has improved the cluster utilization and decreased the imbalance level of distributed computing environment in case of all the three scheduling algorithms.

Keywords Dynamic load balancing · Imbalanced load · Distributed computing Cluster · Scheduling algorithm

1 Introduction

The distributed computing environment is widely used platform for job execution due to the advancement in computing and network technology. It provides platform to access geographically scattered resources by sharing jobs between computational nodes [1, 2]. The geographically scattered computational resources are congregated into virtual groups known as cluster [3]. The user-generated jobs are assigned to

D. Thakor (✉) · B. Patel
Uka Tarsadia University, Bardoli, India
e-mail: devendra.thakor@utu.ac.in

B. Patel
e-mail: bankim.patel@utu.ac.in

clusters for an execution. The biggest issue in distributed computing environment is uneven utilization of clusters [3]. It has been noted in literature that gradually some of the clusters becomes overloaded while the other clusters remain medium loaded or underloaded [4–9]. It is a problem of imbalanced load between clusters which reduces the overall system performance as resources are not utilized properly. The reason behind the problem is post-activeness of load balancing algorithms and job scheduling performs without looking the current status of clusters.

The scheduling algorithms allocate the incoming jobs to different clusters for execution in well-defined pattern. The algorithms consider resource requirements, priority, sequence, deadline, etc. of incoming jobs for scheduling [4]. Load balancing is the process of harmonizing load between clusters of distributed systems. It aims to optimize resource utilization, maximize throughput, minimize response time, and avoid overload of any single resource [1–3, 5, 9].

Existing dynamic load balancing algorithms start balancing activities after system becomes unbalanced. The better approach is to start balancing activities from the beginning by running dynamic load balancing algorithm in parallel with job scheduling algorithm [5]. If the scheduling decision is to be taken by considering current status of clusters, then the problem of imbalanced load between clusters can be resolved. We have designed the pluggable to scheduler dynamic load balancing algorithm (P2S_DLB) in our previous research work [5]. It can be plugged into any scheduling algorithm which helps scheduler to select best suitable cluster by providing current load status of clusters.

The first objective of the study is to check the performance of P2S_DLB algorithm over different well-known scheduling algorithms. The First Come First Serve (FCFS), Shortest Job First (SJF), and Earliest Deadline First (EDF) scheduling algorithms are selected for checking. The second objective is to measure and evaluate the performance of the simulation results of FCFS, SLF, and EDF algorithms with FCFS_DLB, SJF_DLB, and EDF_DLB algorithms, which are the combination of FCFS, SJF, and EDF with P2S_DLB algorithm respectively.

The rest of the paper is organized as follows. In Sect. 2, we have discussed the related work. The system model, list of notations, and evaluation parameters are described in Sect. 3. Section 4 focuses on implementation scenario, simulation setup, and clusters information of input dataset. The steps of P2S_DLB algorithm are stated in Sect. 5. The simulation results and analysis of results are presented in Sect. 6. Finally, conclusion is stated in Sect. 7.

2 Related Work

Many researchers have addressed the issue of imbalanced load between clusters. They came up with different methodologies for balancing load in distributed computing environment which can be found in literature [5–9].

The pluggable to scheduler dynamic load balancing algorithm (P2S_DLB) is proposed in [5]. The algorithm takes current load status of clusters and resource

requirement of incoming jobs as an input and find out best suitable cluster for incoming job execution as an output. The cluster having higher number of free resources is selected first for execution of incoming job. The algorithm has total four levels; in the first level, it tries to achieve more than 50% utilization of every clusters. In the second level, an algorithm going one step ahead and tries to achieve more than 66% cluster utilization. The algorithm tries to achieve 75% and 80% cluster utilization in levels 3 and 4, respectively. The P2S_DLB is simulated in ALEA [12] using real-time dataset over priority scheduling (PS). The simulation results show that P2S_DLB increased the average resource utilization by 16.12% and decreased the average of percentage of imbalance matric to 9.12% which is 67.71% in case of PS.

In paper [6], a semi-distributed load balancing method is proposed for homogeneous system. The new job is assigned to a processing node by considering the present load status of each node. The system is partitioned into a number of clusters, and each cluster has a master to perform local load balancing and also to communicate with global master. The mechanisms of election for new master in case of failure exclude faulty nodes from a cluster is proposed. The proposed load balancing method is scalable and has low message and time complexities. The method of cluster partitioning and load transfer is not addressed in paper.

In [7], a new dynamic task scheduling algorithm for heterogeneous cluster-based HEFT with duplication (CBHD) have been developed. The CBHD algorithm is union of the heterogeneous earliest finish time (HEFT) and the triplet clustering algorithms. The developed CBHD algorithm takes less execution time than both HEFT and triplet cluster algorithms, and it achieves the load balancing which considered one of the main performance factors in the dynamic environment.

The prediction-based dynamic load balancing algorithm (PDLB) is designed in [8]. It predicts the future status of clusters which is playing vital role to balance load between the clusters. The experimental results show that prediction approach improves the performance but still there is a scope of improvement [9].

3 System Model

The distributed computing environment comprises two entities, computational resources and the users. The computational resources are geographically scattered and in high quantity while the users are of different types and connected via networks.

The following properties of the distributed computing environment are considered:

- Incoming jobs are heterogeneous with respect to CPU, memory, and I/O requirements.
- Clusters are heterogeneous with respect to number of machines, cost of processing, processing speed, operating system, memory, and I/O capacity.
- Study is limited to non-preemptive job scheduling.
- Migration of job is not required.

Table 1 List of notations

Notation	Description
n	Total number of jobs
m	Total number of clusters
C_j	jth cluster, where $1 \leq j \leq m$
λ_j	Load imbalance matric of jth cluster
Lj_{max}	Maximum load of jth cluster
$\overline{L}$	Mean load of all clusters
Num_PE_j	Total number of processing elements on jth cluster
$Num\_Free\_PE\_j$	Total number of free processing elements on jth cluster
$Num\_Running\_PE\_j$	Total number of running processing elements on jth cluster
$Num\_Required\_PE\_i$	Number of required processing elements for successful execution of ith job

Table 1 shows the list of notations used in paper. Load imbalance metrics characterize how unevenly work is distributed between clusters. The percent imbalance metric is most commonly used parameter to check level of imbalance. Imbalance matric can be computed using the following equation: [10]. For cluster C_j,

$$\lambda_j = (\frac{Lj_{max}}{\overline{L}} - 1) * 100\%,\tag{1}$$

where Lj_{max} is the maximum load on jth cluster and $\overline{L}$ is the mean load of all clusters. This metric measures the performance lost due to imbalanced load, the performance that could be reclaimed by balancing the load.

4   Implementation Scenario and Input Dataset

Windows 10 on an Intel Core i3-370M processor with 3 GB of RAM and 320 GB of hard disk is used during experiments in ALEA version 2 [11] with JDK 1.8 and JRE 1.8. The goal of the experiment is to check the performance of P2S_DLB with combination of different scheduling algorithms. ALEA is GridSim [12] based job scheduling simulator which in turn is a Java-based simulation tool that provides facilities to model and simulates the entities like users, heterogeneous resources, resource load balancers, and applications.

The real-time dataset from the Gaia cluster log is taken for testing algorithms. The Gaia cluster is one of the four clusters operated by the University of Luxembourg HPC Center (ULHPC) initially released in 2011 [13, 14]. The Gaia is a heterogeneous cluster that has been upgraded several times. The selected dataset contains 3 months of data from May to August, 2014. It is used mainly by biologists working with large data problems and engineering people working with physical simulations.

Table 2 Cluster configuration

Cluster Id	Cluster Name	Nodes	CPUs	Total CPUs
1	gaia-[1–60]	60	12	720
2	gaia-[61–62]	2	12	24
3	gaia-[63–72]	10	12	120
4	gaia-73	1	160	160
5	gaia-74	1	36	36
6	gaia-[75–79]	5	16	80
7	gaia-[80–119]	40	12	480
8	gaia-[120–151]	32	12	384
Total		151	272	2004

Total of 51,987 jobs are created during the 3-month period and submitted to Gaia cluster having 151 nodes, manufactured by Bull and Dell, with a total of 2004 cores for job execution [13, 14].

Table 2 describes the specifications of the Gaia cluster. The Gaia cluster is collection of total 08 heterogeneous clusters. Each cluster has different numbers of machines and processers. It contains total 2004 processors to execute incoming jobs.

5 Pluggable to Scheduler Dynamic Load Balancing Algorithm (P2S_DLB)

The P2S_DLB is doing load balancing activity at the time of job scheduling. The algorithm takes current load status of clusters and incoming job as an input and find out best suitable cluster for incoming job execution as an output. The clusters having higher number of free resources are selected first for incoming job execution. The algorithm achieves better resource utilization and increases the level of load balancing step by step. It dynamically balances load by scheduling incoming job to specific cluster by considering current load status of that cluster. The best part of P2S_DLB is that it balances load without job migrate.

6 Experimental Results and Analysis

The scheduling algorithms (FCFS, SJF, and EDF) and dynamic load balancing algorithms (FCFS_DLB, SJF_DLB, and EDF_DLB) are simulated in predefined environment. The cluster utilization and the percent of imbalance matric are calculated as performance evaluation parameters.

The statistics of each cluster utilization, average cluster utilization, and improvement in average cluster utilization for all algorithms is shown in Table 3. The result

---

Algorithm 1 Pluggable to scheduler dynamic load balancing algorithm (P2S_DLB)

1: **for** $i = 1$ to n **do**
2:    **for** $j = 1$ to m **do**
3:       **if** $Num\_Free\_PE_j > Num\_PE_j/2$ **then**
4:          **if** $Num\_Running\_PE_j \geq Num\_Required\_PE_i$ **then**
5:             Schedule i^{th} job to j^{th} cluster for execution
6:       **else if** $Num\_Free\_PE_j > Num\_PE_j/3$ **then**
7:          **if** $Num\_Running\_PE_j \geq Num\_Required\_PE_i$ **then**
8:             Schedule i^{th} job to j^{th} cluster for execution
9:       **else if** $Num\_Free\_PE_j > Num\_PE_j/4$ **then**
10:         **if** $Num\_Running\_PE_j \geq Num\_Required\_PE_i$ **then**
11:            Schedule i^{th} job to j^{th} cluster for execution
12:      **else if** $Num\_Free\_PE_j > Num\_PE_j/5$ **then**
13:         **if** $Num\_Running\_PE_j \geq Num\_Required\_PE_i$ **then**
14:            Schedule i^{th} job to j^{th} cluster for execution
15:      **else**
16:         **if** $Num\_Running\_PE_j \geq Num\_Required\_PE_i$ **then**
17:            Schedule i^{th} job to j^{th} cluster for execution
18:   update load status of C_j

---

shows that the first cluster which is having highest processing resources is overutilized, while seventh and eight clusters are medium utilized and second to sixth clusters are underutilized in all scheduling algorithms. It indicates uneven utilization of clusters and so the average utilization is only 26.70%, 25.94%, and 26.60% in FCFS, SJF, and EDF scheduling algorithms, respectively. The average utilization is increased to 39.52%, 39.89%, and 41.21% when scheduling algorithm plugged into P2S_DLB. The average utilization is improved by 12.83%, 13.95%, and 14.60% in FCFS_DLB, SJF_DLB, and EDF_DLB compared to FCFS, SJF, and EDF, respectively. The reason behind improvement is even utilization of all clusters which

Table 3 Cluster utilization in percentage

Cluster Id	Cluster Name	FCFS	FCFS_DLB	SJF	SJF_DLB	EDF	EDF_DLB
1	gaia-[1–60]	69.38	51.71	68.83	51.61	69.04	51.36
2	gaia-[61–62]	13.71	33.74	13.09	34.67	16.66	41.33
3	gaia-[63–72]	15.96	36.84	16.03	36.34	15.39	36.54
4	gaia-73	17.89	38.24	18.31	36.61	19.73	37.17
5	gaia-74	16.54	37.08	8.74	38.06	12.19	39.63
6	gaia-[75–79]	13.53	36.05	14.50	37.04	13.16	35.64
7	gaia-[80–119]	42.03	44.43	42.08	44.88	41.80	44.66
8	gaia-[120–151]	33.40	40.67	34.74	40.77	34.04	40.94
Average Utilization		26.70	39.52	25.94	39.89	26.60	41.21
Improvement in Average Utilization		12.83		13.95		14.60	

Table 4 Imbalance matric in percentage

Cluster Id	Cluster Name	FCFS	FCFS_DLB	SJF	SJF_DLB	EDF	EDF_DLB
1	gaia-[1–60]	159.89	30.84	165.34	29.39	159.53	24.64
2	gaia-[61–62]	48.65	14.63	49.54	13.09	37.37	0.29
3	gaia-[63–72]	40.22	6.78	38.20	8.90	42.14	11.33
4	gaia-73	33.00	3.24	29.41	8.22	25.83	9.80
5	gaia-74	38.05	6.17	66.31	4.59	54.17	3.83
6	gaia-[75–79]	49.33	8.78	44.10	7.14	50.53	13.52
7	gaia-[80–119]	57.42	12.42	62.22	12.51	57.14	8.37
8	gaia-[120–151]	25.09	2.91	33.92	2.21	27.97	0.66
Average Imbalance Metric		56.46	10.72	61.13	10.76	56.83	9.06

is achieved because P2S_DLB balances the load between clusters at the time of scheduling. It indicates that the P2S_DLB is performed better compared to traditional scheduling algorithms and performance remains stable with all the three scheduling algorithms.

Table 4 shows the percent of imbalance matric and average of imbalance matric for all algorithms. The percent of load imbalance is more than 50% for FCFS, SJF, and EDF with average imbalance matric value 56.46%, 61.13%, and 56.83%, respectively. The reason behind more amount of load imbalance is overutilization of the first cluster in all scheduling algorithms. The problem of overutilization is because of the scheduling algorithms which schedules all the incoming jobs to the first cluster though the other clusters are free or underutilized. This problem is resolved by combining scheduling algorithms with P2S_DLB which schedules incoming jobs to all clusters by looking current load status of cluster. The proposed technique balances the load between clusters, and because of that the percent of imbalance is very less in FCFS_DLB, SJF_DLB, and EDF_DLB with the average imbalance matric value 10.72%, 10.76%, and 9.06%, respectively. The results indicate that P2S_DLB balances the load effectively with all the three scheduling algorithms.

Figures 1, 3, and 5 show the cluster usage on day-to-day basis in graph format for FCFS, SJF, and EDF scheduling algorithms and Figs. 2, 4, and 6 show same for FCFS_DLB, SJF_DLB, and EDF_DLB dynamic load balancing algorithms. In graphs, the X-axis represents number of days and Y-axis represents the clusters in descending order of total number of CPUs. The total eight rows, which are painted with green, yellow, or red colors and separated by thin white lines, represent the each cluster utilization day-wise in graphs. The rows are painted for 90 days as the selected dataset is of 3 months. The green, yellow, and red colors represent under-, medium, and overutilization of clusters, respectively. In Figs. 1, 3, and 5, most of the portion of the first two rows is painted with red color, and in remaining six rows most of the portion is painted with green color which illustrates the first two clusters are overutilized while remaining six clusters remain underutilized in all three
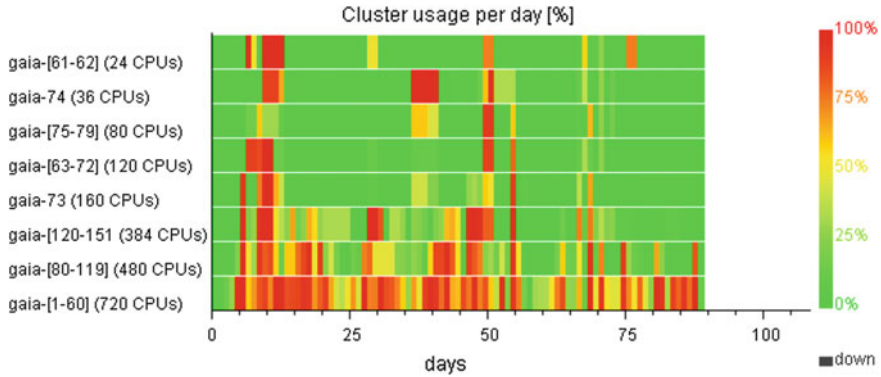
Fig. 1 Cluster utilization per day in FCFS
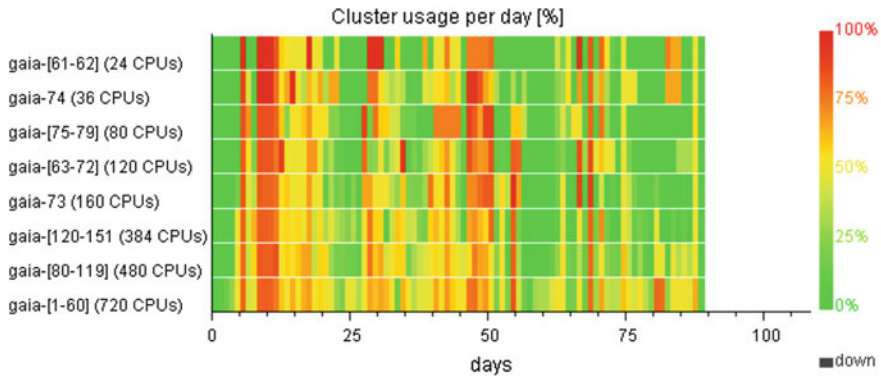
Fig. 2 Cluster utilization per day in FCFS_DLB

scheduling algorithms. In Figs. 2, 4, and 6, most of the portion of all eight rows is painted with yellow or green colors which indicates all eight clusters are medium utilized or underutilized in all three modified scheduling algorithms. The red painted portion is very less in Fig. 2, 4, and 6 which means overutilization of clusters is very less in modified scheduling algorithms. The results show that the idea of balancing load between clusters in parallel with job scheduling works, as the combination of P2S_DLB with scheduling algorithms improves the performance of all three algorithms.

7   Conclusion

The P2S_DLB algorithm plugged into traditional FCFS, SJF, and EDF scheduling algorithms. All three scheduling algorithms without and with combination of
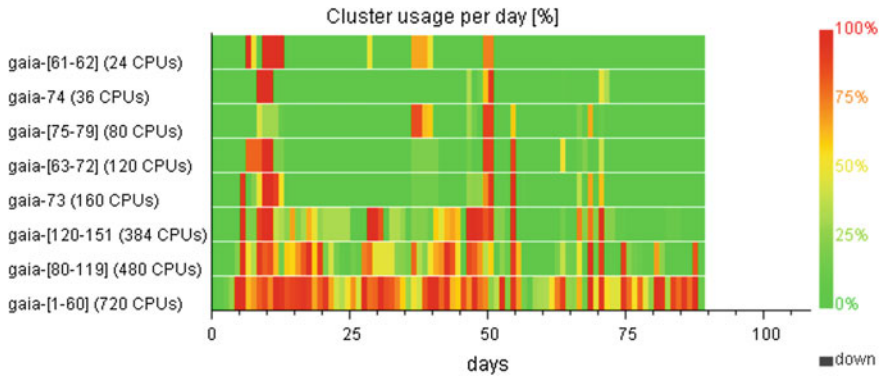
Fig. 3 Cluster utilization per day in SJF
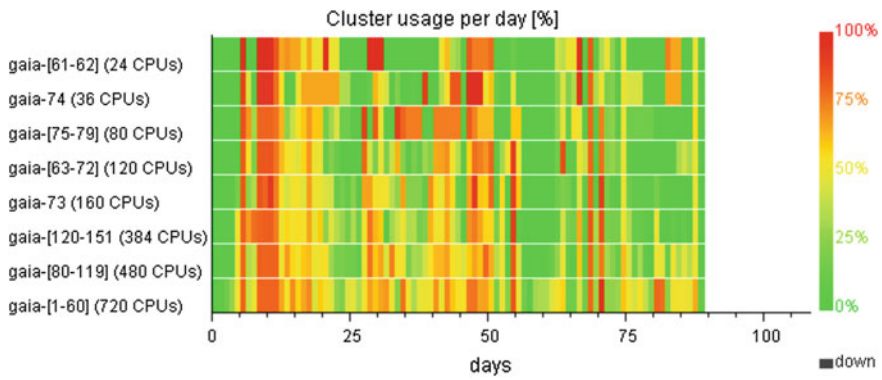
Fig. 4 Cluster utilization per day in SJF_DLB
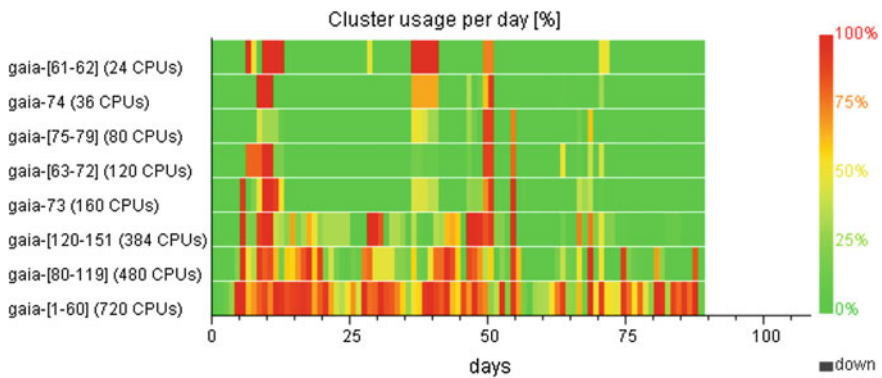
Fig. 5 Cluster utilization per day in EDF
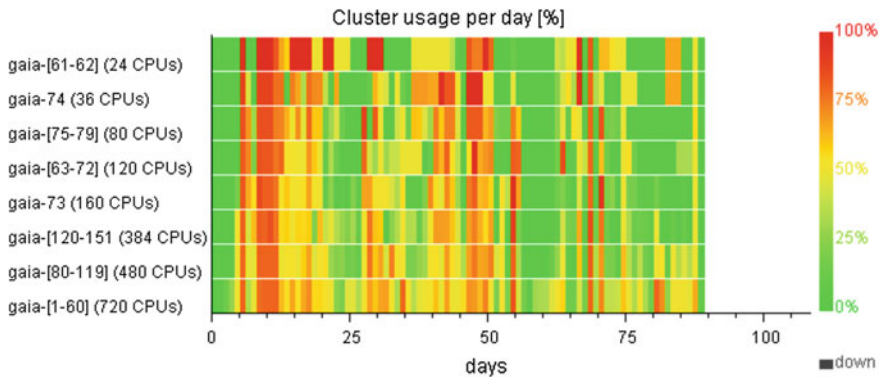
Fig. 6 Cluster utilization per day in EDF_DLB

P2S_DLB are simulated on real-time dataset of 3 months having 51,987 jobs in real-time cluster configuration using ALEA simulator. The cluster utilization and percent of imbalance matric are calculated as an evaluation parameter. The simulation results show that the P2S_DLB increased the average utilization of clusters and decreased the percent of imbalance matric for all the three scheduling algorithms. The performance of P2S_DLB remains stable with all the tree scheduling algorithms. The combination of EDF and P2S_DLB gives the best performance out of three scheduling algorithms.

References

1. Coulouris, G.F., Dollimore, J., Kindberg, T.: Distributed Systems: Concepts and Design. Pearson Education (2005)
2. Liu, M.L.: Distributed Computing: Concepts and Applications, 4th edn. Pearson Education (2009)
3. Buyya, R.: High Performance Cluster Computing: Architecture and Systems, volume i, vol. 1, p. 999. Prentice Hall, Upper SaddleRiver, NJ, USA (1999)
4. Azmi, Z.R.M., Bakar, K.A., Abdullah, A.H., Shamsir, M.S., Manan, W.N.W.: Performance comparison of priority rule scheduling algorithms using different inter arrival time jobs in grid environment. Int. J. Grid Distrib. Comput. **4**(3), 61–70 (2011)
5. Thakor, D., Patel, B.: P2s_dlb: Pluggable to scheduler dynamic load balancing algorithm for distributed computing environment. In: International Conference on Emerging Trends in Expert Applications & Security (ICETEAS), February 2018
6. Datta, L.: A new task scheduling method for 2 level load balancing in homogeneous distributed system. In: 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 4320–4325, March 2016
7. Abdelkader, D.M., Omara, F.: Dynamic task scheduling algorithm with load balancing for heterogeneous computing system. Egypt. Inf. J. **13**(2), 135–145 (2012)
8. Thakor, D., Patel, B.: Pdlb: An effective prediction based dynamic load balancing algorithm for clustered heterogeneous computational environment. In: 5th International Conference on Advanced Computing, Networking, and Informatics, 6 (2017)

9. Thakor, D., Patel, B.: Prediction based dynamic load balancing algorithm for distributed system. Nat. J. Syst. Inf. Technol. **9**(2), 67–76 (2016)
10. Pearce, O., Gamblin, T., de Supinski, B.R., Schulz, M., Amato, N.M.: Quantifying the effectiveness of load balance algorithms. In: Proceedings of the 26th ACM International Conference on Supercomputing. ICS '12, NY, USA, pp. 185–194 (2012)
11. Klusáček, D., Rudová, H.: Alea 2: Job scheduling simulator. In: Proceedings of the 3rd International ICST Conference on Simulation Tools and Techniques. SIMUTools '10, ICST, Brussels, Belgium, pp. 61:1–61:10. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2010)
12. Buyya, R., Murshed, M.: Gridsim: A toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing. Concur. Comput. Pract. Exp. **14**(13–15), 1175–1220 (2002)
13. Emeras, J.: Parallel workloads archive university of luxemburg gaia cluster. http://www.cs.huji.ac.il/labs/parallel/workload/l_unilu_gaia/. Accessed 15 Feb 2016
14. ULHPC: The gaia cluster hpc at university of luxemburg. https://hpc.uni.lu/systems/gaia/. Accessed 15 Feb 2016

Performance Enhancement of Hadoop for Big Data Using Multilevel Queue Migration (MQM) Technique

C. Sreedhar, N. Kasiviswanath and P. Chenna Reddy

Abstract The recent advancements in Hadoop MapReduce scheduling techniques have demonstrated significant outcomes. The continuous tradeoff between the data-job locality and synchronization results in the higher efficiency for the framework. Thus, large number of scientific and enterprise applications have adopted the parallel and synchronized mechanism through Hadoop framework. However, with this adaptation, a large number of datacenter-based nodes are been deployed, significantly causing the increase of energy consumptions. Henceforth, the demand of the recent research is to enhance the overall efficiency of Hadoop jobs and to decrease the energy consumption without degrading the performance. The recent advancements have demonstrated by many strategies by improving the Map and Reduce job allocation techniques; conversely, the same improvement can also be achieved through multilevel queues. Hence, this work constitutes the multilevel queue with custom load balancing to demonstrate the improvement in overall performance of Hadoop job scheduling. The work results in a significant improvement of Hadoop jobs in terms of execution times and energy consumption.

Keywords Hadoop · Big data · Scheduling · MapReduce
Multilevel queue

C. Sreedhar (✉) · N. Kasiviswanath
Computer Science & Engineering Department,
G Pulla Reddy Engineering College, Kurnool, India
e-mail: csrgprec@gmail.com

N. Kasiviswanath
e-mail: nkasiviswanath@yahoo.com

P. Chenna Reddy
Computer Science & Engineering Department,
Jawaharlal Nehru Technological University Anantapur, Anantapur, India
e-mail: pcreddy1@rediffmail.com

1   Introduction

In the recent years, datasets generated by digital devices are huge in terms of volume and distributed across various servers [1]. Hadoop is an open-source platform which can process data in various formats such as structured, semi-structured, and unstructured. A MapReduce job splits the given dataset into chunks of data such that the map task processes in a parallel manner [2]. The framework sorts the outputs of the maps, which are then inputted to the reduce tasks. Typically, both the input and the output of the job are stored in a file system. The framework takes care of scheduling tasks, monitoring them, and re-executing the failed tasks. Hadoop Distributed File System (HDFS) is a distributed file system which has master/slave architecture and can store huge volumes of data [3]. Default queues present in Hadoop job schedulers can no longer cope with several number of jobs/tasks made as a request to Hadoop. Hence, there is a need for multilevel queue mechanism which can process the jobs/tasks in parallel to meet the requirements of data analytics applications such as clustering large datasets, at a faster rate with increased size in data. The jobs submitted by the users in MapReduce framework are divided into map and reduce jobs. The map jobs are handled by process sharing queues, and the reduce jobs are handled by the server sharing queues [4]. After submitting the jobs into the MapReduce framework, the map and the reduce jobs are been allocated to the node level slots for the execution based on the availability [5]. The Hadoop framework encounters three major challenges as locality defined the distance between the job and the data for the job to processes, synchronization defined synchronization of the data from the map and reduce jobs, and finally the fairness defined comparative fairness between the locality and the synchronization [6]. However, the most critical challenge is the fairness of the framework. Many algorithms are been defined in order to solve this problem. Few of them are been understood and analyzed in later part of this work. Hence, this work demonstrates the use of multilevel queues for Yet Another Resource Negotiator (YARN)-based scheduling techniques to enhance the job performance. The other side of the MapReduce framework problem points toward the energy cost efficiency as the increasing number of nodes pushes the framework toward the massive energy-consuming structure [7]. Henceforth, large amount of researches are been carried out in order to find the most energy-efficient job scheduling and migration techniques. The energy efficiency of the Hadoop framework significantly depends on the workload and eventually calculated as amount of work process per energy units [8]. Hence, the demand of the current research is to provide energy-efficient job migration technique. Conversely, only the job migration based on the energy is not sufficient enough to provide the fairness of the job efficiencies; hence, the priority and other factors are also to be considered for the scheduling and migration techniques. There is a need for enhancing the Hadoop scheduling techniques through the use of multilevel queues. Here, the queues or pools can be configured based on the needs and requirements of the problem. Clustering very large datasets is one such problem where huge volumes of datasets are to be grouped based on the

similarities [9]. Large datasets are split into chunks of data and the task of finding similarities or unknown patterns can be achieved using several map and reduce jobs.

The remainder of this study is organized as follows. Various approaches developed recently to improve the efficiency of Hadoop schedulers are discussed in Related Work. Approaches recently developed to improve clustering results are discussed in Related Work. The proposed multilevel queue migration technique is described in Proposed Work. An extensive set of experiments with various configuration sets and comparison of job migration and job reallocation is presented in Experimental Setup and in Results and Discussion section. Finally, conclusions and future work are discussed in Conclusions.

2   Related Work

The FIFO is the default job scheduler in Hadoop framework and schedules the jobs based on the arrival into the queue [10]. Hadoop's built-in scheduler executes the map task in the job with data locality preference [11]. The basic principle of fair scheduler is to identify the available map–reduce slots and then allocate the slots to the queued jobs with a fair sharing [12] mechanism such that all the jobs get same amount of CPU times [13]. The fair scheduler deploys resource pools for each user and sometimes for per cluster. The fair scheduler can limit the number of concurrent jobs per user or per cluster by limiting the jobs per pool [14, 15]. By default, fair scheduler consists of a single default queue.

The fundamental technique used for capacity scheduler is very similar to the fair scheduler with an only significant difference as the jobs are allocated to the queues instead of pools [16]. The configurations of the queues are such that when a job is received by YARN, the YARN searches for the queue with the highest available capacity and assigns the job. The other major techniques for job allocation are the predefined allocation techniques like user- or group-based policies, where the jobs submitted by a specific user or group are allocated to a dedicated queue for the user or group [17].

The work of Nguyen et al. [18] proposed a scheduling technique based on the job priority to reduce the waiting time of the job with higher priority. However, the locality of the job is been underset in this algorithm. In the latter part of this work, the novel algorithm demonstrates the use of job priority with increased locality of the jobs. The work of Zaharia et al. [19] proposed a technique based on the shortest job first algorithm called Longest Approximate Time to End (LATE) for the heterogeneous Hadoop jobs. In this framework, the submitted and slow running jobs are analyzed based on the remaining time to complete and the jobs are been scheduled to higher ranked nodes for faster competition. Nevertheless, the localities of the jobs are certainly been overlooked to prioritize the job competition time. A Self-adaptive MapReduce Scheduling (SAMR) algorithm in heterogeneous environment [20] demonstrates the appropriate utilization of the nodes based on the

identification of the fast and slow nodes. The tasks running slow are migrated to the faster cluster, and hence the time for job completion is reduced. The SAMR is an enhancement of LATE scheduler.

Zaharia et al. [21] demonstrated the job scheduling for the specific jobs where the principle of locality cannot be obtained as the jobs requested at some clusters where the data is not present locally called delay scheduler. This framework works on a simple principle to address the jobs as if the current job in the queue cannot be started locally; then, the job is pushed back into the queue unless the locality of the data is achieved. The major drawback is the starvation issue for the long waiting jobs. This framework handles the starvation issue by starting the jobs locally in case of long waits.

Cassales et al. [22] demonstrated the context-aware scheduler where the map–reduce jobs are started based on the available nodes with regard to availability of CPU or I/O or network capacities. The work demonstrates a good strategy; however, the optimal configuration after the job migration is not been achieved in this framework.

3 Proposed Solution

The first phase of the algorithm analyzes the highest loaded node and migrates the jobs to the available less-loaded node. After identifying the source and destination, the algorithm identifies the jobs to be migrated based on the workload of data nodes. The outcome of this algorithm is to enhance the overall performance of Hadoop scheduling using Multilevel Queue Migration (MQM) technique. The following steps present the description of MQM technique.

The first step is to calculate the load on each node in the Hadoop clusters using Eqs. 1–5. This step must be repeated for all the nodes in the clusters in Hadoop environment where the clusters are considered to be heterogeneous.

$$Phy_{CPUCapacity} = \sum_{i=1}^{n} VM(i)_{CPUCapacity} \tag{1}$$

$$Phy_{MemoryCapacity} = \sum_{i=1}^{n} VM(i)_{MemoryCapacity} \tag{2}$$

$$Phy_{IOCapacity} = \sum_{i=1}^{n} VM(i)_{IOCapacity} \tag{3}$$

$$Phy_{NetworkCapacity} = \sum_{i=1}^{n} VM(i)_{NetworkCapacity} \tag{4}$$

$$\Pi = (Phy_{CPUCapacity} + Phy_{MemoryCapacity} + Phy_{IOCapacity} + Phy_{NetworkCapacity}) \qquad (5)$$

In the second step, the algorithm identifies the highest and lowest loaded node in the Hadoop environment using Eqs. 6 and 7.

$$\Pi_{MAX} = \begin{cases} If\ \Pi_i > \Pi_j, then\ \Pi_{MAX} = \Pi_i \\ Else\ \Pi_j > \Pi_i, then\ \Pi_{MAX} = \Pi_j \end{cases} \qquad (6)$$

$$\Pi_{MIN} = \begin{cases} If\ \Pi_i < \Pi_j, then\ \Pi_{MIN} = \Pi_i \\ Else\ \Pi_j < \Pi_i, then\ \Pi_{MIN} = \Pi_j \end{cases} \qquad (7)$$

Once the source and destination are identified as MAX and MIN, respectively, the identification of job to be migrated is carried out, and thus the identification of the virtual machine holding the job is done using Eqs. 8, 9, and 10.

$$Migratable\_Job_{Priority} = \begin{cases} \exists\ Job, Priority[Job(i)] > Priority[Job(j)], then\ Migratable\_Job \leftarrow Job(i) \\ Else\ Migratable\_Job \leftarrow Job(j) \end{cases}$$
$$(8)$$

$$Migratable\_Job_{Capacity} = \begin{cases} \exists\ Job, Capacity\_Required[Job(i)] > Capacity\_Required[Job(j)], \\ then\ Migratable\_Job \leftarrow Job(i) \\ Else\ Migratable\_Job \leftarrow Job(j) \end{cases}$$
$$(9)$$

$$Migratable\_Job = \begin{cases} If\ Migratable\_Job_{Priority} = Migratable\_Job_{Capacity}, \\ then\ Migratable\_Job \leftarrow Migratable\_Job_{Capacity} \\ Else\ Migratable\_Job \leftarrow Migratable\_Job_{Priority} \end{cases} \qquad (10)$$

After identifying the job to be migrated using the above Eq. 10, the virtual machine running the job is to be identified using Eq. 11.

$$VM(i) \leftarrow Identified\ VM \qquad (11)$$

Once virtual machine is identified, the job is to be migrated from source to destination as identified in the second step; the migration of the job internally takes place based on the migration of virtual machines. After the migration, the new load of the cluster is to be identified using Eq. 12.

$$VM(i) = VM(i)_{CPUCapacity} + VM(i)_{MemoryCapacity} + VM(i)_{IOCapacity} + VM(i)_{NetworkCapacity}$$
$$(12)$$

$$\Pi_{MAX} - VM(i) = \Delta_{Source} \qquad (13)$$

$$\Pi_{MIN} + VM(i) = \Delta_{Destination} \tag{14}$$

After the calculation of the new load using Eqs. 13 and 14, the source and destination nodes must obtain the efficient workload condition, where the loads are nearly equally balanced using Eq. 15.

$$\begin{cases} If \; \Delta_{Source} \approx \Delta_{Destination}, Then \; Migrate \; VM(i) \\ Else \; i = \; \in (n) \end{cases} \tag{15}$$

where n is the total number of virtual machines.

4 Experimental Setup

A Hadoop cluster consists of many machines in a network, each machine is termed as a node, and these nodes communicate to each other over the network. Large dataset is broken into chunks or splits, where each chunk is equal to the block size that is set for the HDFS cluster (which is 64 MB by default). In HDFS, there are two types of nodes: a master node (NameNode) and worker node (SlaveNode). Several experiments were conducted on various configurations to evaluate the efficiency and performance comparison of our proposed solution. The experiments were conducted with the Intel Core i3-5005U CPU 2.66 GHz and RAM capacity of 12 GB and with four processor cores. The slave nodes also referred as datanodes with the configuration of Intel Core i3-5005U CPU 2.66 GHz and with specified RAM capacity of 2 GB and with two processor cores. The allocation rule can be identified here, where the jobs submitted by the respective users will be allocated to the specified queue by default. Multilevel queues can be configured based upon the objective of the problem, for example, clustering. In the problem of clustering, multilevel queues can be formed based upon the number of clusters needed. Multilevel queue configuration, capacity scheduler multilevel queue setup, and fair scheduler multilevel queue setup are discussed in [23].

5 Results and Discussion

This work demonstrates the results of configuration with CS-1U-1J, FS-1U-1J–FIFO and FS-1U-1J–FAIR, where the performance evaluation parameters such as local storage cost, HDFS storage cost, map job waiting time, reduce job waiting time, total job time, and total job waiting time are been compared as shown in Table 1.

Hence, this clearly demonstrates with a proper configuration of the queue, and the waiting times for the jobs are nearly zero. Next, this work demonstrates that the

Table 1 Performance comparison for configuration set-1

Experiment name	Local storage cost (KB)	HDFS storage cost (KB)	Map job waiting time (ms)	Reduce job waiting time (ms)	Total job time (ms)
CS-1U-1J	8494	208	0	0	1011018
FS-1U-1J–FIFO	8493	208	0	0	2194453
FS-1U-1J–FIFO	8493	208	0	0	2510011

Table 2 Performance comparison for configuration set-2 (User-1)

Experiment name	Local storage cost (KB)	HDFS storage cost (KB)	Map job waiting time (ms)	Reduce job waiting time (ms)	Total job time (ms)
CS-2U-2J	8494	208	0	0	1562530
FS-2U-2J–FIFO	8493	208	0	0	2217508
FS-2U-2J–FIFO	4595	123	0	0	1160195

Table 3 Performance comparison for configuration set-2 (User-2)

Experiment name	Local storage cost (KB)	HDFS storage cost (KB)	Map job waiting time (ms)	Reduce job waiting time (ms)	Total job time (ms)
CS-2U-2J	4595	123	0	0	556563
FS-2U-2J–FIFO	4595	123	0	0	1207336
FS-2U-2J–FIFO	8493	208	0	0	1738994

results of configuration set-2, for user-1 and user-2 are shown in Table 2 and Table 3, respectively.

The results clearly demonstrate the reduction of the local and HDFS storage cost for the configuration, where multiple users submit multiple jobs with weighted queue configurations. Next, this work demonstrates the results of configuration with CS-3U-3J, FS-3U-3J–FIFO, and FS-3U-3J–FAIR where the performance evaluation parameters such as local storage cost, HDFS storage cost, map job waiting time, reduce job waiting time, total job time, and total job waiting time are been compared for user-1, user-2, and user-3 and shown in Table 4, Table 5, and Table 6, respectively.

The results clearly demonstrate the reduction of the local for the configuration, where multiple users submit multiple jobs with weighted queue configurations. The waiting time increases significantly for fair scheduler with FIFO configuration and

Table 4 Performance comparison for configuration set-3 (User-1)

Experiment name	Local storage cost (KB)	HDFS storage cost (KB)	Map job waiting time (ms)	Reduce job waiting time (ms)	Total job time (ms)
CS-3U-3J	8494	208	0	0	764193
FS-3U-3J–FIFO	8493	208	77878	0	814368
FS-3U-3J–FAIR	8493	208	0	0	2173455

Table 5 Performance comparison for configuration set-3 (User-2)

Experiment name	Local storage cost (KB)	HDFS storage cost (KB)	Map job waiting time (ms)	Reduce job waiting time (ms)	Total job time (ms)
CS-3U-3J	4595	123	0	0	389306
FS-3U-3J–FIFO	4595	123	0	0	370165
FS-3U-3J–FAIR	4595	123	0	0	1253263

Table 6 Performance comparison for configuration set-3 (User-3)

Experiment name	Local storage cost (KB)	HDFS storage cost (KB)	Map job waiting time (ms)	Reduce job waiting time (ms)	Total job time (ms)
CS-3U-3J	4596	124	0	0	635978
FS-3U-3J–FIFO	4595	124	0	0	1143084
FS-3U-3J–FAIR	4595	124	0	0	1225295

remains zero for other configurations. The second part of the results demonstrates the effect of proposed job migration technique with the comparison of all existing 18 techniques, as shown in Table 7, on the parameters like energy consumption in kWh, execution time—host selection mean in seconds, execution time—VM reallocation mean in seconds, and execution time—total mean in seconds.

Comparative analysis of the energy consumption is been undertaken, where this work demonstrates the energy consumption difference with existing techniques. Comparative analysis of the job reallocation time is been undertaken, where this work demonstrates the job reallocation time difference with existing techniques shown in Table 8.

Next, comparative analysis of the job/host selection time is been undertaken, where this work demonstrates the job/host selection time difference with existing techniques. Hence, the result of job/host selection time shows nearly a significant

Table 7 List of techniques used for performance comparison

Used name in this work	Selection policy	Allocation policy
IQR MC	Maximum Correlation	Inter Quartile Range
IQR MMT	Minimum Migration Time	Inter Quartile Range
LR MC	Random Selection	Local Regression
LR MMT	Minimum Migration Time	Local Regression
LR MU	Minimum Utilization	Local Regression
LRR MC	Maximum Correlation	Robust Local Regression
LRR MMT	Minimum Migration Time	Robust Local Regression
LRR MU	Minimum Utilization	Robust Local Regression
LR RS	Robust Local Regression	Rom Selection
MAD MC	Maximum Correlation	Median Absolute Deviation
MAD MMT	Minimum Migration Time	Median Absolute Deviation
MAD MU	Minimum Utilization	Median Absolute Deviation
MAD RS	Rom Selection	Median Absolute Deviation
THR MC	Maximum Correlation	Static Threshold
THR MMT	Minimum Migration Time	Static Threshold
THR MU	Minimum Utilization	Static Threshold
THR RS	Rom Selection	Static Threshold
IQR MU	Inter Quartile Range	Minimum Utilization
MQM ALGO	Proposed Algorithm Part-1	Proposed Algorithm Part-2

80% improvement over the existing techniques and 66.67% lesser than the existing techniques. Hence, the result of execution time shows nearly a significant 73% improvement over the existing techniques and 87.5% lesser than the existing techniques.

Hence, the result of job reallocation time shows nearly a significant 70% improvement over the existing techniques and 83.35% lesser than the existing techniques.

6 Conclusion

The work demonstrates the existing techniques with understanding of the significance in time and data complexity paradigm like FIFO scheduler, fair scheduler, capacity scheduler, dynamic priority based hybrid scheduler, LATE scheduler, SAMR scheduler, delay scheduler, and context-aware scheduler. This paper presents MQM technique, which can efficiently schedule the jobs using multilevel queues and produce fair amount of test results to prove the novelty and improvement in energy consumption. The result shows nearly a significant 30% improvement of execution times over the existing techniques and 88.89% lesser than the existing techniques along with a significant 70% improvement over the existing

Table 8 Performance comparison for job migration–job reallocation time

Experiment name	No. of hosts	No. of VMs	Total simulation time (s)	Execution time–VM reallocation mean (s)	Change	Change in %	Improvement/ Deterioration
IQR MC	50	50	86400.00	0.00395	0.001220	31	Improvement
IQR MMT	50	50	86400.00	0.00694	0.004210	61	Improvement
LR MC	50	50	86400.00	0.00264	−0.000090	−3	Deterioration
LR MMT	50	50	86400.00	0.00318	0.000450	14	Improvement
LR MU	50	50	86400.00	0.00213	−0.000600	−28	Deterioration
LRR MC	50	50	86400.00	0.00222	−0.000510	−23	Deterioration
LRR MMT	50	50	86400.00	0.00607	0.003340	55	Improvement
LRR MU	50	50	86400.00	0.00672	0.003990	59	Improvement
LR RS	50	50	86400.00	0.00274	0.000010	0	Improvement
MAD MC	50	50	86400.00	0.00736	0.004630	63	Improvement
MAD MMT	50	50	86400.00	0.01822	0.015490	85	Improvement
MAD MU	50	50	86400.00	0.02090	0.018170	87	Improvement
MAD RS	50	50	86400.00	0.01400	0.011270	81	Improvement
THR MC	50	50	86400.00	0.00469	0.001960	42	Improvement
THR MMT	50	50	86400.00	0.00379	0.001060	28	Improvement
THR MU	50	50	86400.00	0.00390	0.001170	30	Improvement
THR RS	50	50	86400.00	0.00241	−0.000320	−13	Deterioration
IQR MU	50	50	86400.00	0.00954	0.006810	71	Improvement
MQM ALG	**50**	**50**	**86400.00**	**0.00273**	0.02130	91	Improvement

techniques for job reallocation time. Significant achievement of this work also includes the coining the novel performance evaluation matrix for comparative understanding of the proposed and existing technique. The final outcome of the work demonstrates the significance of queue scheduling to achieve similar performance. The future work presents testing the proposed solution for large datasets.

References

1. De Camargo, R.Y., Goldchleger, A., Kon, F.: InteGrade: a tool for executing parallel applications. J. Grid Comput. (2005)
2. Dittrich, J., Quiane-Ruiz, J.-A.: Efficient Big Data Processing in Hadoop MapReduce. In: VLDB Endowment, Vol. 5, No. 12 (2012)
3. Pinedo, M.: Scheduling: Theory, Algorithms, and Systems, 3rd edn. Springer Science, Heidelberg (2008)
4. Morton, K., Balazinska, M., Grossman, D.: ParaTimer—a progress indicator for MapReduce DAGs. In: International Conference on Management of Data, pp. 507–518 (2010)
5. Wei, L.: Efficient processing of k nearest neighbor joins using MapReduce. VLDB Endowment **5**(10), 1016–1027 (2012)
6. Dean, J., Ghemawat, S.: MapReduce—simplified data processing on large clusters. In: Symposium on Operating System Design and Implementation (2004)
7. Hadoop. Hadoop home page: http://hadoop.apache.org/
8. Hadoop's Fair Scheduler. https://hadoop.apache.org/docs/r1.2.1/fair_scheduler
9. Kavulya, S.: An analysis of traces from a production MapReduce cluster. In: 10th IEEE/ACM CCGrid, pp. 94–103 (2010)
10. Liu, S., Xu, J., Liu, Z., Liu, X.: Evaluating task scheduling in hadoop-based cloud systems. In: IEEE International Conference on Big Data (2013)
11. Ren, Z., Xu, X., Wan, J., Shi, W., Zhou, M.: Workload characterization on a production Hadoop cluster: a case study on Taobao. In: IEEE International Symposium on Workload Characterization (IISWC2012), pp. 3–13 (2012). https://doi.org/10.1109/iiswc.2012.6402895
12. Pinedo, M.: Scheduling—Theory, Algorithms, and Systems, 3rd edn. Springer Science, Heidelberg (2008)
13. Baptiste, P.: Scheduling equal-length jobs on identical parallel machines. Discrete Appl. Math. **103**(1), 21–32 (2000)
14. Max-Min Fairness (Wikipedia). http://en.wikipedia.org/wiki/Max-min_fairness
15. Isard, M., Prabhakaran, V., Currey, J., Wider, U., Talwar, K., Goldberg, A.: Quincy—fair scheduling for distributed computing clusters. In: ACM SIGOPS, pp. 261–276 (2009)
16. Chen, J., Wang, D., Zhao, W.: A Task Scheduling algorithm for hadoop platform. J. Comput. **8**(4), 929–936 (2013)
17. Tiwari, N.: Scheduling and energy efficiency improvement techniques for Hadoop MapReduce: state of art and directions for future research. Doctoral dissertation, Indian Institute of Technology, Bombay Mumbai
18. Nguyen, P., Simon, T., Halem, M., Chapman, D., Le, Q.: A hybrid scheduling algorithm for data intensive workloads in a MapReduce environment. In: IEEE/ACM Fifth International Conference on Utility and Cloud Computing. UCC'12, pp. 161–168. Washington, DC, USA: IEEE Computer Society (2012)
19. Zaharia, M., Konwinski, A., Joseph, A.D., Katz, R., Stoica, I.: Improving MapReduce performance in heterogeneous environments. In: USENIX Symposium on Operating Systems Design and Implementation (2008)

20. Chen, Q., Zhang, D., Guo, M., Deng, Q., Guo, S.: A self-adaptive MapReduce scheduling algorithm in heterogeneous environment. In: 10th IEEE International Conference on Computer and Information Technology (2010)
21. Zaharia, M., Borthakur, D., Sarma, J.S., Elmeleegy, K., Shenker, S., Stoica, I.: Delay scheduling—a simple technique for achieving locality and fairness in cluster scheduling. In: 5th European Conference on Computer Systems (EuroSys). ACM Press (2010)
22. Cassales, G.W., Charao, A.S., Pinheiro, M.K., Souveyet, C., Steffene, L.A.: Context-aware scheduling for Apache Hadoop over pervasive environments. Comput. Sci. **52**, 202–209 (2015)
23. Sreedhar, C., Kasiviswanath, N., Chenna Reddy, P.: A novel multilevel queue based performance analysis of Hadoop job schedulers. Indian J. Sci. Technol. **9**(44) (2016). https://doi.org/10.17485/ijst/2016/v9i44/96414

Semi-automatic Ontology Builder Based on Relation Extraction from Textual Data

Anjali Thukral, Ayush Jain, Mudit Aggarwal and Mehul Sharma

Abstract This paper proposes a semi-automated tool to build ontology from text. The tool consists of an analyzer to parse the given text and a mapper that maps NLP triple to RDF triple under user supervision. The resulted RDF triple is then validated through "triple validator" for its existence in the ontology. The triple is augmented to the ontology if it does not exist. System learns during this process and provides better mapping suggestions with time, making ontology building faster.

Keywords Ontology builder · Text to ontology · Concept ontology
RDF triple

1 Introduction

Ontology is a description (like a formal specification of a program) of the concepts and relationships that can formally exist for an agent or a community of agents [1]. A domain ontology (or concept ontology) represents the conceptual model underlying a certain domain, describing it in a declarative way [2]. In the present time, majority of the data is in the textual form such as journals, documents, and website. Converting these unstructured data into ontologies requires a lot of time and manual work. So, there is high requirement of some kind of automation to make the task of conversion easier. To fulfill this requirement, we developed a tool that converts text into concept ontology with minimum human intervention. The tool uses three main methods relation extraction, rule-based mapping, and validator and storage which are described in later sections. This paper is organized as follows: Sect. 2 presents the related work on building ontologies. Details of the proposed system are provided in Sect. 3. Section 4 provides performance evaluation of the system, followed by the conclusion in Sect. 5.

A. Thukral (✉) · A. Jain · M. Aggarwal · M. Sharma
Keshav Mahavidyalaya University of Delhi, Delhi, India
e-mail: athukral@keshav.du.ac.in

2   Related Work

A lot of attention has been given toward the building of ontology over the years.
A huge set of efforts have been put in the field of automatic building of ontologies,
mainly focusing on domain (concept) ontology. Maddi et al. [3] presented the
method of extraction of ontologies for text documents using linear algebraic method
called as single-value decomposition for obtaining the concepts from terms and
used bipartite ontology graphs to represent the results. Yeh and Yang [4] proposed a
method of automatic ontology creation for historical documents from digital library
using latent topic extraction and topic clustering. Bedini and Nguyen [5] presented
a framework that evaluates the automation of ontology generation. The work also
provides a comprehensive analysis of existing software. The work of Moreno and
Sanchez [6] provides a methodology to build ontologies automatically and
extracting information from web documents. The work of Gantayat and Iyer [7] for
automatic building of ontology from lecture notes is available at several courseware
repositories. It uses NLP for keyword extraction, term frequency inverse document
frequency (tf-idf) for extracting the concepts from keywords, and "apriori
algorithm" to determine associations among concepts. Gillam and Ahmad [8] have
proposed statistical methods for extracting the concepts from the text. Most of the
work done for building of ontology is for a specific domain, whereas the proposed
tool is a generic system that can build any ontology in less time under the domain
expert (user) supervision, irrespective of its domain. Moreover, final decision is in
the hands of domain expert, so it prevents the insertion of ambiguous and incorrect
information into the ontology. This tool generates standard OWL ontology which
can be accessed and manipulated in ontology editors such as protégé [9].

3   Proposed Approach

The proposed tool is a user-friendly generic semi-automated ontology builder
which uses relation extraction and rule-based mapping procedure to build concept
ontologies from text (Please refer Fig. 1). The ontology builder creates the concept
ontology, besides domain learning during the building process. With time and
enough learning, the system is able to produce concept ontology in less time with
minimal human intervention. It prunes the input texts and converts it into basic
form, i.e., triples (subject, predicate, and object) using relation extraction and adds it
to ontology after validation from user. If builder is not able to recognize the
semantics of the information, user will provide the semantics, which will intern add
to the learning of the software, so that it becomes capable to recognize the similar
semantics of the information in future and is able to add the correct information in
the ontology.

The proposed system contains three modules analyzer subsystem, mapper sub-
system, and triple validator and storage. Home screen of the tool provides the
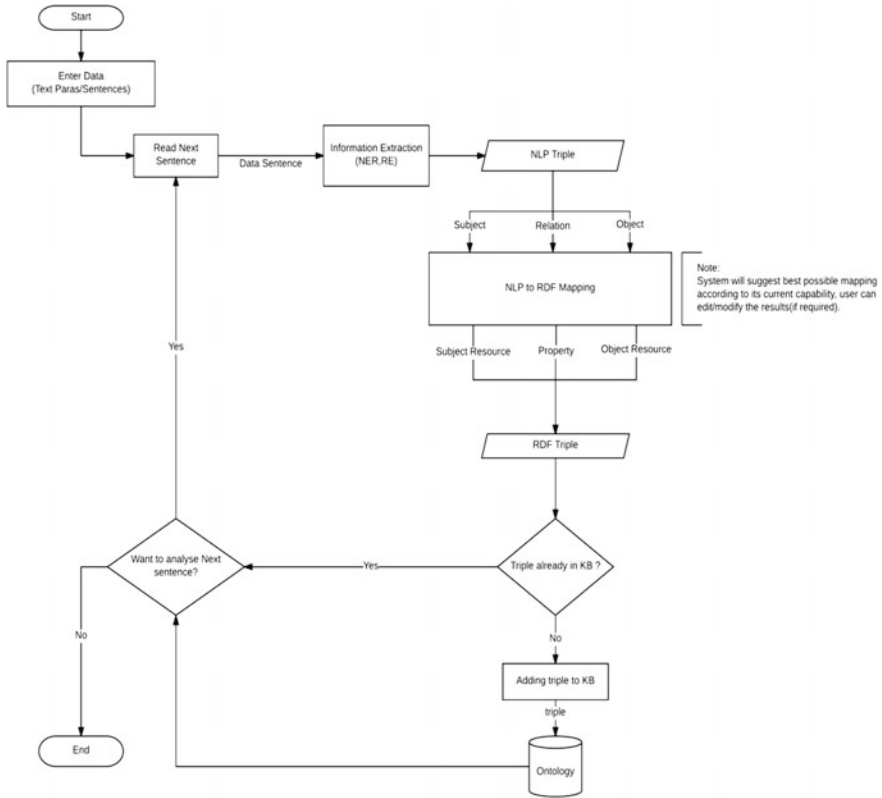
Fig. 1 Proposed system flow

interface for the features like starting a new ontology building session, viewing the current ontology file, importing and exporting of OWL file, rule file, and performance CSV file which are explained below.

3.1 Analyzer Subsystem

Analyzer subsystem performs the task of relation extraction. It processes the input text and presents all the possible relations that can be extracted from the given sentence by passing the text through OpenIE component [10] of "Stanford CoreNLP API [11]". User then selects the best relation among the system suggestions, which is then passed to the mapper subsystem for further processing.
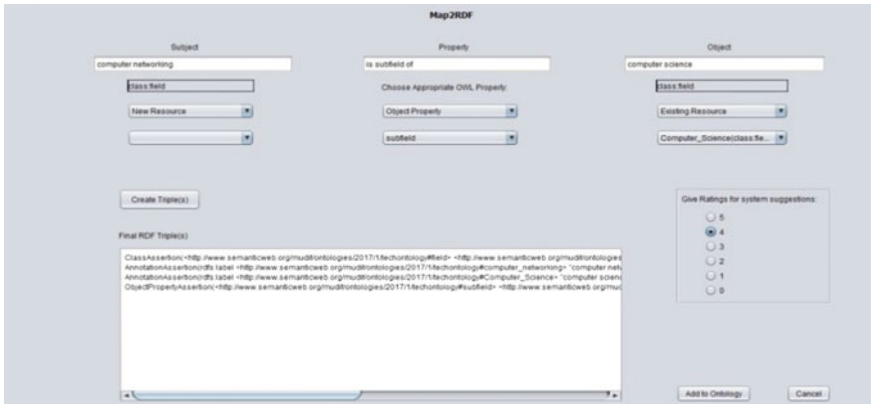
Fig. 2 Mapper screen

3.2 Mapper Subsystem

The mapper uses a relation that was extracted using analyzer, schema of the concept ontology, and the rule file which defines certain mapping rules to map predicate text to an appropriate OWL property. Mapping of subject/object text is done using label annotations of existing individual in the ontology. If the predicate text is mapped to a data property, then object and subject text are mapped to a literal and an individual, respectively. But, both subject and object texts are mapped to an individual if predicate is mapped to an object property.

Rule File
The rule file contains rule of format "Predicate text" → "OWL property IRI". Rules are arranged in the lexical order of predicate text for efficient searching. During the mapping process whenever user explicitly maps predicate text to an OWL property, then corresponding rule is added to the rule file.

Working of Mapper
The mapper first tries to map subject, predicate, and object itself (Ref. Fig. 2). In case the mapper does not find an appropriate solution, it recommends the possible options to be selected by a user manually. After mapping all the components of relation, it creates appropriate OWL triples using OWL API [12]. User is asked to give a feedback from scale of 0–5 as on what level tool was helpful. User feedback is saved in a CSV file with the help of Java CSV API.[1] The resulted triples are then passed to the next module, i.e., Triple validator and storage system.

---

[1]https://www.csvreader.com/java_csv.php.

3.3 Triple Validator and Storage

Validator module checks for the existence of all the triples received from mapper, in the ontology either as it is or in the form of inference from the existing information in ontology using HermiT Reasoner [13]. "Already Exists" status depicts that triple information is already present in the ontology, and "New" status depicts that the triple is added to the ontology as it was not there earlier. After adding all the triples, ontology is checked for the consistency using HermiT Reasoner [13]. In case of inconsistency, changes are not reflected in the original ontology file.

4 Experimental Results

We created ontology on "Technology Classification" for experimental purpose. The ontology classifies various fields and subfields of science and technology. The ontology and relevant text data (collected from various sources) were fed into the system. Experiment was performed over four sessions under supervision of domain expert to display the building process. Growth of ontology, generated rule file, and performance of the system were captured. Ontology graph was created using SOVA plug-in[2] for Protégé ontology editor [9].

4.1 Technology Classification Ontology

Ontology is illustrated in Fig. 3. Ontology schema "Technology Classification" fed to the application was created using protégé ontology editor [9] and was fed to the system as ontology schema.

Figure 4 illustrates the populated "Technology Classification" ontology which was obtained after passing significant amount of text data into the system.

4.2 Rule File Generation

All rules formed during the mapping procedure for the predicate text part of the relation are stored in rule file. The system refers to this rule file for mapping purpose.

---

[2]https://protegewiki.stanford.edu/wiki/SOVA_1.0.0.

Fig. 3 Ontology schema
"Technology Classification"
fed to the application
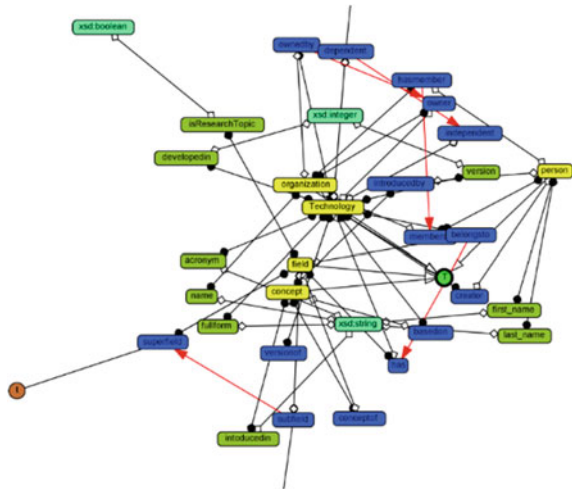
Fig. 4 Ontology obtained
after processing through
application
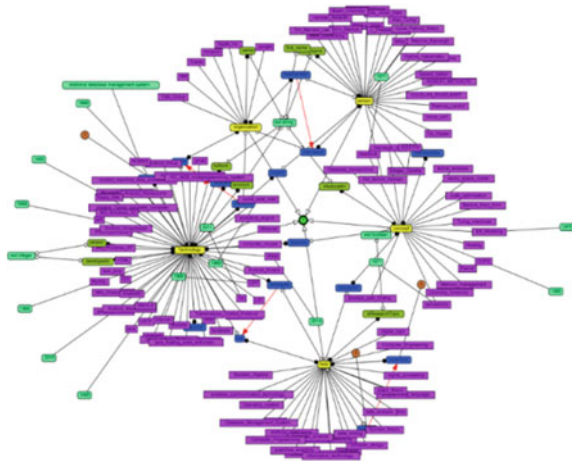
4.3 Performance

A user may follow the guidelines (tabulated in Fig. 5) while providing feedback
regarding the mapping accuracy.

For subject/object, 0, 1, and 2 stand for new resource, existing resource but not
mapped by the system, and existing resource which is mapped by the system,
respectively. Similarly, for predicate, 0 and 1 stand for predicate which was not
mapped to a property and mapped to a property, respectively.

Figure 6 shows the performance of the tool during four sessions. It shows that as
the process of ontology building continues, the system is able to efficiently convert
text to RDF triples, making the ontology building process faster.

Fig. 5 Feedback guidelines
regarding mapping

Subject	Predicate	Object	Rating
0	0	0	0
0	0	1	0
0	0	2	1
0	1	0	1
0	1	1	1
0	1	2	3
1	0	0	0
1	0	1	1
1	0	2	1
1	1	0	1
1	1	1	2
1	1	2	4
2	0	0	1
2	0	1	1
2	0	2	2
2	1	0	3
2	1	1	4
2	1	2	5
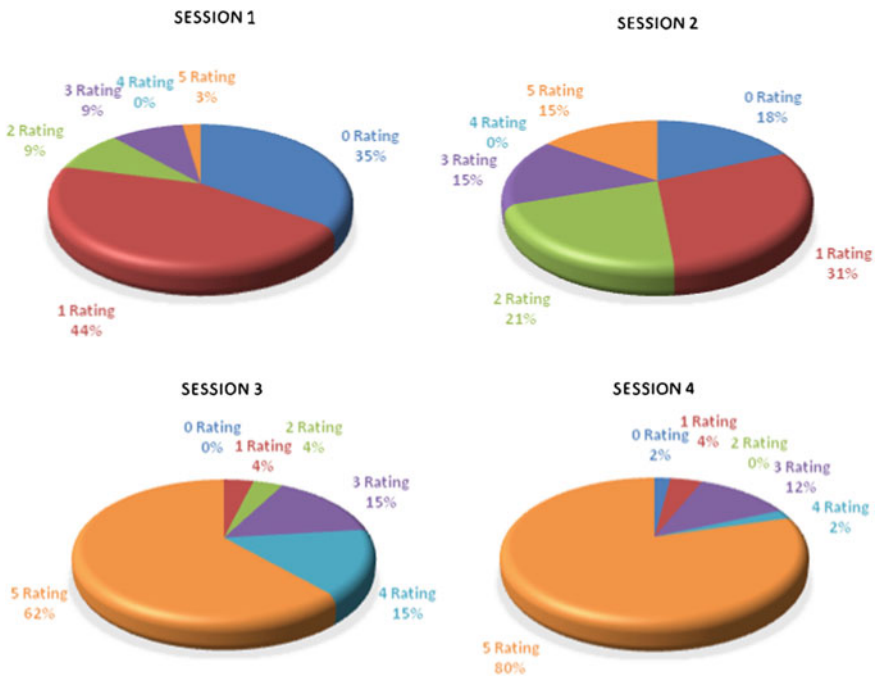
Fig. 6 The tool performance (rating based)

5 Conclusion

A tool for building ontology from text was proposed in the paper. The design of the proposed tool is modular with three components: analyzer subsystem, mapper subsystem, and triple validator and storage. The experiment was performed on technology classification ontology. It was noted that system's performance and results are greatly dependent upon the ontology schema and the input text fed to the tool. The ontology building is a learning process, and the perceived knowledge is a result of user's interaction with the system. It was also noted that whatever be the ontology, after converting a certain amount of text (which covers good amount of domain's knowledge) into triples, the system was able to get trained and was able to make correct recommendations. As the future work, the analyzer subsystem module can be improved to handle complex sentences.

References

1. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing. Int. J. Hum. Comput Stud. **43**(5–6), 907–928 (1995)
2. Bedi, P., Thukral, A., Banati, H., Behl, A., Mendiratta, V.: A multi-threaded semantic focused crawler. J. Comput. Sci. Technol. **27**(6), 1233–1242 (2012)
3. Maddi, G.R., Velvadapu, C.S., De Lamadrid, J.G.: Ontology extraction from text documents by singular value decomposition (2001)
4. Yeh, J.H., Yang, N.: Ontology construction based on latent topic extraction in a digital library. In: International Conference on Asian Digital Libraries, pp. 93–103. Springer, Berlin, Heidelberg (2008)
5. Bedini, I., Nguyen, B.: Automatic ontology generation: state of the art. PRiSM Laboratory Technical Report. University of Versailles (2007)
6. Sanchez, D., Moreno, A.: Creating ontologies from web documents. In: Recent Advances in Artificial Intelligence Research and Development, vol. 113, pp. 11–18. IOS Press (2004)
7. Gantayat, N., Iyer, S.: Automated building of domain ontologies from lecture notes in courseware. In: 2011 IEEE International Conference on Technology for Education (T4E), pp. 89–95. IEEE (2011)
8. Ahmad, K., Gillam, L.: Automatic ontology extraction from unstructured texts. In: On the Move to Meaningful Internet Systems 2005: Coopis, Doa, and Odbase, pp. 1330–1346 (2005)
9. Musen, M.A.: The protégé project: a look back and a look forward. AI Matters **1**(4), 4–12 (2015)
10. Angeli, G., Premkumar, M.J., Manning, C.D.: Leveraging linguistic structure for open domain information extraction. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (ACL 2015) (2015)
11. Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J.R., Bethard, S., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: ACL (System Demonstrations), pp. 55–60 (2014)
12. Horridge, M., Bechhofer, S.: The OWL API: a Java API for OWL ontologies. Semant. Web **2**(1), 11–21 (2011)
13. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z.: HermiT: an OWL 2 reasoner. J. Autom. Reason. **53**(3), 245–269 (2014)

An Ideal Approach for Medical Color Image Enhancement

Dibya Jyoti Bora

Abstract Medical images are poorly illuminated and often suffer from low contrast so it needs enhancement before further processing. In this paper, we have introduced an ideal approach for medical color image enhancement which is based on type-2 fuzzy set with unsharp masking based post-processing. The proposed technique has been tested in various poor contrast medical color images and found the results superior to the other traditional state-of-the-art algorithm.

Keywords Color medical image · Contrast improvement · CIELAB
Image sharpening · Medical color image enhancement · Type-2 fuzzy set
Unsharp masking

1 Introduction

In medical imaging area, the medical images are used for various diagnostic purposes and help to visualize the affected organ of a patient in a better way [1, 2]. But it is found that medical images often suffer from poor illumination due to low contrast and the important regions are not properly visible due to the fuzzy nature of the boundaries [3]. So, a better diagnosis of a patient cannot be expected from these poor quality images. Here arises the need of an efficient enhancement technique for medical images which will have the capability to deal with the vague behavior of medical images. Crisp-based techniques are not capable of dealing with the vagueness property of medical images [3]. Again, enhancement of color medical images needs a special treatment due to the color computations involved [1]. Here, we have to apply the enhancement technique very carefully so that the hue values will not be affected. Keeping these things in mind, in this paper we have introduced an efficient approach for medical color image enhancement. This technique is based

D. J. Bora (✉)
Department of Computer Science & Applications, Barkatullah University,
Bhopal, India
e-mail: research4dibya@gmail.com

on the concept that the enhancement can be achieved without altering the hue values if we target only the intensity channel [4–6]. For this, the first requirement will be a suitable color space with a devoted intensity channel. As RGB does not fit in this condition, we have selected LAB color space for our proposed approach. Then, a type-2 fuzzy set based technique is used for the contrast enhancement of the intensity channel of LAB-converted image. Finally, the extra blurriness that usually appears after fuzzy enhancement is removed by an unsharp masking based technique.

The outline of the remaining portion of the paper is as follows: Sect. 2 reviews previous work done in the field. In Sect. 3, the steps involved in the proposed approach are presented. In Sect. 4 and its subsections, the technical concepts involved in the proposed approach are explained. Experiments are demonstrated in Sect. 5. Finally, the conclusion is drawn in Sect. 6.

2 Related Work and Motivation Toward the Proposed Work

Gu et al. [1], proposed a new enhancement technique for medical color image enhancement by Young–Helmholtz (Y-H) transformation with the adaptive equalization of intensity numbers matrix histogram. They enhance the contrast by adaptive histogram equalization and thereby suppress the noises present in the original image. The inverse transformation is carried out from Y-H to RGB to show the enhanced color image without affecting hue and saturation values. They claim through experiments that the proposed technique is well suitable for diagnosis of medical images as it carries low computational complexity.

Lin and Lin [7] proposed a novel method Fuzzy Automatic Contrast Enhancement (FACE) which first performs fuzzy clustering to segment the input image. The pixels with similar colors in the CIELAB color space are classified into similar clusters with smaller characteristics.

Hsu and Chou [8] introduced a medical image enhancement technique based on modified color histogram equalization where they focus on hue preservation while doing the enhancement. For this, they employed two methods. The first one tries to obtain an equalized color image preserving hue by using the ratio of the original grayscale image and the equalized image. The another one achieves hue preservation by applying the difference between the original gray version of the image and the equalized one obtaining the final equalized color image.

Chaira [3] suggested intuitionistic fuzzy set theory based image enhancement techniques as it takes into account two uncertainties and there is a good possibility of better enhancement of medical images. The author also proposed [9] type-2 fuzzy set based medical image contrast enhancement technique. She used Hamacher T co-norm as an aggregation operator to form a new membership function with proper upper and lower membership functions. The image with the new membership function is found as an enhanced one of the original one.

Zhao and Zhou [10] introduced an efficient unsharp masking based medical image enhancement technique for suppressing background noise where they embed the PLIP multiplication into the unsharp masking framework.

From the literature, it is found that medical images often suffer from vagueness and hence it requires a strong fuzzy set based technique to deal with this high level of uncertainties exist in the intensity channel of the medical image. Type-2 fuzzy set based technique has this special benefit in this case. Unsharp masking filtering technique removes the burliness to a very promising level. Also, only a few good techniques are available for enhancing color medical images. So, it motivates us to develop a color medical image enhancement technique which will have the capability of dealing a high level of uncertainties as well as sharpen the image by removing the blurriness which usually adds a redundant layer of brightness.

3 Steps Involved in the Proposed Algorithm

The proposed algorithm consists of four steps which are stated below:

Step1. *Input poor contrast medical color image. Color space conversion from RGB to LAB is taken place and L-channel is extracted from the converted image.*
Step2. *The L-channel of LAB-converted image obtained at step 1 is passed to Type-2_Fuzzy_Enhnc() for improving the poor contrast.*
Step3. *The enhanced L-Channel obtained at step 2 is passed to the method Unsharp_Mask(). This method will output the sharpen L-Channel by removing the blurriness that may arise in the enhanced L-channel produced by Type-2_Fuzzy_Enhnc() maintain an equilibrium.*
Step4. *The old L-channel obtained at step 1 is replaced by the sharpen L-channel obtained at step 3 to obtain the enhanced LAB image. This enhanced LAB image is converted back to RGB color space to obtain the final enhanced version of the input color Medical Image.*

4 Materials and Methods

In this section, the methods involved in the proposed approach are discussed. They are explained sequentially as per their requirement in the subsections.

4.1 CIELAB Color Space

This is a device-independent color space originally defined by CAE and specified by the International Commission on Illumination [11]. Unlike RGB color space

where every channel is associated with intensity levels, LAB color space has a particular channel L specifically for luminance level. The other two channels are a* and b* which are chromaticity layers. The a* values show where the color falls along the red green axis, and b* values show where the color falls along the blue-yellow axis. In the following diagram (Fig. 1), it is clearly shown that a* negative values indicate green while positive values indicate magenta; and b* negative values indicate blue and positive values indicate yellow.

Originally, LAB color space is based on CIE XYZ color space [12]. [X, Y, Z] can be obtained by the following equations:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

Then it can be converted into LAB color space easily with the following Eqs. (2), (3), and (4):

$$L = \begin{cases} 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 & if \ \frac{Y}{Y_n} > 0.008856 \\ 903.3 & otherwise \end{cases} \tag{2}$$

$$a = 500\left[\frac{X^{\frac{1}{3}}}{X_n} - \frac{Y^{\frac{1}{3}}}{Y_n}\right] \tag{3}$$

$$b = 200\left[\frac{Y^{\frac{1}{3}}}{Y_n} - \frac{Z^{\frac{1}{3}}}{Z_n}\right] \tag{4}$$

We have chosen LAB color space over RGB for the required enhancement task as LAB has the devoted channel for luminance, L-channel. So, first, the input RGB color image is converted into LAB and then the L-channel has been extracted from it (see Fig. 2).

Fig. 1 LAB color space
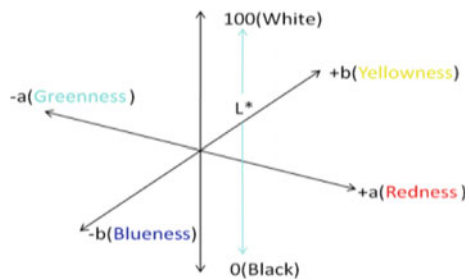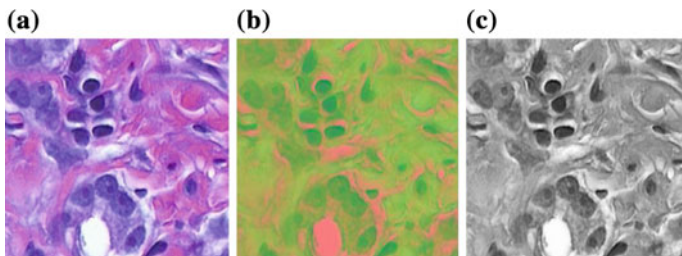
Fig. 2 **a** Original HE-stained tissue image; **b** LAB-converted image; and **c** Extracted L-channel

4.2 Type-2 Fuzzy Set Based Enhancement Technique, Type-2_Fuzzy_Enhnc()

Zadeh [13] introduced fuzzy set where he brought the concept of the membership function. Here, reasoning is based on "approximation" rather than "exact". But, the main problem with a fuzzy set is that there is no definite membership function, so different results may arise according to the different membership functions considered by different researchers for the same problem [9]. In [14], Zadeh introduced the concept of the type-2 fuzzy set by considering fuzziness in the membership function itself. The membership function is now bound to interval based properly bounded by lower and upper bounds. The membership function of the type-1 fuzzy set is blurred to obtain the type-2 fuzzy set (see Fig. 3) [15].

A type-2 fuzzy set A_{TYPEII} can be defined as [9, 14, 16]:

$$A_{TYPEII} = \left\{ x, \widehat{\mu}_A(x, u) | \forall x \in X, \forall u \in J_x \subseteq [0, 1] \right\},$$ (5)

where $\widehat{\mu}_A(x, \mu)$ is the type-2 membership function, J_x is the primary membership function of x, and the upper and lower limits are defined as

$$\begin{aligned} \mu^{upper} &= [J_x]^{\alpha} \\ \mu^{lower} &= [J_x]^{1/\alpha} \end{aligned}$$ (6)

where $0 < \alpha \leq 1$.
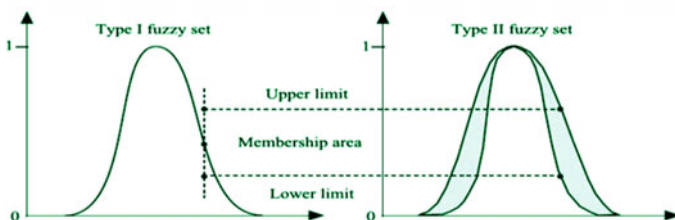
Fig. 3 Type-1 and type-2 fuzzy set membership area

The uncertainty in the primary membership of a type-2 fuzzy set is represented by the footprint of uncertainty (FOU), which can be shown by the following equation:

$$FOU(A_{TYPEII}) = \bigcup_{x \in X} J_x \tag{7}$$

Our proposed technique for type-2 fuzzy set based enhancement, Type-2_Fuzzy_Enhnc() is mainly motivated from the works presented in [9, 16].

First, we have defined the primary membership function J_x as:

$$J_x = \frac{I - I_{\min}}{I_{\max} - I_{\min}}, \tag{8}$$

where I is the gray level of the image (say A) of the range 0 to L − 1, $I_{\min}$ is the minimum, and $I_{\max}$ is the maximum gray value of the image.

We have introduced the following equation for calculating the new membership function:

$$\widehat{\mu}_A = \frac{\mu_{low} \cdot \lambda + \mu_{high} \cdot (1 - \lambda)}{1 - (1 - \lambda) \cdot \mu_{low} \cdot \mu_{high}}, \tag{9}$$

where λ is calculated using the equation below:

$$\lambda = \frac{I_{avg}}{L}, \tag{10}$$

where L is the number of gray levels in the image.

The value of α is calculated on the trial-and-error basis. We have found optimum enhancement result for $\alpha = 0.7$.

In our case, we have enhanced the original L-channel of the LAB-converted image using the equations defined above (Fig. 4).

4.3 *Unsharp Masking,* Unsharp_Mask()

This is a simple but very efficient image sharpening technique [10, 17, 18]. Its working methodology is that it first finds out the blurred or negative version to

Fig. 4 **a** Original L-channel; **b** Enhanced L-channel by the proposed type-2 fuzzy set enhancement technique
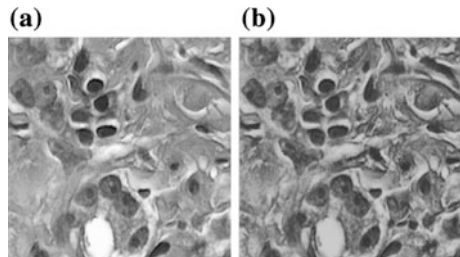
Fig. 5 **a** Enhanced
L-channel by the proposed
type-2 fuzzy set enhancement
technique; **b** Final enhanced
version of (**a**) by unsharp
masking
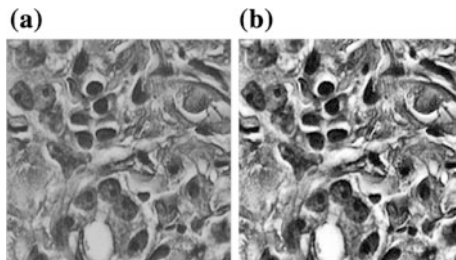
create the unsharp mask of the original image. After that, it subtracts the blurred version (means combines the unsharp mask) from the original image resulting in an edge enhanced sharpen version of the original version. We have employed unsharp masking on the type-2 fuzzy set enhanced L-channel to remove unwanted blurriness that if arise after initial enhancement to finally produce enhanced and edge sharpened version of the same. The steps and equations involved in our unsharp masking technique, Unsharp_Mask(), are illustrated below:

Say, $L(x, y)$ is the current L-channel and $L'(x, y)$ is its smooth version obtained with anisotropic Gaussian smoothing [19] with kernel values along row and column directions, respectively $\sigma_x = 1$, $\sigma_y = 8$.

Then, the edge image $L_e(x, y)$ is given by the following equation:

$$L_e(x, y) = L(x, y) - L'(x, y) \tag{11}$$

Now, the final sharpen image $L_{sharp}(x, y)$ can be obtained by

$$L_{sharp}(x, y) = L(x, y) + c * L_e(x, y), \tag{12}$$

where c > 0.1 is a scaling constant. We have taken c = 0.5 through trial-and-error strategy (Fig. 5).

5 Experiment and Result Discussion

The proposed technique has been implemented in Matlab in a system with an i5 processor and 64 bit Windows 10 operating system. For comparison of results, we have chosen the traditional state-of-the-art technique CLAHE and a recent technique proposed by Ensafi and Tizhoosh [16]. We have tested our technique on several HE stain images out of which results for two images are presented here. HE stain (or HE stain) stands for hematoxylin and eosin stain. These are the widely used stain in medical diagnosis. This type of staining method helps pathologists to distinguish different tissue types. Hematoxylin is a dark blue or violet stain which is basic/positive and eosin is a red or pink stain that is acidic/negative. A combination of hematoxylin and eosin generally produces blues, violets, and reds.
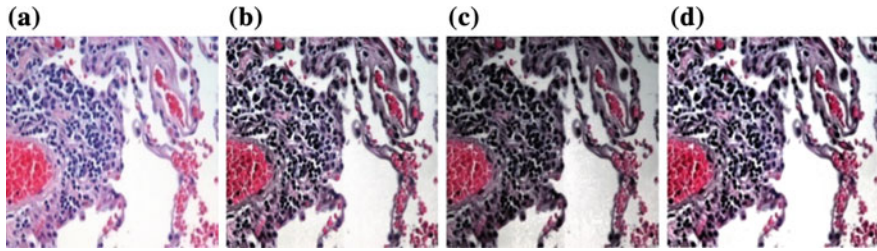
Fig. 6 **a** Original image; **b** Image enhanced by CLAHE; **c** Image enhanced by Ensafi and Tizhoosh [16]; and **d** Enhanced image by proposed technique
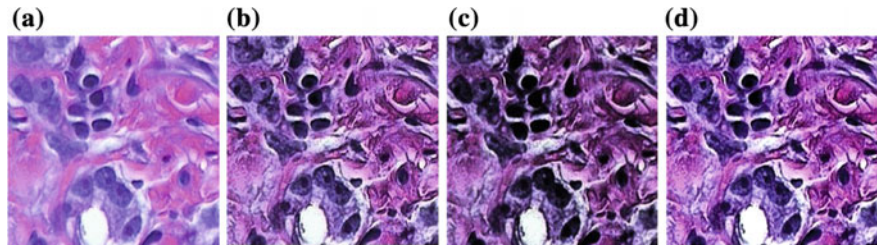
Fig. 7 **a** Original image; **b** Image enhanced by CLAHE; **c** Image enhanced by Ensafi and Tizhoosh [16]; and **d** Enhanced image by proposed technique

Image 1: This image (Fig. 6) is the HE-stained lung tissue sample taken from an end-stage emphysema patient [20, 21]. Here, cell nuclei are blue-purple, red blood cells are red, other cell bodies and extracellular material are pink, and air spaces are white.

Image 2: This (Fig. 7) is an HE-stained image of tissue (Image courtesy of Alan Partin, Johns Hopkins University).

5.1 Subjective Evaluations

Subjective evaluation is carried out on the experimental results by five experts from the field. They have rated the proposed approach based on how much contrast is enhanced, color information restored with fewer color artifacts, brightness preserving, and less appearance of darker regions due to over-amplification of noises present in the homogeneous regions which ultimately determine the overall visual improvement obtained through a color image enhancement technique.

So, from the subjective evaluation (refer Table 1), it is found that our proposed approach succeeds to produce an efficient enhancement for medical color images on the basis of a good rating from all the experts on an average basis. It is clear from the results that the enhanced color image produced by the proposed approach is free

Table 1 Subjective evaluation

Expert	Is visual improvement done after enhancement: Yes/No	Visual improvement rating (value ranges from 1 to 10)	Rating that how much possibility the proposed approach can be applied to medical color image enhancement (value ranges from 1 to 10)
1	Yes	9	9
2	Yes	9	7
3	Yes	8	8
4	Yes	7	9
5	Yes	8	8
Average	*Yes*	*8.2*	*8.2*

from over enhancement, and vagueness is removed to a satisfactory level with clear sharpen boundaries. Contrast is improved better than the same output by CLAHE and type-2 fuzzy set based enhancement technique by Ensafi and Tizhoosh [16].

5.2 Objective Evaluation

To conduct the objective evaluation, we have adopted two metrics: Entropy and Peak Signal-to-Noise Ratio (PSNR). The entropy of an image can be calculated using the following equation:

$$E = -sum(p.*\log_2(p)) \tag{13}$$

The PSNR of an image can be calculated using the following equation:

$$PSNR = 10\log_{10}(MAXi^2/MSE), \tag{14}$$

where MAX$_i$ is the maximum possible pixel value of the image and MSE is the mean squared error which can be calculated using Eq. (15):

$$MSE = \sum_{y=1}^{M}\sum_{x=1}^{N}[I(x,y) - I'(x,y)]^2, \tag{15}$$

where M, and N are the dimensions of the image, I(x, y) is the original image, and I′ (x, y) is the enhanced image.

We have selected entropy and PSNR for the objective evaluation as these two metrics include both the enhancement impact and also the noise sensitivity issue. For a better enhancement, the entropy value should be high [22]. And, for a better suppression of noises, the PSNR value should be high. So, the higher the values of entropy and PSNR, the better the enhancement achieved by an enhancement technique.

Table 2 Objective evaluation

Image no.	Objective evaluation	CLAHE	Ensafi	Proposed approach
Image 1	Entropy	5.9543	6.0312	***7.7984***
	PSNR	7.9014	8.0131	**8.0567**
Image 2	Entropy	4.0106	4.9310	4.8871
	PSNR	22.5043	18.5601	**22.5045**
Average	Entropy	4.98245	5.4811	***6.34275***
	PSNR	15.20285	13.2866	***15.2806***

From Table 2, it is clear that on average, the proposed approach is showing higher values of entropy and PSNR than the other techniques in comparison; this means the proposed approach succeeds to produce better enhancement.

6 Conclusions

The type-2 fuzzy set based enhancement technique for the medical color image is proposed in this paper. For the color computations, LAB color space is chosen because of its unique device-independent nature, possessing a devoted L-channel for luminance value representation. We have introduced a new type-2 fuzzy membership function by which the L-channel of the LAB-converted image is enhanced first. After that, the unwanted blurriness is removed using unsharp masking producing a better enhanced and sharpened version of L-channel. The old L-channel is replaced by this enhanced version to obtain the enhanced LAB image which after converted into RGB color space, the enhanced version of the original color medical image is obtained. The experimental results prove the superiority of the proposed approach through both subjective and objective evaluations. So, the proposed technique can be considered as an ideal one for medical color images to increase the contrast by removing the fuzziness which generally found in these types of images.

References

1. Gu, J., Hua, L., Wu, X., Yang, H., Zhou, Z.: Color medical image enhancement based on adaptive equalization of intensity numbers matrix histogram. Int. J. Autom. Comput. **12**, 551–558 (2015). https://doi.org/10.1007/s11633-014-0871-9
2. Firoz, R., Ali, M., Khan, M., Hossain, M., Islam, M., Shahinuzzaman, M.: Medical image enhancement using morphological transformation. J. Data Anal. Inf. Process. **04**, 1–12 (2016). https://doi.org/10.4236/jdaip.2016.41001
3. Chaira, T.: Medical image enhancement using intuitionistic fuzzy set. In: 2012 1st International Conference on Recent Advances in Information Technology (RAIT) (2012). https://doi.org/10.1109/rait.2012.6194479

4. Bora, D.: AERSCIEA: An efficient and robust satellite color image enhancement approach. In: Proceedings of the Second International Conference on Research in Intelligent and Computing in Engineering (2017)
5. Robinson, G.S.: Color edge detection. In: Proceedings of SPIE Symposium on Advances Image Transmission Techniques, vol. 87, pp. 126–133 (1976)
6. Bora, D.: A novel approach for color image edge detection using multidirectional Sobel filter on HSV color space. Int. J. Comput. Sci. Eng. **5**(2), 154–159 (2017). E-ISSN: 2347-2693
7. Lin, P., Lin, B.: Fuzzy automatic contrast enhancement based on fuzzy C-means clustering in CIELAB color space. In: 2016 12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA) (2016). https://doi.org/10.1109/mesa.2016.7587156
8. Hsu, W., Chou, C.: Medical image enhancement using modified color histogram equalization. J. Med. Biol. Eng. **35**, 580–584 (2015). https://doi.org/10.1007/s40846-015-0078-8
9. Chaira, T.: An improved medical image enhancement scheme using Type II fuzzy set. Appl. Soft Comput. **25**, 293–308 (2014). https://doi.org/10.1016/j.asoc.2014.09.004
10. Zhao, Z., Zhou, Y.: PLIP based unsharp masking for medical image enhancement. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2016). https://doi.org/10.1109/icassp.2016.7471874
11. Hunter, R.: Photoelectric color difference meter*. J. Opt. Soc. Am. **48**, 985 (1958). https://doi.org/10.1364/josa.48.000985
12. Kaur, A., Kranthi, B.: Comparison between YCbCr color space and CIELab color space for skin color segmentation. Int. J. Appl. Inf. Syst. **3**(4), 30–33 (2012)
13. Zadeh, L.A.: Fuzzy sets. Inf. and Cont. **1965**(8), 338–353 (1965)
14. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning-1. Inf. Sci. **8**, 199–249 (1975). https://doi.org/10.1016/0020-0255(75)90036-5
15. Mendel, J., John, R.: Type-2 fuzzy sets made simple. IEEE Trans. Fuzzy Syst. **10**, 117–127 (2002). https://doi.org/10.1109/91.995115
16. Ensafi, P., Tizhoosh, H.: Type-2 Fuzzy Image Enhancement. Lecture Notes in Computer Science, pp. 159–166 (2005). https://doi.org/10.1007/11559573_20
17. 3.9. Unsharp Mask. In: Docs.gimp.org (2017). https://docs.gimp.org/en/plug-in-unsharp-mask.html. Accessed 1 May 2017
18. Spatial Filters—Unsharp Filter. In: Homepages.inf.ed.ac.uk (2017). http://homepages.inf.ed.ac.uk/rbf/HIPR2/unsharp.htm. Accessed 1 May 2017
19. Spatial Filters—Gaussian Smoothing. In: Homepages.inf.ed.ac.uk (2017). http://homepages.inf.ed.ac.uk/rbf/HIPR2/gsmooth.htm. Accessed 4 May 2017
20. Emphysema H and E.jpg. In: En.wikipedia.org (2017). http://en.wikipedia.org/wiki/Image:Emphysema_H_and_E.jpg. Accessed 15 June 2017
21. Rogers, A.: T cells cause lung damage in emphysema. PLoS Med. **1**, e25 (2004). https://doi.org/10.1371/journal.pmed.0010025
22. Bora, D.: Importance of image enhancement techniques in color image segmentation: a comprehensive and comparative study. Indian J. Sci. Res. **15**(1), 115–131 (2017). https://arxiv.org/abs/1708.05081

Score Formulation and Parametric Synthesis of Musical Track as a Platform for Big Data in Hit Prediction

Sunil Karamchandani, Prathmesh Matodkar, Suraj Iyer and Nirav Gori

Abstract In today's entertainment industry which is becoming increasingly competitive, music producers, record labels are striving hard to give the next big hit song and capture the viable music market. We propose to formulate factors and dependency variables which would form the basis of hit prediction in big data environment. The audio features such as pitch and tempo are analyzed in tandem with statistical parameters such as root mean square energy, slope, period frequency, and musical topographies like acousticness, loudness, and instrumentalness. This is a preliminary experiment where the simulated ratings are paralleled with ground truth obtained from Billboard, Spotify, and Radio Mirchi rankings over a period of 5–10 weeks. The paper covers a wide area of tracks from USA, UK, Australia, and India, and proposes to arrive at a consensus to the factors contributing to the success of the track according to their topography. While acousticness plays a vital role in US and India countdowns, British are highly influenced by the danceability and the energy components of the track. The paper provides a cushion for hit prediction classification of musical tracks in big data applications.

S. Karamchandani (✉)
DJSCE, University of Mumbai, Mumbai, India
e-mail: skaramchandani@rediffmail.com

P. Matodkar · S. Iyer · N. Gori
Electronics & Telecommunication Engg, DJSCE,
University of Mumbai, Mumbai, India
e-mail: prathmat@gmail.com

S. Iyer
e-mail: suraj31295@gmail.com

N. Gori
e-mail: gorinirav5@gmail.com

363

1 Introduction

Consign to oblivion Anu Malik, Farah Khan for an ear to audit potential talent in song and music, we suggest to foretell the hit potential of a song in proposing a Hit Conceivable Equivalency (HCE) by forking in statistical parameters with musical topographies. Authors Silk et al. [1] have presented a music browsing service to discover emerging styles in UK using four factors: artist, genre, location, and fans. Statistical data from social media like Twitter mentions and number of hits on SoundCloud have been used to determine the same. However, they have restricted their region of research only to UK and have used non-musical attributes for the analysis. "Popular music estimation based on topic model" [2] proposes a technique for music estimation using audio parameters and time information. Relationship between music trends and corresponding time period has been established using machine learning algorithm. Micro-blogging data via tweets was represented as a time series and correlated with Billboard ratings by the authors E. Zangerle in the paper, "Can microblogs predict music charts?" [3] for predicting the music chart ratings. However, the tweets and chart ratings predicted a mild correlation between them. Hence, results do not consider the demographic variation in Twitter and Billboard consumer data. The authors Yang et al. [4] harp on the success of deep learning by conjoining primitive Mel-spectrogram with Convolution Neural Network (CNN). The aim of their experiment suggests more success with deep learning than convolutional NN. Again, they predict only the popularity of songs in Taiwan. Authors Borg and Hokkanen [5] have employed machine learning algorithm which have been fed with only signal features, such as cepstrum and MFCC coefficient and devoid of any musical features. However, their concentration is focused on pop media genre only.

We have accumulated the factors which would be instrumental in predicting the hit potential of a song. Initially, we have assigned weights to these parameters manually with reference to the Billboard rankings and Spearman rank coefficient correlation method. However, our intention is to assign weights to the parameters dynamically by big data and neural network applications in future.

2 Attribute Extraction

The attributes like acousticness, instrumentalness, loudness, and time signature are obtained from echo-nest API [6]. The audio information with the aid of the appropriate parameters can be extracted with an API call. Echo-nest software comes to our rescue by generating the audio analysis when the song is run through it.

2.1 Statistical Attributes

The following numerical values of the attributes have been formulated to exploit the problem of hit prediction.

2.1.1 Slope

For the given audio signal, we have best curve, and then the derivate of this curve is computed, which gives the value of slope.

2.1.2 Period Frequency

The audio signal or the song is divided into number of frames. We calculate the frequency of the maximum periodicity of these frames and this is estimated using auto-correlation function. In the auto-correlation if there are no peaks, its periodicity is zero.

2.1.3 Entropy

We also calculate the Shannon entropy of the audio signal, i.e., the song to be analyzed. Shannon entropy gives us the amount of uncertainty. The higher the entropy, the more is the noise in the song. We use Eq. (1) to calculate the entropy, where b is the base of the logarithmic function used.

$$H(X) = - \sum_{i=1}^{n} p(x_i) * \log_b p(x_i) \tag{1}$$

2.1.4 RMS Energy

An audio spectrum is comprised of a number of peaks. The RMS energy gives us the effective energy of the song, and statistically it is area under the graph. The RMS can be calculated by the following formula. The global energy of the audio can be calculated by taking the root average of the square of the amplitude, also called root mean square (RMS). It is given by Eq. (2).

$$x_{\text{rms}} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^3 + \cdots + x_n^2}{n}}, \tag{2}$$

where x is the input audio signal. It basically provides us with the energy throughout the song based on its sampled values and also gives the effective power which is independent of the phase of the sample.

Tempo, key, energy, and signature features are extracted from echo-nest API. Let us look for a hit track levels by Avicii. This gives us a song-ID of the track. A profile call is made where we pass on song-ID along with a bucket which characterizes the audio summary parameter [7]. The attribute data of the song is returned in separate variables on extraction from the echo-nest software. Also, the description of each feature is available in echo-nest's documentation as well as below the analysis.

2.2 Musical Topographies

Different musical topographies influence the quality of the sound and are used to enhance the soundtrack. The topographies which dominate the sound quality are considered for the hit prediction analysis and are described in this section.

2.2.1 Danceability

Dance is a vibrant form of expression. It also acts as an indicator of whether the song is likeable or not. Hence, we have considered a factor called the danceability which is given in Eq. (3).

$$Danceability = Sum\,(Tempo * v\,(t)), \tag{3}$$

where v(t) represents the velocity of the beat at sample time t, i.e., beat strength.

2.2.2 Energy

Energy is a normalized measure representing a perceptual measure of intensity to a value below unity. As we know tracks with high energy have a fast, loud, and noisy feel, death metal on the higher end of the scale with a Bach prelude on the other extreme end is example of the energy variations. Energy value is a consequence of dynamic range, perceived loudness, timbre, onset rate, and entropy calculated as in (4).

$$dx = 25.6/nx$$

$$xn = dx/2 * n$$

$$Energy = Sum\,(xn)/n \tag{4}$$

2.2.3 Key

Key signifies the scale or the different pitches that occur along a soundtrack. Each individual scale has been assigned a key, say a key of 1 is allocated to track sung in D-flat.

2.2.4 Liveliness

We have considered this feature which informs us about the involvement of the audience in that particular recording. Tracks when performed live suggest an increase in the liveliness coefficient. A score of 0.8 indicates a lively track. Values below 0.6 most likely suggest studio recordings.

2.2.5 Loudness

Loudness measure is obtained as the amplitudes of the audio waves averaged across the entire time period. It represents a primary psychological correlate of physical strength defined in (5) and (6).

$$\text{Loudness} = (\text{RM} \cdot \text{avg} + \text{Peak} \cdot \text{avg})/2 \tag{5}$$

The typical values of the loudness range from -80 to 0 dB.

2.2.6 Speechiness

The dominance of spoken words in a track is a measure of the speechiness of a track. It is more pronounced in songs which are more like, poetry. The normalized speechiness values above 0.66 are indicative of tracks that may be entirely of spoken words. The median values of speechiness describe tracks which contain equal contribution of both music and speech. Rap music is one such example where speechiness is dominant in layers. Instrumentals can be mapped to lower range of speechiness such as 0.33 and below.

2.2.7 Tempo

Tempo is analogous to the number of beats per minute in a sound track. For a full duration of the song if $V_{\text{inst}}(200 \text{ Hz}) > -6$ dB, then the

$$\text{Tempo} = b/6 \tag{6}$$

Range of tempo analysis ranges from 70 to 180 bpm.

2.2.8 Valence

Valence illustrates the musical positivity in a track as shown in (7). As suggested, high valence tracks are represented in qualities of happy and cheerful tones while tracks with sad, depressive, or even angry tones are an indication of low valence.

$$\text{Valence} = ((\text{loudness}(\text{rms}) * \text{Average segment duration}) + - \sigma) * \mu, \tag{7}$$

where the sign of the standard deviation depends on the mode.

3 Observation and Score Formulation

For understanding the listening and likeability patterns of the listeners of UK, USA, and Australia, we have considered the same set of songs for these three nations. To compute the weights, we needed a reliable source of ranking which we found out in the form of Billboard hot 100 charts for the countries USA and UK and Spotify top 20 charts for Australia. Since India is a culturally diverse country, it is only fair to include local Indian songs which would help us correctly gauge the likeability factor and help us recognize the taste of the Indian audience, for which we have used Radio Mirchi's top 200 chart listings as a reference.

3.1 USA

Score (India) =
(RMS Energy * 100) + Slope + Period Freq + Entropy + Tempo
+ BeatIntensity + Mean + (Energy * 100) + Liveness + (Acousticness * 1000)
+ Time Signature + Key + Instrumentalness + Speechiness + (Danceability *
100) + Valence + Mode + Loudness

From the above formulation of the score for USA, it can be seen that for a song to attract the audiences, acousticness parameter has the highest weight, which means that the song should have less electronic instruments or tweaking and should contain more acoustic instruments like guitar, piano, etc. which are electronically treated. This is followed by danceability and energy with same weightage. Table 1 provides the hit prediction values with attributes for US song database (Table 2).

Table 1 USA song database and score

	Closer	Heathens	Cold water	Let me love you	Cheap thrills	Brocolli
RMS energy	0.2482	0.23679	0.24521	0.2282	0.23249	0.32611
Slope	0.1283	−0.097	0.0322	−0.2549	0.0502	0.0752
Period freq	0.2781	0.1668	0.3034	1.6684	0.8342	0.1517
Entropy	0.9962	0.9964	0.9707	0.9721	0.9692	0.9716
Tempo	95.01	90.024	92.94	199.864	89.972	145.99
Beat intensity	59	68	59	66	64	72
Mean	154.41	150.89	134.29	125	113.76	141.34
Energy	0.524	0.396	0.798	0.718	0.8	0.525
Liveness	0.111	0.105	0.156	0.122	0.0775	0.057
Acousticness	0.414	0.0841	0.0736	0.0784	0.0561	0.236
Time signature	4	4	4	4	4	4
Key	8	4	6	8	6	8
Instrumentalness	0	0.0000358	0	0.000010	0.0000020	0
Speechiness	0.0388	0.0286	0.0432	0.0576	0.215	0.131
Danceability	0.748	0.732	0.608	0.476	0.592	0.886
Valence	0.635	0.547	0.488	0.143	0.736	0.711
Mode	1	0	0	1	0	1
Loudness	−5.599	−9.438	−5.092	−5.309	−4.931	−7.39
Score	**884.03**	**529.80**	**531.85**	**622.37**	**494.23**	**776.75**

Table 2 USA average rank and analysis rank

Relative ranking:						
Rank (weeks)	Closer	Heathens	Cold water	Let me love you	Cheap thrills	Brocolli
1	1	2	3	4	5	6
2	1	2	3	4	5	6
3	1	2	3	5	4	7
4	1	2	3	6	4	7
5	1	3	2	6	4	7
6	1	4	2	6	3	7
7	1	4	3	6	2	7
8	5	4	2	6	1	7
9	4	5	3	0	1	6
10	0	4	2	0	1	5
Average rank	**1.77**	**3.2**	**2.6**	**4.3**	**3**	**6.5**
Analysis rank	**1**	**6**	**5**	**4**	**7**	**2**

3.2   India

Score (India) =

(RMS Energy * 100) + Slope + Period Freq + Entropy + Tempo

   + BeatIntensity + Mean + (Energy * 100) + Liveness + (Acousticness * 1000)

      + Time Signature + Key + Instrumentalness + Speechiness + (Danceability *

100) + Valence + Mode + Loudness

Indian music industry has come a long way from 1950s to 2017 and it has seen many changes in terms of song production, recording, and distribution. Although there is a rise of usage of electronic instruments, samples, and loops in today's song from the score, we can make out that even today acoustcness parameter dominates the most which is followed by energy and danceability. From this, we can conclude that the Americans and Indians tastes of music are analogous (Tables 3, 4, and 5).

Table 3   Indian song database and score

	Ae Dil Hai Mushkil	Toota Jo Kabhi Tara	Channa Mereya	Tere Sang Yaara	Bulleya	Kala Chashma	Jag Ghoomeya
RMS energy	0.2823	0.292	0.287	0.235	0.29088	0.2846	0.2097
Slope	0.0121	−0.1478	0.026	0.028	−0.296	0.026	−0.0126
Period freq	0.1192	0.175	0.256	0.667	0.1854	0.232	0.3337
Entropy	0.969	0.9755	0.9687	0.972	0.9655	0.970	0.9712
Tempo	123.89	139.9	90.06	76.02	86.95	106.04	82.927
Beat intensity	59	68	64	66	64	72	72
Mean	108.817	113.5266	156.1014	129	134.97	105.9	117.12
Energy	0.654	0.672	0.788	0.6	0.769	0.837	0.577
Liveness	0.165	0.219	0.106	0.075	0.112	0.272	0.103
Acousticness	0.713	0.412	0.237	0.649	0.497	0.107	0.49
Time signature	4	4	4	4	4	4	4
Key	3	9	9	8	2	0	11
Instrumentalness	0	0.000016	0.000024	0	0	0.000026	0
Speechiness	0.0679	0.0369	0.0446	0.033	0.046	0.0774	0.0344
Danceability	0.495	0.437	0.476	0.615	0.552	0.806	0.536
Valence	0.344	0.219	0.745	0.559	0.705	0.911	0.618
Mode	0	0	0	1	0	1	1
Loudness	−6.639	−6.339	−4.281	−6.308	−4.504	−4.273	−7.83
Score	**1149.88**	**881.83**	**713.18**	**1074**	**947.335**	**587.02**	**904.54**

Table 4 India average rank and analysis rank

Rank (Weeks)	Ae Dil Hai Mushkil	Toota Jo Kabhi Tara	Channa Mereya	Tere Sang Yaara	Bulleya	Kala Chashma	Jag Ghoomeya
1	1	2		3	5	4	7
2	1	2		3	5	4	8
3	1	2		3	5	4	10
4	1	2	4	3	6	5	14
5	1	2	3	4	5	8	16
Average rank	**1**	**2**	**3.5**	**3.2**	**5.2**	**5**	**11**
Analysis rank	**1**	**5**	**6**	**2**	**3**	**7**	**4**

Table 5 India relative ranking

Relative ranking							
Rank (Weeks)	Ae Dil Hai Mushkil	Toota Jo Kabhi Tara	Channa Mereya	Tere Sang Yaara	Bulleya	Kala Chashma	Jag Ghoomeya
1	1	2		3	5	4	6
2	1	2		3	5	4	6
3	1	2		3	5	4	6
4	1	2	4	3	6	5	7
5	1	2	3	4	5	6	7
Average rank	**1**	**2**	**3.5**	**3.2**	**5.2**	**4.6**	**6.4**

3.3 United Kingdom

Score (UK) =
(RMS Energy * 100) + Slope + Period Freq + Entropy + Tempo + Beat Intensity + Mean + (Energy * 100) + Liveness + Acousticness + Time Signature + Key + Instrumentalness + Speechiness + (Danceability * 100) + Valence + Mode + Loudness

As we analyzed the scores for USA and INDIA, we noticed that the people of these two nations are more acoustically inclined, but what we observe from the score for UK is that here acousticness does not play a major role, and the energy and danceability of the song have equal importance and more prominence over other parameters (Table 6, 7, and 8).

Table 6 UK song database and score

	Say you won't let go	Starboy	Closer	Side to side	Let me love you	Cold water	The greatest
RMS energy	0.188	0.3244	0.2482	0.269	0.2282	0.2452	0.26347
Slope	0.0739	0.2108	0.1283	0.029	−0.2549	0.0322	0.06619
Period freq	0.256	0.2086	0.2781	0.208	1.6684	0.3034	0.8342
Entropy	0.970	0.9704	0.9962	0.967	0.9721	0.9707	0.9689
Tempo	99.26		95.01	159.1	199.86	92.943	192.024
Beat intensity	68	70	59	69	66	59	66
Mean	130.9	155.67	154.4	124.9	125	134.29	139.37
Energy	0.564	0.594	0.524	0.738	0.718	0.798	0.723
Liveness	0.086	0.134	0.111	0.292	0.122	0.156	0.0507
Acousticness	0.693	0.159	0.414	0.040	0.0784	0.0736	0.0109
Time signature	4	4	4	4	4	4	4
Key	10	7	8	6	8	6	1
Instrumentalness	0	0.00002	0	0	0.00001	0	0.0004
Speechiness	0.052	0.28	0.038	0.247	0.0576	0.0432	0.268
Danceability	0.4	0.608	0.748	0.648	0.476	0.608	0.666
Valence	0.449		0.635	0.606	0.143	0.488	0.731
Mode	1	1	1	0	1	0	1
Loudness	−7.444	−7.021	−5.599	−5.883	−5.309	−5.092	−6.164
Score	**423.5**	**385.5**	**470.4**	**525.1**	**544**	**458.33**	**565.40**

Table 7 UK average rank and analysis rank-I

Rank (weeks)	Say you won't let go	Starboy	Closer	Side to side	Let me love you	Cold water	The greatest
1	1	3	2	6	5	7	8
2	1	2	3	5	6	7	8
3	1	3	2	4	7	8	5
4	2	4	3	7	9	11	6
5	2	4	3	9	12	13	8
6	2	4	5	12	14	16	9
7	3	4	8	14	15	18	10
8	2	5	9	14	16	19	11
9	4	8	9	13	15	22	12
10	4	3	10	20	19	24	16
Average rank	**2.2**	**4**	**5.4**	**10.4**	**11.8**	**14.5**	**9.3**
Analysis rank	**6**	**7**	**4**	**3**	**2**	**5**	**1**

Table 8 UK average rank and analysis rank-II

Relative ranking:

Rank (weeks)	Say you won't let go	Starboy	Closer	Side to side	Let me love you	Cold water	The greatest
1	1	3	2	5	4	6	7
2	1	2	3	4	5	6	7
3	1	3	2	4	6	7	5
4	1	3	2	5	6	7	4
5	1	3	2	5	6	7	4
6	1	2	3	5	6	7	4
7	1	2	3	5	6	7	4
8	1	2	3	5	6	7	4
9	1	2	3	5	6	7	4
10	2	1	3	6	5	7	4
Average rank	**1.1**	**2.3**	**2.6**	**4.9**	**5.6**	**6.8**	**4.7**
Analysis rank	**6**	**7**	**4**	**3**	**2**	**5**	**1**

4 Verdict and Conclusion

We have considered a number of parameters that goes into making of song and based on the rakings of the respective songs we analyzed and formulated a score that defines the position of the song on the chart list, which indirectly gives us an idea about the taste of the audience of these three nations. Hence, we can conclude that the audience from India and USA like songs which are more acoustic in nature as well which are highly energetic and have danceability quotient in it, whereas on the other hand the UK audience wants only high energetic and highly danceability quotient songs where acousticness does not play a major role. Also, we plan to create and feed these results which include the parameters and the weights obtained to a big data neural network as a future scope.

References

1. Silk, H., Santos-Rodriguez, R., Mesnage, C., DeBie, T.: Data science for the detection of emerging music styles. In: Extended abstracts for the Late-Breaking Demo Session of the 17th International Society for Music Information Retrieval Conference (2016)
2. Kinoshita, S., Ogawa, T., Haseyama, M.: Popular music estimation based on topic model using time information and audio features. In: IEEE 3rd Global Conference on Consumer Electronics (GCCE), Tokyo, pp. 102–103 (2014)
3. Hupfauf, B., Pichl, M., Specht, G., Zangerle, E.: Can microblogs predict music charts? An analysis of the relationship between #nowplaying tweets and music charts. In: ISMIR (2016)

4. Yang, L.-C., Chou, S.-Y., Liu, J.-Y., Yang, Y.-H., Chen, Y.-A.: Revisiting the problem of audio-based hit song prediction using convolutional neural networks. In: 2017 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 621–625 (2017). ISSN 2379-190X
5. Borg, N., Hokkanen, G.: What makes for a hit pop song? What makes for a pop song? http://cs229.stanford.edu/proj2011/BorgHokkanen-WhatMakesForAHitPopSong.pdf (2011)
6. Echonest's developer.echonest.com and phyechonest python library
7. Herremans, D., Martens, D., Sorensen, K.: Dance hit song prediction. J. New Music Res. **43**(3), 291–302 (2014)

Dynamics of Dust-Ion-Acoustic Anti-kink Waves in a Dissipative Nonextensive e-p-i Dusty Plasma

Jharna Tamang

Abstract Dynamics of the dust-ion-acoustic anti-kink waves (DIAAKWs) in an unmagnetized multicomponent electron–positron–ion (e-p-i) dusty plasma consisting of negatively charged static dust, positively charged positrons following Maxwellian distribution, inertial ions, and q-nonextensive distributed electrons of two distinct temperatures are studied. The ion kinematic viscosity is intended and the reductive perturbation technique (RPT) is employed to obtain the Burgers equation. Using Galilean transformation, the Burgers equation is reduced to a system. Analyzing vector fields and corresponding potential energy of the dynamical system, the stability and instability of the equilibrium points are discussed. The physical parameters $q_1, q_2, \mu_1, \mu_2, \sigma_1, \sigma_2, \eta$, and U affect significantly on the characteristics of the DIAAKWs.

Keywords Dust-ion-acoustic wave $\cdot$ Anti-kink wave $\cdot$ e-p-i dusty plasma Dynamical systems

1 Introduction

For the last few years, the investigation on basic characteristics of nonlinear waves in dusty plasmas [1] gained rapid growing interest due to its important role both in space plasmas and in laboratory plasma. Dusty plasmas composed of ions, free electrons, and micro-sized dust particles which are negatively charged. Due to dissolution and succession, the dust grains have flexible charge and mass. This dynamic nature plays a crucial role to draw attention of the researchers. Many research works were carried out in specific field namely dust-ion acoustic (DIA) waves both theoretically and experimentally. It had been more than a decade, Shukla and Silin [2] theoretically proved that a dusty plasma underpins DIA wave with low frequency. The linear

J. Tamang (✉)
Department of Mathematics, Sikkim Manipal Institute of Technology,
Sikkim Manipal University, Majitar, Rangpo, East-Sikkim 737136, India
e-mail: jharnatamang.12@gmail.com; Jharna_201710008@smit.smu.edu.in

features of the DIA waves were reported both in magnetized [3] and unmagnetized plasmas [4]. The existence of DIA waves was examined experimentally [5, 6] in the literature. Many researchers investigated the features of DIA solitary waves [7, 8] in dusty plasmas.

It is relevant to note that a plasma with effective dissipation causes the formation of shock structures. Recently, the existence of shock waves in a dusty plasma was observed [9] when kinematic viscosity due to ions is taken under consideration. On further investigation, it was shown by Misra et al. [10] that in the rendering of DIA shock waves, the kinematic viscosity due to ions in a dusty plasma plays a vital task. Adhikary [11] reported that the amplitude of the DIA shock waves decreases as dissipation was increased. During the expedition of the international Space Station, Nakamura et al. [12] performed experimental observation for the DIA shock waves in dusty plasmas. Masud et al. [13] studied the influence of kappa-distributed electrons with two different temperatures on DIA shock waves in e-p-i dusty plasmas. Recently, Ema et al. [14] performed theoretical and numerical interpretation on DIA shock structures in unmagnetized dusty multi-ion plasmas consisting q-nonextensive electrons.

Recently, a considerable awareness has been given to a nonextensive generalized Boltzmann–Gibbs–Shannon (BGS) entropy that was first initiated by Tsallis. The nonextensivity of plasma particles is utterly common in space environments, stellar polytropes, hadronic matter, and quark–gluon. The nonextensive behavior of electrons and ions (characterized by a parameter q) has been successfully utilized in plasma physics. It is prominent to note that the nonextensive plasma particles have been successfully engaged in the research field of plasma dynamics. Furthermore, if $q < -1$, the corresponding q-distribution is not normalizable. However, the strength q of nonextensivity varies in $-1 < q < 1$. Furthermore, for $q \geq 1$, the corresponding distribution unveils Maxwellian–Boltzmann velocity function.

Recently, applying theory of planar dynamical systems [15], many researchers [16–18] studied nonlinear waves in plasmas. However, there is no attempt on the study of the Burgers equation in dusty plasmas applying the concept of dynamical systems. That is why, in this present work, dynamics of the dust-ion-acoustic anti-kink waves are studied in a dissipative nonextensive e-p-i dusty plasma implementing the theory of one-dimensional dynamical systems [15].

2 Model Equations

An unmagnetized dusty plasmas that consists of static dust with negative charge, inertial ions, positrons (isothermal), and q-nonextensive electrons with two distinctive temperatures are considered. Basic equations are

$$\frac{\partial n_i}{\partial t} + \frac{\partial}{\partial x}(n_i u_i) = 0, \tag{1}$$

$$\frac{\partial u_i}{\partial t} + u_i \frac{\partial u_i}{\partial x} = -\frac{\partial \phi}{\partial x} + \eta \frac{\partial^2 u_i}{\partial x^2}, \tag{2}$$

$$\frac{\partial^2 \phi}{\partial x^2} = \mu - n_i - \mu_p e^{-\sigma_2 \phi} + \mu_1 n_{e1} + \mu_2 n_{e2}, \tag{3}$$

where n_i denotes the number density of ions, n_{e1} and n_{e2} denotes lower and higher temperature electron number densities, respectively, with $n_{i0} + n_{p0} = n_{e10} + n_{e20} + Z_d n_{d0}$ at the equilibrium, where the addendum "0" denotes for unperturbed quantity, n_{d0} and n_{p0} denotes the unperturbed dust and positron number density, u_i denotes the velocity of the ions, and ϕ denotes electrostatic potential. Here, the following normalizations are used: n_p, n_{e1}, n_{e2}, and n_i are normalized by its equilibrium value n_{p0}, n_{e10}, n_{e20}, and n_{i0}, respectively. u_i is normalized by $C_i = (\frac{k_B T_{e1}}{m_i})^{\frac{1}{2}}$, in which k_B is the Boltzmann constant, m_i is the mass of ions, e is the magnitude of the electron charge, η is the coefficient of viscosity normalized by $m_i n_{i0} \omega_{pi} \lambda_D^2$, $\sigma_1 = \frac{T_{e1}}{T_{e2}}$, $\sigma_2 = \frac{T_{e1}}{T_p}$, $\mu_1 = \frac{n_{e10}}{n_{i0}}$, $\mu_2 = \frac{n_{e20}}{n_{i0}}$, $\mu = \frac{Z_d n_{d0}}{n_{i0}} = 1 + \mu_p - \mu_1 - \mu_2$, and ϕ is normalized by $\frac{k_B T_{e1}}{e}$. It should be noted that $T_{e1}(T_{e2})$ is the lower (higher) electron temperature, T_p is the temperature of the positrons, and T_i denotes the temperature of an ions. The time variable t is normalized to $\omega_{pi}^{-1} = (\frac{m_i}{4\pi n_{i0} e^2})^{\frac{1}{2}}$, where ω_{pi} is ion plasma frequency and the space variable is normalized to the Debye length $\lambda_D = (\frac{k_B T_{e1}}{4\pi n_{i0} e^2})^{\frac{1}{2}}$.

The q-nonextensive electron velocity distribution [19] is $f_e(v) = C_q \{1 + (q_e - 1)$ $[\frac{m_e v^2}{2k_B T_e} - \frac{e\phi}{k_B T_e}]\}^{\frac{1}{(q_e - 1)}}$, where $C_q = n_{e0} \frac{\Gamma(\frac{1}{1-q_e})}{\Gamma(\frac{1}{1-q_e} - \frac{1}{2})} \sqrt{\frac{m_e(1-q_e)}{2\pi k_B T_e}}$ for $-1 < q_e < 1$, and $C_q = n_{e0} \frac{1+q_e}{2} \frac{\Gamma(\frac{1}{q_e-1} + \frac{1}{2})}{\Gamma(\frac{1}{q_e-1})} \sqrt{\frac{m_e(q_e-1)}{2\pi k_B T_e}}$ for $q_e > 1$.

The normalized number densities of q-nonextensive electrons with low and high temperatures are given by

$$n_{e1} = \{1 + (q_1 - 1)\phi\}^{\frac{1}{q_1-1} + \frac{1}{2}}, \tag{4}$$

$$n_{e2} = \{1 + (q_2 - 1)\sigma_1 \phi\}^{\frac{1}{q_2-1} + \frac{1}{2}}, \tag{5}$$

where q_1 and q_2 are real numbers with $q_1 > -1$ and $q_2 > -1$.

3 Burgers Equation

To acquire the Burgers equation, the reductive perturbation technique is applied. The stretching of independent variables is given by

$$\xi = \epsilon(x - Vt), \tag{6}$$

$$\tau = \epsilon^2 t, \tag{7}$$

and we consider expansions of dependent variables as

$$n_i = 1 + \epsilon n_1 + \epsilon^2 n_2 + \epsilon^3 n_3 + \cdots, \tag{8}$$

$$u_i = \epsilon u_1 + \epsilon^2 u_2 + \epsilon^3 u_3 + \cdots, \tag{9}$$

$$\phi = \epsilon \phi_1 + \epsilon^2 \phi_2 + \epsilon^3 \phi_3 + \cdots, \tag{10}$$

where V denotes the phase speed of the DIA waves, and ϵ is a small nonzero parameter which is proportional to the amplitude of the perturbation. Now, substituting Eqs. (4)–(10) into Eqs. (1)–(3) and considering the lowest order of ϵ, the obtained dispersion relation is

$$V^2 = \frac{1}{a + \mu_p \sigma_2}, \tag{11}$$

where $a = \frac{1}{2}\{\mu_2 \sigma_1 (q_2 + 1) + \mu_1 (q_1 + 1)\}$.

Considering next higher order of ϵ and eliminating the perturbed quantities, one can obtain the Burgers equation as

$$\frac{\partial \phi_1}{\partial \tau} + A \phi_1 \frac{\partial \phi_1}{\partial \xi} = B \frac{\partial^2 \phi_1}{\partial \xi^2}, \tag{12}$$

where $A = \frac{3}{2V} - (b - \mu_p \frac{\sigma_2^2}{2})V^3, B = \frac{\eta}{2}$, with $b = \frac{1}{8}\{\mu_2 \sigma_1^2 (q_2 + 1)(3 - q_2) + \mu_1 (q_1 + 1)(3 - q_1)\}$.

4 One-Dimensional Dynamical System

To investigate dust-ion-acoustic anti-kink waves for the Burgers equation (12), we consider a new variable χ in the form

$$\chi = (\xi - U\tau), \tag{13}$$

where U is the speed of the dust-ion-acoustic anti-kink wave. Using $\psi(\chi) = \phi_1(\xi, \tau)$ in the Burgers equation (12) and integrating w.r.t χ once along with initial condition $\psi \to 0, \frac{d\psi}{d\chi} \to 0$ as $\chi \to \infty$ or $\chi \to -\infty$, the Burgers equation (12) reduces to

$$\frac{d\psi}{d\chi} = c\psi^2 - d\psi, \tag{14}$$
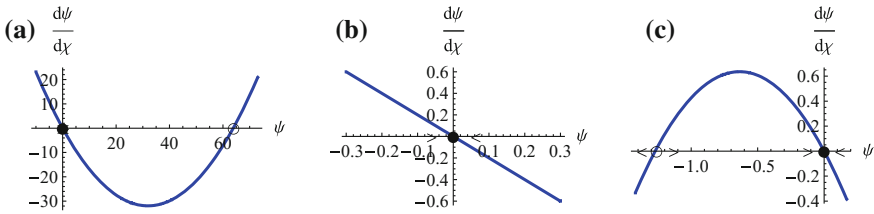
Fig. 1 Vector fields of the Eq. (14) for the different values of q_2 [$q_2 = 1.8$ (**a**); $q_2 = 0.55527$ (**b**); $q_2 = 0.11$ (**c**)] with $q_1 = 0.8$, $\mu_1 = 0.4$, $\mu_2 = 0.04$, $\eta = 0.4$, $\sigma_1 = 0.1$, $\sigma_2 = 0.04$ and $U = 0.4$

where $c = \frac{A}{2B}$ and $d = \frac{U}{B}$. The Eq. (14) represents an one-dimensional dynamical system [15] with equilibrium points at $\psi^* = 0$ and $\psi^* = \frac{d}{c}$.

To investigate the system (14), we illustrate it as a vector field. To abstract the corresponding vector field, it is sufficient to plot $\frac{d\psi}{d\chi}$ versus ψ and then we draw arrows on the ψ-axis to illustrate the corresponding velocity vector at each ψ. The arrows point to the right when $\frac{d\psi}{d\chi} > 0$ and to the left when $\frac{d\psi}{d\chi} < 0$. In Fig. 1a–c, all possible vector fields for system (14) are shown depending on the physical parameters q_1, q_2, μ_1, μ_2, η, σ_1, σ_2, and U. It is necessary to note that the stable equilibrium point is represented by solid black dot and unstable equilibrium point is represented by open black circle. Moreover, in these vector fields (see Fig. 1a–c), the procedure to an equilibrium point is always uniform. Hence, the Burgers equation (12) for dust-ion-acoustic wave does not possess periodic wave solution.

5 Potential Energy

Now we try to speculate the dynamics of the one-dimensional dynamical system (14) established on the physical idea of potential energy. The potential energy corresponding to the dynamical system (14) is given by

$$-\frac{dV}{d\psi} = c\psi^2 - d\psi. \tag{15}$$

Integrating (15) with respect to ψ along with initial condition $\psi \to 0$, $V(\psi) \to 0$ as $\chi \to \infty$ or $\chi \to -\infty$, the potential energy can be attained as

$$V(\psi) = -\frac{c\psi^3}{3} + \frac{d\psi^2}{2}. \tag{16}$$

For this prospect, we present three different graphs of the potential energy $V(\psi)$ of the system (14) in Fig. 2a–c depending on the physical parameters q_1, q_2, μ_1, μ_2, σ_1, σ_2 η, and U. $V(\psi)$ remains constant if a particle is at an equilibrium point where

(a) **(b)** **(c)**
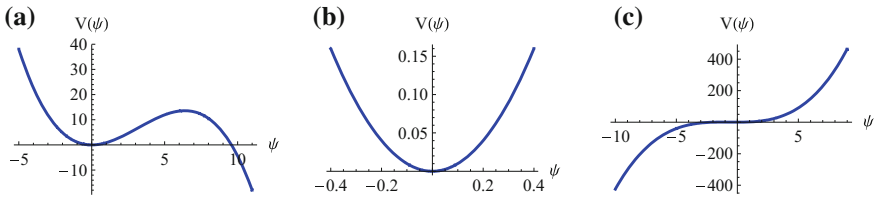
Fig. 2 Graphs of the potential energy $V(\psi)$ of Eq. (14) for the same values of parameters as shown in Fig. 1a, b, c

$\frac{dV}{d\psi} = 0$. For $\frac{dV}{d\psi} = 0$ implies $\frac{d\psi}{d\chi} = 0$, i.e., equilibria occur at the equilibrium points of the vector field. It is essential to note that local minima of $V(\psi)$ corresponds to a stable equilibrium point and local maxima corresponds to unstable equilibrium point.

6 Anti-kink Wave Solution

Integrating the system (14) with the initial condition $\psi \to 0$, $\frac{d\psi}{d\chi} \to 0$ as $\chi \to \infty$ or $\chi \to -\infty$, the anti-kink wave solution for the Burgers equation (12) is obtained as

$$\psi = \frac{d}{2c}\{1 - tanh(\frac{d}{2}\chi)\}. \tag{17}$$

The effect of the physical parameters μ_p, σ_1, σ_2, q_1, U, and η on dust-ion-acoustic anti-kink wave solution is depicted in Figs. 3, 4, 5 and 6.

The nature of the dust-ion-acoustic anti-kink wave profile with positive potential for several values of μ_p with fixed values of μ_1, μ_2, U, σ_1, σ_2, η, q_1, and q_2 is

Fig. 3 Variation of anti-kink profiles with positive potential of Eq. (12) for different values of μ_p with $\mu_1 = 0.4$, $\mu_2 = 0.04$, $\sigma_1 = 0.2$, $\sigma_2 = 2$, $q_1 = 0.8$, $\eta = 0.3$, $U = 0.4$, and $q_2 = 0.7$
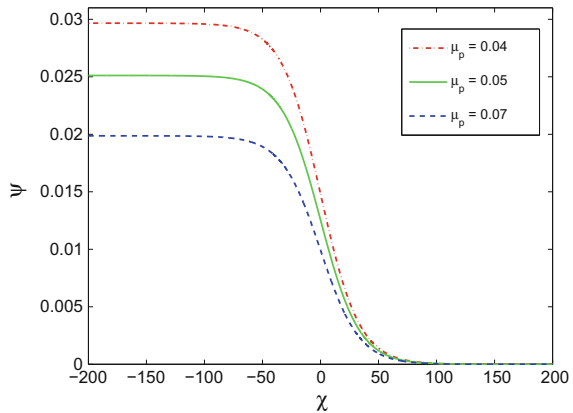
Fig. 4 Variation of anti-kink profiles with positive potential of Eq. (12) for different values of q_1 with $\mu_1 = 0.4$, $\mu_2 = 0.04$, $\sigma_1 = 0.2$, $\sigma_2 = 2$, $q_2 = 0.8$, $\eta = 0.3$, and $U = 0.4$
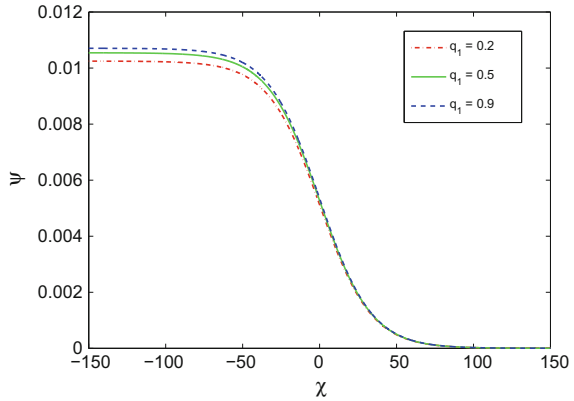
Fig. 5 Variation of anti-kink profiles with positive potential of Eq. (12) for different values of U with $\mu_1 = 0.4$, $\mu_2 = 0.04$, $\sigma_1 = 0.2$, $\sigma_2 = 2$, $q_1 = 0.8$, $\eta = 0.3$, and $q_2 = 0.7$
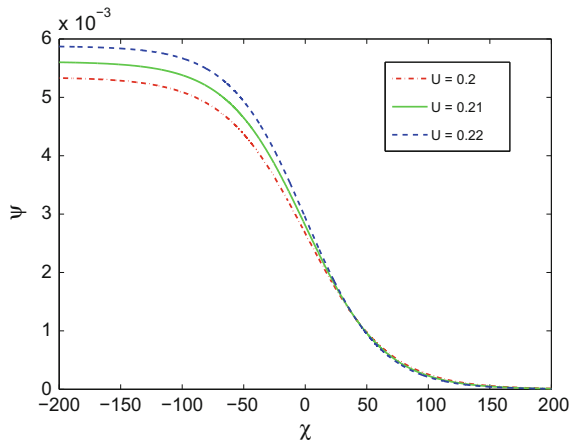
Fig. 6 Variation of anti-kink profiles with positive potential of Eq. (12) for different values of η with $\mu_1 = 0.4$, $\mu_2 = 0.04$, $\sigma_1 = 0.2$, $\sigma_2 = 2$, $q_1 = 0.8$, $U = 0.4$, and $q_2 = 0.7$
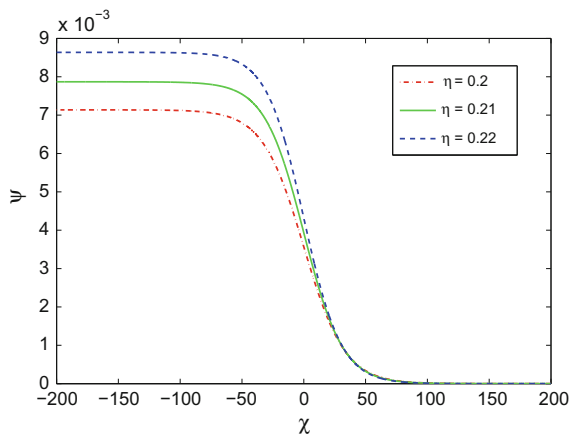
shown in Fig. 3. It is observed that amplitude of the dust-ion-acoustic anti-kink wave decreases and becomes smooth as μ_p is increased. The nature of the dust-ion-acoustic anti-kink wave profile with positive potential for several values of q_1 with fixed values of μ_p, μ_1, μ_2, U, η, σ_1, σ_2, and q_2 is shown in Fig. 4. It is observed that amplitude of the dust-ion-acoustic anti-kink wave with positive potential decreases as q_1 is increased. The properties of the dust-ion-acoustic anti-kink wave profile with positive potential for several values of U with fixed values of μ_p, μ_1, μ_2, η, σ_1, σ_2, q_1, and q_2 are shown in Fig. 5. It is observed that amplitude of the dust-ion-acoustic anti-kink wave with positive potential increases as U is increased. The properties of the dust-ion-acoustic anti-kink wave profile with positive potential for several values of η with fixed values of μ_p, μ_1, μ_2, σ_1, σ_2, q_1, and q_2 are shown in Fig. 6. It is observed that amplitude of the dust-ion-acoustic anti-kink wave with positive potential increases as η is increased. The values of the physical parameters μ_p, μ_1, μ_2, σ_1, and σ_2 taken for numerical simulations in this study are supported by the results [20].

7 Conclusions

Dynamics of the DIAAKWs in an unmagnetized multicomponent e-p-i dusty plasma have been investigated in the light of one-dimensional dynamical system. The ion kinematic viscosity has been considered, and the RPT has been applied to extract the Burgers equation. Employing the Galilean transformation, the Burgers equation has been transformed to a dynamical system. Considering all practicable vector fields, we have examined the dynamical system using the concept of potential energy and identified the stability and instability of the critical points. We observed that the viscous force acting on ion fluid is a root of dissipation and is effective for the evolution of the DIAAKWs. It has been observed that the fundamental features of the DIAAKWs are significantly affected by the influence of the physical parameters q_1, q_2, μ_1, μ_2, σ_1, σ_2, η, and U. The consequences of this investigation may be applied to understand the dynamical behavior of the DIAAKWs in space and laboratory plasma environments.

References

1. Shukla, P.K., Mamun, A.A.: Introduction to Dusty Plasma Physics. IOP Publishing, Bristol (2002)
2. Shukla, P.K., Silin, V.P.: Dust ion-acoustic wave. Phys. Scr. **45**, 508 (1992)
3. Shukla, P.K.: A survey of dusty plasma physics. Phys. Plasmas **8**, 1791 (2001)
4. Shukla, P.K., Rosenberg, M.: Boundary effects on dust-ion-acoustic and dust-acoustic waves in collisional dusty plasmas. Phys. Plasmas **6**, 1038 (1999)
5. Barkan, A., D'Angelo, N., Merlino, R.L.: Experiments on ion-acoustic waves in dusty plasmas. Planet. Space Sci. **44**, 239 (1996)
6. Merlino, R.L., Goree, J.: Dusty plasmas in the laboratory, industry, and space. Phys. Today **57**, 32 (2004)

7. Rahman, O., Mamun, A.A.: Dust-ion-acoustic solitary waves in dusty plasma with arbitrarily charged dust and vortex-like electron distribution. Phys. Plasmas **18**, 083703 (2011)
8. Alinejad, H.: Dust ion-acoustic solitary and shock waves in a dusty plasma with non-thermal electrons. Astrophys. Space Sci. **327**, 131 (2010)
9. Sen, B., Das, B., Chatterjee, P.: Effect of electron inertia on large amplitude solitary waves in presence of kinematic viscosity in dusty plasma. Eur. Phys. J. D **49**, 211 (2008)
10. Misra, A.P., Adhikary, N.C., Shukla, P.K.: Ion-acoustic solitary waves and shocks in a collisional dusty negative-ion plasma. Phys. Rev. E **86**, 056406 (2012)
11. Adhikary, N.C.: Effect of viscosity on dustion acoustic shock wave in dusty plasma with negative ions. Phys. Lett. A **376**, 1460 (2012)
12. Nakamura, Y., Bailung, H., Shukla, P.K.: Observation of ion-acoustic shocks in a dusty plasma. Phys. Rev. Lett. **83**, 1602 (1999)
13. Masud, M.M., Sultana, S., Mamun, A.A.: Effects of double temperature superthermal electrons on dust-ion-acoustic shock waves in electron-positron-ion dusty plasmas. Astrophys. Space Sci. **348**, 99 (2013)
14. Ema, S.A., Ferdousi, M., Sultana, S., Mamun, A.A.: Dust-ion-acoustic shock waves in nonextensive dusty multi-ion plasmas. Eur. Phys. J. Plus **130**, 46 (2015)
15. Guckenheimer, J., Holmes, P.J.: Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields. Springer-Verlag, New York (1983)
16. Selim, M.M., El-Depsy, A., El-Shamy, E.F.: Bifurcations of nonlinear ion-acoustic travelling waves in a multicomponent magnetoplasma with superthermal electrons. Astrophys. Space Sci. **360**, 66 (2015)
17. Saha, A., Chatterjee, P.: Solitonic, periodic and quasiperiodic behaviors of dust ion acoustic waves in superthermal plasmas. Braz. J. Phys. **45**, 419 (2015)
18. Saha, A., Chatterjee, P.: Qualitative structures of electron-acoustic waves in an unmagnetized plasma with q-nonextensive hot electrons. Eur. Phys. J. Plus **130**, 222 (2015)
19. Bains, A.S., Tribeche, M., Gill, T.S.: Modulational instability of ion-acoustic waves in a plasma with a qq-nonextensive electron velocity distribution. Phys. Plasmas **18**, 022108 (2011)
20. Lakhina, G.S., Buti, B.: Generation of a d.c. field by nonlinear electromagnetic waves in relativistic plasmas. Astrophys. Space Sci. **79**, 25 (1981)

Multiple Information Hiding in General Access Structure Visual Cryptography Using Q'tron Neural Network

Sandeep Gurung and Mrinaldeep Chakravorty

Abstract With the advent of the widespread use of Internet for communication, there is an increase in exchange of personal secured information via the Internet wherein a large variety of sensitive information is exchanged between the end users which affect the privacy and safety of the data during storage and communication. Visual cryptography, an example of a secret sharing scheme encrypts the secret digital information into a number of shares wherein the decryption is performed by overlaying the shares generated and by utilizing the HVS (Human Visual System). General access structures efficiently hides the data by defining an access structure of qualified sets which only can produce the hidden information. Since no transmission is noise free, quantum neural network is used to extract the original information even when the information is not clearly visible to the human eye. The paper proposes a methodology to conceal multiple secret in a pair of shares and the Q'tron network to improve the security of information systems.

Keywords Visual cryptography · General access structure · Random grids
Circular girds · Q'tron neural networks · Multiple information hiding

1 Introduction

Secret sharing relates the idea of uniformly distributing sensitive data such as cryptographic keys into a number of fragments to separate trusted parties; the data is decoded if a specified number of these fragments termed as the shares are available. Visual Cryptography (VC), a subset of such a scheme, decodes digital information similarly by stacking the minimum number of noisy shares and utilizes the Human Visual System (HVS) to interpret it. A random collection of noisy shares are

S. Gurung (✉) · M. Chakravorty
Department of CSE, Sikkim Manipal Institute of Technology, Sikkim, India
e-mail: gurung_sandeep@yahoo.co.in

M. Chakravorty
e-mail: mri.mc36@gmail.com

generated from written material (text, notes, pictures, etc.) taken as an input. The ideology was originally designed by Naor and Shamir [1] in the year 1994. The idea generates perfect secrecy through the use of one-time pad which is secure and is simple to use. However, due to pixel expansion the output fragments have larger size in comparison to the original secret affecting the aspect ratio of the shares, and also generating greater traffic on its transmission. The scheme conceals single secret information and employs the usage of a complex codebook for the generation of shares.

Kafri and Keren [2] proposed a novel visual secret sharing (VSS) based on non-expansion of pixels termed as random grids. The identical size of the share and the original image and the simplicity made random grids popular in the domain.

A VC scheme for General Access Structures (GAS) [3, 4] splits the information into a subset of qualified and forbidden set of participants, wherein only the participants belonging to the qualified set can reveal the information encoded. The construction of general access structures can be done by either using the basis set, a minimal qualified set, or by utilizing a collection of maximal forbidden sets.

Multiple information hidings using circular shares have been implemented in [5, 6], where one of the shares superimposed over the other is rotated at various angles to conceal multiple secret data in a share pair.

The model of Q'tron Neural Network (QNN) [7] is based on the local-minima escaping capability. In this model of NN, each neuron, called quantum neuron (Q'tron; for short), is weighted and allowed to have multiple output levels. In a Q'tron NN, the Q'tron is fully connected with symmetric connection strengths, and each Q'tron in general is self-connected with negative connection strength to provide means for negative feedback. By reducing the error function designed for GAS, the model converges to the required solution.

The organization of the paper is done accordingly. Section 1.1 gives a detailed description about random grids. Section 1.2 reviews the backgrounds of general access structures and describes the technique of using cumulative arrays. Section 1.3 reviews the Q'tron network. The proposed methodology is discussed in Sect. 2. Section 3 gives a brief overview about the design strategy for the proposed system. Section 4 gives an analysis based on the experimental results. Section 5 gives the conclusion based on the proposed system.

1.1 Random Grids

A random grid encrypts the secret information into 2D arrays of transparent and opaque pixels. The scheme is simple and does not involve any increase in pixel size. Decryption is done by overlaying the grids similarly as the traditional VC techniques. The initial grid (first) is a combination of random collections of ones and zeros. The second grid is created using the original image (secret) and the initial grid generated as a reference. Shyu [5] extended the basic (2, 2) for higher values of "n" but the clarity of the resultant information decreased with larger values of "n".

Traditional random grids encrypt single information. Chen et al. [8] used rectangular shares and their rotations to hide two secret images. A brute-force attack on the predicted angular rotation of 90°, 180°, and 270° was a major limitation to the approach used. The problem of angular restriction was solved by the introduction of circular random grids as proposed by Hsu et al. [9]. Multiple images are hidden in two circular random grids by rotating one of the grids over various angular rotations over another. These angular rotations refer to the columnar shifts. An authentication model for "n" users is designed in [10] where the required number of authentication messages is limited to n + 1. Chen et al. [6] and Hu and Chang [11] suggested other variations for using circular grids to conceal multiple secret data.

1.2 General Access Structures

General access structures [12] are useful for extracting a secret between multiple parties who work together as a group to obtain a resource. The set of participants who belong to the qualified set domain can access the secret information, whereas stacking of transparencies of the forbidden participants generates meaningless information. The set of all such qualified and unqualified set of participants is called as the access structure of the system.

1.2.1 Construction of VCS Using Cumulative Arrays

Cumulative arrays [3, 12, 13] are generated using the collection of elements in the forbidden set. It is a combination of a cumulative map (β, T) and the participant set P. The map consists of a finite mapping $\beta: P \rightarrow 2T$. It is formulated by utilizing the information of the maximal forbidden sets Z_M [14]. The column index "i" of the cumulative array is assigned a value of zero if it is present in the forbidden set. The disadvantages of the system are that each participant in P tends to get many shares of the scheme, and consequently the information rates of schemes formed in this way are low. However, the cumulative array is in fact a very significant structure in the theory of geometric secret sharing schemes. The cumulative array contains useful information when it comes to considering constructing other geometric schemes. The significance of this is that it may lead to more systematic methods for constructing geometric schemes as opposed to the somewhat ad hoc methods currently being used.

If (Γ_{Qual}, Γ_{Forb}) denote a strong access structure on the set of participants P = {1, …, n}, then ZM the maximal forbidden set can be expressed as

$$ZM = \{B \in \Gamma_{Forb} : B \cup \{i\} \in \Gamma_{Qual} \quad \text{for all} \ \ i \in P \backslash B\}. \tag{1}$$
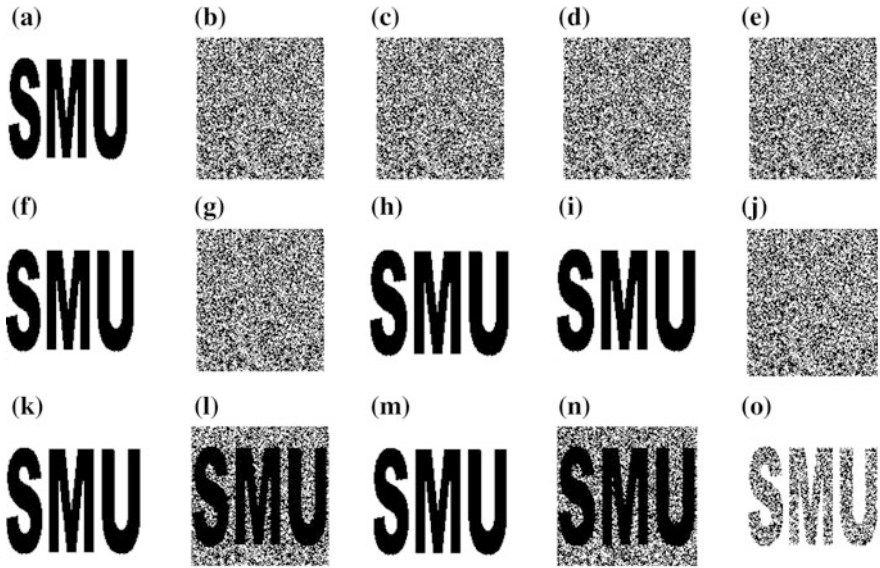
Fig. 1 **a** Original image, **b** share R1, **c** share R2, **d** share R3, **e** share R4, **f** R1 ⊕ R2, **g** R1 ⊕ R3, **h** R1 ⊕ R4, **i** R2 ⊕ R3, **j** R2 ⊕ R4, **k** R3 ⊕ R4, **l** R1 ⊕ R2 ⊕ R3, **m** R1 ⊕ R3 ⊕ R4, **n** R2 ⊕ R3 ⊕ R4, **o** R1 ⊕ R2 ⊕ R3 ⊕ R4

A cumulative array CA is a Boolean matrix such that CA $(i, j) = 1$ if and only if $i \varepsilon F_j$, where F_j is a forbidden set. An implementation of such a scheme is given in Fig. 1.

1.3 Q'tron Neural Network

QNN [14–16] is a modification of the Hopfield neural network. In this model of NN, each neuron, called quantum neuron (Q'tron; for short), is weighted and allowed to have multiple output levels. To help the model escape from infeasible local minima as in Hopfield neural network, noise is injected into each Q'tron whose strength will be controlled by a single parameter called the solution qualifier which is correlated to the solution performance or the goal. The so-built Q'tron is able to achieve local minima free by injecting persistent noise with bounded strength into each Q'tron. Specifically, the network will not be trapped at any local minima and the solution quality will only evolve further once the NN has reached a satisfactory state.

Fig. 2 **a** Input image, **b** noise added to the image, **c** retrieved image

The Q'tron Neural Network algorithm [15]

1. Initialize all weights, i = 1 to n; j = 1 to n;
           $T_{ij} = 0$
2. For each vector to be stored, do Steps 3-5
3. Set activation for each input unit, i = 1 to n;
           $x_i = s_i.$
4. For each output j = 1 to n, set activation,
           $y_j = s_j.$
5. Adjust the weights, i = 1,….n; j = 1,….,n;
           $T_{ij}(new) = T_{ij}(old) + x_i y_j$
 	
   Here, x_i is the transpose of the original matrix.

The energy function of the system may be given as

$$E = Constant * (Input\ Image) - (Output\ Image) \tag{2}$$

The main objective is to monotonically decrease the energy E defined in time until the function E reaches global minima. An implementation is given in Fig. 2.

The number of binary patterns "P" that can be learned in a net having "n" number of neurons is approximated in [14] as

$$P = 0.15 * n \tag{3}$$

2 Proposed Methodology

The basic random grid suggested by Kafri [2] is used for the creation of shares making the size of the input and output shares as identical. General Access Structure (GAS) [3, 4, 12] is utilized for generating different share pairs of forbidden and qualified sets. QNN [14, 16] is modeled to generate the general access structure by carefully formulating the error function. The added noise helps to converge to the solution by escaping from local minima. Circular grids [16, 17] are then used to encrypt multiple secrets using the idea of rotation on the shares belonging to the qualified sets.

3 Design Strategy

A digital binary image is taken as an input. The QNN-based general access structure converges to the solution by carefully minimizing the error function. The energy function is computed using Eq. 2 as E = constant * Input − $\sum$ Forbidden sets, as shown in Fig. 3.

The encryption procedure minimizes the energy function keeping the input parameter intact with changes included in the resultant shares. As the energy monotonically decreases, the system reaches a stable state and a pattern is matched which is present in the network. The decryption procedure works in the reverse manner helping us identify patterns using the Human Visual System (HVS) even when noise is induced during the time of transmission. As an example if the number of forbidden sets is two as denoted by the sets {1, 3} and {2, 4}, then two random grids RG1 and RG2 are generated. The participants 1 and 3 receive random grid RG1, and participants 2 and 4 receive random grid RG2.

The shares of the qualified sets generated by this scheme are then given to a circular grid generator for converting the rectangular grids into circular form. Multiple secrets are encoded by rotating one of the shares at an angular displacement of column/m where "m" denotes the cardinality of the secret information.
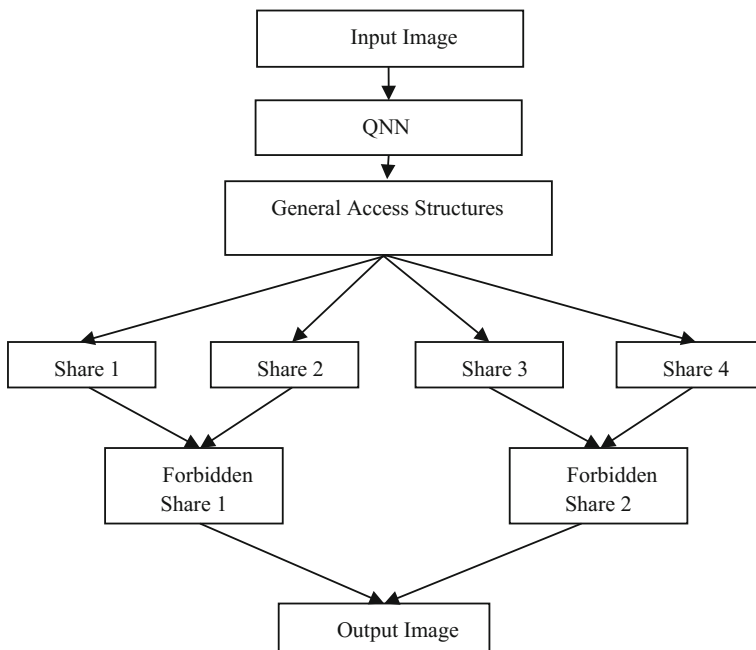
Fig. 3 Design of the proposed system

4 Experimental Results and Analysis

4.1 Experimental Results

The system takes input information of size 64 × 64 pixels. The number of patterns learned is equal to "10" using Eq. 3. These random information are taken as an input and added with external noise and then fed into a VC scheme constructed using cumulative array for GAS. The energy function E is computed as E = 1.5 * Input – ∑ Forbidden sets. The qualified sets are used to recover the first secret image. The remaining secrets are extracted by converting the rectangular grids as in Fig. 4 into circular shares and by rotating one of the shares over another over an angular displacement.

4.2 Analysis—PSNR Calculation and Contrast

The PSNR value is used as a metrics to measure the quality of the output images. Higher PSNR reflects better quality of reconstructed image. It is measured in a logarithmic (dB) scale. The mean squared error between a noise-free monochrome image I of m × n dimension and its noisy approximation K is evaluated as

$$MSE = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]2$$
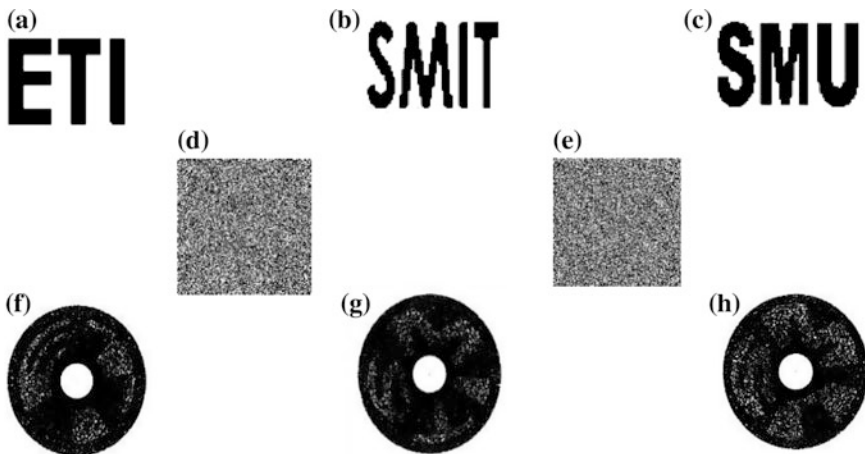
Fig. 4 **a**, **b** and **c** Secret information, **d** and **e** qualified shares 1 and 2, **f**, **g** and **h** circular secret information 1, 2 and 3
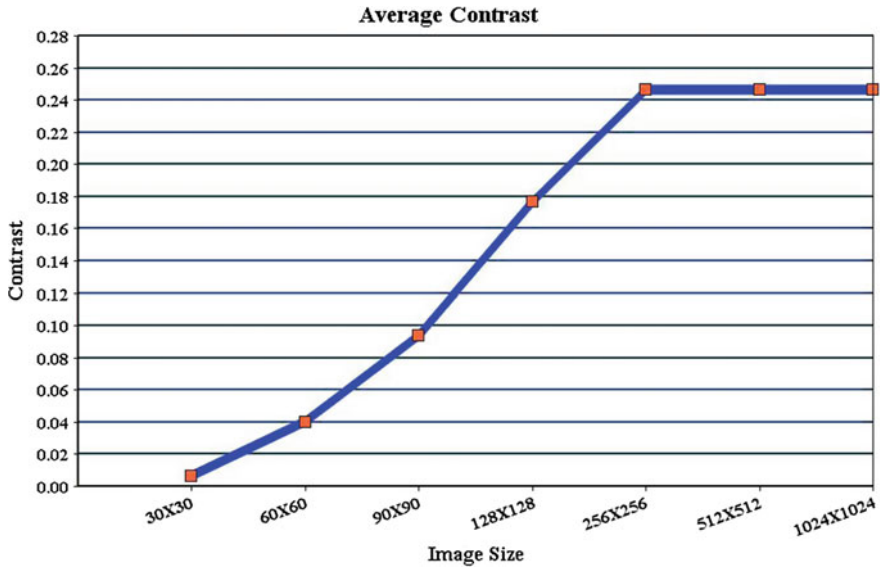
Fig. 5 Contrast calculation for different sized information

Table 1 Comparative analysis of PSNR for three embedded information

Input	Intermediate Shares	Output	PSNR
ETI	**Share 1**		4.1178
SMU			3.8528
SMIT	**Share 2**		3.5664

The PSNR value is then computed as

$$PSNR = 10. \log 10 \, (MAX2I/MSE) = 20. \log 10 \, (MAXI/\sqrt{MSE})$$
$$= 20 \log 10 \, (MAX) - 10. \log 10 \, \sqrt{(MSE)},$$

where MAXI is the maximum value of a pixel in an image. The contrast for images of varying sizes is given in Fig. 5.

The contrast achieved for larger images is close to 26% making the secret information easily visible to the human visual system. The comparative analysis of the PSNR of the secret images is given in Table 1.

5   Conclusion

The paper may be concluded by stating the facts that

1. The systems generate circular random grids without pixel expansion to conceal the angular orientation and also hides compound secret information.
2. The mechanism achieves confidentiality of data and authentication of the end users.
3. The ideology can be upgraded to incorporate encryption of grayscale and colored images.
4. Codebook needless: Traditional approaches fulfill the access schemes of visual cryptography using codebooks. For complex access schemes, a codebook is hardly to be found and/or not existent.
5. Generality: The approach is very general and, hence, can be used to fulfill any access scheme for visual cryptography.

References

1. Naor, M., Shamir, A.: Visual cryptography. In: Advances in Cryptology Eurocrypt (1995)
2. Kafri, O., Keren, E.: Encryption of pictures and shapes by random grids. Opt. Lett. (1987)
3. Kumar, S., Sharma, R.K.: Secret image sharing for general access structures using random grids. IJCA **83**(7) (2013). ISSN 0975-8887
4. Ateniese, G., Blundo, C., De Santis, A., Stinson, D.R.: Visual cryptography for general access structures. Inf. Comput. **129**, 86–106 (1996)
5. Shyu, S.J.: Image encryption by multiple random grids. Pattern Recogn. **42**, 1582–1596 (2009)
6. Chen, J., Chen, T.-S., Hsu, H.-C., Chen, H.-W.: New visual cryptography system based on circular shadow image and fixed angle segmentation. J. Electron. Image **14**(3), 033018 (2005)
7. Chiang, S.C., Yue, T.-W.: Visual cryptography using Q'tron neural networks. Ph.D. Dissertation, Department of Computer Science

8. Chen, T.H., Tsao, K.H., Wei, K.C.: Multiple-image encryption by rotating random grids. In: Proceedings of the 8th International Conference on Intelligent System Design and Applications (2008)

9. Hsu, H.C., Chen, J., Chen, T.S., Lin, Y.H.: Special type of circular visual cryptography for multiple secret hiding. Imaging Sci. J. **55**(3), 175–179 (2007)

10. Gurung, S., Chakravorty, M., Agarwal, A., Ghose, M.K.: Multiple information hiding using circular random grids. Conference Organized by Inderscience Institute of Management and Technology, Bhubaneswar, Odisha, India. Procedia Comput. Sci. **48**, 65–72 (2015)

11. Wu, H.-C., Chang, C.-C.: Sharing visual multi-secrets using circle shares. Comput. Stan. Interfaces **28** 123–135 (2005)

12. Guo, C., Chang, C.C.: A construction for secret sharing scheme with general access structure. J. Inf. Hiding Multimedia Signal Proc. **4**(1), (2013). ISSN 2073-4212

13. Adhikari, A., Dutta, T.K., Roy. B.: A new black and white visual cryptographic scheme for general access structures. In: INDOCRYPT'04. LNCS, vol. 3348, pp. 399–413. Springer (2004)

14. Fausett, L.: Fundamental of Neural Networks—Architectures, Algorithms, and Applications. Pearson Education (2008)

15. Feng, B., Wu, H.C., Tsai, C.S., Chu, Y.P.: A new multi-secret images sharing scheme using Lagrange's interpolation. J. Syst. Softw. **76**(3), 327–339 (2005)

16. Shyu, S.J.: Visual cryptograms of random grids for general access structures. IEEE Trans. Circuits Syst. Video Technol. **23**(3) (2013)

17. Chen, T.-H., Li, K.-C.: Multi-image encryption by circular random grids. Department of Computer Science and Information Engineering, National Chiayi University, Chiayi City 60004, Taiwan (2011)

Overlapping Community Detection Through Threshold Analysis on Disjoint Network Structures

Sudeep Basu and Indrajit Pan

Abstract Distributed approach in a network is a prime attribute to achieve quality throughput. Many real-life infrastructures share such distributed network structures. Recently, researchers are focusing on different prime attributes of these distributed networks and meaningfully analyzing them to retrieve essential information toward throughput enhancement. These structures exhibit different constructs. Some of those are static and some are dynamic. They also contain strategic groups within it. Appropriately, identifying these groups is the key essence of community detection. Present work applies a novel mechanism for graphical analysis of network structures to detect overlapping communities. Experimental findings and comparative analysis with existing methods show efficacy of the present algorithm.

Keywords Betweenness study · Community detection · Disjoint community
Overlapping community · Threshold analysis

1 Introduction

The concept of network is not only confined in computer science but also applicable in other fields like electrical, electronics, production, management, etc. A group of some distinct yet inter-related entities shape these networks. In practical perception, these real-world networks are very complex in nature. It requires in-depth analysis. A deliberate investigation is required for different network types which ensemble computer network, electrical grid, or some virtual network falling under this category. Study of this complex networks is the key task for its next-level betterment in terms of service, quality, and offerings for advanced applications.

S. Basu (✉)
Siliguri Institute of Technology, Siliguri, India
e-mail: basu.sudeep@gmail.com

I. Pan
RCC Institute of Information Technology, Kolkata, India
e-mail: indrajit.pan@rcciit.org

Being inspired from the colloquial behavior of human societies and communities, scientists have provided a deep insight for these complex networks and found community structures as key tool for strategic analysis. A complex network is said to have many community structures if that can be separated in different distinct groups of members.

These groups are identifiable through the behavior of their members. Members of a group seem to have a dense connection or interaction with other members of that group. Incidentally, they maintain a sparse communication with the members belonging to the same network but not to the same group. Community detection has gained due importance, and nowadays it is a part of recent research trends.

Nowadays, researchers are working on different aspects of community structures within a network perspective. This article reports a novel graph-theoretic approach for detecting overlapping community structures in such large networks. Members of such large networks share high degree of association. This makes some entities to be a part of multiple groups or community structures, and the scenario is known as overlapping community.

In literature, not much of works are reported on this. In the following section, some of the recent research trends will be discussed to present the current status of research in this domain.

1.1 Recent Works

A work in [1] represents a clique partitioning approach toward community detection. The approach is termed as clique percolation method. This method suffers a basic inability to categorize a member node into any community structure if that node does not belong to any clique. Hence, the approach is not suitable for large network or any network containing isolated component. A k-clique percolation method is presented in [2]. This approach is reported to retrieve overlapping community structures from a given network.

Another graphical approach is reported in [3]. Mainly, it analyzes the similarity of nodes on the basis of vertex incidence on edges.

An adaptive approximation approach for identifying communities is discussed in [4]. This approach mainly focuses on betweenness and modularity issue. Betweenness of vertices is found to be processed through modularity analysis. Modularity analysis finally returns community structures. This method yields community structure on static network but inadequate to retrieve overlapping scenarios within the members.

A label propagation method is proposed in [5]. This is one of the fastest algorithms proposed in the literature; however, randomness of the method leaves it in inestimable delay in case of dynamic network structures. It produces a progressive consensus mechanism to construct communities. An extension of this method is presented in [6]. Here, each node is associated with a memory. This memory keeps

a track of most frequently referred label in the network. Due to association with memory, this method has shown little improvement over the previous one in [5].

An overlapping community detection approach is reported in [7]. It proposes a microblog-user model. Initially, it frames a microblog-user network which further analyzed in the form of microblog-user model to detect overlapping community structures through similarity analysis.

Some other works reported in [8] approach with the strength of link algorithms and a clique merge method is discussed in [9]. However, none of these two are focused on overlapping community detection.

1.2 Clustering Metrics

Community structure in a network is identified through strength of connections among different members. It requires an analysis on density of connections among these members. Sometimes, it evolves as a ratio of dense connection to sparse connection. There are many techniques for measuring this strength of connection among the members. *Betweenness analysis* is one such competent approach among them which was proposed by Freeman [4]. *Vertex betweenness* and *Edge betweenness* are effective metric for measuring communities in a network [4].

Vertex betweenness is a measurement of a vertex (v) in a network (N) with respect to a pair of vertices v_1 and v_2. A count (c_t) of shortest paths between v_1 and v_2 and a count (c_p) of those which pass through v is taken. Then, the vertex betweenness of v with respect to v_1 and v_2 is represented by a ratio between c_p and c_t. Vertex betweenness of v with respect to whole graph is determined by sum of all such ratios for every possible pair of vertices present in the graph, excluding v.

Edge betweenness concept was proposed on vertex betweenness. An edge betweenness measure for an edge (e) is taken with respect to a pair of vertices v_1 and v_2. It is the number of shortest paths between v_1 and v_2, those of which passes through e. If there is more than one shortest path, then the weight is distributed proportionately.

1.3 Motivation

Betweenness centrality is a linear analysis [4] of finding similarity among members of a network. Very few works are reported in literature on overlapping community detection and it is still an open problem to address.

In this work, a network will be first analyzed for disjoint communities. A novel graph-theoretic analysis will be applied on these disjoint communities to retrieve overlapping communities from that existing network. These disjoint communities will be analyzed for betweenness study. The work will be further extended to form overlapping community structure over the given network.

2   Problem Formulation

A network is represented by a graph G. A graph G contains a set of vertices V and a set of edges E. Initially, on the basis of vertex betweenness, this network will be partitioned into clusters or sub-graphs (G_s^1, G_s^2, G_s^3,…, G_s^k).

Any of those sub-graphs (G_s^n) contains a set of vertices (V_s^n) and edges (E_s^n). A disjoint pair of sub-graph (G_s^i and G_s^j) will be represented by Eq. 1

$$V_s^i \cap V_s^j = E_s^i \cap E_s^j = \phi \tag{1}$$

Final objective is to find out overlapping communities by analyzing the communities. This analysis can be done by Eq. 2, where for any two existing communities (C_s^i and C_s^j),

$$V_s^i \subseteq V_s^j \| V_s^j \subseteq V_s^i \| E_s^i \subseteq E_s^j \| E_s^j \subseteq E_s^i \tag{2}$$

3   Proposed Method

Present method applies the concept of vertex betweenness to compute disjoint community structures. These disjoint communities are strictly nonoverlapping in nature and there is no single common member among any pair of communities. In the process of forming disjoint communities, each node of the network is analyzed for its vertex betweenness value. Then, the nodes are sorted in a descending order of their betweenness values. Nodes or vertices having same betweenness value are grouped together. Each of these groups represents a community within the network and those are nonoverlapping in nature.

After formation of disjoint communities, these communities are taken in pair for analysis to develop overlapping communities. During vertex betweenness analysis, lengths of shortest path between all pair of vertices are determined. A record of maximum length shortest path among all such pairs is taken in (m_{sp}). A threshold for analysis is set in t_h, where

$$t_h = 0.5 \times (m_{sp}) \tag{3}$$

Now communities are taken in pair, where they belong in two disjoint sets. Considering the edges of master network, bipartite connections are introduced between two sets. Now all vertices are individually checked for their incidency (number of edges incident upon that vertex) value in the bipartite graph. If any vertex achieves incidency greater or equal to t_h, then that is merged with another set to form an overlapping community.

3.1 Proposed Algorithm

a. Read the vertex (V) and edge (E) set of given network.
b. For all vertex (v) in V, find out the betweenness (v_b).
c. Sort all vertices (v) in descending order of betweenness (v_b).
d. Form nonoverlapping clusters (C_k) where k = 1 to m, by taking V_i where i = 1 to n from the list S having the same betweenness.
e. Determine threshold (t_h)
f. **loop:** i = 1 to (m−1)

    i. **loop:** j = (i +1) to m

        1. Consider C_i and C_j in set V_1 and V_2, respectively.
        2. Introduce bipartite connections from (E) between V_1 and V_2.
        3. Initialize an overlapping community (C_{v2}) for V_2
            a. $C_{v2} = C_{v2} \cup V_2$
            b. $\forall$ v in V_1, find incidency(v)
            c. **if** incidency(v) $\geq t_h$

                i. $C_{v2} = C_{v2} \cup$ v

                end if
            d. **return** C_{v2} **if** $C_{v2} \neq V_2$

        4. Initialize an overlapping community (C_{v1}) for V_1
            a. $C_{v1} = C_{v1} \cup V_1$
            b. $\forall$ v in V_2, find incidency (v)
            c. **if** incidency(v) $\geq t_h$

                i. $C_{v1} = C_{v1} \cup$ v

                end if
            d. **return** C_{v1} **if** $C_{v1} \neq V_1$

        end loop

    end loop

4 Experimental Results

4.1 Experiment Setup

Proposed algorithm explained in the previous section was implemented in Java platform. The program was executed in Windows 10 platform, on an AMD A4 microprocessor-based chipset board having 4 GB primary memory.

This proposed method was primarily simulated on two different customized networks as illustrated below in demo network I and demo network II. These two networks are discussed in the following section called *Simulator Network*. The purpose of execution on these simulator networks is to judge the effectiveness of the proposed method. Also, the performance testing was done through these dummy runs.

Finally, this proposed method was simulated on a benchmark data suit retrieved from [10]. Test result on the benchmark data shows that present method is working convincingly. Efficiency of this method seems better than the techniques proposed in [4, 8].

4.2 Simulator Network

Demo network I and II: Demo network I is given in Fig. 1 and demo network II is in Fig. 2. Both the graphs in Figs. 1 and 2 contain 6 nodes and 7 edges but the orientation of edges is different. In these figures, nodes represent the members in a network and edges represent the connection in between those members.

After initial run of the proposed algorithm, Table 1 is generated, which yields all vertices of the networks shown in Figs. 1 and 2 along with their vertex betweenness values sorted in descending order. A sample vertex betweenness calculation of vertex "3" of Fig. 1 is shown below:

$\frac{c_p}{c_t}$	1	2	4	5	6
1	0	$\frac{0}{1}$	$\frac{1}{1}$	$\frac{2}{2}$	$\frac{1}{1}$
2	$\frac{0}{1}$	0	$\frac{1}{1}$	$\frac{2}{2}$	$\frac{1}{1}$
4	$\frac{1}{1}$	$\frac{1}{1}$	0	$\frac{0}{1}$	$\frac{0}{1}$
5	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{0}{1}$	0	$\frac{0}{1}$
6	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{0}{1}$	$\frac{0}{1}$	0

$=$

$\frac{c_p}{c_t}$	1	2	4	5	6
1	0	0	1	1	1
2	0	0	1	1	1
4	1	1	0	0	0
5	1	1	0	0	0
6	1	1	0	0	0

$= 12$

Vertex betweenness values recorded in Table 1 help to build disjoint and nonoverlapping communities. Members of each such community are said to be

Fig. 1 Demo network I
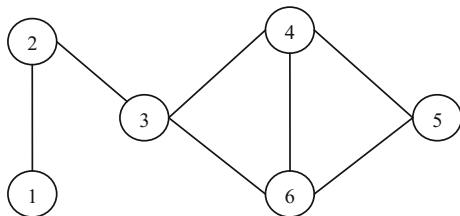
Fig. 2 Demo network II
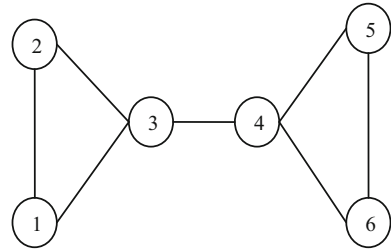
Table 1 Vertex betweenness in descending order for Figs. 1 and 2

Demo network I		Demo network II	
Vertex	Betweenness	Vertex	Betweenness
3	12	4	12
2	8	3	12
6	8	1	0
4	3	2	0
5	3	5	0
1	0	6	0

strongly related and they share same betweenness values. Table 2 contains details of those communities with respect to Figs. 1 and 2.

Information retrieved from Table 2 is further used for feature analysis of the communities. Here, all possible pair of communities in each network is compared for common vertices and disjoint vertices. A ratio between number of common vertices and disjoint vertices are obtained. Finally, these ratio values are compared with threshold as discussed in the previous section to detect overlapping communities. Result set is given in Table 3.

Table 2 Disjoint communities identified from Table 1

Demo network I		Demo network II	
Community	Members	Community	Members
C1	[6]	C1	[6, 7]
C2	[2, 11]	C2	[2, 8, 10, 11]
C3	[7, 10]		
C4	[8]		

Table 3 Overlapping community information in Figs. 1 and 2

Sl. No.	Demo network I	Demo network II
	Overlapping community	Overlapping community
1	[2, 6, 11]	[2, 6–8, 10, 11]
2	[7, 10, 11]	–

Table 4 Overlapping community information on benchmark data [10] and comparison with [7]

Proposed method	[7]	[1]
Overlapping community	Overlapping community	Overlapping community
C5-C7-C29	C1-C6-C12	C1-C6-C12
C1-C6-C12	C5-C7-C29	C5-C7-C9-C11-C29
C2-C3-C4-C8-C15	C13-C23-C27	C2-C3-C4-C8-C15
C7-C9-C11-C19		C13-C23-C27
C13-C23-C27		

4.3 Performance on Benchmark Data

Benchmark data as explained in [10] was simulated using this proposed method. The simulation was carried out in a similar way as explained in above section for Figs. 1 and 2. Apart from that, the methods of [4, 8] were also simulated on benchmark data [10] and the final comparative result is given in Table 4. Table 4 shows that the proposed method finds some excellent result in community detection. This present method is found to be more efficient in detection of communities. Efficacies of the present method are more evident from Table 4 in terms of exploring in detail analysis and record more overlapping features among communities.

5 Conclusion

This work proposes a novel technique to identify overlapping and disjoint communities. In the study of existing research, it was noted that uncovering disjoint communities from a network only divulges limited information regarding existing communities and the members of those communities. The method illustrated by Girvan and Newman [6] only retrieves disjoint communities from a network. This work explores overlapping members across the communities in a network and thus exposes the details of overlapping communities.

References

1. Palla, G., Derenyi, I., Farkas, I., Vicsek, T.: Uncovering the overlapping community structure of complex networks in nature and society. Nature **435**(7043), 814–818 (2005)
2. Derenyi, I., Palla, G., Vicsek, T.: Clique percolation in random networks. Phys. Rev. Lett. 1–4 (2005)
3. Rao, B., Mitra, A.: A new approach for detection of common communities in a social network using graph mining techniques. In: Proceedings of International Conference on High Performance Computing and Applications, pp. 1–6. Bhubaneswar, India (2014)

4. Sutaria, K., Joshi, D., Bhensdadia, C.K., Khalpada, K.: An adaptive approximation algorithm for community detection in social network. In: Proceedings of International Conference on Computational Intelligence and Communication Technology, pp. 785–788. Ghaziabad, India (2014)
5. Raghavan, U.N., Albert, R., Kumara, S.: Near linear time algorithm to detect community structures in large-scale networks. Phys. Rev. E (Physics and Society) **76**(3), 036106. arXiv:07092938 (2007)
6. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. In: Proceedings of the National Academy of Sciences of the United States of America, vol. 99 (12), pp. 7821–7826. PNAS, USA (2002)
7. Gu, Y., Zhang, B., Zou, G., Huang, M., Jiang, K.: Overlapping community detection in social network based on microblog user model. In: Proceedings of International Conference on Data Science and Advanced Analytics, pp. 333–339. Shanghai, China (2014)
8. Basuchowdhuri, P., Chen, J.: Detecting communities using social ties. In: Proceedings of IEEE International Conference on Granular Computing, pp. 55–60. San Jose, CA, USA (2010)
9. Newman, M.E.J.: Modularity and community structure in networks. In: Proceedings of the National Academy of Sciences of the United States of America, vol. 103 (23), pp. 8577–8582. PNAS, USA (2006)
10. Lancichinetti, A,. Fortunato, S.: Community detection algorithms: a comparative analysis. Phys. Rev. E **80**, 056117, Physics and Society (2009)
11. Maity, S., Rath, S. K.: Extended clique percolation method to detect overlapping community structure. In: Proceedings of International Conference on Advances in Computing, Communications and Informatics, pp. 31–37. New Delhi, India (2014)

Chaotic to Periodic Phenomena of Dust-Ion-Acoustic Waves in a Collisional Dusty Plasma

Tushar Kanti Das and Prasanta Chatterjee

Abstract The existence of chaotic and periodic structures of dust-ion-acoustic nonlinear waves (DIAWs) is reported through dynamical system approach in a dusty plasma with dust-ion collision effect. Employing the reduction perturbation technique (RPT), the damped modified Korteweg-de Vries (DmKdV) equations are derived. The nonlinear wave phenomena of perturbed DmKdv equation is studied in presence of an applied external force. Chaotic behavior is found in some critical composition for the said perturbed system. It is seen that dust-ion collisional frequency acts as a controlling parameter to chaos and changes the structure from chaotic motion to periodic motion.

Keywords DIAWs · Collision · Dynamical system · Chaotic · Periodic orbit

1 Introduction

A dusty plasma is defined as the plasma comprising electrons, ions, neutral particles, and micro-sized charged dust particles [1–8]. The study of different kinds of communal processes in dusty plasmas is of considerable significance in apprehending many features of dusty plasmas that are detected in astrophysical environments, for example, comet tails, planetary rings, the lower part of the earth's ionosphere, and experimental plasma [3–6]. There are two type of acoustic modes in dusty plasmas such that dust acoustic mode including mobile dust particles, and dust-ion-acoustic mode with mobile ions and dusts. The DIAW is an ion acoustic mode reformed by

T. Kanti Das (✉) · P. Chatterjee
Department of Mathematics, Siksha-Bhavana, Visva-Bharati University,
Santiniketan 731235, India
e-mail: tushardas2014math@gmail.com

P. Chatterjee
e-mail: prasantachatterjee1@rediffmail.com

the existence of dusts. Shukla and Silin [7] reported theoretically that a dusty plasma comprising negatively charged stationary dust grains follows low-frequency DIAWs for the first time in the literature. Furthermore, Barkan et al. [8] performed an experiment on DIAWs in laboratory plasmas. The dusty plasma is a major interdisciplinary research area to explain the fundamental plasma physics [9, 10].

Shukla and Mamun [11] reported the dust acoustics shock in a strongly coupled dusty plasma. Mamun and Eliasson [12] investigated the DA shocks for strong correlation among arbitrary charged dust particles. Rahaman et al. [13] studied nonplanar dust acoustic waves in an adiabatic dusty plasma. Garaiet et al. [14] observed the velocity shear effect on the longitudinal wave in a dusty plasma. Chakrabarti and Ghosh [15] studied the longitudinal dust acoustic solitary waves in a dusty plasma. In 2000, Ghosh et al. [16] reported the nonlinear features of small amplitude DIASWs in a three-component dusty plasma comprising electrons, ions, and dust particles. They observed that amplitude of the wave exponentially decays with time because of dust-ion collision. Moslem [17] showed the effects of dust-neutral collision for dust acoustic waves (DAWs) in dusty plasmas with ambient magnetic field. It was found that, due to collisions, the dust acoustic damp waves and the damping rate of the waves depend mainly on the collisional frequency. Some authors [18–21] pointed out the influence of dust-ion collision on the wave propagation in a collisional dusty plasma. Recently, few researchers [22–25] studied different features of DIA waves employing the theory of planar dynamical systems.

The ideas for controlling chaos were demarcated by stabilizing a periodic orbit which is not stable. The main goal resided in waiting for an unrefined passage of the chaotic orbit close to the desired periodic behavior, and then employing a small perturbation with good judgment, in order to stabilize all such periodic dynamics. Most recently, chaotic oscillation through quasiperiodic behaviors of the positron acoustic waves had been analyzed by Saha and Tamang [26].

However, no work has been reported to study chaos control in dusty plasma. Here, dust-ion collisional frequency plays the role of controlling parameter to chaos. The study is done in the framework of DmKdV equation. It has been observed that at a critical temperature ratio, the chaos is occurred and dust-ion collisional frequency plays a key role in the transition of DIAWS from chaotic motion to periodic oscillation in an unmagnetized collisional dusty plasma.

The presentation of remaining part is as follows. The governing equations are provided in Sect. 2. In Sect. 3, we have derived the DKdV and DmKdV equations. Discussion about chaotic and periodic structures of the corresponding three-dimensional dynamical system is given in Sect. 4. Conclusions of this study are given in Sect. 5.

2 Model Equations

We contemplate a collisional dusty plasma that consists of negatively charged ions, Gaussian distributed dust particles, and q-nonextensive electrons with density function [27] $f_e(v) = C_q \{1 + (q-1)[\frac{m_e v^2}{2k_B T_e} - \frac{e\phi}{k_B T_e}]\}^{\frac{1}{q-1}}$, where the normalization constant

is followed by $C_q = n_{e0} \frac{\Gamma(\frac{1}{q-1})}{\Gamma(\frac{1}{q-1}-\frac{1}{2})} \sqrt{\frac{m_e(1-q)}{2\pi k_B T_e}}$, for $-1 < q < 1$ and $C_q = n_{e0} \frac{1+q}{2} \frac{\Gamma(\frac{1}{q-1}+\frac{1}{2})}{\Gamma(\frac{1}{q-1})} \sqrt{\frac{m_e(q-1)}{2\pi k_B T_e}}$, for $q > 1$ and other parameters have their usual significance. The basic equations are

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(nu) = 0, \tag{1}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = -3n\frac{\partial n}{\partial x} - v_{id}u - \frac{\partial \phi}{\partial x}, \tag{2}$$

$$\frac{\partial^2 \phi}{\partial x^2} = \delta_1 n_d + \delta_2 n_e - n, \tag{3}$$

where n_d denotes number density of dusts given by $n_d = e^{\sigma \phi}$ with $\sigma = \frac{T_d}{T_i}$ and number density of electrons is given by $n_e = [1 + (q-1)\beta\phi]^{\frac{1}{2}+\frac{1}{q-1}}$ with $\beta = \frac{T_e}{T_i}$. Here, T_i, T_e, T_d denote the temperature of ions, electrons, and dust, respectively. At equilibrium, $n_{i0} = z_d n_{d0} + n_{e0}$ which gives $\delta_1 + \delta_2 = 1$ with $\delta_1 = \frac{n_{d0}z_d}{n_0}$ and $\delta_2 = \frac{n_{e0}}{n_0}$, where n_{i0}, n_{d0} and n_{e0} denote unperturbed number densities of ions, dusts, and electrons, respectively. In this case, z_d denotes the dust charge number and the charge of the dust is taken as $q_d = -ez_d$, where e denotes the elementary charge. The ion density n, velocity u, and electrostatic potential ϕ are normalized to n_{i0}, C_i, and T_i/e, respectively, where $C_i = (\frac{T_i}{m_i})^{\frac{1}{2}}$ denotes ion acoustic speed and m_i denotes mass of ions. Here, v_{id} denotes dust-ion collisional frequency. The variables x and t are normalized to Debye length $\lambda_D = (\frac{T_i}{4\pi e^2 n_0})^{\frac{1}{2}}$ and ω_p^{-1}, respectively, where ion plasma frequency $\omega_p = (\frac{4\pi e^2 n_0}{m_i})^{\frac{1}{2}}$.

3   Derivation of the DKdV and DmKdV Equations

We appoint the RPT to procure the DKdV and DmKdV equations. In case of RPT technique, the normalized model is linearized to get the relation between ω and k which is called the linear dispersion relation where ω is the wavenumber and k is its frequency. This dispersion relation of the linearized system characterizes the stretching of independent variables x and t as

$$\xi = \varepsilon^{1/2}(x - Mt), \tag{4}$$

$$\tau = \varepsilon^{3/2}t, \tag{5}$$

where M denotes the phase velocity of dust-ion acoustic wave and ε denotes a small parameter that measures the weakness of the nonlinearity. The dependent variables are expanded as

$$n = 1 + \varepsilon n_1 + \varepsilon^2 n_2 + \varepsilon^3 n_3 + \cdots, \tag{6}$$

$$u = 0 + \varepsilon u_1 + \varepsilon^2 u_2 + \varepsilon^3 u_3 + \cdots, \tag{7}$$

$$\phi = \varepsilon \phi_1 + \varepsilon^2 \phi_2 + \cdots, \tag{8}$$

$$v_{id} \sim \varepsilon^{3/2} v_{id0}, \tag{9}$$

where v_{ido} is the dust-ion collisional frequency. Substituting Eqs. (6)–(9) along with the stretching (4)–(5) into Eqs. (1)–(3) and equalizing the coefficients of smallest order of ε, we have

$$M^2 = 3 + \frac{1}{a}, \tag{10}$$

where $a = \delta\sigma + \frac{q+1}{2}\beta\delta_2$.

Considering the coefficients of next higher order of ε, we have the damped KdV as

$$\frac{\partial\phi}{\partial\tau} + A\phi_1\frac{\partial\phi_1}{\partial\xi} + B\frac{\partial^3\phi_1}{\partial\xi^3} + C\phi_1 = 0, \tag{11}$$

where $A = \frac{M(12a^2+3a^2-2b)}{2(3a^2+a)}$, $b = \frac{1}{2}\delta_1\sigma^2 + \delta_2\frac{(q+1)(3-q)}{8}\beta^2$, $B = \frac{1}{2a^2M}$, and $C = \frac{v_{id0}}{2}$

When $\sigma = 0.08933$, $q = 0.5$, $\beta = 0.75$, $\delta_1 = 0.45$, and $\delta_2 = 0.55$, there will be a critical point where $A \simeq 0$ that will show a huge extension to the features of traveling wave solution of Eq. (11) that disintegrates the soundness of the RPT. Here, $\sigma, q,\ \delta_1,\ \delta_2,\ \beta$ are known as critical physical parameters with special values $\sigma = 0.08933$, $q = 0.5$, $\beta = 0.75$, $\delta_1 = 0.45$ and $\delta_2 = 0.55$. However, there is a collection of special values of the critical parameters. To receive all information of DIAWs at the critical composition, a higher order nonlinear equation is taken into account. Therefore, we proceed to consider the higher order coefficients of ε. Substituting Eqs. (4)–(5) and expanded dependent variables (6)–(8) with new perturbed quantity $v_{id} \sim \varepsilon^2 v_{id0}$ into the system of equations (1)–(3) and equating the coefficients smallest order of ε, we can develop the dispersion relation as Eq. (10). Thereafter, taking into consideration the coefficients of succeeding higher orders of ε, we have the DmKdV equation as

$$\frac{\partial\phi}{\partial\tau} + A'\phi_1^2\frac{\partial\phi_1}{\partial\xi} + B\frac{\partial^3\phi_1}{\partial\xi^3} + C\phi_1 = 0, \tag{12}$$

where $A' = \frac{3a^2M^2(3b-2a^2)+9ab-3c}{2a^2M}$, $b = \frac{1}{2}\delta_1\sigma^2 + \delta_2\frac{(q+1)(3-q)}{8}\beta^2$, $c = \frac{\delta_1\sigma^3}{6} + \delta_2\frac{(q+1)(3-q)(5-3q)}{48}\beta^3$, $B = \frac{1}{2a^2M}$, and $C = \frac{v_{id0}}{2}$.

4 Chaotic and Periodic Structure at Critical Values

In order to investigate the chaotic and periodic structure at critical values of the DmKdV equation using the theory of dynamical system, we transform the DmKdV equation (12) to an ordinary equation using the transformation

$$\eta = l\xi - \lambda\tau. \tag{13}$$

Here, λ denotes velocity of the traveling wave and l denotes slope of the line wave propagation with respect to ξ-axes. Substituting $\psi(\eta) = \phi_1(\xi, \tau)$ into the DmKdV equation (12), we get

$$\frac{d^3\psi}{d\eta^3} + \frac{A'}{Bl^2}\psi^2\frac{d\psi}{d\eta} - \frac{\lambda}{Bl^3}\frac{d\psi}{d\eta} + \frac{C}{Bl^3}\psi = 0. \tag{14}$$

Considering an external periodic perturbation $f_0 \cos(\omega\eta)$, the Eq. (14) can be converted into the following three-dimensional dynamical system:

$$\begin{cases} \frac{d\psi}{d\eta} = Z_1, \\ \frac{dZ_1}{d\eta} = Z_2, \\ \frac{dZ_2}{d\eta} = -\frac{A'}{Bl^2}\psi^2 Z_1 + \frac{\lambda}{Bl^3}Z_1 - \frac{C}{Bl^3}\psi + f_0 cos(\omega\eta), \end{cases} \tag{15}$$

where f_0 and ω denote strength and frequency of the externally perturbed periodic force. To investigate the influence of dust-ion collisional frequency (ν_{ido}) on the chaotic to periodic phenomena, we apply numerical techniques, as (i) three-dimensional phase portraits analysis, (ii) time series analysis, and (iii) Poincare section. The perturbed system (15) is consisting of different parameters as l, δ_1, δ_2, λ, σ, ν_{id0}, β, q, ω, and f_0.

Phase portrait is a geometric presentation of the qualitatively different trajectories of a dynamical system in its phase space. In this case, we show all feasible phase portraits of the system (15) in the (ψ, Z_1, Z_2)-space. In Fig. 1a, three-dimensional phase portrait for the system (15) is depicted for the critical values $l = 0.4$, $\delta_1 = 0.75$, $\delta_2 = 0.25$, $\lambda = 0.5$, $\sigma = 0.089333$, $\beta = 0.5$, $q = 0.45$, $\omega = 0.5$, $\nu_{id0} = 0$, and $f_0 = 1.0$ showing the chaotic features for the system. In Fig. 1b, we plot the variation of ψ, Z_1, and Z_2 against η for the system (15) with same value of physical parameters as in Fig. 1a. The qualitatively different trajectories, shown in Fig. 1a, and variation of ψ, Z_1, and Z_2 with η in Fig. 1b ignore the periodic motion and exhibit chaotic oscillations. Moreover, a very applicable theory for determining the chaotic attractors is Poincare sections. With the parameter values as in Fig. 1, the poincare section of the system (15) is followed by Fig. 2 which depics that the points are densely scattered and exhibit irregular distributions without any definite pattern. This characterizes the chaotic oscillations of the system (15).
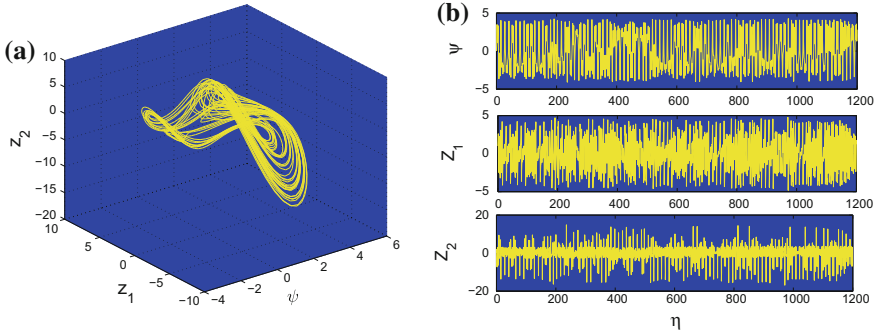
Fig. 1 **a** Three-dimensional phase portrait and **b** variation of ψ, Z_1, and Z_2 with respect to η of the system (15) for $l = 0.4$, $\delta_1 = 0.75$, $\delta_2 = 0.25$, $\lambda = 0.5$, $\sigma = 0.089333$, $\beta = 0.5$, $q = 0.45$, $v_{id0} = 0$, $\omega = 0.5$, and $f_0 = 1.0$

Fig. 2 Poincare section of the system (15) in which the parameter values are followed from Fig. 1
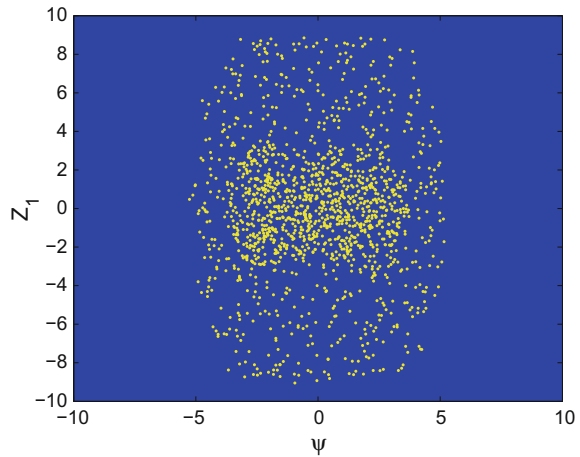
Thus, the system (15) possesses a chaotic motion in absence of the dust-ion collisional frequency (v_{id0}). On the other hand, if we take into account the influence of dust-ion collisional frequency (v_{id0}) with same value of physical parameters as in Fig. 1, then one can observe the qualitative change in the system (15).

In Fig. 3a, we present the three-dimensional phase portrait for the system (15) with $v_{id0} = 0.1$, and other physical parameters are same as in Fig. 1. Changing of ϕ, Z_1, Z_2 with respect to η is shown in Fig. 3b with same parameter values as Fig. 3a. It is important to notice that the dynamical system unveils a periodic motion.

Therefore, the dust-ion collisional frequency (v_{id0}) plays a crucial role in the transition from chaotic structure to periodic structure. This investigation may be competent to fathom the effect of v_{id0} on the features of DIAWs in comet tails and interstellar clouds.
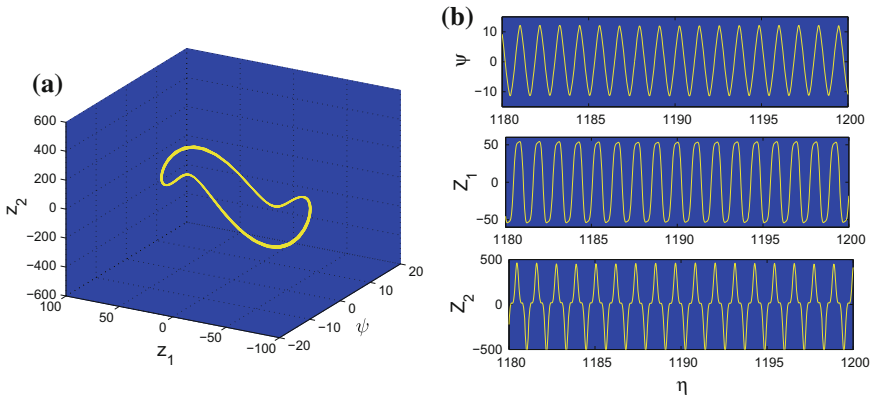
Fig. 3 **a** Three-dimensional phase portrait and **b** variation of ψ, Z_1 and Z_2 with respect to η of the system (15) with $v_{id0} = 0.1$ and other parameters are same as in Fig. 1

5 Conclusions

The chaotic and periodic behaviors of DIAWs in an unmagnetized dusty plasma with dust-ion collision effect, q-nonextensive electrons, isothermal dust grains, and ions have been investigated in the framework of the DmKdV equations in presence of an external periodic force at some critical composition. The effect of collision on the DIAWs has been discussed along with critical parameters with their critical values. As we increase the value of collisional frequency (v_{id0}), dynamical behaviors of the system corresponding to the DmKdV equation change its chaotic behaviors to periodic oscillation in the existence of externally applied force. Thus, the collisional frequency (v_{id0}) plays the role of chaos controlling parameter.

The consequence of this study may helpful to chaos control in various situations, viz., continuous time delay feedback, biological and biochemical systems, digital communication, encoding and decoding massages with chaotic lasers and electronic circuits, etc. [28–34].

References

1. Verheest, F.: Waves in Dusty Space Plasmas. Kluwer Academic, Dordrecht (2000)
2. Shukla, P.K., Mamun, A.A.: Introduction to Dusty Plasma Physics. IOP, Bristol (2002)
3. Barkan, A., D'Angelo, N., Merlino, R.L.: Experiments on ion-acoustic waves in dusty plasmas. Planet Space Sci. **44**, 239 (1996)
4. Wang, X., Bhattacharjee, A., Gou, S.K., Goree, J.: Ionization instabilities and resonant acoustic modes. Phys. Plasmas **8**, 5018 (2001)
5. Mendis, D.A.: Progress in the study of dusty plasmas. Plasma Sources Sci. Technol. **11**, A219 (2002)

6. El-Labany, S.K., Shalaby, M., El-Shamy, E.F., El-Sherif, L.S.: Effects of two-temperature electrons, external oblique magnetic field, concentration of charged dust grains and higher-order nonlinearity on dust ion-acoustic solitary waves in Saturn's E-ring. Planet Space Sci. **57**, 1246 (2009)
7. Shukla, P.K., Silin, V.P.: Dust ion-acoustic wave. Phys. Scr. **45**, 508 (1992)
8. Barkan, A., Merlino, R.L., D'Angelo, N.: Laboratory observation of the dustacoustic wave mode. Phys. Plasmas **2**, 3563 (1995)
9. Hayashi, Y., Tachibana, K.: Observation of Coulomb-crystal formation from carbon particles grown in a methane plasma. Jpn. J. Appl. Phys. **33**, L804 (1994)
10. Thomas, H., Morfill, G.E., Demmel, V., Goree, J., Feuerbacher, B., Mhlmann, M.: Plasma crystal: Coulomb crystallization in a dusty plasma. Phys. Rev. Lett. **73**, 652 (1994)
11. Shukla, P.K., Mamun, A.A.: Dust-acoustic shocks in a strongly coupled dusty plasma. IEEE Trans. Plasma Sci. **29**, 221 (2001)
12. Mamun, A.A., Eliasson, B.P.: Dust-acoustic solitary and shock waves in a strongly coupled liquid state dusty plasma with a vortex-like ion distribution. Phys. Lett. A **332**, 412 (2004)
13. Rahaman, M.S., Shikha, B., Mamun, A.A.: Time-dependent non-planar dust-acoustic solitary and shock waves in strongly coupled adiabatic dusty plasma. J. Plasma Phys. **79**(3), 249–255 (2013)
14. Garai, S., Banerjee, B., Janaki, M.S., Chakrabarti, N.: Velocity shear effect on the longitudinal wave in a strongly coupled dusty plasma. Astrophys. Space Sci **349**, 789–798 (2014)
15. Chakrabarti, N., Ghosh, S.: Longitudinal dust acoustic solitary waves in a strongly coupled complex (dusty) plasma. J. Plasma Phys. **81**, 905810310 (2015)
16. Ghosh, S., Sarkar, S., Khan, M., Gupta, M.R.: Nonlinear properties of small amplitude dust ion acoustic solitary waves. Phys. Plasmas **7**, 3594 (2000)
17. Moslem, M.: Linear and nonlinear properties of dust-acoustic waves in collisional, magnetized dusty plasmas. Phys. Plasmas **10**, 3168 (2003)
18. Zahibo, N., Pelinovsky, E., Sergeeva, A.: Weakly damped KdV soliton dynamics with the random force. Chaos Solitons Fractals **39**, 1645–1650
19. El-Labany, S.K., Shalaby, M., El-Shamy, E.F., Khaled, M.A.: The dust ion acoustic waves propagation in collisional dusty plasmas with dust charge fluctuations: effect of ion loss and ionization. J. Plasma Phys. **77**, 95 (2010)
20. Misra, A.P., Adhikary, N.C., Shukla, P.K.: Ion-acoustic solitary waves and shocks in a collisional dusty negative-ion plasma. Phys. Rev. E **86**, 056406 (2012)
21. Maitra, S., Banerjee, G.: Dust ion acoustic solitary waves in a collisional dusty plasma with dust grains having Gaussian distribution. Phys. Plasmas **21**, 113707 (2014)
22. Samanta, U.K., Saha, A., Chatterjee, P.: Bifurcations of dust ion acoustic travelling waves in a magnetized dusty plasma with a q-nonextensive electron velocity distribution. Phys. Plasmas **20**, 022111 (2013)
23. Saha, A., Chatterjee, P.: Dust ion acoustic travelling waves in the framework of a modified Kadomtsev-Petviashvili equation in a magnetized dusty plasma with superthermal electrons. Astrophys. Space Sci. **349**, 813 (2014)
24. Saha, A.: Solitonic, periodic, quasiperiodic and chaotic structures of dust ion acoustic waves in nonextensive dusty plasmas. Eur. Phys. J. D **69**, 203 (2015)
25. Saha, A., Chatterjee, P.: Electron acoustic blow up solitary waves and periodic waves in an unmagnetized plasma with kappa distributed hot electrons. Astrophys. Space Sci. **353**, 163 (2014)
26. Saha, A., Tamang, J.: Qualitative analysis of the positron-acoustic waves in electron-positron-ion plasmas with deformed Kaniadakis distributed electrons and hot positrons. Phys. Plasmas **24**, 082101 (2017)
27. Bains, A.S., Tribeche, M., Gill, T.S.: Modulational instability of ion-acoustic waves in a plasma with aq-nonextensive electron velocity distribution. Phys. Plasmas **18**, 022108 (2011)
28. Boccaletti, S., Grebogi, C., Lai, Y.C., Mancini, H., Maza, D.: The control of chaos: theory and applications. Phys. Rep. **329**, 103–197 (2000)

29. Garfinkel, A., Spano, M., Ditto, W., Weiss, J.: Controlling cardiac chaos. Science **257**, 1230 (1992)
30. Garfinkel, A., Weiss, J., Ditto, W., Spano, M.: Chaos control of cardiac arrhythmias. Trends Cardiovasc. Med. **5**, 76 (1995)
31. Singer, J., Bau, H.H.: Active control of convection. Phys. Fluids A **3**, 2859 (1991)
32. Singer, J., Wang, Y.Z., Bau, H.H.: Controlling a chaotic system. Phys. Rev. Lett. **66**, 1123 (1991)
33. Colet, P., Roy, R.: Digital communication with synchronized chaotic lasers. Opt. Lett. **19**, 2056 (1994)
34. Alsing, P.M., Gavrielides, A., Kovanis, V., Roy, R., Thornburg, K.S.: Encoding and decoding messages with chaotic lasers. Phys. Rev. E **56**, 6302 (1997)

ICT in Social Development—Context-Sensitive Design Strategies to Develop Mobile Applications for Barefoot Animal Breeders

Divya Piplani, Dineshkumar Singh, Karthik Srinivasan, Vaibhav Lonkar and Sujit Shinde

Abstract Cattle crossbreeding improves the breed quality and milk yield of the progenies and hence improves the nutrition security and livelihood of the farmers. It is a complex data-intensive process and requires the animal breeder to maintain 7–10 forms having 230 data fields. Each data field may have 5–100 options to be filled in, which is a challenging task to be done on a small screen mobile phone by low-literate breeder. We followed an iterative design process to improve the overall user experience. We interacted with the end users, understood their environment, and analyzed the existing methods. We analyzed all the forms and data recorded by them in last 3 years. We analyzed major parameters like cattle species, pregnancy stages, cattle locations, etc., and based on the unique combination of these values, we defined various "contexts". For every context and data field, we identified "context-sensitive" vocabulary (word library) extracted from the available records. So, for the selected data field, we recommended the values based on the context, from the vocabulary. During further iterations, the user choices were also learnt, and values substituted by the users, if any, against the recommendation, were added to the vocabulary. We added autocomplete suggestion feature to make it more user-friendly. We used the word usage count to rank the suggested words with the words used maximum ordered first in the recommendation. This design approach minimizes the user data entry efforts, improves the speed, and reduces error,

D. Piplani (✉) · D. Singh · K. Srinivasan · V. Lonkar · S. Shinde
TCS-Innovation Lab-Mumbai, Tata Consultancy Services, Yantra Park, Thane, India
e-mail: divya.piplani@tcs.com

D. Singh
e-mail: dineshkumar.singh@tcs.com

K. Srinivasan
e-mail: Srinivasan.Karthik@tcs.com

V. Lonkar
e-mail: vaibhav.lonkar@tcs.com

S. Shinde
e-mail: sujit.shinde@tcs.com

especially by the low-literate animal breeders who are not quite comfortable with text typing on smaller screens.

Keywords Rural mobile app · Data management · Design · Autocomplete Context sensitive · Breed improvement · Livestock artificial insemination

1   Introduction

As per a 2016 survey by People Research on India's Consumer Economy (PRICE), 88% of the Indian households have a mobile phone. June 2016 report of Internet and Mobile Association of India (IMAI) mentions that 371 million Indians access Internet using mobile phones [1]. Though the majority of these users are from urban areas, the penetration in the rural areas is also increasing significantly. This plays an important role in bridging the digital divide in India. But the majority of the Internet usage in the rural areas is related to online music download, and just 22% of usage is via mobile apps. One of the bigger challenges being to design the app or services for the rural consumers, whose needs and profile differ drastically from the urban consumers.

Indrani Medhi carried out a detailed designed process involving ethnographic interviews, field trials, and participatory designs, involving 450 h and 400 people from India, the Philippines, and South Africa [2]. We used a similar approach and the participatory design helped in improving the UI of the application, but filling out the data-intensive forms still remained a big challenge.

Kathryn Summers conducted a 4-month research project to design web-based forms for users with lower literacy skills [3]. But the major challenge was the time taken by the users to get computer literate first before using the service. Users had to be trained on using a mouse and/or understanding what a hyperlink is. With the advent of the mobility and smartphone, this has improved and there is no need of a specific expertise for using the mobile phones, even in the villages.

Fulcher and team suggested using an automatic word recognition system for oral data entry [4]. Though this simplifies the process, it requires a good speech recognition system, which is challenging for adaptation to given rural activities or domain in local language.

Psychometric profile of user like age, education level, mobile phone literacy, and their knowledge of English language helps, while designing application for rural users. Since the majority of the users do not know English, it is good to develop the applications in local language. But depending on the nature of the app, this may not be enough. We conducted a study of 25 rural users and observed that 80% of the users struggled with the forms which required them to provide the textual input using the keyboard. Discrepancies were observed in the spelling of uncommon words such as spelling of the farmers or places, cattle breeds or operation codes.

To some extent, keyboard prediction or autocomplete feature makes this task easy [5]. Most keyboard prediction features use text that we enter overtime to build

a custom, local "dictionary" of words and phrases that we have typed repeatedly. It then "scores" those words by the probability by which we would use it again. But these predictions are common to all applications installed on the mobile phone and are not specific to any particular domain or an application. So it is possible that the user types a word, which is specific to an application but since it is not the most frequent or commonly used word, it may not be added in the prediction dictionary. Even if some of the keyboards allow adding it, it may get less weightage and may be pushed to the bottom of the recommendation list. Hence, unless almost half the number of characters of that particular word is typed, it is not displayed in the prediction list. This means we must have an autocomplete or prediction feature that is very specific to given domain-application context. In this paper, we will capture how autocomplete specific to application context helps the rural user to enter/type text inputs. We will also cover some other design practices that need to be followed while designing an application for the rural user, for example, using search option in a dropdown (spinner) and many other web or app form design tools.

2 Background

Around 70 million small and marginal farmers depend on livestock resources for their livelihood [6]. To help these farmers, many government agencies and other organizations like BAIF provide doorstep cattle breeding services through artificial insemination (AI). BAIF provides cow and buffalo breed improvement services through AI to around 60,000 villages across the country [7]. In Maharashtra, BAIF-MITTRA serves around 3800+ villages and covering 2 lakh families [8]. MITTRA appoints and trains barefoot animal breeder, known as AI Technicians (AIT) for this purpose. They are in age group of 21–65 years, with education level from primary till 10th standard. AIT visits farmers and performs AI on their cattle and records each detail related to farmer, cattle, insemination breed, and process. They also record farmer's socioeconomic details, post AI monitoring activities such as pregnancy detection (PD) and calving data. With so many activities across so many villages and farms, AIT finds it difficult to keep track of each activity and use paper-based monthly report. AIT supervisors and MITTRA officials find it difficult to monitor real-time data and issues, as they to collect the paper-based report from the AITs and digitize, which takes months. To overcome these challenges a mobile-based application mKRISHI® Livestock was designed and developed in 2013.

Digital data entry activity was continuously monitored. By implementing Goal, Operators, Method and Selection Rules (GOSM) approach of (Card et al. 1983) [9], we found that average time taken to complete any flow is considerably high and may take up to 30 min for an AIT of age 50 years and above. A focused group discussion (FGD) was organized with the AITs to understand the service flow, the number of forms being filled, seeking feedback and suggestions.

Table 1 Focused group discussion (FGD) feedback results

Major queries	Feedback (Yes %)
Are you able to use the application without any prior training?	100
Are you comfortable using the app in local language (Marathi)?	100
Do you face any difficulties while entering data in the application?	80

Participants for these sessions included experts (veterinarian doctors) from MITTRA too. They were responsible for analyzing data received from the mobile-based application. Other participants include MITTRA AITs who are the actual users of the application and our technical team who conducted this session.

Feedback form included the below questions.

- Are you able to use the application without any prior training?
- Are you comfortable using the app in local language (Marathi)?
- Do you face any difficulties while entering data in the application?
- How much average time do you take to fill any form?

After collecting feedback forms from them, we conducted open forums and encouraged them to share their practical difficulties and suggestions. It was observed that around 80% agents were facing trouble while filling the form especially with the fields requiring text input (Table 1). We sought their feedback on each form, field by field and decided to redesign our application to make it more user-friendly.

3   Design Approach

Based on the feedback we decided to build a domain-specific *contextual* word library supported by the autocomplete feature. Instead of using the keyboard provided prediction dictionary, which could be generic in nature, we decided to personalize the autocomplete word vocabulary for every form.

We analyzed forms in each service. We looked at each data field, its data type, and the range of values which needed to be typed to fill it. We decided to define a "*context*" for each data field. Hence, we divided the complete farm data into various contexts, wherein a context is a form associated with a given breeding operation or workflow. For every context, we created a word (value) library (or vocabulary). This vocabulary was initially empty. Every time when AIT entered a word which does not exist in the vocabulary, it gets added.

During the next iteration for the given context, the word would be suggested from the vocabulary corresponding to each data field and if the user ignores the suggestion and enters a new value, it would get added to the word library.

3.1 Custom Dropdown with Search Option

In order to avoid data typing, we pre-populated the default words in the dropdown. All the possible values were pre-populated and shown in drop-down like menu. The user needed to select the desired value from a set of values in the vocabulary [10]. Hence, it reduced the typing effort. We also provided search option wherein user can search through the dropdown by providing a few initial strokes. Once the user selected an option, we maintained its hit score and the next time the user provided the same keystroke in the search option, based on the hit score, the most frequently used value or the value with highest hit score is shown on top (Fig. 1).

3.2 Custom Autocomplete EditBox

For every text data entry field used in various forms, we designed custom auto-complete EditBox. This custom autocomplete EditBox was built upon Android's autocomplete TextView widgets with custom filtering logic. Autocomplete Text-View provides suggestions automatically when the user is typing; the list of suggestions is displayed in a drop-down menu from which the user can choose an item to replace the content of the edit box.

Autocomplete TextView uses a predefined list of items to be displayed, but in our case, we have kept the list dynamic. When the user enters any new word which was not there previously in the list, that particular word gets added to the list. So next time when the user is typing that newly added word will be shown in the list along with other words. In this way, a dynamic list is kept for each text field in the

Fig. 1 Search text showing the list of the farmers matching with the search criteria, ordered in the sequence of usage
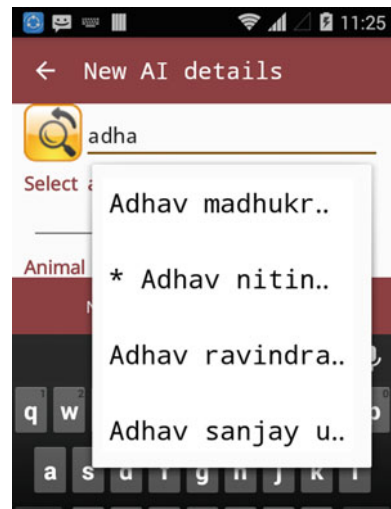
Fig. 2 Batch nos. are shown in the order of their usage in the data entry form. This used to be one of the most typing error content
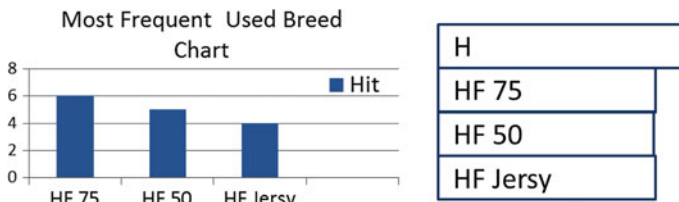
Fig. 3 Initially HF75 was the more frequently used breed with count or hit score of 6

form (Fig. 2). So suppose in a Cattle registration form if we have two text input fields to capture (i) breed name and (ii) receipt number, then separate list will be maintained for breed name and receipt number. In this way, user gets suggestions based on the selected application context.

Apart from showing context-specific dynamic list, we also maintain hit score for the item in each list. Suppose our field agent is entering data for Village A where most commonly used breed name is *HF 75* and he has to register 10 cattle with same breed type. And in our breed name list, we have 50 items stored with starting character "*H*" and 20 items stored with starting character "*HF*". When in breed name field user types character "*H*", all the 50 items will be shown in suggestion list and the user needs to scroll down to select *HF 75* breed name or he has to type next character "*F*" and again scroll down to select the desired breed. However, by maintaining hit score and displaying the list based on his hit score, the user need not
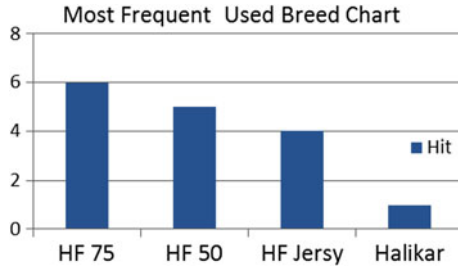
Fig. 4 A new breed Halikar is added by user and its hit score gets set as 1
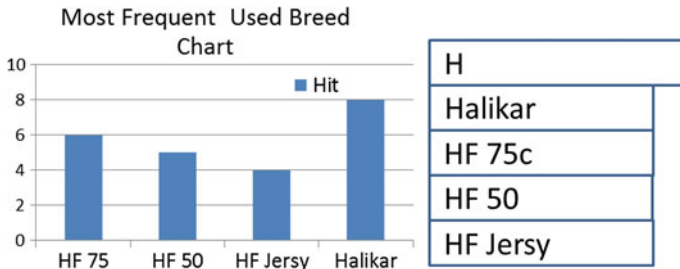
Fig. 5 Overtime more users selected Halikar, thus it got the highest score and became most frequently used breed

scroll or type the second character. The desired breed name would appear by just typing the first character of breed name (Fig. 3). When a new word is added to vocabulary, its hit score is set to one (Fig. 4). Hit score is maintained based on user selection and every time user selects the same text, the hit score kept increasing (Fig. 5).

4 Results

The entire design went through their iterations and the same approach was built for Marathi and English both languages. This provided flexibility to the AITs to use the services in their desired language and personalize the word library as per their need. A word mapping at the backend system helped to move freely from one language to other. User interface design was also customized to suit the user needs. Currently, 260 AITs across Maharashtra are using this service.

5   Conclusion

Defining a domain-specific context, based on the domain parameters and the end users skills and experiences, helps to have a much focused "context-specific" vocabulary. Hence, it personalizes this library to each data form and the words are sorted in the order of their hit score. It minimizes the effort to scroll down through a long list of suggested words. Hence, the data entry process improves. This is further enhanced by remembering the additional user choices or words, which are provided by the end users, if any, against the recommended words by the app. This helps in continuously improving the vocabulary. Autocomplete suggestion feature uses this vocabulary to minimize the data entry efforts, improves the speed, and reduces error, especially by the low-literate animal breeder who is not quite comfortable with text typing on smaller screens. This also helps in avoiding spelling mistakes where the same words may be written differently.

As an extension of this work, a master data vocabulary or library can be dynamically built by storing data values entered by other users from the different devices into data fields of the same form. Other filtering criteria such as location-based context setting can also be added. This approach has been piloted in 13 centers in Maharashtra and around 300 AITs are using this application. Periodic feedback from the user was taken, and it was found that about 90% users found this approach to be faster and user-friendly.

References

1. Internet and Mobile Association of India (IMAI) and Times Internet Limited. A Study of Smartphone Users in India (2016). http://economictimes.indiatimes.com/articleshow/55543589.cms. Accessed 31 Jan 2017
2. Medhi, I., Toyama, K.: User Interfaces for Non-literate and Semi-literate Users. Microsoft Research, India (2016)
3. Summers, K., Langford, J., Jennifer, W., Abela, C., Souza, R.: Designing web-based forms for users with lower literacy skills. Proc. Assoc. Inf. Sci. Technol. **43**(1), 1–12 (2006)
4. Fulcher, J., Gluck, R., Worthy, M., Draisma, K., Vialle, W.: A neural network, speech-based approach to literacy. In: Proceedings of the 1st International Symposium on Information and Communication Technologies, Trinity College Dublin, pp. 40–45 (2003)
5. Kim, D.: Autocompletion and automatic input method correction for partially entered search query. U.S. Patent Application 12/188,163, filed 7 Aug 2008
6. Hegde, N.G.: Livestock development for sustainable livelihood of small farmers. Europe **2682**, 4233 (2006)
7. BAIF Development Research Foundation (2016). www.baif.org.in. Accessed 31 Jan 2017
8. Maharashtra Institute of Technology Transfer for Rural Areas (MITTRA) BAIF—MITTRA (2016). www.mittra.org/. Accessed 31 Jan 2017
9. John, B.E., Kieras, D.E.: The GOMS family of user interface analysis techniques: comparison and contrast. ACM Trans. Comput. Hum. Interact. (TOCHI) **3**(4), 320–351 (1996)
10. Snapper, E.J., Jiggins, J.P., Shyam, B., Partovi, H., Berman, E.R., Freedman, S.J., Allard, J.E., Chang, F.Z., Proteau, S.P., Jorgenson, C.C.: System and method for populating forms with previously used data values. U.S. Patent 7,216,292, issued 8 May 2007

11. Sedgewick, R., Wayne, K.: Algorithms. Addison-Wesley Professional (2011)
12. Mahnke, M., Uprichard, E.: Algorithming the Algorithm. Society of the Query Reader: Reflections on Web Search, pp. 256–271. Institute of Network Cultures, Amsterdam (2014)
13. Comer, R.W., Stein, A.B., Williams, D.R., Jr.: Method and apparatus for suggesting completions for a partially entered data item based on previously-entered, associated data items. U.S. Patent 5,845,300, issued 1 Dec 1998

Multiple Solution Sorting Method Using Translocation

Pranav Kumar and G. Sahoo

Abstract The problems related to different sorting methods over a signed permutation is initiated and raised by comparative genomics. Especially, translocation in comparative genomics is often deal with large-scale mutation between species. Related to that Hannenhalli has proposed the first polynomial time algorithm in 1996 for calculating translocation distance between genome. Here our algorithm improves algorithm given by Bergeron, Stoye and Mixtacki for sorting by translocation problem. Not only that, our algorithm is a betterment of the algorithm given by Tannier et al.

Keywords Translocation · Reversal · Genome rearrangement

1 Introduction

Translocation is nothing but a block interchange of non-empty sets or ends of any two chromosomes. First, the problem of translocation was introduced by Ravi and Kececioglu. Whereas each problem has a sub-problem of translocation distance which was given by Hannenhalli and Pevzner [1] and further improved by Bergeron et al. [2]. Translocation distance is the minimum distance required to convert a sequence of species into another target sequence of species but apart from that without producing a shortest sequence or dealing a shortest distance. Li et al. gave a linear time algorithm. In 2006, Bergeron et al. deal the study of comparative genomics and revisit the problem of sorting multichromosomal genomes, and their real challenge was to propose the algorithm whose validity can be checked by both mathematically and biologically. In 2007, Flato and Shamir used the concept of

P. Kumar (✉) · G. Sahoo
Department of Computer Science and Engineering, Birla Institute of Technology,
Mesra, Ranchi 835215, India
e-mail: Pranav5503.06@bitmesra.ac.in

G. Sahoo
e-mail: gsahoo@bitmesra.ac.in

translocation but having recursive algorithm. Pevzner has shown the relation between translocation and reversal.

In reversal [3], a section of gene sequence from unichromosome is selected and reinserted onto the same position but with sign interchanged, as permutation of genome can be signed or unsigned. If two genome sequences are given π_1 and π_2 transforming shortest sequence of reversal from Ψ_1 into Ψ_2.

In our paper, we are proposing a doubly concept of forming cycle graph not only that we are focusing on generalizing the overlap graph. We have also shown how the same concept of sorting by reversal can be used and extract the chromosomal information to develop a new algorithm, i.e. sorting by translocation, since translocation involves the multichromosome whereas reversal requires only unichromosome notation of overlap graph in a better way.

2   Definitions and Notifications

2.1   Genome

A gene is a different sequence of nucleotides forming a part of chromosome or chromosome is a set of genes whereas different chromosomes combine to form a genome. Example π represents a genome or permutation containing two chromosomes having seven genes:

$$\Pi = \{(-3, 2, 1, -4)(6, -5, 7)\}$$

And for $\pi = (x_1, x_2, \ldots x_n)$ its reverse can be written as:

$$-\Pi = (-x_n, \ldots -x_2, -x_1)$$

As a chromosome does not have orientation, so we can flip the permutation as $-\Pi$. So any two chromosomes are considered to be identical if $\Pi = \Pi$ or $\Pi = -\Pi$. Here we assume that each gene appears exactly once in a genome as each repeated cytosine, guanine, thymine and adenine have given a separate identity. There are two ways to perform translocation depending upon the target genome; these are prefix–prefix [4] and the other one is prefix–suffix. In prefix–suffix as the name specifies, some genes from the starting point of first chromosome are interchanged with the some genes from the ending point of second chromosome, as shown below:

$$\{(X_1, X_2), (Y_1, Y_2)\} \text{ can be written as } \{(Y_2, X_2), (Y_1, X_1)\}$$

Whereas in case of prefix–prefix, some genes from the starting point of both the chromosomes are interchanged as follows:

$\{(X_1, X_2), (Y_1, Y_2)\}$ can be written as $\{(Y_1, X_2), (X_1, Y_2)\}$

So, sorting by translocation problem is the shortest sequence of translocation that transforms one given genome into other or that can convert source genome into target genome and the shortest sequence means within minimum number of steps one can be transformed into other that can be find out using translocation distance problem.

For any chromosome, let $\Pi = (x_1, x_2, x_3, \ldots x_n)$ can be called as Tails(Π) [5], if Tails$(\Pi) = (x_1, -x_n)$. Here all the in between elements are being eliminated whereas starting and ending points of elements are remained as it is of all the chromosomes but the sign changes to the suffix elements of all the chromosome. For example, Tails $\{(1, 2, 3, 4) (5, 6)\}$ can be written as $\{(1, -4, 5, -6)\}$. Here 2, 3 are being eliminated whereas sign of 4 and 6 is changed.

2.2 Cycle Graph

Cycle graph can be formed using three methods, first one in which the signed permutation is shown by positive and negative signs. Steps involved in it are as follows:

- For each element i of the permutation add two vertices $(-i, +i)$ if $(i > 0)$ otherwise $(+i, -i)$ in that order.
- Augment vertices 0 to the left and $-(n + 1)$ to the right.
- Add black edges between adjacent vertices.
- Add grey edges (arc) between vertices $+i$ and $-(i + 1)$.

This gives diverging/converging edges shown by forward or backward arrow, where $0 \leq i \leq n$ (Fig. 1) and the second method in which signed permutation was shown by zero and one symbol at the tail. According to them the rules as follows:

- Set of vertices are shown by $\sum_{i=1}^{n} \{i^0, i^1\}$.
- If grey edge connected outside the chromosome, then

Fig. 1 Construction of cycle graph using signed format
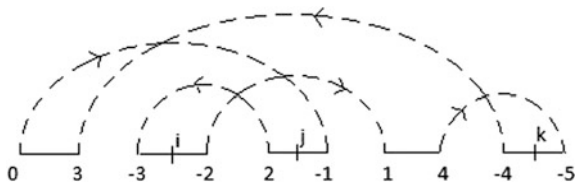
Fig. 2 Circular cycle graph
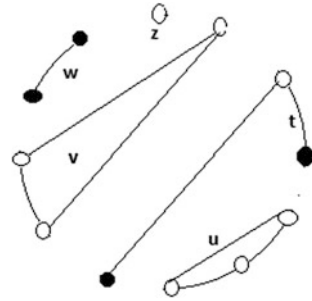
$$Out\{i\} = \begin{vmatrix} i^1 & \text{if i has positive sign} \\ i^0 & \text{otherwise} \end{vmatrix}$$

- If grey edge connected within the chromosome, then

$$In\{j\} = \begin{vmatrix} j^1 & \text{if j has positive sign} \\ j^0 & \text{otherwise} \end{vmatrix}$$

Whereas the third method is given by us in which according to rule these are the following points as follows (Fig. 2):

- In which even if the element i is greater than zero or less than zero, same rule exist.
- For each element i, add two vertices $(2i - 1, 2i)$ in the same order they exist.
- Ignore the leftmost element and rightmost element of each chromosome.
- Start pairing black edges with the adjacent vertices.

Connect grey edges between vertices as in increasing orders (Fig. 3).

Lemma 1 If two cut points are selected on black edge and both on two different cycles, then $\Delta c = -1$ (two cycles are merged) and if on the same cycle, then either $\Delta c = 1$ or $\Delta c = 0$ means that cycle is split into two cycles or there is no change.

Fig. 3 Construct by our doubly concept method
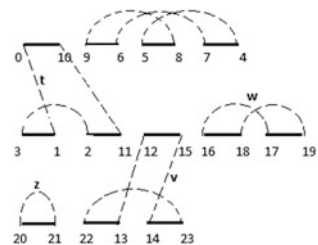
3   Proposed Algorithm

_____

Algorithm 1
_____

1. Loop continues until our source genome converted to target genome
2. If proper translocation exist in permutation Ψ
3. Select each cut points containing grey edge then perform valid proper translocation
4. If no translocation point exist then add a valid bad translocation
5. Return

_____

_____

Algorithm 2
_____

Begin
  d ← {It is the distance on which whole algorithm is dependent, as we are talking about the shortest path to reach to our target genome G(A)}
  t ← { This is the value generated from algorithm 1 as putting each translocation point one by one to form new permutation}
  β ← { *is an (i-1) sequence}*
  Ψ ← {permutation}
  $\delta_{(c1, c2)}$ ← {c1 and c2 be the 2 cut points selected between the chromosomes to interchange all the elements afterword's}
  α ← Append 't'   *[all possible translocation points generated form algorithm 1 are stored in variable α]*
  for each translocation point loop continues till 'd' **do**
          α' ← initialize to zero as it will contain further sequences
          for each β in α **do**
              Ψ' ← Ψ ∘ β *[Here we are applying each translocation points stored in β to our permutation and after resolving grey edges on cycle graph new permutation is generated as Ψ']*
              t ← { This is the new t for optimal sequence of Ψ' *[algorithm 1]*
              ∑ ← t
              for each $\rho_{(e1, e2)}$ ∈ ∑ **do**
                  β' ← β . $\delta_{(c1, c2)}$ *[start concatenating each generated points to form a sequence]*
                  put β' in α'
              end for
          end for
          α ← α'
  end for
  return α   *[α be the final collection of sequence]*
  end
_____

4   Result and Discussion

4.1   Discussion

Example:

$$\Psi = [A: (1\ 3\ 9) \qquad\qquad \Psi' = [A: (1\ 2\ 3\ 4\ 5\ 6)$$
$$B: (7\ 8\ 4\ 5\ 6) \qquad\qquad\quad B: (7\ 8\ 9)$$
$$C: (10\ 2\ 11\ 12\ 13)] \qquad\quad C: (10\ 11\ 12\ 13)]$$

As source genome is given, we have to convert this source genome into the target genome. For converting, first of all, we have to construct the cycle graph of given permutation (Fig. 4).

As in above graph, each element i of the permutation add two vertices (−i, +i) if (i > 0) otherwise (+i, −i) in that order and then as mentioned above in rules ignoring the leftmost and rightmost elements of each chromosome start pairing black edge to the neighbouring vertices. Now connect grey edges from +i to −(i + 1) as shown above with dotted lines.

As in graph, there are two grey edges exist that are external, i.e. connected between different chromosomes. So one by one we have to select two cut points onto the black edge containing external grey edge. Red bar shows that cut points selected are on black edge between 1 and −3 whereas other one is between 2 and −11. So here, we are applying translocation on (1, 2) to form a new permutation.

$$\Psi' = \ [A: (1\ 11\ 12\ 13)$$
$$B: (7\ 8\ 4\ 5\ 6)$$
$$C: (10\ 2\ 3\ 9)$$

Now again, we have to construct a cycle graph with this new generated permutation (Fig. 5).

So again, two cut points selected are on black edge between 1 and −11, and the other one is between 10 and −2.

So here, we are applying translocation on (1, 10) to form a new permutation.

Fig. 4 Specify initial cycle graph
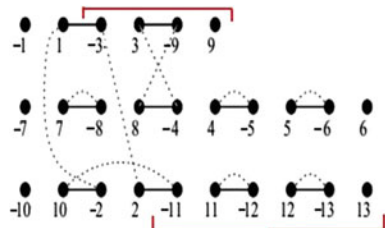
Fig. 5 After applying 1 translocation new graph formed
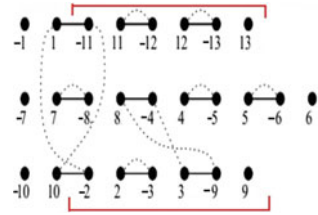
Fig. 6 Further stage cycle graph
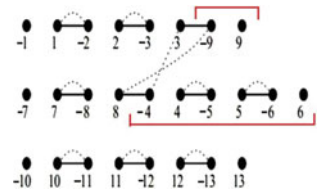
$$\Psi' = \quad [\text{A:} \,(1\ 2\ 3\ 9)$$
$$\text{B:} \,(7\ 8\ 4\ 5\ 6)$$
$$\text{C:} \,(10\ 11\ 12\ 13)$$

Now again, we have to construct a cycle graph with this new generated permutation (Fig. 6).

So again, for the last time, two cut points selected are on black edge between 3 and −9, and the other one is between 8 and −4. So here, we are applying translocation on (3, 8) to form a new permutation.

$$\Psi' = \quad [\text{A:} \,(1\ 2\ 3\ 4\ 5\ 6)$$
$$\text{B:} \,(7\ 8\ 9)$$
$$\text{C:} \,(10\ 11\ 12\ 13)$$

This is final target genome sequence and shows that final stage permutation is equal to our target genome (Fig. 7).

Fig. 7 Target genome
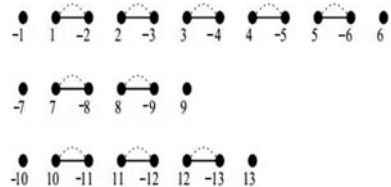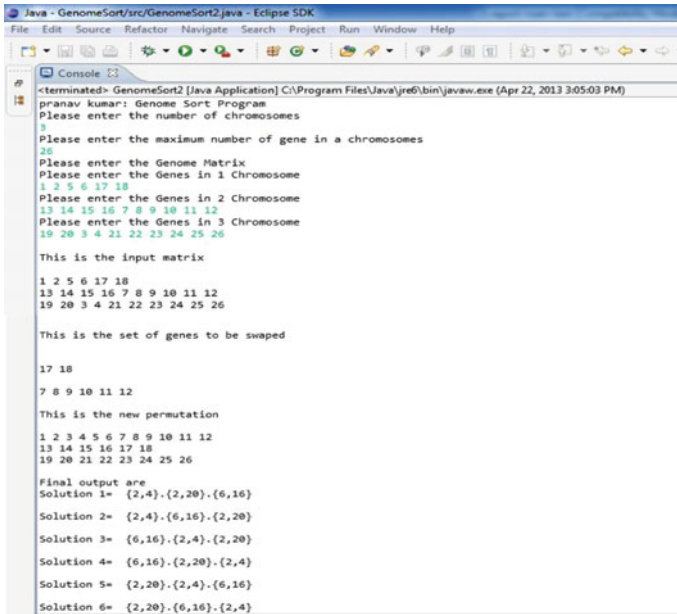
Fig. 8 Implementation of our proposed algorithm

4.2 Result

As in the above example, there can be multiple cut points in different orders; here, we have shown only one but there can be many more. So by applying Algorithm 2, we get different solutions as listed below (Fig. 8):

Solution 1.	X: {1, 2}; Y: {1, 10}; Z: {3, 8}.
Solution 2.	Y: {1, 10}; Z: {3, 8}; X: {1, 2}.
Solution 3.	Z: {3, 8}; X: {1, 2}; Y: {1, 10}.
Solution 4.	Y: {1, 10}; X: {1, 2}; Z: {3, 8}.
Solution 5.	X: {1, 2}; Z: {3, 8}; Y: {1, 10}.
Solution 6.	Z: {3, 8}; Y: {1, 10}; X: {1, 2}.

5 Conclusion

As specified in the paper, for any problem, there can be many more solutions possible because until now, everyone is showing single solution as they have that single solution only. If a person has only one option, he/she may think this is the best option but they may be wrong. That is why we are providing all solution, among which we can select the optimal one.

References

1. Hannenhali, S., Pevzner P.A.: Transforming cabbage into turnip. In: 27th ACM-SIAM Symposium on Theory of Computing, pp. 178–189, Las Vegas, USA (1995)
2. Bergeron, A., Mixtacki, J., Stoye, J.: On sorting by translocations. J. Computat. Biol. **13**, 567–578 (2006)
3. Xiao, Y., Daming, Z.: Sorting genomes by reversals and translocations. In: Asia-Pacific Conference on Information Processing, pp. 391–394 (2009)
4. Braga, M.D.V., Sagot, M., Scornavacca, C., Tannier, E.: Exploring the solution space of sorting by reversals, with experiments and an application to evolution. IEEE/ACM Trans. Computat. Biol. **5**(3), 348–356 (2008)
5. Bafna, Vineet, Pevzner, Pavel A.: Sorting by Transpositions. Soc. Indust. Appl. Math. **11**(2), 224–240 (1998)
6. Braga, M.D.V.: Baobabluna, the solution space of sorting by reversals, Bioinformatics **25**(14) (2009)
7. Amritanjali, Sahoo, G.: Exploring the solution space of sorting by translocations. Procedia Comput. Sci. **11**, 160–168 (2012)
8. Hannenhalli, S.: Polynomial algorithm for computing translocation distance between genomes. Discr. Appl. Math. **71**, 137–151 (1996)
9. Lin, Y.C., Lu, C.L., Chang, H.U., Tang, C.Y.: An efficient algorithm for sorting by block-interchanges and its application to the evolution of vibrio species. J. Comput. Biol. **12**(1), 102–112 (2005)
10. Ozery-Flato, M., Shamir, R.: Sorting by reciprocal translocations via reversals theory. J. Computat. Biol. **14**, 408–422 (2007)
11. Wang, L.S., Zhu, D.M., Liu, X.W., Ma, S.H.: An $O(n^2)$ algorithm for signed translocation. J. Comput. Syst. Sci. **70**, 284–299 (2005)
12. Diekert, V., Rozenberg, G.: The book of Traces. World Scientific, Singapore (1995)
13. Isaac, E., Tzvika, H.: A 1.5-Approximation algorithm for sorting by transpositions and transreversals. Jonassen, I., Kim, J. (Eds.): WABI 2004, LNBI 3240, pp. 50–61 (2004)
14. Isaac, E., Tzvika, H.: A1.375-Approximation algorithm for sorting by transpositions. Casadio, R., Myers, G. (Eds.): WABI 2005, LNBI 3692, pp. 204–215 (2005)

Chaos Control in a Two Prey and One Predator System with Predator Switching

Saheb Pal, Mainul Hossain, Sudip Samanta and Nikhil Pal

Abstract Vance (Am. Nat. 112:797–813, 1978) modeled a food web consisting of two prey and one predator with competing prey species. Gilpin (Am. Nat. 113:306–308, 1979) explored that for biological feasible parameter values, Vance's model can produce chaotic dynamics. In the present paper, we consider a modification of Vance's model by incorporating predator switching into the model for the same set of parameter values considered by Gilpin. We observe that if we increase switching intensity above a threshold value, then the system becomes stable from chaotic oscillations and enhances the persistence of the system.

Keywords Chaos · Predator switching · Bifurcation · Stability

1 Introduction

The discovery of chaotic dynamics in the deterministic nonlinear system motivated the researchers to investigate simple mathematical models and explore the possibilities of chaos and its control mechanisms. Chaos is observed in various branches of research field like, economics, electronics, physics, ecology, climate and weather prediction, etc. Mathematical models on predator–prey, plant–herbivore, plant–pest, resource–prey–predator, nutrient–phytoplankton–zooplankton, and many other systems have played important roles in understanding ecological processes. In continuous time dynamics, Lotka and Volterra proposed a model with one prey and one predator and explained the cycles in biological populations. Subsequently, the

S. Pal · M. Hossain · N. Pal (✉)
Department of Mathematics, Visva-Bharati, Santiniketan, India
e-mail: nikhilpal.math@gmail.com

S. Samanta
Department of Mathematics, King Abdulaziz University, Rabigh, Saudi Arabia
e-mail: samanta.sudip.09@gmail.com

Lotka–Volterra model was modified and extensively studied by many researchers. Gilpin [1] showed that chaos could occur in a model with one predator and two competing prey species [2]. Further studies on the same model by Schaffer [3] have also confirmed that the dynamics are chaotic. However, chaos is rare in real food chain systems. Here we are mentioning few ecological phenomena, e.g., omnivory, toxic inhibition, predator switching, Allee effect, and migration by which chaos can be controlled.

In real predator–prey environment, where several preys are present, the predator prefers to feed preys in a habitat for some duration and then changes its preference to another habitat. Predator switching is a biological phenomenon, where the predator population likes to graze the most available prey in a habitat when other prey population is rare [4]. In nature, many authors observed predator switching phenomenon [4]. Fisher-Piette observed that intertidal snail switches its feeding strategy depending upon the availability of barnacles or mussels [5]. Lawton et al. experimentally showed predator feeding switching in Notonecta and Ischnura [6]. This situation also has been analyzed by mathematical models. The classical Tansky model [7] is essentially based on the relative abundance of two prey populations.

Our aim is to study the impact of predator switching on the dynamics of a three species model with two prey and one predator populations. To attain our goal, we modify the Vance's model [2] by incorporating predator feeding switching into the model and studied the model for the same set of parameter values considered by Gilpin [1].

2   Model

In this section, we modify a two prey and one predator model proposed by Vance [2] by incorporating predator switching behavior. The modified system of equations can be written as follows:

$$
\begin{aligned}
\frac{dx}{dt} &= b_1 x - a_{11} x^2 - a_{12} xy - \frac{a_{13} xz}{1 + c\frac{y}{x}}, \\
\frac{dy}{dt} &= b_2 y - a_{22} y^2 - a_{21} xy - \frac{a_{23} yz}{1 + c\frac{x}{y}}, \\
\frac{dz}{dt} &= -b_3 z + \frac{a_{31} xz}{1 + c\frac{y}{x}} + \frac{a_{32} yz}{1 + c\frac{x}{y}}.
\end{aligned}
\tag{1}
$$

x, y, and z, respectively, denote the prey-I, prey-II, and predator population sizes. The meaning of the parameters, all non-negative, is as follows. b_1 and b_2 are the growth rate of the resources. a_{11} and a_{22} denote intraspecific competition of resources. a_{12} and a_{21} are the competition between the resources. The parameters a_{13} and a_{23} denote the prey consumption rates, while a_{31} and a_{32} represent the predators' assimilation rates. The natural mortality of the predator is b_3. c is the switching intensity.

The parameter c provides continuity of the switching phenomenon [8, 9]. For $c = 0$ (no switching), the system becomes exactly the Vance's two prey and one predator system [2], while for $c = 1$, the system shows perfect switching phenomenon [7].

3 Stability Analysis

The above model is not defined at the origin. In ecological systems, the coexistence equilibrium point always important to infer the dynamical properties of the system. At a coexisting equilibrium point all species exist with positive densities. In this section, our main aim is to study the stability behavior of the coexisting steady state of the system (1). We evaluate the interior equilibrium point $E^*(x^*, y^*, z^*)$, where

$$x^* = \frac{b_3}{\frac{a_{31}u}{c+u} + \frac{a_{32}}{u(1+cu)}}, \quad y^* = \frac{b_3}{\frac{a_{31}u^2}{c+u} + \frac{a_{32}}{(1+cu)}},$$
$$z^* = \frac{c+u}{a_{13}u}(b_1 - a_{11}x^* - a_{12}y^*) = \frac{1+cu}{a_{23}}(b_2 - a_{22}y^* - a_{21}x^*), \quad u = \frac{x^*}{y^*}, \tag{2}$$

where u is the positive root of the polynomial

$$C_1 u^5 + C_2 u^4 + C_3 u^3 + C_4 u^2 + C_5 u + C_6 = 0, \tag{3}$$

with

$C_1 = a_{13}a_{21}b_3 c^2 - a_{13}a_{31}b_2 c^2,$

$C_2 = a_{13}a_{21}b_3 c(2 + c^2) - 2a_{13}a_{31}b_2 c + a_{31}a_{23}b_1 c - a_{11}a_{23}b_3 c + a_{13}a_{22}b_3 c^2 - a_{13}a_{32}b_2(1 + c^2),$

$C_3 = a_{13}a_{21}b_3(1 + 2c^2) - a_{13}a_{31}b_2 - a_{11}a_{23}b_3(1 + 2c^2) - a_{23}a_{31}b_1(1 + c^2)$
$\quad - a_{12}a_{23}b_3 c - 2a_{13}a_{32}b_2 c + a_{13}a_{22}b_3(2 + c^2),$

$C_4 = a_{13}a_{21}b_3 c + a_{13}a_{22}b_3(2 + c^2) + a_{23}a_{32}b_1 - a_{11}a_{23}b_3 c(2 + c^2)$
$\quad - a_{12}a_{23}b_3(1 + 2c^2) + a_{31}a_{23}b_1 c - a_{13}a_{32}b_2(1 + c^2),$

$C_5 = a_{13}a_{22}b_3 c + 2a_{23}a_{32}b_1 c - a_{11}a_{23}b_3 c^2 - a_{12}a_{23}b_3 c(2 + c^2) - a_{13}a_{32}b_2 c,$

$C_6 = a_{23}c^2(a_{32}b_1 - a_{12}b_3).$

By Descarte's rule of sign, the Eq. (3) has a positive real root when $C_1 C_6 < 0$. Hence, the equilibrium point $E^*(x^*, y^*, z^*)$ will be feasible if $x^* < min\{\frac{b_1 u}{a_{11}u + a_{12}}, \frac{b_2 u}{a_{22}u + a_{21}}\}$ and $C_1 C_6 < 0$.

Now, the Jacobian matrix of the system (1) at the interior equilibrium $E^*(x^*, y^*, z^*)$ is

$$J = \begin{pmatrix} W_1 & W_2 & W_3 \\ W_4 & W_5 & W_6 \\ W_7 & W_8 & 0 \end{pmatrix},$$

with

$$W_1 = -a_{11}x^* - \frac{a_{13}cx^*y^*z^*}{(x^* + cy^*)^2}, \ W_2 = -a_{12}x^* + \frac{a_{13}cx^{*2}z^*}{(x^* + cy^*)^2}, \ W_3 = -\frac{a_{13}x^{*2}}{(x^* + cy^*)},$$

$$W_4 = -a_{21}y^* + \frac{a_{23}cy^{*2}z^*}{(y^* + cx^*)^2}, \ W_5 = -a_{22}y^* - \frac{a_{23}cx^*y^*z^*}{(y^* + cx^*)^2}, \ W_6 = -\frac{a_{23}y^{*2}}{(y^* + cx^*)},$$

$$W_7 = \frac{a_{31}x^{*2}z^* + 2a_{31}cx^*y^*z^*}{(x^* + cy^*)^2} - \frac{a_{32}cy^{*2}z^*}{(y^* + cx^*)^2}, \quad \text{and} \quad W_8 = \frac{a_{32}y^{*2}z^* + 2a_{32}cx^*y^*z^*}{(y^* + cx^*)^2}$$

$$-\frac{a_{31}cx^{*2}z^*}{(x^* + cy^*)^2}.$$

The characteristic equation of J at E^* is $\lambda^3 + \sigma_1\lambda^2 + \sigma_2\lambda + \sigma_3 = 0$,
where $\sigma_1 = -(W_1 + W_5)$, $\sigma_2 = W_1W_5 - W_2W_4 - W_3W_7 - W_6W_8$, and $\sigma_3 = W_1W_6W_8 - W_2W_6W_7 - W_3W_4W_8 + W_3W_5W_7$.
Hence, the system (1) around the coexisting equilibrium point will be stable if $\sigma_1 > 0$, $\sigma_3 > 0$ and $\sigma_1\sigma_2 - \sigma_3 > 0$. Also, a Hopf-bifurcation may occur if the following condition holds:

$$\sigma_1\sigma_2 - \sigma_3 = 0.$$

4 Numerical Simulations

In numerical simulations, we consider the following parameter values taken from [1]:

$$b_1 = 1, \quad b_2 = 1, \quad b_3 = 1, \quad a_{11} = 0.001, \quad a_{12} = 0.001, \quad a_{13} = 0.01, \quad (4)$$
$$a_{21} = 0.0015, \quad a_{22} = 0.001, \quad a_{23} = 0.001, \quad a_{31} = 0.005, \quad a_{32} = 0.0005,$$

with initial condition $(x(0) = 0.1, y(0) = 0.1, z(0) = 0.3)$.

For the above set of parameter values (4) with $c = 0$, the system (1) shows chaotic behavior (see, Fig. 1). Our main objective here is to investigate the impact of predator feeding switching on the system's behavior. First, we draw a bifurcation diagram for the y-population of the system (1) around the coexistence equilibrium. The bifurcation diagram (see, Fig. 2) shows that, if the switching intensity increases gradually the system (1) becomes stable from chaotic dynamics through period-halving Hopf-bifurcation. The model (1) exhibits higher periodic and chaotic oscillations for $0 < c < 0.01$, period doubling oscillations for $0.01 < c < 0.016$, limit cycle oscillations for $0.016 < c < 0.1$, and for $c > 0.1$ the system shows stable dynamics. For the parameter choice (4) with $c = 1$, the system (1) has a unique positive interior equilibrium $E^*(375.085, 414.098, 44.3561)$, which is validated by numerical simulation. We have also obtained $\sigma_1 = 0.9109 > 0$, $\sigma_2 = 0.3076 > 0$, $\sigma_3 = 0.1402 > 0$ and $\sigma_1\sigma_2 - \sigma_3 = 0.1399 > 0$ (Routh–Hurwitz stability criterion). The eigenvalues of
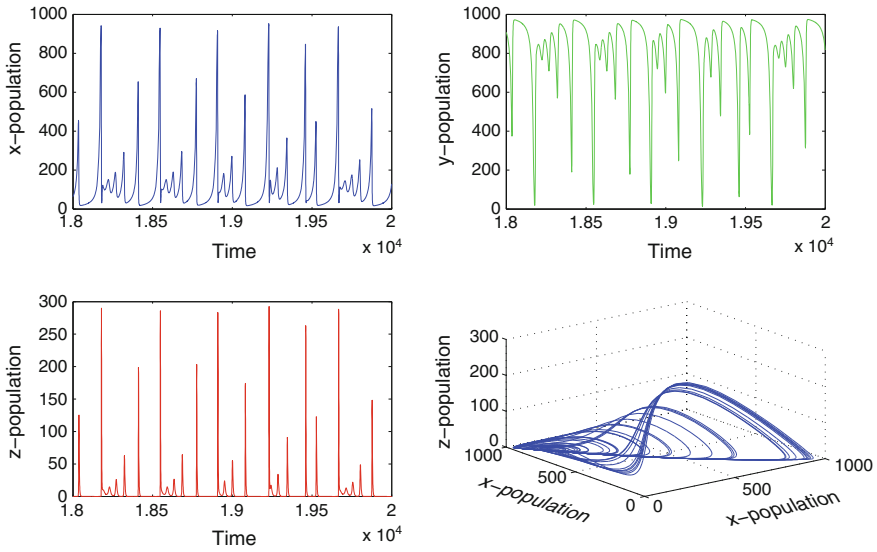
Fig. 1 The figure shows chaotic dynamics of the system (1), where $c = 0$ and the other parameter values are same as in Eq. (4)
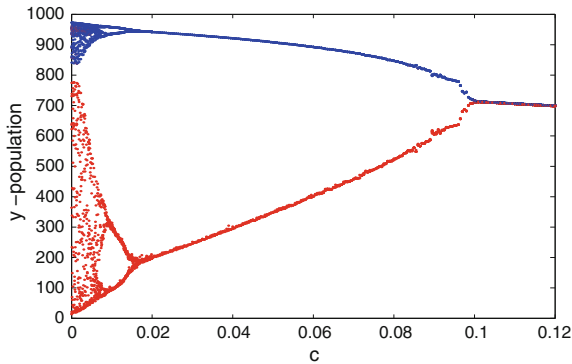
Fig. 2 Bifurcation diagram of the system (1) with respect to the bifurcating parameter c. Here we plot the maximum and minimum values of oscillations in blue and red colors, respectively

(1) are $-0.0804 \pm 0.4248i$ and -0.7500. Hence, the system (1) is stable at E^* (see, Fig. 3). Further, we verify the occurrence of Hopf-bifurcation of the system (1) by Matcont simulation. We observe that Hopf-bifurcation occurs at $c = 0.1$ and the system (1) remains stable for large values of c (see, Fig. 4). We also calculate maximum Lyapunov of the system, which shows that the system becomes stable from chaotic dynamics (see, Fig. 5).
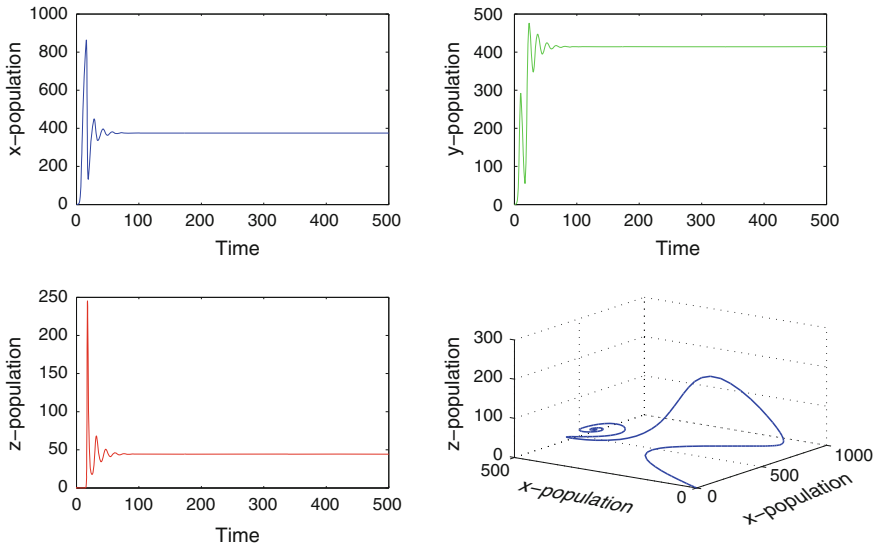
Fig. 3 The figure shows the system (1) is stable for $c = 1$, where the other parameter values are same as in Eq. (4)

Fig. 4 The figure shows the occurrence of Hopf-bifurcation and the system (1) remains stable for large value of c, where the other parameter values are same as in Eq. (4)
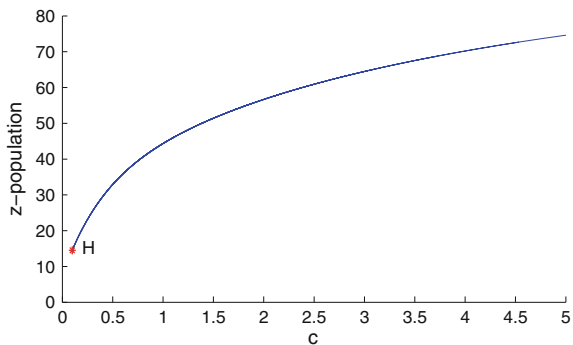
5 Conclusion

Vance [2] modeled a food web consisting of two prey and one predator with competing prey species. Gilpin [1] showed that chaos could occur in a two prey–one predator system. In a food web with two prey and one predator, predator has two types of food sources to graze, which increases the chance of predator switching between two prey. In this paper, we considered a two prey and one predator model [2] by incorporating predator feeding switching in the model system. We observe that if we increase the switching intensity, then the chaotic system becomes stable via period-halving Hopf-bifurcation (chaos–period doubling–limit cycle–stable). A biological clarification could be given as follows. When a specific prey population

Fig. 5 Maximum Lyapunov exponent of the system (1) w.r.t. the switching parameter c
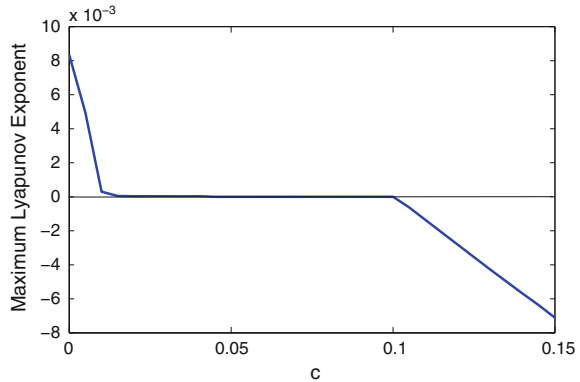
density decreases, due to predation, the predator instead of losing time in searching for the scarce resource, switches its attacks on the next most abundant available prey, and this type of switching phenomenon drives the system toward stability.

References

1. Gilpin, M.E.: Spiral chaos in a predator-prey model. Am. Nat. **113**, 306–308 (1979)
2. Vance, R.R.: Predation and Resource Partitioning in One Predator-Two Prey Model Communities. Am. Nat. **112**, 797–813 (1978)
3. Schaffer, W.M.: Order and chaos in ecological systems. Ecology **66**, 93–106 (1985)
4. Murdoch, W.: Switching in general predators: experiments on predator specificity and stability of prey populations. Ecol. Monogr. **39**, 335–354 (1969)
5. Piette. F.: Sur equilibre de faunas: Interactions des moules des poupres et des cripedaed, C. R. Soc. Biologeogr **92**, 47–48 (1934)
6. Lawton, J., Beddington, J., Bonser, R.: Switching in invertebrate predators. In: Usher, M.B., Williamson, M.H. (eds.) Ecological Stability, pp. 141–158. Chapman & Hall, London (1974)
7. Tansky, M.: Switching effects in prey-predator system. J. Theor. Biol. **70**, 263–271 (1978)
8. Pal, N., Samanta, S., Chattopadhyay, J.: Revisited Hastings and Powell model with omnivory and predator switching. Chaos Solit. Fract. **66**, 58–73 (2014)
9. Chattopadhyay, J., Pal, N., Samanta, S., Venturino, E., Khan, Q.J.A.: Chaos control via feeding switching in an omnivory system. BioSyst. **138**, 18–24 (2015)

An Improved Multi-secret Sharing Visual Cryptography Technique for Color Images Using Sterilization Algorithm

G. D. Dalvi and D. G. Wakde

Abstract Security in the recent technology world is an important issue to be taken care of and to be encountered with preventive measure on various aspects. Secret sharing is a technique used in visual cryptography (VC) which divides the image into many shares and by overlapping those shares original image is revealed, but it creates a threat when an intruder get shares with which the image is going to be decrypted easily. However, in the present project work, an extremely useful bitwise operation is performed on every pixel with the help of key. The key is provided by new concept of sterilization algorithm. Initially, Red, Green, and Blue channels get separated from image and are going to be encrypted on multiple levels using multiple shares; it converts an image into unreadable format and by combining all the shares in proper sequence the original secret image is revealed.

Keywords Bitwise operation · Multi shares · Pixel sharing
Sterilization · Algorithm · Keys · XOR · Encryption

1 Introduction

With the ascending of blended media application, there is a tremendous enthusiasm for transmission and secured securing of information. So, security is fundamental to discover good of fashioned certification. If the information is secured, the gate-crashers may not reshape the data. The best way to deal with oversees true blue and secure transmission of the data changes into a testing issue. Cryptographic frameworks bear the cost of the social occasion and security by reducing the likelihood of enemies [1]. It deals with the framework which is used to refresh the data among real and unlimited structures by using encryption and unscrambling

G. D. Dalvi (✉) · D. G. Wakde
Department of EXTC, Sant Gadge Baba Amravati University, Amravati, Maharashtra, India
e-mail: gopalddalvi09@gmail.com

D. G. Wakde
e-mail: dr_dgwakde@rediffmail.com

system under the centrality of the keys. It gives the substance security and discovers the chance to control. A champion among the most part used a sort of information sharing or question sharing is the visual puzzle sharing. Without including any astonishing figuring's, interpret the puzzle picture evidently by superimposing a qualified subset of offers through the visual befuddle sharing logic. In the emerge condition, there exists Boolean operation of the befuddle picture sharing that beats the drawback of low visual quality and pixel advancement made by the VSS [2].

VC is an imperative puzzle sharing system that finds it is interesting in connection to fundamental cryptography, for the reason that it does not require complex count to translate. Shading VC, rising field, scrambles the shading secret messages into various measures of shading halftone picture shares. A visual information pixel synchronization and goof dispersal procedure enable the encryption of visual data with high bore. Synchronization unwinding the spot of pixels close to the inquiry pictures in the midst of mess up spread produces shares satisfying to human visual system. The turmoil made by the preset a pixel was diffused by the neighbor pixels at whatever point the encryption on share happens [3]. The data covering is the embedding system which the inquiry shares are disguised using some kind of strategies. The mix of the VC with the watermarking framework is to widen the photo abundance and security [4]. In visual cryptography the passed on picture shares are mixed by using RSA figuring. The mix of VC with the complete gathering enters encryption shut in high security while transmitting the photo [5]. The procedure of the creative framework for checking untruthfulness is the request of a few problem pictures with the genuine target that each qualified subsets will reveal the relative daze picture basically, leaving the other riddle pictures new to the unavoidable vendors [6].

2 Literature Review

The basic concept of visual cryptography was invented by M. Naor and A. Shamir [7] in this VC technique. It contains black and white pixel only. It is for only sharing a single secret. The secret images are separated in only two shares, i.e., share 1 and share 2. When these two shares overlapped, it generates the original image. It is used for only black and white image. It is secure and easy to implement.

Nakajima and Yamaguchi [8] presented an innovative method of extended visual cryptography scheme (EVS) in which performing halftoning and encryption simultaneously because of that this method improves the output image quality and gives a better result as compare to previous technology regarding quality of image and security.

P.V. Chavan et al. [9] introduced hierarchical visual cryptography (HVC) for fingerprint-based authentication mechanism in this system the signature of person taken as an input and encrypted by using HVC and created four resultant shares from these four shares any three shares are taken to generate the key share.

It overcomes the problem of repetitiveness during authentication and non-acceptability.

Yadav and Vijay [10] proposed a visual cryptography scheme that encodes T >=2 (T = no of secrets) secrets into two rectangular shares in such way that single share unable to create creates the secret. In this scheme, only one share is required to store with password. In this system, communication and storage overhead are reduced, and it is efficient and more secure multiple secret VC schemes.

Linju and Mathews [11] presented simple block replacement method (SBR) and halftoned images are preprocessed using SBR technique. In this scheme, two secret color images are used, and these are divided into three shares and secret images should be changed into its halftone representation. These systems increase the quality of recovered and improve the other parameter like PSNR, MES, accuracy, and pixel expansion

3 Proposed Work

In this research work, a new concept of sharing the color image at multiple levels has given which provided more security to the encryption. Encryptions perform by separating Red, Green, and Blue channels, and then sterilization algorithm is used. It provides keys which are used to encrypt every component of a pixel. Each level consist database of particular number of shares, by using that database image is encrypted or decrypted. For revealing the original image, all the shares are required to be superimposed using the keys. By stacking shares in proper sequence original image will be obtained. The concept is extremely secure as shares are encrypted at multiple levels using the keys without which one can never decrypt the image.

Algorithm for Encryption:-

Step 1: Select Image.
Step 2: Separate Red, Green, and Blue channels.
Step 3: Each channel is further encrypted into eight shares, i.e., R + G + B = 8 + 8 + 8 = 24 shares. This encryption is done with the key provided by sterilization process.
Step 4: In this level, eight encrypted shares of each channel make group of 3, 3, and 2 shares.
Step 5: Three shares are obtained for each channel from previous step, using this share finally encrypted R, G, and B obtained. This operation is performed at level 3.
Step 6: It checks the condition for encrypting all the 3 shares. Unless all the three shares are encrypted it process in loop.
Step 7: In this last level of encryption, all the three shares get combined and finally encrypted share was obtained. Save the encrypted image to database (Fig. 1).
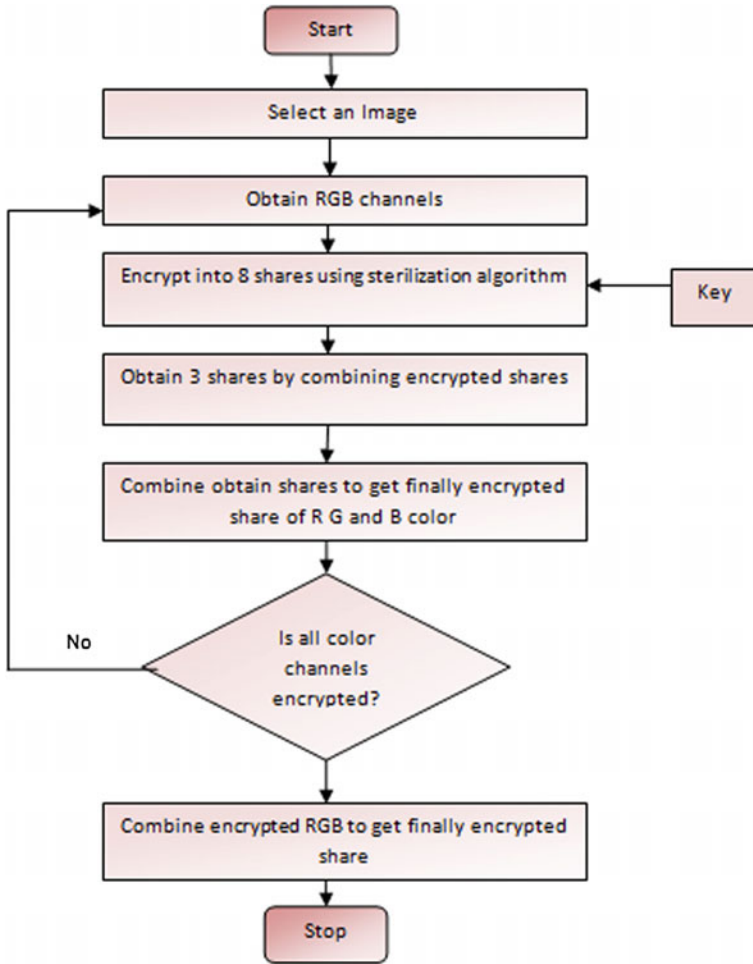
Fig. 1 Flowchart for encryption of an image

Above figure specifies the flow of control once the input has given to the system. Once all the steps of the algorithm are implemented, the last gives the encrypted share.

A. *Input Image*

Selected image should be in .bmp, .jpg, .tiff, etc., format.

B. *Extract RGB Channel*

Let function f(I) will extract Red, Green, and Blue channels of an image

$$f(I) = E(Ir, \ Ig, \ Ib)$$

C. *Preprocess images*

Image preprocessing is a step where an image can be kept ready with necessary modification before applying algorithm over it. In our proposed work, we are doing resize and image enhancement

D. *Position Keys*

Position keys are most vital part of the security of image. Proposed algorithm generates position keys based on permutation comprising digits from 0 to7.

$$P(n, r) = n!/(n - r)!$$

Where, n = Total digits, i.e., 0–8
R = Repetition of digits, i.e., 8

In our proposed method

$$
\begin{aligned}
P(n, r) &= P(8, \ 8) \\
&= 8!/(8-8)! \\
&= 8!/(8-8)! \\
&= 40320
\end{aligned}
$$

Proposed algorithm generates total 40320 keys.
Few position keys are as follows: 12345678, 12345687, and 12345768.

E. *Level 0 Encryption*

Let an image I represented with I H X W. Total pixels in an image are H X W. Total keys required for level 0 encryptions are related to image dimension. One key for one-pixel components. It means three keys for one pixel. Concerning this, an image of H X W dimension need 3 X H X W Keys

Keys		Pixels	
1	→R		
2	→G		Pixel 1
3	→B		
4	→R		
5	→G		Pixel 2
6	→B		
.		R	.
.		G	.
.		B	.
.		.	Pixel H X W
.		.	
40320			

Proposed scheme provides unique key for individual pixel color components. We are using round robin scheduling scheme to allocate keys to pixel components.

if (length(keys)> total selected keys)
  set index =1;
  Key = keys [index];
  index ++;
end;

Below figure show how round robin technique implemented for key allocation to individual pixel color components

Pixels	Keys
1	1
2	2
3	3
4	4
5	
.	
.	
H X W	

		R-8 bits
Pixel		G-8 Bits
		B-8 Bits
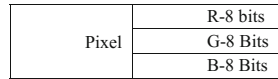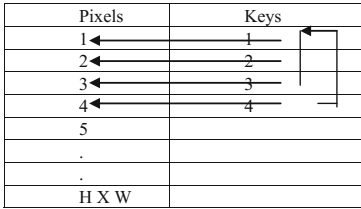
A pixel with its components

Position	8	7	6	5	4	3	2	1
Bit	1	0	1	1	0	0	0	1

8 bit of red component can be represented and position key as 32456187.
Position-based swapping/encryption work by swapping group key position bits as below:

Group	32	45	61	87

Position	8	7	6	5	4	3	2	1
Bit	1	0	1	1	0	0	0	1
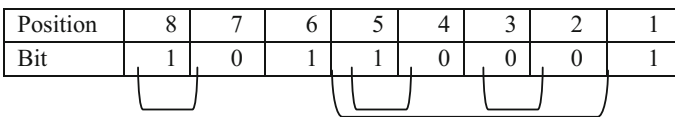
after swapping resultant key.

Position	8	7	6	5	4	3	2	1
Bit	1	0	1	1	0	0	0	1

Strong encryption can predicate with max MSE error which can be measured as

$$MSE = |originalPixel - ResultPixel|$$

In proposed scheme, max MSE is 255 and minimum is 0.

SE	0	Weak encryption
	255	Strong encryption
	−254	Moderate encryption

Other parameters for encryption standard measurement are as follows:

- PSNR,
- Average difference,
- Correlation, and
- Absolute difference, etc.

At level 0, eight different shares obtain from individual bits. We are differentiating image pixels into eight shares I level $0 = d\,I_{h\,x\,w}/d_{bitlength}$

```
if bit == 0
set pixel to black;
else
set pixel to white;
end;
```

$$EncryptedImagePixelsWeight(Eip) = \sum_{i=0}^{7} (i^2)$$

Hence at level 0, output shares are black and white images that create a level of image security.

F. *Level 1 Encryption*

At level 1, shares are integrated to creates another level of shares integration hap pens between share 1, 2, 3 4, 5, 6, 7, and 8

$$L1r = L1g = L1b = f\left\{ \int_{i=0}^{2} Ii \,,\, \int_{i=3}^{5} li \,,\, \int_{i=6}^{7} li \right\}$$

Pixel's integration at level 1 creates RGB shares that again create a security level in facial image authentication. Red component created from share 1, green from share 2, and blue from share 3

Level 1 Share Pixel	Level 0 share
Red ←	1
Green ←	2
Blue ←	3

Total shares at level 1 are 9.

G. Level 2 Encryption

At level 2, shares are integrated to creates another level of shares integration happens between share 1, 2, and 3

$$L2r = L2g = L2b = f\left\{ \int L1r \,,\, \int L1g, \int L1b \right\}$$

Pixel's integration at level 2 creates RGB shares that again enhance a security level in facial image authentication. Red component created from share 1, green from share 2, and blue from share 3

Level 1 Share Pixel	Level 1 share
Red ←	1
Green ←	2
Blue ←	3

Total shares generated at level 2 are 3.

H. *Level 3 Encryption*

It is a final stage of share generation. Shares from level two are integrated to create another level of security shares integration happens between share 1, 2, and 3

$$L3 = f\{\int(L2r, L2g, L2b)\}$$

Pixel's integration at level 3 creates RGB shares that again enhance a security level in facial image authentication. Red component created from share 1, green from share 2, and blue from share3.

Level 1 Share Pixel	Level 2 share
Red	1
Green	2
Blue	3

Total shares generated at level 2 are 3

4   Result Analysis and Discussion

A. *Performance Analysis* (Table 1)

The result shows that as compare to other techniques, the proposed work gives high PSNR value and low MSE which means it proved that decrypted image is approximate equal to the original image, and others parameters like absolute error and cross correlativity are also good and proposed work also improves the visibility (Fig. 2) (Table 2).

Table 1 Comparative table in terms of PSNR, MSE, AE, cross correlation, and visibility

Sr.No	Method	PSNR	MSE	Absolute error	Cross correlation	Visibility
1	Young-Chang	65.34	4.342	67.345	3.234	Strong
2	Sabu M Thampi	54.3523	2.2342	54.3453	5.231	Quite visible
3	Koo Kang	67.3523	5.232	54.3452	2.454	Strong
4	Jagdeep Verma	94.3453	7.2342	54.3445	6.234	Strong
5	Pallavi Chavan	87.33	1.234	54.3532	2.234	Strong
6	Proposed	99.2342	0.232	95.564	1.334	Very strong
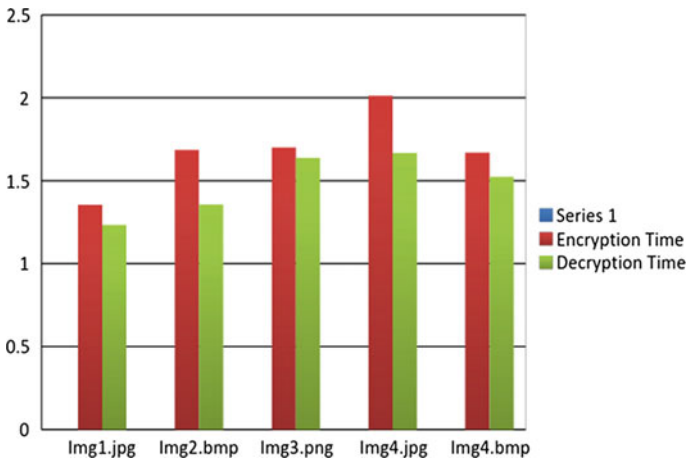
Fig. 2 Indicate the graph of encryption and decryption time

Table 2 Time required for the shares generation and reconstruction of proposed scheme

Name of input image	Encryption time (s)	Decryption time (s)
Img1.jpg	1.3554	1.2342
Img2.bmp	1.6854	1.3562
Img3.png	1.7004	1.6385
Img4.jpg	2.0124	1.6678
Img5.bmp	1.6692	1.5232

5   Conclusions

This research work proposes a novel idea of facial image authentication using sterilization algorithm in VC. In this work new, the concept of sharing the color image at multiple levels has given to provide more security to the encryption. Encryptions performed by separating Red, Green, and Blue channels, and then sterilization algorithm is used. It provides keys which are used to encrypt every component of a pixel. Each level consists of the database of a particular number of shares; by using that database image is encrypted or decrypted. For revealing the original image, all the shares are required to be superimposed using the keys. By stacking shares in proper sequence, original image will be obtained. The concept is extremely secure as shares are encrypted at multiple levels using the keys without which one can never decrypt the image.

References

1. Shamir, A.: How to share secret. Commun. ACM **22**, 612–613 (1979)
2. Nag, A., Biswas, S., Sarkar, D., Sarkar, P.P.: Secret image sharing scheme based on boolean operation. Cybermetics Inf. Technol. **14**, 98–113
3. Kang, I., Arce, G.R., Lee, H.K.: Color extended VC using error diffusion. IEEE Trans. Image Process. **20**, 132–145 (2011)
4. Jena, D., Jena, S.K.: A Novel VC scheme. In: International Conference on Advanced Computer control (2008)
5. Kaur, K., Khemchandani, V.: Securing visual cryptographic shares using public key encryption. In: IEEE 3rd International Conference on Advance Computing Conference (IACC), pp. 1108–1113 (2013)
6. Tsai, D.-S., Chen, T.-H., Horng, G.: A cheating prevention scheme for binary VC with homogeneous secret images. J. Pattern Recognit. **40**(2), 2356–2366 (2007)
7. Naor, M., Shamir, A.: Visual Cryptography. In: Proceedings of the Eurocrypt 94, Perugia, Italy, pp. 1–12. Springer, LNCS 950, May 2010
8. Nakajima, M., Yamaguchi, Y.: Extended visual cryptography for natural images. J. WSCG, v10i2 303–310 (2002)
9. Chavan, P.V.: Signature based authentication using contrast enhanced hierarchical visual cryptography. In: IEEE Students' Conference on Electrical, Electronics and Computer Science, 978-1–2526-1 (2014)
10. Yadav, M., vijay, R.: Efficient multiple secret visual cryptography via CA-Rule 30. In: IEEE International Conference on Communication and Signal Processing, 6–8 Apr 2016, India, pp. 926–930
11. Linju, P.S., Mathews, S.: An efficient interception mechanism against cheating in visual cryptography with non pixel expansion of images. Int. J. Sci. Technol. Res. **5**(01), 102–106 (2016)

A Morphological Color Image Contrast Enhancement Technique Using Hilbert 3D Space Filling Curve

Rajesh Kumar Sinha, Priyambada Subudhi and Susanta Mukhopadhyay

Abstract This paper presents a method based on mathematical morphology for enhancing the contrast of a color image using total ordering with Hilbert space filling curves. The method defines a total ordering of three-dimensional (3D) space (RGB space) using Hilbert 3D space filling curve and then applies morphological operators on the color image to obtain the contrast enhanced image. The output obtained through the above method has been compared with the outputs obtained through marginal morphology and vector morphology based on a distance measure. Experimental results show the efficiency of the proposed method in terms of enhanced contrast and better time complexity.

Keywords Color image morphology · Contrast enhancement · Hilbert curves
Multivalued morphology · Space filling curves

1 Introduction

Image contrast enhancement is an important aspect in the field of image processing and computer vision not only because it improves the visual quality of the image by making it more informative but also it acts as the preprocessing step in many higher level image processing tasks. Several contrast enhancement techniques such as contrast stretching [1], histogram equalization (HE) [2], transform-based methods using wavelet transform [3] and curvelet transform [4], and morphological methods [5] etc. have been proposed to enhance the contrast of an image. Among them, HE has been used most extensively for enhancing the contrast of both gray scale as well as color images. These HE techniques can be applied to the image both globally or locally. Recently, many improvements have been proposed to the HE method such as bi-histogram equalization [6], dynamic HE [7], adaptive HE [8], etc. Although, these improvements are capable of overcoming some of the shortcomings of traditional

R. K. Sinha · P. Subudhi (✉) · S. Mukhopadhyay
Indian Institute of Technology (Indian School of Mines), Dhanbad, India
e-mail: priyambadasubudhi@gmail.com

HE, they still may not produce the best result for all images being ad hoc approaches especially when the global or local histogram is skewed. As it is well known that mathematical morphology takes into account the spatial features in the image, so can overcome these shortcomings.

Mathematical morphology is based on the set theoretic concept of shape and here each object in the image is treated as a set [9]. In the framework of mathematical morphology, the contrast enhanced image can be obtained by combining the top-hat and bottom-hat images with the original image in a prescribed manner [5]. These morphological operations can be performed on grayscale images without any difficulty, but it is a tedious task in case of color images. Several approaches have been proposed in the literature to extend mathematical morphology for color images [10]. Those are mainly categorized as marginal approach and vectorial approach [11, 12]. In this paper, we have used total ordering vectorial approach [13] based on a space filling curve known as Hilbert curve or Peano curve [14] to map the triplet vector (R,G,B) to a single value and then obtain the contrast enhanced image by using morphological operations.

Rest of this paper is organized as follows. In Sect. 2, we have reviewed the background details required in our approach, Sect. 3 presents our proposed method, Sect. 4 includes experimental results and comparisons, and finally, we conclude our paper in Sect. 5.

2 Background

2.1 Ordering in Color Images

As per set theory, mathematical morphology requires an algebraic structure, a set T which must satisfy the following points:

- T is induced by a partial ordering relation "$\leq$".
- For any family of elements in T, there exists a smallest majorant called *sup* (for supremum) and a greatest minorant called *inf* (for infimum).

However in case of color images, both the above points are not feasible as there is no natural ordering of multivariate pixels [11]. Two main strategies have been proposed in literature for ordering multivariate pixels. Those are as below.

2.1.1 Marginal Ordering

In this approach, each component of the color image is processed separately and then combined together to give the final processed image. It does not take into account the inter-component relationship and is more expensive in terms of time as N components take N processing time units. It may also produce false color appearance while

combining the processed components which has been overcome by converting the image into HSI (hue, saturation, and intensity) components and processing only the *I* component keeping the hue and saturation constant [11].

2.1.2 Vectorial Ordering

In this approach, all the components are processed at once as a vector. It maps each multivariate pixel value (triplet pixel in case of RGB image) to a single scalar value then, ordering is done based on this value. Again, there are two techniques of vector ordering. First one is based on the computation of a distance or a cumulative distance between pixels within the filtering window (window of structuring element in mathematical morphology). This cumulative distance is taken as the scalar mapping of this pixel over that region. After this, morphological operators are applied on the image using these scalar values. The other one is to order scalar values using a space filling curve like Hilbert curve. This paper deals with finding a scalar value for each vector pixel and then ordering them using Hilbert 3D space curve.

2.2 *Hilbert Space Curve*

Hilbert Curve is a type of space filling curve which is the orthogonal form of the Peano curve and was first demonstrated by mathematician David Hilbert in the year 1891. It is a continuous curve passing through each point in an *ND* space exactly once and thus producing an 1*D* representation for the data points in the *ND* space. In its original form, it was defined for a 2*D* space and fills a square. In this paper, we have used it for the 3*D* RGB space and the curve is generated recursively as shown in Fig. 1.

(a) First iteration   (b) Second iteration   (c) Third iteration
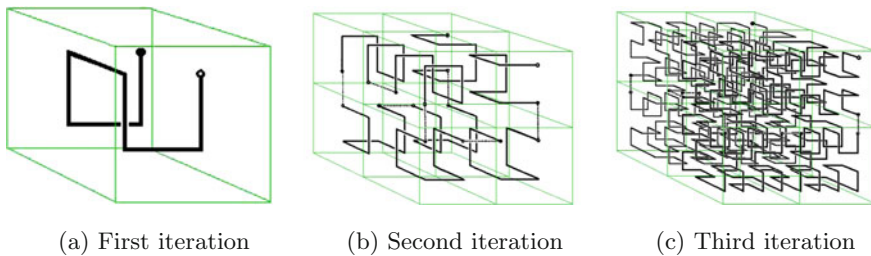
Fig. 1 Hilbert space curve
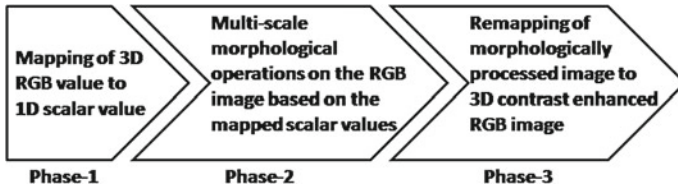
Fig. 2 Overall workflow of our proposed approach

3 Proposed Method

This section presents our proposed color contrast enhancement approach. The overall process is divided into three phases as shown in Fig. 2.

3.1 Mapping of RGB Values to Scalar Value

In this first phase of our method, mapping of the whole RGB space to 1D scalar values is done using Hilbert space filling curves. A simple example of generation of the Hilbert curve and mapping of values for a 3D space of size $2 \times 2 \times 2$ is shown in Fig. 3. It starts from the point $(0, 0, 0)$ and traverses the 3-digit binary strings in such a manner that two consecutive strings differ from each other at a single position by the addition or subtraction of 1 just like gray code strings. Again, it changes its direction each time it encounters a new point on its way. In our proposed approach, we have numbered these changing direction points in an incremental way starting from 0. From the figure, we can clearly obtain the mapping values like as shown in Table 1.

Fig. 3 Mapping of 3D space to 1D space
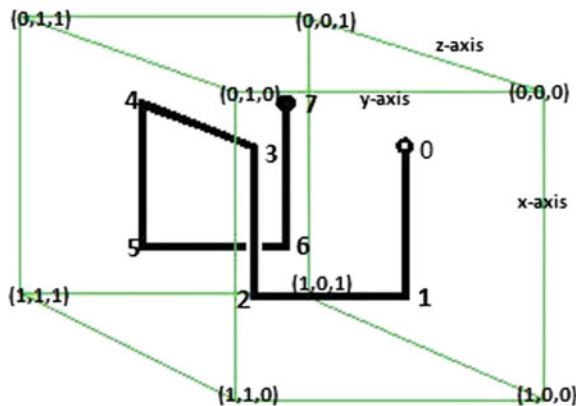
Table 1 Mapping of 3D values to scalar values for a $2 \times 2 \times 2$ cube

3D value	Mapping scalar value
0, 0, 0	0
0, 0, 1	7
0, 1, 0	3
0, 1, 1	4
1, 0, 0	1
1, 0, 1	6
1, 1, 0	2
1, 1, 1	5

It seems to be an easy task for this given cube as there are only 8 points to be mapped. However, in our case, we are interested to generate the 1D Hilbert curve for a cube of size $256 \times 256 \times 256$ (the whole RGB space). So a recursive algorithm is given in Algorithm-1 to generate the corresponding scalar values for each RGB triplet using the same strategy as explained earlier for the $2 \times 2 \times 2$ cube.

3.2 Multi-scale Morphological Contrast Enhancement

As discussed in the previous section, for each possible (r,g,b) value, there is a corresponding scalar value. Let these scalar values are stored in a 3D matrix M. Let us also consider f be an RGB image of size $N \times N \times 3$ and B be a structuring element of size $k \times k$. Suppose C is the region currently occupied by the structuring element B. Then for each triplet (r,g,b) in C, corresponding scalar values are taken from the 3D matrix M. Then erosion and dilation of the RGB image by this structuring element is performed based on the scalar values under the structuring element. In erosion, the center pixel is replaced by the RGB value having smallest mapping value among all the pixels' mapping values in C. Similarly, in case of dilation, the center pixel is replaced with the RGB value corresponding to largest mapping value in C. Those are given by

$$(f \oplus B)(x, y, z) = Arg\{max\{M(r, g, b) | (r, g, b) = f(x - j, y - k, z - l) | (j, k, l) \in B\}\} \tag{1}$$

$$(f \ominus B)(x, y, z) = Arg\{min\{M(r, g, b) | (r, g, b) = f(x - j, y - k, z - l) | (j, k, l) \in B\}\} \tag{2}$$

where $Arg(val) = (r, g, b) \ s.t. \ val = M(r, g, b)$.

Algorithm 1 Hilbert curve generation

Input: Side of cube (s), Starting point of Hilbert curve (x, y, z), three vectors for rotated x ($dx1,dy1,dz1$), y ($dx2,dy2,dz2$) and z ($dx3,dy3,dz3$) axis of the cube

Output: A 3D matrix M

 1: $s \leftarrow 256$, $(x, y, z) \leftarrow (0, 0, 0)$, $(dx1, dy1, dz1) \leftarrow (1, 0, 0)$, $(dx2, dy2, dz2) \leftarrow (0, 1, 0)$, $(dx3, dy3, dz3) \leftarrow (0, 0, 1)$
     ▷ Initial values of the inputs
 2: $m \leftarrow 0$
 3: **procedure** HILBERTCURVE($s, x, y, z, dx1, dy1, dz1, dx2, dy2, dz2, dx3, dy3, dz3$)
 4:     **if** $s == 1$ **then**
 5:         $M(x, y, z) \leftarrow m$
 6:         $m \leftarrow m + 1$
 7:     **else**
 8:         $s \leftarrow s/2$
 9:         **if** $dx1 < 0$ **then**
10:             $x \leftarrow x - s * dx1$
11:         **end if**
12:         **if** $dy1 < 0$ **then**
13:             $y \leftarrow y - s * dy1$
14:         **end if**
15:         **if** $dz1 < 0$ **then**
16:             $z \leftarrow z - s * dz1$
17:         **end if**
18:         **if** $dx2 < 0$ **then**
19:             $x \leftarrow x - s * dx2$
20:         **end if**
21:         **if** $dy2 < 0$ **then**
22:             $y \leftarrow y - s * dy2$
23:         **end if**
24:         **if** $dz2 < 0$ **then**
25:             $z \leftarrow z - s * dz2$
26:         **end if**
27:         **if** $dx3 < 0$ **then**
28:             $x \leftarrow x - s * dx3$
29:         **end if**
30:         **if** $dy3 < 0$ **then**
31:             $y \leftarrow y - s * dy3$
32:         **end if**
33:         **if** $dz3 < 0$ **then**
34:             $z \leftarrow z - s * dz3$
35:         **end if**
36:         HILBERTCURVE($s, x, y, z, dx2, dy2, dz2, dx3, dy3, dz3, dx1, dy1, dz1$)
37:         HILBERTCURVE($s, x + s * dx1, y + s * dy1, z + s * dz1, dx3, dy3, dz3, dx1, dy1, dz1, dx2, dy2, dz2$)
38:         HILBERTCURVE($s, x + s * dx1 + s * dx2, y + s * dy1 + s * dy2, z + s * dz1 + s * dz2, dx3, dy3, dz3, dx1,$ $dy1, dz1, dx2, dy2, dz2$)
39:         HILBERTCURVE($s, x + s * dx2, y + s * dy2, z + s * dz2, -dx1, -dy1, -dz1, -dx2, -dy2, -dz2, dx3,$ $dy3, dz3$)
40:         HILBERTCURVE($s, x + s * dx2 + s * dx3, y + s * dy2 + s * dy3, z + s * dz2 + s * dz3, -dx1, -dy1, -dz1,$ $-dx2, -dy2, -dz2, dx3, dy3, dz3$)
41:         HILBERTCURVE($s, x + s * dx1 + s * dx2 + s * dx3, y + s * dy1 + s * dy2 + s * dy3, z + s * dz1 + s *$ $dz2 + s * dz3, -dx3, -dy3, -dz3, dx1, dy1, dz1, -dx2, -dy2, -dz2$)
42:         HILBERTCURVE($s, x + s * dx1 + s * dx3, y + s * dy1 + s * dy3, z + s * dz1 + s * dz3, -dx3, -dy3, -dz3,$ $dx1, dy1, dz1, -dx2, -dy2, -dz2$)
43:         HILBERTCURVE($s, x + s * dx3, y + s * dy3, z + s * dz3, dx2, dy2, dz2, -dx3, -dy3, -dz3, -dx1,$ $-dy1, -dz1$)
44:     **end if**
45: **end procedure**

In this case, no remapping step is required as we are directly using corresponding RGB value in both the operations and thus the outputs of both the operations are also colored images.

It is well known that opening (closing) of an image can be obtained by sequential operations of erosion (dilation) and dilation (erosion), i.e.,

$$f \circ B = (f \ominus B) \oplus B \tag{3}$$

$$f \bullet B = (f \oplus B) \ominus B \tag{4}$$

The bright and dark top-hat images can be obtained from these opened and closed images highlighting the bright and dark features in the image respectively as follows:

$$f_{btop} = f - (f \circ B) \tag{5}$$

$$f_{dtop} = (f \bullet B) - f \tag{6}$$

Now, the contrast enhanced image is given by

$$f_{enhanced} = f + f_{btop} - f_{dtop} \tag{7}$$

A better enhanced image can be obtained by finding bright and dark features at multiple scales. This can be obtained by performing opening and closing operations not only using the SE B but also its higher order homothetics. The nth order homothetic of a convex SE B can be obtained by dilating B recursively $(n-1)$ times with itself as follows [5].

$$nB = B \oplus B \oplus B ... \oplus B$$

Thus, performing the above operations for n structuring scales, we will get n bright and dark top-hat transforms highlighting the bright and dark features at various scales. So, the final contrast enhanced image is obtained as

$$f_{enhanced} = f + \sum_{i=1}^{n} f_{btop} - \sum_{i=1}^{n} f_{dtop} \tag{8}$$

Here, i denote the scale of the image.

4   Experimental Results and Comparisons

To test the performance of our proposed method, experiments have been conducted on a number of color images (taken from [15, 16]). All the experiments have been performed on a machine with Intel Core i3, 2.3 GHz CPU, 2 GB RAM and

Fig. 4 Contrast enhancement results for different test images, first column: original color images, second column: results with marginal ordering, third column: results with distance-based vectorial ordering, and fourth column: results of proposed Hilbert space filling curve-based mapping

Windows 7 OS. We have compared the results of our proposed approach with the results obtained by following other ordering approaches like marginal ordering and distance-based vectorial ordering while all other steps remain the same. In marginal ordering, first RGB image has been converted to HSI space, and then, morphological operators have been applied on them separately to get the result.

The results with four different color images are shown in Fig. 4. The first column shows the original images, the results with marginal ordering and distance-based ordering are shown through the second and third columns of the figure, and the fourth column shows the results of applying our proposed Hilbert curve-based method.

We can find from these images, the contrast has been considerably enhanced in our proposed method as well as in case of other two approaches compared to the original image. So for the purpose of comparison, it is very important to quantify how much the contrast has been enhanced in all the three methods, and we obtain the statistical measurement of contrast from the corresponding co-occurence matrix. To get the co-occurence matrix for the color image, it has been first converted to YUV space, and then, co-occurence matrix has been computed for the Y-component. From this matrix, contrast can be found out like as follows [17]:

Table 2 Comparison of contrast of output images obtained through different methods

Image	Channel	Original	Marginal	Distance	Proposed
Lenna	R	0.4437	2.3502	2.1601	2.0975
	G	0.4643	1.7402	2.1752	2.0208
	B	0.4323	1.8095	2.2917	2.1282
	Y	0.4368	2.0666	2.2730	2.1650
Parrot	R	0.1767	0.4950	0.7503	0.6095
	G	0.2207	0.5517	0.9712	0.7395
	B	0.1768	0.4616	0.7766	0.6248
	Y	0.2064	0.5502	0.8958	0.7192
Fruits	R	0.2601	0.8217	1.2114	0.9027
	G	0.2710	0.7200	1.2740	0.9348
	B	0.2427	0.6402	1.1630	0.8160
	Y	0.2614	0.7616	1.2711	0.9323
Jelly beans	R	0.1135	0.2726	0.5487	0.3335
	G	0.1178	0.2306	0.5407	0.3450
	B	0.0731	0.1921	0.5129	0.2845
	Y	0.0972	0.2268	0.5475	0.3281

$$Contrast = \sum_i \sum_j (i-j)^2 N_d(i,j)$$

where $N_d(i,j)$ is the normalized co-occurrence matrix defined by

$$N_d(i,j) = \frac{C(i,j)}{\sum_i \sum_j C(i,j)}$$

Here $C(i,j)$ represents number of times the pixel gray level value i co-occurs with value j with some designated spatial relationship in the image. Results are given first for the R, G, and B components separately and then for the Y-component of all the test images in Table 2.

Following observations can be made about the outputs based on the results obtained:

i Contrast of Hilbert curve-based method is better than marginal ordering but not better than distance-based ordering (except for R channel of lena image).

ii Marginal ordering output contains black lines which is not present in other methods.

iii Time complextiy of method presented in this paper is better than distance-based ordering which can be proved as follows.

For total ordering, first we need to compute Hilbert curve which requires $256 \times 256 \times 256$ computations. However, once this pre-computation is done, it requires same time as a grayscale image takes for remaining processing.

Let $k \times k$ be the size of SE and $m \times n$ be the size of image, So

$$(\textit{Time complexity})_{HC} = O(k^2 mn)$$

However, for distance-based ordering, the time taken is more since for every SE, we need to compute the distance for each of its element from every other element present under the SE. So

$$(\textit{Time complexity})_{DO} = O(k^4 mn)$$

It is clear from the above expressions that our proposed method is k^2 times better than the distance ordering-based approach in terms of time complexity.

iv Since marginal approach converts RGB image to HSI and then processes each component separately. So

$$(\textit{Time complexity})_{MO} = O(Nk^2 mn)$$

where N is number of components in the image. So our proposed method is N times better than the marginal approach in terms of time complexity.

Thus, we can summarize here that although our method has not produced the best contrast enhancement, it takes comparatively less time than the other methods.

5 Conclusion

A contrast enhancement method for color images is proposed in this paper using Hilbert space curve-based total ordering and morphological processing. In this method, first we map every RGB value to a scalar value using Hilbert space filling curve and then apply morphological operators to obtain contrast enhanced image. Visual analysis of outputs suggests that it is better than the outputs obtained through marginal approach, but not as good as distance-based approach. However, in terms of time complexity, it is better than both the distance-based approach and the marginal approach. As contrast enhancement is a preprocessing step in many higher level image processing tasks, so it is quite essential to compute it in as efficient time as possible, and our method serves the purpose.

References

1. Gillespie, A.R., Kahle, A.B., Walker, R.E.: Color enhancement of highly correlated images. I. decorrelation and HSI contrast stretches. Remote Sens. Environ. **20**(3), 209–235 (1986)
2. Hummel, R.: Image enhancement by histogram transformation. Comput. graph. Image Process. **6**(2), 184–195 (1977)
3. Xiao, D., Ohya, J.: Contrast enhancement of color images based on wavelet transform and human visual system. In: Proceedings of the IASTED International Conference on Graphics and Visualization in Engineering, pp. 58–63. ACTA Press (2007)
4. Starck, J.-L., Murtagh, F., Candès, E.J., Donoho, D.L.: Gray and color image contrast enhancement by the curvelet transform. IEEE Trans. Image Process. **12**(6), 706–717 (2003)
5. Mukhopadhyay, S., Chanda, B.: A multiscale morphological approach to local contrast enhancement. Signal Process. **80**(4), 685–696 (2000)
6. Kim, Y.-T.: Contrast enhancement using brightness preserving Bi-Histogram equalization. IEEE Trans. Consum. Electron. **43**(1), 1–8 (1997)
7. Abdullah-Al-Wadud, M., Kabir, M.H., Dewan, M.A.A., Chae, O.: A dynamic histogram equalization for image contrast enhancement. IEEE Trans. Consum. Electron. **53**(2) (2007)
8. Huang, S.-C., Cheng, F.-C., Chiu, Y.-S.: Efficient contrast enhancement using adaptive gamma correction with weighting distribution. IEEE Trans. Image Process. **22**(3), 1032–1041 (2013)
9. Dougherty, E.: Mathematical Morphology in Image Processing. CRC Press (1992)
10. Zaharescu, E.: Analysis of morphology-like operators used in color image contrast enhancement. Acta Universitatis Apulensis. Mathematics-Informatics **8**, 448–470 (2004)
11. Lambert, P., Chanussot, J.: Extending mathematical morphology to color image processing. In: Proceedings of the CGIP 2000 (2000)
12. Aptoula, E., Lefèvre, S.: A comparative study on multivariate mathematical morphology. Pattern Recognit. **40**(11), 2914–2929 (2007)
13. Chanussot, J., Lambert, P.: Total ordering based on space filling curves for multivalued morphology. Comput. Imaging Vis. **12**, 51–58 (1998)
14. Stevens, R.J., Lehar, A.F., Preston, F.H.: Manipulation and presentation of multidimensional image data using the peano scan. IEEE Trans. Pattern Anal. Mach. Intell. **5**, 520–526 (1983)
15. University of south california image dataset. http://sipi.usc.edu/database/database.php?volume=miscl
16. Pinteres image dataset. https://in.pinterest.com/trishh47/birds-~-birdhouse-~-bird-feeders-~-birdbath/l
17. Haralick, R.M., Shanmugam, K., et al.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)

Agent-Based Modelling and Simulation of Religious Crowd Gatherings in India

Abha Trivedi and Mayank Pandey

Abstract Management of religious crowd gatherings is a complex and essential task. Lack of planning and management has resulted in many unfortunate incidents and deaths in the past. Though local authorities deploy crowd management personnel to avert and control this crisis, some unforeseen events still lead to crowd panic resulting in a sudden rush of individuals. In this paper, we have utilized Agent-Based Modelling technique to model a near to real scenario of crowd gatherings at Alopi Devi temple of Allahabad, India. We have attempted to incorporate certain heuristics such as closing and opening of exit and entry gates which can be beneficial in controlling crowd-related disaster. These heuristics are then simulated over panic situations to analyse their effects in terms of numbers of victims. We have used NetLogo, an Agent-Based Modelling tool, to design and simulate our model. The simulation results establish the applicability of our methodology.

Keywords Agent-Based Modelling and Simulation · Religious gatherings
Crowd disaster · NetLogo

1 Introduction

Looking back at the history of disasters happened at crowded places, it may be observed that lots of human lives have been lost due to mismanagement and lack of planning [1]. These crowded places include popular religious spots, railway stations, live rock-band concerts, etc. India has also faced such tragedies in the past few decades. Most of these disasters occurred at temples and places of religious gatherings. A common aspect at such places is that generally there is not much crowd

A. Trivedi (✉) · M. Pandey
Motilal Nehru National Institute of Technology Allahabad,
Allahabad 211004, Uttar Pradesh, India
e-mail: abhson1711@gmail.com

M. Pandey
e-mail: mayankpandey@mnnit.ac.in

at these places, but on some auspicious day or event, there is sudden rush of crowd which is difficult to control and manage. At some overcrowded occasions some misleading information, fire or any significant event may result in panicked rush claiming precious human lives [2]. People in crowd have their own cumulative behaviour, and it is difficult to predict their responses to such events. These types of false events can result in confusion among the persons forcing them to change their behaviour drastically [3]. The available infrastructure at these places is capable to handle normal inflow of crowd but is generally unable to handle the unprecedented rush of crowd, which is difficult to predict or speculate. Further, it is very difficult to test and analyse the preventive measures under real-time environment. It is required to model such mass gatherings and analyse behavioural patterns of individuals with the help of modelling and simulation techniques. On the basis of levels of abstraction, there are three different approaches for modelling crowd dynamics, namely, macroscopic, mesoscopic and microscopic [4]. Micro-models work at lower level of abstraction and have the ability to describe behaviours of individuals in fine-grained manner [5].

Agent-Based Modelling and Simulation is an elegant micro-modelling approach to model the behaviour of individuals in a crowd [6]. It is helpful in modelling the real-world scenarios where an individual can be represented as a programmable agent and its behaviour can be simulated according to predefined setup. Further, individuals in crowd possess distinguishing properties which make it possible to model them as agents. Essentially, there exists a synergy between agents forming an Agent-Based Model and behavioural patterns of crowd. This synergy makes Agent-Based Modelling a better candidate to model the crowd behaviour. Many Agent-Based Modelling (ABM) tools are available to effectively perform such experiments [7].

In this paper, Agent-Based Modelling and Simulation (ABMS) has been used to model crowd gathering at Alopi Devi temple of Allahabad, India. We have used NetLogo [8] an open source simulation platform for Agent-Based Modelling developed by Uri Wilensky from Northwestern University, for simulating our model. This simulation platform provides a good interface to design near to real environment. NetLogo has several default primitives which help to model the behaviour of individual agents. We have simulated behaviour of each individual in both normal and panic situations. The effect of panicking behaviour has been analysed and compared in terms of numbers of victims under different scenarios.

The rest of the paper is structured as follows. Section 2 discusses the relevant related works. Section 3 presents the details of our case study of Alopi Devi temple. In Sect. 4, we present the details of simulation experiments and analysis of results for different scenarios. Finally, conclusion is given in Sect. 5.

2   Related Work

Several researches have been made towards crowd-related disasters at places of mass gatherings featuring different aspects of crowd behaviour. Some studies [9, 10] have

discussed the causes and required rectifications in such disasters, whereas some empirical studies [11–13] have given mathematical models on crowd dynamics. In [11], a comparative study on the issues in crowd dynamics and the evacuation dynamics is presented, such as calculation of evacuation time, identification of congestion and corrective actions. A comprehensive study on the methodology of verification and validation of models of crowd dynamics is presented in [12] which suggests certain amendments. In [13], impact on pedestrian flow has been evaluated due to the presence of groups in a crowd of people at places like corridor or T-junctions.

Further, some research efforts [14–17] have utilized ABM technique to study crowd behaviour in emergency evacuations from buildings, public exhibition pavilions, concert houses, or stadiums. In [14], an effort has been made to model human gatherings during Tawaf at Masjid Al-haram in Maqqa. In [15], a model is developed in NetLogo, to simulate Ramy on Jamarat Bridge. An Agent-Based evacuation simulation is performed in [16] to measure efficient evacuation in terms of evacuation rate and collision rate utilizing leader–follower and leaver approach. In [17], a simulation system is designed to understand human social behaviours in emergency. Additionally, some of the efforts [18, 19] have been made towards utilizing agent-based modelling and simulation technique for crowd management and control. In [18], a multi-agent system is developed to simulate long crowded queue at temples to monitor and avoid occurrences of stampedes. In [19], robotic agents are used to detect congestion at stampede prone areas to increase the efficiency of crowd management system.

3 Our Approach

In this section, we present the details of our field survey performed to collect geographical parameters of Alopi Devi temple and its surroundings along with other relevant details. Further, we present our Agent-Based Model which includes types of agent, their behaviour, and interactions with other agents.

3.1 Field Survey: Alopi Devi Temple

Alopi Devi temple is devoted to goddess Durga and has significant religious importance among Hindus. Usually, a huge surge of crowd is experienced at this temple during auspicious occasions. Morning and evening times are the auspicious hours for prayer and offerings. To collect geographical measurements of temple and its surrounding areas, we have used Trimble Juno 3 series [20] Global Positioning System (GPS) unit during the field survey. This temple is built in an area of around $15000\,\text{m}^2$, out of which around $280\,\text{m}^2$ is open space. There are three main gates for entry/exit that are used by the people to reach sanctum sanctorum of the temple. These gates are located at different distances from sanctum sanctorum. The connecting pathways

from these gates to the sanctum sanctorum are around 4 m wide. Further, the temple is reachable via two 6 m wide roads, Matiyara road and Daraganj Ghat road passing by its front side.

3.2 Our Agent-Based Model

In this section, we present our Agent-Based Model which is utilized to simulate crowd gathering at Alopi Devi temple. We have categorized two types of agents termed as '*mobile agents*' and '*fixed agents*'. We have used *mobile agents* to represent pilgrims and *fixed agents* to represent elements of environment. The elements of environment such as roads, ground, walls, gates, etc. are modelled on the basis of their geographical attribute values. With the help of *mobile agents* and *fixed agents*, we have created the virtual world of Alopi Devi temple and its surrounding to simulate near to real scenarios.

Modelling Mobile Agents We have categorized pilgrims as men, women and children along three parameters: height and width [21], walking speed and energy level. As established in [22], there are many factors like age, health, density, etc. that affect speed of pedestrians. An estimated average speed of pedestrians is around 1.34 m/s. Based on these findings we have implemented mobility speed of our agents. These agents have the ability to decide directions at intersections and also show automatic reactions to avoid obstacles and other pilgrims on the way [23]. Figure 1 represents different behaviours possessed by *mobile agents* at any instance of time.

Behaviour of each individual is simulated for normal and panic situations. Under the normal situations, they follow a directed path to reach sanctum sanctorum, perform prayers and offerings, and exits. This disciplined behaviour of *mobile agent* gets transformed into random and wayward rush if certain misleading information spreads leading to panic situations. The panicking behaviour is modelled based on

Fig. 1 Behaviour of mobile agents at any instance
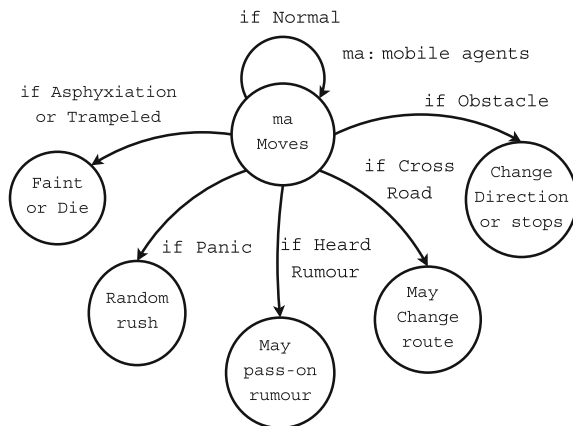
the studies done in [24]. Such panic situation leads to death of *mobile agents* due to asphyxia and trampling. We have utilized concepts of crowd-induced forces, energy level and certain other parameters to model deaths of *mobile agents*. The values for these parameters are set on the basis of empirical studies performed by [2, 25, 26]. They reported that the forces in crowd create huge pressure among individuals. For example, five persons are capable of developing a force of 3430 N on guardrails which is enough for a person to faint or die. In addition, [25] has established that the tolerable force for women is significantly less when compared with men. Further, in [26], a mathematical concept of resultant forces in crowd is formalized as $F = F_d + F_{pp'} + F_o$. The resultant force F is the summation of driving force (F_d), interaction force between two people ($F_{pp'}$) and interaction force between a person and the obstacle (F_o). We have utilized several built-in primitives provided in NetLogo to set different parameters and model the behaviour of *mobile agents*.

4 Simulation: Experiments and Results

We have used NetLogo to implement our simulation environment and Agent-Based Model. Initially, the NetLogo world of Alopi Devi temple and its surroundings with all geographical parameters was set up. We assumed each patch to be equal to 1.5 m^2 inside temple premises and 2 m^2 for other areas in the model. In NetLogo patch is a cell over which agents can move and patches have coordinates. It may be noted that the variations in assumptions of patch sizes are applied to maintain width of different types of roads and area of temple premises.

On this basis, the diameters of national highway, local roads and inner pathways are set to be 4-patches, 3-patches and 2-patches. Figure 2 depicts the snapshot of our simulation setup where it can be seen that roads surrounding the temple, inner
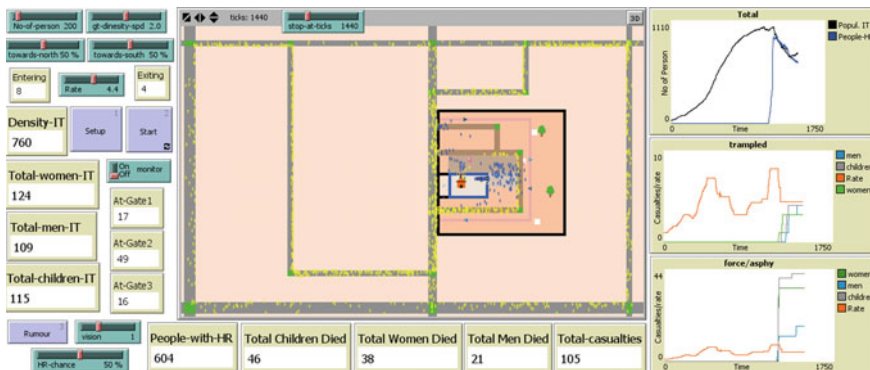
Fig. 2 Snapshot of simulation setup

pathways, boundary wall and temple boundary are set with grey, brown, black and blue colour, respectively.

It has been reported in [25] that crowd with more than 4 persons/m^2 is at high risk of disaster. Based on this study, we have considered maximum 4 persons/m^2 for normal crowd density and 8 persons/m^2 for overcrowded density. The initial positions and speed of *mobile agents* are set randomly. The simulation experiments are run for 1440 ticks (one tick assumed as 1 min) on varying crowd density. Tick is the simulation time step. The experiments are simulated for 5000 agents, approximately.

4.1 Experiments

We have performed experiments on three different scenarios under two different classes of densities. In the first scenario, entry and exit are kept open from all three gates (Gate I, Gate II and Gate III). In the second scenario, entries are kept open from Gate I and Gate III, and exit from Gate II. We have taken this scenario on the basis of location of gates and their distances from sanctum sanctorum. In the third scenario, the points of entry and exit are kept similar to second scenario, and additionally utilized a density threshold inside temple premises. The density threshold is utilized to maintain density at 4 persons/m^2 by restricting entries of pilgrims. We have simulated these scenarios to analyse the effects of panic situations in terms of numbers of victims. The parameters that we have set for a pilgrim to die or faint in overcrowding are density per square metre, availability of space to move and energy level. In the situation of random rush and increased density, resultant force (described in Sect. 4.2) on pilgrims increases. If pilgrims could not find space to move and resultant force goes beyond threshold, their energy level reduces. If the energy level of any pilgrim goes below the threshold, he/she may faint or dies.

4.2 Results

Table 1 shows total men, women and children victims in all three scenarios for both the crowd densities. In the first scenario, serious consequences of crowd disturbances are observed. Effects of crowd disturbances in the second scenario are less in comparison to the first scenario. Further, application of density threshold in the third scenario resulted in approximately 50% reduction of numbers of victims. In Fig. 3, graphs (a) and (b) depict comparative results for all scenarios in normal and heavy crowd densities, respectively.

Table 1 Results of simulation experiments

Simulation for 1440 ticks	Mishap in varying crowd density/m^2	Victims			
		Men	Women	Children	Total
First scenario	4 persons/m^2	46	81	92	219
3Entry/3Exit	8 persons/m^2	66	113	136	315
Second scenario	4 persons/m^2	11	35	36	82
2Entry/1Exit	8 persons/m^2	32	49	72	153
Third scenario	4 persons/m^2	3	11	31	45
2Entry/1Exit WT	8 persons/m^2	22	36	52	110
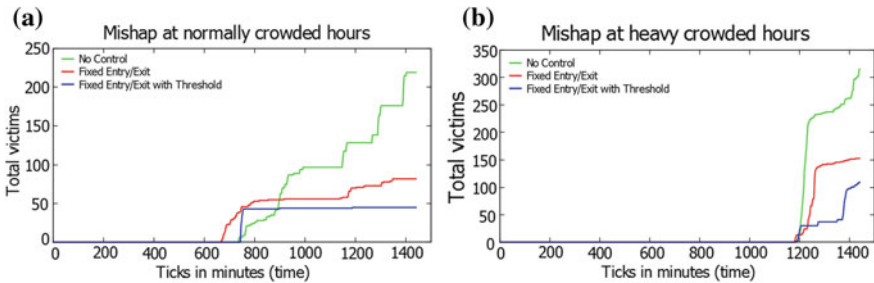
Fig. 3 Victims in **a** normal crowd density and **b** heavy crowd density

5 Conclusion

In this paper, we have presented an Agent-Based Model for implementing near to real scenarios of crowd gatherings at religious places. We have considered the case of Alopi Devi temple situated in Allahabad, India as our simulation environment. We have utilized basic concepts of crowd dynamics to model behaviour of pilgrims arriving at this temple. The purpose of our model is to test and analyse the effect of crowd-induced disturbances in different entry and exit scenarios. The analysis is performed with respect to the numbers of pilgrims who fainted or died. It is observed that count of victims is more when no control is enforced in the entry and exit of pilgrims. After deploying strategies of fixed entry and exit, and density threshold, there is a huge reduction in this count. This establishes the applicability of our proposed model. Currently, we are attempting to use an extended framework to model control strategies being followed by local authorities. Also, we wish to propose new control strategies to remove shortcomings of their strategies. Further, to use our model for larger gatherings in areas such as Maha-Kumbh Mela of Allahabad.

References

1. Illiyas, F.T., Mani, S.K., Pradeepkumar, A.P., Mohan, K.: Human stampedes during religious festivals: a comparative review of mass gathering emergencies in India. Int. J. Disaster Risk Reduct. **5**, 10–18 (2013)
2. Fruin, J.J.: The causes and prevention of crowd disasters. Eng. Crowd Saf. **1**, 10 (1993)
3. Helbing, D., Johansson, A.: Pedestrian, Crowd and Evacuation Dynamics. Springer (2011)
4. Ijaz, K., Sohail, S., Hashish, S.: A survey of latest approaches for crowd simulation and modeling using hybrid techniques. In: 17th UKSIMAMSS International Conference on Modelling and Simulation, pp. 111–116 (2015)
5. Johansson, F.: Microscopic Modeling and Simulation of Pedestrian Traffic (2013)
6. Crooks, A.T., Heppenstall, A.J.: Introduction to Agent-Based Modelling. In: Agent-based Models of Geographical Systems, pp. 85–105. Springer (2012)
7. Macal, C.M., North, M.J.: Tutorial on agent-based modeling and simulation. In: Proceedings of the 37th Conference on Winter Simulation, WSC '05, pp. 2–15. Winter Simulation Conference (2005)
8. Wilensky, U.: Netlogo 5.2 (1999). https://ccl.northwestern.edu/netlogo/
9. Kasthala, S., Lakra, H.S.: Disaster preparedness for mass religious gatherings in India-learning from case studies. In: Second World Congress on Disaster Management (2015)
10. Helbing, D., Mukerji, P.: Crowd disasters as systemic failures: analysis of the love parade disaster. EPJ Data Sci. **1**(1), 1 (2012)
11. Schadschneider, A., Klingsch, W., Klüpfel, H.: Tobias Kretz, Christian Rogsch, and Armin Seyfried. Evacuation dynamics: empirical results, modeling and applications. In: Encyclopedia of Complexity and Systems Science, pp. 3142–3176. Springer (2009)
12. Robert, L., Mycek, M., Porzycki, J., Jarosaw, W.: Verification and validation of evacuation models-methodology expansion proposition. Transp. Res. Procedia **2**, 715–723 (2014)
13. Vizzari, G., Manenti, L., Crociani, L.: Adaptive pedestrian behaviour for the preservation of group cohesion. Complex Adapt. Syst. Model. **1**(1), 1 (2013)
14. Haghighati, R., Hassan, A.: Modeling the flow of crowd during tawaf at masjid al-haram. Jurnal Mekanikal **36**, 02–18 (2013)
15. Ilyas, Q.M.: A netlogo model for ramy al-jamarat in hajj. J. Basic Appl. Sci. Res. (2013)
16. Vani, K., Idicula, S.M.: Agent based evacuation simulation using leader-follower model. Int. J. Sci. Eng. Res. **4** (2013)
17. Pan, X., Han, C.S., Dauber, K., Law, K.H.: A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations. AI and Society (accepted for publication) (2006)
18. Ahuja, G., Karlapalem, K.: Crowd congestion and stampede management through multi robotic agents. CoRR (2015). arXiv:1503.00071
19. Kolli, S., Karlapalem, K.: Mama: Multi-agent management of crowds to avoid stampedes in long queues. In: Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems, AAMAS '13, pp. 1203–1204. Richland, SC (2013). International Foundation for Autonomous Agents and Multiagent Systems
20. Trimble: Transforming the way the world works. Juno series. http://www.trimble.com/mappingGIS/junoS.aspx
21. Butera, L.: Height, power, and gender: politicizing the measured body. Ph.D. thesis, Bowling Green State University (2008)
22. Daamen, W., Hoogendoorn, S.: Experimental research of pedestrian walking behavior. Transp. Res. Rec. J. Transp. Res. Board **20–30**, 2003 (1828)
23. Jarosaw, L.R.W.: Towards realistic and effective agent-based models of crowd dynamics. Neurocomputing **146**, 199–209 (2014)
24. Fedewa, N., Krause, E., Sisson, A.: Spread of a rumor. Society for Industrial and Applied Mathematics, vol. 25. Central Michigan University (2013)
25. Still, G.K.: Introduction to Crowd Science. CRC Press (2014)
26. Helbing, D., Molnár, P.: Social force model for pedestrian dynamics. Phys. Rev. E **51**, 4282–4286 (1995)

An Experimental Study of Scalability in Cross-Domain Recommendation Systems

Akarsh Srivastava, Aman Jain, Ashwin Jayadev, Rajdeep Mukherjee, Shronit Bhargava and Prosenjit Gupta

Abstract Recommender systems attempt to predict the future behavior of a particular individual based on her past preferences. Today any individual may have more than one profile that he/she maintains on various websites, and leveraging all this data on the preference of an individual from various domains (cross-domain) can help us in making better user models that can be used to make better and improved recommendation. A cross-domain recommender system thus aims to improve the recommendation of a target domain extracting and using the metadata from many source domains. Building scalable recommender systems is always a challenge in the presence of Big Data, and this is compounded for cross-domain recommenders. In this paper, we aim to tackle the problem of scalability in cross-domain recommendations. We have performed various experiments to divide the datasets into smaller clusters and then running a recommendation algorithm using the attributes in the dataset to return the best recommendations.

Keywords Cross-Domain recommender systems · Collaborative filtering
Scalability

A. Srivastava (✉) · A. Jain · A. Jayadev · R. Mukherjee · S. Bhargava · P. Gupta
Department of Computer Science and Engineering, NIIT University,
Neemrana, Rajasthan, India
e-mail: akarsh372@gmail.com

A. Jain
e-mail: amanjainddn@gmail.com

A. Jayadev
e-mail: jaydevashwin7@gmail.com

R. Mukherjee
e-mail: rajdeep.mukherjee295@gmail.com

S. Bhargava
e-mail: bhargavashronit@gmail.com

P. Gupta
e-mail: prosenjit_gupta@acm.org

1   Introduction

Over the last two decades, a lot of effort has focused on developing recommender systems [1, 2]. A popular approach in recommender systems is collaborative filtering [1]. In collaborative filtering, we can use a single domain for recommendation or multiple domains. Here domain refers the type of an item. For instance, in our case, we have used two domains for collaborative filtering which are two sources of movies: one source being MovieLens and another being Kaggle. Involving multiple domains in recommendation gives us a better chance at recommending items effectively [3–5]. For instance, if we use movies and books as the two domains then we can recommend movies that have the same genre as the books that the user has rated well and vice versa. Cross-domain collaborative filtering is powerful in improving recommendations.

2   Our Contribution

2.1   Scalability of the System

For the purpose of achieving a scalable recommender system, we had to partition the data to make the search of an item computationally efficient. We used clustering to initially partition the data into clusters and recommending items from those clusters. Clustering helps in greatly reduce our search space [6]. The clustering algorithm used was **K-Means clustering algorithm.** The entire recommendation was therefore broken up into two phases: **Preprocessing,** where we cluster the data and **Querying,** where we actually recommend similar movies from the appropriate cluster. Hence, we are not searching the movie in the entire search space but only a subset of it. We conducted a detailed experimental study with the goal of improving the scalability.

2.2   Numeric and Non-numeric Data Clustering

One issue we tackled was the clustering of non-numeric data. The clustering of numeric data was easy as we had to simply create vectors of the numeric fields in our datasets; however, in case of non-numeric fields, we had to find a way to represent non-numeric fields as vectors. This was achieved using **tf-idf vectorization** of non-numeric data. Due to the inclusion of both numeric and non-numeric data, the recommendation is enhanced. Our final approach of combining both numeric and non-numeric data was to create clusters of non-numeric data first and then to create subclusters of numeric data inside the non-numeric clusters.

3 Experiments

3.1 Datasets Used

For the purposes of experimenting with our methodology, we used datasets from two different sources. The item chosen for experimentation was **movies** (Table 1).

3.2 Methodology

Preprocessing. Recall that our objective was to tackle scalability through partitioning the datasets effectively. To address this issue of scalability of large recommender systems due to their enormous data we need to partition the data so that at a time we can work with smaller portions of the data.

Various techniques were used and analyzed for various ways of partitioning the data. These techniques are illustrated in the various experiments that follow.

Querying. After preprocessing, querying would imply that we query the partitions of the data for similar items which in our case are movies. Since now we are searching for similar movies within a small portion with respect to the entire dataset, our query will be both faster and more effective.

3.3 Experiment 1: Genre and Rating Based Partitioning

A rather naive approach to solve the problem of scalability we used the **genre** of movies in order to partition the entire data and within each genre **ratings of movies** were used to create another level of partitions.

Partitions chosen for ratings for the dataset **MovieLens**
[0–1.5): Partition A
[1.5–3): Partition B
[3–5]: Partition C

Table 1 Datasets used and their specifications

Dataset	Source	Size (MB)	Unique entries	Columns	Description
Movies	MovieLens (Domain 1)	15.4	10197	33	Movies versus movie attributes
Movies	Kaggle (Domain 2)	1.44	5043	28	Movies versus movie attributes
Users	MovieLens	26.7	2114	9	Users versus user ratings for movies in MovieLens

Partitions chosen for ratings for the dataset **Kaggle**:

[0–2.5): Partition A

[2.5–5): Partition B

[5–7.5): Partition C

[7.5–10]: Partition D

Results and Drawbacks. Since genre was the only criteria used for the partitioning of data into partitions, the results were consistent only with respect to the genre and their corresponding rating. One important drawback of this approach is that we were not using other potentially contributing fields of the dataset such as actor names, director names, location of shoot, and the like.

3.4   Experiment 2: Numeric Data Clustering

In an attempt to use maximum possible number of relevant fields to partition the data, we had to come up with a proper partitioning scheme that would use a number of fields together. We, therefore, turned our attention to information retrieval techniques. In Information Retrieval and Machine Learning, a number of classification techniques are used. One such technique used extensively for unsupervised classification is called Clustering.

Method. Following the clustering technique, we first considered it appropriate to consider all the numeric data present in the datasets as it would be easy to consider them as dimensions in the n-dimensional hyperspace. These points can be thought of as n-dimensional vectors. In our case after conversion of the items to vectors, we normalized the vectors and then we used **K-Means clustering algorithm** to cluster those points (representative of the items present in the dataset).

K-Means Clustering Algorithm. K-Means is one of the many clustering algorithms available to cluster the points in an n-dimensional hyperspace. K-means clustering aims to partition x observations (where x is the total number of observations or in our case the total number of movies present in one dataset) into k clusters. K-Means clustering is applied in domains like Information Retrieval, Recommender systems and the like.

Numeric Clustering in MovieLens Dataset. A total of 14 attributes were chosen for numeric clustering in MovieLens Dataset.

All these attributes are considered for the purpose of vectorizing each item. This means that each attribute essentially becomes a dimension in the n-dimensional hyperspace where n is the total number of attributes which in our case is 14.

The scale of all the dimensions may not be the same and hence it is necessary to apply a normalization of the vectors. In our case, we applied L2-normalization in order to normalize the vectors.

After all, points are found and normalized the k-means algorithm is applied in this 14-dimensional hyperspace in order to cluster the points and hence the movies into clusters.

Numeric Clustering in Kaggle Dataset. A total of seven attributes were chosen for numeric clustering in Kaggle dataset.

The clustering was applied in the same way in the Kaggle dataset as well separately.

Results and Drawbacks of this Approach. After the clustering of the two datasets into two separate set of clusters, clusters were stored in a csv file. Although no querying was done on this dataset, it was clear that the optimum result could not be achieved by just using the numeric data present in the datasets. We initially had taken the numeric data because it could be easily converted to vectors without any prior processing. We were not considering some of the more potent fields like actors in the movie, director, genre, and so on.

3.5 Experiment 3: Non-numeric Data Representation and Nearest Neighbor Approach

The technique we used for including non-numeric data was tf-idf vectorization which is popular in Information Retrieval. Inverse document frequency (idf) is used to measure the 'rareness' of a term.

Conversion of Non-Numeric Data to vectors. Non-numeric fields considered in the MovieLens Dataset: **ActorName, Genre, and DirectorName**.

The other non-numeric fields were not chosen as such they had little relevance with respect to similarity of movies and what a typical viewer with a specific choice would like to see. Non-numeric fields considered in the Kaggle dataset are **DirectorName, Actor2Name, Genre, Actor1Name, and Actor3Name**.

Conversion of data in fields to vectors: Conversion of field values to documents. For each item (movie), the values present in the fields were converted to paragraphs such that each value in the field becomes one term in the paragraph and the paragraph essentially became a document.

Generation of a term-document matrix. The term-document matrix can then be generated in the following manner. For instance, considering the MovieLens dataset (Tables 2 and 3):

Mind you that the total number of terms (even for a single movie) here is 14. Needless to say that we are getting an extremely sparse matrix here.

Once we get these term frequencies for each movie (here document) these can be considered as coordinates for an n-dimensional hyperspace where n is the number of terms present in the entire collection and the vectors are documents themselves.

Table 2 Sample data from dataset

Movie name	Actor name	Genre	Director name
Titanic	Leonardo diCaprio, Kate Winslet	Drama, Romance	James Cameron
Tomorrow never dies	Colin Salmon, Geoffrey Palmer	Action, Romance, Thriller	Roger Spottiswoode
Twisted	Richard T Jones, Andy Garcia	Comedy, Drama	Philip Kaufman

Table 3 Term-document matrix

Movie name	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
Titanic	1	1	0	0	0	0	1	1	0	0	0	1	0	0
Tomorrow never dies	0	0	1	1	0	0	0	1	1	1	0	0	1	0
Twisted	0	0	0	0	1	1	1	0	0	0	1	0	0	1

Normalization of the vectors. The vectors are next normalized following the L-2 normalization technique or Euclidean normalization technique.

Querying of movies using nearest neighbor approach. Since all vectors are unit length, the similarity of two items (movies) can simply be computed by figuring out the dot product between the two items. Similarly, now we can use a query movie and find out movie; the query is most similar to by finding out the dot product of that movie with every other movie.

This querying of movies was done in both datasets separately and the final result was simply the combination of the two. If dataset 1 returned three movies: movie1, movie2, and movie3 and dataset 2 returned three movies: movie4, movie5, and movie6, then the final result was simply shown as movie1, movie2, movie3, movie4, movie5, and movie6.

Results and Drawback of this approach. The results obtained were not as convincing as the expected. For instance when the query movie Spiderman was given the top 5 most similar movies were expected to be sequels of Spiderman. However, that was not the case.

Although it seems to be an excellent approach, time is a major factor in reducing the importance and relevance of this approach. Partitioning the data was required especially in order to tackle scalability. However, we were unable to tackle scalability as we were not partitioning the data into appropriate partitions. We were simply searching in the entire collection for better results. This was causing an $O(n)$ algorithm at a high level even during the querying phase. This would quickly become an unfeasible solution when the dataset would be large.

3.6 Experiment 4: Non-numeric Data Clustering and Subsequent Querying

The next obvious approach was to amalgamate the ideas of both non-numeric data representation and clustering for tackling the scalability issue. Hence, we considered K-Means clustering to cluster the data.

Conversion of data to vectors. The conversion of data to vectors was done in exactly the same way as the previous experiment.

Generation of term-document matrix. The term-document matrix was generated in exactly the same way as the previous experiment.

Normalization of vectors. Once again, we used L-2 normalization for the normalization of the vectors in exactly the same way as the previous experiment.

Application of K-Means Clustering. The difference between the previous approach and this one comes here. Here we apply K-Means clustering to the formed vectors in order to cluster the data. Recall that specifics of K-Means clustering have already been discussed in a previous experiment. Once the clusters are formed, it is now easy to search the similar movies among the specific cluster movies. This solution would therefore tackle the issue of scalability. The clustering was done on both datasets separately.

Storage of Clusters. All the movies present in a specific cluster were stored in a csv file where each row of the file was one cluster and the ids of the movies were stored in each row.

Number of Clusters (size of K). Many articles on K-means clustering denote that the number of clusters to be formed or the value of k depends on the data itself. In our case, we found that for the Kaggle dataset number of clusters = 100 yielded optimum results, and in the case of MovieLens dataset, it was also found that number of clusters = 100 yielded optimum results.

Results: Querying and Cluster Evaluation
We have a total of three datasets: the movies dataset of MovieLens, the movies dataset of Kaggle and the user dataset of MovieLens. For the purposes of querying and cluster evaluation, the following steps were employed:

1. A user was selected from the user dataset.

2. The top 3 highest rated movies were selected for that user.

3. The clusters of those movies were found in both datasets separately.

4. For the purposes of cluster evaluation, we return the intersection of the movies that the user has watched and rated and the movies that are present in that cluster. This was done separately for both datasets and for all three movies.

Illustration. If user A has watched and rated the following movies:

userA: {m1: 4, m2: 3.5, m3: 3.5, m4: 3.5, m5: 4, m6: 4, m7: 5, m8: 4, m9: 4, m10: 5, m11: 4.5}

and the top 3 highest rated movies are *{m7: 5, m10: 5, m11: 4.5}*

Let the clusters for these three movies be *{cluster7, cluster88, cluster94}*

Let the movies in the clusters for the MovieLens dataset be the following:

Cluster 7. Contents of this cluster are: *m1, m2, m4, m7, m15, m88, m54*

Cluster 88: Contents of this cluster are: *m10, m78, m41, m6, m84*

Cluster 94: Contents of this cluster are: *m11, m3, m44, m71, m89, m27, m26, m35*

Hence, for the first movie m7, the movies we get after querying are the intersection of the movies the user has watched and rated and the movies present in the cluster of movie m7, in this case, cluster 7. Therefore, we get the set of movies: *{m1, m2, m4, m7, m15, m88, m54}* ∩ *{m1, m2, m3, m4, m5, m6, m7, m8, m9, m10, m11}* = *{m1, m2, m4, m7}.*

The same approach was also followed for Kaggle dataset.

After we get the required movies, we compute the average rating of all movies that we have got as results. Here rating is the user's rating of the movies. Now if the cluster quality is good enough the clusters should contain movies which that same user has preferably rated well which in turn means that when we compute the average rating for each and every movie that we queried, in this case movies: *{m7, m10, m11}*

Illustration. For the first movie m7, we got movies: *{m1, m2, m4, m7}*

These movies have the following ratings given by the user: *{m1: 4, m2: 3.5, m4: 3.5, m7: 5}*; averaging these ratings we get (4 + 3.5 + 3.5 + 3.5)/4 = 4.

This measure gives us an idea of the cluster quality, there may be movies with lower ratings given by the user. But since the cluster should contain similar movies to the movies that the user has rated high, it is obvious that the average rating we got should also be high enough. Therefore, a rating of 4/5 is pretty good for cluster quality.

The above average rating was computed for all three top movies separately to evaluate cluster quality. This same approach was also followed to evaluate cluster quality of the Kaggle dataset.

Fine Tuning Experiment 4: Inclusion of numeric data. A drawback of the previous approach was that we were using only non-numeric data in the previous experiment. This is not leading to full utilization of the datasets we had, as we were not using the numeric data.

This approach involved two levels of clustering. At the first level, clusters were formed using non-numeric data just like in the previous experiment. At the next level, subclusters were formed using the numeric data of the movies present in each cluster. Numeric data was done in exactly the same way which is described in Experiment 2. This was done in both datasets separately.

References

1. Shi, Y., Larson, M., Hanjalic, A.: Collaborative filtering beyond the user-item matrix: a survey of the state of the art and future challenges. ACM Comput. Surv. (2014). http://dx.doi.org/10.1145/2556270
2. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2010)
3. Berkovsky, S.,Kuflik, T., Ricci, F.: Cross-domain mediation in collaborative filtering. User Modeling, Proceedings of the 2007 11th International Conference, UM 2007, Corfu, Greece, 25–29 June 2007. https://doi.org/10.1007/978-3-540-73078-1_44
4. Tang, J., Yan, J., Ji, L., Zhang, M., Guo, S., Liu, N., Wang, X., Chen, Z.: Collaborative users' brand preference mining across multiple domains from implicit feedbacks. In: Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI'11)
5. Zhang, Y., Cao, B., Yeung, D.-Y. Multi-domain collaborative filtering. In: Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (UAI'10) (2010)
6. Sarwar, B.M., Karypis, G., Konstan, J., Riedl, J.: Recommender systems for large-scale e-commerce: scalable neighborhood formation using clustering. In: Proceedings of the 5th International Conference on Computer and Information Technology (ICCIT) (2002)

Document Categorization Using Graph Structuring

Sandipan Sarma, Punyajoy Saha and Jaya Sil

Abstract This paper proposes a document classification model using feature learning (Coates, Demystifying unsupervised feature learning, 2012) [5] approach based on semantics of the documents. In the learning phase, basic vocabulary (BV) for each document class consisting of nouns has been created by proposing a novel approach. The classification phase searches unique words in the BVs and if found, the corresponding sentence becomes a basic sentence (BS). A tree with unique words of the BS is inserted in the respective forest. Associated words of the children are used to continue the tree formation process until no new node is generated in the tree. Finally, we assign the test document to a class which has a clearly dominant percentage of sentences in the respective forest. The proposed algorithm is compared with various feature-based classification models and satisfactory performance has been observed.

Keywords Document categorization · Basic vocabulary · Importance function

1 Introduction

Today's organizations face problems of managing complex textual data, mostly unorganized and unstructured. So explicit knowledge is distributed among different types of documents and it is difficult to localize and identify the required text [10, 11]. Hence, we need to develop methods for document categorization, which exploit inherent semantics of the data by analyzing the unique words of the documents. Various models that exist for categorization of documents are decision trees [13], neural

S. Sarma (✉) · P. Saha · J. Sil
Indian Institute of Engineering Science and Technology, Shibpur,
Botanical Garden, Howrah, West Bengal, India
e-mail: sandipan.sarma7@gmail.com

P. Saha
e-mail: punyajoysaha1998@gmail.com

J. Sil
e-mail: js@cs.iiests.ac.in

networks [12], and k-NN and TF-IDF-based models [14]. The decision tree interprets the datasets by constructing well-defined true/false queries in the form of a tree structure. The decision trees may give more than one possible answers by initializing the root node with different variables. The TF-IDF-based method [14] does not capture spatial information in the text and co-occurrences of words in different documents. The limitations of the existing methods in document classification have been dealt with by proposing an efficient model of document categorization [15] which extracts semantics of data using hierarchical graph structure. This paper aims to learn the relation between the words in the documents of different classes using feature learning approach, similar to human ability to capture context of the documents. The test documents are preprocessed to design the "spanning tables"for unique words containing important information like frequency, connected words, and position of the words. In the learning phase, the BV of each class has been obtained by proposing a new measure which evaluates importance of each word in the documents. Next, to extract semantic of the documents, the relation between the words of [3] the respective BV is learned by developing the co-occurrence and conditional probability matrices. Finally, the test document is assigned to the class represented by the forest containing the maximum number of test sentences [7]. The proposed classifier model, Graph Classification with Basic Quantum (GCBQ), is compared with various feature-based classification models using text corpus of five different classes: Business, Entertainment, Politics, Sports, and Technology. Performance of the classifier has been evaluated and observed that the model works better than the existing models of document categorization with respect to time complexity, and comparable otherwise. This paper has been organized into five sections. Section 2 describes the preprocessing method and the learning scheme of the classifier model. Section 3 presents the classification of the test documents, and results appear in Sect. 4. Finally, conclusions are drawn in Sect. 5.

2   The Methodology

The proposed GCBQ model first describes preprocessing of the training and test documents by removing stop words and singularizing them [1]. The preprocessed training datasets consist of noun words only, assuming that noun-to-noun links (common nouns) [2] represent the strongest relation between two words in a document. The preprocessed test document has been broken into sentences to extract unique words from each sentence. We now consider a three-attribute list for each word in the test document as described by the *basic spanner*, shown in Fig. 1.

   A collection of such basic spanners forms the *spanning table* of the test document, where words are linked with predecessor and successor in the sentences. For each new word in the test sentences, the latest spanning table is searched. If the word is found, its frequency is incremented and the third attribute is appended. Otherwise, we create a corresponding basic spanner, and thus, the complete spanning table of the document is obtained. Consider the following document and its spanning table given in Table 1.
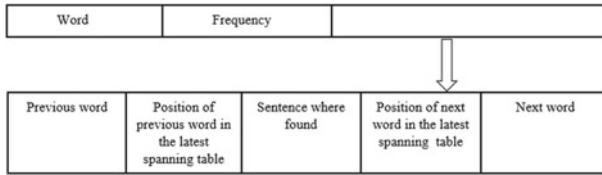
Fig. 1 Elements of a basic spanner

Table 1 The complete spanning table

Word	Frequency	Semantic information
Software	3	(NULL,−1,0,2,patent), (bill,5,1,−1,NULL), (consultation,7,3,2,patent)
Patent	2	(software,1,0,3,law), (software,1,3,5,bill)
Law	2	(patent,2,0,4,axe), (chance,9,2,−1,NULL)
Axe	1	(law,3,0,−1,NULL)
Bill	4	(NULL,−1,1,1,software), (NULL,−1,2,6,round), (patent,2,3,−1,NULL), (NULL,−1,4,10,technology)
Round	1	(bill,5,2,−1,NULL)
Consultation	2	(round,6,2,8,chance), (NULL,−1,3,1,software)
Chance	1	(consultation,7,2,3,law)
Technology	1	(bill,5,4,−1,NULL)

EU software patent law faces axe. The European Parliament has thrown out a bill that would have allowed software to be patented. Politicians unanimously rejected the bill and now it must go through another round of consultation if it is to have a chance of becoming law. During consultation the software patents bill could be substantially redrafted or even scrapped. The bill was backed by technology.

Initially, the spanning table is empty. NULL means no word to the left/right; in such a case, the position of the previous/next word is denoted by −1.

2.1 Learning

In the learning phase, fixed size basic vocabularies of class-specific training documents are created using a basic quantum (BQ) measure. Semantic relationships between the unique words of all the training documents are evaluated using co-occurrence and conditional probability matrices for each document [4].

2.1.1 Basic Vocabulary

The BQ measure of a word is evaluated as product of word frequency in the documents of a particular class and the number of documents containing the word.

BQ signifies how important a word is in the class. The words are ranked in decreasing order using BQ, and the **top five words** are taken as the BV of each class, determined experimentally.

2.1.2 Semantic Analysis

In order to do the semantic analysis [3, 4], we find associations between the unique words of the class-specific training documents by constructing a co-occurrence matrix. Each element of the co-occurrence matrix represents number of occurrence of the words in the respective rows and columns. Next, we construct the conditional probability matrix using the co-occurrence matrix, each element of which represents conditional probability of the co-occurring words, calculated as follows:

$$P(A/B) = P(A, B)/P(B) \tag{1}$$

Two points are considered here: first, if A and B co-occur only once, and both of them individually also occur only once, it is assumed that their relative importance in the class would be very low, hence P(A/B) is neglected, i.e., P(A/B) = 0, and second, as per the formula, P(A/A) = 1, as it is a sure event.

3 Classification

The words of the preprocessed test document are searched in each BV for a match using the basic spanners in the spanning table of that document whose frequency attribute is more than 10% of the highest frequency among all the basic spanners. For classification, a test document is tested against all the classes simultaneously. A forest is maintained for each class, initially empty. For each word in a particular document, the basic words among the sentences in all the currently available basic spanners are searched. If a basic word is found, the sentence containing that word is pushed into the forest of the class in which the basic word is found. These sentences are the *basic sentences* and they represent the *parent sentences*. Since the BV of each class will be updated and new words will be added to it later on, while finding a match for basic words in the vocabulary, at first the **top five words** from each class are taken. If no match is found, then the next five words are considered. This process of selection of basic words continues till the basic spanners are exhausted of words or a match is found. When a match is found, all the words that fall just to the right and left of the basic words in the sentences are extracted; since the basic or parent sentences are directly linked with the class, so the nouns within them are assumed to be related. Then, the relation between each pair of words: the word and its predecessor, and the word and its successor, and the frequency of each new word appearing in the document is obtained. The relation between any two words is found from the conditional probability matrix created during the learning phase. For each

CLASSIFICATION PROCESS FLOWCHART

Fig. 2 Flowchart of the classification process

word found in this way, the sentences where the word has occurred are retrieved from the spanning table. These sentences are then inserted into the forest of the class of the basic words, of which the new words found were children of. That is, these new sentences are the children of their parent sentence. The new words become the secondary vocabulary words. Thus, many trees of sentences are formed in the forests. When this is all over, a list of words under each basic word with their frequency and relation with their right and left words is obtained. The above steps are repeated by considering the secondary vocabulary words as basic words now, assuming they are related to the class of their parent basic words. The process comes to a halt if no new basic word is found. A flowchart of the process is shown in Fig. 2.

Now, an *importance function* is defined as follows:

$$I(\omega, \upsilon) = S \times \omega + P \times \upsilon \tag{2}$$

where ω = weight, and it is the conditional probability of co-occurrence of a word and its parent in the tree, υ = frequency of the word, and S and P are experimentally determined constants (S = 0.053 and P = 2.043). The secondary vocabulary words found are included in the proper vocabularies. All these words, along with the existing words in the respective vocabularies are ranked in decreasing order according to the value returned by the importance function, hence updating the vocabularies. Now, the percentage of total sentences in the document that got inserted into the forest of each class is checked. If there is a class which has a clearly dominant
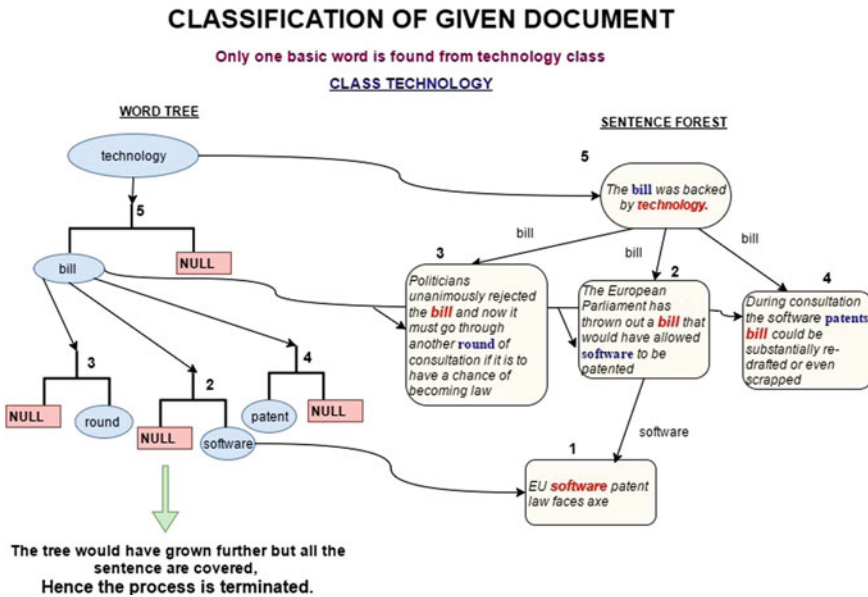
Fig. 3 Classification using technology as the basic word

percentage, then the document is assigned that class. As an example, the process mentioned above is shown in Fig. 3 applied on the sample test document considered earlier.

Now, since there are no more classes with basic word, the class (Technology) having maximum number of sentences is assigned to the document. The words software, patent, bill, and round are inserted into the basic vocabulary of Technology and arranged according to the importance function.

4 Experimental Results

The dataset consists of 2,225 documents from the BBC news [6] website corresponding to the stories in five topical areas from 2004 to 2005: politics, sports, business, entertainment, and technology. Each time a document is tested, the class is declared on the basis of the number of sentences that is inserted in the forest of a specific class. There can be another situation when no word of a particular document matches with any basic word of any class. In that case, the document is not analyzed. If there is no "clearly dominant" forest, then ambiguity may arise. As evident from Table 2,

Table 2 Performance with updated vocabulary and non-updated vocabulary

Class	No. of training documents	No. of test documents	$S^a(uv^b)$	$S(nuv^c)$	$F^d(uv)$	$F(nuv)$	$A^e(uv)$	$A(nuv)$
Politics	291	126	95	86	25	34	6	6
Sports	350	150	109	48	40	94	1	8
Entertainment	270	116	91	79	11	23	14	14
Business	357	153	87	68	39	58	27	27
Technology	280	121	50	37	50	65	21	13

[a]Successes
[b]Updated vocabulary
[c]Non-updated vocabulary
[d]Failures
[e]Ambiguities

Table 3 Perfomance of the GCBQ model

Classes	Precision	Recall	Specificity
Politics	0.77	0.92	0.94
Sports	0.86	0.91	0.96
Entertainment	0.77	0.91	0.95
Business	0.73	0.74	0.93
Technology	0.92	0.50	0.99
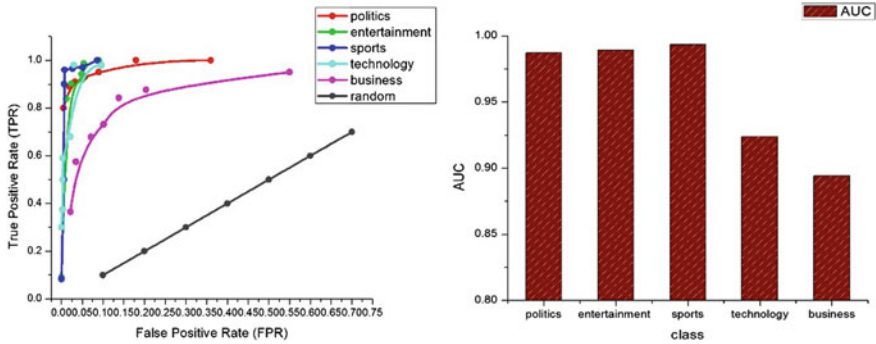
Fig. 4 ROC curves for the five classes and AUC values for different classes

categorization of documents with updated vocabulary is more accurate because the words in secondary vocabulary are strongly related to the BV which increases the chance of matching. Performance of the model is summarized in Table 3. Figure 4 shows the ROC curve and AUC values for different classes of GCBQ. The percentage of sentences considered for success can be varied and used as a threshold to generate the ROC curve [8]. The AUC values are within 0.9 to 1, which is an indication of a good classifier.

5   Conclusions

The objective of this paper is to devise an efficient algorithm to classify an unknown document into one of the given classes while elucidating the word-to-word relation. Document classification appears in many applications, including e-mail filtering, news monitoring, topic modeling, and identification of document genre. GCBQ model performs better than the existing methods like K-NN and TF-IDF while shows comparable performance to the existing decision tree model. The proposed classifier model is scalable and interpretable due to graphical approach in classification. In our future work, we consider other parts of speech in order to extract more information and deal with the problem of overlapping classes.

References

1. Bird, S.: NLTK Documentation (2017)
2. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
3. Bollegala, D., Matsuo, Y., Ishizuka, M.: Measuring Semantic Similarity Between Words Using Web Search Engines, Semantic Web (2007)

4. Bolshakov, I.A., Gelbukh, A.: Two methods of evaluation of semantic similarity. In: Kedad, Z., Lammari, N., Mtais, E., Meziane, F., Rezgui, Y. (eds.) Natural Language Processing and Information Systems. NLDB 2007. Lecture Notes in Computer Science, vol. 4592 (2007)
5. Coates, A.: Demystifying Unsupervised Feature Learning (2012)
6. Cunningham, D.G.: Practical solutions to the problem of diagonal dominance in kernel document clustering. In: Proceedings of 23rd International Conference on Machine learning (ICML'06), pp. 377–384. ACM Press (2006)
7. Dao, N.T., Simpson, T.: Measuring Similarity Between Sentences
8. Fawcett, T.: ROC Graphs: Notes and Practical Considerations for Data Mining Researchers. Intelligent Enterprise Technologies Laboratory, HP Laboratories, Palo Alto (2003)
9. Gomaa, W.H., Fahmy, A.A.: A survey of text similarity approaches. Int. J. Comput. Appl. **68**(13) (2013)
10. Gretarsson, B., et al.: Topicnets: visual analysis of large text corpora with topic modeling. ACM Trans. Intell. Syst. Technol. (TIST) **3**(2), 23 (2012)
11. Lafferty, J.D., Blei, D.M.: Correlated topic models. In: Advances in Neural Information Processing Systems (2006)
12. Li, C.H., Park, S.C.: Artificial Neural Network for Document Classification Using Latent Semantics Indexing. IEEE, Joenju, South Korea (2007)
13. Ramaswamy, S.: Multiclass Text Classification: A Decision Tree Based SVM Approach. University of California, Berkley
14. Trstenjaka, B., Mikac, S.: KNN with TF-IDF based framework for text categorization. Science Direct **69**, 1356–1364 (2014)
15. Tsuda, K., Saigo, H.: Graph classification. In: Managing and Mining Graph Data, pp. 337–363 (2010)

An Empirical Analysis for Predicting Source Code File Reusability Using Meta-Classification Algorithms

Loveleen Kaur and Ashutosh Mishra

Abstract Although various quantifiers of software component reusability have been proposed, these metrics have seldom been utilized in existing literature to analyze source code file reusability transpiring within a single product family. Such metrics can be effortlessly employed to develop reuse prediction models which can support the software practitioners in obtaining information regarding the total cost involved in developing a novel version of a prevailing software or upgrading an existing software version by estimating the total reusable code files in advance without being compelled to scrutinize the complete codebase. In view of this, this research work aims to examine the efficacy of seven meta-classification techniques in the development of such reuse prediction models on four software datasets constructed from four successively released versions of software using appropriate reuse metrics. We also evaluate the predictive performance of these meta-classifiers against the statistical technique of logistic regression and rank these techniques using the Friedman statistical test.

Keywords Software reuse · Meta-classification techniques · Machine learning
Reusability prediction

1 Introduction

A systematic reuse activity incorporates the utilization of previously developed software components/artifacts to create new software systems, thus leading to the overall reduction in the cost, development time, and effort of software development [1]. For instance, in order to identify the actuality of reusability in existing software projects, investigation was performed on numerous sizeable open-source and

---

L. Kaur (✉) · A. Mishra
Department of Computer Science & Engineering, Thapar University, Patiala, India
e-mail: loveleen.kaur@thapar.edu

A. Mishra
e-mail: ashutosh.mishra@thapar.edu

individual software projects bundled with popular distributions of BSD and Linux [2]. The observations depicted that almost fifty percent of the total number of files (5.3 million) scrutinized in the projects have been employed in at least two projects. However, no software metrics were employed by the author to conduct this analysis.

Software metrics for calculating the ease of reusability of software projects are vital for accomplishing "development by reuse" and "development for reuse" [3]. Besides, these reuse metrics could turn out to be contributory in the development of reusability prediction models that could be applied by software developers to gather knowledge pertaining to the aggregate expenditure involved in developing a novel version of an existing software or updating an existing software version by knowing the total code that can be reused, i.e., integrated without change in the new version, in advance without requiring to examine the complete codebase. Since the development of new code would need substantial time and effort, reusability valuation of source code components could aid in minimizing this development effort and time, and provide means to assess the development cost of the new software.

The machine learning (ML) algorithms are being effectively utilized to build potent predictor models on a varied set of areas such as engineering, medicine, geology, etc. [4]. Classification, in ML, is the challenge of allocating a new observation to a certain set of category or class (subpopulation), based on a given training set of data having instances (or observations) whose category belonging is known [4]. An algorithm implementing classification is known as a classifier. Meta-classifiers formulate a collection of classifiers and then classify novel data records on the basis of the combination of the results of these classifiers using certain mechanisms. Investigational results have indicated that meta-classifiers often show more accuracy and more robustness even in the presence of noisy data, and obtain a lesser average error rate as compared to the constituent individual classifiers [5].

Therefore, in this study, we conduct an empirical validation by means of the reusability datasets constructed(with respect to seven randomly selected reuse metrics) from four consecutively released versions of a software developed in Java language to determine the competency of seven meta-classification algorithms for developing version to version source code reusability prediction models. In addition to this, we examine and compare the results of these meta-classifiers with logistic regression (LR) [6] technique using performance indicators such as accuracy and AUC ROC analysis. Lastly, we statistically rank all the techniques used in this study using Friedman's statistical test in order to know which algorithm performs the best.

Though there are studies [7, 8] in the existing literature that assess the ability of classification techniques for change proneness and fault prediction, there has been no study till date performing the statistical comparison of the performance of meta-classifiers with the LR technique to predict version to version source code reusability.

The rest of this paper is organized as follows. The next section delivers a concise summary of the existing literature related to the topic in concern. Section 3 describes

the empirical data collected and independent and dependent variable selected as a part of the research background. Section 4 reports the meta-classification techniques employed and the performance evaluation measures selected to evaluate the meta-classifiers. Section 5 describes the empirical analysis results of the application of the seven meta-classifiers and the LR technique along with the Friedman statistical test results. Section 6 comprises the various threats to validity of our work, followed by the last section, in which we state the conclusions and the future work.

2 Related Work

This section provides a brief summary of the literature related to the software reusability prediction and for detailed reading, one can refer to the review [9], which happens to be the most recent and the sole systematic literature review to be done with respect to reusability metrics and prediction of software components according to established guidelines of systematic literature review. Recently, self-organizing maps (SOM) were employed for clustering the datasets corresponding to the CK metrics values which were gathered from three Java-based projects [10]. It was also recently established that the reusability of a source code class increases reciprocally with the increase in depth of inheritance as well as number of children [11]. Moreover, authors [12] have also considered different variants in reuse (common reuse, high-reuse variation, low-reuse variation, and single-use) to empirically and correctly estimate fault proneness across products and across releases of Software Product Lines. Though many metrics have been proposed for the measurement of the reusability of a software component or a software as a whole in the previous research works [3, 9, 13–15], in the majority of the cases these have been qualitative reuse metrics, the evaluation of which is certainly dependent on individuals. Also, very few of the machine learning techniques (K-means and hierarchical clustering, support vector machines, artificial neural networks, and decision trees) have been explored for reusability prediction which do not include any ensemble or meta-classifiers or LR techniques. Moreover, these articles do not include the comparison of actual metric values against concrete reuse results of a realistic software development environment to validate the reusability prediction. Also, an empirical evaluation of the software reuse transpiring within the same product family and from version to version via the reuse metrics has not been analyzed in the literature examined.

To redress this, we assess a wide range of meta-classification techniques for reusability prediction with the help of four datasets which have been created from four consecutively released versions of a realistic Java-based software using actual values of their reuse results and some randomly selected reuse metrics. We also use Friedman statistical test to allocate statistical ranks to the techniques for the purpose of determining if the selected meta-classifiers significantly outperform the LR technique, with the motive to provide empirical substantiation to evaluate the best version to version reusability prediction model.

3 Research Background

The following subsections constitute an overview of the empirical data collected and the metrics (dependent variables and independent variable) incorporated to form the reusability prediction models.

3.1 Empirical Data Collection

In order to construct the version to version file reusability prediction models, we created four datasets using four consecutive versions of the JFreeChart[1] software, which is a free (LGPL) chart library for the Java platform. The details of the four selected versions have been illustrated in Table 1. Numerical values of seven different and commonly used static code metrics were collected corresponding to each of the Java files existing in each of the four selected versions of JFreeChart using two different open-source static code analysis tools- Stan4J[2] and JHawk 6.1.3.[3] These Java files are basically composed of one or more than one java classes (or interfaces). In those cases where there exists more than one class per file, an aggregate value of the metrics with respect to all the classes is collected. A brief description of the seven selected reuse metrics (Coupling between objects, Efferent Coupling, Depth of inheritance, Lack of cohesion between methods, Number of Calls, Number of methods defined in a file, and Cyclomatic Complexity) has been given in below:

Coupling between objects [16] is stated as the total number of files coupled to a given file. Two source code files are said to be coupled to each other if the methods declared in one makes use of some other code file's instance variables or methods, whereas *Efferent Coupling* [17] only evaluates the total number of external files used by a given file. *Depth of inheritance* [16] is the maximum length of a path from a given source code file to the root code file in the inheritance structure of the given software. *Lack of cohesion between methods* [16] calculates the total of different methods in a given code file that refer a given instance variable. *Number of Calls* is the number of method calls (in statements as well as in logical expressions) in the target file. *Number of methods defined in a file* [18] is the total of methods contained in a given Java code file, and lastly, *Cyclomatic Complexity* [19] measures the number of independent paths through program source code.

We used a clone detection tool called AntiCut&Paste [20] in order to estimate the reusability of individual source code files of each of the versions. Source code files of two successive versions of the selected JFreeChart software were supplied to the tool as an input and it returned those Java files which were discovered to be

---

[1]http://www.jfree.org/jfreechart/.

[2]http://stan4j.com.

[3]JHawk Metrics Tool, http://www.virtualmachinery.com/jhawkprod.htm.

Table 1 Details of the four JFreeChart versions employed in this study

JFreeChart version	Release date	Total size (in LOC)	Total number of Java files	Number of Java files reused into next version
JFreeChart 0.6.0	2001-11-27	16,445	86	67
JFreeChart 0.7.0	2001-12-30	20,723	105	47
JFreeChart 0.7.1	2002-01-25	24,171	128	94
JFreeChart 0.7.2	2002-02-09	24,961	130	112

common to the two releases. Post this step, a binary variable of a "No" or "Yes" was calculated as the reuse statistic and was allocated for every source code file with "Yes" indicating that the file had been reused into next version in its entirety and without any modifications and "No" indicating that the file had not been employed in the next version or was employed with some modification.

3.2 Dependent and Independent Variables

In our study, the binary variable of reusability is selected as the dependent variable which is to be estimated via the independent variables. The independent variables are those for which the results need to be calculated for the prediction of the reuse statistic of a Java file in the next release of the software. Therefore as per our context of application, the independent variables are the seven software metrics discussed in Sect. 3.1.

4 Research Methodology

Having described the construction of the four datasets in Sect. 3, this section discusses the various meta-classification techniques incorporated in this study for the prediction of source code file reusability. The measures undertaken to analyze the performance of the selected meta-classification algorithms have also been stated.

4.1 Meta-Classification Techniques Employed

This section provides a brief description of the seven meta-classifiers [21] (Ada-Boost, Bagging, Filtered, Multi-class (M-Class), Random Sub Space (RSS), Stacking and Voting) incorporated in this study.

AdaBoost or Adaptive Boosting utilizes a sequence of simple weighted classifiers, where each classifier is made to analyze a separate characteristic of the data, to finally generate an all-inclusive classifier, with the help of which there is a high probability of obtaining a low misclassification error rate as compared to an individual classifier. *Bagging* (stands for Bootstrap Aggregation) classifier supplies random subgroups of the original dataset to every base classifiers and then combines their separate predictions (either using averaging or voting) to decide the absolute prediction. The *Filtered* classifier consists of running an arbitrary classifier on data that has been passed through an arbitrary filter which uses some mathematical evaluation (that is based on some intrinsic characteristic of the training set like correlation). *Multi-class (M-Class)* classifier formulates a methodology to convert a given multi-class problem into numerous binary class problems. The metric is evaluated for every class by considering it to be a binary classification problem after combining all the remaining classes as second-class entities. Then the weighted average (weighted by class frequency) or macro average (treat each class equally) metric is obtained by averaging the binary metric over all the classes. However, unlike binary classification problems, here one does not need to select a threshold score to generate predictions. The label or class obtaining the highest predicted score is the predicted answer. The *Random Sub Space (RSS)* classifier is similar to bagging technique except here random subsets of the dataset are drawn as random subsets of the features whereas in bagging the samples are drawn with replacement. *Stacking* is similar to boosting. The difference here is that several classifiers are combined using the stacking method instead of an empirical formula for the weight function and the base learner's predictions are given as input for a meta-level classifier whose output is the final class. In the *Voting* methodology, the base-level classifier's predictions are combined in accordance with a static voting scheme (usually the plurality voting scheme).

4.2 Performance Evaluation Measures

Two performance evaluation measures: Accuracy and Area under the ROC Curve (AUC) are chosen to assess the predictive performance of the selected algorithms against version to version source code file reusability. The *accuracy* of a model is depicted as the ratio of the number of Java files that are predicted with accuracy to the total number of Java files in the version. The *AUC* metric is obtained through the ROC[4] plotting which indicates the optimal cutoff point at which both sensitivity and specificity are maximized. AUC values $>= 0.7$ and < 0.8 exhibit acceptable

---

[4]The ROC curve is a plot of sensitivity (on the y-axis) and 1-specificity (on the x-axis). Several cutoff points between 0 and 1 are chosen during the creation of a ROC curve. The *sensitivity* of the model is calibrated as the percentage of the reused files that were predicted accurately. The *specificity* of the model is calibrated as the percentage of the non-reused files that were predicted accurately. High values are desired for both sensitivity and specificity.

division; AUC values >= 0.8 and < 0.9 exhibit excellent division; and AUC >= 0.9, exhibit outstanding division between the reused and non-reused files by the prediction algorithm.

5 Empirical Analysis

Results of the models constructed using meta-classifiers to predict the version to version file reusability on the four selected versions of JFreeChart are described in this section of the paper. These results were predicted using the WEKA tool which is an open-source tool and is easily obtained on http://www.cs.waikato.ac.nz/ml/weka/. The Naïve Bayes [5] classifier was used as the base classifier/learner. We used the default settings of tool to construct the seven meta-classification models.

5.1 Model Evaluation Results

The columns in Table 2 show the version-wise Accuracy (Acc.) in %, and Area under Curve (AUC) scored by each of the seven meta-classification algorithms on each of the four versions of JFreeChart. In order to acquire a supplementary accurate assessment with respect to the predictive potential of the selected classification models, a k cross-validation [22] of all the models generated in this research was conducted in which the dataset is randomly separated into roughly "k" equivalent subsets and for each assessment, one of the k subsets is employed as the test set and the training set is formed with the residual k−1 subsets. This process is reiterated for all the k subsets. In this study, the meta-classification models generated observations stated in Table 2 were validated with a "k" value of 10.

From the values of the performance measures observed in Table 2, six out of seven models(except the Stacking meta-classifier) exhibit good results, i.e., depicting high scores for both the performance metrics, with accuracy values ranging from 75.8 to 90% and AUC values ranging from 0.64 to 0.88. The obtained results, especially the AUC results show that the six out of seven selected meta-classification algorithms display an acceptable and, in some cases, excellent discrimination between the reused and non-reused files contained in the four JFreeChart versions, thus demonstrating their effectiveness for developing fitting and authentic version to version source code reusability prediction models. The Stacking technique[5] achieves the lowest AUC values (< 0.50). Thus even though it achieves high accuracy values (which is solely due to the correct classification of

---

[5]The results indicated that stacking achieved a sensitivity value of 0 and specificity value of 1 for all the four selected versions, thus indicating that it is not able to classify the reused classes at all and predicts all the classes included in the dataset to be "not reused".

Table 2 Results of the meta-classification analysis

Meta-classifiers	JFreeChart 0.6.0		JFreeChart 0.7.0		JFreeChart 0.7.1		JFreeChart 0.7.2	
	Acc.	AUC	Acc.	AUC	Acc.	AUC	Acc.	AUC
ABM1	80.2	0.70	76.2	0.77	77.3	0.78	83.9	0.83
Bagging	80.2	0.87	78.1	0.82	76.6	0.76	83.9	0.75
Filtered	76.7	0.66	73.3	0.87	74.2	0.72	83.1	0.79
M-Class	80.2	0.86	77.1	0.85	78.1	0.77	83.9	0.79
RSS	80.2	0.64	74.3	0.83	76.6	0.75	80.8	0.79
Stacking	77.9	0.46	55.2	0.46	73.0	0.45	86.2	0.45
Vote	80.2	0.87	77.1	0.88	78.1	0.78	83.9	0.81

Table 3 Binary logistic regression results

JFreeChart 0.6.0		JFreeChart 0.7.0		JFreeChart 0.7.1		JFreeChart 0.7.2	
Acc.	AUC	Acc.	AUC	Acc.	AUC	Acc.	AUC
77.9	0.82	76.2	0.81	75.8	0.79	90	0.87

the reusable classes) the Stacking meta-classifier does not qualify as an efficient predictor of version to version reusability.

We also applied the LR technique on the four selected datasets, the results of which are stated in Table 3.

The results indicate that the LR technique also shows a comparable performance to six out of seven meta-classifiers, especially to the Vote meta-classifier with accuracy values ranging from 73.3 to 83.9% and AUC values ranging from 0.79 to 0.87.

Though there does exist a difference(however minor) in the prediction performance of the models developed via the selected techniques, we needed to ascertain if the difference between them holds statistical significance for which we conduct the Friedman statistical test [23].

Table 4 reports the mean ranks scored by every technique after the use of Friedman test, where the model with the lowest mean rank is the one which performs the worst. The test was based on the AUC results obtained by each of the models on the four selected JFreeChart datasets. According to the results stated in Table 4, the technique with the best performance on all the three datasets is the Vote meta-classifier with a mean rank of 7. This is followed by the LR algorithm which obtains a mean rank of 6, which is closely followed by the M-Class meta-classifier with a mean rank of 5.25. The Stacking meta-classifier is selected as the worst meta-classifier with the lowest mean rank of 1.

The Friedman statistical value with degree of freedom seven was evaluated to be 15.421, which is true for $\alpha = 0.05$. Additionally, a p-value of 0.031 exhibits that the results obtained are true at a 95% confidence interval. Therefore, the null hypothesis of Friedman test that states that all techniques perform the same is rejected. The eight techniques (seven meta-classifiers and LR technique) taken

Table 4 Friedman statistical test results

Algorithm	Mean rank
ABM1	4.88
Bagging	4.38
Filtered	4
M-Class	5.25
RSS	3.50
Stacking	1
Vote	7
LR	6

under consideration are significantly diverse in their performance behavior. The results, however, indicate that only one meta-classifier "Vote" out of the seven meta-classifiers performs better than the LR statistical technique therefore establishing the LR technique to be also effective for the prediction of version to version source code file reusability.

6 Threats to Validity

There can be potential threats to this study like any other empirical study. Our research work focuses only on the estimation of prediction performance of the selected meta-classifiers which was performed using the statistical and machine learning methods, with the metrics as the independent variables and the reuse parameter of Yes/No as the dependent variable. Thusly, the threat to internal validity exists since this research work does not signify to establish change-outcome. The most critical threat to the external validity of our study is that our results may not generalize to a similar sample or new research environment and could be constrained to the surveyed systems, i.e., the four selected JFreeChart versions. In order to ascertain the generalizability of the classification inferences made in our study, the predictive performance of the selected techniques on similar datasets developed using other programming languages need to be evaluated. So this threat exists in the study. The construct validity makes sure that independent and dependent variables properly represent the concepts. The metric data was collected via mature source code analysis and clone detection tools. Although we make no declarations with respect to the accuracy of these software, we suppose that the tools collect the data reliably as they are in fact being employed effectively in practice [24, 25], therefore decreasing the threat to construct validity.

7 Conclusion and Future Work

This research work constitutes the evaluation of seven meta-classification techniques to predict version to version source code file reusability. The empirical validation was done on four datasets created using four consecutively released versions of JFreeChart. To the best of the author's knowledge, no study till now has made use of meta-classifiers for reusability prediction. We further compared the performance of the selected meta-classifiers with the statistical LR technique and ranked the performances of the various algorithms using the Friedman statistical test.

Following are chief conclusions made from this analysis:

All the selected meta-classifiers, except for the Stacking technique, showed reasonably good performances(Accuracy and AUC values ranging from 75.8 to 90% and 0.64 to 0.88) over the four selected versions of the JFreeChart software

and did not depict extremely divergent outcomes as far as the values of the performance measures over the four versions were concerned. The LR technique also showed comparable performance to the meta-classifiers (Accuracy and AUC values ranging from 73.3 to 83.9% and 0.79 to 0.87) for the prediction of version to version source code reusability over the four selected datasets.

Moreover, with the results of the Friedman test, it was statistically clarified, at a confidence interval of 95%, that only one meta-classifier –"Vote" significantly outperforms the LR technique in giving the best performance results, followed by the LR technique. Rest all of the meta-classifiers perform poorly as compared to the LR technique, thus establishing that the LR technique and six classifiers(AdaBoost, Bagging, Filtered, Multi-class (M-Class), Random Sub Space (RSS), and Voting) out of the seven selected meta-classifiers are indeed effective for developing fitting and authentic version to version source code reusability prediction models.

Future work may involve the replication of the selected meta-classifiers on other similar software datasets for the purpose of yielding generalized results. Application of other prediction models like deep learning could also be done to establish their pertinence in the development of software reusability models.

References

1. Mojica, I.J., Adams, B., Nagappan, M., Dienst, S., Berger, T., Hassan, A.E.: A large-scale empirical study on software reuse in mobile apps. IEEE Softw. **31**(2), 78–86 (2014)
2. Mockus, A.: Large-scale code reuse in open source software. In: First International Workshop on Emerging Trends in FLOSS Research and Development, pp. 1–7. IEEE (2007)
3. Washizaki, H., Koike, T., Namiki, R., Tanabe, H.: Reusability metrics for program source code written in C language and their evaluation. Product-Focused Software Process Improvement. Springer, 89–103 (2012)
4. Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: Data Mining: Practical machine learning tools and techniques. 4th edn. Morgan Kaufmann (2016)
5. Kotsiantis, S.B., Zaharakis, I., Pintelas, P.: Supervised machine learning: a review of classification techniques. Informatica **31**, 249–268 (2007)
6. Peng, C.Y.J., Lee, K.L., Ingersoll, G.M.: An introduction to logistic regression analysis and reporting. J. Education. Res. **96**(1), 3–14 (2002)
7. Hall, T., Beecham, S., Bowes, D., Gray, D., Counsell, S.: A systematic literature review on fault prediction performance in software engineering. IEEE Trans. Softw. Eng. **38**(6), 1276–1304 (2012)
8. Moser, R., Pedrycz, W., Succi, G.: A comparative analysis of the efficiency of change metrics and static code attributes for defect prediction. In: Proceedings of the 30th International conference on Software engineering. pp. 181–190. ACM (2008)
9. Mijač, M., Stapić, Z.: Reusability metrics of software components: survey. In: Proceedings of the 26th Central European Conference on Information and Intelligent Systems. pp. 221–231 (2015)
10. Hudaib, A., Huneiti, A., Othman, I.: Software Reusability classification and predication using self-organizing map (SOM). Commun. Netw. **8**, 179–192 (2016)
11. Padhy, N., Satapathy, S., Singh, R.P.: Utility of an Object Oriented Reusability Metrics and Estimation Complexity. Indian J. Sci. Technol. **8**(1), 1–9 (2017)

12. Devine, T., Goseva-Popstojanova, K., Krishnan, S., Lutz, R.R.: Assessment and cross-product prediction of software product line quality: accounting for reuse across products, over multiple releases. Automat. Softw. Eng. **23**(2), 253–302 (2016)
13. Manhas, S., Sandhu, P.S., Chopra, V., Neeru, N.: Identification of reusable software modules in function oriented software systems using neural network based technique. World Acad. Sci. Eng. Technol. **43**, 823–827 (2010)
14. Sharma, A., Grover, P.S., Kumar, R.: Reusability assessment for software components. ACM SIGSOFT Softw. Eng. Notes **34**(2), 1–6 (2009)
15. Shri, A., Sandhu, P.S., Gupta, V., Anand, S.: Prediction of reusability of object oriented software systems using clustering approach. World Acad. Sci. Eng. Technol. **43**, 853–856 (2010)
16. Chidamber, S.R., Kemerer, C.F.: A metrics suite for object oriented design. IEEE Trans. Softw. Eng. **20**(6), 476–493 (1994)
17. Martin, R.C.: Agile Software Development: Principles, Patterns, and Practices. Prentice Hall (2002)
18. Lorenz, M., Kidd, J.: Object-oriented software metrics: a practical guide. Prentice-Hall, Inc. (1994)
19. McCabe, T.J.: A complexity measure. IEEE Trans. Softw. Eng. **4**, 308–320 (1976)
20. ACNP Software. http://www.anticutandpaste.com/antiplagiarist/. Accessed 15 May 2017
21. Rokach, L.: Ensemble-based classifiers. Artif. Intell. Rev. **33**(1), 1–39 (2010)
22. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai **14**(2), 1137–1145 (1995)
23. Gibbons, J.D., Chakraborti, S.: Nonparametric Statistical Inference. International Encyclopedia of Statistical Science. Springer, Berlin Heidelberg (2011)
24. Buchgeher, G., Weinreich, R.: Integrated software architecture management and validation. In: Proceedings of the 3rd International Conference on Software Engineering Advances. pp. 427–436. IEEE (2008)
25. Scandariato, R., Walden, J.: Predicting vulnerable classes in an Android application. In: Proceedings of the 4th international workshop on Security measurements and metrics. pp. 11–16. ACM (2012)

Non-head-on Non-overtaking Collision of Two Solitary Waves in a Multicomponent Plasma

Tapas Kumar Maji, Malay Kumar Ghorui and Prasanta Chatterjee

Abstract The finding in this chapter is that two oblique solitary waves collide in different angles. The KdV equations are derived, and the collisions figures are drawn with different angles. The multicomponent plasma here is constituted by +ve and −ve charged dust and hydrogen ions, and remaining components are hot and cold electrons. The solutions of the KdV equations are calculated by using Hirota's method. The phase shifts are derived. The effects of density ratio and temperature ratio have investigated on the phase shifts. There are many areas to use these results like study on mesosphere, modern research in plasmas, etc.

Keywords Oblique collision · Phase shift · Hirota's bilinear method · KdV equation

1 Introduction

Nowadays, a fruitful space in plasma physics is nonlinear wave propagation. In plasma research, wave–wave collision (head-on, overtaking, oblique) [1–4] has a great importance in the field of plasma dynamics. Dust ion-acoustic solitons (DIASs) [5, 6] describe the valuable features of non-linearity in contemporary plasma research. They appear due to balance between non-linearity and dispersion. Parveen et al. [7] studied face-to-face interaction between a pair of DIAWs and found that for +ive dust charged grains, compressive solitary waves exist and for the opposite dust charged grains compressive as well as rarefactive solitary waves exist. In this paper, multicomponent plasma consists of five components, namely, hydrogen ion, positive dust, negative dust, hot electrons and cold electrons. Research in multicomponent is

T. K. Maji (✉) · M. K. Ghorui
Department of Mathematics, B. B. College, Ushagram, Asansol 713303, India
e-mail: tapasmaji.bwn@gmail.com

T. K. Maji · P. Chatterjee
Department of Mathematics, Siksha-Bhavana, Visva-Bharati University,
Santiniketan 731235, India

a developing diverse field. With various types of masses and sizes, dust particles are available. After electron–ion collision on dusts, photoionization, thermionic emission, etc. dust particles become +ive and −ive. Existence of dust particles causes a significant difference in the physical characteristics such as charge, potential, etc. of the plasmas.

Chatterjee et al. [8] concluded that the structure of the solitary wave rely on the external magnetic field, Mach number, etc. while studying obliquely promulgating ion-acoustic solitary waves using Sagdeev potential in a magnetized plasma considering of anisotropic pressure. Hafez et al. [4] found that in an oblique propagation of ion-acoustic shock waves the phase velocity is decreasing with increase in positron concentration and electron to positron temperature ratio. Maji et al. [9] investigated that the phase shift in an oblique collision of two ion-acoustic waves in a multicomponent plasma is effected by temperature ratio and density ratio. The KdV equation is extracted using extended PLK perturbation method [10, 11] and Hirota's method [12] for soliton solutions.

We arrange this paper into four sections. Introduction part is in Sect. 1; in the second section, we extracted KdV equations. Next, in the third, we keep the numerical results and discussions one- and two-soliton solutions. At the end section, we give the conclusion.

2 Basic Equations

The interested area here is dust ion-acoustic solitary waves in a multicomponent plasma. Dust is weightier than the ion components though both are considered as cold, while both the electrons are governed by [13]

$$n_s = n_{s0}[1 + \frac{e_s \phi}{k_B T_s (k_s - 3/2)}]^{-k_s + 1/2}. \tag{1}$$

In Eq. (1), subscript s denotes the species ($s = se$ for solar electrons and s = ce for cometary photo-electrons). Density is described by n. The subscript 0 denotes the equilibrium value, e_s the charge, T_s the temperature, and k_s the spectral index for the species 's'. k_B is the Boltzmann's constant, and ϕ is the potential.

The normalized basic set of equations are given below:

$$\frac{\partial n_i}{\partial t} + \nabla(n_i u_i) = 0, \tag{2}$$

$$\frac{\partial n_1}{\partial t} + \nabla(n_1 u_1) = 0, \tag{3}$$

$$\frac{\partial n_2}{\partial t} + \nabla(n_2 u_2) = 0, \tag{4}$$

$$\frac{\partial u_i}{\partial t} + (u_i \nabla)u_i = -\alpha_i \beta_i \nabla \phi + \Omega_i(u_{iz}\hat{y} - u_{iy}\hat{z}), \tag{5}$$

$$\frac{\partial u_1}{\partial t} + (u_1 \nabla) u_1 = \nabla \phi - \Omega_1 (u_{1z} \hat{y} - u_{1y} \hat{z}), \qquad (6)$$

$$\frac{\partial u_2}{\partial t} + (u_2 \nabla) u_2 = -\alpha_2 \beta_2 \nabla \phi + \Omega_2 (u_{2z} \hat{y} - u_{2y} \hat{z}), \qquad (7)$$

$$\nabla^2 \phi = n_1 - (1 - z_i u_i + \mu_{ce} + \mu_{se}) n_2 - (1 - z_2 u_2 + \mu_{ce} + \mu_{se}) n_i +$$

$$\mu_{ce} (1 - \frac{\phi}{\sigma_{ce}(k_{ce} - 3/2)})^{-k_{ce}+1/2} + \mu_{se} (1 - \frac{\phi}{\sigma_{se}(k_{se} - 3/2)})^{-k_{se}+1/2} \qquad (8)$$

In Eqs. (2)–(8), n_i, n_1, and n_2 are the hydrogen ion, negative and positive dust number densities, normalized by their equilibrium values n_{i0}, n_{10}, and n_{20}, respectively. u_i, u_1, and u_2 are the hydrogen ion, negatively and positively charged dust fluid speeds normalized by $(\frac{z_1 k_B T_1}{m_1})^{1/2}$. The electric potential ψ is normalized by $(\frac{k_B T_1}{e})$. x and t are normalized by the Debye length $(\frac{z_1 k_B T_1}{4\Pi z_1^2 e^2 n_{10}})^{1/2}$ and $(\frac{m_1}{4\Pi z_1^2 e^2 n_{10}})^{1/2}$, respectively. The cyclotron frequencies of the different species are normalized by the plasma frequency, i.e. $(\frac{4\Pi z_1^2 e^2 n_{10}}{m_1})^{1/2}$. Also $\alpha_i = \frac{z_i}{z_1}, \alpha_2 = \frac{z_2}{z_1}, \beta_i = \frac{m_1}{m_i}, \beta_2 = \frac{m_1}{m_2}, \mu_s = \frac{n_{s0}}{z_1 n_{10}}, \sigma_s = \frac{T_s}{T_1}$, where n_{s0} is the equilibrium density for species s. T_s, T_1 are the temperatures of the species s and the negative dust, respectively. z_i, z_1, z_2 are, respectively, the charge numbers of hydrogen ions, negatively and positively charged dust particles while m_i, m_1, m_2 are the corresponding masses.

Here we think about two infinitesimal as well as finite amplitude wave-like perturbations promulgating obliquely with an angle θ_{11} in the 3D surface with distinct velocities, leaving a phase shift after they came nearer, contact once in a while and then they exit yielding a collision mark on waves causing a phase shift. The phase variables are involved in the asymptotic extension of plasma variables around thermodynamics equilibrium state in a stretched coordinate with the phase variables. This method is known as an extended PLK method. In utilizing this method, we launch additional coordinates as follows:

$$\xi = \varepsilon^{1/2}(l_x x + l_y y + l_z z - \lambda_1 t) + \varepsilon P^{(0)}(\eta, \tau) + \varepsilon^{3/2} P^{(1)}(\xi, \eta, \tau) \dots \qquad (9)$$

$$\eta = \varepsilon^{1/2}(k_x x + k_y y + k_z z - \lambda_2 t) + \varepsilon Q^{(0)}(\xi, \tau) + \varepsilon^{3/2} Q^{(1)}(\xi, \eta, \tau) \dots \qquad (10)$$

$$\tau = \varepsilon^{3/2} t, \qquad (11)$$

where $l_x^2 + l_y^2 + l_z^2 = 1, k_x^2 + k_y^2 + k_z^2 = 1$, where ε is an infinitesimal parameter which characterizes the relative amplitude of the excitation. Here λ_1, λ_2 are the phase velocities to be calculated later. We have to obtain the mutually related dependent functions $P^{(0)}$ and $Q^{(0)}$ referring the phase description of the colliding waves in space–time coordinate. The incipient wave velocities are described by the vectors $\mathbf{r}_{11} = (l_{x1}, l_{y1}, l_{z1})$ and $\mathbf{r}_{22} = (k_{x1}, k_{y1}, k_{z1})$ and the collision acts through an angle θ_{11} give by

$$\cos\theta_{11} = \frac{\mu}{\zeta_1\zeta_2}; \quad \mu = (l_xk_x + l_yk_y + l_zk_z)$$

$$\zeta_1 = (l_x^2 + l_y^2 + l_z^2)^{1/2} \quad ; \zeta_2 = (k_x^2 + k_y^2 + k_z^2)^{1/2}. \tag{12}$$

Here normalized wave numbers are described by ζ_1 and ζ_2. Also $\theta_{11} \neq 0$ or $\cos\theta_{11} \neq 1$. This method would fail if $\theta_{11} = 0$. The expansion of the dependent plasma variables are as follows:

$$n_k = 1 + \varepsilon n_k^{(1)} + \varepsilon^2 n_k^{(2)} + \cdots \tag{13}$$

$$u_{kx} = \varepsilon u_{kx}^{(1)} + \varepsilon^2 u_{kx}^{(2)} + \cdots \tag{14}$$

$$u_{ky} = \varepsilon^{3/2} u_{ky}^{(1)} + \varepsilon^2 u_{ky}^{(2)} + \cdots \tag{15}$$

$$u_{kz} = \varepsilon^{3/2} u_{kz}^{(1)} + \varepsilon^2 u_{kz}^{(2)} + \cdots \tag{16}$$

$$\phi = \varepsilon\phi^{(1)} + \varepsilon^2\phi^{(2)} + \ldots \tag{17}$$

where $k = i$ (ions), 1 (negative dust), and 2 (positive dust).

Substituting Eqs. (9)–(11) and (13)–(17) into Eqs. (2)–(8), we obtain the following equations; comparing the orders of ε, we obtain, after tedious calculations, the KdV equation:

$$\frac{\partial\phi_{1\xi}}{\partial\tau} + A\left(\phi_{1\xi}\frac{\partial\phi_{1\xi}}{\partial\xi}\right) + B\frac{\partial^3\phi_{1\xi}}{\partial\xi^3} = 0, \tag{18}$$

where $A = \frac{A_{11}+A_{22}}{2\lambda_1 l_x^2(1+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i)}$, with $A_{11} = 3l_x^4(-1 + (1 - z_i\mu_i + \mu_{ce} + \mu_{se})\alpha_2^2\beta_2^2 + (1 - z_2\mu_2 + \mu_{ce} + \mu_{se})\alpha_i^2\beta_i^2)$, $A_{22} = \lambda_1^4\left(\frac{\mu_{ce}(\kappa_{ce}-1/2)(\kappa_{ce}+1/2)}{\sigma_{ce}^2(\kappa_{ce}-3/2)^2}\right.$

$+ \left.\frac{\mu_{se}(\kappa_{se}-1/2)(\kappa_{se}+1/2)}{\sigma_{se}^2(\kappa_{se}-3/2)^2}\right)$, and $B = \frac{\lambda_1^3(1+(\frac{1-l_x^2}{\Omega_1^2})+(1-z_i\mu_i+\mu_{ce}+\mu_{se})((\frac{(1-l_x^2)\alpha_2\beta_2}{\Omega_2^2}))+(1-z_2\mu_2+\mu_{ce}+\mu_{se})(\frac{(1-l_x^2)\alpha_i\beta_i}{\Omega_i^2}))}{2l_x^2(1+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i)}$,

$\lambda_1^2 = \frac{l_x^2+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2 l_x^2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i l_x^2}{\frac{\mu_{ce}(\kappa_{ce}-1/2)}{\sigma_{ce}(\kappa_{ce}-3/2)}+\frac{\mu_{se}(\kappa_{se}-1/2)}{\sigma_{se}(\kappa_{se}-3/2)}}$.

Similarly for η terms we get,

$$\frac{\partial\phi_{1\eta}}{\partial\tau} + A_1\left(\phi_{1\eta}\frac{\partial\phi_{1\eta}}{\partial\eta}\right) + B_1\frac{\partial^3\phi_{1\eta}}{\partial\eta^3} = 0, \tag{19}$$

where $A_1 = \frac{A_{33}+A_{44}}{2\lambda_2 k_x^2(1+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i)}$, $A_{33} = 3k_x^4(-1 + (1 - z_i\mu_i + \mu_{ce} + \mu_{se})\alpha_2^2\beta_2^2 + (1 - z_2\mu_2 + \mu_{ce} + \mu_{se})\alpha_i^2\beta_i^2)$, $A_{44} = \lambda_2^4\left(\frac{\mu_{ce}(\kappa_{ce}-1/2)(\kappa_{ce}+1/2)}{\sigma_{ce}^2(\kappa_{ce}-3/2)^2}\right.$ +

$\left.\frac{\mu_{se}(\kappa_{se}-1/2)(\kappa_{se}+1/2)}{\sigma_{se}^2(\kappa_{se}-3/2)^2}\right)$, $B_1 = \frac{\lambda_2^3(1+(\frac{1-k_x^2}{\Omega_1^2})+(1-z_i\mu_i+\mu_{ce}+\mu_{se})((\frac{(1-k_x^2)\alpha_2\beta_2}{\Omega_2^2}))+(1-z_2\mu_2+\mu_{ce}+\mu_{se})(\frac{(1-k_x^2)\alpha_i\beta_i}{\Omega_i^2}))}{k_x^2(1+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i)}$.

$\lambda_2^2 = \frac{k_x^2+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2 k_x^2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i k_x^2}{\frac{\mu_{ce}(\kappa_{ce}-1/2)}{\sigma_{ce}(\kappa_{ce}-3/2)}+\frac{\mu_{se}(\kappa_{se}-1/2)}{\sigma_{se}(\kappa_{se}-3/2)}}$.

Therefore, we obtain the following equations:

$$\frac{\partial P^{(0)}}{\partial \eta} = C\phi_{1\eta} \tag{20}$$

$$\frac{\partial Q^{(0)}}{\partial \eta} = D\phi_{1\xi}, \tag{21}$$

where $C = \frac{C_{11}+C_{22}}{\lambda_1(\lambda_2 l_x^2 - l_x k_x \lambda_1)((1+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i))}$ with $C_{11} = l_x^4(1 +$
$(1 - z_i\mu_i + \mu_{ce} + \mu_{se})\alpha_2^2\beta_2^2 + (1 - z_2\mu_2 + \mu_{ce} + \mu_{se})\alpha_i^2\beta_i^2)$, $C_{22} = \lambda_1(\frac{\mu_{ce}(\kappa_{ce}-1/2)(\kappa_{ce}+1/2)}{\sigma_{ce}^2(\kappa_{ce}-3/2)^2} +$
$\frac{\mu_{se}(\kappa_{se}-1/2)(\kappa_{se}+1/2)}{\sigma_{se}^2(\kappa_{se}-3/2)^2}))$ and $D = \frac{D_{11}+D_{22}}{\lambda_1(\lambda_2 k_x^2 - l_x k_x \lambda_1)((1+(1-z_i\mu_i+\mu_{ce}+\mu_{se})\alpha_2\beta_2+(1-z_2\mu_2+\mu_{ce}+\mu_{se})\alpha_i\beta_i))}$, with
$D_{11} = k_x^4(1 + (1 - z_i\mu_i + \mu_{ce} + \mu_{se})\alpha_2^2\beta_2^2 + (1 - z_2\mu_2 + \mu_{ce} + \mu_{se})\alpha_i^2\beta_i^2)$,
$D_{22} = \lambda_1(\frac{\mu_{ce}(\kappa_{ce}-1/2)(\kappa_{ce}+1/2)}{\sigma_{ce}^2(\kappa_{ce}-3/2)^2} + \frac{\mu_{se}(\kappa_{se}-1/2)(\kappa_{se}+1/2)}{\sigma_{se}^2(\kappa_{se}-3/2)^2}))$.

Equations (18) and (19) represent two KdV equations. Theses two waves are promulgating in the ξ and η directions, respectively, with an angle θ_1. One-soliton solution of the KdV equation (18) is extracted by using Hirota's method [12]

$$\phi_{1\xi} = \frac{12B}{A} \frac{\partial^2}{\partial \xi^2}(logf); \tag{22}$$
$$f = 1 + e^{\theta_1}; \quad \theta_1 = kB^{-1/3}\xi - k^3\tau + \delta.$$

The leading phase changes due to the collision can be calculated from Eqs. (20) and (21) produce the main phase changes for the collisions. Thus, from Eq. (21), we have

$$\frac{\partial Q^{(0)}(\xi, \tau)}{\partial \xi} = \frac{12BD}{AC} \frac{\partial^2}{\partial \xi^2}(logf),$$
$$or \ Q^{(0)}(\xi, \tau) = \frac{12BD}{AC} \frac{\partial}{\partial \xi}(logf) = \frac{12B^{2/3}D}{AC} \frac{ke^{\theta_1}}{1+e^{\theta_1}}, \tag{23}$$

and the corresponding phase shift is

$$\triangle Q^{(0)} = \varepsilon(k_x x + k_y y + k_z z - \lambda_2 t)|_{\xi=-\infty,\eta=0} - \varepsilon(l_x x + l_y y + l_z z - \lambda_1 t)|_{\xi=\infty,\eta=0}$$
$$= \varepsilon^2 Q^{(0)}(\infty, \tau) - \varepsilon^2 Q^{(0)}(-\infty, \tau)$$
$$= \frac{12\varepsilon^2 DB^{2/3}}{AC}k. \tag{24}$$

Similarly, the other phase shift is

$$\triangle P^{(0)} = -\frac{12\varepsilon^2 DB^{2/3}}{AC}k. \tag{25}$$

Phase shifts in Eqs. (24) and (25) are similar to that in the investigations in different plasma models but the approaches are different. Again, each of the KdV equations given by (18) and (19) has many of soliton solutions. We consider here two-soliton solutions of each of the KdV equations. Using Hirota's method [12] two-soliton solution of the KdV, Eq. (18) is given by

$$\phi_{1\xi} = \frac{12B}{A} \frac{\partial^2}{\partial \xi^2}(logg); \tag{26}$$
$$g = 1 + e^{\theta_1} + e^{\theta_2} + a_{12}e^{\theta_1+\theta_2},$$
$$\theta_i = k_i B^{-1/3}\xi - k_i^3 \tau + \delta_i \quad i = 1, 2,$$
$$a_{12} = (k_1 - k_2)^2/(k_1 + k_2)^2.$$

As in the case of one-soliton solution from Eq. (18), we have

$$\frac{\partial Q^{(0)}(\xi, \tau)}{\partial \xi} = \frac{12BD}{AC} \frac{\partial^2}{\partial \xi^2}(logg),$$
$$\Rightarrow \quad Q^{(0)}(\xi, \tau) = \frac{12BD}{AC} \frac{\partial}{\partial \xi}(logg),$$
$$= \frac{12B^{2/3}D}{AC} \frac{k_1 e^{\theta_1} + k_2 e^{\theta_2} + a_{12}(k_1 + k_2)e^{\theta_1+\theta_2}}{1 + e^{\theta_1} + e^{\theta_2} + a_{12}e^{\theta_1+\theta_2}}, \tag{27}$$

and the corresponding phase shift is

$$\triangle Q^{(0)} = \varepsilon(k_x x + k_y y + k_z z - \lambda_2 t)|_{\xi=-\infty,\eta=0} - \varepsilon(l_x x + l_y y + l_z z - \lambda_1 t)|_{\xi=\infty,\eta=0}$$
$$= \varepsilon^2 Q^{(0)}(\infty, \tau) - \varepsilon^2 Q^{(0)}(-\infty, \tau)$$
$$= \frac{12\varepsilon^2 DB^{2/3}}{AC} \frac{a_{12}(k_1 + k_2)}{a_{12}}$$
$$= \frac{12\varepsilon^2 DB^{2/3}}{AC}(k_1 + k_2). \tag{28}$$

Similarly, the other phase shift is

$$\triangle P^{(0)} = -\frac{12\varepsilon^2 DB^{2/3}}{AC}(k_1 + k_2). \tag{29}$$

3 Numerical Results and Discussion

In the present paper, we examine thoroughly the oblique collision of two DIAWs in a multicomponent plasma composed of hydrogen ions and hot electrons, a pair

(a)

(b)
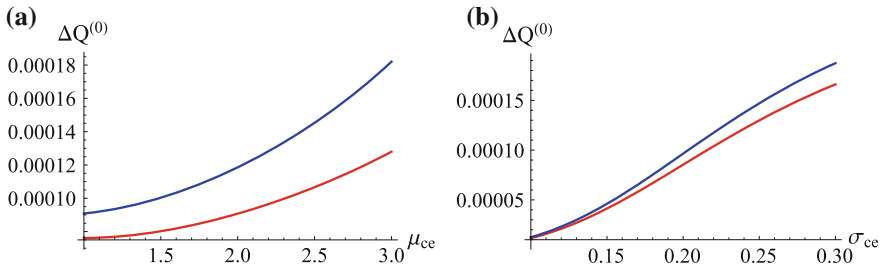
Fig. 1 **a** Variation in phase shift $\triangle Q^{(0)}$ with μ_{ce} when $K1 = 1; K2 = 2; \delta1 = 1.2, \delta2 = 1.2; \varepsilon = 0.01; l_x = 0.1; l_y = 0.3; l_z = 5.33818; k_x = 0.499179; k_y = 0.2; k_z = 0.7; \Theta_{11} = 11/14; \alpha1 = 2.2; \alpha2 = 2; z_i = 0.002, \mu_i = 100, \mu_2 = 50, \alpha_2 = 1, \beta_2 = 2, \alpha_i = 2, \beta_i = 1, \mu_{se} = 0.5, \sigma_{ce} = 0.2, \sigma_{se} = 0.2, \kappa_{ce} = 3.5, \kappa_{se} = 4.5, \Omega_1 = 1, \Omega_2 = 2, \Omega_i = 3$. **b** Variation in phase shift $\triangle Q^{(0)}$ with σ_{ce} when $K1 = 1; K2 = 2; \delta1 = 1.2, \delta2 = 1.2; \varepsilon = 0.01; l_x = 0.1; l_y = 0.3; l_z = 5.33818; k_x = 0.499179; k_y = 0.2; k_z = 0.7; \Theta_{11} = 11/14; \alpha1 = 2.2; \alpha2 = 2; z_i = 0.002, \mu_i = 100, \mu_2 = 50, \alpha_2 = 1, \beta_2 = 2, \alpha_i = 2, \beta_i = 1, \mu_{se} = 0.5, \sigma_{se} = 0.2, \kappa_{ce} = 3.5, \kappa_{se} = 4.5, \Omega_1 = 1, \Omega_2 = 2, \Omega_i = 3$

of oppositely charged heavier dust(ions) and colder electrons. The PLK method is being taken for deriving the two-sided KdV equations and then Hirota's method is being used for deriving its solutions which give the phase shifts for single, double and triple solitons for the two KdV equations. We investigated that the phase shifts are varied due to change density ratio μ_{ce}, temperature ratio σ_{ce}. The phase shifts are increasing moderately with the increase in density ratio μ_{ce}. It is shown in Fig. 1a. Lastly, the phase shifts are increasing rapidly with the increase in temperature ratio σ_{ce}. Figure 1b exhibits this.

The two solitons move in the same direction with an angle θ_{11}. We consider here one and two solitons of each KdV equations. The two solitons for a fixed KdV equation promulgate in the same direction, and the faster moving soliton, at last, overtakes the slower one. This can be realized from the Fig. 2a, b.

In Fig. 2a, b, we investigate the variation in electrostatic potential with space coordinate $r_{11} = (l_{x1}x + l_{y1}y + l_{z1}z), r_{22} = (k_x x + k_y y + k_z z)$ and the time variable t. In Fig. 2a, we have taken the values $\varepsilon = 0.01, l_x = 0.1, l_y = 0.3, l_z = 5.33818, k_x = 0.499179, k_y = 0.2, k_z = 0.7$ and $\theta^{11} = 11/14$. And for Fig. 2b, we have $\varepsilon = 0.01, l_x = 0.1, l_y = 0.3, l_z = -0.658389, k_x = 0.202032, k_y = 0.2, k_z = 0.7,$ and $\theta_{11} = 22/21$. From Fig. 2a, b we realize that at the time of oblique collision a statistical structure of the two DIAWS has been created and then a new non-linear wave been initiated at the time of collision. It is observed that for some time interval during oblique collision, one practically motionless composite structure is formed both its amplitude and width are greater than of the colliding solitary waves.
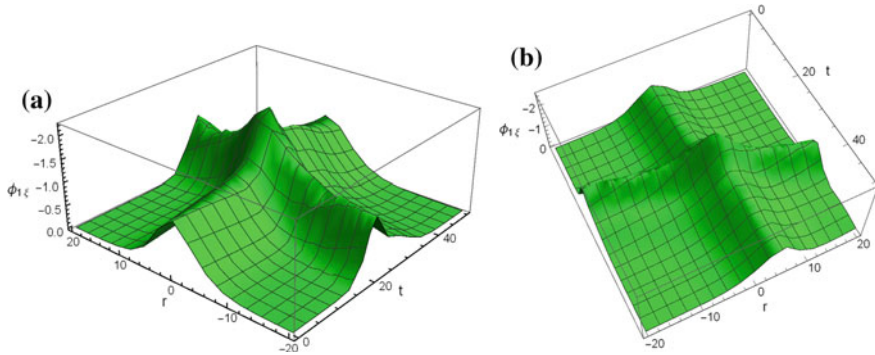
Fig. 2 **a** Oblique collision of two solitary waves when $k_1 = 1; k_2 = 2; k_3 = 5; k_4 = 6; \delta 1 = 1.2, \delta 2 = 1; \varepsilon = 0.01; l_x = 0.1; l_y = 0.3; l_z = 5.33818; k_x = 0.499179; k_y = 0.2; k_z = 0.7; \Theta_{11} = 11/14; \alpha 1 = 2.2; \alpha 2 = 2; z_i = 0.002, \mu_i = 100, \mu_2 = 50, \alpha_2 = 1, \beta_2 = 2, \alpha_i = 2, \beta_i = 1, \mu_{se} = 0.5, \mu_{ce} = 0.5, \sigma_{ce} = 0.2, \sigma_{se} = 0.2, \kappa_{ce} = 3.5, \kappa_{se} = 4.5, \Omega_1 = 1, \Omega_2 = 2, \Omega_i = 3$. **b** Oblique collision of two solitary waves when $K1 = 1; K2 = 2; K3 = 5; K4 = 6; \delta 1 = 1.2, \delta 2 = 1; \varepsilon = 0.01; l_x = 0.1; l_y = 0.3; l_z = -0.658389; k_x = 0.202032; k_y = 0.2; k_z = 0.7; \Theta_{11} = 22/21; \alpha 1 = 2.2; \alpha 2 = 2; z_i = 0.002, \mu_i = 100, \mu_2 = 50, \alpha_2 = 1, \beta_2 = 2, \alpha_i = 2, \beta_i = 1, \mu_{se} = 0.5, \mu_{ce} = 0.5, \sigma_{ce} = 0.2, \sigma_{se} = 0.2, \kappa_{ce} = 3.5, \kappa_{se} = 4.5, \Omega_1 = 1, \Omega_2 = 2, \Omega_i = 3$

4 Conclusion

We have worked on the oblique collision of two dust ion-acoustic waves (IAWs) in a multicomponent plasma system consisting of hot electrons, cold electrons, positive dust, negative dust and hydrogen ion. We have used the PLK method to derive the two-sided KdV equations and using Hirota's method we derive soliton solutions. The phase shifts are increasing moderately with the increase in density ratio μ_{ce}. Lastly, the phase shifts are increasing rapidly with the increase in temperature ratio σ_{ce}. Finally, we observed that at the time of oblique collision a statistical structure of the two DIAWS has been created and then a new non-linear wave been initiated at the time of collision. Thus after collision of DIAWs, both the amplitude and width are larger than that of the colliding solitary waves.

References

1. Roy, K., Chatterjee, P., Roychowdhury, R.: Phys. Plasmas **21**, 104509 (2014)
2. Roy, K., Maji, T.K., Ghorui, M.K., Chatterjee, P., Roychowdhury, R.: Astrophys. Space Sci. **352**, 151–157 (2014)
3. Gardner, C.S., Greener, J.M., Kruskal, M.D., Miura, R.M.: Phys. Rev. Lett. **19**, 1095 (1967)

4. Hafez, M.G., Roy, N.C., Talukdar, M.R., Hossain Ali, M.: Astrophys. Space Sci. **361**, 312 (2016)
5. El-Labany, S.K., El-Taibany, W.F., Behery, E.E., Fouda, S.M.: Phys. Plasmas **24**, 033713 (2017)
6. Xue, J.K.: Chin. Phys. Soc. **15**(3), 562 (2006)
7. Parveen, S., Mahmood, S., Adnan, M., Quamar, A.: Phys. Plasmas **23**, 092122 (2016)
8. Chatterjee, P., Saha, T., Ryu, C.M.: Phys. Plasmas **15**, 123702 (2008)
9. Maji, T.K., Ghorui, M.K., Saha, A., Chatterjee, P.: Braz. J. Phys. **47**, 295–301 (2017)
10. Sue, C.H., Mirie, R.M.: J. Fluid Mech. **509** (1980)
11. Jeffery, A., Kawahawa, T.: Asymptotic Methods in Nonlinear Wave Theory. Pitman, London (1982)
12. Hirota, R.: The Direct Method in the Soliton Theory. Cambridge University Press, Cambridge (2004)
13. Sijo, S., Manesh, M., Sreekala, G., Neethu, T.W., Renuka, G., Venugopal, C.: Phys. Plasmas **22**, 123704 (2015)

Signed Product and Total Signed Product Cordial Labeling of Cartesian Product Between Balanced Bipartite Graph and Path

Sumonta Ghosh and Anita Pal

Abstract Nowadays, signed product cordial labeling and total signed product cordial labeling play an important role in graph labeling. For the graph $G = (V, E)$, all the vertices are label by 1 or -1 with the restriction that the difference between total number of vertices labeled by 1 and -1 is less than equal to 1, where edge of the graph $G = (V, E)$ labeled by the multiplication of the label of the end vertices also with the restriction that the difference between total number of edges labeled by 1 and -1 is less than equal to 1. We investigate some interesting result under the above labeling scheme on some new type of complex graph. In this paper, we apply signed product cordial labeling and total signed product cordial labeling on the graph obtained by Cartesian product between balanced bipartite graph $K_{n,n}$ and path P_r. We have shown that the time complexity of the algorithms is superlinear.

1 Introduction

Graph labeling has so many different applications in real world. Gallian [1] has published a productive survey on entire graph labeling and their different applications. According to Hegde [2], graph labeling problem is described in such a way where a set of values from which the labels are taken under a restricted environment and a procedure by which each edge label by a value under some efficient condition. Beineke and Hegde [3] explained graph labeling is an extreme limit between graphs structure and number theory.

S. Ghosh (✉) · A. Pal
Department of Mathematics, NIT Durgapur, Durgapur 713209, West Bengal, India
e-mail: mesumonta@gmail.com

A. Pal
e-mail: anita.buie@gmail.com

The technique by which a graph is labeled can be applied on coding theory, missile guidance code, etc., also the labeling techniques can be successfully applied on X-ray crystallography and communication network also. Cahit [4, 5] has given the idea of both graceful [6] and harmonious [7] labelings. Cahit [5] proved the followings: tree is a cordial graph, K_n is cordial graph if and only if $n \leq 3$, $K_{m,n}$ is cordial graph $\forall m, n$, friendship graph is also a cordial graph with some restrictions, all fans are cordial, etc.

The main idea of signed product cordial labeling was initiated by Babujee and Shobana [8] and he already proved that most of the graphs follow signed product cordial labeling. This concept is explored by many researchers like Ulaganathan et al. [9], Ho et al. [10], and Santhi [11]. The concept of cordial language and cordial numbers introduced by Babujee and Shobana [8]. Most of the researchers pay their attention in signed product cordial labeling on simple network as well as complex network. We investigate some interesting result on the above labeling on some new complex graphs.

As time goes, utilization of systems becomes very high, which experienced wider and complex network structure. Connection of different types of network model plays a vital role in real life, so the product of two existing network model gives a complex network structure with the facility of single integrated network. Such complex networks become more reliable (possibly at high cost). In this paper, we mainly concentrate on signed product and total signed product cordial labeling of Cartesian product between balanced bipartite graph and path.

The remaining part of this paper is arranged in the following way. Section 2 contains some important and related definitions, and Sect. 3 presents algorithms to label Cartesian product between balanced bipartite graph and path, proof of correctness of each algorithm, followed by conclusion.

2 Preliminaries

Definition 1 A vertex labeling of a graph $G = (V, E)$, where V denote the set of vertices and E denote the set of edges. f be the function where $f : V \rightarrow \{1, -1\}$ with induced edge labeling $f^* : E \rightarrow \{1, -1\}$ defined by $f^*(xy) = f(x)f(y)$ is called a *signed product cordial labeling* if the difference between total number of vertices labeled by 1 and total number of vertices labeled by -1 is less than equal to 1, and the difference between total number of edges labeled by 1 and total number of edges labeled by -1 is less than equal to 1, i.e., $|v_f(1)v_f(-1)| \leq 1$ and $|e_f^*(1) - e_f^*(-1)| \leq 1$. f is called *total signed product cordial labeling* if $|e_f^*(-1) + v_f(-1) - e_f^*(1) - v_f(1)| \leq 1$, a graph which follows *total signed product cordial labeling* is known as *total signed product graph*.

Definition 2 *Cartesian product* of two graphs $G = (V, E)$ with vertex set V and edge set E and $H = (V', E')$ with vertex set V' and edge set E' is the *Cartesian product* between two set of vertices $V(G) \times V'(G')$ denoted by $G \times H$, where (u, u') and (v, v')

are the order pair of the *Cartesian product* and will be adjacent in $G \times H$ if and only if either

1. $u = v$ and u' and v' are adjacent in G', or
2. $u' = v'$ and u and v are adjacent in G.

Commutative law holds for the above graphs.

Definition 3 A graph $G = (V, E)$ with vertex set V and edge set E is called a *complete bipartite graph* if its vertices can be divided into two subsets V_1 and V_2 such that no end points are in the same set for each edge $e \in E$, and each vertex of $V_1(V_2)$ is connected with all vertices of $V_2(V_1)$. Here $V_1 = \{X_{11}, X_{12}, \ldots, X_{1m}\}$ contains m vertices and $V_2 = \{Y_{11}, Y_{12}, \ldots, Y_{1n}\}$ contains n vertices. A *complete bipartite graph* with $|V_1| = m$ and $|V_2| = n$ is denoted by $K_{m,n}$. When $|V_1| = |V_2|$ then *complete bipartite graph* is known as *balanced bipartite graph* and which is denoted by $K_{n,n}$.

Definition 4 A *walk* of length k in a graph $G = (V, E)$ is an alternating sequence of vertices and edges, $v_0, e_0, v_1, e_1, v_2, \ldots, v_{k-1}, e_{k-1}, v_k$ which begins and ends with vertices. If the graph is directed, then e_i is a directed arc from v_i to v_{i+1}, and it may form a *trial* if any edge is traversed at most once.

A *trial* is a *path* in which all vertices (except possibly the first and last) are distinct.

3 Labeling of Cartesian Product Between Balanced Bipartite Graph and Path

3.1 Algorithm

We consider balanced bipartite graph $K_{n,n}$. Assuming that each $K_{n,n}$ having two set of vertices X_i and Y_i for $i = 1, 2, 3, \ldots, r$ and $|X_i| = |Y_i| = n$ where the vertices within the set are not connected but every vertex of X_i is connected with every vertices of Y_i. Already it is clear that X_i has n vertices and the vertices are $X_i = \{x_{i1}, x_{i2}, x_{i3}, \ldots, x_{ij}, \ldots, x_{in}\}$ similarly $Y_i = \{y_{i1}, y_{i2}, y_{i3}, \ldots, y_{ij}, \ldots, y_{in}\}$.

Here we propose three algorithms, Algorithms 1 and 2 are for signed product cordial labeling and total signed product cordial labeling of Cartesian product between $K_{n,n} \times P_3$ and $K_{n,n} \times P_r$, respectively (n is odd and $r > 3$), and Algorithm 3 is for signed product cordial labeling of Cartesian product between $K_{n,n} \times P_r$ (n is even). For Cartesian product between $K_{n,n} \times P_r$ we have to draw r copies of $K_{n,n}$. Algorithm 2 is working for path length $r > 3$ and it will label the graph obtained by Cartesian product between $K_{n,n} \times P_r$ upto multiple of four copies of $K_{n,n}$ and for rest of the copies of $K_{n,n}$, we follow Algorithm 1. Algorithm 3 can be applied for any path length r and even value of n.

Throughout the algorithm, we consider $f(v_i)$ as label of the vertex v_i. According to the drawing of Cartesian product between $K_{n,n} \times P_r$, it is clear that there will be r copies of $K_{n,n}$. Basically, we have to label each copy of $K_{n,n}$ by 0 and 1 starting from 1st, 2nd, 3rd, ... upto rth copy.

3.2 Algorithm

3.2.1 The Proof of Correctness of the Algorithm 1 Is Given Below

Theorem 1 *Algorithm 1 correctly labels the graph $K_{n,n} \times P_3$ by signed product and total signed product cordial labeling.*

Proof Considering the graph $G = K_{n,n} \times P_3$ when n is odd, where G has total number of vertices is equal to $6n$ and total number of edges is equal to $3n^2 + 4n$. According to Algorithm 1, total number of vertices label by -1 is $3n$ and total number of vertices label by 1 is $3n$, i.e., $V_f(-1) = V_f(1) = 3n$. So $|V_f(-1) - V_f(1)| = |3n - 3n| = 0 \leq 1$. Now total number of edges label by -1 is $\frac{(3n^2+4n-1)}{2}$ and total number of edges label by 1 is $\frac{(3n^2+4n+1)}{2}$, i.e., $|e_{f*}(-1) - e_{f*}(1)| = |\frac{(3n^2+4n-1)}{2} - \frac{(3n^2+4n+1)}{2}| = 1 \leq 1$, which follow the restriction of signed product cordial labeling. Now $|e_{f*}(-1) + v_f(-1) -$

Algorithm 1 Algorithm for signed product and total signed product cordial labeling of cartesian product between $K_{n,n} \times P_3$ (SPTSPCLBP_3)

Input: A Cartesian product between $K_{n,n}$ and P_3 i.e., $G = (K_{n,n} \times P_3)$, where n is odd.
Output: Signed product and total signed product cordial labeling of the graph $G = (K_{n,n} \times P_3)$.
Step 1 for $i = 1$ to 3
Step 2 if $i = 1(mod\ 2)$ then
Step 3 for $j = 1$ to n

1. $f(x_{ij}) = -1$ if $sum(i, j) = 0(mod\ 2)$.
2. $f(y_{ij}) = 1$ if $sum(i, j) = 0(mod\ 2)$.
3. $f(x_{ij}) = 1$ if $sum(i, j) = 1(mod\ 2)$.
4. $f(y_{ij}) = -1$ if $sum(i, j) = 1(mod\ 2)$.

end of loop.
Step 4 if $i = 0(mod\ 2)$ then
Step 5 for $j = 1$ to $(n - 1)$

1. $f(x_{ij}) = 1$ if $sum(i, j) = 0(mod\ 2)$.
2. $f(y_{ij}) = 1$ if $sum(i, j) = 0(mod\ 2)$.
3. $f(x_{ij}) = -1$ if $sum(i, j) = 1(mod\ 2)$.
4. $f(y_{ij}) = -1$ if $sum(i, j) = 1(mod\ 2)$.

end of loop.
Step 6 $f(x_{in}) = -1$ and $f(y_{in}) = 1$
end of loop.
Stop.

$e_{f^*}(1) - v_f(1)| = |\frac{(3n^2+4n-1)}{2} + 3n - \frac{(3n^2+4n+1)}{2} - 3n| = 1 \leq 1$ follows the restriction of total signed product cordial labeling.

3.3 Algorithm

3.3.1 The Proof of Correctness of the Algorithm 2 Is Given Below

Theorem 2 *Algorithm 2 correctly labels the graph $K_{n,n} \times P_r$ (where n is odd) by signed product and total signed product cordial labeling.*

Algorithm 2 Algorithm for signed product and total signed product cordial labeling of cartesian product between $K_{n,n} \times P_r$ (SPTSPCLBP(O))

In this algorithm we consider the graph obtained by the Cartesian product between balanced bipartite graph $K_{n,n}$ with P_r where n is odd and $r > 3$. Here we first check the value r, then calculate $r \equiv q(mod\ 4)$ and $t = (r - q)$.

Input: A Cartesian product between $K_{n,n}$ and P_r i.e., $G = (K_{n,n} \times P_r)$, where n is odd.

Output: Signed product and total signed product cordial labeling of the graph $G = (K_{n,n} \times P_r)$.

Step 1 for $i = 1$ to t

Step 2 if $i = 1(mod\ 2)$ then

Step 3 for $j = 1$ to $(n - 1)$

1. $f(x_{ij}) = -1$ if $sum(i, j) = 0(mod\ 2)$.
2. $f(y_{ij}) = 1$ if $sum(i, j) = 0(mod\ 2)$.
3. $f(x_{ij}) = 1$ if $sum(i, j) = 1(mod\ 2)$.
4. $f(y_{ij}) = -1$ if $sum(i, j) = 1(mod\ 2)$.

   end of loop.

Step 4 $f(x_{in}) = -1$ and $f(y_{in}) = 1$.

Step 5 else if $i = 0(mod\ 4)$ then

Step 6 for $j = 1$ to n

1. $f(x_{ij}) = 1$ if $sum(i, j) = 0(mod\ 2)$.
2. $f(y_{ij}) = 1$ if $sum(i, j) = 0(mod\ 2)$.
3. $f(x_{ij}) = -1$ if $sum(i, j) = 1(mod\ 2)$.
4. $f(y_{ij}) = -1$ if $sum(i, j) = 1(mod\ 2)$.

   end of loop.

Step 7 else

Step 8 for $j = 1$ to n

1. $f(x_{ij}) = -1$ if $sum(i, j) = 0(mod\ 2)$.
2. $f(y_{ij}) = -1$ if $sum(i, j) = 0(mod\ 2)$.
3. $f(x_{ij}) = 1$ if $sum(i, j) = 1(mod\ 2)$.
4. $f(y_{ij}) = 1$ if $sum(i, j) = 1(mod\ 2)$.

   end of loop.

end of loop.

Step 9 For rest of the copy of $K_{n,n}$ call Algorithm 1 Stop.

Table 1 Vertex label matrix of $K_{5,5} \times P_7$

−1	1	1	1	−1	1	−1	−1	−1	1	−1	−1	−1	1
1	−1	−1	−1	1	−1	1	1	1	−1	1	1	1	−1
−1	1	1	1	−1	1	−1	−1	−1	1	−1	−1	−1	1
1	−1	−1	−1	1	−1	1	1	1	−1	1	1	1	−1
−1	1	1	1	−1	1	−1	−1	−1	1	−1	−1	−1	1

Table 2 Some edge label matrix of $K_{5,5} \times P_7$

−1	1	−1	1	−1	−1	1	1	−1	1	−1	1	−1	1
1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	−1	1
−1	1	−1	1	−1	−1	1	1	−1	1	−1	1	−1	1
1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	−1	1
−1	1	−1	1	−1	−1	1	1	−1	1	−1	1	−1	1

Proof Considering the graph $G = K_{n,n} \times P_r$ when n is odd, where G has total number of vertices is equal to $2nr$ and total number of edges is equal to $n^2 r + 2nr - 2n$. According to Algorithm 2, total number of vertices label by −1 is nr and total number of vertices label by 1 is nr, i.e., $V_f(-1) = V_f(1) = nr$. So $|V_f(-1) - V_f(1)| = |nr - nr| = 0 \leq 1$. Now for r is odd total number of edges label by 1 is $\frac{(n^2 r + 2nr - 2n - 1)}{2}$ and total number of edges label by −1 is $\frac{(n^2 r + 2nr - 2n + 1)}{2}$, i.e., $|e_{f^*}(1) - e_{f^*}(-1)| = |\frac{(n^2 r + 2nr - 2n - 1)}{2} - \frac{(n^2 r + 2nr - 2n + 1)}{2}| = 1 \leq 1$, which follow the restriction of signed product cordial labeling. Now $|e_{f^*}(-1) + v_f(-1) - e_{f^*}(1) - v_f(1)| = |\frac{(n^2 r + 2nr - 2n + 1)}{2} + nr - \frac{(n^2 r + 2nr - 2n - 1)}{2} - nr| = 1 \leq 1$ follows the restriction of total signed product cordial labeling. For r is even total number of edges label by 1 is $\frac{(n^2 r + 2nr - 2n)}{2}$ and total number of edges label by −1 is $\frac{(n^2 r + 2nr - 2n)}{2}$, i.e., $|e_{f^*}(1) - e_{f^*}(-1)| = |\frac{(n^2 r + 2nr - 2n)}{2} - \frac{(n^2 r + 2nr - 2n)}{2}| = 0 \leq 1$, which follow the restriction of signed product cordial labeling. Now $|e_{f^*}(-1) + v_f(-1) - e_{f^*}(1) - v_f(1)| = |\frac{(n^2 r + 2nr - 2n)}{2} + nr - \frac{(n^2 r + 2nr - 2n)}{2} - nr| = 1 \leq 1$ follows the restriction of total signed product cordial labeling (Tables 1 and 2).

3.4 Algorithm

3.4.1 The Proof of Correctness of the Algorithm 3 Is Given Below

Theorem 3 *Algorithm 3 correctly labels the graph $K_{n,n} \times P_r$ (where n is even) by signed product and total signed product cordial labeling.*

Algorithm 3 Algorithm for signed product and total signed product cordial labeling of cartesian product between $K_{n,n} \times P_r$ (SPTSPCLBP(E))

Input: A Cartesian product between $K_{n,n}$ and P_r i.e. $G = (K_{n,n} \times P_r)$, where n is even.

Output: Signed product and total signed product cordial labeling of the graph $G = (K_{n,n} \times P_r)$.

Step 1 for $i = 1$ to r

Step 2 if $i = 1(mod\ 2)$ then

for $j = 1$ to n

1. $f(x_{ij}) = 1$ if $sum(i,j) = 0(mod\ 2)$.
2. $f(x_{ij}) = -1$ if $sum(i,j) = 1(mod\ 2)$.
3. $f(y_{ij}) = -1$ if $sum(i,j) = 0(mod\ 2)$.
4. $f(y_{ij}) = 1$ if $sum(i,j) = 1(mod\ 2)$.

   end of loop.

Step 3 if $i = 0(mod\ 2)$ then

for $j = 1$ to n

1. $f(x_{ij}) = -1$ if $sum(i,j) = 1(mod\ 2)$.
2. $f(x_{ij}) = 1$ if $sum(i,j) = 0(mod\ 2)$.
3. $f(y_{ij}) = -1$ if $sum(i,j) = 1(mod\ 2)$.
4. $f(y_{ij}) = 1$ if $sum(i,j) = 0(mod\ 2)$.

end of loop.

end of loop.

Stop.

Proof Considering the graph $G = K_{n,n} \times P_r$ when n is even, where G has total number of vertices is equal to $2nr$ and total number of edges is equal to $n^2r + 2nr - 2n$. According to Algorithm 3, total number of vertices label by -1 is nr and total number of vertices label by 1 is nr, i.e., $V_f(-1) = V_f(1) = nr$. So $|V_f(-1) - V_f(1)| = |nr - nr| = 0 \le 1$. For any value of r total number of edges label by 1 is $\frac{(n^2r+2nr-2n)}{2}$ and total number of edges label by -1 is $\frac{(n^2r+2nr-2n)}{2}$, i.e., $|e_{f^*}(1) - e_{f^*}(-1)| = |\frac{(n^2r+2nr-2n)}{2} - \frac{(n^2r+2nr-2n)}{2}| = 0 \le 1$, which follow the restriction of signed product cordial labeling. Now $|e_f^*(-1) + v_f(-1) - e_f^*(1) - v_f(1)| = |\frac{(n^2r+2nr-2n)}{2} + nr - \frac{(n^2r+2nr-2n)}{2} - nr| = 1 \le 1$ follows the restriction of total signed product cordial labeling.

4 Conclusion

In this article, we have developed three algorithms of signed product and total signed product cordial labeling of the graph obtained by Cartesian product between balanced bipartite graph and path. All the above algorithms will work successfully. We have also analyze time complexity of the above algorithms and time complexity of the signed product and total signed product cordial labeling of the graph obtained by

the Cartesian product between $K_{n,n} \times P_r$, i.e., balanced bipartite and path is $\bigcirc(nr)$. Labeling of a complex graph like Cartesian product between two graphs is an interesting job, so for the researcher, it is a challenge to label more complex graph with a less time complexity which indeed gives more flexibility in communication system or frequency assignment. Researchers may also incorporate the above algorithms in computer memory by writing computer programming to analyze the time complexity more efficiently. It is really a challenging work for us to label a complex graph by cordial, signed product cordial, total signed product cordial, prime cordial, etc.

References

1. Gallian, J.A.: A dynamic survey of graph labeling. Electron. J. Comb. **17**, 6 (2010)
2. Hegde, S.M.: On Multiplicative Labelings of a Graph. Labeling of Discrete Structures and Applications, pp. 83–96. Narosa Publishing House, New Delhi (2008)
3. Beineke, L.W., Hegde, S.M.: Strongly multiplicative graphs. Discuss. Math. Graph Theory **21**, 63–75 (2001)
4. Cahit, I.: Cordial graphs a weaker version of graceful and harmonious graphs. Ars Comb. **23**, 201–207 (1987)
5. Cahit, I.: On cordial and 3-equitable labellings of graphs. Util. Math. **37**, 189–198 (1990)
6. Rosa, A.: On certain valuations of the vertices of a graph. Theory of graphs. In: International Symposium, pp. 349–355 (1967)
7. Graham, R.L., Sloane, N.J.A.: On additive bases an harmonious graphs. SIAM J. Algebraic Discrete Methods **1**(4), 382–404 (1980)
8. Babujee, J.B., Shobana, L.: Cordial languages and cordial numbers. J. Appl. Comput. Sci. Math. **13**(6), 9–12 (2012)
9. Ulaganathan, P.P., Selvam, B., Vijaya, K.P.: Signed product cordial labeling in duplicate graphs of bistar, double star and triangular ladder graph. Int. J. Math. Trends Technol. (IJMTT) **33**(1), 19–24 (2016)
10. Ho, Y.S., Lee, S.M., Shee, S.C.: Cordial labelings of unicyclic graphs and generalized Peterson graphs. Congr. Numer. **68**, 109–122 (1989)
11. Santhi, M.: Some star and bistar related signed product cordial graphs. Int. J. Math. Arch. **6**(10), 232–236 (2015)

A Survey on Detection and Mitigation of Interest Flooding Attack in Named Data Networking

Sandesh Rai and Dependra Dhakal

Abstract Named Data Networking (NDN) is an instantiation of Content-Centric Networking (CCN) that focuses on the limitation of the current working IP-based Internet architecture. Like any other networks, NDN suffers from many threats that include denial-of-service attack (DoS) or distributed Dos (DDoS). To exhaust the NDN router resources or content provider, DDoS attack can be triggered. This survey paper addresses the interest flooding attack that is one of the DDoS attack types in NDN which tries to overflow the Pending Interest Table of the NDN router and exponentially decreases the NDN router performance.

Keywords CCN · NDN · DDOS · Pending interest table

1 Introduction

There are many limitations and challenges of the current working Internet architecture, and there are many challenges that Internet is coming across. So for supporting the increasing demands of the current Internet in the field of security, mobility, reliability, performance, etc., it has become much more difficult than when it used to be. Recently, there are going numbers of research on the networking architecture field to design next-generation Internet architecture to overcome the abovementioned challenges. And one the ongoing research is Named Data Networking (NDN) that focuses on the limitation of the IP-based Internet architecture. NDN works on the concept of the Content-Centric-Networking (CCN) where data are focuses much more than the location. Content plays a key role in this architecture. In NDN approach, each and every piece of packet is named. NDN is funded by the

S. Rai (✉) · D. Dhakal
Computer Science & Engineering, Sikkim Manipal Institute of Technology,
Rangpo, Sikkim, India
e-mail: chinangkhung@gmail.com

D. Dhakal
e-mail: ddhakal@gmail.com

United State National Science Foundation (NSF) under the Future Internet Architecture Program (IFA). NDN focuses on the two major aspects of the networking, i.e., security and privacy. As the time fly's, distributed denial-of-service (DDOS) attack has become common and more dangerous than it used to be, and it becomes most dangerous threats on the Internet. They are very difficult to detect to avoid and to mitigate. For any new architecture, it should able to detection and mitigation of the DDoS attack or at least minimize the effectiveness is the must. NDN is very efficient for providing content for the legit user but remain inefficient to the malicious user.

DDOS attack in NDN uses various number of host computers and various number of Internet connections. It is distributed across the whole wide world. There is not much difference between a DoS attack versus a DDOS attack. In DoS, the victim computer will be overflooded by thousands number of resources requests. In the DDoS attacks process, the malicious node in the network sends interest request, a very large number of zombie request for the attack to take place. A malicious node attacks the network by requesting resources in a high frequency, in the form of interest packets with spoofed names or without spoofed name. These high-frequency numbers of interest overload the bandwidth of the network hampering the performance and exhaust the router's resources. This type of attack is known as Interest Flooding Attack (IFA) and this paper exclusively focuses on this problem and their proposed countermeasures.

2   Interest Flooding Attacks

An NDN router route the content by looking up the Pending Interest Table (PIT) information. PIT contains all the pending interest request and it interface that I came from. The adversary node uses the state information of the PIT to perform DDoS in the router. An adversary node can perform DDoS attack in three different ways. (1) *Static DDoS,* (2) *Dynamic DDoS, and* (3) *Non-existing DDoS.* Static DDoS is least dangerous attack the three different types of attack. Basically, it tries to attack the network infrastructure. It is limit so Content Store provide and mitigation method. On the hand, Dynamic DDoS is very dangerous. The attack request is dynamic. It tries to deplete the network performance bandwidth and state of the PIT. The entire request tries to reach the content provider. So the Content Store is not enough to provide countermeasure to it. Non-existing DDoS as the name implies, the entire request is not satisfiable. Since it does not exit the entire request that router forward will roam around the network until it expires and hence depleting the network bandwidth, performance, and router resources. A build in cache cannot provide a countermeasure. All abovementioned different types of DDoS attack are distributed in nature and attacker send a very huge number of fake interest request in the network. The malicious user uses PIT feature to perform interest flooding attack in router. The malicious user issues a huge set of fake interest packet with spoof name of the content in the network. Its main aims are to overwhelm the PIT entries in router and to slough the content provider. Since
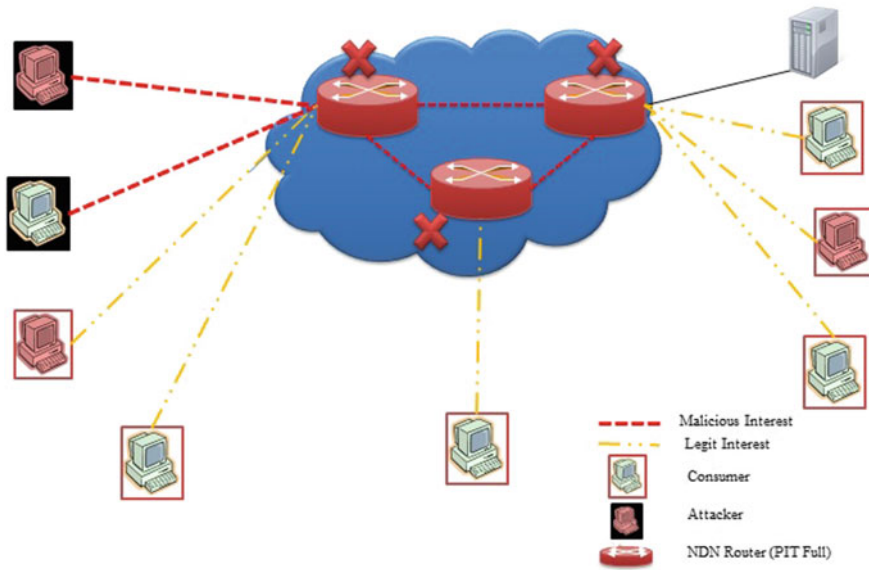
Fig. 1 Example of interest flooding attack

the entire interest request will be fake, the router will not be able to satisfy the request and that interest will be added to the PIT. Since the request is in huge number, it quickly overwhelms the limited router PIT space. This happens until the PIT is full.

And if the PIT is full, all legit incoming interest request will be dropped as all the PIT space is used up by fake interest request of the attacker. The interest will remain in the PIT for as much as possible, which will definitely exhaust the router memory and router resources on routers. This is the goal of interest flooding attack (Fig. 1).

3 Architecture

Named data networking is an ongoing research architecture that is for future whose intention is to mix the current Internet architecture and its usages. However, the architecture design and principles of the NDN are the derivation from of today's Internet architecture [1].

The current growth of the Internet is due to the thin waist of the architecture that is shown in Fig. 2 of hourglass architecture that allows both the upper and bottom layer to provide services independently.

The NDN architecture uses the same hourglass shape as shown in Fig. 2, but changes the thin waist by using data directly rather than its location.

Node in NDN uses two different packets for communication, i.e., interest packet and data packet. A node sends interest packet for asking resources in the network
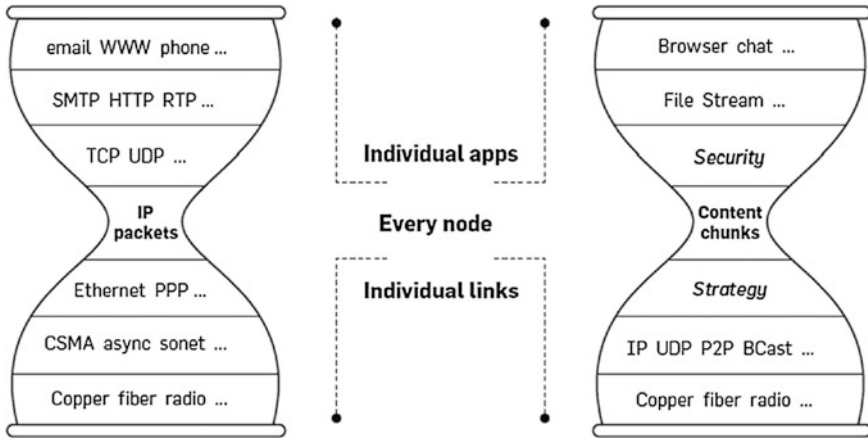
Fig. 2 NDN hourglass architectures [1]

and any node that has those resources send back a data packet. A satisfying node can be a simple node or can be a router. Different fields of a data packet [2].

Any node that either is a simple NDN node or router having the requesting data that satisfies it, a data packet is issued by the satisfied router [3] or a node. An NDN router maintain three different data structure and one forwarding policies for interest packet and data packet forwarding (Fig. 3).

- (i) **Content Store (CS)**: Recently used data are store.
- (ii) **Forward Information Base (FIB)**: Routing table of name of the data and it guides interests toward data producers.
- (iii) **Pending Interest Table (PIT)**: Store unsatisfied data request. It records the requested data name.
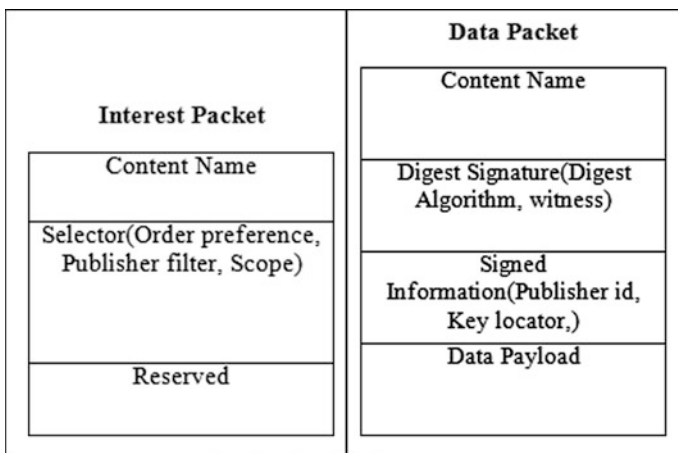
Fig. 3 NDN packet

Whenever an NDN router receives an interest packet form a node and if it is not in its Content Store (CS) and there are no pending interests entry for the same interest request in its PIT, then that interest is added in PIT and looking to FIB it forwards that interest. For each and every interest, in PIT, NDN router stores information, like the interest name and the interface name from which it arrived. If the PIT already contains the same interest, than the router drops that interest and stores only the interface from which it was arrived. And when the data packet arrives in the router, it forward that data packet from which that interest came from by looking into PIT. After that, the interest request is deleted from PIT and that data is cache in CS.

4 Objectives

The main objective of this paper is to presents an NDN focused detailed overview of DDOS attack in the NDN architecture and how to deal with attacks. It offers a literature review of different papers and then a summary comparing the different approaches present currently.

5 Related Works

In this part, a number of detection and mitigation countermeasures of interest flooding are discussed and analyzed. It is aimed for providing a good understanding for the readers.

Gasti et al. [4] explained the DoS and DDoS attack in the Named Data Networking. Basically, the authors explained the attack resilience and two different types of DDoS, i.e., (1) interest flooding and (2) cache poisoning attack. Gasti also proposed two different detection and mitigation methods, i.e., (1) router statistics and (2) pushback approach.

Afanasyev et al. [5] also explained different types of DDoS attack in Named Data networking. The paper presents the feasibility of interest flooding attack and the requirement for the attack to be performed. The authors proposed a token bucket mitigation approach for detection and mitigation of attack. It further modifies the token bucket mitigation approach into three different approaches, i.e., (a) token bucket with per interface fairness, (b) satisfaction-based pushback, and (c) satisfaction-based interest acceptance. For detection and mitigation of attack, all three approaches exploit its own state information. Among the three proposed approaches, satisfaction-based pushback mechanism effectively stops the attack in the victim router.

Campagno et al. [6] explained the interest flooding attack in the context of non-existing type of interest attack. The authors strictly focus on mitigation approaches only for non-existing type of interest flooding attack in NDN. The proposed approach is based on frequency of the interest request. The approach is called Poseidon.

Dai et al. [7] explained the interest flooding attack, their effect, and their requirements. The authors proposed a solution based on combination of router state and content provider. The solution is known as interest trace back. For every PIT entries, the algorithm generates a spoof data packet to satisfy the interest and to trace the interest originator. Due to generation of the spoof data packet, the algorithm generates overhead and traffic in the network causing depletion of bandwidth and performance. The limitation of the algorithm is that it takes the interest which is longest among PIT entries as malicious request. Any router running the interest traceback algorithm drops any long interest request to the router.

Choi et al. [8] provide an overview of the attack which is of non-existing type of attack. The authors explained clearly about the impact of the attack in the network performance and in the quality of service in Named Data Networking (NDN).

Karami et al. [9] proposed a hybrid proactive and intelligent algorithm for detection and mitigation of interest flooding attack. The algorithm is divided into two phases, i.e., (1) detection and (2) reaction. Multiobjective evolutionary optimization and Radial Basis function is used by algorithm for detection purpose. And for mitigating, an adaptive reaction mechanism is used.

Vassilakis et al. [10] proposed an IFA mitigation mechanism. The algorithm is based on anomalous user behavior. The algorithm classifies the router into two types, i.e., (1) Edge router: Router that is directly connected to user. (2) Core routers: Router that is directly connected to other router. The mechanism are divide into three phases, i.e., (1) attack detection phase: Here the edge router tries to detect any anomalous behavior of the node and classify as suspicious or attacker. (2) Rate reduction and blocking phase: Here the edge router block the identify attacker and reduces the interest request rate in the interface. (3) Attack notification: Here the edge router notifies other edge routers.

Wang et al. [11] proposed a cooperative-filter mechanism to detect and mitigate the interest flooding attack. Every NDN router running cooperative-filter mechanism detects IFA based on fuzzy logic and mitigate by filtering the entire incoming interest request between the cooperative routers using pushback message.

Ding et al. [12] proposed a cooperative detection and protection mechanism for interest flooding attack. The mechanism maintains a vector space model to obtain the spatial distance for every router and Markov transition is applied to compute the interest state for determining whether to forward such interest or not.

Salah et al. [13] proposed a lightweight coordinated defense mechanism against the IFA in NDN. This mechanism has three major components, i.e., (1) Domain Controller (DC): It maintains all the packet and corresponding state information from all the router. (2) NDN Router (NRs): It is normal NDN routers. (3) Monitoring Routers (MRs): This router monitors the NDN routers. It checks the interest satisfaction ratio per name-space and name-prefix of the NDN routers. And periodically, this state information is given to DC which return a feedback whether there is attack or not. The attack detection is carried out in two levels: (1) in a wide range, it is done by DC based on the satisfaction ratio, and (2) at MR observing the satisfaction ratio of the interest.

6 Analysis of Survey

The following table shows the analysis of the all the paper and comparison related only to the project. The table is a comparison of different papers which are written by well-known publishers. The analysis tries to provide a possible research gap that is present in the paper (Table 1).

Table 1 Comparison of different NDN-related papers

Approach	Detection	Trigger	Mitigation	Percentage	Topology tested
Token bucket with per interface fairness	In router	Token threshold	Round-robin in all interface	15–20%	ISP topology
Satisfaction-based acceptance	Per interface	Ratio of interest and data	Limit the interest	15–25%	ISP topology
Satisfaction-based pushback	Per interface	Ratio of interest and data	Distributed pushback	30–100%	ISP topology
Interest traceback	In router	PIT size threshold	Downstream traceback	–	Rocket fuel topology
Poseidon local	Per interface	Ratio of interest and data	Limit the interest and send alert message	More than 80%	AT&T topology
Poseidon distributed	Per interface	Ratio of interest and data	Limit the interest and send alert message	More than 80%	AT&T topology
Hybrid multiobjective RBF-PSO method	In router	Proactive prediction	Limit the interface for adversary	80–90%	DFN-like topology
IFA mitigation mechanism	In router	Anomalous user behavior	Rate reduction and blocking interface	More than 80%	Randomly generated network topology
Cooperative-filter	In router	Prediction based on fuzzy logic	Pushback alert message	–	Large-scale network topology
Cooperative detection and protection mechanism	In router	Prediction with granularity	Discard the packet	More than 80%	Small-scale tree topology, rocket fuel's AT&T topology
Lightweight coordinated defense mechanism	In router	Satisfaction interest ratio	Discard the packet and exchange attack information	More than 78%	AS-3967 topology

7   Conclusion

In this paper, we try to explain for both a comparable analysis of countermeasures of interest flooding. In future work, we will investigate the explained approaches more deeply and also try to argue the future work of those approaches. Furthermore, we will extend our analysis, to cover more complex application scenarios and also to complement simulations by long-range real-world experiments.

References

1. Pan, J., Paul, S., Jain, R.: A survey of the research on future internet architectures. Commun. Mag. IEEE (2011)
2. Hoque, A., Amin, S.O., Alyyan, A., Zhang, B., Zhang, L., Wang, L.: NLSR: Named-data link state routing protocol. In: Proceedings of the 3rd ACM SIGCOMM Workshop7Information-Centric Networking, ACM, pp. 15–20 (2013)
3. Ghali, C., Tsudik, G., Uzun, E.: Elements of trust in named-data networking. ACM SIGCOMM Comput. Commun. Rev. ACM **44**(5), 1–9 (2014)
4. Gasti, P., Tsudik, G., Uzun, E., Zhang, L.: DoS and DDoS in named data networking. In: 22nd International Conference on Computer Communications and Networks (ICCCN), pp. 1–7 (2013)
5. Afanasyev, A., Mahadevan, P., Moiseenko, I., Uzun, E., Zhang, L.:Interest flooding attack and countermeasures in named data networking. In: IFIP Networking Conference, pp. 1–9 (2013)
6. Compagno, A., Conti, M., Gasti, P., Tsudik, G.: Poseidon: mitigating interest flooding DDoS attacks in named data networking. In: 38th Conference on Local Computer Networks (LCN), IEEE, pp. 630–638 (2013)
7. Dai, H., Wang, Y., Fan, J., Liu, B.: Mitigate DDoS attacks in NDN by interest traceback. In: Conference on Computer Communications Workshops (INFOCOM WKSHPS), IEEE pp. 381–386 (2013)
8. S. Choi, K. Kim, S. Kim, and B.-H. Roh.: Threat of DoS by interest flooding attack in content-centric networking, In: International Conference on Information Networking (ICOIN), pp. 315–319 (2013)
9. Karami, A., Guerrero-Zapata, M.: A hybrid multiobjective RBFPSO method for mitigating DoS attacks in named data networking, Neurocomputing **151**, 1262–1282 (2015)
10. Vassilakis, V.G., Alohali, B., Moscholios, I., Logothetis, M.: Mitigating distributed denial-of-service attacks in named data networking. In: Proceedings of 11th Advanced International Conference on Telecommunications (AICT), Brussels, Belgium (pp. 18–23) June 2015
11. Wang, K., Zhou, H., Qin, Y., Zhang, H.: Cooperative-Filter: countering Interest flooding attacks in named data networking. Soft. Comput. **18**(9), 1803–1813 (2014)
12. Ding, K., Liu, Y., Cho, H.H., Chao, H.C., Shih, T.K.: Cooperative detection and protection for interest flooding attacks in named data networking. Int. J. Commun. Syst. Oct 1, 2014
13. Salah, H., Julian, W., Thorsten, S.: Lightweight coordinated defence against interest flooding attacks in NDN. In: 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 103–104. IEEE (2015)
14. Jacobson, V., Mosko, M., Smetters, D., Garcia-Luna-Aceves, J.: Content-centric Networking, Whitepaper, Palo Alto Research Center, pp. 2–4 (2007)

15. Jacobson, V., Smetters, D.K., Thornton, J.D., Plass, M.F., Briggs, N.H., Braynard, R.L..: Networking named content. In: Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies, ACM (2009)
16. Zhang, L., Estrin, D., Burke, J., Jacobson, V., Thornton, J.D., Smetters, D.K., Zhang, B., Tsudik, G., Massey, D., Papadopoulos, C. et al.: Named Data Networking (NDN) Project, RelatorioTecnico NDN-0001, Xerox Palo Alto Research Center-PARC (2010)
17. Jacobson, V., Burke, J., Zhang, L., Zhang, B., Claffy, K., Krioukov, D., Papadopoulos, C., Wang, L., Yeh, E., Crowley, P.: Named Data Networking (NDN) Project 2013–2014 Report. http://named-data.net, Annual Progress Report (2014)
18. Aamir, M. Zaidi, S.M.A.: Denial-of-service in content centric (named data) networking: a tutorial and state-of-the-art survey. Secur. Commun. Netw. **8**(11), 2037–2059 (2015)
19. Wählisch, M., Schmidt, T.C., Vahlenkamp, M.: Backscatter from the data plane threats to stability and security in information-centric networking. CoRR, abs/1205.4778 (2012)
20. Content Centric Networking (CCNx) Project. http://www.ccnx.org
21. Afanasyev, A., Moiseenko, I., Zhang, L.: ndnSIM: NDN simulator for NS-3. Technical Report NDN-0005, University of California, Los Angeles (2012)
22. Wang R, Jia Z, Ju L.: An entropy-based distributed DDoS detection mechanism in software-defined networking. In: Trustcom/BigDataSE/ISPA, vol. 1, pp. 310–317 (2013)
23. Kumar, K., Joshi, R.C., Singh, K..: A distributed approach using entropy to detect DDoS attacks in ISP domain. In: International Conference on Signal Processing, Communications and Networking, ICSCN'07, pp. 331–337 (2007)
24. Feinstein, L., Schnackenberg, D., Balupari, R., Kindred, D.: Statistical approaches to DDoS attack detection and response. In: DARPA Information Survivability Conference and Exposition, 2003. Proceedings vol.1, pp. 303–314 (2003)
25. Krishnan, R., Krishnaswamy, D., Mcdysan, D.: Behavioral security threat detection strategies for data center switches and routers. In: 2014 IEEE 34th International Conference on Distributed Computing Systems Workshops (ICDCSW), pp. 82–87 (2014)
26. Zhang, Y.: An adaptive flow counting method for anomaly detection in SDN. In: Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies, pp. 25–30 (2013)

An Integrated TOPSIS Approach to MADM with Interval-Valued Intuitionistic Fuzzy Settings

Animesh Biswas and Samir Kumar

Abstract In this paper, the three-parameter characterization of intuitionistic fuzzy sets and normalized hamming distance are employed to develop mathematical programming-based TOPSIS techniques in interval-valued intuitionistic fuzzy settings. A pair of linear fractional programming models are generated which are simplified for producing intervals to measure relative closeness coefficients of alternatives. Possibility degree matrix is obtained by pairwise comparisons of closeness coefficients and optimal degrees are estimated for final ranking of alternatives. The proposed approach is illustrated through a numerical example.

Keywords Intuitionistic fuzzy sets · Interval-valued intuitionistic fuzzy numbers
TOPSIS · Mathematical programming · Possibility measure

1 Introduction

Hwang and Yoon [10] developed a technique for order preference by similarity to ideal solution (TOPSIS) technique as a very prominent multiple attribute decision-making (MADM) method. The fundamental idea of the TOPSIS is that the chosen alternative should have the shortest distance from the positive ideal solution (PIS) and the longest distance from the negative ideal solution (NIS). Chen [7] utilized the concepts of fuzzy set theory [19] to extend the TOPSIS in fuzzy settings. Atanassov [1] introduced intuitionistic fuzzy (IF) sets (IFSs) as a generalization of fuzzy set theory [2–6, 13]. Atanassov and Gargov [2] advanced the concepts of

A. Biswas (✉)
Department of Mathematics, University of Kalyani, Kalyani 741235, India
e-mail: abiswaskln@rediffmail.com

S. Kumar
Department of Mathematics, Acharya Jagadish Chandra Bose College,
Kolkata 700020, India
e-mail: kumarsamir2007@gmail.com

interval-valued IFSs (IVIFSs) as a further generalization of IFSs. Chen [8] extended the TOPSIS in interval-valued IF (IVIF) settings by employing the concepts of inclusion comparison possibilities to develop an index for an inclusion-based closeness coefficient for ranking the alternatives. Zhao [21] developed an optimization model for determination of criteria weights and presented a TOPSIS method for solving MADM problems in IVIF environment using weighted hamming distance between each alternative and PIS and NIS. Zeng and Xiao [20] presented a hybrid IF ordered weighted averaging weighted averaging (OWAWA) distance TOPSIS method for IF multiple-criteria decision-making problems. Szmidt and Kacprzyk [14] proposed the four basic distances between the IFSs by taking into account the three-parameter characterization of IFS, which were the hamming distance, the normalized hamming distance, the Euclidean distance and the normalized Euclidean distance. Li [12] developed a linear programming approach to solve MADM problems in IVIF settings by defining weighted absolute distance between two IFSs with respect to IF weights employing two-parameter characterization of IFS. Szmidt and Kacprzyk [14] established that any extension of fuzzy distance measure in the IF settings should contain all the three parameters membership, non-membership and hesitancy degrees.

Clearly, there is no scope for hesitation in Zadeh's fuzzy sets [19] binding the decision makers between dual choices of membership and non-membership while the three-parameter characterization of an IFS A assigns to each element $x \in X$ a membership degree $\mu_A(x)$, a non-membership degree $\nu_A(x)$ and a hesitancy degree $\pi_A(x)$ [14, 17] giving more freedom to decision makers for better formation of opinions regarding a given decision situation. The points $\langle 1, 0, 0 \rangle$, $\langle 0, 1, 0 \rangle$ and $\langle 0, 0, 1 \rangle$ correspond, respectively, to full acceptance, full rejection and absolute hesitation regarding the choice of a decision maker.

In this paper, the normalized hamming distance [14] between two IFSs is utilized to develop a mathematical programming-based TOPSIS technique by employing the three-parameter characterization of IFS to solve MADM problems in IVIF settings.

2 Preliminaries

The notion of IVIFSs and the possibility degree measures for comparing two IVIF numbers (IVIFNs) are briefly revisited in this section.

Definition 2.1 [2]. An IVIFS, A, on a universe of discourse $X(\neq \varnothing)$ is defined as a set $A = \{\langle x, \mu_A(x), \nu_A(x)\rangle : x \in X\}$ satisfying the condition $0 \leq \sup \mu_A(x) + \sup \nu_A(x) \leq 1$, for each $x \in X$, where $\mu_A : X \to D[0, 1]$ and $\nu_A : X \to D[0, 1]$, respectively, represent membership and non-membership functions and $D[0, 1]$ denotes the set of all closed subintervals of [0, 1].

The values $\mu_A(x)$ and $\nu_A(x)$ denote, respectively, the interval degree of membership and non-membership to the set A for each element $x(\in X)$. For simplicity, let $\mu_A^l(x), \mu_A^u(x)$ and $\nu_A^l(x), \nu_A^u(x)$ be, respectively, the infimum and supremum of $\mu_A(x)$ and $\nu_A(x)$, for each $x \in X$. Hence, the IVIFS A becomes $A = \{\langle x, [\mu_A^l(x), \mu_A^u(x)], [\nu_A^l(x), \nu_A^u(x)]\rangle : x \in X\}$ with $0 \leq \mu_A^u(x) + \nu_A^u(x) \leq 1$, $\mu_A^l(x) \geq 0$ and $\nu_A^l(x) \geq 0$.

Definition 2.2 [2]. Let IVIFS(X) denote the set of all IVIFSs in X. Let $A, B \in$ IVIFS(X). Then $A \subseteq B$ if and only if $\mu_A^l(x) \leq \mu_B^l(x), \mu_A^u(x) \leq \mu_B^u(x)$ and $\nu_A^l(x) \geq \nu_B^l(x), \nu_A^u(x) \geq \nu_B^u(x), \forall x \in X$.

An associated function $\pi_A : X \to D[0, 1]$ for IVIFS A in X is also defined by

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) = [\pi_A^l(x), \pi_A^u(x)]$$
$$= [1 - \mu_A^u(x) - \nu_A^u(x), 1 - \mu_A^l(x) - \nu_A^l(x)], \quad \forall x \in X$$

Here, $\pi_A(x)$ represents the interval hesitancy degree of the element x in A.

Xu and Yager [17] called the ordered triple $\alpha = \langle \mu_\alpha, \nu_\alpha, \pi_\alpha \rangle$ an intuitionistic fuzzy number (IFN), where $\mu_\alpha \in [0, 1], \nu_\alpha \in [0, 1], \pi_\alpha \in [0, 1]$ satisfying $0 \leq \mu_\alpha + \nu_\alpha \leq 1$, and $\pi_\alpha = 1 - \mu_\alpha - \nu_\alpha$. Thus the greatest and smallest IFNs are, respectively, given by $\alpha^+ = \langle 1, 0, 0 \rangle$ and $\alpha^- = \langle 0, 1, 0 \rangle$.

Definition 2.3 [16]. Assume that $A = \{\langle x, \mu_A(x), \nu_A(x)\rangle : x \in X\}$ be an IVIFS. Then, the ordered pair $\langle \mu_A(x), \nu_A(x) \rangle$, for each $x \in X$, is called an IVIFN.

For ease of computation, let $\alpha_j = \langle [a_j^l, a_j^u], [b_j^l, b_j^u] \rangle$ denote an IVIFN. The basic operations for combining the two IVIFNs are given by the following definition.

Definition 2.4 [16]. Suppose that $\alpha_j = \langle [a_j^l, a_j^u], [b_j^l, b_j^u] \rangle$ $(j = 1, 2)$ be two IVIFNs and $\lambda > 0$. Then,

$$\alpha_1 + \alpha_2 = \langle [a_1^l + a_2^l - a_1^l a_2^l, a_1^u + a_2^u - a_1^u a_2^u], [b_1^l b_2^l, b_1^u b_2^u] \rangle$$
$$\lambda \alpha_1 = \langle [1 - (1 - a_1^l)^\lambda, 1 - (1 - a_1^u)^\lambda], [(b_1^l)^\lambda, (b_1^u)^\lambda] \rangle$$

Xu and Yager [17] also called the ordered triple $\langle \mu_A(x), \nu_A(x), \pi_A(x) \rangle$ an IVIFN for each $x \in X$. For convenience, an IVIFN is denoted by $A = \langle \mu_A, \nu_A, \pi_A \rangle$.

It is found that $A^+ = \langle [1, 1], [0, 0], [0, 0] \rangle$ and $A^- = ([0, 0], [1, 1], [0, 0])$ represent the greatest and smallest IVIFNs, respectively.

Definition 2.5 [9]. For two interval numbers $a = [a^l, a^u]$ and $b = [b^l, b^u]$ having interval lengths $L(a) = a^u - a^l$ and $L(b) = b^u - b^l$, respectively, the possibility degree of $a \geq b$ is defined as $p(a \geq b) = \min\{\max\{\frac{a^u - b^l}{L(a) + L(b)}, 0\}, 1\}$.

Definition 2.6 [11]. Assume that $\alpha_i = \langle [a_i^l, a_i^u], [b_i^l, b_i^u] \rangle = \langle a_i, b_i \rangle$ $(i = 1, 2)$ be two IVIFNs and $L(a_i) = a_i^u - a_i^l$, $L(b_i) = b_i^u - b_i^l$ be their interval lengths, then the degree of possibility $p(\alpha_1 \geq \alpha_2)$ of $\alpha_1 \geq \alpha_2$ is defined as

$$p(\alpha_1 \geq \alpha_2) = \min \left\{ \max \left\{ \frac{a_1^u - a_2^l + b_2^u - b_1^l}{L(a_1) + L(a_2) + L(b_1) + L(b_2)}, 0 \right\}, 1 \right\}.$$

3 MADM Problems with Three-Parameter Characterization of IVIFNs

Assume that $A = \{\langle \mu_A(x_j), \nu_A(x_j), \pi_A(x_j) \rangle : x_j \in X\}$ and $B = \{\langle \mu_B(x_j), \nu_B(x_j), \pi_B(x_j) \rangle : x_j \in X\}$ be two IFSs in a universe of discourse $X = \{x_1, x_2, \ldots, x_n\}$ representing the weights of importance of all elements $x_j \in X$. Then, the normalized hamming distance [14] between IFSs A and B is defined as

$$d(A, B) = \frac{1}{2n} \sum_{j=1}^{n} \left(|\mu_A(x_j) - \mu_B(x_j)| + |\nu_A(x_j) - \nu_B(x_j)| + |\pi_A(x_j) - \pi_B(x_j)| \right) \quad (1)$$

Let $A = \{A_1, A_2, \ldots, A_m\}$ be a set of alternatives which are evaluated on the set of criteria $X = \{X_1, X_2, \ldots, X_n\}$ by a decision maker generating IVIF decision matrix $D = (d_{ij})_{m \times n} = \langle [\alpha_{ij}^l, \alpha_{ij}^u], [\beta_{ij}^l, \beta_{ij}^u], [\gamma_{ij}^l, \gamma_{ij}^u] \rangle_{m \times n}$, where $0 \leq \alpha_{ij}^l \leq \alpha_{ij}^u \leq 1$, $0 \leq \beta_{ij}^l \leq \beta_{ij}^u \leq 1$, $0 \leq \gamma_{ij}^l \leq \gamma_{ij}^u \leq 1$ with $\alpha_{ij}^u + \beta_{ij}^u \leq 1$. Now, let the decision maker evaluates criteria assigning the IVIFN weight $w_j = \langle [\omega_j^l, \omega_j^u], [\rho_j^l, \rho_j^u], [\sigma_j^l, \sigma_j^u] \rangle$ to the jth criterion $X_j \in X$, where $0 \leq \omega_j^l \leq \omega_j^u \leq 1$, $0 \leq \rho_j^l \leq \rho_j^u \leq 1$, $0 \leq \sigma_j^l \leq \sigma_j^u \leq 1$ with $\omega_j^u + \rho_j^u \leq 1$. Thus, the IVIF criteria weight vector becomes

$$\tau = (\tau_1, \tau_2, \ldots, \tau_n) = (\tau_j)_{1 \times n} = \left(\langle [\omega_j^l, \omega_j^u], [\rho_j^l, \rho_j^u], [\sigma_j^l, \sigma_j^u] \rangle \right)_{1 \times n}.$$

Using Definition 2.6 and Eq. (11), let the crisp weights of criteria be $w = (w_1, w_2, \ldots, w_n)$.

Using Definition 2.4, the weighted IVIF decision matrix is generated as

$$D_w = \left(d_{ij}^w \right)_{m \times n} = \left(\langle [\mu_{ij}^l, \mu_{ij}^u], [\nu_{ij}^l, \nu_{ij}^u], [\pi_{ij}^l, \pi_{ij}^u] \rangle \right)_{m \times n}$$

where $d_{ij}^w = w_j d_{ij}$ $(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$, $\pi_{ij}^l = 1 - \mu_{ij}^u - \nu_{ij}^u$ and $\pi_{ij}^u = 1 - \mu_{ij}^l - \nu_{ij}^l$.

4 TOPSIS Approach to Solve MADM Problems in IVIF Settings with Three-Parameter Characterization of IFSs

Without loss of generality, let $X_B = \{X_1, X_2, \ldots, X_r\}$ and $X_C = \{X_{r+1}, X_{r+2}, \ldots, X_n\}$, respectively, represent sets of benefit criteria and cost criteria. Then, the IVIF PIS (IVIFPIS), A^+, and IVIF NIS (IVIFNIS), A^-, may be considered as

$$A^+ = \left(A_1^+, A_2^+, \ldots, A_n^+\right),$$

where

$$A_j^+ = \begin{cases} \langle [1,1], [0,0], [0,0] \rangle, & \text{for } j = 1, 2, \ldots, r \\ \langle [0,0], [1,1], [0,0] \rangle, & \text{for } j = r+1, r+2, \ldots, n \end{cases}$$
$$A^- = \left(A_1^-, A_2^-, \ldots, A_n^-\right),$$

where

$$A_j^- = \begin{cases} \langle [0,0], [1,1], [0,0] \rangle, & \text{for } j = 1, 2, \ldots, r \\ \langle [1,1], [0,0], [0,0] \rangle, & \text{for } j = r+1, r+2, \ldots, n \end{cases}$$

Let $\mu_{ij} \in \left[\mu_{ij}^l, \mu_{ij}^u\right]$, $\nu_{ij} \in \left[\nu_{ij}^l, \nu_{ij}^u\right]$, $\pi_{ij} \in \left[\pi_{ij}^l, \pi_{ij}^u\right]$, then using Eq. (1), the normalized hamming distance between an alternative $A_i \in A (i = 1, 2, \ldots, m)$ and the IVIFPIS A^+ with respect to all the criteria $X_j \in X (j = 1, 2, \ldots, n)$, may be expressed as

$$
\begin{aligned}
d(A_i, A^+) &= \frac{1}{2n} \left(\sum_{j=1}^r \left[(1 - \mu_{ij}) + \nu_{ij} + \pi_{ij} \right] + \sum_{j=r+1}^n \left[(\mu_{ij} + (1 - \nu_{ij}) + \pi_{ij} \right] \right) \\
&= \frac{1}{n} \left(\sum_{j=1}^r (1 - \mu_{ij}) + \sum_{j=r+1}^n (1 - \nu_{ij}) \right) \left(\text{Since } \pi_{ij} = 1 - \mu_{ij} - \nu_{ij} \right)
\end{aligned}
\tag{2}
$$

Similarly, using Eq. (1), the normalized hamming distance between an alternative $A_i \in A (i = 1, 2, \ldots, m)$ and the IVIFNIS A^- with respect to all the criteria $X_j \in X (j = 1, 2, \ldots, n)$ becomes

$$
\begin{aligned}
d(A_i, A^-) &= \frac{1}{2n} \left(\sum_{j=1}^r \left[\mu_{ij} + (1 - \nu_{ij}) + \pi_{ij} \right] + \sum_{j=r+1}^n \left[(1 - \mu_{ij}) + \nu_{ij} + \pi_{ij} \right] \right) \\
&= \frac{1}{n} \left(\sum_{j=1}^r (1 - \nu_{ij}) + \sum_{j=r+1}^n (1 - \mu_{ij}) \right) \left(\text{Since } \pi_{ij} = 1 - \mu_{ij} - \nu_{ij} \right)
\end{aligned}
\tag{3}
$$

Using Eqs. (2) and (3), the relative closeness coefficients of alternatives $A_i \in A(i = 1, 2, \ldots, m)$ is given by

$$C_i(\mu_{ij}, \nu_{ij}, \pi_{ij}) = \frac{d(A_i, A^-)}{d(A_i, A^+) + d(A_i, A^-)}, \quad \text{for } i = 1, 2, \ldots, m$$

i.e.

$$C_i(\mu_{ij}, \nu_{ij}, \pi_{ij}) = \frac{\frac{1}{n}\left(\sum_{j=1}^{r}(1-\nu_{ij}) + \sum_{j=r+1}^{n}(1-\mu_{ij})\right)}{\frac{1}{n}\left(\sum_{j=1}^{r}(1-\mu_{ij}) + \sum_{j=r+1}^{n}(1-\nu_{ij})\right) + \frac{1}{n}\left(\sum_{j=1}^{r}(1-\nu_{ij}) + \sum_{j=r+1}^{n}(1-\mu_{ij})\right)}$$

$$= \frac{\sum_{j=1}^{r}(1-\nu_{ij}) + \sum_{j=r+1}^{n}(1-\mu_{ij})}{\sum_{j=1}^{n}(1+\pi_{ij})} \left(\text{As } \pi_{ij} = 1 - \mu_{ij} - \nu_{ij}\right)$$

$$= \frac{M_i}{N_i}(\text{say}), \quad \text{for } i = 1, 2, \ldots, m \tag{4}$$

Thus from Eq. (4),

$$\log C_i = \log M_i - \log N_i \tag{5}$$

Using Eq. (5), it can be shown that $C_i(\mu_{ij}, \nu_{ij}, \pi_{ij})$ is a monotonically increasing function with respect to μ_{ij} $(j = 1, 2, \ldots, r)$ and monotonically decreasing function with respect to $\mu_{ij}(j = r+1, r+2, \ldots, n)$. Also $C_i(\mu_{ij}, \nu_{ij}, \pi_{ij})$ is a monotonically decreasing function with respect to $\nu_{ij}(j = 1, 2, \ldots, r)$ and monotonically increasing function with respect to $\nu_{ij}(j = r+1, r+2, \ldots, n)$, as $\log C_i$ and C_i possess identical monotonic behaviour.

As the values of $C_i(\mu_{ij}, \nu_{ij}, \pi_{ij})$ or C_i depend on the values $\mu_{ij} \in \left[\mu_{ij}^l, \mu_{ij}^u\right]$, $\nu_{ij} \in \left[\nu_{ij}^l, \nu_{ij}^u\right]$, $\pi_{ij} \in \left[\pi_{ij}^l, \pi_{ij}^u\right]$, therefore C_i should be an interval $\left[C_i^l, C_i^u\right]$, where C_i^l and C_i^u represent infimum and supremum of C_i, respectively. The values of C_i^l and C_i^u may be estimated using the following linear fractional programming models:

$$C_i^u = \max\left\{\frac{\sum_{j=1}^{r}(1-\nu_{ij}) + \sum_{j=r+1}^{n}(1-\mu_{ij})}{\sum_{j=1}^{n}(1+\pi_{ij})}\right\}$$

$$\text{subject to } \begin{cases} \mu_{ij}^l \leq \mu_{ij} \leq \mu_{ij}^u \ (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n) \\ \nu_{ij}^l \leq \nu_{ij} \leq \nu_{ij}^u \ (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n) \\ \pi_{ij}^l \leq \pi_{ij} \leq \pi_{ij}^u \ (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n) \end{cases} \tag{6}$$

and

$$C_i^l = \min\left\{ \frac{\sum_{j=1}^r \left(1 - \nu_{ij}\right) + \sum_{j=r+1}^n \left(1 - \mu_{ij}\right)}{\sum_{j=1}^n \left(1 + \pi_{ij}\right)} \right\}$$

$$\text{subject to} \quad \begin{cases} \mu_{ij}^l \leq \mu_{ij} \leq \mu_{ij}^u (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n) \\ \nu_{ij}^l \leq \nu_{ij} \leq \nu_{ij}^u (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n) \\ \pi_{ij}^l \leq \pi_{ij} \leq \pi_{ij}^u (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n) \end{cases} \tag{7}$$

Solving (6) and (7) it can be found that

$$C_i^u = \frac{\sum_{j=1}^r \left(1 - \nu_{ij}^l\right) + \sum_{j=r+1}^n \left(1 - \mu_{ij}^l\right)}{\sum_{j=1}^r \left(2 - \mu_{ij}^u - \nu_{ij}^l\right) + \sum_{j=r+1}^n \left(2 - \mu_{ij}^l - \nu_{ij}^u\right)} \tag{8}$$

$$C_i^l = \frac{\sum_{j=1}^r \left(1 - \nu_{ij}^u\right) + \sum_{j=r+1}^n \left(1 - \mu_{ij}^u\right)}{\sum_{j=1}^r \left(2 - \mu_{ij}^l - \nu_{ij}^u\right) + \sum_{j=r+1}^n \left(2 - \mu_{ij}^u - \nu_{ij}^l\right)} \tag{9}$$

Thus, the intervals of relative closeness coefficients of m alternatives $A_i(i = 1, 2, \ldots, m)$ are obtained as $C_i = \left[C_i^l, C_i^u\right]$. 'Alternative A_i is not inferior to A_k' is denoted by $A_i \succcurlyeq A_k$ whose possibility degree can be measured by that of $C_i \geq C_k$, C_i and C_k being intervals of relative closeness coefficients of alternatives A_i and $A_k(i, k = 1, 2, \ldots, m)$, respectively. Thus, the possibility degree of $A_i \succcurlyeq A_k$, using Definition 2.5, is given by

$$p_{ik} = p(A_i \succcurlyeq A_k) = p(C_i \geq C_k) = \min\left\{ \max\left\{ \frac{C_i^u - C_k^l}{L(C_i) + L(C_k)}, 0 \right\}, 1 \right\} \tag{10}$$

where $C_i = \left[C_i^l, C_i^u\right]$, $C_k = \left[C_k^l, C_k^u\right]$, $L(C_i) = C_i^u - C_i^l$ and $L(C_k) = C_k^u - C_k^l$.

The possibility degree matrix obtained by pairwise comparison of relative closeness coefficients C_i takes the form $P = (p_{ij})_{m \times m}$ which is an additive reciprocal fuzzy matrix.

Xu and Da [15] proposed the optimal degree formula to estimate crisp criteria weights or performance values of alternatives using possibility degree matrix P as

$$\xi_i = \frac{1}{m(m-1)} \left(\sum_{k=1}^n p_{ik} + \frac{m}{2} - 1 \right), \quad \text{for } i = 1, 2, \ldots, m \tag{11}$$

4.1 IVIF TOPSIS Method

Based on the previous deductions, the TOPSIS techniques for solving the MADM problems in IVIF settings are outlined as follows:

Step 1. By Definition 2.6, the additive reciprocal fuzzy preference relation $P = \left(p(\tau_i \geq \tau_j)\right)_{n \times n} = \left(p_{ij}\right)_{n \times n}$ from IVIF criteria weight vector $\tau = (\tau_1, \tau_2, \ldots, \tau_n) = (\tau_j)_{1 \times n}$ is generated.

The crisp weight w_j for the IVIFN weight τ_j $(j = 1, 2, \ldots, n)$ may be obtained by the optimal degree formula given by Eq. (11).

Thus, the crisp criteria weight vector $w = (w_1, w_2, \ldots, w_n)$ with $w_j \in [0, 1]$ and $\sum_{j=1}^{n} w_j = 1$ is obtained.

Step 2. For an IVIF decision matrix $D = \left(d_{ij}\right)_{m \times n} = \langle \left[\alpha_{ij}^l, \alpha_{ij}^u\right], \left[\beta_{ij}^l, \beta_{ij}^u\right], \left[\gamma_{ij}^l, \gamma_{ij}^u\right] \rangle_{m \times n}$, the weighted IVIF decision matrix is obtained using Definition 2.4 as $D_w = \left(d_{ij}^w\right)_{m \times n}$, where $d_{ij}^w = w_j d_{ij}$ $(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$. Let $D_w = \left(d_{ij}^w\right)_{m \times n} = \left(\langle \left[\mu_{ij}^l, \mu_{ij}^u\right], \left[\nu_{ij}^l, \nu_{ij}^u\right], \left[\pi_{ij}^l, \pi_{ij}^u\right] \rangle\right)_{m \times n}$, where $\pi_{ij}^l = 1 - \mu_{ij}^u - \nu_{ij}^u$ and $\pi_{ij}^u = 1 - \mu_{ij}^l - \nu_{ij}^l$.

Step 3. Using Eqs. (8) and (9), the intervals of relative closeness coefficients of m alternatives $A_i (i = 1, 2, \ldots, m)$ are given by $C_i = \left[C_i^l, C_i^u\right]$.

Step 4. Using Eq. (10), the possibility degree matrix for pairwise comparison of intervals of relative closeness coefficients C_i of alternatives $A_i (i = 1, 2, \ldots, m)$ is obtained as $P = \left(p_{ij}\right)_{m \times m}$.

Step 5. The performance scores as optimal degrees ξ_i for alternatives $A_i (i = 1, 2, \ldots, m)$ are evaluated using Eq. (11).

Step 6. The alternatives $A_i (i = 1, 2, \ldots, m)$ are ranked in decreasing order of the optimal degrees ξ_i.

5 Illustrative Example

In this section, an example has been adapted from Li [12] which also is an adaptation of Ye [18] for an MADM problem for illustrating the application of the proposed IVIF TOPSIS. The problem is related to an investment company which wishes to invest the capital in the four companies (alternatives) A_1 (car company), A_2 (food company), A_3 (computer company), and A_4 (arms company). An expert evaluates these alternatives relative to three criteria X_1 (risk analysis), X_2 (growth analysis), and X_3 (environmental impact analysis) generating IVIF decision matrix

$$D = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{ccc} X_1 & X_2 & X_3 \\ \left(\begin{array}{ccc} \langle [0.4, 0.5], [0.3, 0.4], [0.1, 0.3] \rangle & \langle [0.4, 0.6], [0.2, 0.4], [0.0, 0.4] \rangle & \langle [0.1, 0.3], [0.5, 0.6], [0.1, 0.4] \rangle \\ \langle [0.6, 0.7], [0.2, 0.3], [0.0, 0.2] \rangle & \langle [0.6, 0.7], [0.2, 0.3], [0.0, 0.2] \rangle & \langle [0.4, 0.7], [0.1, 0.2], [0.1, 0.5] \rangle \\ \langle [0.3, 0.6], [0.3, 0.4], [0.0, 0.4] \rangle & \langle [0.5, 0.6], [0.3, 0.4], [0.0, 0.2] \rangle & \langle [0.5, 0.6], [0.1, 0.3], [0.1, 0.4] \rangle \\ \langle [0.7, 0.8], [0.1, 0.2], [0.0, 0.2] \rangle & \langle [0.6, 0.7], [0.1, 0.3], [0.0, 0.3] \rangle & \langle [0.3, 0.4], [0.1, 0.2], [0.4, 0.6] \rangle \end{array} \right) \end{array}$$

The IVIFN criteria weight vector is

$$\tau = (\langle [0.1, 0.4], [0.2, 0.55], [0.05, 0.7] \rangle, \langle [0.2, 0.5], [0.15, 0.45], [0.05, 0.65] \rangle,$$
$$\langle [0.25, 0.6], [0.15, 0.38], [0.02, 0.6] \rangle)$$

Using Definition 2.6, the possibility degree matrix obtained by pairwise comparison of criteria is

$$P = \left(p(\tau_i \geq \tau_j) \right)_{3 \times 3} = \begin{array}{c} \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{array} \begin{array}{ccc} \tau_1 & \tau_2 & \tau_3 \\ \left(\begin{array}{ccc} 0.5 & 0.36 & 0.27 \\ 0.64 & 0.5 & 0.41 \\ 0.73 & 0.59 & 0.5 \end{array} \right) \end{array}$$

Using Eq. (11), the crisp criteria weight vector becomes $w = (w_1, w_2, w_3) = (0.27, 0.34, 0.39)$.

Using Definition 2.4, the weighted IVIF decision matrix is obtained as

$$D_w = \left(d_{ij}^w \right)_{4 \times 3} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{ccc} X_1 & X_2 & X_3 \\ \left(\begin{array}{ccc} \langle [0.129, 0.171], [0.722, 0.781] \rangle & \langle [0.159, 0.268], [0.579, 0.732] \rangle & \langle [0.040, 0.130], [0.763, 0.819] \rangle \\ \langle [0.219, 0.278], [0.648, 0.722] \rangle & \langle [0.268, 0.336], [0.579, 0.664] \rangle & \langle [0.181, 0.375], [0.407, 0.534] \rangle \\ \langle [0.092, 0.219], [0.722, 0.781] \rangle & \langle [0.210, 0.268], [0.664, 0.732] \rangle & \langle [0.237, 0.300], [0.407, 0.625] \rangle \\ \langle [0.278, 0.352], [0.537, 0.648] \rangle & \langle [0.268, 0.336], [0.457, 0.664] \rangle & \langle [0.130, 0.181], [0.407, 0.534] \rangle \end{array} \right) \end{array}$$

where the third parameter as interval-valued $\left[\pi_{ij}^l, \pi_{ij}^u \right]$ is given by $\pi_{ij}^l = 1 - \mu_{ij}^u - \nu_{ij}^u$ and $\pi_{ij}^u = 1 - \mu_{ij}^l - \nu_{ij}^l$ $(i = 1, 2, 3, 4; j = 1, 2, 3)$.

Case 1. *Let* $X_B = \{X_1, X_2, X_3\}$, *i.e. All criteria are benefit criteria.*

Using Eqs. (8) and (9), the intervals of relative closeness coefficients of alternatives are obtained as

$$C_1 = [0.200, 0.278], C_2 = [0.317, 0.405]$$
$$C_3 = [0.259, 0.353], C_4 = [0.332, 0.429]$$

Using Eq. (10), the possibility degree matrix for pairwise comparison of alternatives is

$$P = \left(p_{ij} \right)_{4 \times 4} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{cccc} A_1 & A_2 & A_3 & A_4 \\ \left(\begin{array}{cccc} 0.5 & 0 & 0.1105 & 0 \\ 1 & 0.5 & 0.8022 & 0.3946 \\ 0.8895 & 0.1978 & 0.5 & 0.1099 \\ 1 & 0.6054 & 0.8901 & 0.5 \end{array} \right) \end{array}$$

Using Eq. (11), the performance scores of alternatives $A_i(i=1,2,3,4)$ as optimal degrees ξ_i are estimated as $\xi_1 = 0.134, \xi_2 = 0.308, \xi_3 = 0.225, \xi_4 = 0.333$.

Thus, the ranking of alternatives becomes $A_4 \succ A_2 \succ A_3 \succ A_1$.

Case 2. Let $X_B = \{X_2\}$ and $X_C = \{X_1, X_3\}$. Risk and environmental impact analyses can be considered as cost criteria while growth analysis as benefit criteria.

Using Eqs. (8) and (9), the intervals of relative closeness coefficients of alternatives are obtained as

$$C_1 = [0.592, 0.665], C_2 = [0.501, 0.695]$$
$$C_3 = [0.513, 0.602], C_4 = [0.502, 0.590].$$

Using Eqs. (10) and (11), the ranking of alternatives becomes $A_1 \succ A_2 \succ A_3 \succ A_4$.

Remark 5.1 In the above illustrative example, it is found that the choices of criteria in the proposed TOPSIS approach have great impact on the ranking of alternatives. Li [12] did not distinguish the nature of criteria and ranked the alternatives as $A_2 \succ A_4 \succ A_3 \succ A_1$ which has some resemblance with the ranking $A_4 \succ A_2 \succ A_3 \succ A_1$ under *Case 1* based on the proposed approach. The deviations could be further explained by the fact that the advanced method followed three-parameter characterization of IFS in contrast with two-parameter characterization by Li [12]. Also, weighted distance in Li [12] is a pseudo-metric while normalized hamming distance [14] is a metric.

6 Concluding Remarks

Szmidt and Kacprzyk [14] demonstrated that the third parameter π (intuitionistic index) does have an influence on distance measure in IF settings. So, normalized hamming distance containing three parameters is used in this study. Mathematical programming-based TOPSIS technique is presented for solving MADM problems in IVIF environment providing a pair of linear fractional programming models resulting in intervals for closeness coefficients of alternatives. The possibility degree matrix for pairwise comparisons of closeness coefficients is generated and optimal degrees are estimated for ranking of alternatives. The ranking order of alternatives resulted from the proposed model has some resemblance with that of Li [12] and the deviations are explained. The simple calculations involved in the proposed model makes it worth for many applications. Hence, the advanced approach has the power to model imprecision and uncertainty of real-life situations in more reliable and effective manner transforming it into more realistic model. For future study, the weighted absolute distance between IFSs in three-parameter formulation can be used to extend TOPSIS in IVIF ambience extensively.

References

1. Atanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)
2. Atanassov, K.T., Gargov, G.: Interval-valued intuitionistic fuzzy sets. Fuzzy Sets Syst. **31**(3), 343–349 (1989)
3. Biswas, A., De, A.K.: A fuzzy programming method for solving multiobjective chance constrained programming problems involving log-normally distributed fuzzy random variables. Lect. Notes Comput. Sci. **7677**, 442–450 (2012)
4. Biswas, A., Dewan, S.: Priority based fuzzy goal programming technique for solving fractional fuzzy goals by using dynamic programming. Fuzzy Inf. Eng. **4**, 165–180 (2012)
5. Biswas, A., Modak, N.: Using fuzzy goal programming technique to solve multiobjective chance constrained programming problems in a fuzzy environment. Int. J. Fuzzy Syst. Appl. **2**, 71–80 (2012)
6. Biswas, A., Modak, N.: A fuzzy goal programming technique for multiobjective chance constrained programming with normally distributed fuzzy random variables and fuzzy numbers. Int. J. Math. Oper. Res. **5**, 551–570 (2013)
7. Chen, C.T.: Extension of the TOPSIS for group decision-making under fuzzy environment. Fuzzy Sets Syst. **114**(1), 1–9 (2000)
8. Chen, T.Y.: The inclusion-based TOPSIS method with interval-valued intuitionistic fuzzy sets for multiple criteria group decision making. Appl. Soft Comput. **26**, 57–73 (2015)
9. Facchinetti, G., Ricci, R.G., Muzzioli, S.: Note on ranking fuzzy triangular numbers. Int. J. Intell. Syst. **13**, 613–622 (1998)
10. Hwang, C.L., Yoon, Y.: A state of the art survey. In: Multiple Attribute Decision Making: Methods and Applications. Springer, Berlin (1981)
11. Kumar, S., Biswas, A.: Use of possibility measures for ranking of interval-valued intuitionistic fuzzy numbers in solving multicriteria decision making problems. Commun. Comput. Inf. Sci. **776**, 153–167 (2017)
12. Li, D.F.: Linear programming method for MADM with interval-valued intuitionistic fuzzy sets. Expert Syst. Appl. **37**, 5939–5945 (2010)
13. Pal, B.B., Biswas, A.: Fuzzy goal programming approach for solving bilevel decentralized planning problems. Int. J. Math. Sci. **6**, 391–401 (2007)
14. Szmidt, E., Kacprzyk, J.: Distances between intuitionistic fuzzy sets. Fuzzy Sets Syst. **114**(3), 505–518 (2000)
15. Xu, Z.S., Da, Q.L.: A possibility based method for priorities of interval judgment matrices. Chin. J. Manag. Sci. **11**, 63–65 (2003)
16. Xu, Z.S.: Methods for aggregating interval-valued intuitionistic fuzzy information and their application to decision making. Control Decis. **22**(2), 215–219 (2007)
17. Xu, Z., Yager, R.R.: Dynamic intuitionistic fuzzy multi-attribute decision making. Int. J. Approx. Reason. **48**(1), 246–262 (2008)
18. Ye, J.: Multicriteria fuzzy decision-making method based on a novel accuracy function under interval-valued intuitionistic fuzzy environment. Expert Syst. Appl. **36**, 6899–6902 (2009)
19. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
20. Zeng, S., Xiao, Y.: TOPSIS method for intuitionistic fuzzy multiple-criteria decision making and its application to investment selection. Kybernetes **45**(2), 282–296 (2016)
21. Zhao, X.: TOPSIS method for interval-valued intuitionistic fuzzy multiple attribute decision making and its application to teaching quality evaluation. J. Intell. Fuzzy Syst. **26**(6), 3049–3055 (2014)

Bounding Stability in Formal Concept Analysis

Bikram P. Bhuyan, Arindam Karmakar and Shyamanta M. Hazarika

Abstract In Formal Concept Analysis, stability is an important utility measure to rank concepts. However, computation of stability is considered to be a hard problem. Efficient algorithms having good bounds to estimate stability holds promise. In this paper, an effective graph-based technique is proposed to estimate stability. Our estimation algorithm has a polynomial time complexity of $O(|A|^2)$ where $|A|$ is the number of vertices.

Keywords Formal concept analysis · Algorithms · Graph algorithm · Estimation algorithm · Combinatorial problem

1 Introduction

Formal Concept Analysis (FCA) acquires grouping called formal concepts from Boolean data comprising objects and attributes [1]. A formal concept is a pair comprising a. extent—containing the objects which belong to the concept, and b. intent—include all the properties and meanings which apply to all objects in the extent [2]. The set of all formal concepts of a given context forms a complete lattice, called the concept lattice. However, such approaches using concept lattices often encounter the problem of the exponential growth rate of concepts in the lattice with the size of

B. P. Bhuyan (✉)
Asian Institute of Management and Technology, Guwahati, Assam, India
e-mail: bikram23bhuyan@gmail.com

A. Karmakar
Biomimetic and Cognitive Robotics Lab, Computer Science and Engineering,
Tezpur University, Tezpur, India
e-mail: arindam@tezu.ernet.in

S. M. Hazarika
Department of Mechanical Engineering, Indian Institute of Technology
Guwahati, Guwahati, India
e-mail: s.m.hazarika@iitg.ernet.in

the context [3, 4]. A measure to rank concepts is often used; thereby only relevant concepts are retained.

In this regard, 'stability' was introduced as one such ranking measure [1]. However, the computation of stability is proven to be an NP-complete problem [5, 6]. One of the best known algorithms was proposed by Obiedkov et. al [7] whose complexity is itself quadratic to the size of the lattice. Therefore, efficient algorithms to estimate stability hold promise. Although there exist many algorithms to estimate stability in literature [5, 8, 9], there lacks an efficient method having good bounds to estimate stability index. In this paper, a new technique is proposed to estimate the stability of a formal concept exploiting results on bounds for counting cliques in spectral graphs.

2 Terminology and Main Definitions

2.1 Formal Concept Analysis

Formal Concept Analysis was first developed by Wille [10] and is implemented in many fields like psychology, sociology, anthropology, medicine, biology, linguistics, computer sciences, mathematics and industrial engineering [11]. Here we briefly present some basic definitions of FCA [10, 12]. The root of FCA is a formal context from which concepts are generated and are organized in a hierarchical order in the form of a lattice. Formal context is a triple (G, M, I) where G denotes the set of objects, M is the set of attributes and I is a binary relation between G and M, i.e. $I \subseteq G \times M$. $(g, m) \in I$ indicates that object g has the property (attribute) m. If $A \subseteq G$ and $B \subseteq M$, then the following derivation operators produce the Galois connection:

$$A' = \{m \in M | (g, m) \in I, \forall g \in A\} \tag{1}$$

(the set of attributes common to the object set in A)

$$B' = \{g \in G | (g, m) \in I, \forall m \in B\} \tag{2}$$

(the set of objects common to the attribute set in B)

A formal concept of the formal context (G, M, I) is a pair (A, B) where $A \subseteq G$ and $B \subseteq M$, satisfying $A' = B$ and $B' = A$. We also have a property $A'' = A$ and $B'' = B$ where $(.)''$ is the closure operator. Thus, a formal concept consists of a set A (extent of the concept) and a set B (intent of the concept) such that all the objects in A share a common attribute set B.

A partial order relation $\leq$, showing the subconcept–superconcept hierarchy is present in the set $\beta(G, M, I) = \{(A, B) | A' = B \text{ and } B' = A\}$. It is defined by $(A_1, B_1) \leq (A_2, B_2)$ iff $A_1 \subseteq A_2$ and $B_2 \subseteq B_1$. A complete lattice is formed by $\beta(G, M, I)$ satisfying $\leq$.

2.2 Stability

Stability was first introduced in [5, 6], it was later revised in [7, 13] and is defined as follows:

Definition 1 Let $\Gamma = (G, M, I)$ be a formal context and (A, B) be a formal concept of Γ; then the stability of the concept $\gamma(A, B)$ is defined as

$$\gamma(A, B) = \frac{|\ \{X \subseteq A | X' = B\}\ |}{2^{|A|}} \tag{3}$$

Stability measures how much a concept intent depends on objects (extent) of the concept. It is to be noticed that stability index is between 0 and 1.

As enumerating stability is an NP-complete problem, estimation to compute stability was done by Buzmakov et al. in their paper [8]. They obtained an estimated lower bound and upper bound on the value of the stability index and merged both together. This amalgamation was termed is '*bounding method*'. But the tightness of the bound cannot be ensured. Monte Carlo approximation was first discussed in [9] but it was found to be less efficient than the bounding method.

It is to be mentioned that the state of art methods exploit the ordering property of the lattice where the lattice is created at the outset, and then the stability of a superconcept is estimated with the help of the subconcepts. It would be better if only the stable concepts are projected at the onset, and only then, the lattice is generated with only those stable concepts.

Our approach to estimate stability involves the reduction of the stability function, i.e. the set $\{X \subseteq A | X' = B\}$ of a concept (A, B) to a graph problem to estimate the stability index of that concept without taking the subconcepts into consideration.

## 3 Our Approach to Estimate Stability

The stability index is mapped to a graph F. A reduced graph $\overline{F}$ of F is taken by deleting some particular vertices and edges. More than one connected component could be a result of the action. Our goal is now to find the total number of cliques in each connected component and finally summing them to get the total number of cliques in $\overline{F}$.

Before proceeding, we would like to give symbols which we will be frequently using. For a given concept (A, B) in a context $\gamma = (G, M, I)$:

1. A is a set of objects $\{a_1, a_2, \ldots, a_n\}$ and A_k is the subset of k objects of A.
2. $|A|$ denotes the cardinality of A.
3. 2^A denotes the power set of A.
4. B is the set of attributes/properties common to the set A.

Fig. 1 Mapping of the
stability index of concept
({1,2,3} {a,b,g}) of the
context 'Living Beings and
Water' [12] to a graph is
shown, where 'solid circles'
represent V_{stb},
'dashed-circles' represent
$\overline{V}_{stb}$, 'solid lines' represent
E_{stb} and 'dashed-lines'
represent $\overline{E}_{stb}$
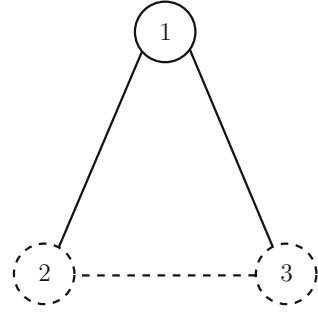
Lemma 1 *Let $\Gamma = (G, M, I)$ be a formal context and (A, B) be a formal concept of Γ; then the stability of the concept (A, B); can also be expressed as*

$$\gamma(A, B) = 1 - \frac{|\ \{X \subseteq A | X' \neq B\}\ |}{2^{|A|}} \tag{4}$$

We now map our problem of computing the stability of a concept (A, B) to a graph problem.

Let $F = (V, E)$ be a graph having the ordered pair (V, E) as the set of vertices and edges, respectively. For any concept (A, B), each object a_i of A correspond to a vertex V of the graph F. Therefore vertex set of F is $V = \{a_1, a_2, \ldots, a_n\}$. All the vertices of F are mutually connected to form a complete graph.

We now have the following definitions:

Definition 2 A vertex $a_i \in V_{stb}$, iff $a_i' = B$. If $a_i' \neq B$ then $a_i \in \overline{V}_{stb}$ (refer Fig. 1).

Definition 3 An edge $(a_i, a_j) \in E_{stb}$, iff $\{a_i, a_j\}' = B, i \neq j$. Otherwise; $(a_i, a_j) \in \overline{E}_{stb}$ (refer Fig. 1).

The set of $\overline{V}_{stb}$ vertices can further be classified into two sets- $\overline{V}_{stb_e}$ and $\overline{V}_{stb_n}$. $\overline{V}_{stb_e}$ are incident to at least one $\overline{E}_{stb}$ edge. And $\overline{V}_{stb_n}$ are the $\overline{V}_{stb}$ vertices which are not adjacent to any $\overline{E}_{stb}$ edge.

We state a proposition from Hui-lai Zhi et al. in [14].

Proposition 1 [14] *Let $\Gamma = (G, M, I)$ be a context, and (A, B) be a concept of the Γ. For any $s \subseteq A$, if $s' = B$ then $g' = B$ for all $s \subseteq g \subseteq A$.*

Observation 1 *Edges between a V_{stb} vertex and any other vertex are always E_{stb} edge.*

It should be noted that an E_{stb} edge may also exist between two $\overline{V}_{stb}$ vertices.

The vertices and edges can also be termed as 1-cliques and 2-cliques, respectively.

A *clique* can be defined as a maximal complete subgraph of a given graph [15]. A clique of size k is called a k-clique. 1-cliques corresponds to vertices, 2-cliques to edges and 3-cliques to 3-cycles (triangles).

Observation 2 *Let (A, B) be a concept of Γ and A_k is a subset of k objects of A. If there is at least one E_{stb} edge in the k-clique formed from A_k, then $A'_k = B$.*

Observation 3 *Let the graph F be reduced to $\overline{F}$ by deleting all V_{stb} vertices and E_{stb} edges. For any subset A_k of the reduced graph, $A'_k \neq B$.*

Observation 4 *After deletion of V_{stb} vertices and E_{stb} edges; the stability index computed from the reduced graph $\overline{F} \subseteq F$ will be same as the stability index calculated from the complete graph.*

Now, in the subgraph $\overline{F}$ consisting of $\overline{V}_{stb}$ and $\overline{E}_{stb}$ as the set of vertices and edges, respectively, there could be more than one connected components. A *connected component* of an undirected graph is a subgraph, satisfying the criteria where edges connect any two vertices of the subgraph and the same is not connected to any other vertices belonging to the supergraph.

There is a linear time algorithm to compute the connected components in a graph [16]. And for each concept, we have to find the $\overline{V}_{stb}$ vertices and $\overline{E}_{stb}$ edges which will form one or more connected components graph and for each such component, we have to find the total number of cliques which will finally result in the total number of cliques in the whole graph. But finding the total number of cliques in a graph takes exponential time [15]. Bollobás and Nikiforov [17], Nikiforov [18] have provided an estimate of the largest eigenvalue of the adjacency matrix of a graph based on the number of cliques using the idea of number of walks.

We have the following result by Nikiforov [17]:

Result 1 [17] *Let $\overline{F}$ be a graph, where $\hat{Z}_i$ denote the count of all i-cliques of $\overline{F}$ and λ be the largest eigenvalue of the adjacency matrix of $\overline{F}$. Then for every graph $\overline{F}$ and $r \geq 2$*

$$\lambda^{r+1} \leq (r+1)\hat{Z}_{r+1} + \sum_{s=2}^{r}(s-1)\hat{Z}_s\lambda^{r+1-s} \tag{5}$$

Lemma 2 *Given any formal context $\Gamma = (G, M, I)$ where (A, B) is a concept of the Γ, then the set ($\{X \subseteq A | X' \neq B\}$) of the concept (A, B) can always be reduced to a unique graph which will consist of one or many n-vertex, connected component graph.*

Theorem 1 *Given any formal context $\Gamma = (G, M, I)$ and (A, B) is a concept of the Γ; the stability index $\gamma(A, B)$ is at most $1 - \frac{\{|\hat{Z}(min)| + |\overline{V}_{stb_n}| + 1\}}{2^{|A|}}$ where $|\hat{Z}(min)|$ is the minimum number of cliques of the graph $\overline{F}$.*

Theorem 2 *Given any formal context $\Gamma = (G, M, I)$ and (A, B) is a concept of the Γ; the stability index $\gamma(A, B)$ is at least $1 - \frac{\{|\hat{Z}(max)| + |\overline{V}_{stb_n}| + 1\}}{2^{|A|}}$ where $|\hat{Z}(max)|$ is the maximum number of cliques of the graph $\overline{F}$.*

Theorems 1 and 2 give new upper and lower bounds to the stability index. We now present the procedure formally in Algorithm 1.

Algorithm 1 Stability Index

Compute stability indexes of all concepts of a given formal context $\Gamma = (G, M, I)$
Input: Formal Context
Output: Stability Indexes of all concepts

1: **procedure** CALCULATE–STABILITY
2:      **for** each concept $\in \Gamma$ **do**
3:          Find the set of $\overline{V}_{stb}$ ($\overline{V}_{stb_e}$ and $\overline{V}_{stb_n}$) and $\overline{E}_{stb}$ at the generation time of each concept.
4:          Find the connected components (CC) of the graph taking $\overline{V}_{stb_e}$ and $\overline{E}_{stb}$ as the set of vertices and edges, respectively [16].
5:          **for** each connected component (CC) **do**
6:              Create adjacency matrix D of the vertices.
7:              Compute λ.
8:              Compute $\hat{Z}(max_i)$ and $\hat{Z}(min_i)$
9:          **end for**
10:         Compute $\hat{Z}(max)$ and $\hat{Z}(min)$
11:         Compute the upper bound and lower bound of stability index as in Theorems 1 and 2.
12:     **end for**
13: **end procedure**

---

4   Time Complexity and Performance Analysis

The asymptotic time complexity of creating a lattice of formal concepts using FCA algorithms like Fast Close by One (FCbO) is $O(|G|^2|M|L)$ [19], where $|G|$ is the number of objects, $|M|$ is the number of attributes and L is the number of concepts. The time complexity of finding connected components in a graph is $O(|V| + |E|)$ [16]. Also, the time complexity of finding the largest eigenvalue of an adjacency matrix is $O(|V|^2)$ where $|V|$ and $|E|$ are the number of vertices and edges in a (V, E)-graph.

Thus, our proposed method has a time complexity (finding the stability of all concepts in the lattice) of $O(|G|^2|M|L + L|A|^2)$ where $|A|$ is the number of vertices in the reduced graph F of a concept. And the time complexity of finding stability of a single concept is $O(|A|^2)$.

The bounds were tested (Figs. 2 and 3) against the '*bounding method*' [8] discussed earlier. It was found that our algorithm results in stricter bounds in case of small concepts having high stability whereas the bounding method is not giving good estimation. With the increase in the number of concepts (large concepts), it is seen that both bounds are portraying relatively similar results. But in case of concepts having very low stability or as in our case after the graph formation, the eigenvalue is large ($\lambda = n - 1$) being a complete graph (K_n); our bounds do not give good results (because of the characteristic polynomial). So if ($\lambda = n - 1$), we do not compute the stability, as it will be a very low and of little use.
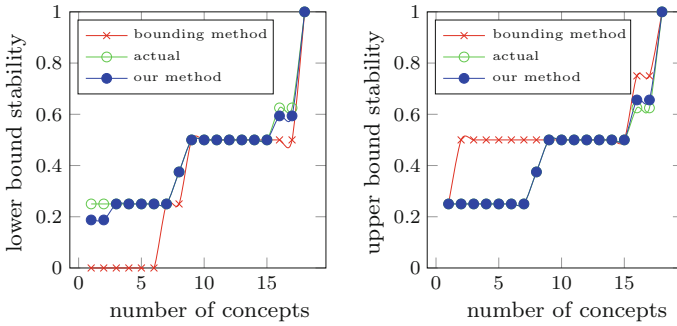
Fig. 2 Lower bound and upper bound comparison of the context 'Living Beings and Water' [12]
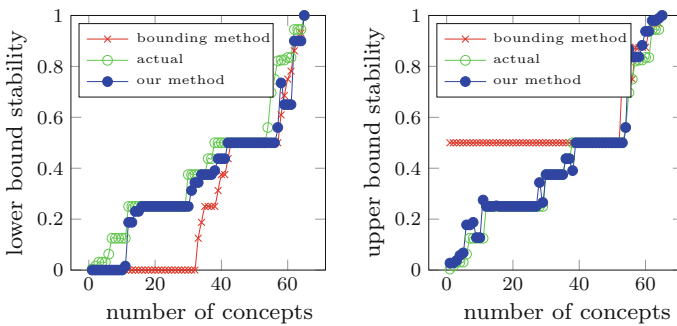

Fig. 3 Lower bound and upper bound comparison of the context 'Tea Ladies' [20]

5 Conclusion

Estimation of stability index for a formal concept is a hard problem. We have put forward an algorithm exploiting results from spectral graph theory to estimate stability. Its strength lies in the fact that it does not take the ordering property of the lattice into consideration and each concept is treated individually. Our algorithm has a polynomial time complexity of $O(|A|^2)$ where $|A|$ is the number of vertices in the reduced graph F of a concept.

As the estimation of the stability index is found to be dependent on the estimation of total clique counts, so in future, better estimation of the latter will result in even better bounds for stability.

References

1. Belohlavek, R., Trnecka, M.: Basic Level in Formal Concept Analysis: Interesting Concepts and Psychological Ramifications. In IJCAI, pp. 1233–1239, June, 2013
2. Wille, R.: Formal concept analysis as mathematical theory of concepts and concept hierarchies. Formal Concept Anal. **3626**, 1–33 (2005)
3. Carpineto, C., Romano, G.: Concept Data Analysis: Theory and Applications. Wiley (2004)
4. Kuznetsov, S.O.: On computing the size of a lattice and related decision problems. Order **18**(4), 313–321 (2001)
5. Kuzuetsov, S.O.: Stability as an Estimate of thie Degree of Substantiation of Hypotheses Derived on the Basis of Operational, Similarity (1990)
6. Kuznetsov, S.O.: On stability of a formal concept. Ann. Math. Artif. Intell. **49**(1), 101–115 (2007)
7. Roth, C., Obiedkov, S., Kourie, D.G.: On succinct representation of knowledge community taxonomies with formal concept analysis. Int. J. Found. Comput. Sci. **19**(02), 383–404 (2008)
8. Buzmakov, A., Kuznetsov, S.O., Napoli, A.: Is concept stability a measure for pattern selection? Procedia Comput. Sci. **31**, 918–927 (2014)
9. Babin, M., Kuznetsov, S.: Approximating concept stability. Form. Concept Anal. 7–15 (2012)
10. Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Ordered Sets, pp. 445–470. Springer, Netherlands (1982)
11. Wolff, K.E.: A first course in formal concept analysis. StatSoft **93**, 429–438 (1993)
12. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer Science & Business Media (2012)
13. Roth, C., Obiedkov, S., Kourie, D.: Towards concise representation for taxonomies of epistemic communities. Lect. Notes Comput. Sci. **4923**, 240–255 (2008)
14. Zhi, H.L.: On the calculation of formal concept stability. J. Appl. Math. (2014)
15. Moon, J.W., Moser, L.: On cliques in graphs. Israel J. Math. **3**(1), 23–28 (1965)
16. Hopcroft, J., Tarjan, R.: Efficient algorithms for graph manipulation (No. STAN-CS-71-207). Department of Computer Science, Stanford University, California (1971)
17. Bollobs, B., Nikiforov, V.: Cliques and the spectral radius. J. Comb. Theory Ser. B **97**(5), 859–865 (2007)
18. Nikiforov, V.: Some inequalities for the largest eigenvalue of a graph. Comb. Probab. Comput. **11**(2), 179–189 (2002)
19. Kuznetsov, S.O., Obiedkov, S.A.: Comparing performance of algorithms for generating concept lattices. J. Exp. Theor. Artif. Intell. **14**(2–3), 189–216 (2002)
20. Wille, R.: Concept lattices and conceptual knowledge systems. Comput. Math. Appl. **23**(6–9), 493–515 (1992)

Feature Extraction Using Fuzzy Generalized Two-Dimensional Inverse LDA with Gaussian Probabilistic Distribution and Face Recognition

Aniruddha Dey, Shiladitya Chowdhury and Jamuna Kanta Sing

Abstract This paper proposes a feature extraction technique called Gaussian probabilistic fuzzy generalized two-dimensional Fisher's inverse linear discriminant analysis (GPFG-2DILDA) method based on fuzzy set theory, Gaussian probabilistic distribution information and inverse LDA. Like the FG-2DLDA, the proposed GPFG-2DILDA method also maximizes class separability along x- and y-axes directions simultaneously. The proposed method first calculates fuzzy membership matrix by fuzzy k-nearest neighbour (Fk-NN) algorithm. These values are combined with the training samples to obtain the global mean and class-wise mean training images. Thereafter, fuzzy membership values are integrated into intra-class and inter-class scatter matrices along both (x- and y-) directions. Similarly, Gaussian probabilistic distribution information is incorporated into the intra-class scatter matrices. Finally, by solving the eigenvalue problems of these scatter matrices, we find the optimal Gaussian-fuzzy inverse projection vectors, which actually used to generate more discriminant features and to solve the binary classification problem. The GPFG-2DILDA method has been evaluated on the AT&T face database to demonstrate the efficacy of the proposed method over some state-of-the-art face recognition methods.

A. Dey (✉) · J. K. Sing
Department of Computer Science & Engineering, Jadavpur University, Kolkata, India
e-mail: anidey007@gmail.com

J. K. Sing
e-mail: jksing@ieee.org

S. Chowdhury
Department of Master of Computer Application, Techno India, Kolkata, India
e-mail: dityashila@yahoo.com

1 Introduction

The feature extraction and face recognition technique has emerged as an active research area in the last couple of decades in the field of computer vision, pattern recognition and biometrics [1, 2]. Face recognition problem can be expressed as the classification of one or more test images by comparing with faces stored in databases [1]. The face recognition task is still very challenging under uncontrolled environment and different issues related to human face, i.e. facial expression (unhappiness, happiness) and facial pose, occlusion (wearing beards, moustaches and glasses), illumination variance, etc. The subspace-based methods, like principal component analysis (PCA), linear discriminant analysis (LDA) or Fisher's linear discriminant (FLD) and their variants, which use eigenfaces and/or Fisherfaces to compute features, fall under this category. In particular, LDA has been proposed as a better alternative to the PCA to provide class discrimination information [3, 4]. The main objective of the LDA is to find best discrimination of vectors among the classes by maximizing the between-class differences and minimizing the within-class ones [3]. The disadvantage of LDA technique is that it suffers from the 'small sample size (SSS)' problem [4]. The aforementioned problem mainly occurs in case of few numbers of samples than the sample dimension. The dimension of face images is generally very high which leads the singular within-class scatter matrix, which makes the FLD method infeasible. The SSS problem in LDA can be solved by sampling down the face images into smaller size. Two-dimensional FLD (2DFLD) method maximizes the class separability in one direction (row or column) at a time [4, 5]. The significant characteristic of 2DFLD method is that it directly works on the 2D image matrices. This method is more efficient feature extraction and face recognition technique than PCA and 2DPCA methods [2]. The projection vectors are extracted from both directions of the training face images simultaneously, by the G-2DFLD method [4]. The discriminant feature matrices are found by linearly projecting a face image matrix on both directions mentioned above. Therefore, the discriminative information is maximized by this method among the classes while minimizing it within a class [4, 5].

Recently, researchers proposed fuzzy-based approaches, such as fuzzy k-nearest neighbour (Fk-NN), fuzzy Fisherface, fuzzy two-dimensional Fisher's linear discrimination (F-2DFLD) and fuzzy two-dimensional principal component analysis (F-2DPCA) for facial feature extraction and face recognition [4–10]. Keller et al. [6] presented the fuzzy k-nearest neighbour (Fk-NN) approach, which fuzzifies the class assignment. The well-established Fisherface method was further extended by combining fuzzy logic which produces a significant result in face recognition technique [5]. The fuzzy 2DFLD (F-2DFLD) is an extension of the fuzzy Fisherface [5]. The scatter matrices were redefined by introducing membership values into each training sample. Yang et al. proposed feature extraction using fuzzy inverse FDA [7]. The Fk-NN was also incorporated into fuzzy inverse FDA for calculating membership degree matrices. Recently, the Gaussian probability distribution is incorporated into the definition of intra- and inter-class scatter matrices [8].

In this study, we have introduced a GPFG-2DILDA method where the Gaussian membership grades of the training samples are incorporated. More, specifically, the proposed method is influenced or extended from our previously presented method FG-2DLDA. In this context, it may be noted that the membership grades of the training samples are obtained by means of the fuzzy k-NN method. Thereafter, the mean training image matrices (global and class-wise) are calculated using the membership values. The intra- and inter-class scatter matrices are found distinctly from x- and y-directions from the face image. Gaussian PDF information is incorporated into the intra-class scatter matrices. Finally, inverse LDA is calculated to solve binary classification problem which is existed for overlapping sample. In this regard, the above scatter matrices are more influenced by each training sample.

The rest of the paper is organized as follows. Section 2 defines the review of FG-2DLDA method. We proposed the idea and described proposed GPFG-2DILDA method in Sect. 3. The empirical results on the face database (ORL) and comparing with other methods are given in Sect. 4. Finally, concluding remarks are summarized in Sect. 5.

2 Fuzzy G-2DLDA Method

The FG-2DLDA method assumes that the face image can be a part of several classes with varying membership values. The fuzzy k-NN [6] method is used to obtain the membership values which is contributed to spawn the fuzzy class-wise mean images and fuzzy global mean image [4]. Then, the scatter matrices (inter- and intra-class) along x- and y-axes to find the mean image. The projective matrices are then found by evaluating the aforementioned scatter matrices.

The scatter matrices are then redefined by integrating the fuzzy membership (M_F) with the training images [4]. The eigenvalue problems of these scatter matrices are solved to generate the optimal fuzzy projection vectors. So, the fuzzy membership matrix (M_F) using the Fk-NN [6] can be illustrated as given below:

$$M_F = [\nu_{ci}]; \quad c = 1, 2, 3, \ldots, C; \quad i = 1, 2, 3, \ldots, N \tag{1}$$

Let the training set contains N images of C classes (subjects) and each one is denoted as $X_i (i = 1, 2, 3 \ldots, N)$ having dimension as $r \times s$. The cth class C_c, has total N_c images and satisfies $\sum_{c=1}^{C} N_c = N$. For an image X, the FG-2DLDA-based features in the form of 2D matrix of size $u \times v$ are generated by projecting it onto the optimal fuzzy projection matrices and can be achieved by the following linear transformation [4]:

$$Y^f = \left(P_{opt}^f\right)^T X \left(Q_{opt}^f\right) \tag{2}$$

Step 1: Compute the global mean training image $(\boldsymbol{\psi})$ and class-wise mean training image $(\boldsymbol{\psi}_c)$. The fuzzy membership degrees are incorporated into the training images to obtain the global- and class-wise mean images [4].

Step 2: Calculate the fuzzy image scatter matrices, i.e. intra-class and inter-class, along row $\left(\boldsymbol{\xi}_{rb}^f, \boldsymbol{\xi}_{rw}^f\right)$ and column $\left(\boldsymbol{\xi}_{cb}^f, \boldsymbol{\xi}_{cw}^f\right)$ directions with dimensions $r \times r$ and $s \times s$, respectively [4].

Step 3: Compute two Fisher's criteria (objective function) along row $\left(J^f(P)\right)$ and column $\left(J^f(\boldsymbol{Q})\right)$ directions [4].

Step 4: Normalize the eigenvalues $\left(\boldsymbol{\xi}_{rb}^f \left(\boldsymbol{\xi}_{rw}^f\right)^{-1}\right)$ and eigenvectors $\left(\boldsymbol{\xi}_{cb}^f \left(\boldsymbol{\xi}_{cw}^f\right)^{-1}\right)$ to find the optimal projection vectors, $\boldsymbol{P}_{opt}^f$ and $\boldsymbol{Q}_{opt}^f$ [4].

Step 5: Sort the eigenvalues in descending order. The eigenvectors are rearranged according to the eigenvalues [4].

3 Gaussian Probabilistic Fuzzy G-2DILDA Method

Human faces are highly variable in nature, especially under different environmental conditions, such as illumination, pose, etc. As a result, sometimes, images of a person may look alike to that of a different person. In addition, variability among the images of a person may rise quite significantly. Therefore, fuzzy class assignment for the training images may be suitable instead of binary class assignment. The proposed GPFG-2DILDA method is basically based on the concept of fuzzy class assignment, where a face image belongs to different classes as characterized by its fuzzy membership values. The idea of fuzzification using fuzzy *k*-nearest neighbour (F*k*-NN) was conceived by Keller et al. [6] and found to be more effective. In the present study, the fuzzy membership values are incorporated into the training images to obtain global- and class-wise mean images, which in turn used to form fuzzy between-class and fuzzy within-class scatter matrices. Gaussian PDF information is incorporated into the intra-class scatter matrices. Therefore, these scatter matrices yield useful information regarding the association of each training image into several classes. The optimal Gaussian-fuzzy 2D projections vectors are obtained by evaluating the eigenvalue problems of these scatter matrices. Inverse LDA is calculated to solve binary classification problem. Finally, the GPFG-2DILDA-based features are extracted by projecting a face image onto these optimal Gaussian-fuzzy 2D inverse projection vectors. The different steps of the GPFG-2DILDA method are presented in detail in the following subsections.

3.1 Creation of Membership Matrix Using Gaussian Probability Density Function

The proposed method, for each class, generates the membership values are obtained using the Gaussian probability density function (PDF). The training set can be redefined as follows:

Let, $X = \{X_i\}(i = 1, 2, \ldots, N)$ represents the training set where N is the number of total images of C different classes $\{C_c\}(c = 1, 2, \ldots, C)$. The cth class C_c, has total N_c images and satisfies $\sum_{c=1}^{C} N_c = N$. Moreover, the mean image for class C_c can be represented as $\nu_c = \frac{1}{N_c} \sum_{i=1|X_i \in C_c}^{N} X_i$. As discussed earlier, the membership grades are calculated by the underlying Gaussian PDF [8],

$$\zeta_{ci} = \frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{\|X_i - \nu_c\|^2}{2\sigma_c^2}} \tag{3}$$

where ζ_{ci} represents the membership grade of the image (X_i); σ_c denotes the mean of Euclidean distance (inter) between the images of C_c class.

The Gaussian membership matrix (G_M) finally is obtained by means of membership values every sample for all classes [8]. The equation can be expressed as follows:

$$G_M = [\zeta_{ci}]; \quad c = 1, 2, \ldots, C, \quad i = 1, 2, \ldots, N \tag{4}$$

3.2 Creation of Membership Matrix Using Fk-NN

Fk-NN-based decision algorithm has been performed for assigning membership values (degree) to the training images [5, 6]. This method redefines the membership values of the labelled face images. When, all of the neighbours belong to the ith class which is equal to the class of jth image under consideration, then $n_{ij} = k$ and μ_{ij} returns 1, making membership values for the other classes as zero. So, the fuzzy membership matrix F_M using the Fk-NN can be illustrated as given below:

$$F_M = [\mu_{ci}]; \quad c = 1, 2, 3, \ldots, C; \quad i = 1, 2, 3, \ldots, N \tag{5}$$

3.3 GPFG-2DILDA Algorithm

The next step of the proposed GPFG-2DILDA method is to integrate the fuzzy membership values with the training images and redefine the scatter matrices along

x- and *y*-axes directions. Finally, generate the optimal fuzzy projection vectors by solving the eigenvalue problems of these scatter matrices.

Let the training set contains N images of C classes (subjects) and each one is denoted as $X_i (i = 1, 2, 3 \ldots, N)$ having dimension as $r \times s$. The *c*th class C_c, has total N_c images and satisfies $\sum_{c=1}^{C} N_c = N$. Let $\bar{\tau}$ is the global mean image and $\bar{\tau}_c$ is the mean image of the *c*th class. Fuzzy membership degrees are incorporated into the training images to get global mean image and class-wise mean images, which are defined as follows [4]:

$$\bar{\tau}_c = \frac{\sum_{i=1}^{N} \boldsymbol{\mu}_{ci} X_i}{\sum_{i=1}^{N} \boldsymbol{\mu}_{ci}}; \quad c = 1, 2, 3, \ldots, C \tag{6}$$

$$\bar{\tau} = \frac{\sum_{c=1}^{C} \sum_{i=1}^{N} \boldsymbol{\mu}_{ci} X_i}{\sum_{c=1}^{C} \sum_{i=1}^{N} \boldsymbol{\mu}_{ci}} \tag{7}$$

Calculate the Gaussian-fuzzy image scatter matrices, i.e. intra-class and inter-class, along row $\left(FS_{rb}^{gp}, FS_{rw}^{gp}\right)$ and column $\left(FS_{cb}^{gp}, FS_{cw}^{gp}\right)$ directions with dimensions $r \times r$ and $s \times s$, respectively. The Gaussian probabilistic membership values are integrated within row- and column-wise fuzzy image scatter matrices $\left(FS_{rw}^{gp}, FS_{cw}^{gp}\right)$. These Gaussian-fuzzy scatter matrices $\left(FS_{rb}^{gp}, FS_{rw}^{gp}, FS_{cb}^{gp}, FS_{cw}^{gp}\right)$ are expressed as follows:

$$FS_{rb}^{gp} = \frac{1}{N} \sum_{c}^{C} N_c^{f} (\bar{\tau}_c - \bar{\tau})(\bar{\tau}_c - \bar{\tau})^{T} \tag{8}$$

$$FS_{rw}^{gp} = \frac{1}{N} \sum_{c}^{C} \sum_{i \in c}^{N} \zeta_{ci} (X_i - \bar{\tau}_c)(X_i - \bar{\tau}_c)^{T} \tag{9}$$

$$FS_{cb}^{gp} = \frac{1}{N} \sum_{c}^{C} N_c^{f} (\bar{\tau}_c - \bar{\tau})^{T}(\bar{\tau}_c - \bar{\tau}) \tag{10}$$

$$FS_{cw}^{gp} = \frac{1}{N} \sum_{c}^{C} \sum_{i \in c}^{N} \zeta_{ci} (X_i - \bar{\tau}_c)^{T}(X_i - \bar{\tau}_c) \tag{11}$$

The inverse Fisher's criteria (objective function) $J_{FG-2DILDA}^{gp}(P)$ and $J_{FG-2DILDA}^{gp}(Q)$ along *x*- and *y*-axes directions are defined as follows:

$$J_{FG-2DILDA}^{gp}(P) = \arg\min_{P_f^{gp}} \frac{\left|(P_f^{gp})^{T} FS_{rw}^{gp} P_f^{gp}\right|}{\left|(P_f^{gp})^{T} FS_{rb}^{gp} P_f^{gp}\right|} \tag{12}$$

and

$$J_{FG-2DILDA}^{gp}(\boldsymbol{Q}) = \arg \min_{Q_f^{gp}} \frac{\left|(\boldsymbol{Q}_f^{gp})^T \boldsymbol{FS}_{cw}^{gp} \boldsymbol{Q}_f^{gp}\right|}{\left|(\boldsymbol{Q}_f^{gp})^T \boldsymbol{FS}_{cb}^{gp} \boldsymbol{Q}_f^{gp}\right|} \qquad (13)$$

The ratio is maximized in the above two Eqs. (12) and (13) when the column vectors of the projection matrix $\boldsymbol{P}_f^{gp}$ and $\boldsymbol{Q}_f^{gp}$ are the eigenvectors of $\boldsymbol{FS}_{rw}^{gp}\left(\boldsymbol{FS}_{rb}^{gp}\right)^{-1}$ and $\boldsymbol{FS}_{cw}^{gp}\left(\boldsymbol{FS}_{cb}^{gp}\right)^{-1}$, respectively.

For an image X, the GPFG-2DILDA-based features in the form of 2D matrix of size $u \times v$ is generated by projecting it onto the optimal Gaussian-fuzzy projection matrices and can be achieved by the following linear transformation:

$$Y_0 = \left(\boldsymbol{FP}_{opt}^{gp}\right)^T \boldsymbol{X} \left(\boldsymbol{FQ}_{opt}^{gp}\right) \qquad (14)$$

4 Empirical Results

The performance of the GPFG-2DILDA method has been evaluated over the AT&T face databases [11]. The average recognition rate of the proposed method is compared with some competent methods to demonstrate its superiority. Furthermore, we have tested the performance of the proposed method on the ORL face database [11]. The database includes the images of 40 different individuals each having 10 dissimilar images with 400 images in total. In our study, we have selected the s images randomly from the image pool for each individual to generate the training set and the rest of the images were taken for creating test set. Therefore, the training and test set comprises completely distinct set of images. The values of s are taken as 4, 5, and 6 form different pairs of training and test sets. Figure 1 represents the face images of an individual.

Furthermore, the proposed GPFG-2DILDA method is also evaluated for $s = 4$, 5, and 6 by varying feature matrices shown in Fig. 2 on the AT&T face database using RBF-NN [3]. The proposed method yields the best average recognition rates of 96.03%, 98.12%, and 98.88% for $s = 4$, 5, and 6, respectively.

The GPFG-2DILDA method is compared with the FG-2DLDA [4], FLDE [9], G-2DFLD [4], C-FLDA [10], RF-LDA [10], 2DFLD [4] and 2DPCA [4] methods. Table 1 illustrates the best average recognition rates achieved by this algorithm for

Fig. 1 Image samples of an individual from the AT&T database

Fig. 2 Average recognition rates of the GPFG-2DILDA method on the AT&T database by varying the feature matrices for different values of s
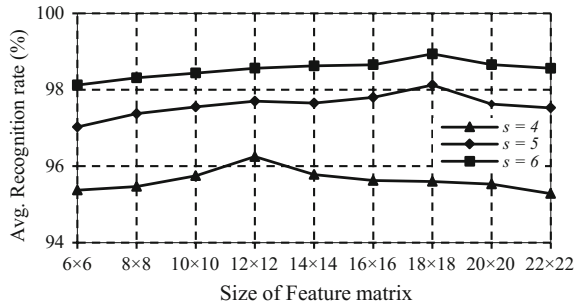
Table 1 Comparison in terms of average recognition rates (%) on the AT&T database

Method	Avg. recognition rates (%)		
	$s = 4$	$s = 5$	$s = 6$
GPFG-2DILDA	**96.03**	**98.12**	**98.88**
FG-2DLDA [4]	–	98	–
FLDE [9]	93.8	95.5	98.1
G-2DFLD [4]	95.94	97.68	98.72
C-FLDA [10]	92.77	96.5	98.81
RF-LDA [10]	92.5	96.47	98.62
2DFLD [4]	95.08	97.50	98.26
2DPCA [4]	94.33	96.83	97.72

different combinations of training and test set on the AT&T face database. The results demonstrate that the GPFG-2DILDA method outperforms the other methods.

5 Conclusion

This study intended to propose a new feature extraction method, namely, the Gaussian probabilistic fuzzy generalized two-dimensional Fisher's inverse linear discriminant analysis (GPFG-2DILDA) method. This method is based on the fuzzy set theory, Gaussian probabilistic distribution information and inverse LDA. Like FG-2DLDA, the GPFG-2DILDA methods maximize class separability along horizontal and vertical directions, simultaneously. The membership values are obtained by the fuzzy k-nearest neighbour (Fk-NN) algorithm and the values are combined with the training samples to obtain the global mean and class-wise mean training images. The fuzzy membership values and Gaussian PDF information are integrated into intra-class and inter-class scatter matrices along both (x- and y-axes) directions. The eigenvalue problems of these scatter matrices are then solved to find the optimal Gaussian-fuzzy inverse projection vectors, which actually used to generate more discriminant features and solve binary classification problem.

The GPFG-2DILDA method is validated on the ORL face database. In this regard, the proposed method outperforms some state-of-the-art face recognition methods.

References

1. Zhou, H., Mian, A., Wei, L., Creighton, D., Hossny, M., Nahavandi, S.: Recent advances on singlemodal and multimodal face recognition: a survey. IEEE Trans. Hum. Mach. Syst. **44**, 701–716 (2014)
2. Poon, B., Amin, M.A., Yan, H.: Performance evaluation and comparison of PCA based human face recognition methods for distorted images. Int. J. Mach. Learn. Cybernet. **2**, 245–259 (2011)
3. Sing, J.K., Basu, D.K., Nasipuri, M., Kundu, M.: Face recognition using point symmetry distance-based RBF network. Appl. Soft Comput. **7**, 58–70 (2007)
4. Dey, A., Sing, J.K.: Face recognition by Fuzzy generalized 2DFLD method and RBF neural network classifier. In: Proceedings of the WCI 2015, pp. 1–6 (2015)
5. Yang, W., Wang, J., Ren, M., Yang, J.: Fuzzy 2-dimensional FLD for face recognition. J. Inf. Comput. Sci. **4**, 233–239 (2009)
6. Keller, J.M., Gray, M.R., Givens, J.A.: A fuzzy *k*-nearest neighbor algorithm. IEEE Trans. Syst. Man Cybernet. **15**, 580–585 (1985)
7. Yang, W., Wang, J., Ren, M., Zhang, L., Yang, J.: Feature extraction using fuzzy inverse FDA. Neurocomputing **72**, 3384–3390 (2009)
8. Sing, J.K.: A novel Gaussian probabilistic generalized 2DLDA for feature extraction and face recognition. In: Proceedings of the CGVIS 2015, pp. 258–263 (2015)
9. Huang, P., Yang, Z., Chen, C.: Fuzzy local discriminant embedding for image feature extraction. Comput. Electr. Eng. **46**, 231–240 (2015)
10. Song, X.N., Zheng, Y.J., Wu, X.J., Yang, X.B., Yang, J.Y.: A complete fuzzy discriminant analysis approach for face recognition. Appl. Soft Comput. **10**, 208–214 (2010)
11. AT&T face database. AT&T Laboratories, Cambridge, U.K. [Online]. http://www.uk.research.att.com/facedatabase.html

Secure and Efficient Data Sharing with User Revocation in Cloud

Nalini Sri Mallela and Nagaraju Devarakonda

Abstract In the recent trends, secure and efficient plan is required to get the access control of shared data in distributed systems. Different ABE schemes are provided for secure data sharing but, in CP-ABE scheme, a set of attributes are associated with each user the data on to user was encrypted based on access structure formed from the set of attributes. Any user who satisfies the access structure of the encrypted text they can only the decrypt the encrypted text. In data outsourcing environment, if you use this attribute-based encryption (ABE) results some challenges in case of attribute and user revocation. In smart grid provides sensitive data sharing policies and schemes because it has to deal with sensitive information and maintains flexibility for giving information about the policies used for the data protection or authentication details of data owner and receiver. A movement toward secure and efficient data sharing with fine grain access control. We propose a scheme that enables the "secure data sharing and key distribution" with communication channels, and the users can obtain their secret keys from the "key generation center (KGC)" and our scheme can accomplish "fine grain access control" and the user in the users' list can use the data available in the cloud and restricted access is only allowed for the revoked users. Our proposed scheme enables the secure user revocation, key distribution, data confidentiality, and fine grain access control in cloud.

Keywords Fine grain access control · User revocation · Key distribution
Data confidentiality · Cloud computing

N. S. Mallela (✉) · N. Devarakonda
Lakireddy Bali Reddy College of Engineering, Mylavaram, India
e-mail: mallela.nalini@gmail.com

N. Devarakonda
e-mail: dnagaraj_dnr@yahoo.co.in

563

1   Introduction

In the recent trends, we live in the information world. The data may be personal data or organizational data people are interested to store the data on the centralized servers may be easy usage or cost reduction. People always wish to make their data as private and accessible only to the trusted users. The security issues are playing more important role in current trend of information world and become more urgent in the cloud computing where users, organizations, and business may outsource their data including the different types of data and personal data or any sensitive information to obtain the security of data from the cloud providers and other nodes in the network different security techniques are needed. To achieve the better security we can use any security service depending on the requirement.

In the recent trends in addition to the storage, the outsourcing requires effective security for the data with access control techniques that are cryptographically efficient [1, 2]. The attribute-based encryption (ABE) is able reach these requirements [3–6]. ABE can able to perform the access control over encrypted text using different access policies and set of attributes. Ciphertext-policy ABE (CP-ABE) is capable in ascendable way of data encryption that encryptor fixes the set of attributes and then the decryptor needs to have these attributes to decrypt the encrypted text [5]. To share the private information secretly, smart grids use the x.509 certificates for the proper identification and authentication checking process [2]. It requires more time and processing for updating the keys securely or revoke the users. For revoked users also keys should be provided.

For achieving the fine grain access control, cryptographic technique available is attribute-based encryption (ABE) [2–4, 7]. CP-ABE (ciphertext-policy attribute-based encryption) enables the sender to fix the set of attributes in the attributes list. The receiver should contain these selected set of attributes to decrypt the encrypted text [3].

2   Related Work

Yu et al. [8] introduced a new encrypted access control scheme. In attribute-based encryption (ABE) scheme, the user key and encrypted text are associated with set of attributes. Only if at least some attributes matched between the user key and the encrypted text, then the user can decrypt the encrypted text. The successive works are focused on the multiparty communication. Hur and Noh [1] considered the issues in data outsourcing. But secure data outsourcing requires proper authentication verification by the defined policies and updates in the policies are also required. The simple ABE scheme achieves the fine grain access control by the dual encryption mechanism with group key distribution in each attribute group. Hur [2] introduced secure data sharing among smart grids with hidden policies. Smart grids always take the sensitive information to share between different parties with proper

authentication by using the tokens, which are still a problem in sharing the encrypted data. Huang et al. [5] proposed a multiparty access control model for distributing encrypted text and provided the propagator to update the policy of encrypted text if the existing policy attributes are satisfied. Nowadays, mainly organizations' (government or private) data is outsourced by using a cloud. Samanthula et al. [6] proposed a method for outsourcing the data with proper identity which achieves confidentiality. CP-ABE partial decryption of the ciphertext with the token is allowed in this scheme and the result is stored in the storage center and sends to the receiver. With help of secret key of the receiver, the partially decrypted text will convert to plain text which can be readable by the receiver. Here we are concentrating against data protection in cloud. But we get out the decrypted data, since data is not in our protection scope. In this paper, our scheme provides secure and efficient data sharing and it achieves anti-collusion, fine grain access control, secure user key revocation, and secure key distribution in cloud.

3 System Model and Design Goals

Contributions of our scheme

- Secure communication channels are used for the distribution of keys. The user can securely get their keys from the key generation center with approval of authentication server.
- The registered data owners can easily upload their data in encrypted form into the cloud and we can obtain the fine grain access control with data owners list in server. Only registered data owners are included in the list.
- The user can obtain corresponding private keys from the key generation center. By using these keys user can able to retrieve decrypted data from cloud.
- If the user revocation is required then user approaches the revocation center and produces proper ID and request for user revocation key.
- This information updates in server and then data owner list is updated.
- Then, the above procedure was repeated for secure data upload and retrieval for revoked user.

3.1 System Model

System model includes six entities: user, cloud, data owner, authentication server, key generation center, and user revocation center (Fig. 1).

- **Cloud** supervised by the service providers, provides storage space for data owners to load their files on pay as you way. But cloud is untrusted because cloud providers may be compromised and will attempt to take the content of the files.
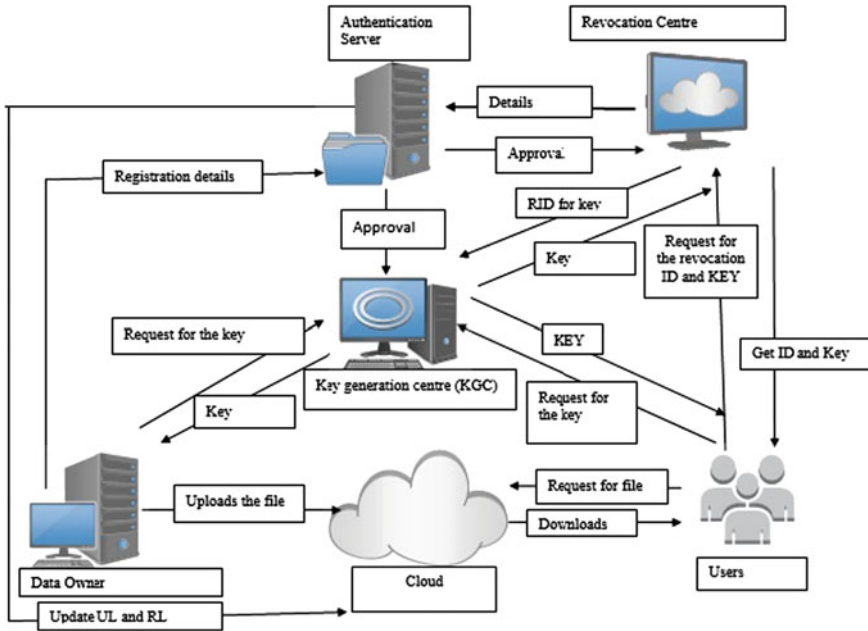
Fig. 1 System model

- **User** set of registered users access the data from the cloud and may be share with the others.
- **Data owner** who upload the data files into the cloud with endorsement of the authentication server.
- **Key Generation Center** generates the keys for user and the data owner by considering approval from the authentication server.
- **Authentication server** checks the user and data owner authentication details for key generation and generation of revocation ids.
- **User revocation center** generates the revocation ids for users who need the revocation by checking all the details.

3.2   Design Goals

Our proposed scheme includes fine grain access control, collusion attack, and secure user revocation scheme.

- *Fine grain access control*, any user in the user list can utilize the source in the cloud and cloud will not provide access privilege for the revoked users.

- *Collusion attack,* original data files will not reveal to the revoked users even they contrive with the untrusted cloud because they all stored in the secret form.
- *Secure user revocation scheme*, it provides the flexibility for users will not update their keys even new user added to the list or a user revoked from the list.

4 The Proposed Scheme

4.1 Scheme Description

Our scheme includes existing user registration, key generation, data owner uploads the data file, user revocation, new user registration, and data file download (Fig. 2).

4.1.1 Existing User Registration and Key Generation

First user sends IDu, Ru, Nu, AC, Tu as request to the authentication. By the approval of AS, the KGC generates the Tgt_u and (Pu, Qu) and sends to the user and cloud. All the details of user store in the cloud as a user list UL. By carrying attributes like ticket, identity of user and IDu, Tgtu, Ru and Qu user request for the file access from the cloud if these attributes satisfies the cloud allowed for data file access (Fig. 3).

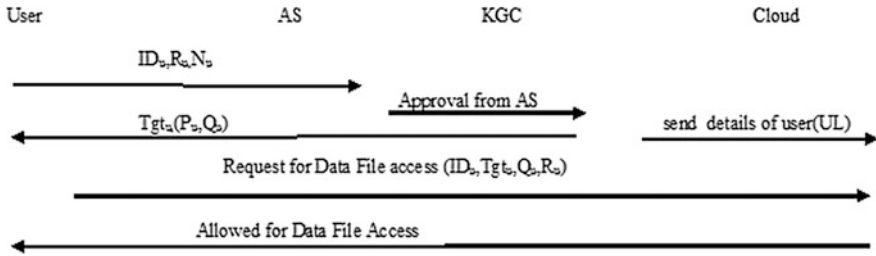| Notation	Description
ID_u	Identity of the User
R_u	Random number allocated for the user
N_u	Network address of the user
Tgt_u	Ticket allocated for user by KGC
P_u	Public key for the user
Q_u	Private key for the User
UL	Users List
RL	Revoked Users List
Tu	Time stamp
AC	User Account

Fig. 2 Notations used in our scheme

Fig. 3 Registration for existing user

4.1.2 Data Owner Uploads the Data File

Data owner sends IDdw and Rdw to the authentication server and gets approval from AS in form of ticket (Tgtdw) and sends approval message to the KGC and then KGC generates a key Sdw to the data owner. The data file is encrypted with the secret key by the data owner. The encrypted data file was uploaded into the cloud (Fig. 4).

4.1.3 User Revocation

User approaches the revocation center by producing its IDu and Ru. These details were forwarded to the authentication Server and checks authenticity details and lifetime of the user includes in the revocation list RL and update in user list UL. It sends approval to KGC and revocation center. User was assigned by unique revocation Id (RIDu). By producing this id user request the revocation center for revocation token (RTu). By carrying this token user request KGC and obtain the key for accessing data file. Here, by unique token and unique revocation id, we can able to avoid the collusion attack between the revoked users (Fig. 5).
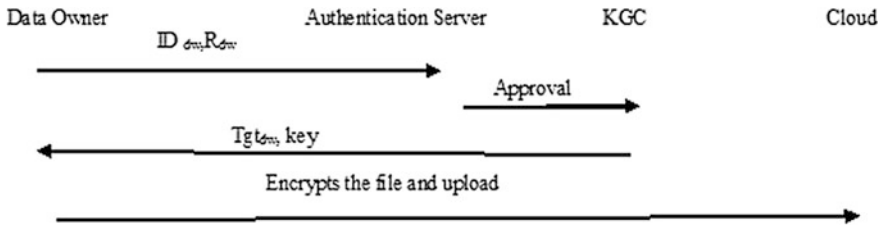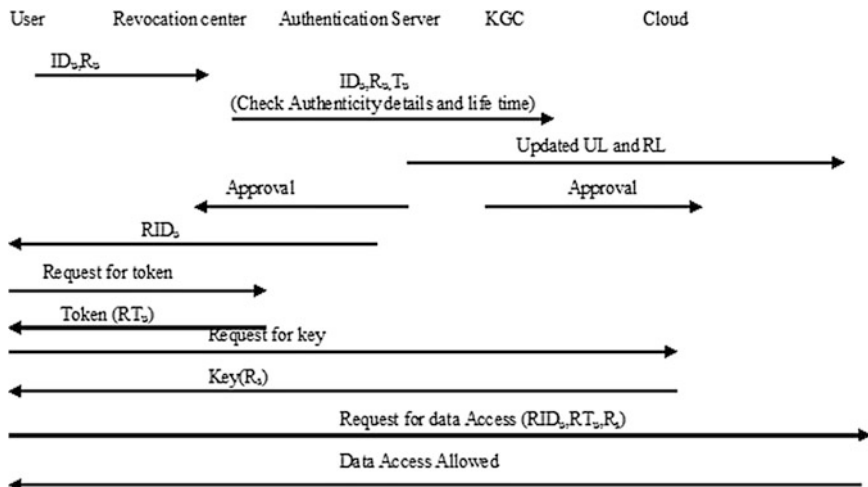

Fig. 4 Uploading the data file

Fig. 5 User revocation

4.1.4 New User Registration

New users Registered with their identity IDu+1 in cloud. Same actions are performed as existing user registration. The users' list was updated and sent to the KGC, which is illustrated in Fig. 6.

4.1.5 Data File Download

User downloads the data files from the cloud. Data owner encrypted the file with the key (Q_u) (from KGC) and uploaded into the cloud. User wants download the file user produces the key (from KGC) and the cloud checks key and time stamp in the user list if it is available cloud provides the file in the encrypted form to the user. User decrypts file with his key (P_u) (Fig. 7).
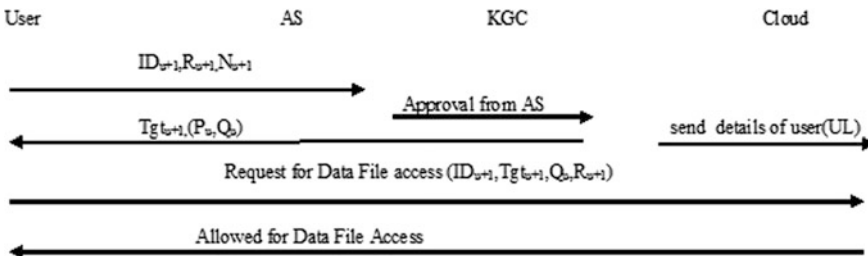
Fig. 6 Registration of new user

Fig. 7 File download
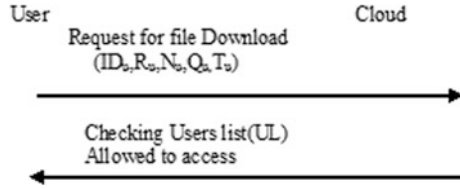
5   Security Analysis

In this section, we prove the security of our scheme in terms of key distribution, data confidentiality, and fine grain access control.

5.1   Key Distribution

In the proposed scheme, corresponding keys are generated by the KGC (key generation center) and communicated securely through the secure channels. Data owner keys, user keys, and revocation keys all are generated by the KGC only.

In registration process, user sends ID_u, R_u, N_u, and $R_u \in Z^*$. The KGC computes the corresponding values or keys P_u, Q_u by verifying the details of the user. P_u is the public key for the user, and Q_u is the private key for the user. These two keys securely distributed to the users

$$P_u = f\ (ID_u,\ Ru,\ Nu,\ Tu)$$

$$Q_u = f\ (ID_u,\ Ru,\ Nu,\ Tu).$$

The attacker bypassed the authentication checking process then also he would not get the data. Because he needs the value Random number Ru but Ru is unique to the particular user. For other people same random number was not generated. So the attacker unable to get the keys. If attacker tries to learn anything, it has other network address. So attacker unable to get the keys

$$N_u \in U^*_w$$

5.2   Data Confidentiality

Our proposed scheme provides "data confidentiality". The cloud was incapable to learn the content of data file because the encrypted file was uploaded and key was not given to the cloud.

$$DF = Encry \ (DF \| K_{dw})$$

The cloud was incapable to learn the content of data file because the data file which is loaded by the data owner is in encrypted form. The data file was encrypted by data owner keys. The keys are secure not known by the cloud so unable to get the file

$$DF = Encry \ (DF) \| K_{dw} \| ID_{dw} \| T_{dw}$$

If the collusion occurs in that case also the cloud was unable to read the content of data files. When the user is revoked, the authentication server updates the user list and prepares newly revoked users' list RL. Both lists are updated in the cloud. When the user is revoked new revocation token was assigned to the user with that token user gets the key from the KGC and data file encrypted and stored in the cloud. The key is given to the user only not the cloud. So cloud was unable to learn the content of data file. With the help of unique token and revocation ID no collusion was happened

$$DF = Encry \ (DF) \| \ K_{ru} \| \ K_{dw} \| T_u \| \ RID_u \| Token_{ru}$$

5.3 Fine Grain Access Control

The proposed scheme achieves "fine grain access control" with clear differentiation between data owners, users, and revocation users.

Data owners are having unique identity and the keys; with these attributes, data owners upload the file securely in encrypted form. The encryption keys are obtained from the KGC. Users are having unique identity and keys from the KGC and unique random number so the keys of users are secure and no third party person will not the data. Authenticated users only access the data files. All the authentication details will be verified by the authentication server (AS). So finally, our scheme achieves fine grain access control

$$P_u = f \ (ID_u, R_u, N_u, T_u)$$

$$Q_u = f \ (ID_u, R_u, N_u, T_u)$$

$$N_u \in U^*_w$$

$$DF = Encry \ (DF \| K_{dw})$$

$$DF = Encry \ (DF) \| K_{dw} \| ID_{dw} \| T_{dw}$$

$$DF = Encry \ (DF) \| K_{ru} \| K_{dw} \| T_u \| \ RID_{u \|} \ Token \ _{ru.}$$

Table 1 Comparative analysis

	Key distribution	Access control	User revocation	Anti-collusion attack	Data confidentiality
ABE		√			
ABEDO		√	√	√	
Our scheme	√	√			√

5.4 Comparative Analysis

Our proposed scheme achieves protected key distribution, data confidentiality, Fine grain access control, User revocation for understanding clearly the advantages and security of our proposed scheme all details were mentioned in the (Table 1). The √ in the blank means the scheme can achieve the corresponding goal.
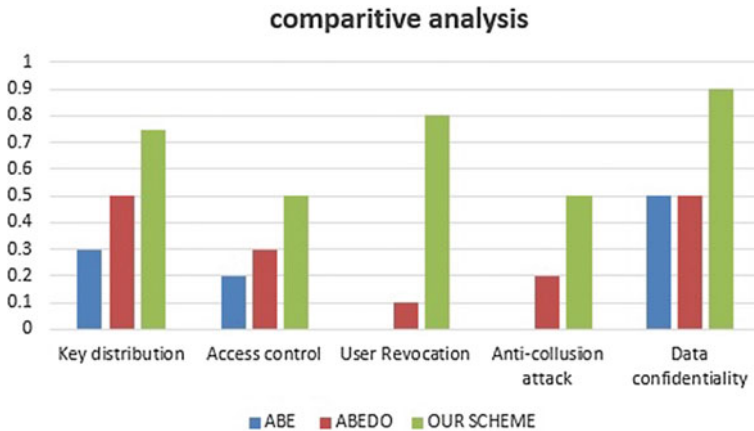
6 Conclusion

In this paper, we proposed a secure and efficient data sharing and anti-collusion in cloud. In our proposed scheme, users obtain the secret keys securely from the key generation center and authentication approval from the authentication server. Our scheme supports the user revocation and anti-collusion in cloud. New users added to the users' revocation list were updated and existing users' keys are not affected. Our scheme accomplish secure user revocation and revoked cannot able to get the data in cloud even with the compromised cloud.

References

1. Hur, J., Noh, D.K.: Attribute-based access control with efficient revocation in data outsourcing systems member. IEEE Trans. Parallel Distrib. Syst. **22**(7) (2011)
2. Hur, J.: Attribute-based secure data sharing with hidden policies in smart grid. IEEE Trans. Parallel Distrib. Syst. **24**(11) (2013)
3. Attribute-based data sharing to provide priority based security to improve the performance efficiency. Int. J. Comput. Trends Technol. (IJCTT) **14**(3) (2014)
4. Qinlong, H., Zhaofeng, M.A., Yixian, Y., Xinxin, N., Jingyi, F.: Improving Security and Efficiency for Encrypted Data Sharing in Online Social Networks Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing 100876, China
5. Huang, X., Liu, J.K., Tang, S., Xiang, Y., Liang, K., Xu, L., Zhou, J.: Cost-effective authentic and anonymous data sharing with forward security. IEEE Trans. Comput. **64**(4) (2015)
6. Samanthula, B.K., Elmehdwi, Y., Jiang, W.: K-nearest neighbor classification over semantically secure encrypted relational. IEEE Trans. Knowl. Data Eng. **27**(5) (2015)
7. Hur, J.: Improving security and efficiency in attribute-based data sharing. IEEE Trans. Knowl. Data Eng. **25**(10) (2013)
8. Yu, S., Wang, C., Ren, K.: Attribute based data sharing with attribute revocation In: ASIACCS'10, Beijing, China, 13–16 Apr 2010. Copyright 2010. ACM. ISBN 978-1-60558-936-7
9. Qinlong, H., Zhaofeng, M.A., Yixian Y., Jingyi, F., Xinxin, N.: EABDS: Attribute-based secure data sharing with efficient revocation in cloud computing (1. Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China) (2. National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing 100876, China) Chin. J. Electron. **24**(4) (2015)
10. Waters, B.: Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization. In: Proceedings of International Conference Practice Theory Public Key Cryptography Conf. Public Key Cryptography, pp. 53–70. 2008
11. Boneh, D., Boyen, X., Goh, E.: Hierarchical identity based encryption with constant size ciphertext. In: Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 440–456. 2005
12. Delerablee, C., Paillier, P., Pointcheval, D.: Fully collusion secure dynamic broadcast encryption with constant-size Ci-pher-texts or decryption keys. In: Proceedings of the International Conference on Pairing-Based Cryptography, pp. 39–59. 2007
13. Zhu, Z., Jiang, Z., Jiang, R.: The attack on mona: secure multi-owner data sharing for dynamic groups in the cloud. In Proceedings of the International Conference on Information Science and Cloud Computing, 7 Dec 2013, pp. 185–189
14. Zhou, L., Varadharajan, V., Hitchens, M.: Achieving secure role-based access control on encrypted data in cloud storage. IEEE Trans. Inf. Forensics Secur. **8**(12), 1947–1960 (2013)
15. Zou, X., Dai, Y.-S., Bertino, E.: A practical and flexible key management mechanism for trusted collaborative computing. In: Proceedings of the IEEE Conference on Computer Communications, pp. 1211–1219. 2008
16. Zhu, Z., Jiang, R.: A secure anti-collusion data sharing scheme for dynamic groups in the cloud. IEEE Trans. Parallel Distrib. Syst. **27**(1) (2016)

An Adaptive Cluster-Based Ensemble Learner for Computational Biology

Niti Jain and Ambar Maini

Abstract In quantitative biology, discovering a class when presented with a large bimolecular dataset poses a big problem. However, ensemble learning approach has been helpful in various complex areas of decision-making. So, in this paper, we propose a cluster-based ensemble learner called adaptive cluster-based ensemble learner (ACEL) which incorporates the prior knowledge of the datasets into the cluster ensemble framework. ACEL computes the cluster boundaries using three diverse clustering algorithms to obtain clusters for classification decision. ACEL learns by transforming the obtained clusters into rules and performing adaptive rule tuning to optimize the classification decision. The cluster-based classification results are then processed using majority voting algorithm. The proposed approach is compared with other supervised benchmark algorithms using seven problems from the field of biology. The experiments performed on benchmark datasets show that ACEL works effectively in classifying datasets.

Keywords Clustering · Adaptive learner · Ensemble · Supervised · Classification

1 Introduction

Adaptive learning is more effective than traditional non-adaptive learning algorithms and is better suitable for large-scale data [3]. In any expert system, before arriving at a conclusion, opinions from all the experts are taken into consideration and then the final decision is made. This is the principle behind ensemble learning [7]. In applications where the size of data is too large for a single classifier to analyse, ensemble

N. Jain (✉) · A. Maini
National Institute of Technology, Raipur, India
e-mail: nitijain305@gmail.com

A. Maini
e-mail: ambarmaini1993@gmail.com

systems partition the data into subsets where each classifier works on a subset of dataset and further combines the results using the existing approaches like majority voting, weighted majority voting, etc. [5]. There are two ways of combining the classifiers: classifier fusion and classifier selection [6]. In classifier fusion approach, all individual classifiers are trained on the whole dataset. Examples of this include bagging predictors and boosting [8]. In classifier selection approach, each individual classifier performs its best in some part of total dataset. There are two major components of any ensemble system. The first component is making a diverse ensemble. The second component is used to combine the output of decisions of the single classifiers.

In the real world, data consist of classes with overlapping boundaries. Excessive training will help solve this problem, but it will result in overfitting which will lead to misclassifications of testing data. Whereas learning generalized boundaries will not lead to overfitting but it will misclassify the overlapping patterns. Therefore, we opt to use clustering. Clustering makes it easy to learn the decision boundaries. Organization of data into groups is one of the fundamental methods of learning.

For each problem, let $x = \{x_1, x_2 \ldots x_n\}$ be a set of input vector in R^p and $y = \{y_1, y_2 \ldots y_n\}$, for a system given by S, where S transmutes x to y

$$y = S(x) \tag{1}$$

Here, $x = \{x_1, x_2 \ldots x_p\} \in R^p$ is an input vector and $y = \{y_1, y_2 \ldots y_r\} \in R^r$ is the output vector of a system. The purpose of our experiment is to identify a classification system that builds S to explain the given input–output data (x, y).

We present adaptive cluster-based ensemble learner (ACEL) in the following sections. Section 2 gives the literature review. Section 3 gives the description about proposed approach. Section 4 summarizes the experimental setup. Section 5 summarizes the results. Section 6 presents the conclusions and the future work in this field.

2 Literature Review

Clusters are dense regions which are separated by low-density regions in feature space. Several Bayesian approaches are used for data clustering like undirected graphical model. Ensemble classifier combines the result of various diverse base classifiers [13]. Diversity is a property used to define ensemble classifiers. Greater diversity is observed when incorrect decisions made by one of the classifiers are handled by the other classifiers. This results in uniform distribution of errors. To combine the results of base classifiers, various methods have been proposed including majority voting, weighted majority and decision template. Among the several

existing approaches for ensemble learning, boosting and bagging have been used to a greater extent [9, 10]. In bagging, the base classifiers learn on data subsets drawn randomly from entire training set, and the results are combined by majority voting [2]. In boosting, re-sampling of data instances is performed. The new learners work on the instances that are difficult to classify by the previous number of the ensemble. This mechanism encourages the construction of complementary learners.

Lately novel cluster ensemble technique, CE-GMDH was brought forward that comprises three parts: one initial approach, one conveying function and one external condition [11]. Experimentations were performed by CE-GMDH on artificial and real data. Yu et al. [14] has suggested a feature assortment oriented semi-supervised cluster ensembling technique for clustering of tumour obtained from instances of biomolecular datasets. A progressive semi-supervised clustering ensemble technique with arbitrary subspace method, limitation propagation and progressive ensemble member selection technique was brought forward by [15]. Alves et al. [1] has developed a methodology of ensembling by using multiple particle swarm optimization and demonstrated its ability to solve problems of computation biology.

3 Proposed Approach

In our proposed approach, we have performed homogeneous clustering for partitioning the patterns belonging to a single class only. Fixed number of rules for each class (here, one rule for each class) is used. Every class is represented by a combination of rules. To generate initial rule base, the training data are clustered using three different categories of clustering method k-means, fuzzy c-means (FCM) and particle swarm optimization (PSO) based clustering algorithm [4]. In order to catch each aspect of data learning process, three different varieties of clustering algorithms which are of different nature and can cluster using different approaches have been used.

Every single cluster represents a thick region in the input dataset which is depicted by the related cluster centroid. Every individual cluster is thereafter transformed into a rule, after which we perform adaptive rule tuning process to minimize the error function. The proposed cluster-based classification produces three diverse base classifiers. These base classifier's results are joined using majority voting algorithm which is used for class prediction. Majority voting technique is used to predict the ultimate classification decision. The majority voting algorithm can be represented as

$$\sum_{t=1}^{T} d_{t,J}(x) = max_{j=1,2,\ldots,c} \sum_{t=1}^{T} d_{t,j} \tag{2}$$
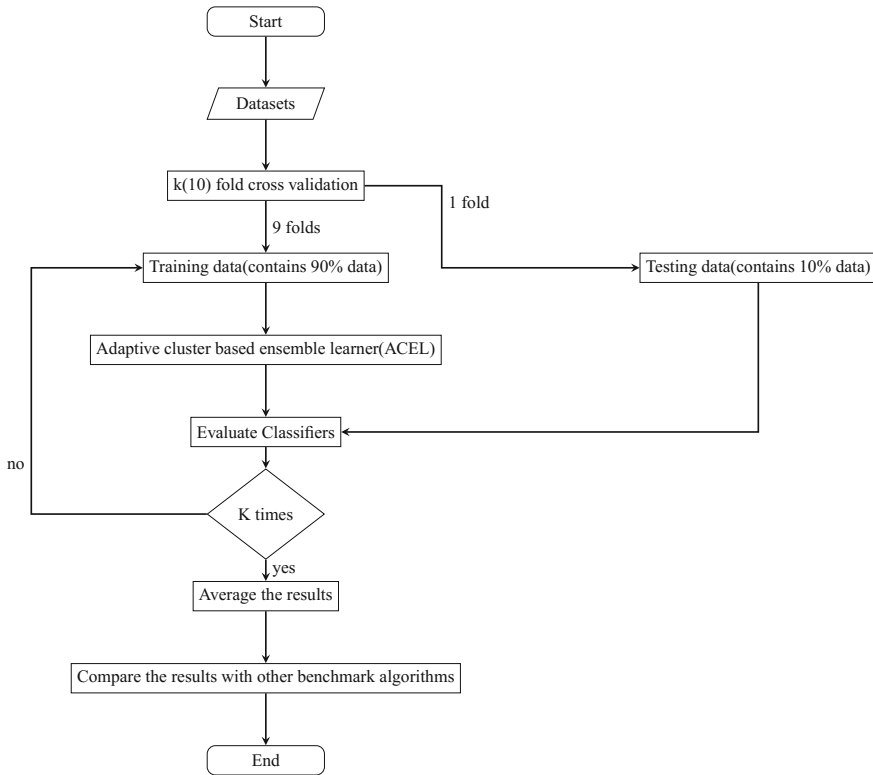
Fig. 1 Proposed approach

In situations where individual classifier decisions are not dependent on each other, it can be observed that majority voting combination technique will lead to a performance and accuracy improvement. The classification results are compared with other standard algorithms J48, Adaboost, SMO, Naive Bayes and Random Forest. Figure 1 depicts the proposed approach.

Adaptive learning should be stopped when the training error (or misclassification) reaches an acceptable level. The main goal of rule tuning is to remove irrelevant data associated with the cluster and to include the data which belong to the cluster. Rule tuning potentially increases the predictive power of the rule, helping to avoid overfitting to the training data. As soon as the misclassification reaches to zero or maximum iteration gets completed, the construction of the final rule, i.e. the rule tuning procedure completes.

The strategy for the rule tuning procedure is based on the concept of best centroid for minimizing sum of squared error (SSE), and it can be obtained by the mean of the points in the cluster.

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} (c_i - x)^2 \qquad (3)$$

Let C_i be the ith cluster, x is a point in C_i and c_i is the mean of the ith cluster. In order to find the best centroid which minimizes sum of squared error(SSE) to zero can be performed by the following differentiation for kth centroid c_k. m_k is the number of objects in kth cluster

$$\frac{\delta}{\delta c_k} SSE = \frac{\delta}{\delta c_k} \sum_{i=1}^{K} \sum_{x \in C_i} (c_i - x)^2$$
$$\frac{\delta}{\delta c_k} SSE = \sum_{i=1}^{K} \sum_{x \in C_i} \frac{\delta}{\delta c_k} (c_i - x)^2 \qquad (4)$$
$$\frac{\delta}{\delta c_k} SSE = \sum_{x \in c_k} 2 * (c_k - x_k)$$

Equating sum of squared error (SSE) to zero,

$$\frac{\delta}{\delta c_k} SSE = 0 \qquad (5)$$

Now combining (4) and (5)

$$\sum_{x \in c_k} 2 * (c_k - x_k) = 0 \qquad (6)$$

$$m_k c_k = \sum_{x \in c_k} x_k \qquad (7)$$

$$c_k = \frac{1}{m_k} \sum_{x \in c_k} x_k \qquad (8)$$

Hence, it can be observed that the best centroid for minimizing the SSE of a cluster is mean of all the points in a cluster.

We have followed the same principle and tuned centre accordingly in order to minimize the misclassification accuracy. The training and prediction methodology of rule tuning for adaptive learning is presented in Algorithm 1.

---

1 Algorithm: Rule Tuning(C,N) ;

   input : Centres obtained from diverse clustering algorithm C, Number
           of features in the training set N
   output: Rule Tuned Centers

2 // Tuning process is repeated until
     classification error is satisfactory for
     each feature of training set

3 **for** $i \leftarrow 1$ **to** N **do**

4   Compute mean m_i and standard deviation sd_i corresponding to i^{th}
       feature of training set;

5   Choose tuning parameter n_m;

6   Set initial value I_v to Minimum $(m_i\text{-}sd_i, c_i)$ and Max to $(m+sd_i)$;

7   Compute the error E_0^x and misclassification M_0^x for initial rule;

8   **while** $(I_v < Max)$ **do**

9 $I_v = I_v + n_m$;

10   Compute the error for the new rule base;

11   **if** $E_t^x > E_{t-1}^x$ **then**

12 $R^t \leftarrow R^{t-1}$ // since the error is increased,
                  we restore the values corresponding
                  to the base rule

13   **end**

14   **if** $M_t^x = 0$ *or* $E_t^x \approx 0$ **then**

15   Stop

16   **end**

17 $t \leftarrow t + 1$;

18   **end**

19 **end**

Algorithm 1: Rule tuning

---

4 Experimental Setup

In order to validate the proposed approach, the experiments were conducted on the benchmark datasets obtained from the UCI machine learning repository [12]. The ACEL algorithm has been applied on seven benchmark datasets. The datasets used are Iris, Thyroid, Balance Scale, Vertebral Column, Haberman's Survival, Liver Disorder and Diabetes.

Table 1 Datasets used

Dataset	Instances	Attributes	Classes
Iris	150	4	3
Thyroid	215	5	3
Balance Scale	625	4	3
Vertebral Column	310	6	3
Haberman's Survival	306	6	2
Liver Disorder	345	6	2
Diabetes	768	8	2

The result of the proposed approach has been compared with benchmark algorithms like J48, Adaboost, SMO, Naive Bayes and Random Forest. Table 1 gives a summary of datasets. The experiment was performed in MATLAB.

5 Results and Discussion

We have trained our learner ACEL on seven biological datasets taken from UCI machine learning repository [12]. To test cluster-based ensemble learner, we compare the classification results by ACEL and by other standard algorithms J48, Adaboost, SMO, Naive Bayes and Random Forest. Table 2 gives the accuracy corresponding to each dataset.

For the Thyroid dataset, ACEL gives 93.46 accuracy, performing similar to other state-of-art algorithms J48, Adaboost and outperforming SMO by almost 5%. For the Iris dataset, ACEL outperformed other algorithms and performed similar to SMO with an accuracy of 96.67. For the dataset Liver disorder, ACEL performs better than Naive Bayes and SMO by at least 15%. Haberman's Survival dataset using ACEL achieved 72.84 accuracy performing better than Random Forest by at least 8%. The algorithm gives an accuracy of 84.83 for the dataset Vertebral Column, which is

Table 2 Results in terms of classification accuracy

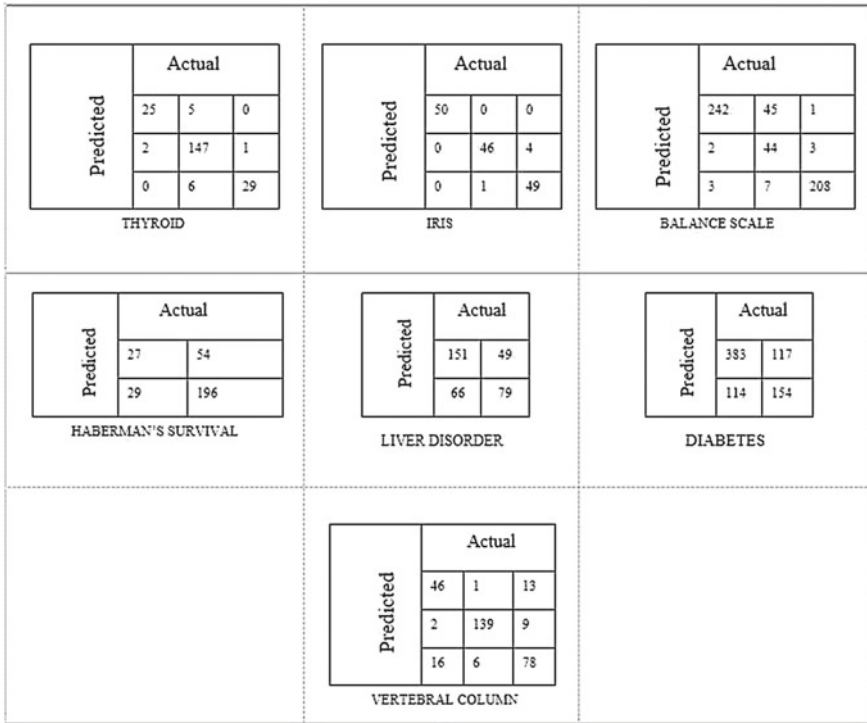| Algorithm	Thyroid	Iris	Liver disorder	Haberman's survival	Diabetes	Vertebral column	Balance scale
ACEL	93.46	96.67	66.6	72.84	69.91	84.83	79.04
J48	92.09	95.33	68.4	72.87	73.83	81.61	76.64
Adaboost	93.48	97.33	66.66	75.16	74.35	77.42	72.32
RF	95.35	94	74.49	66.99	74.349	84.19	81.6
NB	96.74	95.33	54.2	74.5	76.3	83.22	90.4
SMO	89.76	96.67	57.97	73.53	77.34	74.52	87.68

Fig. 2 Confusion matrices

better than Adaboost by almost 9% and SMO by more than 12%. In Balance Scale dataset, our proposed learner performs better than J48 and Adaboost algorithm.

The results obtained after performing ACEL algorithm suggest that applying clustering on the datasets and transforming the cluster to rules followed by adaptive tuning on these clusters optimizes the classification decision. The experiments performed on benchmark datasets show that ACEL works effectively in classifying datasets.

Figure 2 shows the confusion matrix when proposed ACEL algorithm is applied to each dataset. The actual and predicted labels give the number of instances that are correctly classified or misclassified. Confusion matrix can be used to describe the performance of a classification model. In confusion matrix, the count of true positive and true negative indicates how well a classification model works. In the figure, we can see that count of TP and TN is high for Iris, Thyroid, Balance Scale, Haberman's Survival and Vertebral Column suggesting high performance of ACEL.

6 Conclusions and Future Work

We have presented a novel cluster-based ensemble learner (ACEL) based on the principle of cluster ensemble learning along with rule tuning, which gives better adaptability and improved accuracy. The proposed algorithm has been evaluated on biological experimental datasets. The results of the experiments have shown that our ensemble approach has given comparable results to individual learners. The evidences from the experimental results show that adaptive cluster ensemble learning process using tuning improves accuracy to a greater extent. In our future research, we would like to focus on finding the optimal number of clusters and other critical issues in ensemble classification like integration mechanism and robustness.

References

1. Alves, P., Liu, S., Wang, D., Gerstein, M.: Multiple-Swarm ensembles: improving the predictive power and robustness of predictive models and its use in computational biology. IEEE/ACM Trans. Comput. Biol. Bioinform. **5963**, 1–1 (2017). https://doi.org/10.1109/TCBB.2017.2691329
2. Dudoit, S., Fridlyand, J.: Bagging to improve the accuracy of a clustering procedure. Bioinformatics **19**, 1090–1099 (2003). https://doi.org/10.1093/bioinformatics/btg038
3. Giotis, I., Petkov, N.: Cluster-based adaptive metric classification. J. Neurocomputing **91**, 33–40 (2012). https://doi.org/10.1016/j.neucom.2011.10.018
4. Jain, A.: Data clustering: 50 years beyond K-means. Pattern Recognit. Lett. **31**, 651–666 (2010). https://doi.org/10.1016/j.patrec.2009.09.011
5. Kittler, J.: Combining classifiers: a theoretical framework. Pattern Anal. Appl. **1**, 18–27 (1998). https://doi.org/10.1007/BF01238023
6. Kuncheva, L., Ll, G., Kingdom, U., Duin, R.: Decision templates for multiple classifier fusion: an experimental comparison. Pattern Recognit. **34**(2), 299–314 (2001). https://doi.org/10.1016/S0031-3203(99)00223-X
7. Polikar, R.: Ensemble based systems in decision making. IEEE Circuits Syst. Mag. **6**(3), 21–45 (2006). https://doi.org/10.1109/MCAS.2006.1688199
8. Quinlan, J.: Bagging, boosting, and C4.5. Proc. Thirteen. Natl. Conf. Artif. Intell. **5**, 725–730 (1996). https://doi.org/10.1023/A:1018054314350
9. Rodrguez, J., Maudes, J.: Boosting recombined weak classifiers. Pattern Recogn. Lett. **29**, 1049–1059 (2008). https://doi.org/10.1016/j.patrec.2007.06.019
10. Skurichina, M., Duin, R.: Bagging, boosting and the random subspace method for linear classifiers. Pattern Anal. Appl. **5**, 121–135 (2002). https://doi.org/10.1007/s100440200011
11. Teng, G., He, C., Xiao, J., He, J., Zhu, B., Jiang, X.: Cluster ensemble framework based on the group method of data handling. Appl. Soft Comput. **43**, 35–46 (2016). https://doi.org/10.1016/j.asoc.2016.01.043
12. UCI Machine Learning Repository. http://archive.ics.uci.edu/ml/
13. Xu, L., Krzyzak, A., Suen, C.: A method of combining multiple classifiers and their applications to handwriting recognition. IEEE Trans. Syst. Man Cybern. **22**, 418–435 (1992). https://doi.org/10.1109/34.368145

14. Yu, Z., Chen, H., You, J., Wong, H., Liu, J., Li, L., Han, G.: Double selection based semi-supervised clustering ensemble for tumor clustering from gene expression profiles. IEEE/ACM Trans. Comput. Biol. Bioinforma. **11**, 727–740 (2014). https://doi.org/10.1109/TCBB.2014.2315996
15. Yu, Z., Member, S., Luo, P., You, J., Wong, H., Leung, H., Wu, S.: Incremental semi-supervised clustering ensemble for high dimensional data clustering. Tkde **28**, 701–714 (2016). https://doi.org/10.1109/TKDE.2015.2499200

Text Document Analysis Using Map-Reduce Framework

K. V. Kanimozhi, P. Prabhavathy and M. Venkatesan

Abstract Due to the advance Internet and increasing globalization, the electronics forms of information grow in a rapid manner. Extracting the useful hidden information from those multiple documents is a recent challenge. Hence, efficient and automated clustering algorithm which is effective in identifying topics plays the main role in information retrieval. In this paper, the analysis regarding the large unstructured text document corpus using our proposed map-reduce algorithm has been performed, and the results show the advantage of the proposed method by detecting clusters of document features within less computation time and provides premier solution for increasing the precision rate of retrieval in information extraction.

Keywords Text documents · Similarity · Map-reduce · Text clustering

1 Introduction

Due to the fast Internet growth and advancement of digital technology, a huge amount of data is generated every day from various sources. Nowadays, analyzing the unstructured text data is a new field and it is a challenge to mine meaningful information from it. Since nature of the text is unstructured and massive in size, the traditional techniques and algorithm will not be applicable. Hence, parallelizing the existing algorithms using the big data tools and new automatic model is required to

K. V. Kanimozhi (✉) · P. Prabhavathy
VIT University, Vellore 632404, India
e-mail: kani_kayal@rediffmail.com

M. Venkatesan
NIT Surathkal, Mangalore 575025, India

mine the valuable information from unstructured text. Today, World Wide Web and social networks produce huge amounts of data daily. The social networks datas are like chat messages, twitter dataset, newsfeeds, forum messages, reviews about the products, movies, books, and news articles. Traditional machine learning algorithms are not scalable for these huge datasets. Finding documents by matching, clustering, and finding topics on this unstructured text pose latest challenges in various applications like pattern recognition, information retrieval, recommendation systems, intrusion detection, text summarization, and social media data analysis. The paper is ordered as follows: Sect. 2 provides the outline of some existing works regarding text mining and the topic modeling. Section 3 includes the problem statement and the proposed solution for topic modeling by our combined approach algorithm for unstructured text documents in big text datasets using map-reduce framework. In Sect. 4, the experiments are done on real-world data and evaluate the results of the algorithm and describe the explanations of proposed algorithm result compared with existing methodologies. Finally, the conclusion is discussed in Sect. 5.

2   Background Study

Topic modeling and the clustering of document corpus are mutually associated to each other. Topic modeling can project documents into a topic space which make easy and effective document clustering. Wei et al. [1] used a latent semantic indexing-based MLDC method to perform the clustering of multilingual documents, even though they achieve the satisfactory result but not able to handle the document management problems. Clifton et al. [2] infer the topics from the documents based on frequent item mining concept but cannot find the new knowledge when new information or new items found in a dataset is added to the topic, so scalability problem occurs and also performance is affected with incremental mining. Blei et al. [3] use LDA model for allocation of the words in a document to topics, and hence, many to many relationships are performed between topics and words and also between document and topics. Nagwani [4] again used k-means and LDA method for summarizing the text and topic modeling based on map-reduce framework. Kontostathis et al. [5] presented the performance of latent semantic indexing for the search and retrieval application. Xie et al. [6] proposed a multigrain clustering topic model (MGCTM) which integrates document clustering and topic modeling into a unified framework and jointly performs the two tasks to achieve the overall best performance and model tightly couples two components: a mixture component used for discovering latent groups in document collection and a topic model component used for mining multigrain topics including local topics specific to each cluster and global topics shared across clusters.

And they proved that topic modeling-based clustering methods including LDA + K-means, LDA + Naive, CTM, and MGCTM are generally better than K-means, normalized cut and factorization-based methods. This corroborates our assumption that topic modeling can promote document clustering. Ferrara et al. [7] clustering framework adopts a novel pre-clustering procedure, namely, protomeme detection, aimed at identifying atomic tokens of information inside tweets. Due to its efficiency, this strategy should be particularly well suited to work in streaming scenarios. Goldszmidt et al. [8] uses probabilistic-based framework for information retrieval and proved that the results are superior to other information retrieval methods. Aggarwal [9] discussed the various clustering algorithms used for document clustering. Benghabrit [10] proposed algorithm selects iteratively relevant features and performs clustering until convergence and results show that the good performance. Zhao et al. [11] mainly discussed the challenge of using the EM algorithm and the Map-Reduce framework to train PLSA model using map-reduce technique. Tripathy et al. [12], Tang et al. [13], and Sadeghian et al. [14] discussed the algorithm for various real-time data; Kanimozhi and Venkatesan [15–18] discussed various text clustering using map-reduce programming. Generally, in English like languages to discover the meaning of documents is a problem since each word represents multiple meanings so techniques like latent semantic analysis called latent semantic indexing method face difficulties in handling multiple meanings for one word because it is meant for only the direct mapping like one word to one meaning. Even though this method reduces the dimension, it is not probabilistic and fails to handle noise. But probabilistic latent semantic analysis method is used for analyzing documents to discover the underlying meaning or concepts of those documents. Hence, we use probabilistic-based latent semantic analysis model is used to solve this type of problem and to perform better topic modeling since it handles polysemy. It allows for the estimation under circumstances of uncertainty. And also, this method provides a uniform system for integrating and reasoning over heterogeneous information. But probabilistic latent semantic analysis method is also computationally expensive if we model the documents directly to millions of documents. The time required by increases significantly with increase in the number of documents. Hence, we combine cosine similarity with probabilistic latent semantic analysis method to solve the problem.

3 Problem Statement and Proposed Solution

The main problems faced by text processing are, namely, the huge volume of documents and the large size of text documents features. To enhance the efficiency of clustering and to reduce the time and space complexity a topic modeling is proposed by cosine similarity-based probabilistic latent semantic analysis and implemented using map-reduce techniques. The proposed system is divided into

two phases: In the first phase the input as search, query terms are compared with the document corpus and cosine similarity is calculated. After the calculation is performed, the similar documents are obtained. In the second phase, the similar documents are taken as an input, and PLSA technique is applied to the similar documents and the topics are generated. Then, the documents are ranked accordingly.
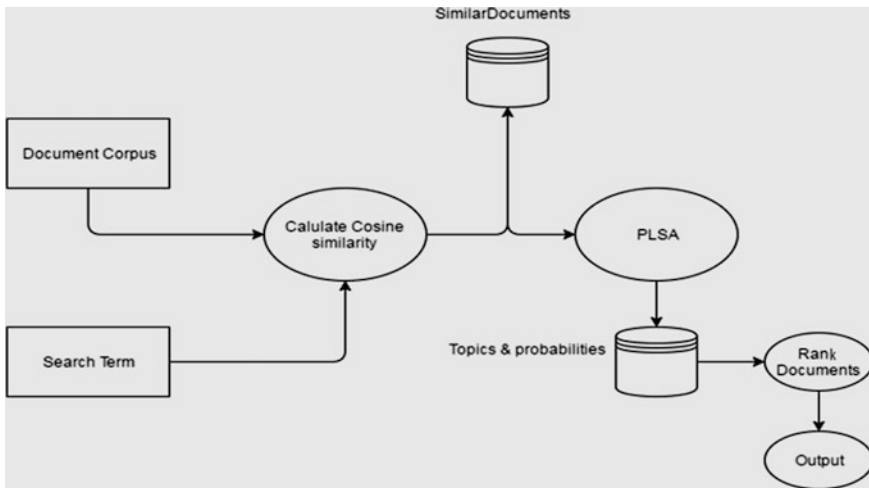
3.1 Map-Reduce Framework

To manage the large text documents and to solve the scalability problem we use the map-reduce programming. The proposed method works in parallel manner at different stages with less computation speed and reduces the huge dimensionality hence algorithm works in different levels and clusters many documents in parallel manner, which leads to reduced computation speed and huge dimensionality reduction.

3.2 Text Preprocessing

Data preprocessing is an important method in data mining which converts involves transforming unprocessed data into an understandable format. Since the real-time data is often incomplete, inconsistent, and noisy, we need to solve those issues by data preprocessing and convert to processed format to yield efficient and accurate result. The preprocessing contains the following functions. Here we perform two main preprocessing steps for preparing dataset like tokenization and stop word removal method in our implementation. Using the map-reduce programming our algorithm does the preprocessing step for tokenization and stop words removal for our dataset collection. The first step is called tokenization which divides the every line of text document into individual words called tokens. Next is stop words removal where the words which do not contain important significant information or occur so often that in text collection are removed. In this paper, stop word list are stored in a separate file and it is verified with all the text documents using the stop word removal method, and thus, those unnecessary stop words are eliminated from those documents.

3.3 Proposed Methodology

Topic modeling using probabilistic approach is used to discover the latent semantics from large text collection. Whenever this probabilistic-based approach is used for topic modeling it is computationally expensive, so whenever number of document increases, time required to process those documents are also increases. Therefore in order to reduce the computational time, we proposed cosine similarity-based probabilistic latent semantic analysis approach to compute the similarity between those documents and the input search terms and then once the similar documents are found we applied probabilistic latent semantic analysis on those limited number of documents that need to be processed and also ranked the documents according to their similarity with the incoming search term.

3.3.1 Cosine Similarity PLSA-Based Text Analysis

In Stage 1: The inputs as search query terms are compared with the document corpus and cosine similarity is calculated. After the calculation is performed the similar documents are obtained. In Stage 2, the similar documents are taken as an input, and PLSA technique is applied to the similar documents and the topics are generated. Then, the documents are ranked accordingly shown in Fig. 1.

Fig. 1 Cosine similarity PLSA-based text analysis

Algorithm for Topic Modeling Using Cosine PLSA Using Map-Reduce Framework:

Input: Document parser and search query terms from local system to hadoop file system.
Output: Topics
Method:
Begin of the Algorithm
Step 1: For each document D in document set do
    1.1 Method to read files and tokenize the input files into individual terms and store in termsDocsArray which hold all terms of each document in an array.
        End for
Step 2: Method to create term vector according to its term frequency inverse document frequency score.
      For all string terms do
    2.1 Calculate term frequency from doctermsArray and return term frequency.
    2.2 Calculate inverse document frequency from termsDocsArray and terms and return inverse document frequency
    2.3 Calculate term frequency inverse document frequency based term frequency and inverse document frequency and return term frequency inverse document frequency.
        End for
Step 3: Method to calculate cosine similarity between all the documents using term frequency inverse document vectors and output the similar documents.
Step 4: Method to remove the stop words from those documents
Step 5: Train the list of documents using PLSA with maximum iteration
    5.1 Train the list of documents using PLSA with maximum iteration by EM algorithm to estimate the parameters
    5.2 Use E-step to calculate the posterior probability
        $(z|d,w,\&)$,$\&$is model params$(p(z|d),p(w|z))$,
        $p(z|d,w,\&)=p(z|d)*p(w|z)/sum(p(z'|d)*p(w|z'))$, z' represent all possible topic.
    5.3 Find document Topic Probabilities and document Term Topic Probabilities
    5.4 Use M-step to update
        $p(w|z),p(w|z)=sum(n(d',w)*p(z|d',w,\&))/sum(sum(n(d',w')*p(z|d',w',\&)))$
        d' represent all documents , w' represent all vocabulary
    5.5 Find topic Term Probabilities
    5.6 Update $p(z|d),p(z|d)=sum(n(d,w')*p(z|d,w'\&))/sum(sum(n(d,w')*p(z'|d,w',\&)))$
      where w' represent all vocabularies and z' represent all topics
    5.7 Update document Topic Probabilities
    5.8 Return topic Term Probabilities and all words
Step 6: Return Topic Word Probabilities
End of the algorithm.

In probabilistic latent semantic analysis, we aim to discover topics which are not apparent, by considering different documents which may have words in common. Our proposed algorithm works well when documents in the corpus are large. Using the map-reduce paradigm we created a method which initially reads all the input

documents and stored in an array format. And the first step of preprocessing is tokenization is performed from all the input documents in which it produces the output as individual terms and stored in an array variable called term document array where it holds all terms of every document in an array. Then, the term frequency is calculated from the array. Term frequency is the frequency of every word occurred in each document is calculated. And inverse document frequency is the measure of the rareness of a term and inverse document frequency helps to identify what it is that makes a given document special. On a large document the frequency of the terms will be much higher than the smaller ones, hence normalized the document based on its size. Then, inverse document frequency is calculated from term document array and terms by combining the above two like term frequency and inverse document frequency term frequency-inverse document frequency. Between all documents the cosine similarity between the query input and all documents are calculated using the term frequency-inverse document vectors. Once the similar documents are calculated the stop words are removed by stop word removal method for all those documents. Hence, we get the output as similar documents without stop word removal. Then, the probabilistic latent semantic analysis algorithm is applied to those minimum documents instead of whole input documents hence reduces the computation time, so time required by the algorithm is very less. After calculating probability for all documents regarding all the probable topics, highest probability words or terms occurred in the documents are considered and the highest probable documents regarding every topic are ranked in order of probability values.

4 Experiments and Results Analysis

In this section, we have implemented the algorithms of PLSA and cosine similarity-based PLSA separately for twitter dataset and the performance evaluation are performed in map-reduce parallel programming on a Hadoop environment.

From our experiment, we tested the scalability of proposed method by increasing the number of documents with respect to the number of nodes, and we found that in Fig. 2, the graph linearly increases to the number of text documents with maximum number of nodes. Computational time decreases with respect to number of nodes is shown in Fig. 3.

4.1 *Comparison Between PLSA and Cosine PLSA*

We have performed a comparative analysis on the performance of traditional probabilistic latent semantic analysis with the combined system of cosine similarity with probabilistic latent semantic analysis on the same Twitter Dataset. The time required to run the twitter dataset for probabilistic latent semantic analysis is 11 min
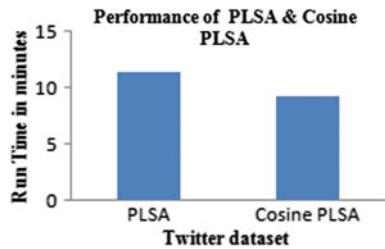
Fig. 2 Scalability with
respect to the number of
nodes
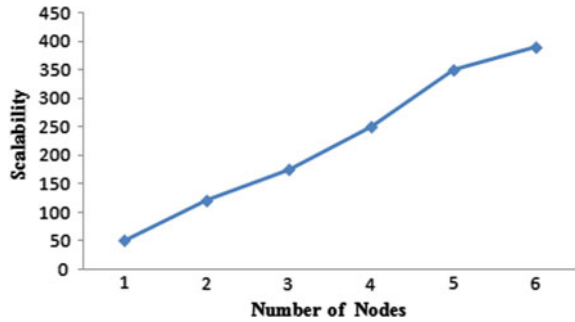
Fig. 3 Computation time
decreases with respect to
number of nodes
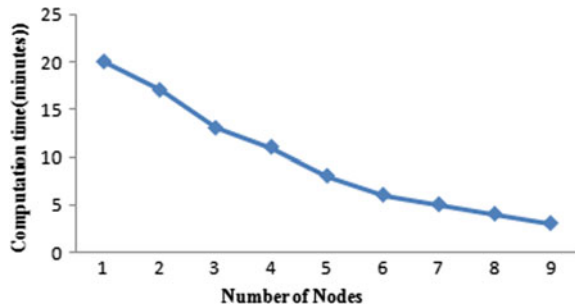
Fig. 4 Run time of PLSA
and cosine PLSA

46 s and the time taken to run the same twitter dataset using our proposed cosine
similarity with probabilistic latent semantic analysis is 9 min 29 s. After performing
the comparative analysis between these two methods on same datasets, we conclude
that our proposed cosine similarity probabilistic latent semantic analysis performs
better by taking less time for execution of the same twitter dataset than PLSA and
the execution time between these two approaches is shown in Fig. 4.

5   Conclusion

In this paper, we have mainly implemented and analyzed the performance of PLSA
and cosine similarity-based PLSA on the same document corpus in a Hadoop
framework. We have seen a considerable difference regarding the computation time

of these two approaches that only PLSA algorithm for topic modeling takes more computation time and as per our expectations, the approach where the combination of the cosine similarity-based PLSA takes less computation time than the single PLSA approach. And we thus reduce the huge dimensionality using our approach and results in generating topics, which leads to effective clustering in large text document corpus. Future work shall consider the possibility of search results by ranking the documents.

References

1. Wei, C.-P., Yang, C.C., Lin, C.-M.: A latent semantic indexing—approach to multilingual document clustering. Sci. Direct. Decis. Support Syst. **45**, 606–620 (2008)
2. Clifton, C., Cooley, R., Rennie, J.: TopCat: data mining for topic identification in text corpus. IEEE Trans. Knowl. Data Eng. **16**(8) (2003)
3. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
4. Nagwani, N.K.: Summarizing large text collection using topic modeling and clustering based on Map Reduce framework. J. Big Data, 2–6 (2015)
5. Kontostathis, A., Pottenger, W.M.: A framework for understanding latent semantic indexing (LSI) performance. Inf. Process. Manag., 56–73 (2006)
6. Xie, P., Xing, E.P.: Integrating document clustering and topic modeling. In: Proceedings of the Twenty Ninth Conference on Uncertainty in Artificial Intelligence (2013). UAI-P-2013-PG-694-703
7. Ferrara, E., JafariAsbagh, M., Varol1, O., Qazvinian, V., Menczer, F., Flammini, A.: Clustering memes in social media (2013)
8. Goldszmidt, M., Sahami, M.: A probabilistic approach to full text document clustering. Technical Report, Stanford Info Lab. Digital Libraries, 1998
9. Aggarwal, C.C., Zhai, C.X.: A survey of text clustering algorithms. In: Mining Text Data, pp 77–128. Springer (2012). https://doi.org/10.1007/978-1-4614-3223-4_4
10. Benghabrit, A., Ouhbi, B., Behja, H., Frikh, B.: Text clustering using statistical and semantic data. In: Computer and Information Technology (WCCIT), 22–24 June (2013). https://doi.org/10.1109/wccit.2013.6618782.ieee
11. Zhao, Y., Chen, Y., Liang, Z., Yuan, S., Li, Y.: Big Data processing with probabilistic latent semantic analysis on Map Reduce. In: International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery. IEEE (2014). https://doi.org/10.1109/cyberc.2014.37. 978-1-4799-6236-5/14
12. Tripathy, R.M., Sharma, S., Joshi, S., Mehta, S., Bagchi, A.: Theme based clustering of tweets. In: Proceedings of the 1st IKDD Conference on Data Sciences, March 21–23, pp 1–5 (2014). https://doi.org/10.1145/2567688.2567694. ISBN: 978-1-4503-2475-5
13. Tang, Z., Jiang, L., Yang, L., Li, K., Li, K.: CRFs based parallel bio medical named entity recognition algorithm employing Map Reduce framework. Cluster Comput. **18**, 493–505 (2015). https://doi.org/10.1007/S10586-015-0426-z
14. Sadeghian, A.H., Nezamabadi-pour, H.: Document cluster ing using gravitational ensemble clustering. In: International Symposium on Artificial Intelligence and signal processing. IEEE (2015). 978-1-4799-8818-1/15/$31.00
15. Kanimozhi, K.V., Venkatesan, M.: Survey on text clustering techniques. Adv. Res. Electr. Electron. Eng. **2**(12), 55–58 (2015)

16. Kanimozhi, K.V., Venkatesan, M.: Unstructured data analysis—a survey. Int. J. Adv. Res. Comput. Commun. Engi. **4**(3) (2015). https://doi.org/10.17148/ijarc015.4354. ISBN: 2278-1021.
17. Kanimozhi, K.V., Venkatesan, M.: Big text datasets clustering based on frequent item sets—a Survey. Int. J. Innov. Res. Sci. Engi. **2**(5) (2016)
18. Kanimozhi, K.V., Venkatesan, M.: A novel Map-Reduce based augmented clustering algorithm for big text datasets. In: Data Engineering and Intelligent Computing: Proceedings of IC3T 2016, 2017 May 31, vol. 542, p. 427. Springer (2017)

WAPiS: WhatsApp Pattern Identification Algorithm Indicating Social Connection

Sawan Kalra, Rahul Johari, Sonika Dahiya and Poonam Yadav

Abstract Today social networking has emerged as a prominent area of networking. It provides a platform where people can exchange information, sentiments, and expressions. Taking this as a core, the current research work was initiated wherein WhatsApp chats of different individuals were taken and effort was made to identify social inclination/pattern between them. The newly proposed WAPiS Algorithm: WhatsApp Pattern Identification Algorithm indicating Social Connection has been formulated to achieve this. The algorithm has been designed and developed using C#–MongoDB combination and the results indicate the existence or identification of the social contact between the WhatsApp text of the sampled individuals.

Keywords Big data · Data mining · MongoDB · WhatsApp

1 Introduction

Big data is the term for a collection of datasets that are large and complex and therefore it becomes difficult to process using in hand database management tools or traditional data processing applications. Nowadays, most of the tasks have been

S. Kalra (✉) · R. Johari · P. Yadav
USICT, Guru Gobind Singh Indraprastha University, New Delhi, India
e-mail: sa1kalra@gmail.com

R. Johari
e-mail: rahul.johari.in@ieee.org

P. Yadav
e-mail: yadavpoonam96@gmail.com

S. Dahiya
Delhi Technological University, New Delhi, India
e-mail: sonika.dahiya@dtu.ac.in

shifted to Internet-based solutions rather being manual. Whether it is bill payments, food ordering, tax payment, online shopping, ticket booking, or some kind of entertainment or social website. The data generated by the use of these services provided on the Internet is increasing exponentially. Hence, many times this data becomes difficult to manage. One such example of big data is the WhatsApp chats, used in current research work to showcase how such large amounts of data can be handled and information can be extracted.

The large number of social networks tools and applications, forums, online marketplaces, and additional online resources publish content, which can be used for understanding the mindset of the consumers and the common public on social events like political movements, company policies, marketing strategies, product choices. Such data can be analyzed using an amalgamation of data mining, deep learning, web mining, and text mining techniques in a variety of real-life applications. The current research work aims to provide a case study that searches through keywords in the WhatsApp chats and gives the information about the social interests of different people. For instance, a teenager uses smileys or emoticons a lot more than an adult. Also, the social interests can be of many types: politics, sports, news, movies, festivals, etc. Taking chats of 4 different people, which may belong to same social groups or may be random at all, which means having no kind of link with each other, or may have some kind of social inclination toward each other knowingly or unknowingly. With the time period of a week, the messages exchanged between two people (referring to one chat) may be focused on only one topic or may contain many topics, for instance, the chat may be focused on the admission procedure in any school\college\university or the chat may be consisting of various topics like recently released movie gossip, planning of a tour, any natural calamity, political talk, any ongoing tournament like Indian Premier League or Champions Trophy or Wimbledon Cup or may be the Olympics, doctor's appointment, which differs from person to person. This project tries to find connections between these chats and devise a social inclination between the two persons. To manage this huge data or big data various tools are available, two major tools among them are MongoDB and MyHadoop.

MongoDB has been used as it is a cross-platform, open-source, document-driven database which provides good performance, 24×7 availability, and thorough scalability of data. It has no concept of table, schemas, SQL, or rows. It does not possess transactions, ACID compliance, joins, foreign keys, or any of the other property that serves as a reason to cause a lot of problems. To the best of our knowledge and vision, this is a first of its kind of work in this direction.

This paper is organized in the following manner. In Sect. 2, we highlight previously done related work on social media data mining. In Sect. 3, we explain our

motivation and in Sect. 4, we outline our methodology adopted, before we proceed with a presentation of our algorithm formulation in Sect. 5. We present simulation environment in Sect. 6. The simulation of the WAPiS is presented in Sect. 7. Finally, conclusion and future work are in Sect. 8 followed by references.

2 Related Work

In [1], the author(s) have presented the detailed account of the benefits and drawback of big data tools and techniques. In [2], the author(s) have designed a workflow to integrate both qualitative analysis and data mining techniques. They focused on students' Twitter posts to explore the insights in their educational experiences. In [3], the author(s) states that it is difficult for traditional databases and architecture to modify, grill, and then structure the data that the Internet user generating daily on chatting, liking, poking, tweeting on social media which has become a vast source of unstructured data. In [4], the author(s) proposed a framework in order to improve the throughput of the system. In this paper, two issues related to fast access to real-time data are presented; the first issue is indexing and second was handling of real-time content spikes. Based on these concepts, RTSEs (Real-Time Search Engines) were analyzed. In [5], the author(s) explored an unsupervised topic modeling approach. In [6], the author(s) have presented a review paper of various fuzzy clustering algorithms. It includes description of FCM, IFCM, T2FCM, DOFCM, KIFCM, FCM-sigma, KFCM, KT2FCM, etc. In this paper, the author(s) have compared the performance of various algorithms on noise-free and noisy data, also linearly and nonlinearly separable data.

3 Motivation

WhatsApp is a cross-platform instant messaging application and whether one uses Android or BlackBerry or iOS. Hence, the motivation to mine the conversations and figure out social interests as the better part of everyone's life is now spent on communicating via these messages.

The social conduct and the topics of interest and the gravity or the seriousness of the topics provided viable information which helped to establish a finer understanding of the customer base, civilian base, or just the humans in general. This could further mine more applications in the corporate world once chats are analyzed for sentiment or social interests that bind a certain category of people together in

social groups, which could further lead to more applications. This unexplored area motivated us to work in this direction and attain the plausible results.

4 Methodology Adopted

WhatsApp chats of 4 different persons over a week were taken. The chats were received in text file format that is for 4 WhatsApp chats were stored in 4 different text files. Now, these chats were imported into Microsoft Excel, and then separate every single chat into different columns that were date, time, meridian, and message. Then, we saved this file in CSV (Comma Separated Values delimited) file format.

C#–MongoDB Connectivity program was written to connect to the database.

Now after saving the WhatsApp Chats in CSV file format, each file was imported into the MongoDB database using the following command:

```
mongoimport -d database -c collection --type csv --file
filename .csv --headerline
```

5 Algorithm Formulated

Text analytics (text mining) indicate the extraction of the information from textual data. Information extraction (IE) techniques extract/fetch structured data from unstructured text. WAPiS: WhatsApp Pattern indicating Social Connection—In this a framework/strategy is presented and results are showcased that show how popular social media tools like WhatsApp chat can relate two persons by their common interest on the basis of words they used during the chat. These are the various steps that are involved in the process of data mining from data collection to decision making:

5.1 Working Steps of WAPiS: WhatsApp Pattern Indicating Social Connection

1. To collect sample WhatsApp chat of 4 persons for a defined date and time, let us say of 1 week.
2. Take first WhatsApp chat:

   2.1. Start scanning from the first word to the EOF.
   2.2. To ignore/eliminate sentence connecting words like: I, We, are, and, for, if, etc.

2.3. To ignore all the words whose word length (W_L) $\leq$ 6.

2.4. Push the selected words into first stack.

3. Repeat step 2 for remaining three WhatsApp chats numbered 2–4.

4. At the end of step 3:

Person A	Person B	Person C	Person D
Blessings	• Captain	• Business	• Amazing
Celebration	• Cricket	• Celebration	• Bollywood
• Clothes	• Family	• Exercise	• Contestant
• Crackers	• Festival	• Government	• Creative
• Deepawali	• Friends	• Holidays	• Entertainment
• Festival	• Gurpurab	• Morning	• Latest
• Greetings	• Pakisthan	• Purchasing	• Lifestyle
• Holiday	• Practical	• Technology	• Trailer
• Relatives	• Preparation	• Tension	• Training
• Shopping	• Tournament	• Thanksgiving	• Wishing

Now we have 4 stacks, one for each person that contain at least 10 unique words whose length is greater than six.

5. From Stack 1, pick up the first word and search for its existence in Stack 2, 3, and 4.

6. If suppose the word occurrence is not found in Stack 2 and 3 then in:

6.1. Pick up the second word from Stack 1 and search for its occurrence in Stack 4.

6.2. If 6.1 is successful then:

Choices/interest of WhatsApp chat of first person matches with fourth person.

Else

Goto step 5, pick up second unique word from stack 1 and search for its existence in Stack 2, 3, and 4.

7. To conclude, the objective is to determine social pattern amongst the selected sample dataset: four WhatsApp chats.

5.2   *Mathematical Modeling of WAPiS Algorithm*

Notations:

WAP$_i$: WhatsApp Chat of ith person for a defined date and time, let's say of 1 week.
n: number of WhatsApp Chat, one per person.
Stack$_i$: is a stack belonging to ith WhatsApp post, used for storing keywords and frequency of occurrence of each keyword.

Procedure :

Step 1. Set all Stack$_i$ Queue, say Q to ϕ.
Step 2. Collect Sample WhatsApp Chat of n-persons for a defined date and time, say of a week.
Step 3. For i=1 to n
   3.1. Start scanning from first letter of WAP$_i$.
      3.1.1. Find each word in the scan and check length of word, say w$_d$ and l$_w$ respectively.
        3.1.2. IF (l$_w$>=6 && w$_d$ ∈ Stack$_i$)

           increment count of w$_d$ in Stack$_i$

             ELSE IF (l$_w$>=6 && w$_d$ ∉ Stack$_i$)

          Add w$_d$ to Stack$_i$
          Set count of w$_d$ to 0
            END IF
          END IF
Step 4. For j=2 to n
   Queue←EnQueue(all C(n,j) combinations)

Step 5. Combination, C←DQueue()
   5.1. Find common words among stacks in Combination C.
   5.2. Compute count of occurrence of common words.
   5.3. IF (this count of common words is greater than a threshold value, say α)
      then
      we can say interest of WhatsApp Chat of considered stacks matches, and thus, the interest of respective people matches.
       ElSE
      drop this combination C and all other combinations that are a superset of C from Queue Q.
      END IF
   5.4. Repeat step 5 until Queue ≠ ϕ

6   Simulation Environment

6.1   Visual Studio

Visual Studio is Microsoft's IDE tool for creating, documenting, designing, running, and debugging programs typed in C#. Visual Studio also provides editing tools for manipulating numerous types of files. Visual Studio is an effective and efficient tool for creating critical and noncritical applications.

6.2   MongoDB

It is a cross-platform, open-source, document-based database which provides, high throughput, high availability, and easy scalability of data. It employs the concept of collection and document. It has no concept of table, schemas, SQL, or rows. It does not have transactions, ACID compliance, foreign keys, joins, or any of the other database features that usually cause a lot of problems.

7   Simulation Performed

The WAPiS Algorithm was designed and developed by writing the code in C# (C Sharp Programming Language) with database connectivity done with MongoDB. Figures 1, 2, and 3 depict the same.
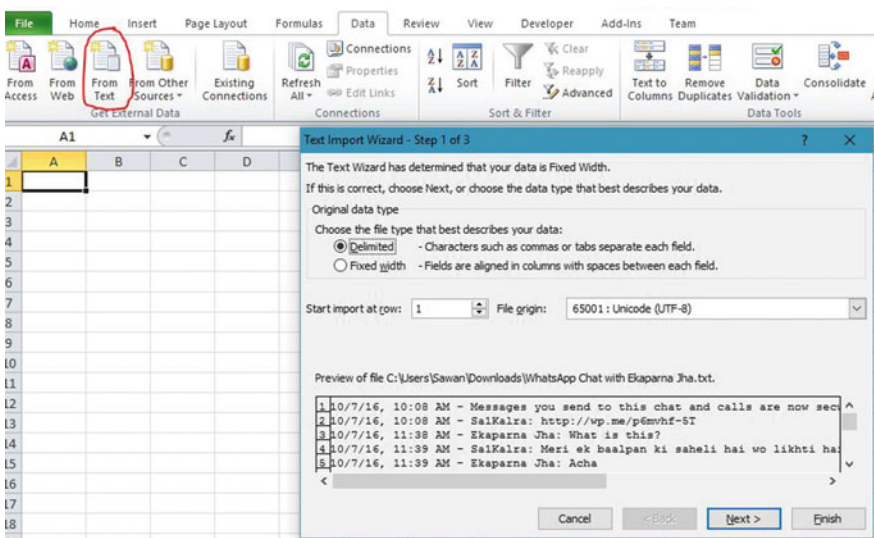
Fig. 1 Snapshot indicating the identification of the keyword in chat
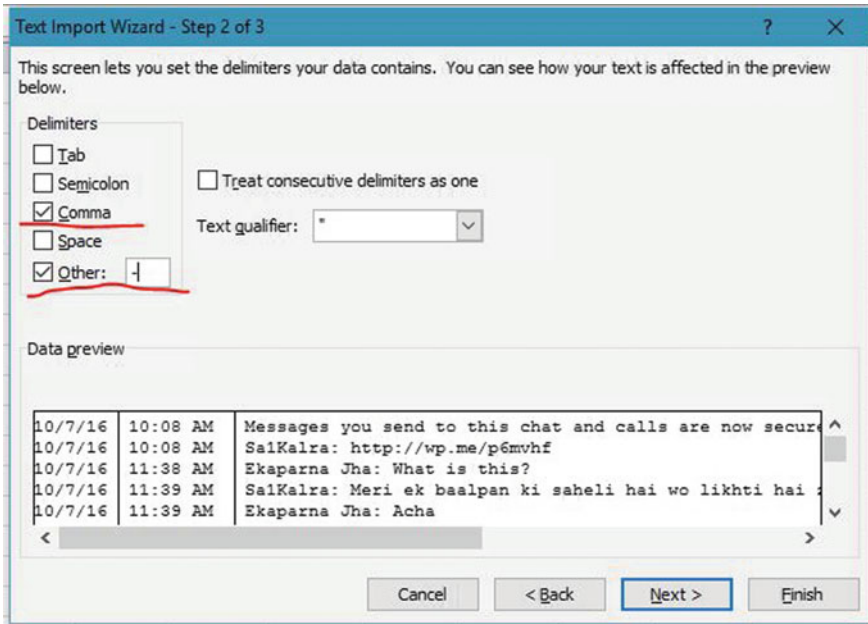
Fig. 2 Snapshot indicating the searching of the keyword in other person's chat
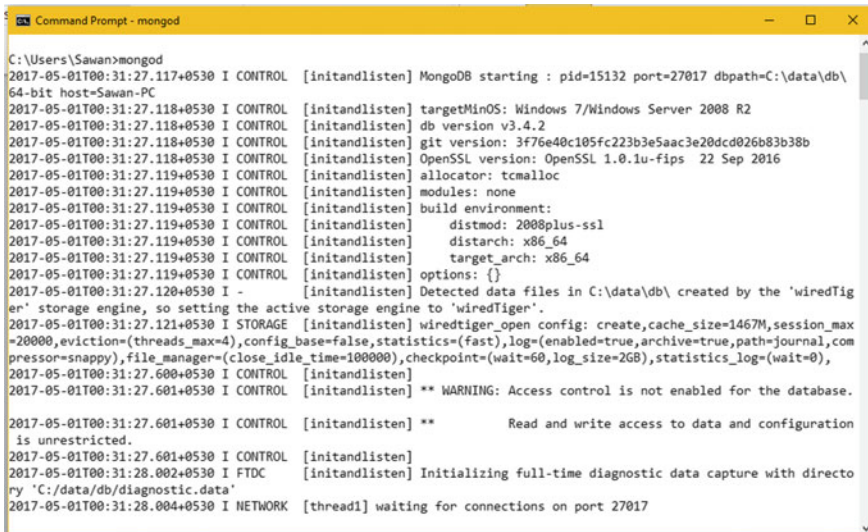
Fig. 3 WAPiS algorithm database connectivity to MongoDB

8   Conclusion and Future Work

WhatsApp chats were imported in MongoDB, and for the future work it is proposed to decrypt the newly encrypted chats directly rather than getting the WhatsApp chats in the text files, import the chats obtained, and then apply the data mining techniques with a more efficient algorithm which will reduce space and time complexities.

Next step that we would be taking is to take the chats of more people based on different criteria like people belonging different geographical regions, people with different interests, people among same institution or company or job profile, etc.

This current research work can further be used to Twitter's tweets dataset and over Facebook's posts dataset and it can be enhanced to apply opinion mining techniques and sentiment analysis.

References

1. Bajaj, S., Johari, R.: Big data: a boon or bane-the big question. In: 2016 Second International Conference on Computational Intelligence and Communication Technology (CICT), pp. 106–110. IEEE (2016)
2. Chen, X., Vorvoreanu, M., Madhavan, K.: Mining social media data for understanding students' learning experiences. IEEE Trans. Learn. Technol. **7**(3), 246–259 (2014)
3. Gole, S., Tidke, B.: A survey of big data in social media using data mining techniques. In: 2015 International Conference on Advanced Computing and Communication Systems. IEEE (2015)
4. Quazilbash, N.Z., Qadri, S.M.H., Khoja, S.: Improved user RTSE experience on the web through fast retrieval of social media content. In: 2012 15th International Multitopic Conference (INMIC). IEEE (2012)
5. Rohani, V.A., Shayaa, S., Babanejaddehaki, G.: Topic modeling for social media content: a practical approach. In: 2016 3rd International Conference on Computer and Information Sciences (ICCOINS). IEEE (2016)
6. Gosain, A., Dahiya, S.: Performance analysis of various fuzzy clustering algorithms: a review. Proc. Comput. Sci. **79**, 100–111 (2016)

A New DNA Cryptography Based Algorithm Involving the Fusion of Symmetric-Key Techniques

Animesh Hazra, Soumya Ghosh and Sampad Jash

Abstract Today, with the increase in technology and computational power everything is getting digitalized. More and more private data is being loaded into our computers. These data are sent to various websites for suitable applications. Now, the problem arises while transferring these data. Sending these data through insecure channels like the Internet without using proper encryption methods can lead to data loss or data hacking. In order to protect this data, various encryption methods are being implemented. One such encryption methodology used in this arena is DNA cryptography. This paper gives a brief overview of DNA cryptology along with a new algorithm based on the fusion of symmetric-key cryptography, DNA nucleotides, and XOR operation is proposed. This algorithm is very much efficient and one main striking feature of the algorithm is the security which can be set as per sender requirements.

Keywords Data security · Decryption · DNA cryptography
Encryption · Substitution

A. Hazra (✉)
Department of Computer Science & Engineering,
Jalpaiguri Government Engineering College, Jalpaiguri 735102, West Bengal, India
e-mail: hazraanimesh53@gmail.com

S. Ghosh · S. Jash
Department of Information Technology, Jalpaiguri Government Engineering College,
Jalpaiguri 735102, West Bengal, India
e-mail: baban.jgec14@gmail.com

S. Jash
e-mail: sampadjash2013@gmail.com

1   Introduction

Cryptography [1] can be described as the art of accomplishing security with the help of encoding messages in order to make them non-readable. The technique of decoding the non-readable messages into readable one is known as cryptanalysis and it can be said that cryptology is defined as the combination of cryptography and cryptanalysis. Before the age of computers, the art of cryptography was performed by manual techniques. Some of the techniques involved are substitution technique and transposition technique. A few examples of substitution cipher are Vigenere cipher, modified version of Caesar cipher, monoalphabetic cipher, polygram substitution cipher. Some transposition techniques include rail-fence technique, simple columnar transposition technique, Vernam cipher technique, etc. The introduction of computer made it possible to use more complex cryptographic algorithms. Technical terms like encryption and decryption came into existence. Encryption is defined as the process involved in the encoding of plain text to cipher text and decryption is the reverse process of encryption.

Encryption and decryption have two major aspects. One is the algorithm that is involved in the encryption process and another one is the key used in the algorithm. Key is the most important attribute on which the security of the encryption technique depends. It can be classified into two categories. When the same key is used for both encryption and decryption then it is known as a symmetric key, whereas in asymmetric key different keys are used for encryption and decryption algorithm. There are various methods to implement these techniques like data encryption standard (DES), international data encryption algorithm (IDEA), RC4, RC5, Blowfish, advanced encryption standard (AES), RSA, ElGamal cryptography, etc. There are some challenges involved in traditional cryptography. One such problem is the sharing of the key because it may fall prey to attacks like eavesdropping or man in the middle. It has been found out that the ultimate power of a cipher depends on three important factors, i.e., the size of the key, the infrastructure on which the key is running, and last one is the algorithm on which it is designed. Today the most secure key length is 2028 bits. A famous symmetric key algorithm known as DES, which uses a 5-bit key size is no more considered as safe. The increase in computational power has compromised the safety of the traditional cryptographic techniques. This problem is solved with the use of DNA cryptography.

2   DNA Cryptography

DNA computing can be defined as a new technique for securing information with the help of biological structures. Adleman [2] can be considered as the pioneer of the DNA computing. He used this in 1994 for solving complex algorithmic problems. Now, it is discovered that DNA can be used to store and transmit data.

There are several advantages of DNA computing out of which a few of them are discussed as follows:

a. Minimum storage requirement: In a compact volume large amount of data can be stored. Calculations suggest that 1 g of DNA contains 10^8 terabytes of data.
b. Speed: DNA computing techniques are almost 100 times faster than the modern-day fastest computer.
c. Minimum power requirements: DNA computing needs very less or no power at all compared to modern-day computers.

The process of hiding data in terms of DNA sequence is known as DNA cryptography. At present, the work in the field of DNA cryptography is focused on the use of DNA sequences to encrypt binary data in some form or other. In the near future, DNA computing will become as one of the leading techniques for securing information. Some of the DNA cryptographic techniques based on which the proposed algorithm is designed are discussed as follows.

2.1 DNA Digital Coding

It is the technique of mapping the 4 different bases of DNA that are A, C, T, and G with 0 and 1. Plain text messages can be easily encoded using this scheme. There are 24 such patterns possible but only 8 unique combinations are considered which fit the complementary rule. This will be clear with the following example. Suppose someone wants to send the number "97" using DNA encoding. He or she can convert "97" to binary form by breaking 9 to 4-bit binary form 1001 and 7 is converted into 0111. Then both the binary forms of 9 and 7 are joined together. The resulting binary number will be 10010111.

Starting from the left-most bit, two consecutive binary digits are taken and converted to corresponding DNA nucleotide bases by following the scheme described in Table 1. In this way finally, the number "97" will be encoded as "CTTG". Now the encrypted message "CTTG" will be sent through a channel to the receiver. The receiver then decodes it and extracts the original message.

Table 1 Conversion scheme for binary form to DNA nucleotide

Binary form	DNA nucleotide
00	A
01	T
10	C
11	G

2.2 Hao's Permutation and Fractal Sequence

Hao et al. [3] designed and proposed a DNA fractal sequence representation approach. A complete genome of length N is given which may be circular or a linear DNA sequence composed of N letters from the alphabets A, C, G, and T. The authors designed a mapping technique that maps the four letters A, C, G, and T to the base-4 number system. α: {G, C, A, T} → {0, 1, 2, 3}. The coordinates for the counter of the first K-string are calculated using (1) and (2).

$$x = \sum_{i=1}^{k} 2^{k-i} [\alpha(s_i) \gg 1] \qquad (1)$$

$$y = \sum_{i=1}^{k} 2^{k-i} [\alpha(s_i) \& E] \qquad (2)$$

In (1) and (2) E means a binary number 1, "≫1" indicates the right shift by one bit, symbol "&" is bitwise AND operator and s stands for the chosen DNA sequence. Here, α is a mapping function which converts the selected sequence into its corresponding base-4 number system. This permutation approach is robust and provides high security, particularly in image encryption. Using this approach authors have proposed a new key generation process for image encryption [4].

2.3 Polymerase Chain Reaction (PCR)

PCR is a technique generally applied in the amplification of DNA with the help of molecular biology. The steps followed in PCR operation are denaturation, which is followed by primer annealing and primer extension. In denaturation, a two single-stranded DNA is formed from a double-stranded molecule. This is done by heating the sample at a very high temperature of about 94–96 °C. Then primer annealing is done at about 50–65 °C, where the primers which are designed to amplify the DNA regions are attached to the complementary sequences. Finally, primer extension is performed. In this step, the temperature is again raised to about 72 °C. Here, nucleotides are added to the strand of a short primer on the base of original DNA strand using polymerase enzyme.

3 Literature Survey

Several research papers are there which explore DNA-based cryptographic techniques. In this section, the summaries of some of them are presented in a nutshell.

Fu et al. [5] implemented a new DNA cryptographic algorithm. The method is based on DNA fragment assembly. The authors have implemented the features of DNA digital coding, DNA molecular key, and some software techniques in their algorithm. This technique follows the concept of symmetric key cryptography. The encryption mechanism is done here using DNA digital coding.

Verma et al. [6] have proposed an index-based DNA encryption algorithm. The plain text messages are encrypted into DNA sequences with the help of index of strings and block cipher. The sequences are given an index number which acts as the ciphertext. It is very hard for an attacker to guess the real message.

In [7], authors have implemented a method which is based on the asymmetric key. The algorithm uses PCR amplification along with DNA digital coding and digital synthesis for encryption of the plain text. PCR amplification is added here to provide security and safeguard during the communication phase. This encryption scheme has high confidential strength.

Yupeng et al. [8] designed a method on index-based encryption algorithm which follows symmetric-key cryptography. The proposed technique is based on the index of string and block cipher. The plaintext is encrypted twice with the help of the key. XOR operation is used after the key generation. The chaos mapping added in this algorithm increases the mathematical security. Key generation ensures that the huge key space is fully optimized and can prevent extensive attacks.

In [9], the authors have proposed a DNA-based encryption method using big data. When a large amount of data is to be stored using big data then the use of this encryption technique is suggested. The designed method uses PHP language and a DNA encoding table for the encryption process. So far big data analysis is facing many challenges. Therefore, this encryption process can be used to solve some of the problems in big data analysis.

Reza et al. [10] have designed a new method to encrypt information with DNA-based cryptography. It is performed on five main stages. Concepts of data preprocessing and data post-processing are used in this technique. This algorithm is very much secured since it has a double layer of security and is based on Vigenere cipher.

Tanaka et al. [11] proposed a DNA cryptographic algorithm which is based on the one-way public key. The two keys are generated using ODN mixture. One of the public keys helps in encoding the plain text in a DNA sequence. In order to decode the DNA sequence, PCR amplification with the help of a secret sequence is implemented. This is an asymmetric method which has a high level of security but very costly to implement.

In networking and data communication the main concern is security. Mobile networks are becoming vulnerable day by day. In [12] the authors have proposed a method to secure mobile networks through DNA-based cryptography. DNA encoding, mRNA sequence, amino acid, and some mathematical concepts are used for encryption procedure. The proposed technique is secured enough to endure the brute force method as the permutations used in this methodology are very strong. This process can be applied as a hardware solution.

4   Proposed Methodology

This section gives an idea about the working mechanism of the proposed methodology. The algorithm is based on symmetric-key technique which is designed with the fusion of XOR operation and substitution method. It is described below in detail.

4.1   Proposed Algorithm

Step 1: Initially the plaintext P which is to be sent is decided by the sender.

Step 2: P is converted to ASCII code and then to its corresponding 8-bit binary value P_B.

Step 3: Now, a DNA sequence D is randomly chosen which will act as the key. D is converted to binary form using the conversion scheme sited in Table 2 and named as D_B.

Step 4: D_B and P_B are XOR-ed with each other and the resultant binary value is termed as R.

$$R = D_B \oplus P_B \tag{3}$$

If D_B and P_B are not of the same size then extra 0 s are added to the left side of D_B or P_B whichever is smaller in size. The number of zeros added here is equal to the difference between the number of bits present in D_B and P_B.

Step 5: Then DNA sequence D is complemented using Table 3 and named as K. At the same time R is complemented using one's complement operation and the resulting value is titled as $\bar{R}$.

Step 6: Then K is encrypted as per the rules explained in Sect. 4.2 and is termed as K_E.

Step 7: Now both $\bar{R}$ and the encrypted DNA sequence K_E are sent together to the receiver through various insecure channels like Internet.

Step 8: To decrypt the message at the receiver end the reverse process of encryption should be followed.

Table 2 Scheme for converting DNA nucleotide into corresponding binary form

DNA nucleotide	Binary form
A	00
T	01
C	10
G	11

Table 3 Rule for complementing the DNA sequence

Original DNA nucleotide (D)	Complemented DNA nucleotide (K)
A	T
T	A
C	G
G	C

4.2 Encryption Scheme for Complemented DNA Nucleotide (K)

Step 1: The total number of DNA nucleotides present in K is counted and then divided by 3. If the remainder is zero then K remains unchanged, otherwise "X" is added to the right side of K. The number of "X" added is equal to the difference between 3 and the remainder.

Example 1 If K is ACTGACTAG, then there are 9 DNA nucleotides. Now dividing the total number of nucleotides present in K by 3, we get 0 as remainder hence nothing will be done with the DNA sequence.

Example 2 If K is ACTG, then there are 4 DNA nucleotides. Now dividing the total number of nucleotides present in K by 3, 1 will be returned as remainder. Hence, the new DNA sequence is ACTGXX (the number of X added is equal to the difference between 3 and the remainder).

Step 2: Now, the DNA nucleotides of the chosen DNA sequence are to be grouped in such a manner that each group has exactly three nucleotides. The grouping starts from the leftmost side nucleotide. Then using Table 5 decimal values are assigned to each group.

Example 3 If ACTGACGTGCXX is chosen as the DNA sequence then in the following fashion the grouping will be done (Table 4).

Table 4 Example for explaining Step 2

Group name	DNA nucleotides present in the group	Decimal values associated with each group (from Table 5)
Group 1	ACT	10
Group 2	GAC	66
Group 3	GTG	71
Group 4	CXX	59

Table 5 Permutation table for converting DNA nucleotides into decimal values

DNA nucleotide	Decimal value	DNA nucleotide	Decimal value	DNA nucleotide	Decimal value	DNA nucleotide	Decimal value
AAA	1	TAA	22	CAA	43	GAA	64
AAT	2	TAT	23	CAT	44	GAT	65
AAC	3	TAC	24	CAC	45	GAC	66
AAG	4	TAG	25	CAG	46	GAG	67
ATA	5	TTA	26	CTA	47	GTA	68
ATT	6	TTT	27	CTT	48	GTT	69
ATC	7	TTC	28	CTC	49	GTC	70
ATG	8	TTG	29	CTG	50	GTG	71
ACA	9	TCA	30	CCA	51	GCA	72
ACT	10	TCT	31	CCT	52	GCT	73
ACC	11	TCC	32	CCC	53	GCC	74
ACG	12	TCG	33	CCG	54	GCG	75
AGA	13	TGA	34	CGA	55	GGA	76
AGT	14	TGT	35	CGT	56	GGT	77
AGC	15	TGC	36	CGC	57	GGC	78
AGG	16	TGG	37	CGG	58	GGG	79
AXX	17	TXX	38	CXX	59	GXX	80
AAX	18	TAX	39	CAX	60	GAX	81
ATX	19	TTX	40	CTX	61	GTX	82
ACX	20	TCX	41	CCX	62	GCX	83
AGX	21	TGX	42	CGX	63	GGX	84

Step 3: After getting the decimal values for the corresponding groups, they are concatenated sequentially from left to right to get an intermediate string.

However, this string represents a set of decimal numbers.

Example 4 After execution of step 3 using the DNA sequence in *Example 3,* the string "10667159" will be generated by following the way described in Table 6.

Table 6 Example for explaining Step 3

Group name	DNA nucleotides present in each group	Decimal values associated with each group (from Table 5)	String after concatenation
Group 1	ACT	10	10
Group 2	GAC	66	1066
Group 3	GTG	71	106671
Group 4	CXX	59	10667159

Table 7 Substitution table for conversion of decimal value to binary value

Decimal value	0	2	4	6	8	9	7	5	3	1
Binary value	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001

Step 4: Now using the following substitution table the decimal values are substituted into binary values (K_E) (Table 7).

Example 5 Finally, the DNA sequence K will be converted to the encrypted form (K_E) as follows.

$$10667159 \rightarrow 1001 - 0000 - 0011 - 0011 - 0110 - 1001 - 0111 - 0101$$

5 Discussion

5.1 Illustration of the Proposed Methodology

If sender wants to send the plaintext "B" with the following chosen DNA sequence ACTGACTGA using this algorithm then the following steps should be followed.

Step 1: Initially the plaintext "B" is converted to ASCII value followed by its binary value in the following way (Table 8).

Step 2: The chosen DNA sequence ACTGACTGA is converted into binary form using Table 2, which generates the binary value 001001110010011100.

Step 3: Now, binary value of the DNA sequence and binary value of the plaintext are XOR-ed with each other. After executing the above procedure as explained in step 4 of Sect. 4.1, R is generated as 001001110011011110.

Step 4: Next, the DNA sequence is complemented using Table 3, which produces the resulting sequence as TGACTGACT and it is called as K. Simultaneously R is complemented into R̄ which is 110110001100100001.

Step 5: Then K is encrypted using the scheme described in Sect. 4.2 and its value is 100001001110000100100000.

Step 6: Finally, both R̄ and K_E will be sent through various unsecured channels like Internet.

Table 8 Example for converting the plain text to ASCII value and then to binary form

Plain text	ASCII value	Binary form
B	66	01000010

Step 7: At the receiver end, initially the encrypted key K_E is decrypted using reverse steps, then with the help of the key $\bar{R}$ is decrypted and hence the original message is recovered.

This algorithm is implemented on Windows 10 operating system with 64-bit platform using Java programming environment.

5.2   Security Analysis of the Proposed Methodology

- Two levels of security is implemented in this algorithm since two levels of substitution is applied on the key.
- Depending on the number of nucleotides present in the selected DNA sequence the security of the key varies. Calculations show that for trying 4^x times only once the key can be broken where x is the number of nucleotides. If the value of x is more than 9 then the attacker would require a million chances to decode the key. It can be concluded from the above fact that the algorithm is very much secured from a dictionary attack.
  Explanation: Four different nucleotides, i.e., A, T, C, and G can be chosen in order to make the DNA sequence. Therefore, according to permutation rule if six is the length of the DNA sequence then total number of possibilities of forming the sequence will be $4 \times 4 \times 4 \times 4 \times 4 \times 4 = 4^6$. Therefore, for n nucleotides the total number of permutations will be 4^n i.e., the complexity of permutation will be exponential in nature.
- Another level of security is incorporated in the encryption scheme during substitution of decimal values of DNA sequence into the binary sequence as the substitution occurs in an scrambled fashion. It is practically very hard for the attacker to guess it, which would take 16 guesses to find the correct binary sequence for one decimal value. Hence, for 10 different numbers, there will be $^{16}P_{10}$, i.e., more than 29 billion chances are required to obtain the encrypted key (K_E).
- This algorithm is highly flexible because the sender can encrypt the plain text using the same key for the whole message. On the other hand, the sender can add enhanced security by encoding each word with different keys. But in that case, the time complexity will increase rapidly. Depending on the requirement of sender, the algorithm can be modified accordingly.
- The complement operations which are applied here also act as a layer of security and making the algorithm even more difficult for the attacker to decipher the message.

6 Conclusion

The advantage of this algorithm is that it is a fusion of many encryption methods. Substitution approach along with XOR operation provides an added security. The permutation table, i.e., Table 5 is the backbone for encrypting the key of this proposed methodology. Mathematical analysis shows that once in a million attempts, the key can be deciphered. Additional advantages of this technique are its user friendliness and effectiveness. Depending upon the need, sender can set the security of the algorithm and adjust the time complexity accordingly. Implementation of the proposed algorithm on a DNA microchip will require very high tech laboratories and is cost effective. Future work regarding this algorithm involves analyzing and decreasing the time and space complexity in real time. This is a symmetric-key algorithm, using the same logic new hybrid cryptographic models can be designed and developed.

References

1. Kahate, A.: Cryptography and Network Security, 3rd edn. New Delhi (2016)
2. Adleman, L.M.: Molecular computation of solutions to combinatorial problems. Science **266** (5187), pp. 1021–1025 (1994)
3. Hao, L.B., Lee, H.C., Zhang, S.Y.: Fractals related to long DNA sequences and complete genomes. Chaos Solitons Fractals **11**, pp. 825–836 (2000)
4. Zhang, Q., Zhou, S., Wei, X.: An efficient approach for DNA fractal-based image encryption. Appl. Math. Inf. Sci. **5**(3), pp. 445–459 (2011)
5. Zhang, Y., Fu, B., Zhang, X.: DNA cryptography based on DNA fragment assembly. Information Science and Digital Content Technology (ICICDT). In: IEEE International Conference, vol. 1, pp. 179–182 (2012)
6. Verma, A.K., Dave, M., Joshi, R.C.: Securing ad hoc networks using DNA cryptography. In: IEEE International Conference on Computers and Devices for Communication, pp. 781–786 (2006)
7. Cui, G., Qin, L., Wang, Y., Zhang, X.: An encryption scheme using DNA technology. In: Bio-Inspired Computing: Theories and applications, IEEE International Conference, pp. 37–42. Adelaide, SA, Australia (2008)
8. Yupeng, Z., Yu, Z., Zhong, W., Sinnot, R.O.: Index based symmetric DNA encryption algorithm. In: IEEE, 4th International Congress on Image and Signal Processing, pp. 2290–2294 (2011)
9. Kannadasan, R., Basha, M.S., Emerson, I.A.: Survey on molecular cryptographic network DNA (MCND) using BIG DATA. In: 2nd International Symposium on Big Data and Cloud Computing (ISBCC) Procedia Computer Science, vol. 50, pp. 3–9 (2015)
10. Reza, M., Torkaman, N., Kazazi, N.S.: A method to encrypt information with DNA-based cryptography. Int. J. Cyber Secur. Digit. Forensics (IJCSDF) **4**(3), pp. 417–426 (2015)
11. Tanaka, K., Okamoto, A., Saito, I.: Public-key system using DNA as a one-way function for distribution. Biosystem **81**(1), pp. 25–29 (2015)
12. Singh, H., Chugh, K., Dhaka, H., Verma, A.K.: DNA based cryptography: an approach to secure mobile networks. Int. J. Comput. Appl. **1**(19), pp. 77–80 (2010)

Enhanced Surveillance Using Integration of Gait Analysis with Iris Detection

Divya Abhilash and Divya Chirayil

Abstract The programmed way of establishing and validating the existing person upon their corporal and observable characters are termed as biometric technology. Due to the accuracy of the iris recognition, it becomes more dominant in the available biometric techniques. The current study aims at a new technology for iris recognition which helps us to identify a human by the iris from various places. This method is more valid and protected when compared with the other biometric technologies. In this biometric, the human characters are used which will not change during the lifetime of that particular individual. The time taken for identification of individual human is very less. Iris recognition uses the uniqueness of the eye and the information is stored in the iris database. The movement of the human individual is recognized by the gait analysis and iris recognition is more liable in the existing biometric systems. The study mainly focuses on the iris preprocessing, edge detection, and feature extraction, finally gait and iris fusion classification techniques in the research area. Without the help of a particular individual, we can use the gait analysis for iris recognition. The security places such as banks, places which are used in elections, military installations and even airport, where more restriction to provide the details of the human use this biometric technology.

Keywords Human identification · Biometric recognition · Challenge–response test · Iris image acquisition · Normalization · Feature extraction

D. Abhilash (✉) · D. Chirayil
Department of EXTC, Pillai HOC College of Engineering & Technology,
Rasayani, Maharashtra, India
e-mail: abhilashdivya96@gmail.com

D. Chirayil
e-mail: dchirayil@mes.ac.in

1   Introduction

For the various applications in the developing fields, the use of biometric systems became more important. The technique such as fingerprint, palm print, hand geometry is already used in applications. There are few drawbacks in these systems: (1) Clarity of the image is very low when it is taken from a far distance. (2) A collaboration with the individual is necessary for getting the correct results [1]. Gait recognition aims essentially to address this problem by identifying people based on the way they walk. Gait the way of walking, can be captured from a distance and provide more efficient means of verification. In this, we propose a fusion of gait—a behavioral trait and iris—a physiological trait for identity verification for low-bandwidth video surveillance scenarios. To make feature extraction easier and simpler in gait recognition process, variable 2-DPCA and Manhattan distance classifier is used. The recognition performance of dynamic features extracted from different parts of human body in an appearance based set up [2]. Iris verification and the manipulation process easily converts the captured image into codes for the present user. The major advantages are an easy and accurate prediction and less negative results. At present, the human motion analysis and biometric are more important for the security purpose such as military, banks, parks, airports, etc. For this reason, this study becomes more important.

2   Related Work

Using different classifiers, gait recognition is improved which is studied by Nancy [3]. The paper deals with the study of the gait recognition system with many approaches. Correct Classification Rate (CCR) is resulted due to the application of the MDA (Multilinear Discriminant Analysis) and LDA (Linear Discriminant Analysis) technique. The image obtained by the gait pal and pal entropy uses the principal component analysis and Support Vector Machine (SVM) for the gait recognition Jeevan et al. [4].

The study of the gait recognition system done by the Robindeepkaur and Kakkar [5] for the classification of the methods is based on the model-free and model-based methods. The live human activities shown in the video format is applied in the Design and Implementation of Gait Recognition System by the Alese et al. [6].

The Mixture of Gaussian model (MoG) and medial axis transformation techniques are mostly used for the identification of the present methods. The gait signature is obtained by the difference in the limbs, hands, and body posture. Outermost Contour Principal Component Analysis (PCA) is the method followed for the efficient gait recognition for the reduction of the dimension was studied by the Liu et al. [7]. The MDA with ENN, BPNN, and SVM are the methods used for the classification of the used signatures. Component analysis is the other technique used for the gait recognition which is studied by the Ali et al. [8].

A comparison study was done for the available gait recognition, PCA with the proposed study. The three methods of walk styles are tested. They are slow, fast, and with a ball. The combination of the width and shape of the binary silhouette is used for the identification of the individual which is analyzed by the Bhuyan and Jagan [9].

For the removing of the noise in the background, the filtering operation is done. For the calculation of the background information, Gaussian Mixture Model is used. The dimensionalities in the data are minimized by the PCA. PCA and LDA are implemented for the gait recognition by Cheng et al. [10]. The measurements of the two sequences are done by the Normalized Euclidean Distance (NED) and the recognition is done by the K-nearest neighbor classification (KNN). Correct classification rates for the side view, oblique view, and frontal view are 65, 63.75, and 77.5% which has the NN and STC, 65, 66.25, and 85% with NN and NED, and 75, 81.25% with ENN and NED [11].

3   Proposed Work

3.1   Overview

The particular sole's growth was characterized by the gait behavior. It depends on the external structure of the individual. It mainly involves the human body structure, motion regulating system, behavioral and psychological activities during the movement of a person. The present study detailed on the gait training set and testing sample for the gait recognition which is compared with the iris recognition biometrics. The training set and testing the particular samples are performed after the loading of the training set and samples. Silhouette conversion is obtained from the silhouette samples and iris recognition.

3.2   Gait Recognition

Gait will acquire the effective way for the recognition of the individual during the motion. The current study is focused on the effective algorithm for the mass and boundary-related studies of the individual. The variation in the gait cycle is mainly due to the gait recognition of the individuals. The extraction of the correct parameters of the gait recognition process is used as the substitute for the drawbacks caused by these issues. Variable 2-DPCA and Manhattan distance classifiers are used as a solution for this problem.
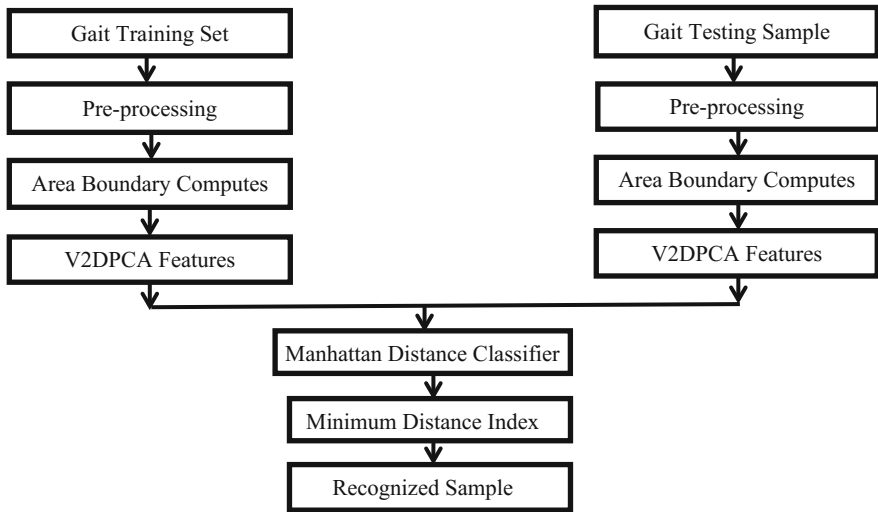
Fig. 1 Gait architecture

3.3 V2DPCA Feature

Gait training set and testing sample utilized for the gait recognition are detailed in this thesis. The samples are utilized for the study after the insertion of the training set and testing samples for the processing. The silhouette samples are used for the conversion of the silhouette in the area of the boundary (Fig. 1).

Different variable factors are used for the calculation of the V2DPCA for various gait cycles. V2DPCA has low computational complexity, enhances the recognition accuracy, and compresses the memory.

3.4 Manhattan Distance Classifier

The differences in the data points are important for the measurement of the distance. The two images that have similar measurements are alike and different objects may vary. The scope of the study is to calculate the distance between the images. The distance in between the training set and the samples are calculated by the Manhattan distance classifier. The distance that shows less difference between each other shows the increased matching cycle.

3.5 Basic Steps in Iris Recognition

The iris recognition involves five steps: image acquisition, iris segmentation/localization, iris normalization, feature extraction, and matching/classification. All the steps are independent and affected by the individual activities. The separation is more important as it may affect the consequent steps in iris recognition [12] (Fig. 2).

3.5.1 Iris Image Acquisition

By acquiring the image of the particular person in the imaging system with the cameras and the sensors are used. When capturing an iris image in unconstrained conditions and without the cooperation of user, the image quality can be highly degraded by poor focus, off-angle view, motion blur, specular reflection (SR), and other artifacts. This will degrade the recognition accuracy. To overcome these problems, K-means clustering algorithm and circular Hough transform was used to localize iris boundary [13]. These cameras and sensors will have a visible light or near-infrared light. The final images are stored on the computer hard disk or server as a bitmap or jpeg format.
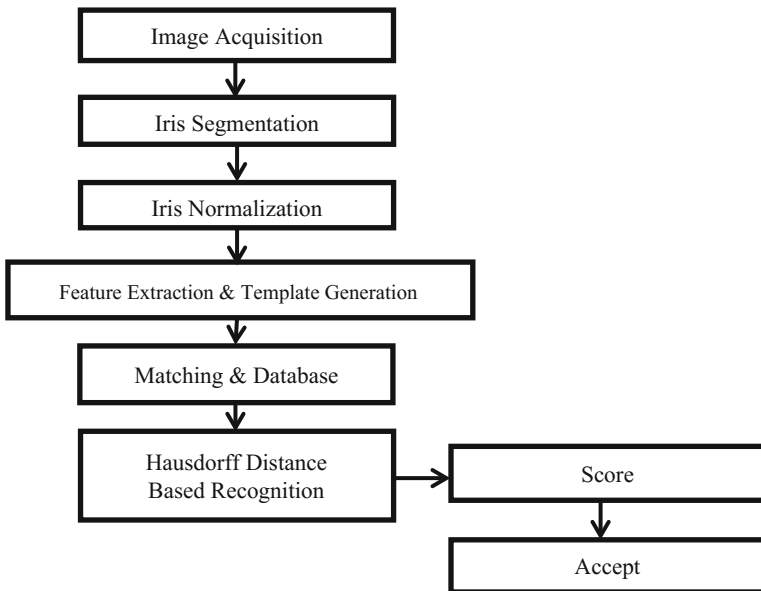
Fig. 2 Iris architecture

3.5.2 Iris Segmentation

The process that refers to the complete identification of the boundaries of the iris and the pupil of an iris in their eye images are shown in the iris segmentation. The perfect information of the individual is identified by the extraction of the process. The separation from the other region such as the parts outside the iris (eyelids, eyelashes, and skin) is done in order to avoid the errors. The image clarity depends on the eye image clarity. The main aim of the separation of the images is to find out the iris boundaries and pupil boundaries and to make the template very clear. The eye image is fitted to the iris region to identify the inside and outside borders needed for the pupil and sclera. The middle point of the radius of the pupil and iris was determined. By this occlusion can be neglected. Eyelids, eyelashes, and their reflections are also identified. The work of separation starts at the Iris where the image is very dark. A boundary is fixed for the iris and various techniques are applied for the detection of the noises from various region of the iris part.

3.5.3 Iris Normalization

This method is very helpful in the determination of the iris region to neglect the dimension and blur images.

3.5.4 Feature Extraction and Code Generation

The iris region and the codes for their features are extracted in the normal sites. The algorithm is used for the template regions. The data that is inserted for the process is reduced to the data that are represented as the final set. Using the human silhouette, a boundary box is created for the complete size length. Gait sequence is segmented into different cycles. The various cycles are calculated with the frame numbers in two rest positions. The rest position that is available are person's rest, left foot forward, rest, right foot forward, rest which is known as gait cycle. The centroid is developed due to segmentation of the central line in the silhouette. To calculate the Hausdroff distance, we use

$$d_H(X, Y) = \max\left\{\sup_{x \in X} \inf_{y \in Y} d(x, y), \inf_{y \in Y} \sup_{x \in X} d(x, y)\right\}$$

where X and Y are two nonempty subsets.

3.5.5 Fusion

Identification is done using iris template. The templates are used for the recognition of the differences between the measurements of the two regions. The templates are stored in the database. The present study will be proposed on the images of the
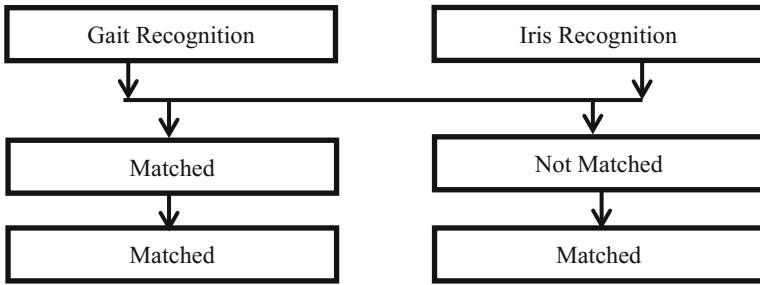
Fig. 3 Comparing the two recognition fusion scores

individuals which are obtained from the gait database that already exists. The general classification of the templates is Hausdorff distance, normalized correlation, and Manhattan distance classifier (Fig. 3).

4 Result and Discussion

Table 1 shows the comparison of different gait recognition processes with the proposed V2DPCA and Manhattan distance classifier method.

Accuracy sensitivity and specificity were used to evaluate the performance and they were defined as the follows respectively:

Table 1 Correlation of various methods

Method	Recognition rate
Inter-valued type	Good
	Drawback: In this method, we must find axis of least inertia for extracting features from the silhouettes
HMM	Low
	Drawback: This method cannot recognize gait with accuracy using frontal view efficiently
Correlation analysis	Low
	The motion of two legs during the human walking process is most important
Normalization, PCA, Euclidean distance	Moderate
	This technique recognizes a person by using gait motion capture data
V2DPCA, Manhattan distance	High accuracy
	It requires less memory and increased efficiency given the processes take place in smaller dimensions

Fig. 4 Classification techniques
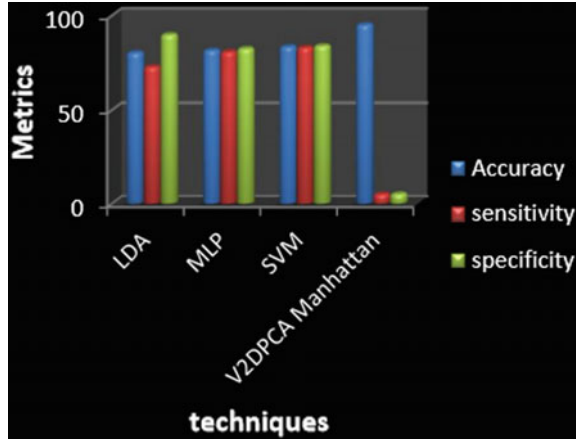
$$Accuracy = \frac{TR + TL}{(TR + TL + FR + FL)} * 100,$$

$$Sensitivity = \frac{TR}{TR + FL} * 100,$$

$$Specificity \; = \frac{TL}{TL + FR} * 100,$$

where TR denotes the number of true right-side gait patterns that correctly iden-tifies. FR is the number of false right-side gait patterns recognized. FL is the number of identification of a left-side gait pattern as right-side, and TL is the number of true left-side gait patterns.

Figure 4 shows the precise, specific, and liable classifier with the LDA and MLP and SVM and V2DPCA Manhattan technique. In this, the V2DPCA Manhattan showed more liable results.

5   Gait and Iris Dataset

CASIA Gait Database and CASIA IrisV4 are provided by The Institute of Automation Chinese Academy of Sciences. CASIA Gait Dataset B is a large multi-view dataset containing 124 subjects from 11 views and CASIA IrisV4 contains six subsets.

6 Conclusion

The development in the recent research works in the computer vision showed an increased interest in the human identification. The main scope of the gait recognition is to recognize the people while they are in movement. The present study showed accuracy with the different gait cycles in the gait and iris recognition system. The accuracy mainly depends on the different gait cycles. The present study showed the gait and iris recognition using the V2DPCA and Manhattan distance classifier for good results.

References

1. Tao, D., Li, X., Wu, X., Maybank, S.J.: General tensor discriminant analysis and gabor features for gait recognition. IEEE Trans. Pattern Anal. Mach. Intell. **29**(10), 1700–1715 (2007)
2. Amin, T., Hatzinakos, D.: Determinants in human gait recognition. J. Inf. Secur. **3**, 77–85 (2012)
3. Nancy, T.S.: Improved gait recognition using different classifiers. Int. J. Comput. Appl. Eng. Technol. **3**(2), 193–198 (2014). ISSN: 2277-7962
4. Jeevan, M., Jain, N., Hanmandlu, M., Chetty, G.: Gait Recognition Based on Gait Pal and Pal Entropy Image. IEEE (2013)
5. Robindeepkaur, Kakkar, N.: Person tracking using gait. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(8) (2013). ISSN: 2277 128X
6. Alese, B.K., Mogaji, S.A., Adewale, O.S., Daramola, O.: Design and implementation of gait recognition system. Int. J. Eng. Technol. **2**(7) (2012). ISSN: 2049-3444
7. Liu, L., Yin, Y., Qin, W., Li, Y.: Gait recognition based on outermost contour. Int. J. Comput. Intell. Syst. **4**(5), 1090–1099 (2011)
8. Ali, H., Dargham, J., Ali, C., Moung, E.G.: Gait recognition using principle component analysis. In: The 3rd International Conference on Machine Vision, ICMV 2011 (2010). ISBN: 978-1-4244-8889-6
9. Bhuyan, M.K., Jagan, A.: Person identification using gait by combined features of width and shape of the binary silhouette. Int. J. Comput. Inf. Eng. **4**(4) (2010)
10. Cheng, Q., Fu, B., Chen, H.: Gait recognition based on PCA and LDA. Int. Comput. Sci. Comput. Technol. 124–127 (2009). ISBN 978-952-5726-07-7
11. Wang, L., Tan, T., Ning, H., Hu, W.: Silhoutte analysis-based gait recognition for human identification. IEEE Trans. Pattern Anal. Mach. Intell. **25**(12) (2003)
12. Proença, H., Alexandre, L.A.: Toward noncooperative iris recognition: a classification approach using multiple signatures. IEEE Trans. Pattern Anal. Mach. Intell. **29**(4), 607–612 (2007)
13. Gupta, R., Kumar, A.: An effective segmentation technique for noisy iris images. Int. J. Appl. Innov. Eng. Manag. **2**(12), 118–125 (2013)
14. Timm, F., Barth, E.: Accurate eye centre localisation by means of gradients. Int. Conf. Comput. Theory Appl. 125–130 (2011)

Effect of Tapering on Swimming Efficiency of Flagellated Microswimmer at Low Reynolds Number

T. Sonamani Singh and R. D. S. Yadava

Abstract The swimming efficiency of a microswimmer composed of a rigid spherical head fixed to an elastic filament with tapered cross section is analyzed in Stokes flow regime. The small-amplitude harmonic base-angle actuation and linear taperings are assumed. The formulation of elastohydrodynamic equations and solution procedure are adapted from the theory of swimming developed by Lauga. The numerical results on the propulsion speed and the energy efficiency of the swimmer are presented as a function of the filament length and taper ratio. The results demonstrate that by introducing a linear filament tapering both the propulsion speed and the energy efficiency can be enhanced with an appropriate trade-off with the filament length.

Keywords Microswimming · Elastohydrodynamics · Tapered tail elastic swimmer · Locomotion at low Reynolds number

1 Introduction

Artificial microswimming is a biomechanics-inspired interdisciplinary subject. It deals with the self-propulsion and navigation of micron-scale objects in fluid media at low Reynolds number. Besides understanding the navigation of microorganisms, there is substantial research interest in this topic due to its significance for the developments in micro/nanorobotics and smart drug delivery systems [1].

T. Sonamani Singh (✉) · R. D. S. Yadava
Sensors & Signal Processing Laboratory, Department of Physics, Institute of Science,
Banaras Hindu University, Varanasi 221005, India
e-mail: sonamani.2065@gmail.com

R. D. S. Yadava
e-mail: ardius@bhu.ac.in; ardius@gmail.com

The microorganisms with flagella (whip-like structure attached to the body) swim by periodic lashing against the fluid medium. The flagellar waving causes the fluid to exert propulsive thrust on the cell body due to viscous drag effects in Stokes flow. The developments in this field are inspired mainly by the locomotion of flagellar microorganisms at low Reynolds number [2, 3]. An artificial microswimmer consists of an analogous structure: a long elastic filament analogous to flagellum or tail is appended to a rigid body referred to as the head. In the simplest form, the waving of the filament is actuated by an excitatory source at the head–tail interface. This is known as boundary actuation [3]. In another actuation mechanism, the tail is modeled as an active filament with a distributed mechanism analogous to the tubular flagellum in some microorganisms. In this, a pair of concentric biological tubules is connected via molecular motors (called dynein). The back and forth motion of which cause the sliding of tubules past each other resulting in bending motion of the flagellum [3].

Here, we consider the case of microswimming by boundary actuation. The problem has two aspects to deal with: the elastohydrodynamics of the filament that generates propulsive thrust due to viscous drag effects in the Stokes flow regime and the kinematics of the head motion. The theoretical formulation of this problem has been developed in [4, 5] by considering filament under small-amplitude transverse boundary actuation with no twisting. The filaments are taken to be of a homogeneous elastic material with the uniform cross section. The elastohydrodynamic equations are solved for the shape of the filament as a function of position and time, which is then used to calculate the thrust exerted by the fluid on the filament. The swimming kinematics is worked under the force-free Stokes flow condition. The swimmer speed and swimming efficiency are calculated to reveal the dependencies on swimmer length and actuation frequency besides other intrinsic parameters associated with filament material and fluid [5]. Also, in [5] scaling arguments have been used to present a more generalized treatment that asserts that both the elastohydrodynamic and the kinematic aspects must be solved simultaneously. This leads to an integro-differential equation instead of the usual hyper-diffusion equation. The analyses led to the conclusion that there exist some optimal dimensions for the swimmer head and the filament length that will yield large swimming speed with optimal energy efficiency.

In this work, we explore another dimensional feature of the filament which implements filament tapering. Adapting from the swimming theory in [5], we analyze here the influence of cross-sectional variation in the swimmer speed and energy efficiency. Recently, we analyzed a simpler case of tapered filament under planar actuation (that is, motion confined to a plane) with filament dynamics and swimmer kinematics done separately [6]. Here, we extend this work to a generalized 3D boundary actuation of the tapered filament based on the integro-differential equation in [5] that accounts for the coupling between dynamical and kinematical aspects.

2 Theory of Microswimming

2.1 Microswimmer

The prototype microswimmer is composed of an elastic filament (tail) attached to a spherical rigid body head as shown in Fig. 1. Lauga has developed the theory of microswimming for this type of swimmer having a uniform cross section [5]. In the present work, we consider the filament to be linearly tapered (cross section varying linearly along the length) and analyze its effect on the propulsion speed and energy efficiency. The spherical head in the present analysis makes a special case of the swimmer in [5]. The swimming occurs due to hydrodynamic drag generated along the filament length while flexural bending vibrations are actuated in the transverse direction. This is implemented by embedding a periodic rotational actuation mechanism in the swimmer body, which entails a periodic variation of the slope (or the base angle) at the head–tail junction. The low Reynolds number swimming in Stokes flow condition is assumed. The actuation at the head–tail junction generates bending waves propagating down the filament length. The propulsion occurs by longitudinal component of hydrodynamic drag distributed along the filament and balanced at every point by the filament elasticity. In the analysis presented below, we adapted the Lauga's theoretical formulation and incorporate the effects of filament tapering. The influence of filament tapering on propulsion speed and energy efficiency are numerically calculated and compared with the uniform cross-section filament.

The swimmer dimensional parameters and the coordinate systems used are defined in Fig. 1. The coordinate system $\{\hat{e}_x, \hat{e}_y, \hat{e}_z\}$ is fixed at the head–tail junction, hence it moves with the swimmer body. The coordinate system $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ is fixed in the laboratory frame. The x-direction ($\hat{e}_1$ and $\hat{e}_x$) is taken along the mean
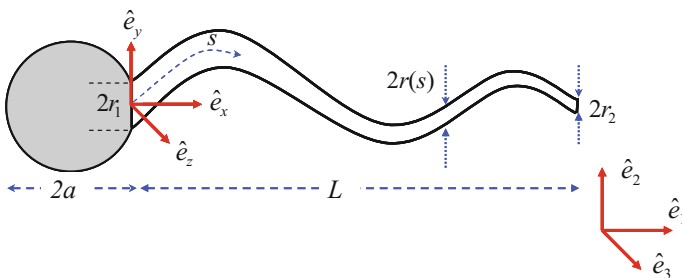
Fig. 1 Dimensional parameters and coordinate system for the elastic filament swimmer with tapered cross section: $\{\hat{e}_x, \hat{e}_y, \hat{e}_z\}$ is fixed with the swimmer body and $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ is fixed in the laboratory frame, r_1 and r_2 are respectively the head-end and the tail-end filament radii, s is the curvilinear distance along the centerline from the head end, $r_2 \leq r(s) \leq r_1$ is the filament radius at s, a is the spherical head radius and L is the filament length (assumed to be inextensible for small-amplitude oscillations)

position of the filament pointing away from the head. The periodic base-angle actuation occurs in $(\hat{e}_y, \hat{e}_z)$ plane, and is defined at angular frequency ω as

$$y(0, t) = 0, \qquad \frac{\partial y}{\partial x}(0, t) = \varepsilon \cos \omega t, \tag{1}$$

$$z(0, t) = 0, \qquad \frac{\partial z}{\partial x}(0, t) = \delta \varepsilon \sin \omega t, \tag{2}$$

where ε denotes the slope amplitude in $(\hat{e}_x, \hat{e}_y)$ plane and $\delta \varepsilon$ in the $(\hat{e}_x, \hat{e}_z)$ plane; $0 \leq \delta \leq 1$ is a measure of the three dimensionality of the actuation. No twisting of the filament is considered.

2.2 Dynamical Equations

The mathematical formulation is done first in the body-fixed frame, and then transformed to the laboratory frame. Let $\vec{r}(s, t)$ be the position vector of the centerline of the filament with s being the arc length ($0 \leq s \leq L$) and L is the filament length (assumed to be inextensible). The elastic energy of the tapered elastic filament associated with pure bending (no torsion) is given by [5]

$$G = \frac{1}{2} \int_0^L \left[A \left(\frac{\partial^2 \vec{r}}{\partial s^2} \right)^2 + \sigma \left(\frac{\partial \vec{r}}{\partial s} \right)^2 \right] ds, \tag{3}$$

where $A(s) = EI(s)$ is the bending stiffness with E the Young's modulus and I the second moment of area, and $\sigma(s, t)$ is the Lagrangian multiplier enforcing the condition of local inextensibility

$$\frac{\partial}{\partial t} \left(\frac{\partial \vec{r}}{\partial s} \cdot \frac{\partial \vec{r}}{\partial s} \right) = 0 \tag{4}$$

The bending stiffness is now position dependent due to tapering. For a cylindrical filament with circular cross section, the bending stiffness is given by

$$A(s) = \frac{E\pi(r(s))^4}{4}. \tag{5}$$

Considering linear taper and cross-sectional radius at head end ($s = 0$) to be r_1 and at tail end ($s = L$) to be r_2 the tapered radius of the filament can be written as

$$r(s) = r_1 + m_l s \quad \text{with } m_l = \frac{(r_2 - r_1)}{L}. \tag{6}$$

From (3) the following elastohydrodynamic equation of motion is obtained by applying the resistive-force theory (RFT) for the local viscous drag and by using the calculus of variation [5]

$$[\xi_{\|}\hat{t}\hat{t} + (I - \hat{t}\hat{t})] \cdot \vec{u} = -\frac{\partial^2}{\partial s^2}\left(A\frac{\partial^2 \vec{r}}{\partial s^2}\right) + \frac{\partial}{\partial s}\left(\sigma\frac{\partial \vec{r}}{\partial s}\right) \tag{7}$$

where $\vec{u}$ is the local instantaneous velocity along the filament, $\hat{t}$ is the local tangent unit vector ($\hat{t} = \partial\vec{r}/\partial s$), I is the identity matrix, and $\xi_{\|}$ and $\xi_{\perp}$ are the tangential and the normal drag coefficients respectively. Equations (4) and (7) along with the actuation and the boundary conditions define the filament swimming problem.

In Lauga's paper, the base-angle actuation is visualized to be caused by the rotation of the head around the head–tail junction as a hinge. This is intrinsically associated with a three-dimensional translational velocity ($\vec{U}$) and a rotational velocity ($\vec{\Omega}$) of the head center of mass. Denoting the three-dimensional instantaneous velocities in components form as $\vec{u} = \{u_x, u_y, u_z\}$, $\vec{U}(t) = \{U_x, U_y, U_z\}$ and $\Omega(t) = \{\Omega_x, \Omega_y, \Omega_z\}$ the relation among the components are given as

$$u_x = U_x + z\Omega_y - y\Omega_z, \tag{8}$$

$$u_y = U_y + \frac{\partial y}{\partial t} + x\Omega_z - z\Omega_x, \tag{9}$$

$$u_z = U_z + \frac{\partial z}{\partial t} + y\Omega_x - x\Omega_y. \tag{10}$$

To proceed further, the problem is simplified by making the small-slope approximation for the filament displacements. This allows us to make approximations $\vec{r}(s,t) \approx x\hat{e}_x + \vec{r}_{\perp}$, where $\vec{r}_{\perp} = y(x,t)\hat{e}_y + z(x,t)\hat{e}_z$ and $s \approx x$. Then, by retaining only the leading order terms the dynamical equations (4) and (7) yield the following four coupled equations for components

$$\xi_{\|}u_x + (\xi_{\|} - \xi_{\perp})\frac{\partial y}{\partial x}u_y + (\xi_{\|} - \xi_{\perp})\frac{\partial z}{\partial x}u_z = \frac{\partial\sigma}{\partial x} \tag{11}$$

$$(\xi_{\|} - \xi_{\perp})\frac{\partial y}{\partial x}u_x + \xi_{\perp}u_y + (\xi_{\|} - \xi_{\perp})\frac{\partial y}{\partial x}\frac{\partial z}{\partial x}u_z = -\frac{\partial^2}{\partial x^2}\left(A\frac{\partial^2 y}{\partial x^2}\right) + \frac{\partial}{\partial x}\left(\sigma\frac{\partial y}{\partial x}\right), \tag{12}$$

$$(\xi_{\|} - \xi_{\perp})\frac{\partial z}{\partial x}u_x + (\xi_{\|} - \xi_{\perp})\frac{\partial y}{\partial x}\frac{\partial z}{\partial x}u_y + \xi_{\perp}u_z = -\frac{\partial^2}{\partial x^2}\left(A\frac{\partial^2 z}{\partial x^2}\right) + \frac{\partial}{\partial x}\left(\sigma\frac{\partial z}{\partial x}\right), \tag{13}$$

$$\frac{\partial^2 \sigma}{\partial x^2} - \sigma \left(\frac{\xi_\parallel}{\xi_\perp} \right) - \frac{\partial^3}{\partial x^3} \left(A \frac{\partial^2 \vec{r}}{\partial x^2} \right) \cdot \frac{\partial \vec{r}}{\partial x} + \left(1 - \frac{\xi_\parallel}{\xi_\perp} \right) \left(\frac{\partial^2}{\partial x^2} \left(A \frac{\partial^2 \vec{r}}{\partial x^2} \right) \right) \cdot \frac{\partial^2 \vec{r}}{\partial x^2} = 0. \quad (14)$$

These are analogous to the Eqs. (19a)–(19d) in [5], but modified in the terms involving the position-dependent filament radius $r(x)$. In the case of no tapering, these equations become identical to that in [5].

Under small-slope approximation, the boundary conditions at the two ends of the filament also get somewhat simplified. Up to leading order terms these are stated as: at $s = 0$,

$$\vec{F}_{ext} = \frac{\partial}{\partial x} \left(A \frac{\partial^2 \vec{r}_\perp}{\partial x^2} \right) - \sigma \frac{\partial \vec{r}_\perp}{\partial x} - \sigma \hat{e}_x, \ T_{ext,y} = A \frac{\partial^2 z}{\partial x^2}, \ T_{ext,z} = -A \frac{\partial^2 y}{\partial x^2} \qquad (15)$$

at $s = L$,

$$\vec{F}_{ext} = -\frac{\partial}{\partial x} \left(A \frac{\partial^2 \vec{r}_\perp}{\partial x^2} \right) + \sigma \frac{\partial \vec{r}_\perp}{\partial x} + \sigma \hat{e}_x, \ T_{ext,y} = -A \frac{\partial^2 z}{\partial x^2}, \ T_{ext,z} = A \frac{\partial^2 y}{\partial x^2}, \qquad (16)$$

where $\vec{F}_{ext}$ and $\vec{T}_{ext}$ are externally applied forces and torques.

2.3 Free Swimming

The free swimming of the filament means there are no external forces and torques applied to the swimmer. The elastic forces and torques at the head–tail junction must, therefore, be balanced by the hydrodynamic forces and the torques generated on the head due to the intrinsic rotational actuation mechanism of the swimmer. The latter calculated in the form of resistance matrices about the center of mass of the head when transformed to the origin of the body-fixed frame (head–tail junction point) make $\vec{F}_{ext}$ in (15). Since the tail end is free, the $\vec{F}_{ext}$, $T_{ext,y}$ and $T_{ext,z}$ there will be zero. Thus, we obtain the boundary equations (17a)–(17e) analogous to the Eqs. (26a)–(26e) in [5] at $s = 0$ (modified due to tapering), and Eq. (18) analogous to the Eq. (27) in [5] at $s = L$:

$$R_\parallel^{FU} U_x(t) = \sigma(0, t) \qquad (17a)$$

$$R_\perp^{FU} U_y(t) - a R_\perp^{FU} \Omega_z(t) = -\frac{\partial}{\partial x} \left(A \frac{\partial^2 y}{\partial x^2}(0, t) \right) + \sigma(0, t) \frac{\partial y}{\partial x}(0, t) \qquad (17b)$$

$$-a R_\perp^{FU} U_y(t) + (R_\perp^{L\Omega} + a^2 R_\perp^{FU}) \Omega_z(t) = A \frac{\partial^2 y}{\partial x^2}(0, t) \qquad (17c)$$

$$R_\perp^{FU} U_z(t) + a R_\perp^{FU} \Omega_y(t) = -\frac{\partial}{\partial x}\left(A \frac{\partial^2 z}{\partial x^2}(0,t)\right) + \sigma(0,t)\frac{\partial z}{\partial x}(0,t) \tag{17d}$$

$$a R_\perp^{FU} U_z(t) + (R_\perp^{L\Omega} + a^2 R_\perp^{FU})\,\Omega_y(t) = -A\frac{\partial^2 z}{\partial x^2}(0,t) \tag{17e}$$

$$\frac{\partial^2 y}{\partial x^2}(L,t) = 0,\ \frac{\partial^3 y}{\partial x^3}(L,t) = 0,\ \frac{\partial^2 z}{\partial x^2}(L,t) = 0,\ \frac{\partial^3 z}{\partial x^3}(L,t) = 0,\ \sigma(L,t) = 0 \tag{18}$$

where $(R_\parallel^{FU}, R_\perp^{FU})$ and $(R_\parallel^{L\Omega}, R_\perp^{L\Omega})$, respectively, are the force–velocity and the torque–rotation rate resistivities acting on the swimmer in the center of mass frame and a is the distance of the center of mass from the point of actuation [5].

Next, the equations of motion (11)–(14) were non-dimensionalized following the procedure in [5]. For this, the non-dimensionalized variables and coefficients were defined as

$$x \leftarrow x/l_\omega,\ \dots y \leftarrow y/l_\omega,\ \dots z \leftarrow z/l_\omega,\ \dots t \leftarrow t\omega^{-1},\ \dots A' \leftarrow A/A_o,\ \dots$$
$$F_{ext} \leftarrow F_{ext}/\xi_\perp l_\omega^2\,\omega,\ \dots T_{ext} \leftarrow T_{ext}/\xi_\perp l_\omega^3\,\omega,\ \dots R^{FU} \leftarrow R^{FU}/\xi_\perp l_\omega,\ \dots R^{L\Omega} \leftarrow R^{L\Omega}/\xi_\perp l_\omega^3 \tag{19}$$

with

$$A_o = E I_o = E\pi\,r_1^4/4,\ \dots l_\omega = (A_o/\omega\xi_\perp)^{1/4} \tag{20}$$

Note that (11) and (14) are $O(\varepsilon^2)$ and the leading order in (12) and (13) are $O(\varepsilon)$. Therefore, for calculations of the shape of the filament only the leading order terms in (12) and (13) will be retained. In non-dimensionalized form, these become

$$\frac{\partial y}{\partial t} + \frac{\partial^2}{\partial x^2}\left(A'\frac{\partial^2 y}{\partial x^2}\right) = (C_1 + xC_2)\left[\frac{\partial}{\partial x}\left(A'\frac{\partial^2 y}{\partial x^2}\right)\right]_{x=0} + (C_3 + xC_4)\left(\frac{\partial^2 y}{\partial x^2}\right)_{x=0} \tag{21}$$

$$\frac{\partial z}{\partial t} + \frac{\partial^2}{\partial x^2}\left(A'\frac{\partial^2 z}{\partial x^2}\right) = (C_1 + xC_2)\left[\frac{\partial}{\partial x}\left(A'\frac{\partial^2 z}{\partial x^2}\right)\right]_{x=0} + (C_3 + xC_4)\left(\frac{\partial^2 z}{\partial x^2}\right)_{x=0} \tag{22}$$

with

$$C_1 = -\frac{(R_\perp^{L\Omega} + a^2 R_\perp^{FU})}{R_\perp^{L\Omega} R_\perp^{FU}},\ C_2 = \frac{a}{R_\perp^{L\Omega}},\ C_3 = -\frac{a}{R_\perp^{L\Omega}},\ C_4 = -\frac{1}{R_\perp^{L\Omega}} \tag{23}$$

Assuming the solutions of (21) and (22) to be of the form $y(x,t) = \varepsilon\mathrm{Re}\{e^{-it}\zeta(x)\}$ and $z(x,t) = \delta\varepsilon\,\mathrm{Re}\{ie^{-it}\zeta(x)\}$, we obtain the following differential equation for $\zeta(x)$:

$$\left\{-i+A'\frac{d^4}{dx^4}+2\frac{\partial A'}{\partial x}\frac{d^3}{dx^3}+\frac{\partial^2 A'}{\partial x^2}\frac{d^2}{dx^2}\right\}\zeta=(C_1+C_2 x)\left[A'\frac{d^3\zeta}{dx^3}+\frac{\partial A'}{\partial x}\frac{d^2\zeta}{dx^2}\right]_{x=0}$$
$$+(C_3+C_4 x)\left[\frac{d^2\zeta}{dx^2}\right]_{x=0}$$

$$(24)$$

The solution of (24) yields $y(x,t)$ and $z(x,t)$. Making substitution of these in (11), then integration and use of (17a) results in an explicit expression for U_x in terms of $\zeta(x)$ as

$$U_x = \langle U_x \rangle + U_x' \qquad (25)$$

where $\langle U_x \rangle$ represents the steady part and U_x' represents the oscillating part of the longitudinal motion calculated as

$$\langle U_x \rangle = \frac{\varepsilon^2(1+\delta^2)}{2(L+R_{\parallel}^{FU})}\mathrm{Re}\left\{\frac{\Omega_z}{\varepsilon}\int_0^L \zeta^* dx + (1-\gamma)\int_0^L \left(A\zeta_x\zeta_{xxxx}^* + 2A_x\zeta_x\zeta_{xxx}^* + A_{xx}\zeta_x\zeta_{xx}^*\right)dx\right\}$$

$$(26)$$

$$U_x' = \frac{\varepsilon^2(1-\delta^2)}{2(L+R_{\parallel}^{FU})}\mathrm{Re}\left\{e^{-it}\left[\frac{\Omega_z}{\varepsilon}\int_0^L \zeta dx + (1-\gamma)\int_0^L \left(A\zeta_x\zeta_{xxxx} + 2A_x\zeta_x\zeta_{xxx} + A_{xx}\zeta_x\zeta_{xx}\right)dx\right]\right\}$$

$$(27)$$

In these equations, the subscripted A and ζ are the derivatives with respect to x, where the number of times x has been used represent the order of derivative, and Ω_z is the z-component amplitude of the rotation rate.

Note that $\langle U_x \rangle$ in (25) gives the steady propulsion speed in the body-fixed frame. As in [5], this result can be transformed to the laboratory frame as

$$\langle U_1 \rangle = \langle U_x \rangle + \frac{1}{2}\mathrm{Im}\left\{\hat{U}_z\Omega_y^* + \hat{U}_y^*\Omega_z\right\} \qquad (28)$$

Even though this expression appears similar to that for non-tapered filament in [7] the expression for $\langle U_x \rangle$ given by (26) contains the effect of tapering.

The efficiency of the microswimmer is defined as the ratio of the useful work (that is, the work necessary to move the entire swimmer at the steady speed $\langle U_1 \rangle$) to the total work done by the swimmer. This is obtained as

$$\eta = \frac{\left(R_{\parallel}^{FU}+\frac{L}{\gamma}\right)\langle U_1 \rangle^2}{\varepsilon^2(1+\delta^2)(R_{\perp}^{FU}|\bar{U}_y - a\bar{\Omega}_z|^2 + R_{\perp}^{L\Omega}|\bar{\Omega}_z|^2 + I)}. \qquad (29)$$

where

$$I = \int\limits_0^L \left\{ A^2 |\zeta_{xxxx}|^2 + 4A_x^2 |\zeta_{xxx}|^2 + A_{xx}^2 |\zeta_{xx}|^2 + 4AA_x \zeta_{xxxx}\zeta_{xxx}^* + 2AA_{xx}\zeta_{xxxx}\zeta_{xx}^* + 4A_x A_{xx}\zeta_{xxx}\zeta_{xx}^* \right\} dx$$

(30)

3 Results and Conclusion

The numerical calculations proceeded by first solving (24) for $\zeta(x)$ subject to the boundary conditions defined by (1), (2) and (18). This allowed calculations in (25)–(27) which were then used to calculate the mean propulsion speed and the swimming efficiency by using (28) and (29) respectively. The calculations were done in MATLAB by discretizing the filament and using the function '*bvp4c*' designed for solving the boundary value problems. The filament and fluid parameters used in numerical calculations are: $E = 1$ MPa, $\mu = 0.001$ Pas, $a = 15$ μm, $\varepsilon = 0.4$ rad and $r_1 = 0.2$ μm. The dependencies on the filament length and the taper ratio were analyzed by varying L and the tail-end radius r_2. The results on the propulsion speed and swimmer efficiency in the normalized forms are shown in Fig. 2 as functions of L with r_2 varied as a parameter. For comparison, the results for the uniform cross section (no taper) are also included which are similar to that in Luaga's paper [5].

It can be noted from Fig. 2a that the propulsion speed is enhanced by tapering almost in proportion to the taper ratio. As in the uniform filament case, there is an optimum length for which the propulsion speed is optimal, and its value shifts towards the lower side with increasing taper ratio.

The results on the energy efficiency of the swimmer shown in Fig. 2b are qualitatively similar to that for the uniform filament. However, there are significant quantitative differences. There are two length values for which the efficiency curves show maxima. These maxima shift towards smaller lengths as the taper ratio is increased. An interesting feature is that by introducing small tapering ($r_2 = 0.15$ μm curve) there is large enhancement in the efficiency for all the filament lengths studied. This trend continues with increasing the taper ratio for large length filaments ($L > 62$ μm). However, for smaller length filaments the efficiency first increases and then decreases as the tapering ratio is increased. Another interesting aspect is that the difference between the two maxima reduces along with a simultaneous broadening of the peak at higher length.

Looking at both the results together one can notice that it is possible to trade-off the filament length and the taper ratio for attaining optimal swimming speed and efficiency. If a large propulsion speed is desired it can be achieved by the small length filaments with large taper ratio, and if the energy efficiency is premium it can

Fig. 2 a Variation of propulsion speed with length for different tapering factors. **b** Variation of efficiency with length for different tapering factors. The parameters used in the simulation are $E = 1$ MPa, $\mu = 0.001$ Pas, $a = 15\ \mu$m, $\varepsilon = 0.4$ rad and $r_1 = 0.2\ \mu$m

(a)

$r_2 = 0.05\ \mu$m
$r_2 = 0.1\ \mu$m
$r_2 = 0.15\ \mu$m
Uniform

$\dfrac{\langle U_1 \rangle}{\varepsilon^2}$

$L\ (\mu$m$)$ $\times 10^2$

(b)

$r_2 = 0.05\ \mu$m
$r_2 = 0.1\ \mu$m
$r_2 = 0.15\ \mu$m
Uniform

$\dfrac{\langle \eta \rangle}{\varepsilon^2}$ (%)

$L\ (\mu$m$)$ $\times 10^2$
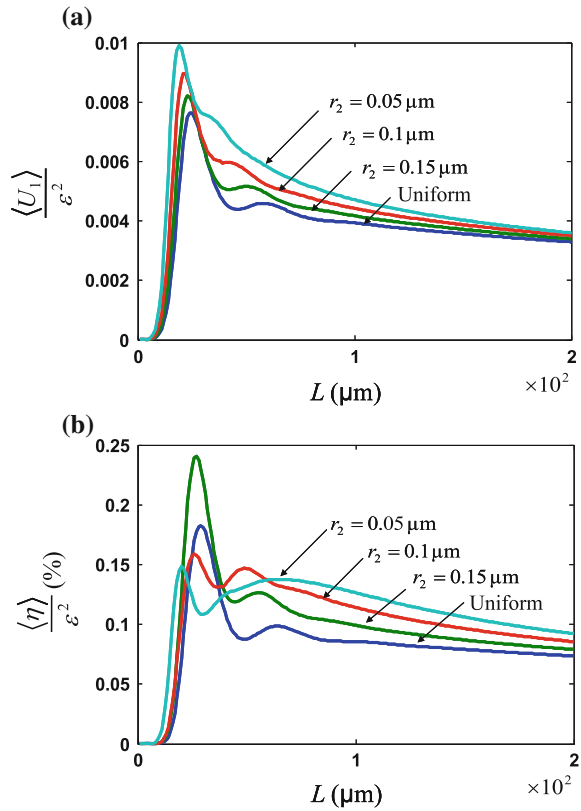
be achieved by the small length filaments with small taper ratio; and, if it is desired to have both at moderately high level then the large filaments with large taper ratio can be employed.

References

1. Li, J., Avila, B.E.F., Gao, W., Zhang, L., Wang, J.: Micro/Nanorobots for biomedicine: delivery. Surg. Sens. Detoxif. Sci. Robot. **2**, 1–9 (2017)
2. Pak, O.S., Lauga, E.: Theoretical model of low-Reynolds-number locomotion. In: Fluid-Structure Interactions in Low-Reynolds-Number Flows, RSC Soft Matter Series, vol. 4. Royal Society of Chemistry, Cambridge, UK (2016)
3. Lauga, E., Powers, T.R.: The hydrodynamics of swimming microorganism. Rep. Prog. Phys. IOP **72**, 1–36 (2009)

4. Wiggins, C.H., Goldstein, R.E.: Flexive and propulsive dynamics of elastica at low Reynolds number. Phys. Rev. Lett. **80**(17), 3879–3882 (1998)
5. Lauga, E.: Floppy swimming: viscous locomotion of actuated elastica. Phys. Rev. E **75** (041916), 1–16 (2007)
6. Singh, T.S., Yadava, R.D.S.: Effect of tapering on elastic filament microswimming under planar body actuation, Biomed. Phys. Eng. Express IOP **4**(015019), (2018)

Field Monitoring Using IoT: A Neural Network Approach

**Ram Krishna Jha, Santosh Kumar, Kireet Joshi
and Rajneesh Pandey**

Abstract A healthy mind requires healthy body and healthy body needs healthy food. Therefore, to provide healthy food to the rapidly growing population of a country is a challenge within the limited fertile land. To fulfill the healthy food demand requires the high production of yield. To achieve high production, it is essential for the farmers to monitor the field from time to time for a good yield, as a little miss can cause disaster and the entire efforts could go waste. The manual monitoring of the field is quite tough and expensive. Therefore, in this paper Internet of Things (IoT) applications are addressed to monitor the field. An Arduino microcontroller board with the soil moisture, temperature, and humidity sensors is used to collect the data from the remote field on the fly. The data once received is analyzed by applying cascade forward and function fitting neural network. Further, the data is tested against an already trained dataset of the field in normal conditions collected over a period of 1 year. The test data, when applied to the trained data, provides a dataset which is used for analysis of the ideal condition of the field. In case of alarming changes in the field properties, a signal is generated and the farmer is informed to take necessary action. A preventive action can be taken to save the crop and also maintain the productivity of the field.

Keywords Soil moisture · Temperature · Humidity · ANN

R. K. Jha (✉) · S. Kumar · K. Joshi · R. Pandey
Graphic Era University, 566/6 Bell Road, Clemet Town, Dehradun, India
e-mail: vickeyjha22@gmail.com

S. Kumar
e-mail: amu.santosh@gmail.com

K. Joshi
e-mail: joshikireet@gmail.com

R. Pandey
e-mail: Rajneesh.p22@gmail.com

1 Introduction

IoT empowers people to carry out routine activities using the Internet and thus spare time and cost making them more productive. The development of IoT-based infrastructure for precision agriculture is one such application. Agriculture is considered as the basis of life as it is the main source of food grains and other raw material for human beings. In most part of the country, farmers use traditional ways of farming. These traditional ways of farming are not helping them to increase the yield of crops, on contrary, the yield is decreasing day by day as the resources required like field and water are decreasing. Automation can help in improving the ecological environment. As the global climate is changing, understanding of the small-scale environment for crop growth is needed along with the wide range of research and study of crop growth [1, 2]. A convenient move can be made if the information is accessible on time.

Furthermore, it provides a thought to the agriculturist what action should be taken. This paper focuses on the IoT-based field monitoring system, which would help the farmer to monitor their fields remotely. The important part of this paper is Zigbee-based remote monitoring system. Second, it has smart irrigation to control the real field of the data. Third, it can monitor the temperature, soil moisture, and humidity of the field. All the things are controlled by the remote device or PC connected to the Internet and the operations have performed by interfacing sensors (temperature, humidity, and soil moisture).

Rest of the paper is organized as follows: Sect. 2 describes the statement of the problem, Sect. 3 describes the proposed solution, Sect. 4 explores implementation and result analysis and finally, the work is concluded in Sect. 5.

2 Problem Statement

The main problem is monitoring of various parameters and their control using Internet of Things (IoT). This is divided into the following subproblem as follows:

  i. To monitor and control various parameters such as temperature, moisture, and humidity using Arduino microcontroller and Internet of Things (IoT).
 ii. To develop the neural network-based predictive system.

2.1 Objective

The main objective is to access the data from the remote field so that timely action if required can be taken. It could be possible to skip the manual monitoring and thus utilize the extra time for other fruitful things and keep an eye on the other important

areas where needed. The collected data can be used to take decisions for alternate crops. A field without altering its mineral content could also be utilized for plantation, cutting down the use of fertilizers.

3 Proposed Solution

The arrangement of the irrigation system is to provide and keep up the perfect condition as far as soil moisture and temperature for the ideal growth of the crops. Utilizing electronic gadgets, for example, smartphones and remote PCs, the client can sign into the cloud storage to fetch the sensor information. The client can monitor the crops and control the water pump and fans in the control board of the UI. There are a few points of interest that can be highlighted in this system. First, it can save time and water as you do not need to go the farm to do the watering (Fig. 1).

In the proposed diagram, temperature, humidity, and soil moisture sensors are deployed on Arduino microcontroller board. The board through a radio model is then connected to the system, where a code is written to store and monitor the data. The architecture is such that with a click of a button a farmer should be able to not only monitor the data but also take timely action. A farmer would himself be able to check diverse parameters of the soil like salinity, acidity, and moisture from time to time. By utilizing this system, they can not only switch on their pump manually [3]
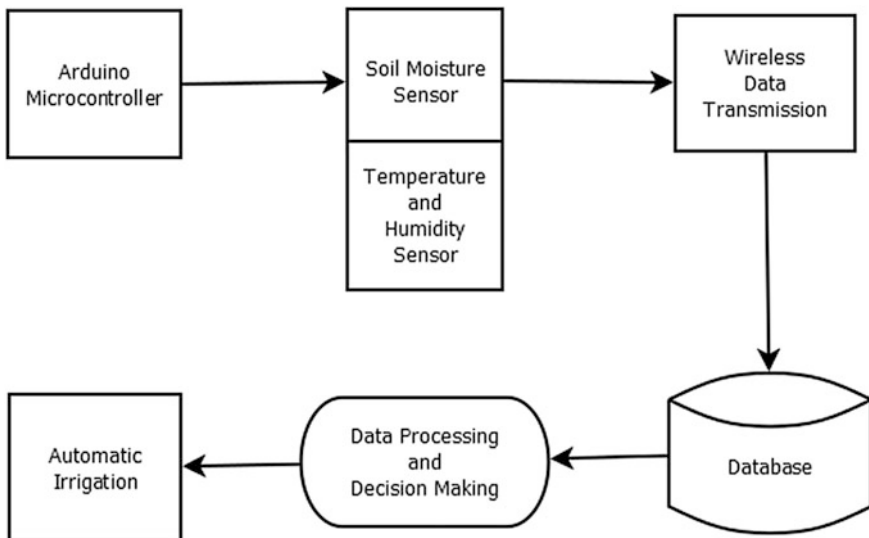
Fig. 1 Block diagram of the proposed system

sitting cozily in their homes but also the system would be able to automatically turn on the pump if put on auto mode.

3.1 Conceptual Design

In Fig. 2, the proposed decision system works as follows: when power is on, the microcontroller unit activates the peripherals for controlling, detecting, and establishing communication from the various connected devices as well as sensors to collect soil moisture, temperature, and humidity from the field. Further, the collected data is sent to the cloud server. The user can directly login from the PC or mobile. A user interface has a decision system in which there is a probability of where the irrigation is required or not. All the parameters are balanced back to the permissible levels. It can be all balanced by the sensor. Then, again and again, sensor measures the parameters repeatedly. After that, the process will end.

3.2 Hardware Description

Arduino modules play a vital role in the development of the wireless system in the agriculture for measuring and checking of the fundamental quality soil parameters. In Arduino, a wireless sensor module connects the soil moisture sensor and DHT11 sensor for capturing data such as temperature, humidity, and moisture of the soil. The wireless communication module connects the base station as well as mobile device/laptop for accepting, receiving, processing, and storing data [4–6]. Arduino board is outfitted with powerful ATmega microcontrollers and normally, fills in as the "mind" of the robot. To attach sensors with the board a shield board is used. The shield board extends the capabilities of Arduino. The Arduino board is kept over a shield board and according to the pins connected. The pins of the sensors are plugged into the shield board. Once a proper connection of the sensors with the pins of Arduino is established, the data acquisition process starts. To capture soil moisture data, a soil moisture sensor is plugged into the soil. In moist soil, resistance is less which has a good conductivity of current as compared with dry soil. Therefore, it helps to detect the moisture contained in the soil.

To get humidity and temperature of the soil, DHT11 temperature and humidity sensors are used in the present work. The aggregate sum of water vapor in the air is characterized as a measure of humidity. The humidity of the soil varies relatively as the temperature varies in the observation. The measure of water droplets in the air increases after irrigation and has effects in temperature drop which results in more humidity in the environment. So, an appropriate action could be initiated by the farmer after receiving regular interval of temperature and humidity readings from the crop field to increase the productivity of the crop.

Fig. 2 Flowchart of the IoT decision system
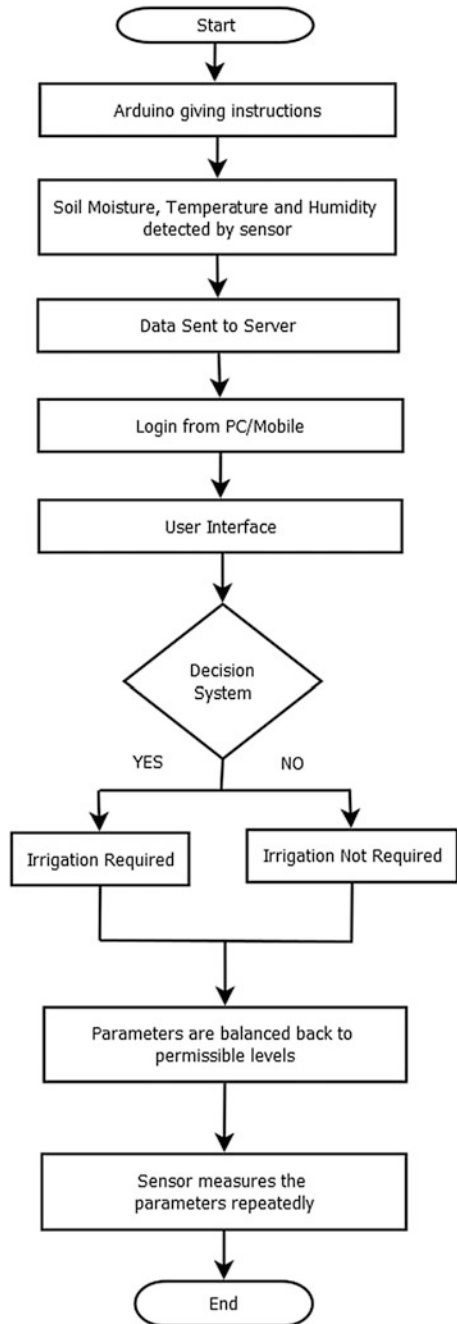
4 Implementation and Result Analysis

Figure 3 shows sensors deployed in the field and then the results are carried out and further analyzed in the form of temperature, moisture and humidity from the wheat crop field. A prediction algorithm (ANN) is used to analyze the performance, error, regression, and training state in the present work.

Figure 4 shows a plot of temperature received over 120 days at 09:00 AM, 02:00 PM, 06:00 PM, 09:00 PM, 12:00 Midnight, and 05:00 AM, respectively, on a

Fig. 3 Sensor deployed in the field
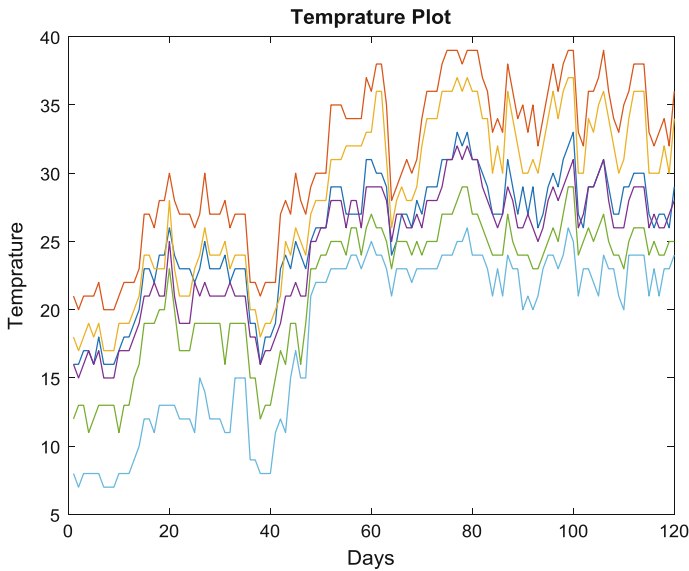
Fig. 4 Temperature monitoring of 120 days

regular interval basis. It is clear from the plot how the temperature varies. The temperate rises but remains more or less constant between 40 and 120 days

Figure 5 shows a plot of soil moisture data received over 120 days. The plot shows the variations in the moisture. Values falling below 10 can be considered as an alarm.

Figure 6 shows a plot of humidity of the site over 120 days. It is observed that as the temperature increases the humidity decreases alarmingly. Near 110 days the humidity falls down to around 35. An alert alarm could be sent to the farmer for watering the fields. It could also be possible as the system can be in auto mode in which the pump could start on its own. There are certain limitations of the sensors considered for testing and the same might be replaced in real-time applications with more robust sensors having Mil Grade standards. The current sensors may not work properly due to heating of the sensor and airflow. At some time the sensor touched 100% humidity which may be due to ground conditions. The proposed neural network is trained first on ideal conditions and then tested over the real-time data; the network eliminates the discrepancies in the data collected and provides ideal parameters for the region and climate.

In the present work temperature, moisture, humidity of 120 days are captured from the wheat crop field and analysis is carried out to predict the soil conditions to take necessary action by a farmer after generating an alarm. The workflow for the neural network design process has seven primary steps and they are collect data, create the network, configure the network, initialize the weights and biases, train the network, validate the network, use the network from the training window. It generates four plots such as performance, training state, error histogram, and
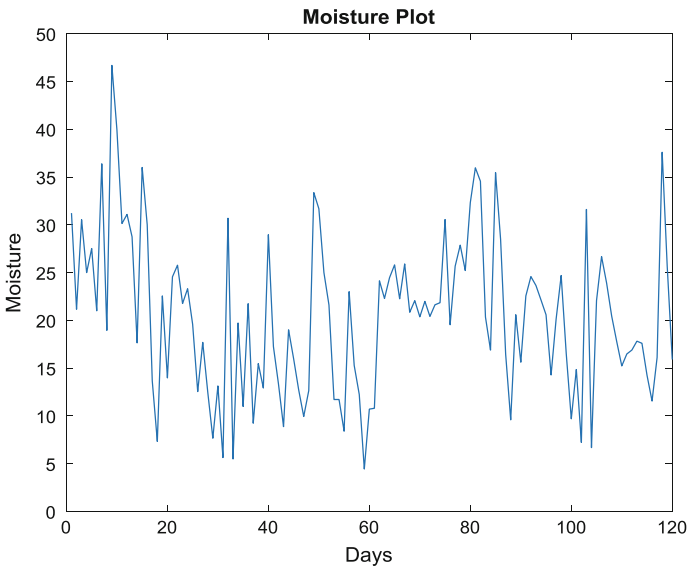
Fig. 5 Soil moisture monitoring for 120 days
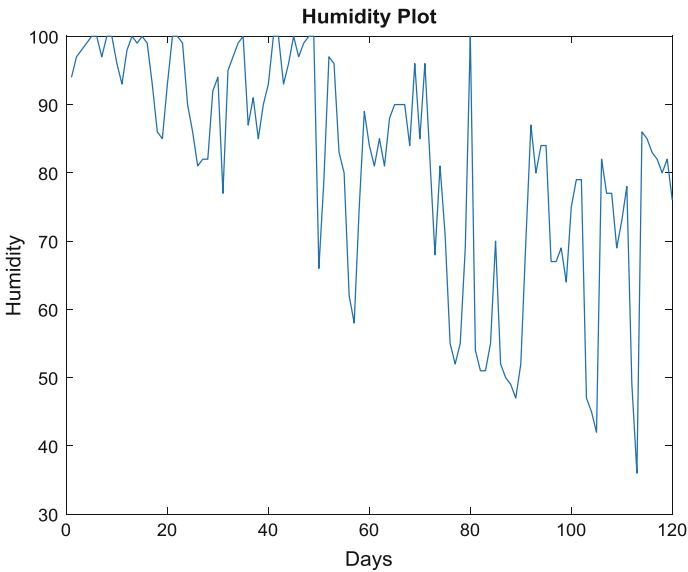
Fig. 6 Humidity monitoring for 120 days

regression. The performance plot demonstrates the estimation of the execution work versus the emphasis number. It plots preparing, approval, and test exhibitions. The training state plot demonstrates the advance of other training factors, for example, the gradient magnitude, the quantity of approval checks, and so forth. The error histogram plot demonstrates the distribution of the network mistakes. The regression plot demonstrates a relapse between network outputs and network targets. Further, the function fitting neural network and cascade forward neural network is applied. Function fitting is the way toward training a neural system on an arrangement of contributions to request to create a related arrangement of the target outputs. After that, a system is built with the coveted hidden layers and the training algorithm. Once the neural system has fit the information, it shapes a speculation of the information I/O relationship. The trained network then can be utilized to produce outputs for data sources. The function cascade forward net makes cascade forward networks. These are like feed forward network yet incorporate a weight association from the contribution to each layer, and from each layer to the progressive layers. For instance, a three-layer arrange has associations from layer 1 to layer 2, layer 2 to layer 3, and layer 1 to layer 3. The three-layer organization additionally has associations from the contribution to each of the three layers. The extra associations may enhance the speed at which the network takes the desired relationship. It is being trusted that a comprehension of the ideal models and their application is fundamental to the agreeable and effective utilization of this toolkit.

Figure 7 shows that the tr.best_epoch, the iteration at which the validation execution achieved a base information. The training proceeded for 6 more iterations
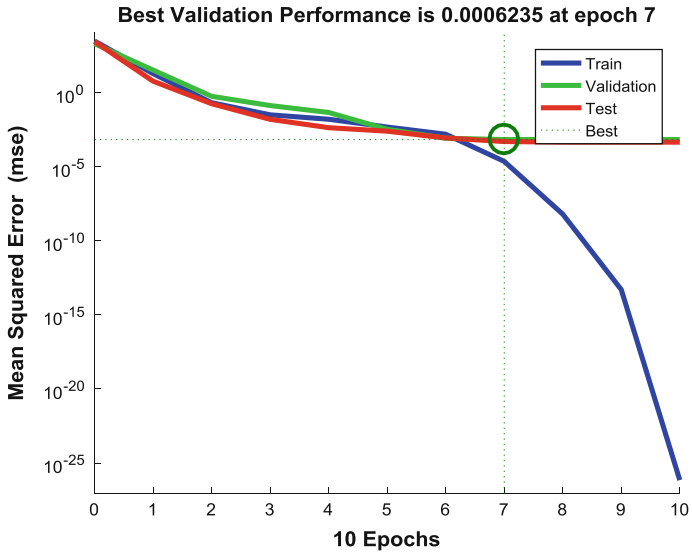
Fig. 7 Best validation performance
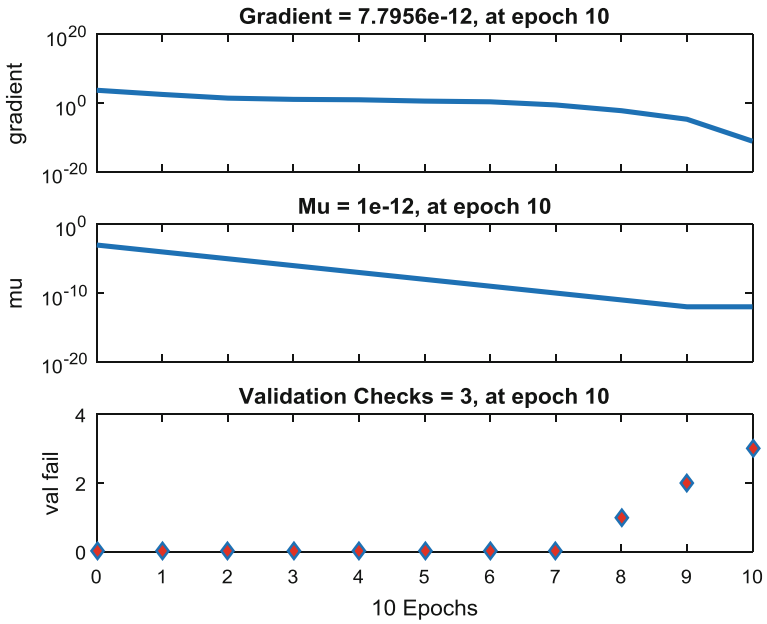
Fig. 8 Training state

before the training stopped. This figure does not show any major issues with the training. The validation and test bends are fundamentally the same. On the off chance that the test bend had expanded fundamentally before the approval bend expanded, at that point it is conceivable that some over-fitting may have happened.

Figure 8 shows that the subsequent stage in validating the network is to make a regression plot, which demonstrates the connection between the outputs of the network and the objectives. If that the training are best, the network outputs and the objectives would be precisely equivalent, however, the relationship is once in a while idealize in practice.

Figure 9 shows a normal condition where a close watch is required by the farmer as the data is not very close to 45°. A signal can be sent to the farmer to closely monitor the fields.
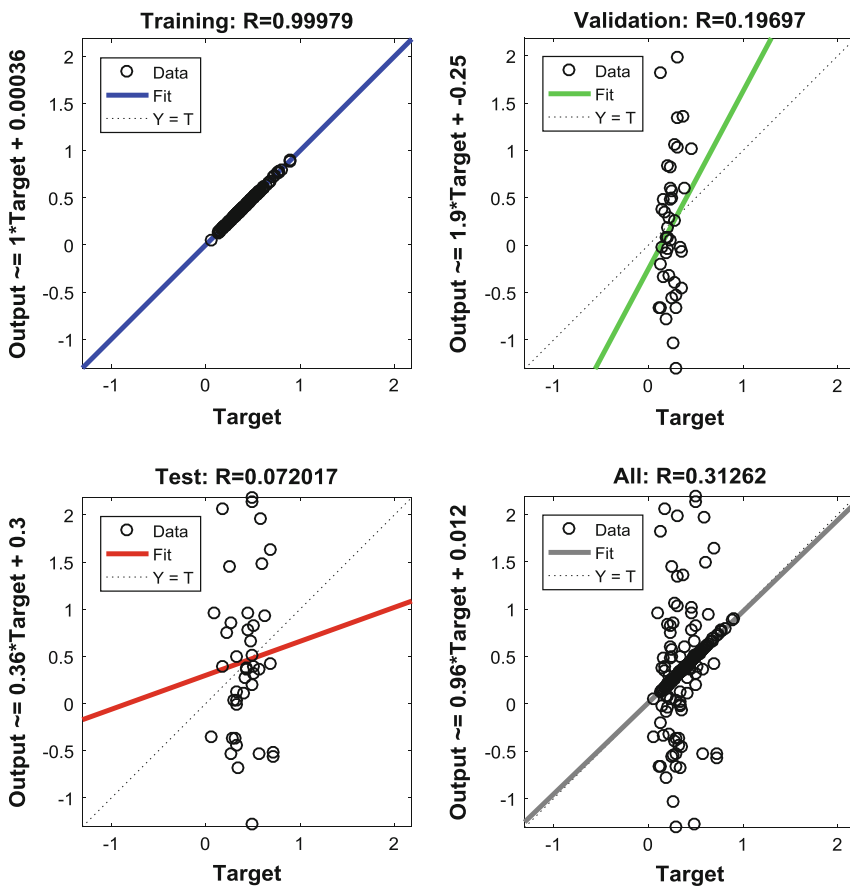
Fig. 9 Regression plot
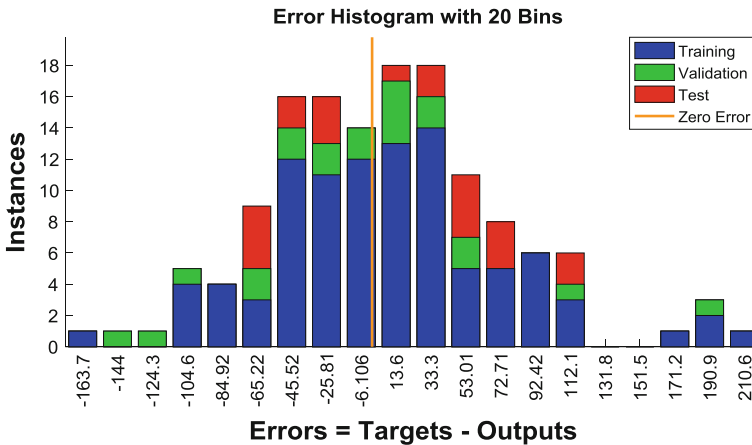
Fig. 10 Error histogram (color figure online)

The outcome appears in the accompanying Fig. 10. The axes speak to the training, validation and testing information. The blue bars represent the training data, the green bars represent validation data and the red bars represent the testing data. It is basically used in the verification of the network performance. In this case, it is observed that while most error falls between −65.22 and 112. There is a training plot which has training error 210.6, validation error −144 to −124.3 and validation point error 13 and 6 respectively.

5   Conclusion

The test carried out using the proposed system provides us with up-to-date information of the field. The trained dataset is created by collecting the data from the field of 120 days. The current dataset which is collected by the proposed sensors in the field is tested with the trained dataset and the resultant dataset gives us a fair idea about the current position of the filed as compared with the same, a year back during the same period. A timely action could be taken by the farmer for not only saving the crop but also to increase the productivity of the crop. The results received gives the near accurate position of the field to the farmer and also send a signal to the farmer for necessary action.

In future to get a more accurate result, a dataset of more than 5 years could be considered. The entire system can be integrated to build a solution to help the farmers where field and crop could be closely monitored and the application is connected with automatic sprinklers to maintain the water and moisture level of the field and the crop.

References

1. Rajalakshmi, P., Devi Mahalakshmi, S.: IoT based crop-field monitoring and irrigation automation
2. Georgieva, T., Paskova, N., Gaazi, B., Todorov, G., Daskalov, P.: Design of wireless sensor network for monitoring of soil quality parameters. In: 5th International Conference "Agriculture for Life, Life for Agriculture"
3. Chavan, C.H., Karande, P.V.: Wireless monitoring of soil moisture, temperature & humidity using Zigbee in agriculture. Int. J. Eng. Trends Technol. (IJETT) **11**(10) (2014)
4. Na, A., Isaac, W., Varshney, S., Khan, E.: An IoT based system for remote monitoring of soil characteristics. In: 2016 International Conference on Information Technology (InCITe)—The Next Generation IT Submit (2016)
5. Gondchawar, N., Kawitkar, R.S.: IoT based smart agriculture. Int. J. Adv. Res. Comput. Commun. Eng. **5**(6) (2016)
6. Lee, M., Hwang, J., Yoe, H.: Agricultural production system based on IoT. In: 2013 IEEE 16th International Conference on Computational Science and Engineering (2013)

Nature-Inspired Optimization Techniques in VANETs and FANETs: A Survey

Parampreet Kaur and Ashima Singh

Abstract In recent years, Vehicular Ad hoc Networks (VANETs) and Flying Ad hoc Networks (FANETs) are evolving rapidly. VANETs and FANETs are special types of Mobile Ad hoc Networks (MANETs). VANET uses vehicles as mobile nodes for the communication. VANET provide communication during emergency situations like accidents to avoid its possibility by sending alert messages to the drivers. FANET is a collection of unmanned aerial vehicles that communicate without any predefined infrastructure. FANET being the most searched and researched topic nowadays is finding its scope in flying objects like drones used for military applications such as border surveillance and for civil applications such as disaster management, traffic monitoring. In VANETs and FANETs, routing is challenging when Quality of Service (QoS) parameters needs to be satisfied. In this paper, VANETs and FANETs' routing protocols implementing optimization techniques (like Ant colony optimization, bee colony optimization, and particle swarm optimization) are surveyed. The differences between VANETs and FANETs are clarified first and then routing protocols for VANETs and FANETs are discussed.

Keywords VANET · FANET · Quality of service parameters
Routing protocols · Optimization techniques

1 Introduction

MANET is a basic category of a wireless ad hoc network. In MANET, mobile nodes are connected to form a self-configured and dynamic network without any predefined infrastructure. If the vehicles are used as nodes in MANET, it becomes the advanced version, i.e., VANET. In VANETs, the vehicles communicate with

P. Kaur (✉) · A. Singh
Thapar University, Patiala, India
e-mail: parampreetkaur93@gmail.com

A. Singh
e-mail: ashima@thapar.edu

other vehicles on road called as Vehicle-to-Vehicle (V2 V) communication or with the roadside units called as Vehicle to Infrastructure (V2I) communication [1]. The other advancement in ad hoc networks is FANET. FANET uses Unmanned Aerial Vehicles (UAV) as nodes and two-ray ground model or Nakagami-m model as radio propagation model for communication [2]. Wi-Fi radio interface using 802.11n and IEEE 802.11b radio equipment mounted on UAVs are suitable in FANETs' environment [2]. FANET gained popularity because of its application in military and in natural calamities like earthquake and flood-affected areas. In FANET, UAVs fly independently without carrying any human personnel [2]. In a wireless ad hoc network, the routing protocols are used to decide how the nodes will route the packets from source to destination. For analyzing the performance of routing protocols efficiently, it is necessary to consider QoS parameters. There is a need to apply optimization techniques on routing to uncover the challenge of satisfying QoS requirement. The effective optimization techniques reduce the delay, increase the packet delivery ratio, and consider other QoS parameters when working with routing protocols.

Nature-inspired optimization is a major area of research in a wireless ad hoc network. Various optimization techniques are applied on routing protocols in VANETs and FANETs. This paper presents a rigorous review of VANET and FANET routing protocols implementing nature-inspired optimization techniques.

1.1 Motivation for Research

Optimization techniques are applied to routing protocols to enhance the performance. This study focuses on optimization techniques applied to different routing protocols in ad hoc networks. We identified the need for survey on optimization techniques after viewing research in nature-inspired optimization techniques in VANETs and FANETs. Accordingly, we have summarized the research which is available on different techniques and routing protocols.

1.2 Difference Between VANET and FANET

Similar to VANET, nodes in FANET are mobile in nature and self-organized during communication. Having these characteristics, still, FANET possess some distinguishing characteristics and is considered as a particular class of MANET [2].

Node Mobility It makes a major difference between VANET and FANET as nodes are road vehicles and aerial vehicles respectively (Fig. 1). In FANET, the nodes are more mobile in nature than that of VANET.

Dynamic Topology Due to highly mobile nodes in FANET, the topology changes frequently and there is a disconnection between nodes communication. In VANET,
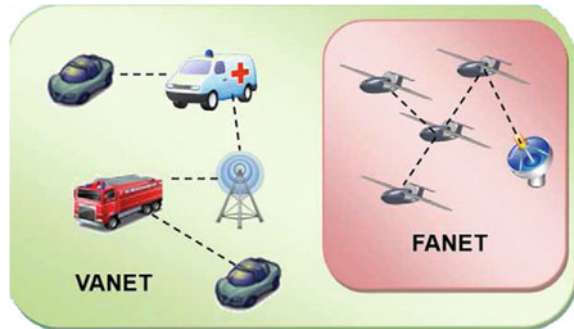
the network scenario depends on sparse and dense traffic. In sparse traffic, topology is more dynamic than in dense traffic.

Node Density It is described as the average count of nodes for some particular zone. In FANETs, the aerial vehicles are more dispersed than vehicles in VANET. This results in much low node density in FANET than in VANET.

Mobility Model In VANET, vehicles either move on a highway or a road. As a result, mobility models of VANETs are predicted easily. In FANET, if UAVs move in a predefined path, the mobility model is regular. In some applications of FANET because of the weather changes flight path is changed. Because of this various FANET mobility models such as random waypoint, Gauss–Markov, etc., are defined by Kumari et al. [3].

2 VANET

In this section, the various VANET routing protocols are discussed in detail and optimization techniques used in VANETs are defined. Further, the survey of existing VANET protocols proposed using optimization techniques is done. Table 1 describes the comparison of VANET routing protocols on the basis of optimization techniques and QoS parameters.

2.1 Routing Protocols in VANETs

In VANETs, there are following five types of protocols' categorization: topology routing, geo-cast routing, position routing, broadcast routing, and cluster routing [4].

Topology Routing It uses link information to forward a packet from the origin node to target. It is further divided into proactive (table-driven) and reactive

Table 1 Relative comparison of VANET routing protocols

VANET protocol	Routing approach	Optimization technique			QoS parameters								
		ACO	BCO	PSO	PDR	End-to-End delay	Routing overhead	Through put	Jitter	Energy consumption	Bandwidth consumption	Packet loss	Hop count
AODV, GSR, DSR [6]	Topology	–	–	–	✓	–	–	–	–	–	✓	–	✓
AODV, DSR, OLSR [7]	Topology	–	–	–	✓	✓	–	✓	–	–	–	–	–
DSR [9]	Reactive	✓	–	–	✓	✓	–	✓	✓	✓	–	–	–
AODV extension [10]	Reactive	✓	–	–	–	–	✓	–	–	–	–	–	–
QoS_OLSR [11]	Cluster	✓	–	–	✓	✓	✓	–	–	–	–	–	–
MAZACORNET [12]	Hybrid (proactive, reactive)	✓	–	–	✓	✓	–	–	–	–	–	–	–
DYMO [13]	Reactive	✓	–	–	–	✓	–	–	–	–	–	–	–
Bee-optimized Fuzzy GRP [15]	Geographic	–	✓	–	–	✓	–	✓	–	–	–	–	–
QoSBeeVanet [16]	Topology	–	✓	–	✓	✓	–	–	–	–	–	–	–
CBQoS_Vanet [17]	Cluster	–	✓	–	✓	✓	✓	–	–	–	–	–	–
HyBR [18]	Hybrid (topology, geographic)	–	✓	–	✓	✓	✓	–	–	–	–	–	–
GeoPSO [20]	Geo-cast	–	–	✓	✓	–	✓	–	–	–	–	–	–
PSO_DS [21]	Topology, position	–	–	✓	–	✓	✓	–	–	✓	–	✓	–
PSO-based HyBR [22]	Hybrid	–	–	✓	✓	✓	✓	–	–	–	–	–	–

(on-demand) routing. Temporally Ordered Routing Protocol (TORA), Ad hoc on-Demand Distance Vector (AODV) and Dynamic Source Routing (DSR), [5] are examples of reactive routing protocols and Fisheye State Routing (FSR) [5] and Global State Routing (GSR) [4] are examples of proactive protocols.

The topology routing protocols for VANETs are used for performance analysis by the following authors. Kumar and Rani [6] experimentally analyzed topology-based AODV, GSR, and DSR protocols in VANET using simulation tool ns-2. The author experiments on various simulation parameters: packet delivery rate, latency, bandwidth consumption, average number of hops. From the results, it is concluded that AODV, GSR, and DSR perform differently with distance and speed and no single routing protocol meet all traffic requirements. M. Ben et al. [7] experimentally analyzed the AODV, OLSR, and DSR in urban scenarios. These protocols are analyzed with simulation tool ns-2 using metrics like node density, node mobility. On varying node density and on varying vehicle speed it is experimentally proved that OLSR outperforms both AODV and DSR.

Geo-cast Routing Geo-cast is location-based routing and packets are delivered within a specific geographic area from the origin node to other nodes. In geo cast, directed flooding of packets is done to reduce congestion in a network which is created when packets are flooded to every node. Various Geo-cast protocols are ROVER (Robust Vehicular Routing), IVG (Inter-Vehicle Geo-cast), and DTSG (Dynamic Time-Stable Geo-cast Routing) [5].

Position-based Routing It does route discovery by using GPS which determines position information of all the nodes. In this routing, a routing table is not maintained as routing information is collected from GPS tool for routing. The various position-based protocols are GPSR (Greedy Perimeter Stateless Routing), GSR (Geographic Source Routing) [5].

Broadcast Routing It is flooding-based routing that forwards the information to a maximum number of nodes whenever an accident occurs. Broadcast routing is used in an emergency to deliver messages quickly and efficiently. Some broadcast routing protocols are Distributed Vehicular Broadcast Protocol (DV-CAST), Preferred Group Broadcast (PGB), Secure Ring Broadcasting (SRB), and BROAD-COM [5].

Cluster-based Routing It depends on the cluster and position for routing. In cluster-based routing, there is a unique cluster-head, responsible for communication between nodes of a cluster and between nodes of different clusters. The various cluster protocols are Cluster-Based Directional Routing Protocol (CBDRP), Cluster-Based Location Routing (CBLR) [4].

It is challenging to satisfy all the QoS requirements like low delay, high data packet delivery ratio, low energy consumption, and low routing overhead. To overcome this problem, various nature-inspired optimization techniques (ACO, PSO, and BCO) are applied on protocols for routing.

2.2 Ant Colony Optimization

This technique is motivated by colonies of ants to give approximate optimization. The communication within the ants is possible because of pheromone trails that facilitate the ant to search the shortest path from the nest to the food position [8].

The ant colony optimization on routing protocols of VANETs is applied by the following authors. A. R. Deshmukh and S. S. Dorle [9] proposed a swarm-based optimization to upgrade the performance of DSR in VANET. Ant colony optimization is applied on DSR routing protocol and performance analysis is done using different QoS parameters like delay, throughput, jitter, energy, packet delivery ratio. The results obtained using ns-2 shows ant colony optimization improved various QoS parameters for VANETs. R. K. Chauhan and A. Dahiya [10] have done AODV enhancement by ant colony to avoid the communication delay which appears because of the usual breakage of connection in VANETs network topology. This proposed algorithm has two events—route establishment and route repair. AODV combination with ACO increases the performance by reducing the routing overhead and avoiding the frequent path loss. Wahab et al. [11] proposed QoS-OLSR protocol to solve the clustering problem in VANET. The protocol maintains the stability of the vehicular network during communication and achieves the quality of service requirements. The proposed protocol is composed of three components: QoS-based clustering using ant colony optimization, Multipoint Relay (MPR) recovery algorithm to deal with link failures and cheating prevention mechanism to guarantee fair MPRs selection procedure. The simulation results show that QoS-OLSR protocol maintains the vehicular network stability, lessens the delay, increase the PDR, and reduce the communication overhead. Rana et al. [12] developed a hybrid, ant colony-based routing algorithm. In this algorithm, vehicles are divided into zones. Proactive routing is applied to search a route in a zone and reactive routing search route among zones. This developed algorithm is more effective for delay and packet delivery ratio when the comparison is done with other existing VANET algorithms and is suitable for a dense network. A. M. Oranj et al. [13] developed a new routing algorithm which is formed on basis of DYMO (Dynamic MANET On-demand) protocol and ant colony optimization. The implementation is done using ns-2 and performance is evaluated using delay time and path reliability parameters. The experimental results show that in proposed routing algorithm path searching accuracy is increased and the delay time is decreased when compared with AODV routing protocol.

2.3 Bee Colony Optimization

It is a random search strategy. It is having a random probability distribution called as stochastic technique. BCO technique finds the optimum solution for

combinatorial optimization problems using the way in which bees look for food in nature [14].

The analysis of routing protocols of VANETs based on bee colony optimization is done by the following authors. Saravanan and Arunkumar [15] proposed Bee-optimized Fuzzy Geographical Routing Protocol (GRP). The geographical routing protocol is enhanced by applying fuzzy controller and rules optimization is done with BSO. BSO algorithm, find a possible solution based on food source position and solution fitness is obtained using nectar value of food source. The simulation results show that proposed GRP's throughput increased, retransmission increased, end-to-end delay decreased as the comparison is with GRP and fuzzy-based GRP. Bitam and Mellouk [16] proposed QoSBeeVanet. It is a quality of service-based multipath routing protocol in VANETs. From the simulation results obtained in ns-2 simulator, it is concluded that QoSBeeVanet gives better performance as compared to DSDV and AODV routing protocols for delay and packet delivery ratio metrics. Fekair et al. [17] proposed CBQoS-Vanet. It is a cluster-based artificial bee colony optimized algorithm for the quality of service routing protocol in VANET. The protocol is dependent on the mentioned techniques: clustering algorithm to improve the transfer of the routing data and artificial bee colony algorithm to find a perfect route from the origin node to target node. The simulation outcomes proved that proposed protocol improves the PDR, the network overhead, and delay. Bitam et al. [18] proposed hybrid bee swarm routing. It is formed by the combination of topology and geographic routing. The authors evaluate the performance on basis of packet delivery ratio, delay, and network load. The experimental outcomes proved that proposed protocol outperforms when compared to GPSR, AODV routing protocol.

2.4 Particle Swarm Optimization

This strategy is based on the population and created by Kennedy and Eberhart. This technique is based on the social behavior of schooling fish and flocking birds while searching for food [19].

The particle swarm optimization is applied on routing protocols of VANETs by the following authors. Kaiwartya and Kumar [20] proposed GeoPSO protocol, developed by applying PSO on geo-cast routing. GeoPSO protocol is simulated using ns-2 to experiment on network load and packet delivery ratio parameters. The experimental outcomes proved that GeoPSO protocol gives better performance compared to P-GEDIR protocol. K.D. Kalambe et al. [21] proposed the combination of DREAM + SIFT (DS) protocol and also PSO strategy is applied on the proposed DS and DSDV protocol. The author experiments on parameters like energy, delay, packet loss, control overhead, and network load using ns-2. The experimental results proved PSO_DS is more effective when the comparison is done with DS, DSDV, and PSO_DSDV. Kaur and Kahlon [22] proposed PSO-based hybrid routing protocol for VANETs. The new protocol is formed by

the combination of particle swarm and bee swarm protocol. The authors did the simulation on ns-2 simulator using parameters delay and packet delivery ratio. From the experiments, it is proved that proposed routing protocol shows better performance when compared with AODV.

3   FANET

In this section, the FANET routing protocols and various optimization techniques used in FANETs are discussed. Also, the survey of FANET protocols proposed using optimization techniques is done. Table 2 describes the comparison of FANET routing protocols on the basis of optimization techniques and QoS parameters.

3.1   Routing Protocols in FANET

The five major categories of FANET protocols are Proactive protocols, Reactive protocols, Hybrid protocols, Position/Geographic-based protocols, Hierarchical protocols [23].

Proactive Protocols In proactive routing, tables are used to store latest routing details of all the nodes in the network. In this routing, the shortest direction is most likely to get because each node is having connectivity graph. Some proactive protocols are FSR, OLSR [23].

Reactive Protocols It is also referred to as on-demand routing protocol which generates routes between nodes only if required. Reactive routing uses RouteRequest and RouteReply messages for communication between nodes. Examples of reactive protocols are AODV, DSR.

The reactive protocols for FANETs are used for performance analysis by the following authors. Singh and Verma [24] experimentally analyzed the FANET routing protocols to enhance the performance of ad hoc networks. The authors analyzed on AODV, DSDV, and OLSR routing protocols. The simulation is done using ns-2 for FANET environment. The authors used packet delivery ratio, delay, and throughput parameters for analyzing protocols performances. From the simulation, it is clear that OLSR gives improved results as compared to AODV and DSDV. Rosati et al. [25] proposed OLSR extension using Global Positioning System (GPS), i.e., predictive OLSR (P-OLSR) for FANETs. The simulation outcomes show that P-OLSR performs better than OLSR when network topology changes. Singh and Verma [26] applied OLSR routing protocol under various mobility models in FANETs using parameters delay, throughput, and packet delivery. From the simulation results, it is concluded that under pursue mobility model OLSR performs better in comparison with Manhattan, RPGM, and random waypoint mobility model.

Table 2 Relative comparison of FANET routing protocols

FANET routing protocol	Routing approach	Mobility model used	Optimization technique			QoS parameters					Simulation results and analysis
			ACO	BCO	PSO	PDR	Delay	Routing load	Throughput	Hop count	
AODV, DSDV, OLSR [24]	Reactive, Proactive	–	–	–	–	✓	✓	–	✓	–	OLSR gives improved results compared to AODV and DSDV
OLSR [26]	Proactive	RPGM, Manhattan, Random Waypoint, Pursue	–	–	–	✓	✓	–	✓	–	OLSR perform better under pursue mobility model as compared to other models
AODV, OLSR, HWMP [27]	Proactive, Reactive, Hybrid	Gauss–Markov	–	–	–	✓	–	✓	–	✓	HWMP routing protocol gives high PDR
Hybrid routing scheme [28]	Unicast, Geo-cast	–	–	–	–	✓	–	–	–	✓	Performance of FANETs increased by proposed hybrid routing scheme
MPGR [29]	Geographic	Gauss–Markov	–	–	–	✓	✓	–	–	–	MPGR gives better performance compared to AODV and GPSR
AntHocNet, AODV, DSDV, DSR [23]	Topology	Random Waypoint	✓	–	–	✓	✓	✓	–	–	AntHocNet performs better than AODV and DSDV protocol
APAR [30]	Reactive	–	✓	–	–	✓	✓	✓	–	–	APAR gives improved PDR, delay, and routing load
BeeAdHoc, AODV, DSDV, DSR [31]	Topology	Random waypoint	–	✓	–	–	✓	✓	✓	–	BeeAdhoc is more effective than AODV, DSDV, and DSR
AntHocNet, BeeAdhoc, AODV, DSDV, DSR [32]	Topology	Random waypoint	✓	✓	–	–	✓	✓	✓	–	AntHocNet and BeeAdHoc gives better results compared to AODV, DSDV, and DSR

Hybrid Protocols It is a protocol formed by combining both proactive routing and reactive routing protocol. It overcomes the limitation of extra time needed to discover route in reactive protocol and huge overhead of control messages in a proactive protocol. Zone Routing Protocol (ZRP) and Hybrid Wireless Mesh Protocol (HWMP) [23] are some examples of hybrid protocols.

The hybrid protocols for FANETs are used by the following authors for performance analysis. Vasiliev et al. [27] compare AODV, OLSR, and HWMP (Hybrid Wireless Mesh Protocol) routing protocols in FANETs. The author experiments on quality of service metrics, i.e., hop count, packet delivery ratio, overheads using ns-3 simulation tool. The Gauss–Markov mobility model used in simulation results that HWMP routing protocol gives high packet delivery ratio. Gankhuyag et al. [28] developed a hybrid scheme having features of unicast and geo-cast routing. The simulation of proposed scheme results in the increase of performance of FANETs when compared with conventional AODV scheme.

Hierarchical Protocols In this category hierarchy model is used in routing. In this protocol, the proactive and reactive routing is chosen on the hierarchy-level basis. In this scheme, first, some proactive routes are used and then reactive protocols are used at a lower level. Some examples of hierarchical routing are Landmark Routing Protocol (LANMAR), Distance Routing Effect Algorithm for Mobility (DREAM).

Position/Geographic-based Protocols It uses real location information of the nodes for routing problems. Each node calculates the position through GPS or by any other positioning facility. Some examples of geographic-based protocols are GPSR, Geographic Position Mobility Oriented Routing (GPMOR). Lin et al. [29] proposed a geographical mobility prediction routing protocol, MPGR in ad hoc UAVs networking. The simulation is based on Gauss mobility model and the outcome shows that MPGR gives better performance when the comparison is done with AODV and GPSR for delay and packet delivery ratio.

3.2 Ant Colony Optimization

The ant colony optimization on routing protocols of FANETs is applied by the following authors. Maistrenko et al. [23] used the ant colony optimization technique to solve the routing problems for FANETs. The comparison of AntHocNet protocol with AODV, DSDV, DSR protocols is done using network simulator ns-2. The performance is analyzed using parameters such as packet delivery ratio, average delay, and overheads. The simulation outcomes show that AntHocNet protocol gives better performance when compared to AODV and DSDV protocol for packet delivery and delay. Yu et al. [30] proposed APAR algorithm, polymorphism-aware routing algorithm. APAR algorithm combines ant colony optimization and DSR algorithm. In results, it is shown that APAR performs better compared to traditional algorithms for packet delivery ratio, delay, and load.

3.3 Bee Colony Optimization

The analysis of routing protocols of FANETs based on bee colony optimization is done by the following authors. Leonov [31] explains the bee algorithm and also the routing process in the ad hoc network based on the bee algorithm. The author compared AODV, DSDV, and DSR protocols with BeeAdHoc protocol using delay, throughput, and routing load metrics in ns-2 simulator. From the experimental results, it is concluded that BeeAdHoc is more effective than AODV, DSDV, and DSR. Leonov [32] performed experimental analysis on AntHocNet and BeeAdHoc protocol for solving the routing problems in FANET. The experiments are done using ns 2. The simulation study shows that AntHocNet and BeeAdHoc algorithm give more effective results when compared with AODV, DSDV, and DSR.

4 Conclusion

Optimization techniques such as ACO, BCO, and PSO are used in VANETS and FANETs to improve various "Quality of Service" parameters like delay, packet delivery ratio, jitter, energy consumption, routing load, throughput, bandwidth consumption, and packet loss. In this paper, VANET and FANET routing protocols are defined and also a comprehensive review of optimization techniques is done. From the survey, it is concluded that in FANETs a few routing protocols are proposed using optimization techniques as compared to VANETs.

References

1. Dua, A., Kumar, N., Bawa, S.: A systematic review on routing protocols for vehicular ad hoc networks. Veh. Commun. 1, 33–52 (2014). https://doi.org/10.1016/j.vehcom.2014.01.001
2. Bekmezci, I., Sahingoz, O.K., Temel, Ş.: Flying ad-hoc networks (FANETs): a survey. Ad Hoc Netw. 11, 1254–1270 (2013). https://doi.org/10.1016/j.adhoc.2012.12.004
3. Kumari, K., Sah, B., Maakar, S.: A survey: different mobility model for FANET. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 5, 1170–1173 (2015)
4. Patel, D., Faisal, M., Batavia, P., Makhija, S., Mani, M.: Overview of routing protocols in VANET. Int. J. Comput. Appl. 136, 4–7 (2016). https://doi.org/10.5120/ijca2016908555
5. Kaur, A., Arora, D.: Survey over VANET routing protocols for vehicle communication. Int. J. Electr. Electron. Eng. 1, 1–6 (2014)
6. Kumar, S., Rani, S.: A study and performance analysis of AODV, DSR and GSR routing protocols in VANET. Int. J. Comput. Appl. 96, 48–52 (2014). https://doi.org/10.5120/16826-6586
7. Ahmed, M.B., Boudhir, A.A., Bouhorma, M., Ahmed, K.B.: Performance study of various routing protocols in VANET case of study. Int. J. Futur. Gener. Commun. Netw. 7, 231–240 (2014). https://doi.org/10.14257/ijfgcn.2014.7.6.21

8. Blum, C.: Ant colony optimization: introduction and recent trends. Phys. Life Rev. **2**, 353–373 (2005). https://doi.org/10.1016/j.plrev.2005.10.001

9. Deshmukh, A.R., Dorle, S.S.: Performance improvement of dynamic source routing (DSR) protocol using ant colony optimization for vehicular ad-hoc network (VANET). IJSR—Int. J. Sci. Res. **5**, 171–173 (2016)

10. Chauhan, R.K., Dahiya, A.: AODV extension using ant colony optimization for scalable routing in VANETs. J. Emerg. Trends Comput. Inf. Sci. **3**, 241–244 (2012)

11. Wahab, O.A., Otrok, H., Mourad, A.: VANET QoS-OLSR: QoS-based clustering protocol for vehicular ad hoc networks. Comput. Commun. **36**, 1422–1435 (2013). https://doi.org/10.1016/j.comcom.2013.07.003

12. Rana, H., Thulasiraman, P., Thulasiram, R.K.: MAZACORNET: mobility aware zone based ant colony optimization routing for VANET. In: 2013 IEEE Congress on Evolutionary Computation, Cancun, Mexico, pp. 2948–2955 (2013). https://doi.org/10.1109/cec.2013.6557928

13. Oranj, A.M., Alguliev, R.M., Yusifov, F., Jamali, S.: Routing algorithm for vehicular ad hoc network based on dynamic ant colony optimization. Int. J. Electron. Electr. Eng. (2016). https://doi.org/10.18178/ijeee.4.1.79-83

14. Nikolić, M.C.A., Teodorović, D.: Empirical study of the bee colony optimization (BCO) algorithm. Expert Syst. Appl. **40**, 4609–4620 (2013). https://doi.org/10.1016/j.eswa.2013.01.063

15. Saravanan, P., Arunkumar, T.: Bee optimized fuzzy geographical routing protocol for VANET. Int. J. Comput. Electr. Autom. Control Inf. Eng. **8**, 2222–2228 (2014)

16. Bitam, S., Mellouk, A.: QoS swarm bee routing protocol for vehicular ad hoc networks. In: 2011 IEEE International Conference on Communications (ICC), Kyoto, Japan, pp. 1–5 (2011). https://doi.org/10.1109/icc.2011.5963424

17. Fekair, M.E.A., Lakas, A., Korichi, A.: CBQoS-Vanet: cluster-based artificial bee colony algorithm for QoS routing protocol in VANET. In: 2016 International Conference on Selected Topics in Mobile & Wireless Networking (MoWNeT), Cairo, Egypt, pp. 1–8 (2016). https://doi.org/10.1109/mownet.2016.7496597

18. Bitam, S., Mellouk, A., Zeadally, S.: HyBR: a hybrid bio-inspired bee swarm routing protocol for safety applications in vehicular ad hoc networks (VANETs). J. Syst. Archit. **59**, 953–967 (2013). https://doi.org/10.1016/j.sysarc.2013.04.004

19. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: IEEE International Conference on Neural Networks. Perth, WA, Australia, pp. 1942–1948 (1995)

20. Kaiwartya, O., Kumar, S.: Geocasting in vehicular adhoc networks using particle swarm optimization. In: International Conference on Information Systems and Design of Communication. Lisbon, Portugal, pp. 62–66 (2014). https://doi.org/10.1145/2618168.2618178

21. Kalambe, K.D., Deshmukh, A.R., Dorle, S.S.: Particle swarm optimization based routing protocol for vehicular ad hoc network. Int. J. Eng. Res. Gen. Sci. **3**, 1375–1382 (2015)

22. Kaur, A., Kahlon, N.K.: Swarm based enhanced hybrid routing protocol in VANETs. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **5**, 310–316 (2015)

23. Maistrenko, V.A., Alexey, L.V., Danil, V.A.: Experimental estimate of using the ant colony optimization algorithm to solve the routing problem in FANET. In: 2016 International Siberian Conference on Control and Communications (SIBCON), Moscow, Russia, pp. 1–10 (2016). https://doi.org/10.1109/sibcon.2016.7491805

24. Singh, K., Verma, A.K.: Experimental analysis of AODV, DSDV and OLSR routing protocol for flying adhoc networks (FANETs). In: 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, pp. 1–4 (2015). https://doi.org/10.1109/icecct.2015.7226085

25. Rosati, S., Kruzelecki, K., Heitz, G., Floreano, D., Rimoldi, B.: Dynamic routing for flying ad hoc networks. IEEE Trans. Veh. Technol. **65**, 1690–1700 (2016)

26. Singh, K., Verma, A.K.: Applying OLSR routing in FANETs. In: 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies,

Ramanathapuram, India, pp. 1212–1215 (2014). https://doi.org/10.1109/icaccct.2014.7019290

27. Vasiliev, D.S., Meitis, D.S., Abilov, A.: Simulation-based comparison of AODV, OLSR and HWMP protocols for flying ad hoc networks. Lecture Notes in Computer Science Internet of Things, Smart Spaces, and Next Generation Networks and Systems 245–252 (2014). https://doi.org/10.1007/978-3-319-10353-2_21

28. Gankhuyag, G., Shrestha, A.P., Yoo, S.-J.: Robust and reliable predictive routing strategy for flying ad-hoc networks. IEEE Access **5**, 643–654 (2017). https://doi.org/10.1109/access.2017.2647817

29. Lin, L., Sun, Q., Li, J., Yang, F.: A novel geographic position mobility oriented routing strategy for UAVs. J. Comput. Inf. Syst. 709–716 (2012)

30. Yu, Y., Ru, L., Chi, W., Liu, Y., Yu, Q., Fang, K.: Ant colony optimization based polymorphism-aware routing algorithm for ad hoc UAV network. Multimed. Tools Appl. **75**, 14451–14476 (2016). https://doi.org/10.1007/s11042-015-3240-y

31. Leonov, A.V.: Application of bee colony algorithm for FANET routing. In: 2016 17th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM), Erlagol, Russia, pp. 124–132 (2016). https://doi.org/10.1109/edm.2016.7538709

32. Leonov, A.: Modeling of bio-inspired algorithms AntHocNet and BeeAdHoc for flying ad hoc networks (FANETS). In: 2016 13th International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE), Novosibirsk, Russia, pp. 90–99 (2016). https://doi.org/10.1109/apeie.2016.7806891

Implementation of Cloud-Assisted Secure Data Transmission in WBAN for Healthcare Monitoring

Sohail Saif, Rajni Gupta and Suparna Biswas

Abstract This work presents the cloud-assisted secure WBAN for healthcare application. There are various security issues associated with WBAN, which need to be solved to provide a secure real-time health monitoring system. Through this implementation, the patient's vital signals can be accessed in a secure manner in real time remotely by sensors and networks without visiting doctor's clinic or hospital. Here, we provide the cost-effective solution for the transmission of the patient's health data to doctor with proper confidentiality, authenticity, freshness, and security using cloud computing. In this work, the biosignals of patients and doctors are used to provide authenticity and vital signals are encrypted by using Advanced Encryption Standard (AES) for the secure m-health application. We have experimentally analyzed the average end-to-end delay for secure healthcare application is 14.59 and 19.31 ms in off-peak hours and peak hours respectively. This delay is only 5.84% in off-peak hours and 7.72% in peak hours of permissible delay of 250 ms for medical application.

Keywords WBAN · Cloud computing · Security · AES

S. Saif (✉) · R. Gupta · S. Biswas
Department of Computer Science and Engineering, Maulana Abul Kalam
Azad University of Technology, Salt Lake Sector-1, BF-142, Kolkata
700064, West Bengal, India
e-mail: sohailsaif7@gmail.com

R. Gupta
e-mail: rajniec75@gmail.com

S. Biswas
e-mail: mailtosuparna@gmail.com

1 Introduction

With the increase in the average age of the person, the healthcare costs augments. This critically affects the quality of life. It demands approachable, affordable, reliable, and secure solutions for healthcare systems [9]. There are a number of diseases like heart attack, asthma, diabetes, etc., which are the basic causes of sudden death of a number of people as they are not diagnosed and treated at the right time [7]. Wearable and implantable sensors are placed in or around the human body for health monitoring and analysis leads to the development of Wireless Body Area Networks (WBANs) used for purposes such as monitoring traffic, crops, infrastructure, and health.

For the healthcare application, numerous biological sensors monitor the vital signals and send the information to a low- energy constraint and high-capacity nodes. This node sends an information of a person or patient to the caretaker or a medical practitioner to ensure the timely medical diagnosis. In contrast to the wired network, the wireless communication provides more mobility to the users [8]. The purpose of WBANs is to provide high-speed, reliable, and secure Anytime, Anywhere, Always (AAA) communication. Besides that, it also provides the cost-effective solution for monitoring the patients. To satisfy the Quality of Service (QoS) for healthcare application there are various challenges associated with WBAN in terms of scalable, affordable, reliable, and secure network and through this work we have satisfied these requirements.

During the post-medical treatment, if the patient is unable to visit doctor's clinic or health centers or if the clinic is far away from his/her location then we need some remote access of patient's vital information instead of physical access. The cloud helps us to provide this feature, by this, the doctor can track the patient's vital signals remotely in a real-time manner [15]. Besides this, the patient's vital information can be stored in the server for a period of time. Hence, the doctor can examine the history of a patient remotely, and diagnose the diseases and provides necessary advice. In this work, the patient's vital signals are encrypted using Advanced Encryption Standard (AES), which provides the confidentiality of the data [2]. The biosignals of the patients and doctors are used for the authenticity [4, 14]. Hence, by incorporating the security module in the cloud, it became the secure, cost-effective, and easily available infrastructure for remote access of patient health data.

The WBAN is basically the three-tier communication architecture shown in Fig. 1. It consists of intra-BAN communication, inter-BAN communications, and beyond BAN communications. Beyond BAN communication concatenates the personal device to the Internet, here we are using the cloud server with our implemented security module for secure communication between the patient and the doctor as shown in Fig. 2. The various attacks [5] associated with data transmission are passive information gathering, subversion of node, false node and malicious data, man-in-the-middle attack, spoofing, Sybil attack, sinkhole attack, etc.

In this work, the man-in-the-middle attack, passive information gathering, and sinkhole attacks are targeted. Through this work, secure module incorporating

Fig. 1 Architecture of WBAN [6]
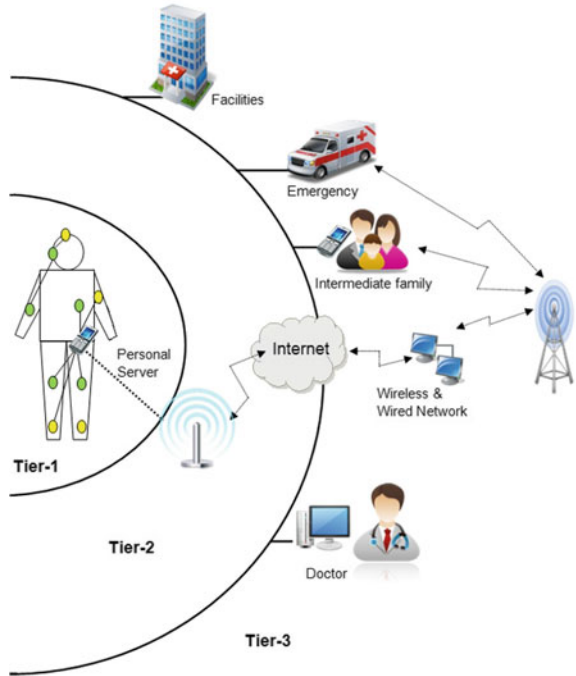
Fig. 2 Cloud-assisted secure WBAN at tier-3
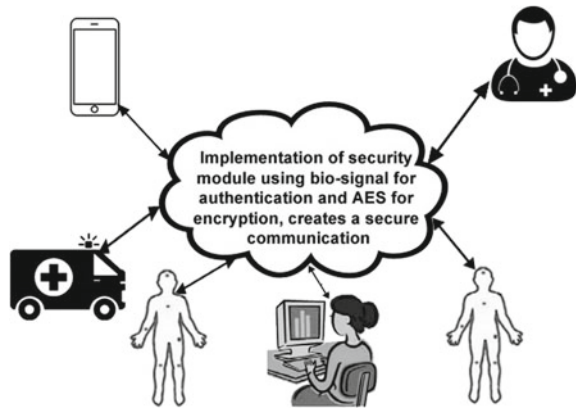
encryption technique [1] with the biosignals implemented in the cloud environment and end-to-end delay is analyzed in the real time. Section 2 gives the brief overview of the related work and Sect. 3 discusses the problem identification. Section 4 describes the implemented solution with the algorithm and Sect. 5 presents the experimental setup. Section 6 analyzes the results and finally Sect. 7, concludes the work.

Table 1 Comparative literature review

Authors, Year	Security aspects				Implementation	Delay
	Confidentiality	Authentication	Integrity	Freshness		
Zhou et al. [15], 2015	✓	✓	×	×	×	×
Liu et al. [3], 2014	✓	✓	×	×	✓	✓
Zheng et al. [14], 2017	✓	✓	×	×	×	✓
Hong et al. [4], 2016	✓	✓	×	×	✓	✓
Raghav et al. [11], 2012	✓	✓	×	✓	×	×
Implemented	✓	✓	✓	✓	✓	✓

2 Related Work

The vital signal information of the human body transmitted using WBAN is critical in nature. If any intruder modifies these data, then it causes adverse effect on the patient's life. Therefore, the security is the prime concern for the health signal transmission. The researches proposed various cryptographic scheme used for the encryption of the vital signals to provide the confidentiality to the data [11]. Some articles present the use of biosignal for authentication with confidentiality [4, 12, 13]. Few of the works which focused on the various security aspects for WBAN application are briefly discussed in Table 1.

3 Problem Address

The tier-3 communication of WBAN incorporated Internet, healthcare service provider, and users for monitoring the health signals. Due to open wireless network communication, they are vulnerable to security attacks. So the various security aspects needed for secure WBANs communication for tier-3 are as follows:

- Confidentiality of vital signal is required to overcome the disclosing or an unauthorized access.
- Integrity of health data is needed to validate the data received from the actual data sender, and if any data may change in the middle of communication, it can be havoc to the patient's life.
- Freshness of vital signal is required to maintain the data sent to the doctor should be with current vital signals.

- Availability is needed to ensure that the patient's information is accessible to the doctor within the permissible time limit of 250 ms for medical application [6].
- The security solutions should be usable and easy to implement.

4 Proposed Algorithm and Implementation of Solution

To resolve the problem addressed above, we have implemented the security module in the cloud which establishes the secure communication between the patient and the caretaker or the doctor. The Algorithm 1 is for secure communication:

Algorithm 1: Algorithm of cloud-assisted WBAN

1: BEGIN:
2: Procedure: Vital signals encryption and transmission
3: **if** *doctor request = true* **then**
4:   **if** *doct_biosignal_pat = doct_biosignal_send* **then**
5:     **for** $i = 1$ to n **do**
6:       $a[i] = $ *vital signals of the patient's measured*
7:       $enc[i] = $ *encryption of the* $a[i]$ *using AES*
8:     **end for**
9:     *Send the* $enc[i]$ *and*
10:    *Send the pat_biosignal_send at the doctor send*
11:   **else**
12:     *Request decline*
13:   **end if**
14: **else**
15:   *Sends biosignals after regular interval*
16: **end if**
17: Procedure: Decryption of vital signals
18: **if** *pat_biosignal_send = pat_biosignal* **then**
19:   **for** $i = 1$ to n **do**
20:     $dec[i] = enc[i]$ *decryption using secret key*
21:   **end for**
22: **else**
23:   *The patient biosignal is not matched*
24: **end if**
25: END

The details of the symbols used in this algorithm are as follows:

1. doct_biosignal_pat: Biosignal of doctor stored at the cloud.
2. doct_biosignal_send: Biosignal of doctor sent during the request.
3. pat_biosignal: Biosignal of patient stored at the cloud.
4. pat_biosignal_send: Biosignal of patient sent during the transmission.
5. n: Number of vital signals.

- In the first level, password is used for the validation.
- Then, the patients and the doctors biosignal are used for the authentication.
- After authentication, in the third level the vital signals of the patients are encrypted using AES with 128-bit block and 128-bit symmetric keys to provide the confidentiality to the vital signals [10].
- The current vital signals are sent and stored in the server for analysis by the doctor to incorporate the freshness of the data.

5   Experimental Setup Using Cloud Server

A cloud server is a virtual private server, whose working principle and functionality are same as a normal server but it is physically located in another space. It is hosted and delivered through primarily an Infrastructure as a service (IAAS)-based cloud model. The cloud server is accessed remotely through cloud service provider. These servers can be purchased/rented for a particular of time depending on the capacity, speed, etc. The WBAN tier-3 is designed using PHP with MySql. The interface designing and functionality of the experiment is provided by PHP, JavaScript, AJAX, etc. The cloud hosting service is provided by BigRock Cloud service provider. When the wireless networks demand is less, then it is considered as off-peak hours which is normally at night and when the traffic is high in the network then it is considered as peak hour or peak busy hour. So, for the complete analysis, the experiment is performed in both off-peak hours and peak hours. Tables 2 and 3 show the resource utilization and simulation environment during the data transmission between doctor and patient through a cloud server. Table 2 reflects that the resource utilization is more in the peak hours compared to off-peak hours. However, during both the experiment duration still minimum resources are required, so the implementation can be easily adopted without any specific consideration.

Table 2   Resource utilization

Description	Usage off-peak hour	Usage peak hour	Limit	Fault
CPU usage	14.5%	22%	100%	0
I/O usage (Kbps)	1230	2090	8192	0
IOPS	2	16	1024	0
Entry processes	2	4	40	0
Number of processes	9	23	100	0
Physical memory usage	92.30 MB	292 MB	2.00 GB	0

Table 3 Simulation environment

Description	Value/name
Server location	Burlington, Massachusetts, United States (US)
Available CPU	2.30 GHZ Dual core
Available memory	2 GB Ram
Storage	5 GB HDD in RAID
Bandwidth	10 GB
Apache version	2.2.3.2
PHP version	5.4.45
MYSQL version	5.5.55–38.8
Architecture	x84_64
Operating system	Cent OS 6.5
Experiment duration	600 min [300 min(off-peak hour) + 300 min (peak hours)]
Experiment start time	00:10 AM (IST) [off-peak hour]
Experiment stop time	05:10 AM (IST) [off-peak hour]
Experiment start time	11:00 AM (IST) [peak hour]
Experiment stop time	04:00 PM (IST) [peak hour]

6 Result and Analysis

We have considered different attack scenarios by randomly simulating the various type of attacks such as man-in-the-middle attack, spoofing, etc. We have performed the experiment in three stages with 5 runs of 15 min each for both peak hour and off-peak hour. For the analysis, the vital signals of patient like heart beat rate, blood pressure, glucose level, body temperature, and hemoglobin level were transmitted in real time. In the first stage, the vital signals are transmitted to the doctor's end without any security and without any attack with an interval of 1 min. The end-to-end delays are shown in Fig. 3a, b, signifies the time needed for transmission of health data of patient using cloud server. The results of Fig. 3a show that about 0.2–0.4 ms is used to fetch health data from the sink, 0.4–0.6 ms to save that data to the cloud database server, and 3–3.5 ms used to retrieve the data from the cloud server during off-peak hours. Similarly, the results of peak hours (Fig. 3b) show that there is an increase in data read, data fetch as well as save time compared to off-peak hours results. The end-to-end delay in each run is slightly varying, as the load shared by the cloud is changing with time. So, average time required for the plain text transmission is 3.80 ms for off-peak hours while 4.90 ms for peak hours. In the second stage, the vital signals are transmitted to the doctor's end without any security and with an attack with an interval of 1 min. In this, the man-in-the-middle attack is periodically and randomly targeted. The average results of 5 runs for the periodic attack
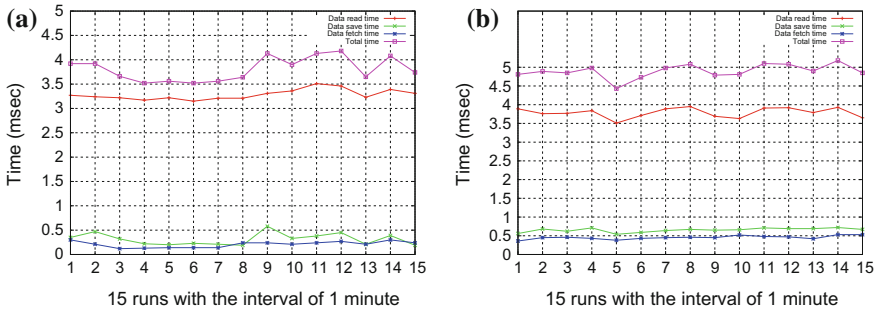
Fig. 3 End-to-end delay of vital signals transmission in plain text. **a** For off-peak hour. **b** For Peak hour
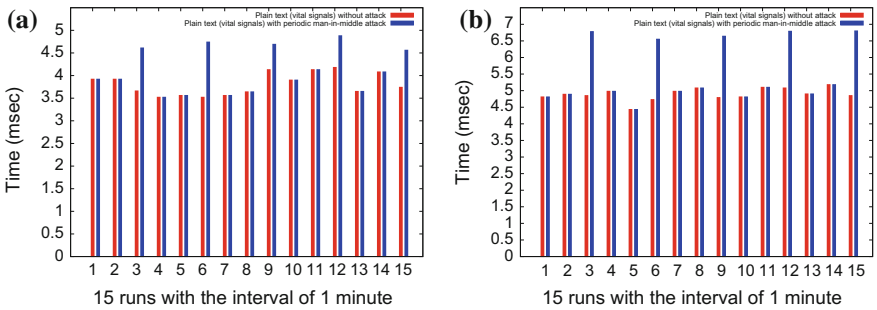
Fig. 4 End-to-end delay of vital signals transmission in plain text with periodic man-in-the-middle attack. **a** During off-peak hours. **b** During peak hours

where intruder reads every third transmitted information and again retransmits it during off-peak hours and peak hours shown in Fig. 4a and Fig. 4b respectively. Due to this attack, the transmission time increases to 4.5–5 ms for off-peak hours and 6.5–7 ms for peak hours. Similarly, the random man-in-the-middle is also implemented and tested with 5 runs. The average results of random attack for off-peak hour and peak hour are shown in shown in Fig. 5a and Fig. 5b respectively. The random attack results show that the increase in end-to-end delay transmission by ≈1.5–2 ms compare to without attack on that time. The increase in the transmission time reflects the attack during transmission.

In the third stage, the vital signals are transmitted to the doctor's end with authenticity and confidentiality. Figure 6a, b shows the transmission time required for each process according to the design technique for both the off-peak hour and peak hour. The doctor's finger validation time reflects the time required to compare the doctor's biosignal (fingerprint) to the biosignal store in the cloud when doctor wishes to see the patient health data. Data read time means the time required to read vital signal of patients from the sink.
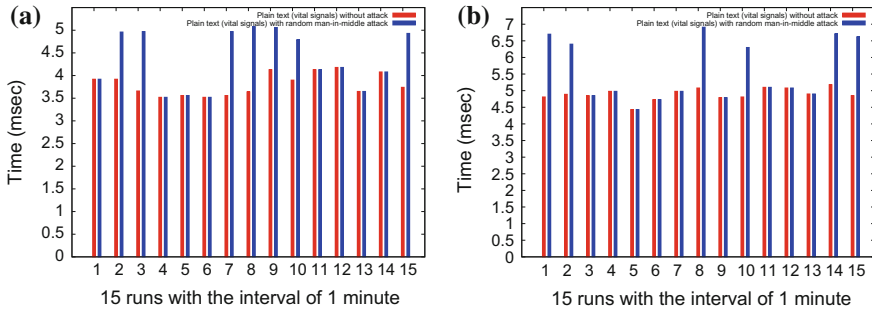
Fig. 5 End-to-end delay of vital signals transmission in plain text with random man-in-the-middle attack. **a** During off-peak hours. **b** During peak hours
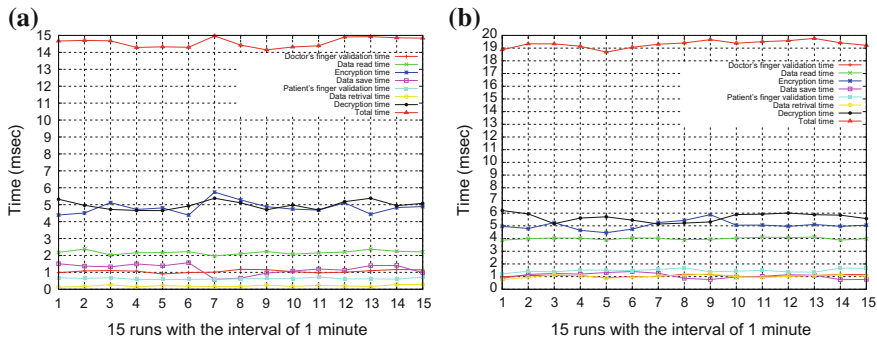
Fig. 6 End-to-end delay of cipher text using AES and authentication using biosignal. **a** For off-peak hour. **b** For peak hour

Encryption time shows the time required to encrypt the data with 128-bit block and 128-bit random secret key generation. Data Save time is the time required to store patient health data in encrypted format, secret key, and patient biosignal. Patient finger validation time is the time required to compare the received patient's biosignal (fingerprint) with the stored biosignal of the patient. Retrieve time is the time required to fetch the encrypted patient health data from cloud database server after the biosignal validation is done. The decryption time includes the time required to authenticate the received secret key with the already stored secret key and the decryption of cipher text to plain text. The average of total time required for complete process during off-peak hour is 14.59 ms (Fig. 6a) and during peak hour is 19.31 ms (Fig. 6b). Hence, the latency during off-peak hour is only 5.84% of 250 ms and during peak hour is only 7.72% of 250 ms permissible delay limit for healthcare application.

7 Conclusion

This work designed and implemented the cloud-assisted secure WBAN communication for healthcare application. The security measures adopted in this work, make it difficult for the intruder to intercept any node governed by the authentication of the biosignal's of the patients and the doctors. The biosignal of individual is unique in nature. The AES ciphering technique overcomes the various wireless attack. When the combined techniques are implemented in the cloud server, then it provides data confidentiality, data authentication, data integrity, freshness, and availability.

References

1. Ali, A., Khan, F.A.: Key agreement schemes in wireless body area networks: taxonomy and state-of-the-art. J. Med. Syst. **39**(10), 115 (2015)
2. Douglas, S.: Advanced encryption standard. Rivier Acad. J. **6**(2), 1–14 (2010)
3. Liu, J., Zhang, Z., Chen, X., Kwak, K.S.: Certificateless remote anonymous authentication schemes for wireless body area networks. IEEE Trans. Parallel Distrib. Syst. **25**(2), 332–342 (2014)
4. Liu, H., El Lazkani, E.: Biometric inspired mobile network authentication and protocol validation. Mob. Netw. Appl. **21**(1), 130–138 (2016)
5. Mainanwal, V., Gupta, M., Upadhayay, S.K.: A survey on wireless body area network: security technology and its design methodology issue. In: 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1–5. IEEE (2015)
6. Movassaghi, S., Abolhasan, M., Lipman, J., Smith, D., Jamalipour, A.: Wireless Body Area Networks: A Survey. IEEE Commun. Surv. Tutor. pp. 1–29 (2013)
7. Mukhopadhyay, S.C.: Wearable sensors for human activity monitoring: a review. IEEE Sens. J. **15**, 1321–1329 (2015)
8. Negra, R., Jemili, I., Belghitha, A.: Wireless body area networks: applications and technologies. In: Procedia Computer Science. In: The Second International Workshop on Recent Advances on Machine-to-Machine Communications, vol. 83, pp. 1274–1281 (2016)
9. Pathania, S., Bilandi, N.: Security issues in wireless body area network. Int. J. Comput. Sci. Mob. Comput. **3**(4), 1171–1178 (2014)
10. Saleem, S., Ullah, S., Yoo, H.S.: On the security issues in wireless body area networks. JDCTA **3**(3), 178–184 (2009)
11. Sampangi, V.R., Dey, S., Urs, R.S., Srinivas, S.: A Security Suite for Wireless Body Area Networks (2012). arXiv:1202.2171
12. Venkatasubramanian, K.K., Banerjee, A., Gupta, K.S.: EKG-Based Key Agreement in Body Sensor Networks. IEEE, 1–6 (2008),
13. Venkatasubramanian, K.K., Banerjee, A., Gupta, KS.: PSKA: usable and secure key agreement scheme for body area networks. IEEE Trans. Inf. Technol. Biomed. **14**(1), 60–68 (2010)
14. Zheng, G., Fang, G., Shankaran, R., Orgun, M., Zhou, J., Qiao, L., Saleem, K.: Multiple ECG fiducial points-based random binary sequence generation for securing wireless body area networks. IEEE J. Biomed. Health Inf. **21**(3), 655–663 (2017)
15. Zhou, J., Cao, Z., Dong, X., Xiong, N., Vasilakos, V.A.: 4S: a secure and privacy-preserving key management scheme for cloud-assisted wireless body area network in m-healthcare social networks. Inf. Sci. **314**, 255–276 (2015)

A Study of Android Application Execution Trends and Their Privacy Threats to a User with Respect to Data Leaks and Compromise

Utkarshni Sharma and Divya Bansal

Abstract Android is used as an operating system by a vast range of devices such as mobile phones, tablets, TV sets, watches, etc., providing users with a collection of android applications to, carry through, their daily needs. Android security has been built upon a "Permission-based mechanism", which gives access to the critical information, of an Android device to an application, and often does not allow the user to understand the privacy implications of those accepted permissions. User's overriding desire to complete the primary task to install an Android application reduces their attention to the installation-time permission requests. This study on Android applications will help the users to gain better understanding of the required and used permissions in Android environment, giving them a preferable user control to effectively maintain confidentiality, integrity, availability, security, and privacy of user's data.

1 Introduction

Android Google play store forechecks every application before publishing in its store, yet some black hat developers make their way out and pass through those tests. Each application has very limited capabilities by default, and require permissions to access sensitive data or services on device. User's overriding desire to complete the primary task to install Android application reduces their attention to those installation-time permission requests. In the work of Gu J. et al. the effects of permission sensitivity, permission justification, and app popularity on Android users' privacy concerns

U. Sharma (✉) · D. Bansal
PEC University of Technology, Chandigarh, India
e-mail: utkarshni.sharma@gmail.com

D. Bansal
e-mail: divya@pec.edu.in

and download intention were investigated. This was being done considering only one application "Delicacy" (provides restaurant descriptions and customer's review) during its download time by mangling with its design on a sample size of 286 subjects. The study did not include the effect of user's comments and developer's information on the subjects. This study on Android applications is done in five sections on selected applications, Sect. 1 determines the relationship between an Android application and an Android user, based on five metrics. Section 2 demonstrates log analysis of applications through "Network log" during their normal usage and after force stopping them on the device. Section 4 discusses results, informing users of what mistakes they mostly make, which could get their data compromised and what auxiliary measures can be considered to be secure. Section 4 depicts comparison of different Android versions based on some standard parameters. At last, Sect. 5 set forth conclusion, future work, and references. This study on Android applications will help user's to gain better understanding of required and used permissions in Android environment and giving a preferable user control as to which permission of an application should be allowed or disabled, to effectively maintain confidentiality, integrity, availability, security, and privacy of user's data.

2   Related Work

Many researchers investigated the arising issues in Android security, including coarse granularity of permissions, incompetent permission administration, insufficient permission documentation, overclaim of permissions, permission escalation attack, and TOCTOU (Time of Check to Time of Use) attack to illustrate the relationship among these issues, and the existing countermeasures to address them. One of them included effects of permission sensitivity, permission justification, and app popularity on Android users' privacy concerns and download intention, considering only one application "Delicacy" (providing restaurant descriptions and customer's review) during its download time by mangling with its design on a sample size of 286 subjects. The study did not included effect of user's comments and developer's information on the subjects. Also, theoretical study on modeling antileak sensitive data firewall for the static and dynamic behavior analysis of an Android application was proposed using API Hooking method based on Cydia substrate framework. This framework had to work for Android version 2.3 to Android version 4.3, but when verified manually the framework did not work on the given Android version, giving the model a major disadvantage. A set of contrast permission patterns were aimed to detect the difference between clean and malicious applications on the basis of static analysis of two sets of permissions: required and used permissions. Data was collected through static analysis (feature extraction using reverse engineering tools) on 1227 applications, due to large number of data set there was an issue of redundant samples. There are other features from the applications that could be included like dynamic analysis (behavior monitoring during application execution), user's comments, application popularity, developer's information, etc., that could be used to

improve the detection ratio of clean and malicious applications. Applications were characterized for each category of application based on rating, and proposed privacy score determining: normal application and abusive or malicious or risky applications, which showed pattern deviation from the set pattern. The study used publically available data set and filtered track of permissions embedded in the Android 4.4 operating system. Forensic study acquired and analyzed user's interaction, device-stored data and network traffic of 20 applications (4 out of 20 : Snapchat, Tinder, Wickr, BBM were not included) under category of chatting, dating, and messaging on the basis of application popularity. They were able to find the messaging application's vulnerabilities in terms of how they store and transmit data. Also, their was a survey conducted to understand the beliefs, attitudes, behaviors, and expectations of mobile users toward privacy and energy issues, and their behavioral intention on what kind of countermeasures adopt to reduce their privacy leakage.

3 System Design

This work on Android applications determines relationship of Android application and an Android user, based on five metrics. Also, a comparison of different Android versions based on some standard parameters is being done to depict their security levels. The system consists of three phases: training phase, testing phase, and pattern analysis (Figs. 1, 2, 3, 4 and 5).
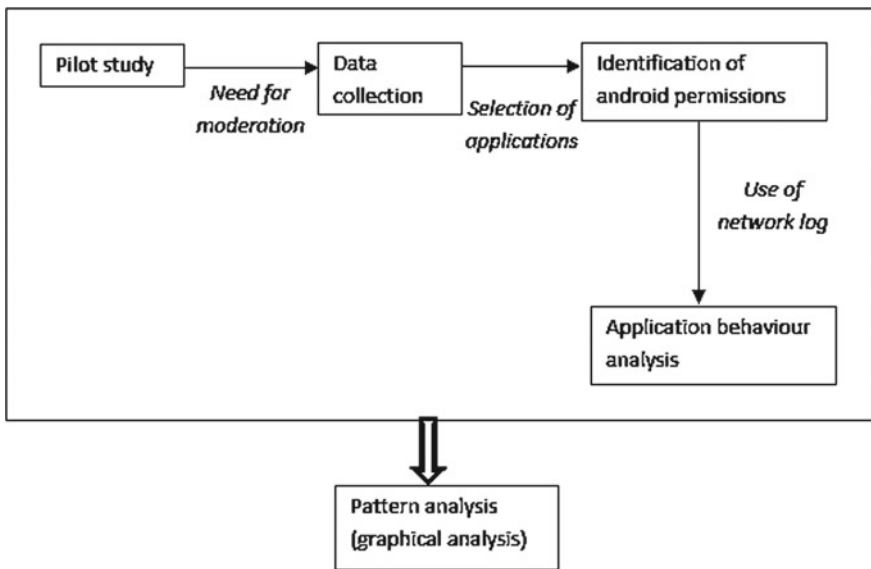
Fig. 1 System process model
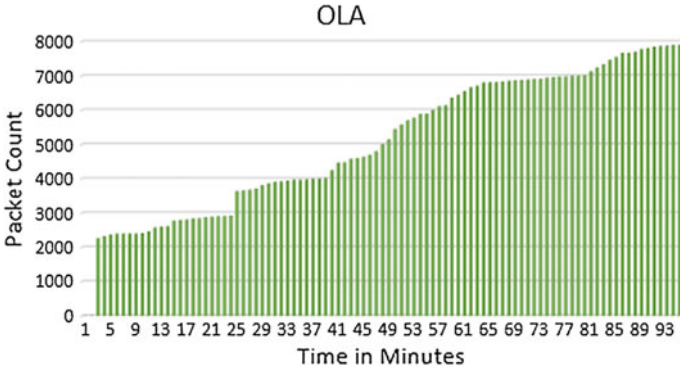
Fig. 2 Ola app in its default running state
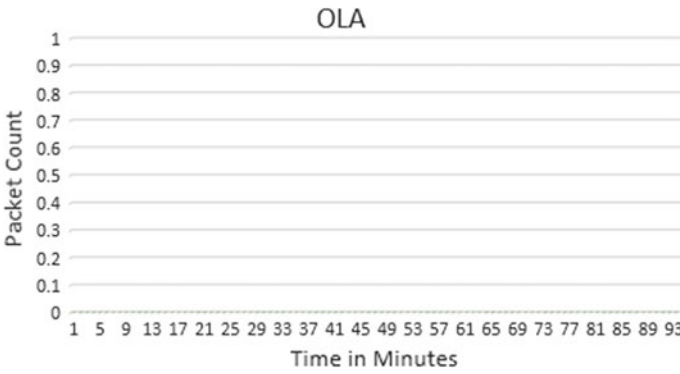
Fig. 3 Ola app after forcibly stopping from its default running state
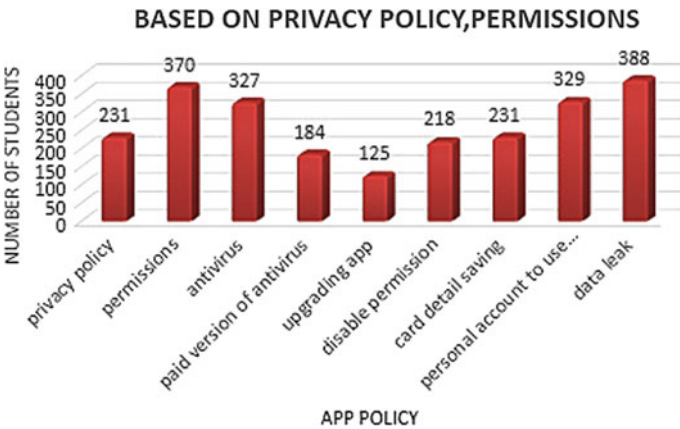
Fig. 4 Graphical representation based on application popularity
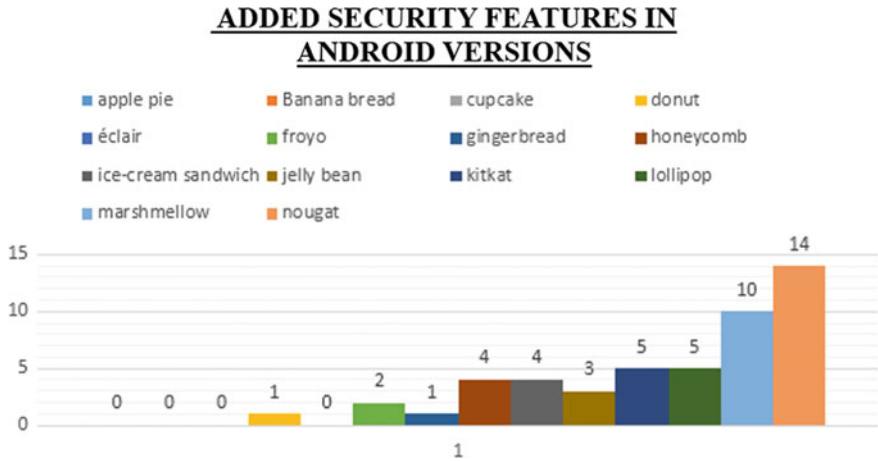
Fig. 5 Security levels in Android versions

3.1 Data Collection

Data collection was done in PEC University of Technology and Panjab University Chandigarh, consisting of undergraduate (B.Tech., B.Ph., Biophysics, Biochem, B.A., HSC, B.Sc.), postgraduate (M.Tech., M.Ph., M.A., M.Sc., M.Phil.) and Ph.D. students under the age group between 18 years and 50 years. The purpose and the significance of the study were explained to the students.

4 Result and Analysis

4.1 Android Permissions

First objective of the study was to identify all the permissions which were asked prior to the installation of the Android applications. In order to achieve this, 33 Android applications were selected from different categories to analyze their permissions prior to their installation which are to be accepted by the user in order to download and install a particular application on the device. In total, 58 permissions were noted, out of which 34 were termed as normal and rest 24 as dangerous. Based on dangerous permissions, all those applications are as follows:

1. Requesting permissions between 0 and 6 were put under the category of Normal App, i.e., they were causing negligible harm to a person's confidential data, security, and privacy.

2. Requesting permissions between 6 and 10 were put under the category of Risky App, i.e., they were causing a moderate amount of harm to a person's confidential data, security, and privacy.
3. Requesting permissions between 10 and 16 were put under the category of Dangerous App, i.e., they were causing a significant amount of harm to a person's confidential data, security, and privacy.

4.2 Execution Style Analysis of Android Applications on Device

All the 33 applications were successfully downloaded and installed on the Android device with Android version (5.1). An app "Network Log"? was installed to monitor the behavior of the applications during their usage and after force stop. The time period selected for behavior analysis was of 3 months. After force stopping an application, there was a significant dip in the amount of network traffic generated, as compared to traffic they generate while they were kept on running at the background even after closing the app. Following are the graphs of the Android application for two instances. First, while they are normally by the user and second, after forcibly stopping them from their running state:

1. Ola cabs

Similarly these observations were observed in all the 33 applications, that after force stopping an application packet count of some of the apps went to zero and some showed a significant dip, in the traffic (packet count) generation by the applications, on the device, such as in Google Maps, Facebook, Instagram, etc. Also, it was observed that between WhatsApp and Google Allo, Google generated less traffic (packet count) as compared to WhatsApp.

4.3 Pattern Analysis

From the data collected, the following graphs were generated based on five criteria.

4.3.1 Based on Application Popularity

This graph holds results on the basis of application popularity on Google Play store as well as among the users.

Results of this study showed that 79.4% users considered app rating for download, 60.9% considered number of downloads of an app before downloading it, 75.8% considered user's information, and 43.2% considered developer's information about a

particular application present on Google Play store before downloading an application. Although knowing that alone app rating or number of downloads or user's information about an app or developer's information of an application does not provide better understanding of an app still users do not consider them before downloading an application. Similarly, all the four criteria based on privacy policy and permissions, based on execution style during run and download time, based on performance and based on data security,generated results which helped to figure out what are the user's perceptions regarding privacy issues and what mistakes they mostly make to get their data compromised without being notified.

4.3.2 Comparison of Different Android Versions Based on Security Features

All the Android versions from its very first version Apple pie (1.0) till date Nougat (7.1.1) are compared on the basis of user/development and security features. Following graphical analysis depicts more clearly the comparison of different versions of Android according to their security features:

Out of all the versions of Android, the latest version Nougat is the safest with most of the security features in it. Gingerbread was targeted more with malware more than any other Android version. The majority of attacks were carried out by exploiting Java vulnerabilities and more than 50% of all attacks were SMS Trojans that steal money from users by sending text messages to premium-rate services.

5 Conclusion

This study aimed to perform the study of Android application execution trends of applications and how users respond toward the privacy issues and what leads to their data compromise. The results have shown that although some of the users are aware of implication of the risk of installing an application to their data but still they do not take any action on their end to somehow protect their data from getting compromised. Two main factors that contributed toward this change is the open-source nature of the platform and the flexibility provided to users and developers. This is an extension of Gu J. et al. where only application popularity and performance were considered as only parameters considering only one application. Through this work a user will become more conscious before downloading and installing Android applications on device. If in any case an application hinders the normal working of a device, over-claims permissions, or promotes inappropriate advertisements, that should be considered as not legitimate and action should be taken immediately. This will give a preferable user control as to which permission of an application should be allowed or disabled, to effectively maintain confidentiality, integrity, availability, security, and privacy of user's data.

References

1. Fang, Z., Han, W., Li, Y.: Permission based android security: issues and countermeasures. Comput. Secur. **30**(43), 205–218 (2014)
2. Wang, W., Wang, X., Feng, D., Liu, J., Han, Z., Zhang, X.: Exploring permission-induced risk in android applications for malicious application detection. IEEE Trans. Inf. Forensics Secur. **9**(11), 1869–1882 (2014)
3. Armando, A., Carbone, R., Costa, G., Merlo, A.: Android permissions unleashed. In: IEEE Computer Security Foundations Symposium, 13 Jul 2015, pp. 320–333
4. Chmielarz, W.: Study of smartphones usage from the customer's point of view. Proced. Comput. Sci. **31**(65), 1085–1094 (2015)
5. Gu, J., Xu, Y.C., Xu, H., Zhang, C., Ling, H.: Privacy concerns for mobile app download: an elaboration likelihood model perspective. Decis, Support Syst (2016)
6. Ammari, N., El Mrabti, A.A., El Kalam, A.A., Ouahman, A.A.: Firewall anti-leak of sensitive data. Proced. Comput. Sci. **31**(83), 1226–1231 (2016)
7. He, Y., Li, Q.: Detecting and defending against inter-app permission leaks in android apps. In: IEEE International Performance Computing and Communications Conference (IPCCC), 9 Dec 2016, pp. 1–7
8. Singh, P., Tiwari, P., Singh, S.: Analysis of malicious behavior of android apps. Proced. Comput. Sci. **1**(79), 215–220 (2016)
9. Android Security: https://source.android.com/security/reports/GoogleAndroidSecurity2016 ReportFinal (2016). cited 10 Apr 2017
10. The Seven Deadly Android Permissions: How to Avoid the Sin of Slothful Preparedness. http://www.makeuseof.com/tag/sevendeadlyandroidpermissions. Cited 28 Feb 2017

Three-Dimensional Design of a New Maglev Vehicle and a Study of It Using Computer Vision

Kuldip Acharya and Dibyendu Ghoshal

Abstract A Maglev car travels by floating on a magnetic track with the attractive and repulsive force of magnets to create both lift and propulsion. The paper presents outlines of an application on magnetic levitation transportation technology to design a new car and study of it utilizing computer vision. This vehicle is specially designed to give high comfort to the physically challenged people and to move in a profoundly secure condition. The vehicle is planned with a full resting beds compartment and can move at a very high speed. Inductrack a frictionless, attractive levitation framework is proposed to be used here for super speeds. It offers a smooth contactless trip to the goal. Helper components are just required when the vehicle needs to back off and needs to stop at a station. The financially savvy Halbach arrays of magnets are utilized to produce high magnetic fields to suspend and move the vehicle. Computer animation of the whole circumstance shows up and observed it to be a good vehicle with high speed. Computer vision system study is utilized to detect cars, monitoring vehicle, analysis of car video information for a safe journey, and to avoid accidents. Tracking a car in a vulnerable environment is difficult due to noise generation. To overcome this problem, a method is proposed for noise removal from the car videos for tracking it in a vulnerable environment. It is found that the proposed methods give better result compared to the existing one. Thus, it shows that computer vision can greatly help the Maglev vehicle to provide more safety, high alerts, and high comfort to the passengers.

K. Acharya (✉)
Computer Science and Engineering Department, National Institute of Technology,
Agartala, India
e-mail: kuldip.acharjee@gmail.com

D. Ghoshal
Department of Electronics and Communication Engineering,
National Institute of Technology, Agartala, India
e-mail: tukumw@gmail.com

1 Introduction

Magnetic levitation (Maglev) [1] is a train transportation technology which applied magnetic levitation to transport train, and vehicles without making physical touch with the ground, unlike conventional trains. Maglev automobile travels along a guideway supplied with magnets to create both lift and propulsion. Maglev trains flow more smoothly and speedy than wheeled transportation structures [1]. There are many road transportation systems available for both short or a long route within a state or from a state to another state. However, the time and journey are taken by the buses, cars, autos, and the traditional rails are not suitable for all the passengers. On the other side, electrodynamic system Maglev or electromagnetic Maglev rails expenditure is not always affordable for the general passengers. The physically challenged passengers always a need a safer, quicker journey with high-comfort traveling experience. This scenario introduces a new small vehicle based on the Inductrack technology of Maglev transportation system. In the proposed design, a magnetically suspended vehicle is introduced. In the present study, the vehicle is pulled by magnetic power for transporting merchandise, and passengers. The goal of the configuration of magnetic levitation (Maglev)-based vehicles is to carry individuals and especially passengers who have physical problems. The vehicle seats are planned such that physically challenged individuals who have a problem with their legs or hands can consume additional room and can put their legs on feet rest and can utilize different parts of the vehicle. The visually impaired people can use Braille internet system. It is seldom found where general auto rickshaws, cars, rails, airplanes give this kind of facility to the differently able people groups. Maglev vehicle serves the passenger to get sufficient comfort during the journey. Many research works are going on the Maglev vehicles. The computer design and animation of the Maglev vehicle are done by Autodesk Maya [2] student edition to illustrate the proposed framework. The present study shows how the proposed transportation structure can be a better voyaging vehicle to the future world. Vision-related problems of real-life scenarios, analysis of computer-generated graphics, models, animation, and so forth can be solved using Computer Vision System Toolbox [3] algorithms and functions. Computer vision makes it easy to implement the feature detection, extraction, and matching; object detection and tracking; motion estimation; and video processing efficiently in a short period of time. Computer vision-related simulations and statistical analysis is implemented in MATLAB [4] software.

2 Review Work

Magnetic levitation [1] is a technology by which anybody or object can be floated against the gravitational force. The principle lies in the theory of superconductivity and cryogenics. It has been found that when the temperature of a certain object, called superconductors can be lowered nearly at 0 Kelvin or slightly more than that, the substance loses all magnetic property from it and it is prevalent in the dia-magnetic material body. At this stage, the theoretically infinite amount of current can be passed through it as conductivity becomes infinite. Based on this principle, any superconductor object can be floated and moved along a guiding track which is also a superconductor. The speed of the movement can be controlled by varying the amount of current, i.e., the strength of the magnetic field causing the supercon-ductivity. Among three common types of magnetic levitation system, electromag-netic and electrodynamic systems are costly. Rather Inductrack system is cheaper and can be used as an industrial alternative in vehicle technology. In the Electro-magnetic suspension (EMS), magnetic levitation (Maglev) system powered elec-tromagnets are utilized and in the electrodynamic system (EDS) superconductors are used [1]. To keep up the superconductivity cryogenic framework [5] needs to cool the superconducting magnets which are exceptionally costly. The great physicist Klaus Halbach [6] configures arrays of permanent magnets so that the intensity of the magnetic field becomes very stronger on one side of the array and nearly canceling the magnetic field of another side of the array. Halbach array is made of the neodymium–iron–boron alloy [7] to generate a higher magnetic field. The speed of Inductrack Maglev is approximately estimated to be 350 km/h [8]. Halbach arrays of permanent magnets are utilized in this Inductrack Maglev vehicle to travel miles in a short timeframe. The infrastructure expense of a Maglev train with the electrodynamic structure is high. Here, we use Inductrack magnetic levi-tation method which is in excessive contrast with general rail framework, however, less expensive than Electrodynamic System (EDS) Maglev system. The Maglev transportation system is monitored by the computerized tracking system to report data to control room [9]. These data analyses determine the speed and position of the Maglev train. Tracking and monitoring a vehicle in the vulnerable environment may become difficult. Computer vision may help in this condition by reducing noises from video signals, enhancing the video frames for clear visibility for better tracking, monitoring and to controlling the overall transportation system. A vision based magnetic levitation device can be used to measure the location of the object as opposed to light-based sensors. To do that vision based position sensor evolved. The image processing methods acquired the location information which adjusts the coil current to create a force on the object [10]. Computer vision strategies can track human, density and evaluation of crowd [11].

3   The Design of the Proposed Vehicle

The track of vehicle is proposed to make arrays of the unpowered loop of insulated wires. When the proposed Inductrack Maglev vehicle pushes ahead, the magnets in the Halbach arrays induce currents in the track's cables, which create an electromagnetic field that repels the arrays of permanent magnets. The framework's completely frictionless nature, implying that accomplishing levitation requires no control systems to look after stability. When the vehicle is in rest at a station, no levitation happens, and assistant wheels bolster the vehicle. When the Inductrack vehicle car pushes ahead, the magnets in the Halbach arrays in the track, create an electromagnetic field. This magnetic field levitates the vehicle. Linear synchronous motor [1] is used to propel the vehicle. Initially, external power is needed until the vehicle levitates. At the point when the vehicle runs the magnetic field from these permanent magnets induces repelling currents in the track. The magnets in the Inductrack are let down in such a manner to create one-side magnetic fields and the intensity sufficiently strong to drag the vehicle. The intensity of the side magnets varies in such a manner that it will add some magnetic force on the levitating track. The alternating assembly of the magnets gives rise to a forward propagating pulling force which helps to strengthen the upward levitating force. There are auxiliary tires attached to the system for the safe landing. We propose to use powerful batteries for the secure reinforcement framework.

3.1   Gaussian Mixture Model and Computer Vision Study on the Proposed System

This example shows how to detect and count cars in a video sequence using foreground detector based on Gaussian Mixture Models (GMMs) [12]. The utilization of the foreground detection method, motion-based multiple object tracking, and blob analysis [13] method are some special features in computer vision to search and count cars in a video sequence. In the proposed work, 3D models and animation of the vehicle based on the concept of magnetic levitation transportation technology is designed and animated in Maya software. The work focuses on the computer vision study of various scenarios. The Kalman filter [14] object predicts an object's probable future location and can be used for tracking and reduce noise in the detected location.

   The computer-generated three-dimensional (3D) design, gives an overview how the system will look like and how it will operate. The principle of Inductrack operation adopts a particular setting of the permanent magnets of fixed magnetic intensity. The speed of the introduced inducted vehicle can be varied by tuning the current passing through the solenoids. At the time when a certain break is needed, a reverse magnetic force is necessary to create by changing the direction of the current through a solenoid. Figure 1 shows the view of the proposed Maglev

Fig. 1 Inductrack vehicle running in a city
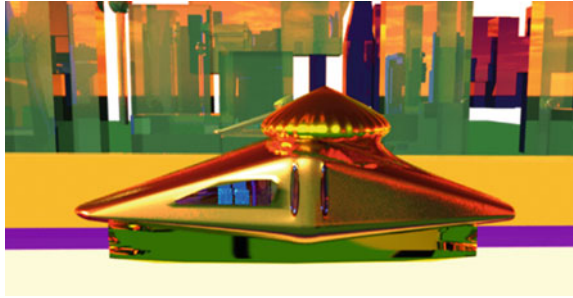
Fig. 2 Vehicle interior room 3D modeling
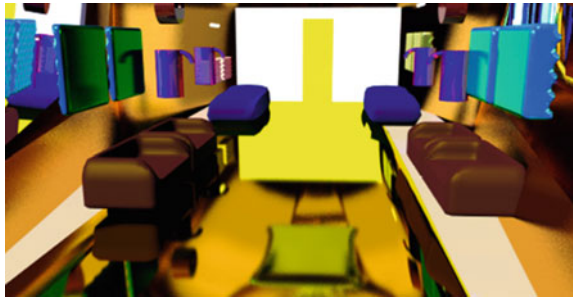
vehicle, where we can see a vehicle is running on a Maglev track on the city roads. By using this vehicle, the traveler can move from a city to another city or states in a short period comparing other road vehicles like cars, autos, trucks, and buses. Figure 2 shows the interior design of the Maglev vehicle room. The room designed with the necessary equipment for the physically challenged person (Figs. 3, 4 and 5), (Table 1).

Fig. 3 Motion object tracking using Kalman filter
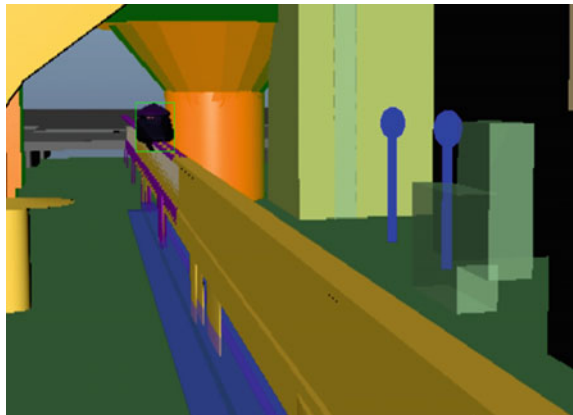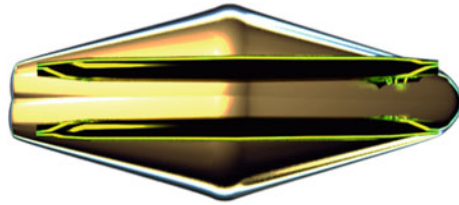
Fig. 4 Bottom side of vehicle
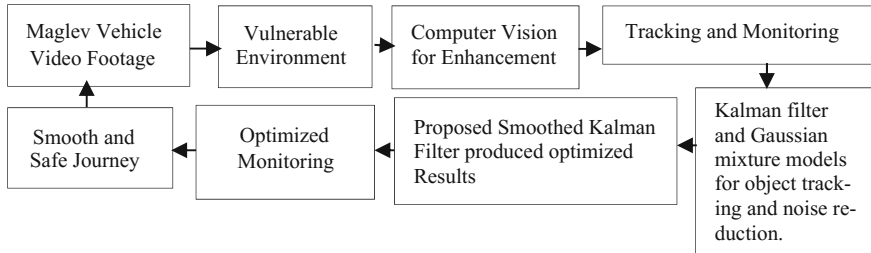
Fig. 5 Flowchart of the proposed design

Table 1 Maglev vehicle distinctive parts and their guesstimate values

Maglev vehicle parts	Approximate values and Guesstimate properties
Components can use to manufacture the vehicle body	Steel, aluminum, copper, fibers, glass, rubber, and special plastic
Vehicle dimension in length × width × height (L × W × H)	50 feet, 15 feet, 25 feet
Vehicle track	17 miles long. made of Halbach array, plastic, iron, and steel
Pillars height	10 feet

In Table 2, the average mean values, average standard deviation, and average variance have been shown. The result has been produced by Matlab program. The animated files of the proposed vehicle have been designed from scratch in Autodesk

Table 2 Statistical information of maglev vehicle animation videos

Video file	Average mean	Average SD	Average variance
Vehicle 1.avi	0.2737	0.1641	0.0269
Vehicle 2.avi	0.2839	0.1683	0.0281
Vehicle 3.avi	0.2636	0.1568	0.0234
Vehicle 4.avi	0.2731	0.1615	0.0229
Vehicle 5.avi	0.2747	0.1651	0.0279

Maya 2016 student version software. This is a conceptual design of a Maglev vehicle mimicking the magnetic levitation train concept.

4 Proposed Method

Synthesize a random signal that has a value of 1 and is corrupted by a zero-mean Gaussian noise with standard deviation of 0.1. Removing noise from a random signal which is corrupted by a zero-mean Gaussian noise have been done [15] through Kalman filter. In the proposed method, we have smoothed the produced result which is obtained by removing noise from the signal by using a Kalman filter.

$$\alpha = 0.99; s = 0; s = \left((\alpha)^* t + (1 - \alpha)^* s\right). \tag{1}$$

Here, α (alpha) is an adjusting parameter to optimize the overall performance. It is used to filter the calculated value. The exponential moving average function has been applied to the predefined calculated value. Here, 't' denotes the result obtained from noise cleaning through Kalman filter. The value of alpha is assumed here as 0.99 for optimal result.

5 Results and Analysis

Figures 6, 7, 8 and 9, shows the simulated results and which is based on the parameters in Table 3. The original signal, noisy signal, filtered signal which is already implemented [15] and the smoothed filtered signal by our proposed

Fig. 6 Signal noise removing using Kalman filter with length 100
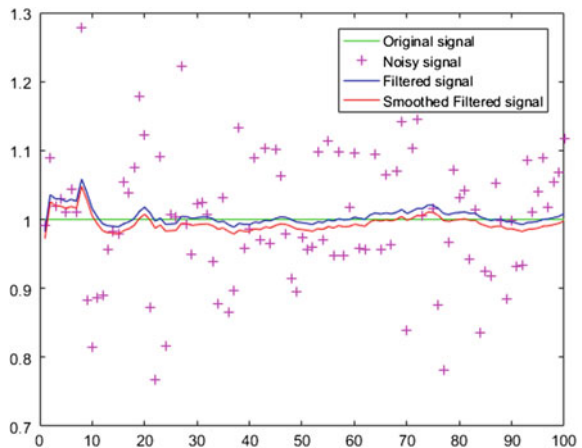
Fig. 7 Signal noise removing using Kalman filter with length 200
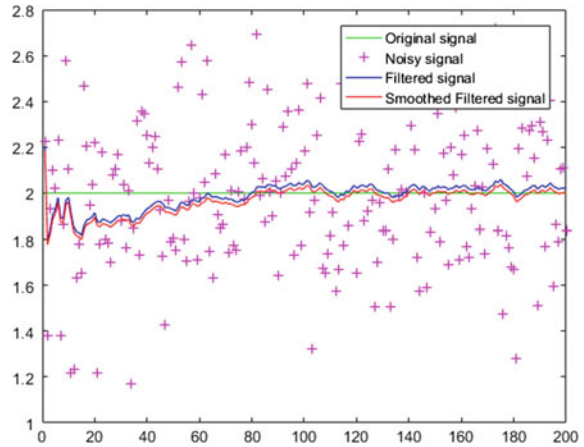
Fig. 8 Signal noise removing using Kalman filter with length 400
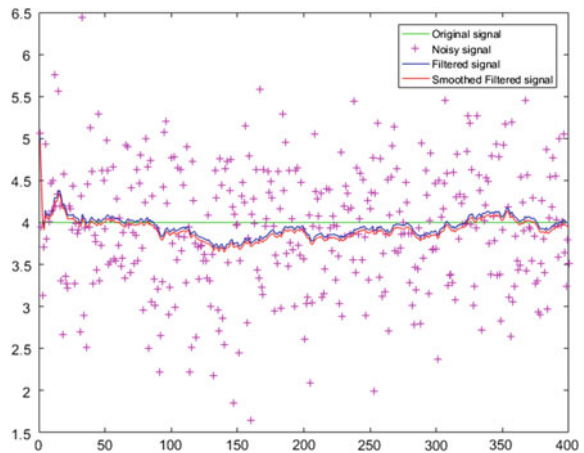
methods have been shown in Figs. 6, 7, 8 and 9. We have run the simulation in MATLAB for five times and get the simulation results. All the produced results clearly showing that smoothed filtered signal is giving optimum result. The Peak Signal-to-Noise Ratio (PSNR) in decibels (dB) mean squared normalized error performance function (MSE) and MAXERR is the maximum absolute squared deviation of the data. The results are showing that the proposed method is giving the slightly better result. Thus, it proving the proposed method is optimal. The simulation parameters are predefined. It is observed that the overall performance degrades when the value of the random signal, standard deviation, and length increases (Table 4).

Fig. 9 Signal noise
removing using Kalman filter
with length 500
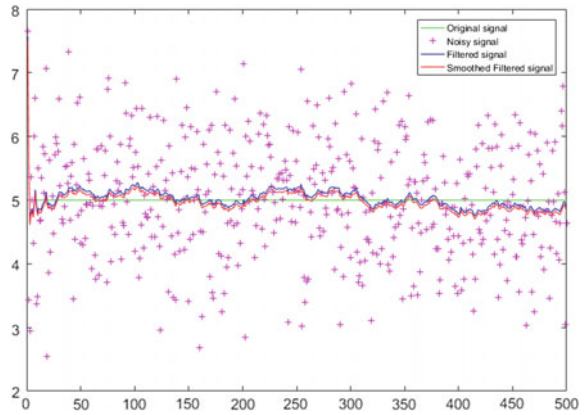
Table 3 Simulation
parameters

Random signal	Standard deviation	Length
1	0.1	100
2	0.3	200
3	0.5	300
4	0.7	400
5	0.9	500

6   Conclusion

Construction of magnetic levitation-based vehicles like car, train, etc., is a costly
affair, hence after meticulous design, it is always safe to have a prior understanding of
the proposed system before actual manufacturing. Computer vision-based analysis
may be thought of a good process for having an insight into the proposed system. In
addition to, all minor facilities can be provided with best available ease which can be
visualized beforehand with the help of computer vision-based design. Magnetic
levitation-based vehicles move at a very high speed and passengers need to be
provided with adequate safety measures by using comfortable and reliable safety
belts. The vehicles move along a fixed track and thus the possibility of head-on
collision or any other accident with other vehicles may be excluded. We have tested
the result using PSNR, MSE, and MAXERR and it is observed that the proposed
methods give the best results. Computer vision can help to control the speed of the
vehicle and automatic scene detection in both inside and outside of the vehicle to
provide safety instruction to the drivers, officials, staffs, and passengers. Thus,
computer vision may be greatly utilized to offer a stable and safe traveling experience
in Maglev transportation, especially in vulnerable environments.

Table 4 Simulation results

Random signal	Standard deviation	Length	Kalman filter [] PSNR	Kalman filter [] MSE	Kalman filter [] Maxerr	Proposed method PSNR	Proposed method MSE	Proposed method Maxerr
1	0.1	100	68.518	0.0091	0.27873	85.753	0.0001	0.0478
2	0.3	200	47.704	1.1033	1.717	48.424	0.9347	1.1811
3	0.5	300	41.815	4.2813	3.5017	42.181	3.9357	2.3123
4	0.7	400	38.498	9.1888	5.4448	38.83	8.5126	3.96
5	0.9	500	35.847	16.919	6.6812	36.14	15.816	6.4996

References

1. Lee, H.-W., Kim, K.-C., Lee, J.: Review of maglev train technologies. IEEE Trans. Magn. **42**(7), 1917–1925 (2006)
2. Palamar, T.: Mastering Autodesk Maya 2016. Autodesk Official Press. Wiley, pp. 1 (2015)
3. Computer Vision System Toolbox. https://in.mathworks.com/help/vision/index.html?search Highlight=Computer%20Vision%20System%20Toolbox&s_tid=doc_srchtitle. Accessed 04 Apr 2017
4. Guide, M.U.S: The Mathworks. Inc., Natick, MA 5, P. 333 (1998)
5. Benedict, B.A., Lester, J.M., Linenberger, D.D.: Cryogenic cooling system. U.S. Patent No. 4,825,667 2 May 1989
6. Post, R.F.: The Inductrack Approach to Magnetic Levitation (PDF)
7. Halbach, K.: Design of permanent multipole magnets with oriented rare earth cobalt material. Nucl. Instrum. Methods **169**(1), 1–10 (1980)
8. Post, R.F.: Maglev: a new approach. Sci. Am. **282**(1), 82–87 (2000)
9. Liu, Z., Long, Z., Li, X.: Maglev Trains: Key Underlying Technologies, p. 90
10. Kizir, S.,Küçükyıldız, G., Ocak, H., Bingül, Z.: Vision based magnetic levitation system. In: 20th Signal Processing and Communications Applications Conference (SIU), Mugla, 2012, pp. 1–4. (2012)
11. Silveira Jacques, Jr. J.C., Musse, S.R., Jung, C.R.: Crowd analysis using computer vision techniques. IEEE Signal. Process. Mag. **27**(5), 66–77 (2010)
12. Detecting Cars Using Gaussian Mixture Models. https://in.mathworks.com/help/vision/examples/detecting-cars-using-gaussian-mixture-models.html. 30 June 2017
13. Vision: BlobAnalysis System object. https://in.mathworks.com/help/vision/ref/vision.blobanalysis-class.html?s_tid=srchtitle. 30 June 2017
14. Motion-Based Multiple Object Tracking. http://in.mathworks.com/help/vision/examples/motion-based-multiple-object-tracking.html. 30 June 2017
15. Welsh, G., Gary, B.: An introduction to the Kalman Filter. University of North Carolina at Chapel Hill Chapel Hill NC 95, 95–041 (1995)

Combining GMM-Based Hidden Markov Random Field and Bag-of-Words Trained Classifier for Lung Cancer Detection Using Pap-Stained Microscopic Images

Moumita Dholey, Maitreya Maity, Atasi Sarkar, Amita Giri, Anup Sadhu, Koel Chaudhury, Soumen Das and Jyotirmoy Chatterjee

Abstract Lung cancer is a malignant tumour having uncontrolled lung cell growth. Papanicolaou (Pap)-stained cell cytology from Fine Needle Aspiration Cytology (FNAC) is the most followed approach for lung cancer diagnosis. However, the manual assessment of cytopathology slides under light microscopy is time consuming and suffers from feature ambiguities including inter-observer variability. Here, an automated computer vision approach is presented for identifying and classifying cancerous cell nuclei from pap-stained microscopic image of lung FNAC sample. The proposed methodology adopted Gaussian mixture model-based hidden Markov random field technique to segment cell nucleus. Later, bag-of-visual words model was used for nucleus classification, where scale-invariant feature transform feature were extracted from segmented nucleus for training a random forest classifier model. The adopted nucleus segmentation-cum-classification model was able to precisely segment the nucleus and classify them in two class, viz. Small Cell Lung Cancer (SCLC)

M. Dholey (✉) · M. Maity · A. Sarkar · K. Chaudhury · S. Das · J. Chatterjee
School of Medical Science and Technology, Indian Institute of Technology Kharagpur,
Kharagpur, West Bengal 721302, India
e-mail: dholey.moumita5@gmail.com

K. Chaudhury
e-mail: koel@smst.iitkgp.ernet.in

S. Das
e-mail: sou@smst.iitkgp.ernet.in

J. Chatterjee
e-mail: jchatterjee@smst.iitkgp.ernet.in

A. Giri
Calcutta National Medical College and Hospital, Kolkata, West Bengal 700014, India
e-mail: amita_giri@mail.com

A. Sadhu
EKO Diagnostic Centre, Medical College, Kolkata 700073, India
e-mail: sadhujee1@gmail.com

and Non-small Cell Lung Cancer (NSCLC). The segmentation process achieves a sensitivity of 98.88% and specificity of 97.93%. And also, the nucleus classification model was able to perform with a sensitivity of 97.31%, specificity of 99.54%, and accuracy of 98.78%.

1 Introduction and Literature Survey

Lung cancer, a fatal disease, is triggered by abnormal cell progression in the lung. According to American Cancer Society (2015), in the United States of America, the number of new cancer cases and deaths due to cancer are 1,658,370 and 589,430 respectively. Among these, the number of newly diagnosed lung cancer cases and deaths are 221,200 and 158,040 respectively [1]. For conventional microscopic assessment, Pap-stained FNAC smears from pulmonary lesions is very powerful diagnostic technique. However, the manual cytopathological assessment process suffers from several challenges like a time-consuming process, inter-observer variabilities, and feature ambiguities. To overcome these burdens, automated image analysis becomes the need of time. In a study for cervical dysplasia detection proposed by Bora et al. [2], authors have applied discrete wavelet transform, median filter, bit plane slicing, K-means clustering and Maximally Stable External Region(MSER) for nucleus segmentation. In another study by Bora et al. [3], deep CNN and Maximal Information Compression Index (MICI) were used for feature extraction and feature selection respectively. For classification, Softmax Regression classifier was used. In the classification study of breast cancer by Abdel-Zaher and Eldeib [4], a deep belief network (unsupervised) path with backpropagation (supervised) path were implemented and DBN-NN performed the best. Another study was done on Pap-stained breast FNAC slides by Kashyap et al. in [5]. To achieve uniform image intensity background correction was done. Image analysis was executed by using Nikon Imaging Software (NIS) (Version 4.00) and SPSS. Kecheril et al. [6] have proposed a study for lung cancer detection from sputum cytology images by using hessian determinant maximization, K-means clustering for segmentation. Yu et al. [7] have performed a histopathological study [6] on lung cancer prognosis. The authors have applied Otsu thresholding- based pipeline method for nucleus segmentation and Elastic net-Cox proportional hazards models (R package glmnet) for classification. Javaid et al. [8] developed another automated approach for lung cancer detection from CT image, where they have used contrast stretching, intensity thresholding, K-means clustering, 2D and 3D features and SVM classifier.

Though a huge development has been observed in Computer-Aided Diagnosis (CAD) tools for different cytological cancer detection (cervical, breast), for lung cancer diagnosis this is still lacking. To fulfill the aforesaid gap, specifically, CAD tool for lung cancer cytology has been targeted to develop.
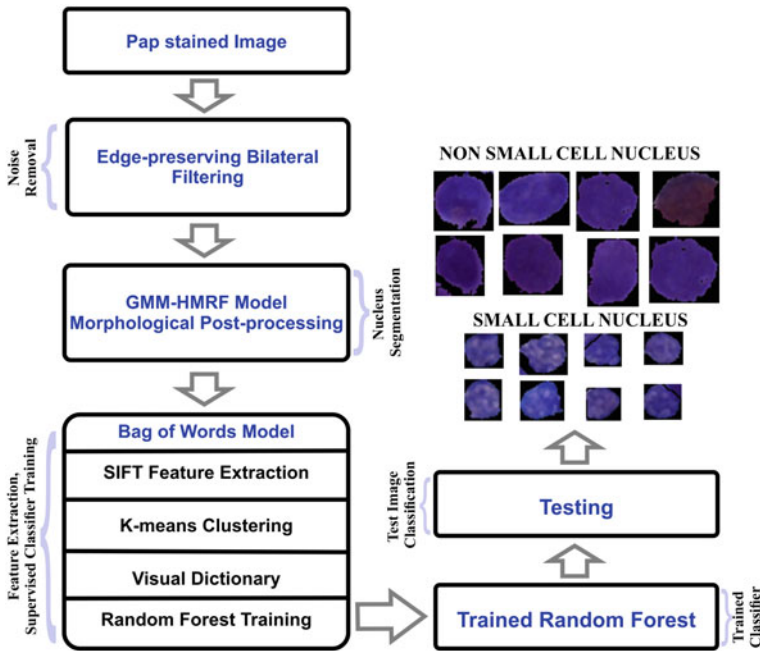
Fig. 1 Workflow diagram of the proposed study

2 Materials and Methods

2.1 Abstract View of the Adopted Algorithm

The workflow diagram of the proposed lung cancer detection algorithm is presented in Fig. 1. The adopted algorithm was designed to segment nucleus from Pap-stained microscopic images and classify them into SCLC and NSCLC types.

2.2 Sample Image Database Development

The pathological samples were collected from EKO Diagnostic Center, Medical College and Hospital, Kolkata, India. Institutional ethical clearance was taken for this study. Multiple images were captured from Pap- stained slides under $20x$ objective magnification (total $200x$ microscopic magnification). Total 120 samples were collected and 600 images were captured for sample database development. The developed image database was provided to multiple pathologists for ground truth image generation (Fig. 2).
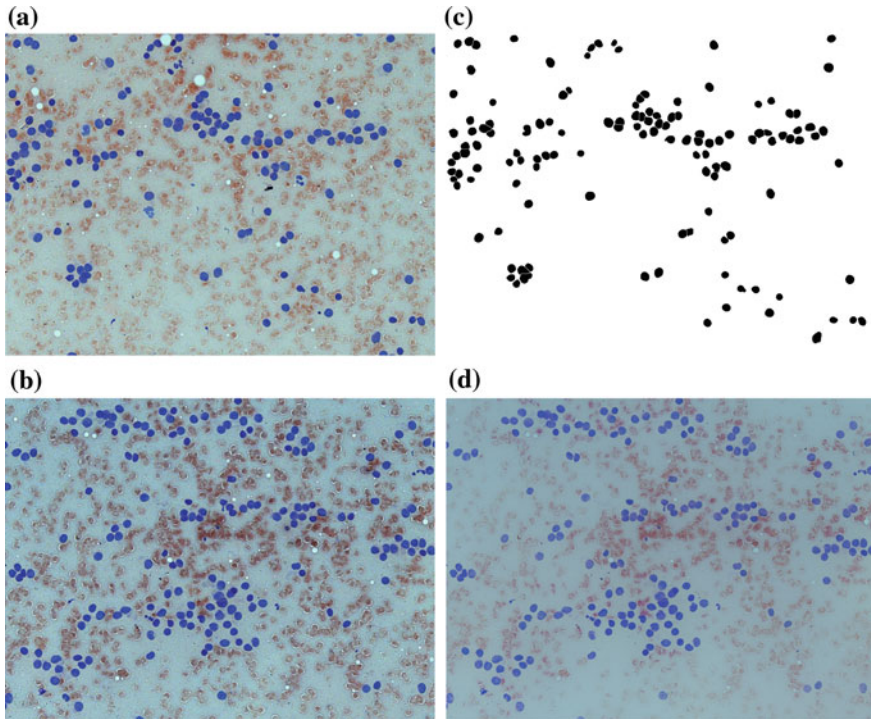
Fig. 2 Microphotographs of Pap-stained FNAC smears (objective magnification *20x*); **a**, **b** Original Pap smear image, **c** Ground truth image, **d** Noise-filtered image

2.3   Edge-Preserving Bilateral Filtering

Bilateral filter, a nonlinear filter, works as an image smoothing filter with edge preservation properties. The noise-removed image outputs are shown the Fig. 2. According to the formulation proposed by Tomasi and Manduchi [9], the bilateral filter is described. Let, $I(v) =$ input image. At $v = (x, y)$ coordinate, the bilateral filter can be defined as

$$\hat{I}(v) = \frac{\sum_{q \in N(v)} W_c (\|q - v\|) W_{sim} (|I(v) - I(q)|) I(q)}{\sum_{q \in N(v)} W_c (\|q - v\|) W_{sim} (|I(v) - I(q)|)}, \tag{1}$$

where $N(v)$ represents the neighbourhood of v. This smoothing filter can be described as a standard Gaussian filter with σ_c : $W_c(x) = e^{\frac{-x^2}{2\sigma_c^2}}$, and a feature preserving similarity weight function with σ_{sim} by penalizing large intensity variation, be $W_{sim}(x) = e^{\frac{-x^2}{2\sigma_{sim}^2}}$.

2.4 Nucleus Segmentation

Gaussian mixture model-based hidden Markov random field (GMM-HMRF) model [10] was adopted for nucleus segmentation. The representation of a Gaussian Mixture Model having g components be as follows:

$$G\left(v \mid \mu_i, \Sigma_i\right) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_i|^{\frac{1}{2}}} exp\left\{-\frac{1}{2}\left(v - \mu_i\right)'\Sigma_i^{-1}\left(v - \mu_i\right)\right\}. \tag{2}$$

GMM model determines the unknown parameters

$$\Theta = \left\{w_1, w_2, ..., w_C; \mu_1, \mu_2, ..., \mu_C; \Sigma_1, \Sigma_2, ..., \Sigma_C\right\}. \tag{3}$$

Hidden Markov random field and its Expectation Maximization (HMRF-EM) is considered to find out the Θ. The HMRF-EM algorithm can be described as follows:

1. Start with $\Theta^{(0)}$
2. Calculate $P^{(t)}\left(y_i|x_i, \theta_{x_i}\right)$ (likelihood distribution).
3. Use $\Theta^{(t)}$ (current parameter set) for label estimation by MAP estimation:

$$\begin{aligned}
X^{(t)} &= \underset{X \in \chi}{argmax}\left\{P\left(Y|X, \Theta^{(t)}\right)P(X)\right\} \\
&= \underset{X \in \chi}{argmin}\left\{U\left(Y|X, \Theta^{(t)}\right) + U(X)\right\}
\end{aligned} \tag{4}$$

4. Posterior distribution calculation by applying Bayesian rule as follows:

$$P^{(t)}\left(l|y_i\right) = \frac{G\left(y_i; \theta_l\right)P\left(l|x_{N_i}^t\right)}{P^{(t)}\left(y_i\right)} \tag{5}$$

Here, $x_{N_i}^t$ represents the neighbourhood configuration of $x_i^{(t)}$.

5. By using $P^{(t)}\left(l|y_i\right)$, update the parameters as $\mu_l^{(t+1)}$ and $\left(\sigma_l^{(t+1)}\right)^2$.

2.4.1 Morphological Post-processing

The unwanted small foreground areas were discarded using morphological opening operation [11]. The splitting of paired nuclei regions is done by watershed algorithm [11]. Connected components labelling was used to label each independent foreground nucleus. Therefore, the original colour cancerous nucleus can be easily segmented out from the original image using the labelled position (Fig. 3).
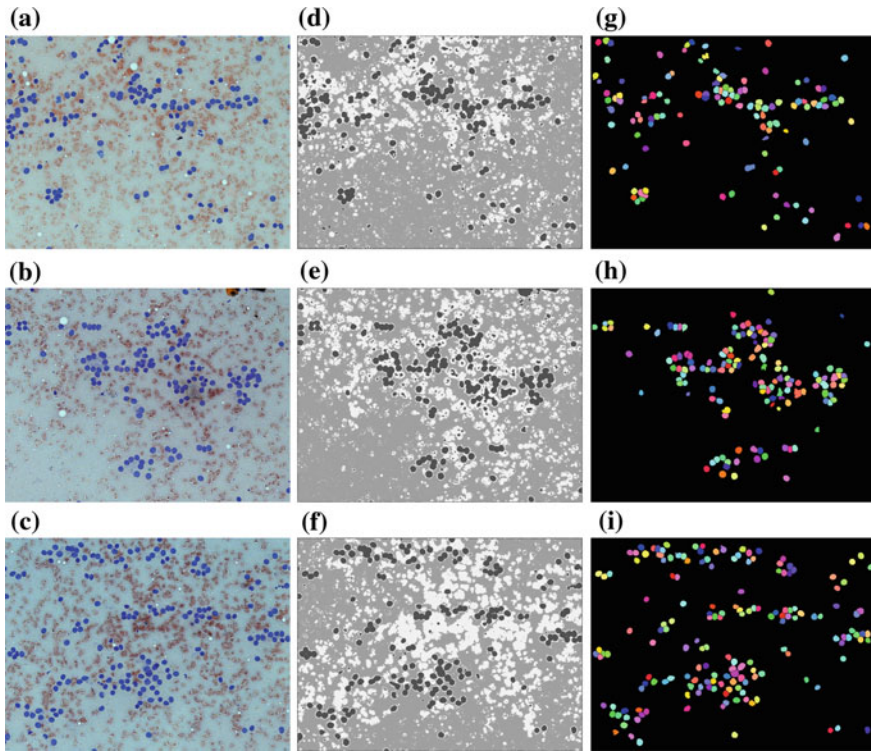
Fig. 3 **a–c** Original image (objective magnification *20x*), **d–f** Corresponding clustered image, **g–i** Corresponding nucleus segmented label image

2.5 Nucleus Classification

The adopted classification model is derived from well known document classification algorithm called Bag of Words (BOW) [12]. The major parts of BOW model are described in the below section.

2.5.1 SIFT Feature Extraction

The Scale-Invariant Feature Transforms (or SIFT) feature in image, provide an object feature set without being affected by any complications like rotation or object scaling [13]. To extract these features, the SIFT approach uses the 4-stage filtering technique, viz. scale-space extrema detection, keypoint localization, orientation assignment and keypoint descriptor.

2.5.2 K-Means Clustering

In K-means, sample objects are divided into $k = 10$ number of clusters [14]. Initially, a cluster centre is selected randomly for each cluster. The goal is to minimize the distance between the points and assigned centroid for each clusters. In an iterative process, the cluster centroids are modified and later fixed after it converged. The membership function used to calculate the distance is as follows:

$$arg \min_{C} = \sum_{c=1}^{k} \sum_{X \in C_i} \|X - \mu_i\|^2,$$ (6)

where input object is $X = \{x_1, x_2, , x_3, ..., x_n\}$, μ is centroid of each cluster C_i and C denotes total dataset divided into $\{C_1, C_2, ..., C_k\}$ clusters.

2.5.3 Construction of Visual Dictionary

Here, an image is represented by SIFT feature keypoints. Therefore, each individual image can be described as weighted combination of k-clustered feature descriptors or visual words. In other way, each image could be represented by a frequency histogram of visual words. An illustrative visualization of the above described process is presented in the Fig. 4.

2.5.4 Random Forest Training

After generating the visual dictionary, a supervised classifier model, i.e. random forest was trained with the developed dictionary. However, during training, the prior information about the class of each nucleus was also provided to the classifier. Ran-
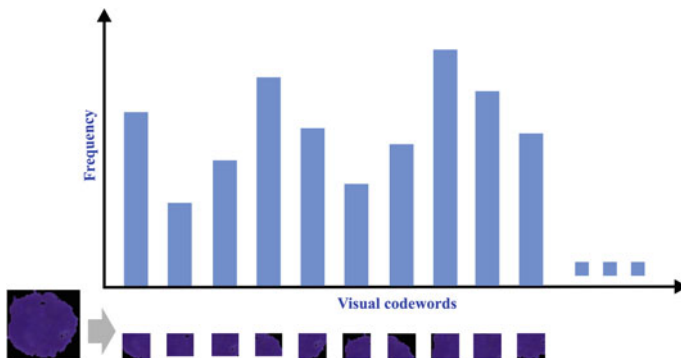
Fig. 4 Bag-of-visual words dictionary

dom Forest [14] is an ensemble learning method for supervised classification problem. It takes a decision by averaging the prediction of multiple decision trees. Bagging or bootstrap aggregating technique was adopted for sub-sampling the dataset with controlling the overfitting.

3 Results

The GMM-HMRF model-based nucleus-segmented images were compared with ground truth images for performance evaluation. Total four segmentation evaluation metrics, viz. Dice coefficient (DC), Jaccard coefficient (JC), Sensitivity (SE) and Specificity (SP) were considered here to validate the results. Also, the nucleus segmentation process passes through multiple image reconstruction phases. It may lead to structural dissimilarity with the original nucleus. Hence, two image quality metrics, viz. Structural Similarity Index measure (SSIM) [15] and Multi-scale Structural Similarity Index Measure (MSSIM) [16] were considered to evaluate the structural quality of segmented nucleus with respect to ground truth nucleus. The segmentation evaluation results are presented in Table 1. After nucleus segmentation, the segmented nucleus were labelled into two classes, i.e. SCLC and NSCLC. Totally 4920 nucleus images were generated where SCLC and NSCLC were distributed in 2:4 ratio. The BOVW model was adopted for nucleus classification. For classification performance analysis, the segmented nucleus were divided into two sets where 60% nucleus images were used for training and rest 40% nucleus image were kept for testing. The SIFT feature was extracted from those training nucleus set, and the visual dictionary was developed from the extracted feature database. The Random Forest classifier model was trained using the dictionary. The trained Random forest was evaluated on the testing nucleus for performance analysis. The confusion matrix for the random forest is shown in Table 2. Based on the results, the BOVW-based random forest classifier model achieved a sensitivity of 97.31%, specificity of 99.54%, the accuracy of 98.78%, F1-score of 98.19% and Mathews correlation coefficient of 97.28%. After successful lung cancer nucleus classification, the same image set was given to two pathologists. The degree of experts agreement was calculated by Cohen's Kappa index [17] and shown in Table 3. The Kappa score is measured

Table 1 GMM-HMRF-based nucleus segmentation performance evaluation

Image	JC	DC	SE	SP	SSIM	MSSIM
Image 1	0.9831	0.9878	0.9882	0.9758	0.9964	0.9918
Image 2	0.9867	0.9731	0.9885	0.9825	0.9921	0.9931
Image 3	0.9887	0.9817	0.9897	0.9795	0.9949	0.9978

Table 2 Confusion matrix of random forest classifier

Actual class	Predicted class		
	SCLC	NSCLC	Row Total
SCLC	650	6	656
NSCLC	18	1294	1312
Column total	668	1300	1968

Table 3 Calculation of Kappa scoring

Sample: 706		First expert		Total
		SCLC	NSCLC	
Second expert	SCLC	654	2	656
	NSCLC	5	1307	1312
Total		659	1309	1968
Agreement		654	1309	1963
By chance		219.59	872.61	1092.2

Table 4 Comparative study between the proposed study with prior works

Literature	Performance
Bora et al. [2]	Accuracy: 98.11%, Precision: 98.38% in smear level and 99.01% in cell level
Bora et al. [3]	Accuracy: 90–95%
Wang et al. [18]	Accuracy: 96.16%, SE: 99.05%, SP: 93.33%
Proposed	Segmentation- SE: 98.88%, SP: 97.93%, Classification- SE: 97.31%, SP: 99.54%

99.2%. A comparative study was also made where the performance of previously reported works were compared with the proposed work. The study is shown Table 4. Based on the results, the proposed methods works superior than other referenced works.

4 Conclusion

A fully automated framework for nucleus classification of SCLC and NSCLC from Pap-stained cytological image is proposed here. This study combines two different methods where unsupervised Gaussian process is adopted for nucleus segmentation and vocabulary based supervised random forest classifier is used for nucleus classification. The segmentation process achieves sensitivity of 98.88% and specificity of 97.93%. Also, the nucleus classification model was able to perform with sensitivity

of 97.31%, specificity of 99.54%, and accuracy of 98.78%. The Kappa score is measured 99.2%. The proposed algorithm can be implemented in a CAD tool, which will definitely assist experts for lung cancer diagnosis.

References

1. Siegel, R.L., Miller, K.D., Jemal, A.: Cancer statistics, 2015. CA Cancer J. Clin. **65**(1), 5–29 (2015)
2. Bora, K., Chowdhury, M., Mahanta, L.B., Kundu, M.K., Das, A.K.: Automated classification of pap smear images to detect cervical dysplasia. Comput. Methods Progr. Biomed. **138**, 31–47 (2017)
3. Bora, K., Chowdhury, M., Mahanta, L.B., Kundu, M.K., Das, A.K.: Pap smear image classification using convolutional neural network. In: Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, p. 55. ACM (2016)
4. Abdel-Zaher, A.M., Eldeib, A.M.: Breast cancer classification using deep belief networks. Expert Syst. Appl. **46**, 139–144 (2016)
5. Kashyap, A., Jain, M., Shukla, S., Andley, M.: Study of nuclear morphometry on cytology specimens of benign and malignant breast lesions: a study of 122 cases. J. Cytol. **34**(1), 10 (2017)
6. Kecheril, S.S., Venkataraman, D., Suganthi, J., Sujathan, K.: Automated lung cancer detection by the analysis of glandular cells in sputum cytology images using scale space features. Signal Image Video Process. **9**(4), 851–863 (2015)
7. Yu, K.H., Zhang, C., Berry, G.J., Altman, R.B., Ré, C., Rubin DL, Snyder M (2016) Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat. Commun. **7**
8. Javaid, M., Javid, M., Rehman, M.Z.U., Shah, S.I.A.: A novel approach to cad system for the detection of lung nodules in ct images. Comput. Methods Progr. Biomed. **135**, 125–139 (2016)
9. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: 1998. Sixth International Conference on Computer Vision, pp 839–846. IEEE (1998)
10. Wang, Q.: Gmm-based hidden markov random field for color image and 3d volume segmentation (2012). arXiv:12124527
11. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital Image Processing Using MATLAB. McGraw Hill Education (2010)
12. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: Workshop on Statistical Learning in Computer Vision, ECCV, Prague, vol. 1, pp. 1–2 (2004)
13. Lowe, D.G.: Object recognition from local scale-invariant features. In: 1999 Proceedings of the Seventh IEEE International Conference on Computer Vision, vol 2, pp. 1150–1157. IEEE (1999)
14. Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann (2016)
15. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, vol 13, no 4, pp. 600–612 (2004)

16. Rouse, D.M., Hemami, S.S.: Analyzing the role of visual structure in the recognition of natural image content with multi-scale SSIM, Proc. SPIE Vol. 6806, Human Vision and Electronic Imaging (2008)
17. Byrt, T., Bishop, J., Carlin, J.B.: Bias, prevalence and kappa. J. Clin. Epidemiol. **46**(5), 423–429 (1993)
18. Wang, P., Hu, X., Li, Y., Liu, Q., Zhu, X.: Automatic cell nuclei segmentation and classification of breast cancer histopathology images. Signal Process. **122**, 1–13 (2016)

Assessment of Segmentation Techniques for Chronic Wound Surface Area Detection

Maitreya Maity, Dhiraj Dhane, Chittaranjan Bar, Chandan Chakraborty and Jyotirmoy Chatterjee

Abstract A skin ulcer is a clinical pathology of localized damage to skin and tissue instigated by venous insufficiency. Precise identification of wound surface area is one of the challenging tasks in the dermatological evaluation. The assessment is carried out by clinicians using traditional approach of scales or metrics through visual inspection. The manual assessment leads to intra-observer variability, subjective error and time complexity. This paper evaluates the performances of supervised and unsupervised segmentation techniques used for wound area detection. The unsupervised methods used for evaluation were namely K-means, Fuzzy C-means and Gaussian mixture model. On the other part, random forest was implemented for supervised classification. Several filtering methods were used to generate image feature set from wound images to train random forest. The Gaussian mixture model with classification expectation–maximization clustering method achieved the highest weighted sensitivity of 95.91% and weighted specificity of 96.7%. The comparative study shows the superiority of proposed method and its suitability in wound segmentation from normal skin.

Keywords Chronic wound · Segmentation · Clustering · Classification Performance

M. Maity (✉) · D. Dhane · C. Chakraborty · J. Chatterjee
School of Medical Science and Technology, Indian Institute of Technology Kharagpur,
Kharagpur, West Bengal, India
e-mail: maitreya.maity@gmail.com

C. Bar
Department of Dermatology, Midnapore Medical College,
Midnapore, West Bengal, India
e-mail: drcrbar@gmail.com

1   Introduction

A chronic ulcer is one of the critical global healthcare problems. From a survey, it is expected that worldwide 380 million people will suffer from it [1]. According to United Nations survey in 2009, it is reported that chronic wounds affect more than 2 billion European populations, and related annual treatment cost is nearly 8 billion. About 4.5 per 1000 population suffers from chronic wound in Indian population [2].

Traditionally, the wound is evaluated by simple visual grading and spatial measurements [3–6]. However, the human vision lacks precision and consistency. Hence, the determination of slow and minute changes appears to be difficult [7]. Moreover, accurate wound measurement is an essential task in chronic wound treatment as wound contraction and change in tissue composition are indicators of the healing progression [8]. Currently, clinicians depend on traditional contact-type wound assessment method consisting of ruler and transparency tracing [9]. These methods are less precise, tedious, and time-consuming.

Precise wound surface area or boundary identification is one of the most important tasks reported in the literature. Duckworth et al. [10] used laser light for marking wound boundaries thereby measuring wound area. They used iterative edge detection algorithm for wound boundary detection. Few studies have reported wound segmentation using region growing technique [11], and differential evolution clustering algorithm [12]. Unsupervised segmentation algorithms such as colour arrangement code, graph-based image segmentation, mean shift, and J-SEG method was compared [13]. It was found that J-SEG outperforms others with a segmentation accuracy of 73.1%. Dhane et al. [14] have applied spectral approach for clustering to segment wound from non-wound region using optical wound images with a segmentation accuracy of 86.73%. Kolesnik and Fexa [15] have established a 3D colour histogram-based SVM classifier to segment wound region from image. Treuillet et al. [16] have developed 3D regeneration model from standard 2D images to the mimic wound area and volume. Wannous et al. [17] have computed 3D model for wound measurements using computer vision techniques. Veredas et al. [18] have reported a machine vision method for ulcer circumference detection using statistical features in the color models.

Both, unsupervised and supervised techniques were explored previously by many researchers for wound boundary detection. However, a thorough investigation to find the best approach for identifying wound area is required. The present study has evaluated both supervised and unsupervised segmentation approaches to determine the best-suited method for wound segmentation. In the following sections, three clustering methods and one supervised classification technique are described.

2 Materials and Methods

2.1 Wound Image Collection

The chronic ulcer images were collected from Midnapur Medical College and Hospital, India (MMCH). Institutional ethical clearance was taken for this study. The distance between camera lens and wounds surface was approximately 15–20 cm. The dimensions of each raw image are 4320×3240 pixels; however, the images were manually cropped to get maximum wound area with the minimum background. The acquired images were homogenized for colour intensity variations using colour correction software. Some images were also taken from online library named *Medetec* [19]. During Sample collection, total 68 images were captured from 11 patients with diabetic ulcer, 24 images from 4 patients with pressure ulcer, 34 images from 6 patients with tropical ulcer and 124 images from 21 patients with surgical wounds.

2.2 Multi-expert Wound Region Labelling

The images were given to three clinicians for generating the ground truth image. The ground truth was reproduced by marking the exact wound periphery. The periphery marking was made final only if two out of three experts agrees upon the exact periphery. Few examples of multi-expert labelling and weighted ground truth image are presented in Fig. 1.

2.3 Clustering Approach for Wound Region Segmentation

2.3.1 K-Means Clustering

K-means clustering [20] is a data clustering method, where each data item belongs to a cluster with minimum distance from mean of the cluster. The membership function used to calculate the distance is as follows:

$$W = \sum_{i=1}^{M \times N} \sum_{k=1}^{K} \left\| I_i - \mu_k \right\| \tag{1}$$

where I_i is the pixel in the image ($M \times N$) and μ_k is the cluster center for kth cluster. Membership function W at each iteration reaches a minimum threshold.
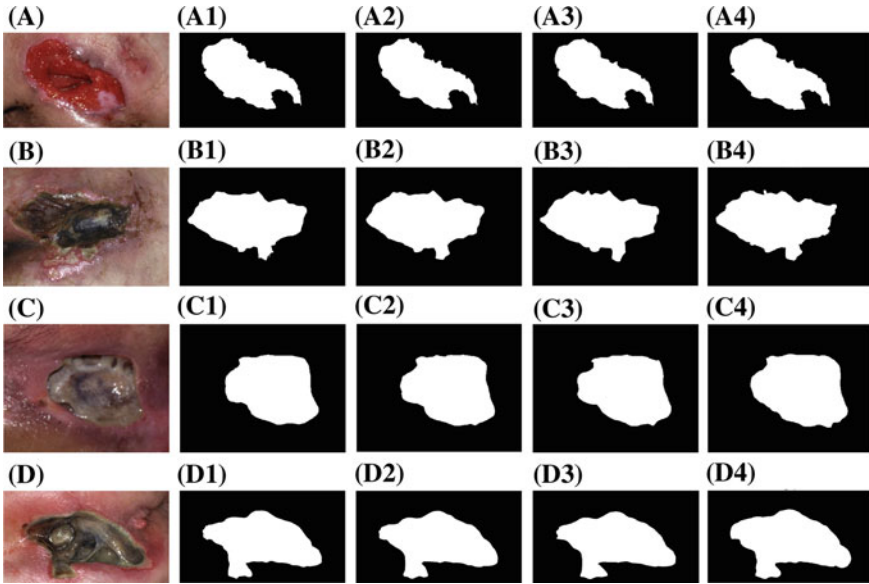
Fig. 1 (A–D) Original image, (A1:D1) labelled by expert 1, (A2:D2) labelled by expert 2, (A3:D3) labelled by expert 3, (A4:D4) weighted final ground truth image

2.3.2 Fuzzy C-Means Clustering

Fuzzy C-means (FCM) [20] is a data clustering algorithm where each data item (univariate or multivariate) belongs to every cluster with a certain weight. In FCM based segmentation process, every pixel value associated with a fuzzy membership value. Let I be an image of dimension $M \times N$, where pixels $I_{i,j}$ are to be divided into c clusters. $\|*\| \to \mathbb{R}$ is any norm expressing the Euclidean similarity between any pixel value and the cluster centre. The cost function defined for the iterative optimization is defined by

$$W = \sum_{i=1}^{M \times N} \sum_{k=1}^{C} (\mu_{ki}^{w'}) \left\| I_{k,i} - M_k \right\|^2, \tag{2}$$

where μ_{ki} represents the membership of the ith pixel for kth cluster. The parameter w' is the weighing exponent controlling the fuzziness of the cluster. $M = m_1, m_2, ..., m_K$ are the cluster centres. Each iteration updates the membership function value and cluster center as,

$$\mu_{ki} = \frac{1}{\sum_{k=1}^{C} \left\{ \frac{\|I_{ki} - M_i\|}{\|I_{ki} - M_k\|} \right\}^{\frac{2}{w'} - 1}} \; ; \; M_i = \frac{\sum_{i=1}^{M \times N} \mu_{ki}^{w'} I_{ki}}{\sum_{i=1}^{M \times N} \mu_{ki}^{w'}} \tag{3}$$

2.3.3   Gaussian Mixture Model

The Gaussian Mixture Model (GMM) [20] is a parametric probability density function characterized as a weighted sum of linear Gaussian density functions. It can be defined as

$$p(x \mid \theta) = \sum_{i=1}^{M} w_i G(x \mid \mu_i, \Sigma_i), \tag{4}$$

where x is the D-dimensional data vector, $w_{i=1,2,\ldots,M}$ are the mixture weights, and G are the Gaussian models where the mean μ and covariance matrix Σ are components of Gaussian models. The total of mixture weights is always 1. Each D-variate Gaussian density function can be formed as

$$G(x \mid \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{D}{2}} \, |\Sigma_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(x - \mu_i)' \Sigma (x - \mu_i) \right\} \tag{5}$$

GMM determines the unknown parameters θ. Expectation Maximization (EM) and Classification Expectation Maximization (CEM) [21], parameter estimation algorithm are considered in the this study to find θ. In the proposed study three GMM, i.e. GMM with EM (GMM-EM), GMM with CEM (GMM-CEM) and multivariate GMM with CEM (MGMM-CEM) were considered. Here, two number of clusters for K-means and FCM and two Gaussian models were considered.

2.4   Classification Approach for Wound Region Segmentation

2.4.1   Feature Extraction

Multiple filtering algorithms were considered and applied on the original colour images to generate corresponding filtered images [22]. Then, the pixel values for each spatial position were extracted from all filtered images. Accordingly, the extracted pixel values were categorized based on their ground truth label. This feature extraction model was applied on all images to develop a labelled feature database. The mechanism of the proposed feature extraction model is described in the Fig. 2. The details of the considered 79 features are described in Table 1.

2.4.2   Random Forest Classifier

Random Forest (RF) [20, 23] is an ensemble estimator where multiple decision trees fitted to take an average decision. Multiple decision trees are performed on subsamples of the dataset to improve the predictive accuracy and control overfitting. The equal sized subsamples are drawn with replacement or bootstrap aggregating.
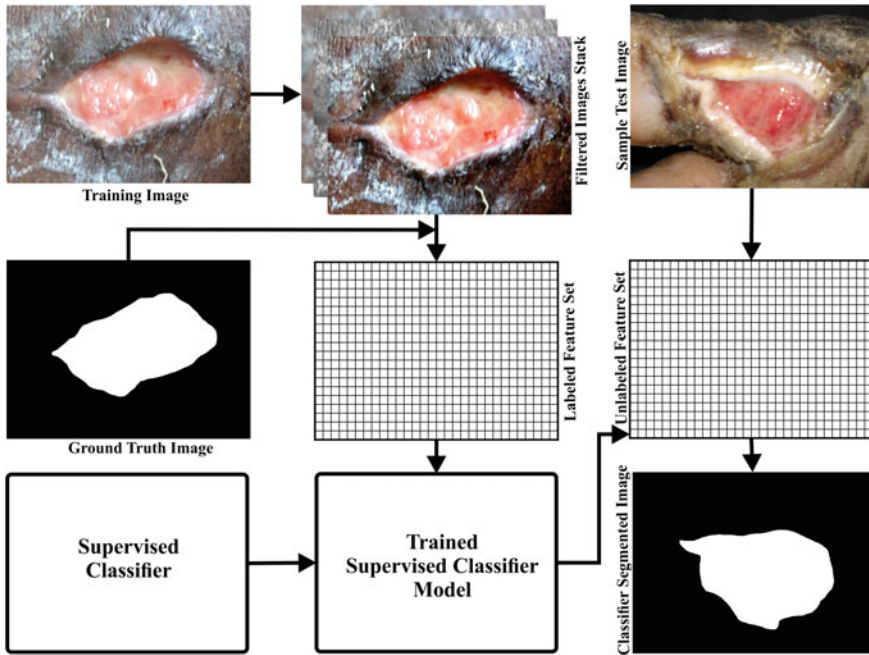
Fig. 2 Abstract mechanism of the supervised classifier-based segmentation

Table 1 Details of feature extraction algorithms for classification

Feature	Description	Count
Gaussian blur	Blurred versions of varying $\sigma = 1, 2, 4, 8, 16$ are generated to compute pixel value at each point. Blurring will help to neighbour pixels homogeneous	5
Sobel filter	Calculate gradient at each pixel using Sobel filtering with varying $\sigma = 0, 1, 2, 4, 8, 16$	6
Hessian matrix	Calculate a Hessian matrix at each pixel and compute module, trace, determinant, first eigenvalue, second eigenvalue, orientation, Gamma-normalized square eigenvalue difference and square of Gamma-normalized eigenvalue difference	48
Membrane detectors	19×19 Zero matrix kernel is rotated in 30 positions. Then it is convolved with the image to generate a Z-projected single image of six types. (sum, mean, standard deviation, median, maximum, and minimum of the pixels in each image)	6
Gaussians difference	Calculates two Gaussian blur images from the original image and subtracts one from the other	10
Colour	Hue, Saturation, Brightness and RGB average	4

2.4.3 Post-processing

Due to oversegmentation or false classification, unwanted artefacts persist in the segmentation outputs. Using morphological post-processing like area opening, area closing and binary filling,such noises were eliminated.

3 Results and Discussion

Both supervised and unsupervised techniques were evaluated for wound surface area segmentation. The accuracy of the considered algorithms was performed by four performance metrics, viz. Sensitivity (SE), Specificity (SP), F-Measure (FM) and Jaccard score (JC). The metrics were computed by comparing the ground truth and segmented binary image. The comparative analysis results were reported in Table 2, where performance measures of four random images were mentioned. The segmented image outputs were also shown in the Fig. 3. For classifier-based segmentation process, a training feature dataset was developed. The statistics of some important features are presented in Table 3. Based on the comparison results, MGMM-CEM gave the highest accuracy. We have applied the considered clustering methods on the developed database. Based on average results, the selected MGMM-CEM clustering model achieved weighted 95.91% sensitivity and weighted 96.7% specificity to segment the wound region. The estimated parameter for MGMM-CEM is as follows,

Table 2 Performance evaluation of clustering methods

Image	Metric	K-means	FCM	GEM	CEM	MCEM	RF
I1	**JC**	0.9147	0.9142	0.9665	0.9648	**0.9818**	0.9541
	SE	0.9172	0.9167	0.9784	0.9743	**0.9968**	0.9796
	SP	0.9947	0.9947	0.9765	0.9811	**0.9708**	0.9913
	FM	0.9554	0.9552	0.9829	0.9821	**0.9908**	0.9765
I2	**JC**	0.8258	0.9614	0.5640	0.9613	**0.9625**	0.8878
	SE	0.9843	0.9841	0.6933	0.9840	**0.9775**	0.9364
	SP	0.7658	0.9712	0.72	0.9712	**0.9809**	0.9811
	FM	0.9046	0.9803	0.7212	0.9782	**0.9809**	0.9405
I3	**JC**	0.9781	0.9781	0.9702	0.9702	**0.9748**	0.9102
	SE	0.9936	0.9936	0.9971	0.9972	**0.9967**	0.9394
	SP	0.9867	0.9867	0.9767	0.9767	**0.9812**	0.9891
	FM	0.9889	0.9889	0.9849	0.9849	**0.9872**	0.9530
I4	**JC**	0.8213	0.8214	0.8088	0.8107	**0.8259**	0.8155
	SE	0.8342	0.8352	0.8463	0.8451	**0.8352**	0.9743
	SP	0.9797	0.9799	0.9019	0.9101	**0.9865**	0.9804
	FM	0.9055	0.9055	0.8943	0.8954	**0.9077**	0.9067
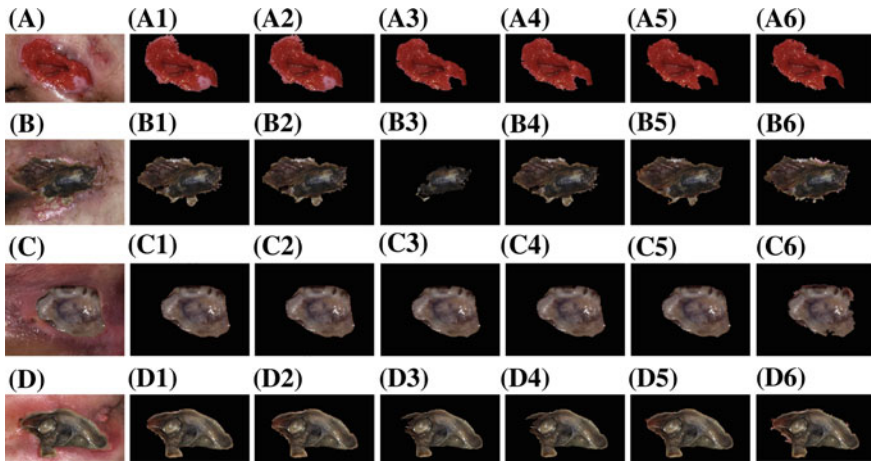
Fig. 3 Segmentation results using clustering and classification methods [A–D: Original image, A1–D1: K-means, A2–D2: Fuzzy C-means, A3–D3: GMM-EM, A4–D4: GMM-CEM, A5–D5: MGMM-CEM, A6–D6: Random Forest

Table 3 Statistics of few significant features

Feature	Skin region (Class 1)	Wound region (Class 2)
Saturation	108.58 ± 26.516	193.76 ± 22.698
Hue	22.64 ± 58.76	12.25 ± 50.09
Gaussian Blur ($\sigma = 1$)	105.66 ± 31.25	65.17 ± 19.54
Gaussian Blur ($\sigma = 2$)	105.66 ± 31.09	65.19 ± 18.05
Gaussian Blur ($\sigma = 4$)	105.66 ± 30.89	65.21 ± 16
Gaussian Blur ($\sigma = 16$)	105.63 ± 29.95	65.43 ± 12.17
Diff Gaussian ($\sigma = 16, 8$)	0.92 ± 1.65	1.29 ± 3.42
Diff Gaussian ($\sigma = 8, 4$)	0.75 ± 1.35	1.2 ± 4.38
Sobel filter ($\sigma = 1$)	22.41 ± 18.09	31.03 ± 45.22
RGB average	105.65 ± 31.26	65.18 ± 19.83
Sobel filter ($\sigma = 4$)	12.93 ± 10.45	19.34 ± 28.57
Brightness	140.68 ± 33.23	127.33 ± 22.151

$$w = [0.2, 0.7] \, ; \mu = \begin{bmatrix} -2.5 & 0.7 \\ 4.2 & -1.2 \\ 4.3 & -1.2 \\ -0.3 & 0.1 \\ -0.1 & 0.04 \end{bmatrix} ; \sum_{Nucleus} = \begin{bmatrix} 0.8 & 0.5 & 0.3 & -0.02 & 0.04 \\ 0.5 & 0.8 & 0.9 & -0.04 & -0.01 \\ 0.3 & 0.9 & 1.6 & -0.02 & 0.02 \\ -0.02 & -0.04 & -0.02 & 0.2 & 0.1 \\ 0.04 & -0.01 & 0.02 & 0.1 & 0.3 \end{bmatrix} ;$$

$$\sum_{Background} = \begin{bmatrix} 7.1 & -0.8 & -1.2 & 0.7 & -1.1 \\ -0.8 & 2.9 & 0.5 & 0.1 & -0.2 \\ -1.2 & 0.5 & 2.5 & -0.1 & 0.1 \\ 0.1 & 0.1 & -0.1 & 0.9 & -0.06 \\ -1.1 & -0.2 & 0.1 & -0.06 & 0.9 \end{bmatrix}$$

4 Conclusion

In this paper, supervised and unsupervised segmentation techniques were evaluated to validate their performances in skin ulcer area detection. The captured wound images with uncontrolled illumination were corrected for colours using 24-bit colour corrector. The colour constancy image was then labelled by multiple experts for exact wound boundary. The K-means, fuzzy C-means and variation of Gaussian mixture models were applied on the colour homogenized images. Meanwhile, the image features were extracted by applying several filtering techniques. The random forest-supervised classifier was trained to classify the wound region. Based on the comparison, Gaussian mixture model performed best with highest accuracy. This comparative study elucidates the superiority of the GMM in delineating the wound boundary, as compared with the state-of-the-art methods in wound bed detection. The outcomes of this assessment study will help to choose target region segmentation method for the development of a complete automated wound diagnosis algorithm.

References

1. Cho, N.H., Whiting, D., Guariguata, L., Montoya, P.A., Forouhi, N., Hambleton, I., et al.: IDF Diabetes Atlas. International Diabetes Federation, Brussels, Belgium (2013)
2. Kailas, A., Chong, C.C., Watanabe, F.: From mobile phones to personal wellness dashboards. IEEE Pulse **1**(1), 57–63 (2010)
3. Kecelj Leskovec, N., Perme, M.P., Jezeršek, M., Mozina, J., Pavlović, M.D., Lunder, T.: Initial healing rates as predictive factors of venous ulcer healing: the use of a laser-based three-dimensional ulcer measurement. Wound Repair Regen. **16**(4), 507–512 (2008)
4. Lubeley, D., Jostschulte, K., Kays, R., Biskup, K., Clasbrummel, B.: 3D wound measurement system for telemedical applications. Biomedizimische Technik **50**(1), 1418–19 (2005)
5. Chang, A.C., Dearman, B., Greenwood, J.E., et al.: A comparison of wound area measurement techniques: visitrak versus photography. Eplasty **11**(18), 158–66 (2011)
6. Little, C., McDonald, J., Jenkins, M., McCarron, P.: An overview of techniques used to measure wound area and volume. J. Wound Care **18**(6), 250–253 (2009)
7. Pavlovčič, U., Diaci, J., Možina, J., Jezeršek, M.: Wound perimeter, area, and volume measurement based on laser 3D and color acquisition. Biomed. Eng. Online **14**(1), 1 (2015)
8. Hansen, G.L., Sparrow, E.M., Kokate, J.Y., Leland, K.J., Iaizzo, P.A.: Wound status evaluation using color image processing. IEEE Trans. Med. Imaging **16**(1), 78–86 (1997)
9. Krouskop, T.A., Baker, R., Wilson, M.S.: A noncontact wound measurement system. J. Rehabil. Res. Dev. **39**(3), 337 (2002)
10. Duckworth, M., Patel, N., Joshi, A., Lankton, S.: A clinically affordable non-contact wound measurement device (2007)
11. Mesa, H., Veredas, F.J., Morente, L.: A hybrid approach for tissue recognition on wound images. In: International Conference on Hybrid Intelligent Systems, pp. 120–125. IEEE (2008)
12. Aslantas, V., Tunckanat, M.: Differential evolution algorithm for segmentation of wound images. In: IEEE International Symposium on Intelligent Signal Processing, pp. 1–5. IEEE (2007)

13. Deng, Y., Manjunath, B.: Unsupervised segmentation of color-texture regions in images and video. IEEE Trans. Pattern Anal. Mach. Intell. **23**(8), 800–810 (2001)
14. Dhane, D.M., Krishna, V., Achar, A., Bar, C., Sanyal, K., Chakraborty, C.: Spectral clustering for unsupervised segmentation of lower extremity wound beds using optical images. J. Med. Syst. **40**(9), 207 (2016)
15. Kolesnik, M., Fexa, A.: Multi-dimensional color histograms for segmentation of wounds in images. In: International Conference on Image Analysis and Recognition, pp. 1014–1022. Springer (2005)
16. Treuillet, S., Albouy, B., Lucas, Y.: Three-dimensional assessment of skin wounds using a standard digital camera. IEEE Trans. Med. Imaging **28**(5), 752–762 (2009)
17. Wannous, H., Lucas, Y., Treuillet, S.: Enhanced assessment of the wound-healing process by accurate multiview tissue classification. IEEE Trans. Med. Imaging **30**(2), 315–326 (2011)
18. Veredas, F.J., Mesa, H., Morente, L.: Efficient detection of wound-bed and peripheral skin with statistical colour models. Med. Biol. Eng. Comput. **53**(4), 345–359 (2015)
19. Medetec Medical Images: http://www.medetec.co.uk/files/medetec-images.html (2016)
20. Alpaydin, E.: Introduction to Machine Learning, 2nd edn. The MIT Press (2010)
21. Li, J.: Clustering based on a multilayer mixture model. J. Comput. Graph. Stat. **14**(3), 547–568 (2005)
22. Arganda-Carreras, I., Kaynig, V., Schindelin, J., Cardona, A., Seung, H.: Trainable weka segmentation: a machine learning tool for microscopy image segmentation (2014)
23. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)

Improved Decision-Making for Navigation Scenarios

Khyati Marwah and J. S. Sohal

Abstract Nowadays, intelligent systems are quite common to assist the driver in its navigation chores. But, if accurate information is not provided to the driver on time then it is of no use and the delay can even lead to disastrous consequences. To improve behavioral realism in real life navigation scenarios we need intelligent and speedy decisions. Such decisions rely on the information being extracted from real-time images. We need to select the best available features/information that can be used as neural network inputs and ultimately predict the next move of the vehicle. Little work has been done on selection of best features of navigation images and research needs to be done in this gray area as its results have direct impact on the classification accuracy and generalized performance of automotive navigation planning. This article proposes a novel approach to find the best possible set of features from real-time navigation images by using Boruta Algorithm and Earth Algorithm so as to improve the prediction power of vehicle to either move or stop in the next course of action. The results obtained were cross-validated by using three classifiers: Logistic Regression, Support Vector Machines (SVM), and K-Nearest Neighbor on the basis of parameters: Classification accuracy and prediction performance. It was identified that the proposed model maintained the classification accuracy and performed with more superiority by getting rid of the irrelevant features of images and thereby reducing the training time as compared to

K. Marwah (✉)
IKG Punjab Technical University, Kapurthala, India
e-mail: er.khyati@gmail.com

J. S. Sohal
LCET, Ludhiana, Punjab, India
e-mail: jssohal2001@yahoo.co.in

computations done on the basis of original feature set. The prediction speed of the proposed model was found to be much better than the model without feature selection. The accuracy of this novel approach was also found to be improved by few ensembles using Generalized Linear Model (GLM) wrapper.

Keywords Decisions · Navigation · SVM · Logistic regression
KNN · GLM

1 Introduction

Decision-making systems are needed to deal with the uncertainty and imprecision of the day-to-day real-world problems. We need to take a number of decisions related to navigation while commuting in real time. In its simplest form of navigation, a moving vehicle needs to decide its next course of action which can be either move or stop. This computation is based on the information being obtained from a series of real-time navigation images. Extracting the best informative features needed for the same is a challenging task. The objective of feature selection is three-fold [1]: improving the prediction performance of the predictors, providing faster and more cost-effective predictors as it reduces the measurements and the storage requirements, and providing a better understanding of the underlying process that generated the data by facilitating better data visualization. Feature selection algorithms are divided into three categories namely filter methods, wrapper methods and embedded methods [2]. This research paper presents a novel approach to find the best possible set of features from real-time navigation images so as to improve the prediction performance of vehicle.

2 Database Acquisition

The database of around hundred images of lanes located in suburban area was collected for this study. Most of the existing systems have used simulators or freely available datasets for the purpose of training and testing of their prediction models. But the proposed system tries to induce realism by considering real-time images of suburban Indian straight and curved roads. The data was assembled by clicking images and videos of busy roads in the vicinity of authors working place. Thus the dataset includes the photorealistic intermingling images of buildings, vehicles, tree markings etc. Image processing techniques [3] are needed to improve the data quality and visual appearance of database images.

3 Proposed Algorithm for Improved Decision Making in Navigation

It compromises of three steps as follows:

3.1 Image Processing Tasks

First and foremost task is to get database images ready for feature extraction. For this, we convert the RGB image into grayscale as comparing the sections in such images is quite simpler and their intensity is usually sufficient to get edges of objects in images. The grayscale image is then passed through the segmentation process. At this step, the image is partitioned into multiple segments so as to change the representation of input image into more meaningful form as per our interest. Segmented regions are clustered by using K-Mean Clustering in order to identify the occupied and unoccupied areas of the processed images. Here the images are partitioned into K mutually exclusive clusters so that objects within each cluster are as close to each other as possible and at the same time, as far as possible from objects of other clusters. The process is then succeeded by masking stage in which a mask is prepared to separate the region of interest from its background. The last step is to perform Morphological Processing which is required to test whether the element fits or hits the neighborhood. These preprocessing steps remove the imperfections of input images and thus make the basis for extraction of features and thereby creating a feature vector.

3.2 Feature Extraction and Creation of Feature Vector

An intelligent navigation system needs to take a decision by perceiving the input image just as the real human eye does, so we need to explore minute details of the image. These minute properties make the basis for classifying the various components of image and thus enable us to extract useful information needed to make decisions. The required information can be best explored by the texture analysis of an image.

Texture Analysis: Texture analysis [4] defines the exact texture of image. It is a set of metrics related to the spatial arrangement of color or intensities of an image which makes it easier to quantify the different gray levels within an image in terms of intensity and distribution. In this research paper, we will be using a number of matrices for detailed feature analysis of database images. A brief description of this matrix along with their list of metrics is given in Table 1. The list of features obtained after Texture Analysis along with their description is given in Table 2.

Table 1 Matrix's for texture analysis: description and metrics obtained

S. No.	Name of texture features matrix used	Description	Metrics calculated
1	SGLDM: Spatial Gray-level Dependence Matrices [5]	Second-Order Statistical Method Focuses on spatial relationship and distribution of different gray-level regions of interest	Angular Second Moment Contrast Correlation Homogeneity Entropy
2	GLDS: Gray-level Difference Statistics	Second-Order Statistical Method Estimates gray-level difference in images	Angular Second Moment Contrast Correlation Homogeneity Entropy
33	FOS: First-Order Spectrum	Deals with First-Order Statistics over the input image	Mean Variance Skewness Kurtosis
4	FPS: Fourier Power Spectrum	Based on Power Spectrum over Fourier Transform	Periodicity
5	SFM: Statistical Feature Matrix	Measures Statistical properties of pixel pair at several distances	Coarseness Contrast Roughness Periodicity
6	TEM: Laws Texture Energy Measure [6]	Statistical measure of amount of variation within fixed size window Involves Average Gray level (L), Edges (E), Spots (S)	TEM at different kernels: LL, EE, SS, LE, ES, and LS
7	FDTA: Fractal-Dimensional Texture Analysis [7]	Fractals refer to self-similar and irregular geometric primitives It gives information about Geometric Structure	Roughness Smoothness Area Solidity

In a similar way, a feature vector was created for 100 input images. Flowchart showing the order of execution of first two steps is shown in Fig. 1

3.3   Computation of the Best Feature Set

Feature Selection is an important aspect of the model building. It helps in making the predictive models free from unwanted correlated variables which may decrease their efficiency. The prediction model for feature selection was trained by inducing

Table 2 Features obtained after texture feature analysis

Attribute	Description
F1	Angular second moment
F2	Contrast
F3	Correlation
F4	Sum of squares
F5	Inverse difference moment (Homogeneity)
F6	Sum average
F7	Sum variance
F8	Sum entropy
F9	Entropy
F10	Difference variance
F11	Difference entropy
F12	Information measures of correlation1
F13	Information measures of correlation2
F14	Mean
F15	Variance
F16	Median
F17	Mode
F18	Skewness
F19	Kurtosis
F20	Energy
F21	Entropy
F22	Radial sum
F23	Angular sum
F24	Coarseness measure
F25	Contrast measure
F26	Periodicity measure
F27	Roughness measure
F28	Texture energy from LL kernel
F29	Texture energy from EE kernel
F30	Texture energy from SS kernel
F31	Average texture energy from LE and EL kernels
F32	Average texture energy from ES and SE kernels
F33	Average texture energy from LS and SL kernels
F34	Fractal dimension texture analysis1
F35	Fractal dimension texture analysis2
F36	Area

response: 1 for move condition and response: 2 for stop condition against the calculated 36 features/predictors. To find out the best decisive features for navigation, we have passed the feature vector obtained in earlier step through execution of following algorithms.
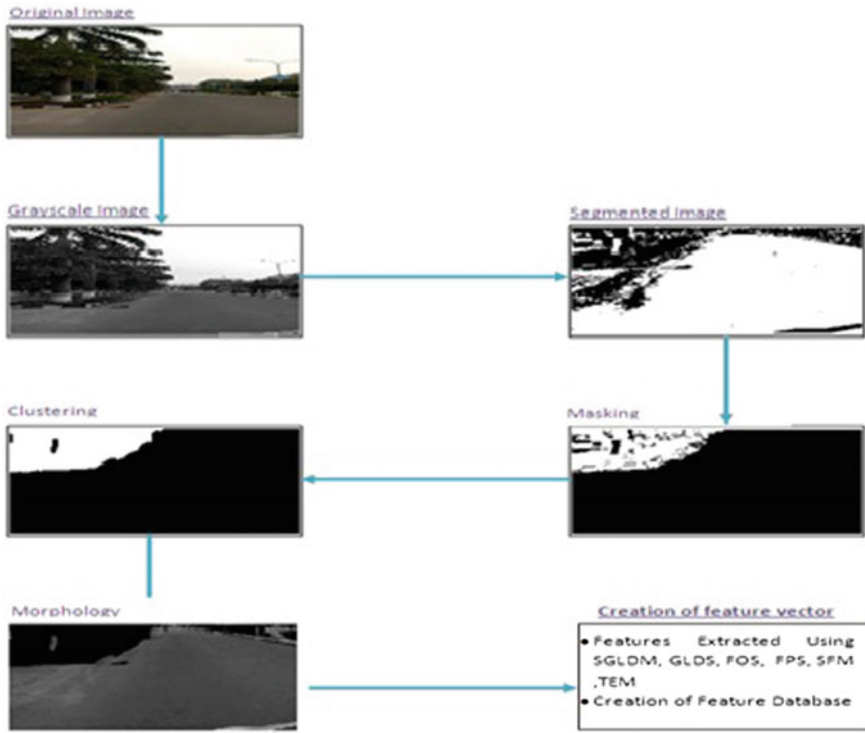
Fig. 1 Flowchart showing execution of step 1 and step 2 of the proposed algorithm

1. Boruta Algorithm [8]: It works as a wrapper algorithm around Random Forest. First, it adds randomness to the given data set by creating shuffled copies of all features (which are called shadow features). Then, it trains a random forest classifier on the extended data set and applies a feature importance measure (the default is Mean Decrease Accuracy) to evaluate the importance of each feature where higher means more important. At every iteration, it checks whether a real feature has a higher importance than the best of its shadow features and constantly removes features which are deemed highly unimportant.
2. Earth: MARS Model [9]: It is multivariate in nature as it able to generate a model based on several input variables Here the flexible models are generated in passes by adjusting the model. The estimation of the relationship between independent and dependent variables is done by regression analysis. It uses a piecewise polynomial function named as Spline.

The execution of Boruta and Earth Algorithm resulted into the selection of some attributes that were considered important by both of them. List of commonly selected important attributes is mentioned in Table 3.

Table 3 Important features
deduced as crux of Boruta and
Earth Algorithms

Attribute	Description of attribute
F3	Correlation
F8	Sum entropy
F9	Entropy
F11	Difference entropy
F12	Information measures of correlation1
F19	Kurtosis
F25	Contrast measure
F31	Average texture energy from LE and EL kernels
F33	Average texture energy from LS and SL kernels
F35	Fractal-dimension texture analysis 2
F36	Area

4 Validation of Proposed Model

To validate the selection of features obtained, we need to identify the proportion of the instances/predictors whose response a classifier can correctly predict. The proposed feature selection model is hereby executed for a number of classifiers. The performance was judged on the basis of **classification accuracy and prediction speed** through the following classifiers and ensembles [10–14]:

Logistic Regression: It uses the logistic function that is the cumulative logistic distribution to estimate the probabilities needed to measure the relation between dependent and independent variables:

Support Vector Machines: It is based on supervised learning methods for outlier's detection. Decision planes are created to find out the decision boundaries

Nearest Neighbor Classifier: It is a nonparametric method for classification. In this, the function is approximated locally and all computations are deferred till classification

Ensembles: It encompasses a weighted composure of a number of a number of classification algorithms in order to achieve better prediction performance than the constituent algorithms.

To determine the accuracy, we need to identify the proportion of instances whose class the classifier can correctly predict. For this, a dataset is taken that consists of instances whose class (stop/move) we already know. We ask the classifier to classify each of these instances in turn. Then we compare its prediction with the actual class of the instance. The proportion of correct classifications acts as an estimate of the accuracy of the classifier. The prediction speed refers to the time taken by classifier during training and testing phases. The output of these classifiers for the original feature vector and the selected feature vector is represented in Table 4. Results of Ensembles: Generalized Linear Model (GLM), Gradient Boosting Model (GBM), and Deep Learning Model (DL) using GLM wrapper are shown in Table 5.

Table 4 Classifier's output

S. N	Model name	Results with original feature vector of 36 attributes			Results with the selected feature vector of 11 attributes		
		Accuracy (%)	Prediction speed (obs/s)	Training time (s)	Accuracy (%)	Prediction speed (obs/s)	Training time (s)
1	**Logistic Regression**	*85*	*440*	*11.048*	*84*	*1100*	*1.9606*
2	**SVMs**						
2.1	Linear SVM	81	480	3.3583	80	1600	1.6415
2.2	Quadratic SVM	85	2000	2.1708	87	990	1.1981
2.3	Cubic SVM	84	2200	2.7081	85	2400	2.9939
	Average Values	83.33	1560	2.7457	84	1663.33	1.9445
3	**Nearest Neighbour**						
3.1	Fine KNN	88	1200	5.4062	86	1200	1.7679
3.2	Medium KNN	80	1300	4.8647	83	650	2.2239
3.3	Coarse KNN	74	510	5.8399	74	590	2.0049
3.4	Cosine KNN	82	1600	5.6082	82	1900	2.1469
3.5	Cubic KNN	81	700	6.9564	82	1200	2.5557
	Average Values	*81*	*1062*	*5.735*	*81.4*	*1108*	*2.1398*

Table 5 Output of ensembles using GLM wrapper

Model name	Accuracy with original feature vector (%)	Accuracy with selected feature vector (%)
Generalized Linear	62.5	**70**
Gradient Boosting	77.5	**81.2**
Deep Learning	76.2	**81.6**

5   Discussions and Conclusions

This research paper has focused to provide a sound judgment with improved response time in navigation scenario to keep abreast with current automotive motion planning by considering the best available decisive features. The proposed
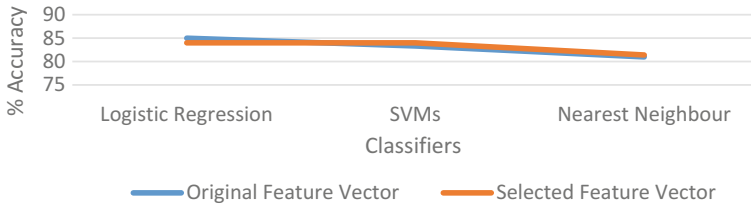
Fig. 2 Comparison of prediction accuracies
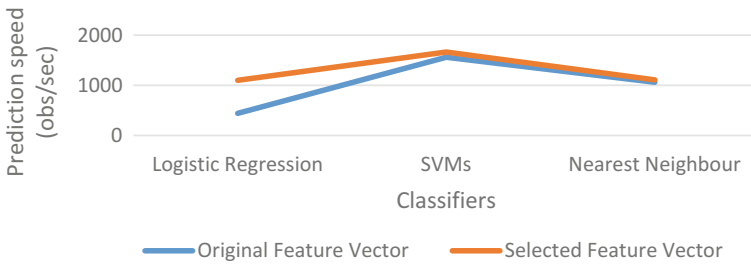
Fig. 3 Comparison of prediction speed

model maintained its accuracy (Fig. 2) in approximately same ration as the model without feature selection. This implies that the discarded features had a low impact on accuracy and were wasting the resources (time and memory). The model performed with more superiority as the prediction speed (Fig. 3) was found to be improved in case of all classifiers. Furthermore, the training time (Fig. 4) required training the feature selection model came out to be comparatively very less than the model without feature selection.

The proposed best feature selection technique could be embedded in the latest automotive technologies and thereby can prove as powerful ensemble by improving the response time (improving prediction speed and decreasing training time) of the navigation system with an average accuracy of around 83%.
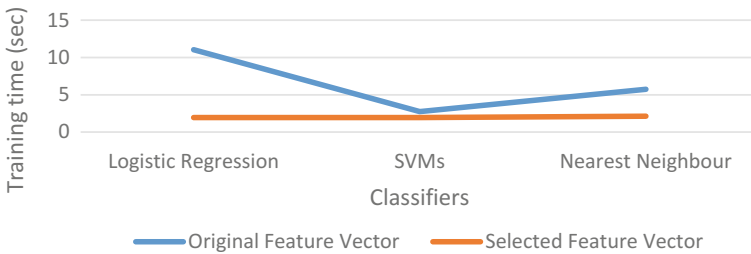
Fig. 4 Comparison of training time

References

1. Guyon, I., Elisseeff A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**, 1157–1182 (2003)
2. https://www.Analyticsvidhya.Com/Blog/2016/12/Introduction-To-Feature-Selection-Methods-With-An-Example-Or-How-To-Select-The-Right-Variables. Accessed 2016/12
3. Krig, S.: Image pre-processing. In: Computer Vision Metrics, pp. 39–83. Springer (2014)
4. Materka, A., Strzelecki, M.: Texture Analysis Methods—A Review, Technical University Of Lodz, Institute Of Electronics, Cost B11 Report, Brussels (1998)
5. Sassi, B., Sellami, L., et. al.: Improved spatial gray level dependence matrices for texture analysis. Int. J. Comput. Sci. Inf. Technol. (IJCSIT) **4**(6) (2012)
6. Elysia, A.S., Welsey, J.: Mammogram classification using law's texture energy measure and neural networks. Procedia Comput. Sci. **59**, 92–97 (2015)
7. Srinivasan, G., Shobha, G.: Statistical texture analysis. In: Proceedings Of World Academy Of Science, Engineering And Technology (2008)
8. Kursa, M., Jankowski, A., et.al.: Boruta-a system for feature selection. Fundamenta Informaticae **101**, 271–285 (2010)
9. Doksum, K., Tang, S., Tsui, K.: Nonparametric variable selection: the Earth algorithm. J. Am. Stat. Assoc. **103**(484), 1609–1620 (2012)
10. Ma, L., Fu, T., et al.: Evaluation of feature selection methods for object-based land cover mapping of unmanned aerial vehicle imagery using random forest and support vector machine classifiers. ISPRS Int. J. Geo-Inf. **6**(2), 51 (2017)
11. Mills, P.: Efficient statistical classification of satellite measurements. Int. J. Remote Sens. 6109–6132(2011)
12. Antonogeorgos, G., Panagiotakos, B., et al.: Logistic regression and linear discriminant analyses in evaluating factors associated with asthma prevalence among 10- to 12-years-old children: divergence and similarity of the two statistical methods. Int. J. Pediatr. **2009**, 6 (2009)
13. https://Github.Com/H2oai/H2o-3/Tree/Master/H2o-R/Ensemble
14. Rokach, L.: Ensemble-based classifiers. Springer Science+Business Media B.V. (2009); Artif. Intell. Rev. **33**, 1–39 (2010)

Pixel-Based Supervised Tissue Classification of Chronic Wound Images with Deep Autoencoder

Maitreya Maity, Dhiraj Dhane, Chittaranjan Bar, Chandan Chakraborty and Jyotirmoy Chatterjee

Abstract With the extensive use of machine vision methodologies, computer-assisted disease diagnosis has become a popular practice for the medical professionals. Detailed analysis of wound bed area and precise identification of the wound tissue regions are the most desirable aspects of an automated wound assessment applications. This study proposes a supervised wound tissue classification method, where a deep neural network classifier model is trained by the colour, texture and statistical features which are extracted from different tissue regions. The proposed classification process considers three types of tissue, viz. granulation (red), necrotic (black) and slough (yellow) and a total of 105 features are used for the classification. A pixel-based feature extraction approach is implemented to extract features from the tissue region, where a mask window of size 9×9 runs over each pixel of the tissue regions for feature extraction. The proposed deep neural network achieves accuracy 99.997215%, sensitivity 99.998006%, specificity 99.996625% and F-Measure 99.997316%.

Keywords Chronic wound · Pixel-based classification · Feature extraction

1 Introduction

Chronic wound or ulcer is a complicated disease condition of localised injury of the skin and its tissues have physiological impaired healing response. The wound becomes chronic if it fails to restore an anatomical and functional integrity in an orderly and timely reparative process over a period of three months [1]. Wound area and tissue composition within the wound are the significant prognosticators of wound

M. Maity (✉) · D. Dhane · C. Chakraborty · J. Chatterjee
School of Medical Science and Technology, Indian Institute of Technology Kharagpur,
Kharagpur, West Bengal, India
e-mail: maitreya.maity@gmail.com

C. Bar
Department of Dermatology, Midnapore Medical College,
Midnapore, West Bengal, India
e-mail: drcrbar@gmail.com

healing; however, precise measurement with an unassisted perception remains a trivial assignment for dermatologists as it requires continuous evaluation. Currently, there exist contact and non-contact methods of wound area measurement. The traditional contact wound area measurement technique includes manual and digital planimetry using transparency tracing, colour dye injection and alginate moulds [2]. These methods are time consuming, inaccurate and unreliable for complex wounds. Hence, it is a non-idealistic clinical approach. The non-contact methods include 2D or 3D digital imaging, thermal imaging, elliptical methods, simple ruler and stereophotogrammetry (SPG) [3]. These methods are more relevant since they predict ulcer healing based on tissue composition and wound surface area.

Digital image processing and machine learning techniques have been applied in several studies to address the wound tissue classification problem. Berriss et al. [4] proposed tissue segmentation method where segmentation of granulation tissue was done by using 3D RGB histogram followed by a clustering technique. The proposed method was restricted to be generalized due to the seed selection procedure, the cluster creation order, less resolution of RGB image, and inadequate wound images. Another study has applied adaptive spline method for wound segmentation [5]. The slough tissue in pressure ulcer was segmented by changing the H value in HSI colour space. Then it was followed by quantitative wound slough measurement. Belem proposed a semi-automated wound segmentation process [6]. Multi-dimensional histogram sampling technique for automated segmentation of wound area and SVM for classification were reported by Kolesnik and Fexa [7]. Another attempt was made by [33], where unsupervised segmentation was used to identify wound-bed region at first. Later SVM classifier was introduced using $La * v$ and $Lc * h$ colour spaces. In another study proposed by Galushka et al. [8] where texture features and RGB colour features were used for classification of wound tissues using case-based classifier (CBR).

Quantitative assessment of wound is a potential research strategy that can be employed for combating traditional dermatology problems. Therefore, precise measurement tissues within it are very essential in monitoring the healing process of wound and assessment of therapeutic prognosis. In the proposed study, we have attempted to develop an automated wound tissue region segmentation using supervised autoencoder.

2 Materials and Methods

2.1 Sample Wound Image Database

An image database of colour images of chronic wounds was collected from the Midnapur Medical College and Hospital, India (MMCH). Institutional ethical clearance was taken for this study. Multiple digital cameras like SONY CyberShot DSC W520 digital camera and ASUS ZenPhone2 were used to capture the images. The dimen-

sions of each raw image were chosen as 4320 × 3240 pixels; however, the images were manually cropped to get maximum wound area with the minimum background. The images were taken under uncontrolled lighting conditions which lead to unstandardised colour effects. A 24-bit Macbeth colour checker was introduced near the wound surface during the image acquisition. Multiple wound images were also taken from online library named Medetec [9]. During sample collection, total 68 images were captured from 11 patients with diabetic ulcer, 24 images from 4 patients with pressure ulcer, 34 images from 6 patients with tropical ulcer and 124 images from 21 patients with surgical wounds.

2.2 Multi-expert Tissue Region Labelling

The sample images were provided to different clinicians for labelling tissue regions. Wound bed region is a mixture of granulation, slough and necrotic tissues. A GUI-driven software was developed to help clinicians for labelling tissue regions on the images easily. The screenshot of the developed wound tissue labelling software interface is presented in Fig. 1. The labelled images were later collected from the software database and compared to generate final ground truth image.
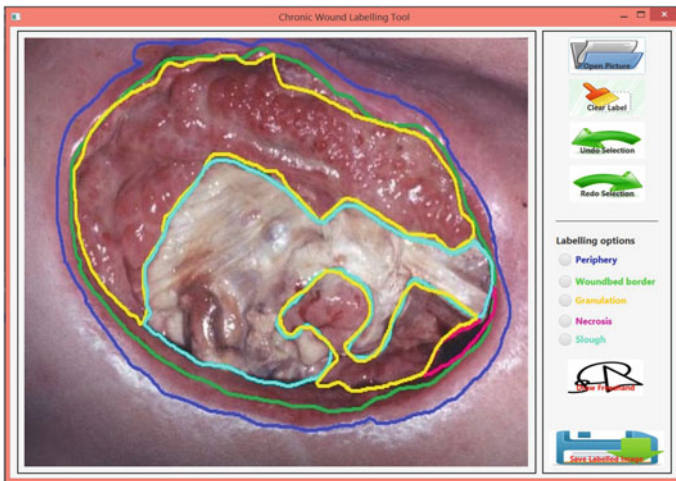
Fig. 1 Software interface for tissue region labelling
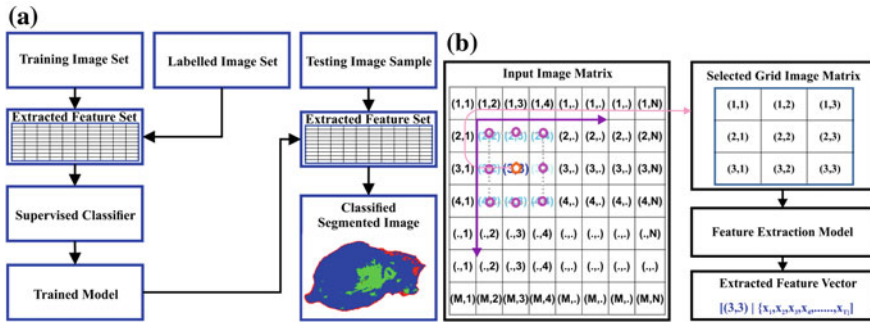
Fig. 2 **a** Workflow diagram of the proposed tissue classification process, **b** Pixel-based feature extraction flowchart

2.3 Pixel-Based Tissue Classification

The proposed methodology considered the only wound bed region as an input. It means the region of interest can only consist of Granulation, Necrotic, and Slough. The proposed algorithm worked in two phases. First, a trained, supervised deep neural network model was developed. And in the second phase, the same trained model was used to classify a new test wound image. The workflow diagram the proposed methodology is shown in Fig. 2a. The integral components of the pixel-based tissue classification approach are described below subsequently.

2.3.1 Feature Extraction Mechanism

Window-based feature extraction mechanism was designed to calculate features from images. The proposed feature extraction model ran over pixel-by-pixel and extracted features for each pixel. The abstract mechanism of the proposed feature extraction model is illustrated in Fig. 2b. An odd-numbered window box was considered which was masked over each pixel, and the targeted pixel was always placed in the centre position of the mask window. At each iteration, the fixed sized window moved by one step as its centre pixel changed to the next pixel position. In this manner, a mask window runs all over the image. At each pixel point, a small image grid was generated from the mask window area. The small image grid represented that particular centre pixel position. The size of the window was 9×9 and it was selected based on hit-and-trail. Eventually, a two-dimensional feature dataset was generated from an image.

2.3.2 Colour and Texture Feature

Visual interpretation is the key to wound diagnosis approach where colour and texture features of the affected area are observed precisely by medical professionals.

Table 1   Considered colour and texture feature information

Feature category	Extracted features	Count
Entropy	Shannon [10], Renyi's [11], Havarda-Charvat [12], Kapur [13], Yager [14]	5
Statistics	Mean, Standard deviation, Skewness, Kurtosis, Energy	6
Co-occurrence Matrix (GLCM) [15]	19 different features from co-occurrence matrix in 0°, 45°, 90° and 135°	19
Run length Matrix (GLRLM) [16]	11 different features from co-occurrence matrix in 0°, 45°, 90° and 135°	44
Local binary pattern (LBP) [17]	Three radius (R = 1, 2, 3) and corresponding pixel count (P = 8, 12, 24) to compute mean, variance, entropy and energy from uniform LBP	12
Colour	Mean, Standard deviation of different channel from RGB, HSV and HSL transformed image	19

Therefore, texture, grey and colour features were extracted from the labelled wound tissue regions. A total 105 different features were considered here. The details of considered image features are described in Table 1.

2.3.3   Deep Autoencoder

A deep network [18] is a neural network, where activation functions computed by lower layered neurons were re-used multiple times in the next-layered neurons for the computation of final decision function. Autoencoder is an optimization problem to find cost function, where the training was done without class label information. The cost function measures the error between input vector x and output $\tilde{x}$.

$$Mapping or Encoding : z^{(1)} = h^{(1)} \left(W^{(1)}x + b^{(1)} \right) \tag{1}$$

$$Decoding : \tilde{x}^{(2)} = h^{(2)} \left(W^{(2)}x + b^{(2)} \right) \tag{2}$$

Stochastic gradient descent-based cost function optimization

$$Cost Function : J = \sum_{i=1}^{L} \left(\tilde{x}^{(i)} - x^{(i)} \right)^2 \tag{3}$$

where superscript (1 and 2) were the first and second hidden layer. W is the weight matrix and b is bias vector. L is the number of hidden layers.

The training phase of the deep autoencoder was started with unsupervised autoencoder learning. Here, totally ten hidden layers were considered, and linear transfer function was considered for decoding of mapped to the output vector. The value of the L_2 weight regularizer set to 0.001, sparsity regularizer set to 4 and sparsity proportion set to 0.05. The features computed from trained first autoencoder was transferred to second autoencoder for learning. Rest of the nine autoencoders also generated features in a similar manner. The ten autoencoder and features were stacked together and generate a trained deep network for the proposed study.

2.3.4 Deep Network Training

In this part, the steps of deep neural network training procedure were discussed. The workflow of the classifier training process is presented in Fig. 3a. Let an image and its corresponding labelled image of size $M \times N$ was considered. And total T number of features were considered in the extraction model. The feature extraction model could extract a total of $M \times N$ feature vector from the given image where each vector would be the size of T. At the same time, from the ground truth image, a column vector for class labels was also generated of dimension $MN \times 1$. Finally, the labelled feature dataset will be constructed of dimension $MN \times (T + 1)$. Next, the labelled feature dataset was used for training with deep learning network model.

2.3.5 Wound Tissue Classification

The abstract block diagram of the classification process described in Fig. 3b. The same feature extraction model was used to generate feature dataset for the test image. The trained deep neural network model would help to predict the class label of each instance of the unlabelled test feature dataset. Using the predictions class label values, pixels were classified accordingly.
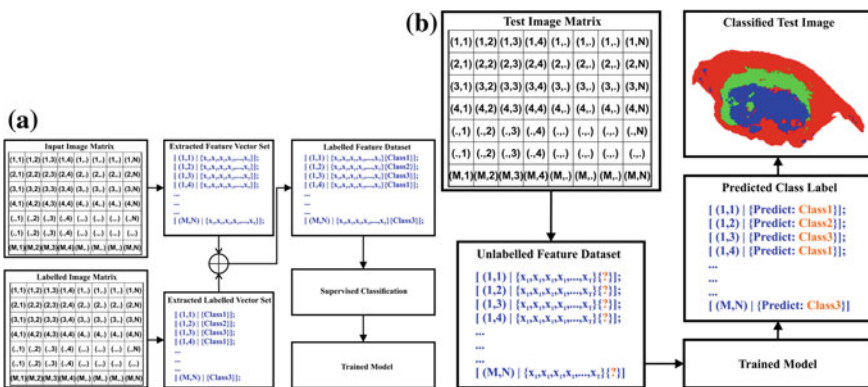
Fig. 3 **a** Classifier training process, **b** Test image classification process
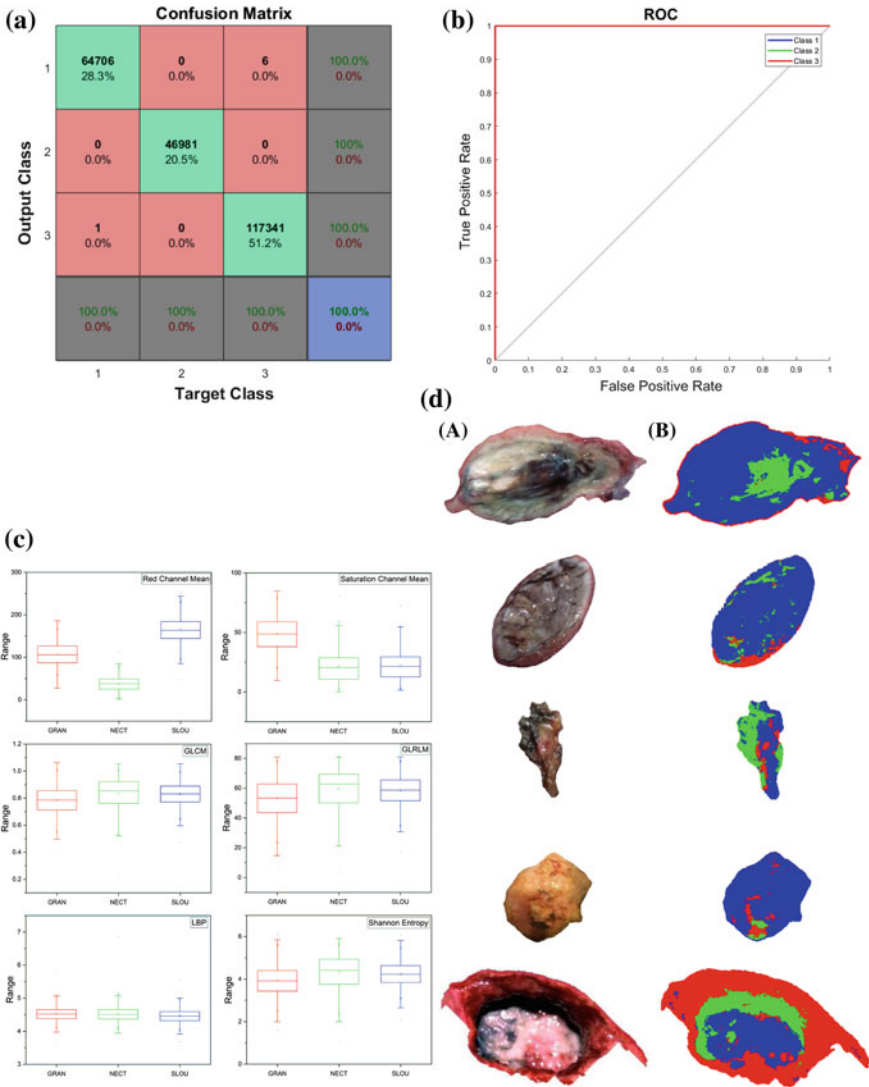
Fig. 4 **a** Confusion Matrix of deep network, **b** ROC curve, **c** Bar graph of extracted feature (GRAN = Granulation, NECT = Necrotic, SLOU = Slough), **d** Classification: (A) Original wound image, (B) Classified image

Table 2 Performance measures of deep network

Performance metrics	Measures (%)
Accuracy	99.997215
Sensitivity	99.998006
Specificity	99.996625
F-Measure	99.997316

3 Results and Discussion

The performance of the classification process was evaluated by five metrics, i.e. sensitivity, specificity, accuracy, F-Measure and Receiver Operating Characteristics (ROC) curve. The feature extraction model was able to generate a feature dataset of size 1465832×106. The statistics of some features are presented in Fig. 4c. The large feature dataset was randomly partitioned into training (66%) and testing (34%) dataset. The classification performance results are given in Table 2. In addition, the confusion matrix of classification on 34% test dataset is shown in Fig. 4a. Along with performance measure, the performance of the deep neural network is also plotted as a ROC curve in Fig. 4b. The AUC for this study is close to 1. An experiment was carried out where several wound images were randomly selected from the database, and their classification results were observed. Tissue classification of five random wound images is shown Fig. 4d. The granulation tissue region marked as red, necrotic tissue region marked as green and slough tissue region marked as blue in the classified image.

4 Conclusion

In this proposed study, a supervised tissue classification framework was proposed. Pixel or region-based feature extraction mechanism is introduced in the proposed methodology, where totally 105 features were extracted at each pixel point. The deep neural network model is trained by the developed labelled feature dataset successfully. The performance of the trained deep learning model is very accurate toward the classification of three types of wound tissue with an accuracy of 99.99%. However, such high accuracy was possible due to considering only wound bed region where the presence of any other types tissue objects is negligible. The proposed classification algorithm can be incorporated in a computer-aided diagnosis tool for automated wound monitoring.

References

1. Lazarus, G.S., Cooper, D.M., Knighton, D.R., Margolis, D.J., Percoraro, R.E., Rodeheaver, G., Robson, M.C.: Definitions and guidelines for assessment of wounds and evaluation of healing. Wound Repair Regen. **2**(3), 165–170 (1994)
2. Gethin, G., Cowman, S.: Wound measurement comparing the use of acetate tracings and VisitrakTM digital planimetry. J. Clin. Nurs. **15**(4), 422–427 (2006)
3. Bowling, F., King, L., Fadavi, H., Paterson, J., Preece, K., Daniel, R., Matthews, D., Boulton, A.: An assessment of the accuracy and usability of a novel optical wound measurement system. Diabet. Med. **26**(1), 93–96 (2009)
4. Berriss, W., Sangwine, S.: A colour histogram clustering technique for tissue analysis of healing skin wounds. In: International Conference on Image Processing and Its Applications, vol. 2, pp. 693–697. IET (1997)
5. Hoppe, A., Wertheim, D., Melhuish, J., Morris, H., Harding, K., Williams, R.: Computer assisted assessment of wound appearance using digital imaging. In: Proceedings of the 23rd Annual IEEE International Conference of the Engineering in Medicine and Biology Society, vol. 3, pp. 2595–2597. IEEE (2001)
6. Belem, B.: Non-invasive wound assessment by image analysis. Ph.D. thesis, University of Glamorgan (2004)
7. Kolesnik, M., Fexa, A.: Multi-dimensional color histograms for segmentation of wounds in images. In: International Conference on Image Analysis and Recognition, pp. 1014–1022. Springer (2005)
8. Galushka, M., Zheng, H., Patterson, D., Bradley, L.: Case-based tissue classification for monitoring leg ulcer healing. In: IEEE Symposium on Computer-Based Medical Systems (CBMS'05), pp. 353–358. IEEE (2005)
9. Medetec Medical Images: http://www.medetec.co.uk/files/medetec-images.html (2016)
10. Hammer, D., Romashchenko, A.E., Shen, A., Vereshchagin, N.K.: Inequalities for Shannon entropies and Kolmogorov complexities. In: IEEE Conference on Structure in Complexity Theory, pp. 13–23. IEEE (1997)
11. Rrnyi, A.: On measures of entropy and information. In: Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 547–561 (1961)
12. Havrda, J., Charvát, F.: Quantification method of classification processes. Concept of structural a-entropy. Kybernetika **3**(1), 30–35 (1967)
13. Kapur, J.N., Sahoo, P.K., Wong, A.K.: A new method for gray-level picture thresholding using the entropy of the histogram. Comput. Vis. Graph. Image Process. **29**(3), 273–285 (1985)
14. Huang, L.K., Wang, M.J.J.: Image thresholding by minimizing the measures of fuzziness. Pattern Recogn. **28**(1), 41–51 (1995)
15. Haralick, R.M., Shanmugam, K., et al.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **6**, 610–621 (1973)
16. Galloway, M.M.: Texture analysis using gray level run lengths. Comput. Graph. Image Process. **4**(2), 172–179 (1975)
17. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–987 (2002)
18. Le, Q.V., et. al.: A tutorial on deep learning part 2: autoencoders, convolutional neural networks and recurrent neural networks (2015)

A CBIR Technique Based on the Combination of Shape and Color Features

Sumit Kumar, Jitesh Pradhan and Arup Kumar Pal

Abstract In CBIR techniques, image retrieval based on object-based features are more precise to retrieve appropriate relevant images. So, in this paper, a CBIR technique is proposed using extracted combined shape and color features from image object region. In this particular work, some significant statistical parameters are calculated from image object or shape region by gray-level co-occurrence matrix and simultaneously, color features are extracted from the color object using color autocorrelogram. Initially, RGB color images are transformed into YCbCr color space, and subsequently, the active contour is employed on Y-component to obtain the foreground and the background regions. Shape or object feature is located in the foreground region of Y-component and gray-level co-occurrence matrix provides some statistical parameters. We have also computed some statistical parameters from the background region to improve the image retrieval performance. Afterward, an intermediate color object image is reconstructed by combining foreground image region along with chrominance components for deriving the prominent color information. We have employed color autocorrelogram over this newly constructed intermediate image. Finally, all the computed features are combined together to form the ultimate feature vector. The proposed technique is tested over two benchmark databases, i.e., Corel-1K and GHIM-10K and we have achieved satisfactory results in object-based images.

Keywords Active contour · Color autocorrelogram · CBIR
GLCM

S. Kumar (✉) · J. Pradhan · A. K. Pal
Department of Computer Science and Engineering, Indian Institute of Technology
(Indian School of Mines), Dhanbad 826004, Jharkhand, India
e-mail: sumitkumar@cse.ism.ac.in

J. Pradhan
e-mail: jitpradhan02@gmail.com

A. K. Pal
e-mail: arupkrpal@gmail.com

1 Introduction

In the present digital era, the volume of multimedia data is increasing rapidly due to the advancement of image capturing devices and availability of high-speed Internet. So in the present scenario, the organization of multimedia data in suitable databases is important but also retrieving the desired result from the stored databases are equally important. Images are an integral part of any multimedia system as a single image persists a lot of information which could be significant for individual or public interest. The conventional retrieval approach is not fruitful since the image descriptive keyword is not sufficient to interpret the actual image content. So the visual content-based image retrieval is the proper solution to search the more relevant images. This area is known as Content-Based Image Retrieval (CBIR) [1]. The purpose of CBIR process is to compute a feature vector using the extracted significant information like color, texture and/or shape of the image. This process is known as image feature extraction and the CBIR performance depends on the formation of the suitable feature vector. This process is employed on the query image as well as on the database images to derive the feature vector of the query image and feature database respectively. Afterwards, query image feature vector is matched with all the feature vectors present in the database to retrieve some top images.

In CBIR, color, texture, and shape [2] are the key contents of any image retrieval process. Color is one of the significant features as it is independent of the size and orientation of the image content. Color histogram [3] is widely used tool to extract color information since it gives the global color distribution information of the image. But two entirely different images can have the same histogram so to overcome this, the multiresolution histogram has been proposed in [4]. Dominant Color Descriptor (DCD) [5] is used in image retrieval process. It provides salient color distribution of an entire image or region of interest but the drawback of DCD is it does not resemble with human perception. Yang [6] proposed a quantized DCD method called Linear Block Algorithm (LBA) and show that LBA is efficient in computation. Lu [7] presented a CBIR system which uses color distribution for global features of the image and image bitmap for incorporating the local features. In CBIR, shape features are also another important component in retrieving object-based images. In general, the user prefers to search images based on some items or things, for example, red rose, blue car, etc. and in this particular retrieval process, the object-based image searching is more relevant. Shape features [8] are mostly categorized into two parts, i.e., region-level information and boundary-level information. Our literature survey found some existing shape-based techniques. Gagaudakis [9] suggested a CBIR technique to incorporate shape. Zhang [10] suggested a CBIR technique to extract image shape features using generic Fourier transformation.

The combination of multiple image features improves the retrieval accuracy since a particular image feature is not adequate to represent images properly. So in literature, several techniques are found with multiple image features. Wang [11] presented

a technique to extract color, texture and shape image features. Kumar [12] presented a CBIR technique using EDH and active contour. Yue [13] presented a CBIR technique which combines color and texture features. Color histogram and co-occurrence matrix to extract color and texture features respectively. Varish [14] proposed a CBIR technique based on color and texture features. Dual-tree complex wavelet transformation and color image histogram are deployed in their technique. The combined color and texture features are used in CBIR technique in the wavelet domain.

In this paper, our objective is to search the images based on object categories. As discussed earlier, the object image retrieval process is preferred by most of the users. So an object-based CBIR is proposed where object-based features are extracted in two phase. In the first phase, the GLCM is employed to extract some statistical parameters from object region and in the second phase, color information is imagined only from the color-based object region using color autocorrelogram. Finally, all extracted features are combined to form image feature which is used in CBIR process.

The rest of the paper is structured as follows: Sect. 2 presents the proposed feature extraction algorithm. In Sect. 3, the experimental results and performance analysis are discussed. Section 4 concludes the paper followed by the corresponding references.

2 Proposed CBIR Technique

The intention of the proposed work is to develop an object-based image retrieval technique. The technique is developed in two phase, i.e., feature extraction scheme and similarity measurement. For the first phase, the intensity image is taken from YCbCr color space as a form of Y-component. In this work, we have used active contour [15] on Y-component to identify the foreground and background image region. Figure 1 shows the intermediate outcomes after applying the active contour on a particular image.

Afterwards, GLCM is applied on foreground image and corresponding statistical parameters are taken as components of a feature vector. We have also considered some statistical parameters from entire background region directly and appended with the feature vector. In this work, instead of computing the color features from the entire image, an intermediate object-based color image is constructed. Later, color autocorrelogram [16] is used to find the significant color components from the earlier generated intermediate image. The overall image feature extraction is efficient in computational aspect since it works only on the selected region, i.e., object instead of the entire image. Figure 2 depicts the overall feature extraction process. The algorithmic steps are summarized in Algorithm 1. In the second phase, user provides a query image which undergoes a certain feature extraction process and generates the query feature vector. Now, a stored database in the user's system undergoes the same feature extraction process and results in feature database. Now the query feature vector is compared with each part of the feature database
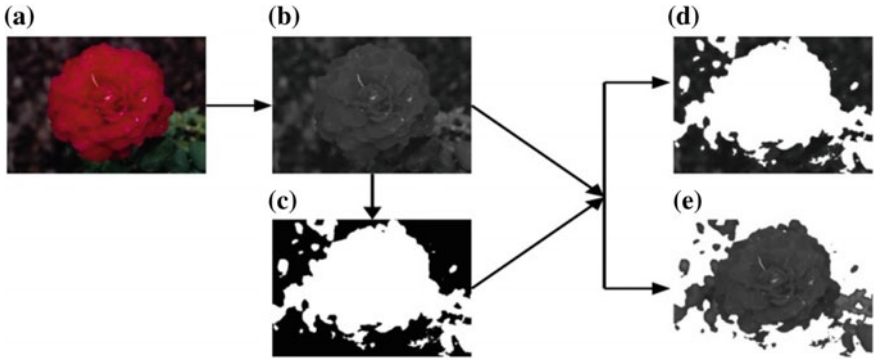
Fig. 1 a Original image. b Y-component. c Active contour result. d Background region.
e Foreground region
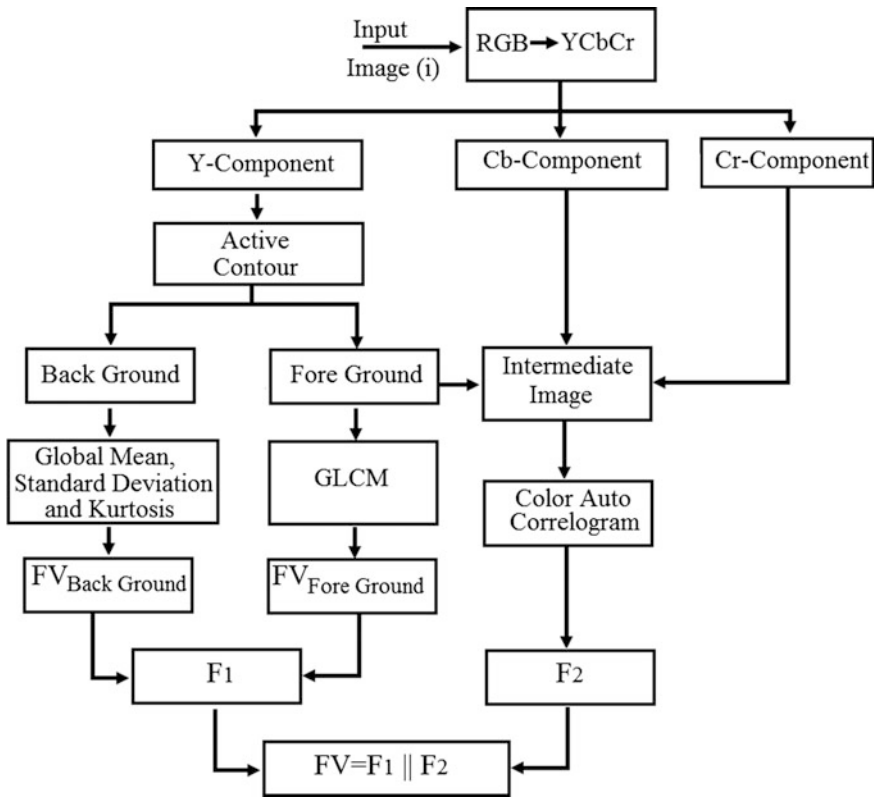
Fig. 2 Feature extraction procedure

according to the selected similarity measure. Then accordingly based on similarity, the topmost relevant images are retrieved.

Algorithm 1: **Feature Extraction Process**

Begin

Step-1: Convert the input RGB image into corresponding YCbCr color space image and decompose it into its principal components, i.e., Y, Cb, Cr.

Step-2: Perform active contour to get foreground and background region from the Y-component.

Step-3: Apply GLCM on foreground region to compute energy, contrast, correlation and homogeneity components.

Step-4: Compute mean, standard deviation, and kurtosis from background region.

Step-5: Combine the features of Step-3 and Step-4 to get feature vector, F1.

Step-6: Select only the Cb and Cr components from foreground region and construct the intermediate object-based color image (As shown in Fig. 3).

Step-7: Apply color autocorrelogram on the intermediate object-based color image to get feature vector, F_2.

Step-8: Merge F_1 and F_2 to achieve final feature vector, FV.

End
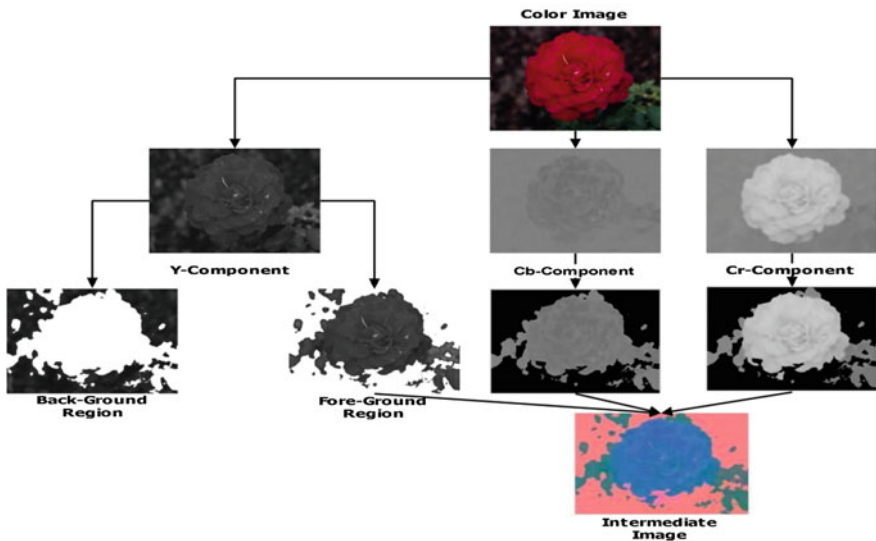
Fig. 3 Intermediate image generation

3   Experimental Results

In this section, we have presented the results that we have received after the simulation of the described algorithm using MATLAB platform. To display our results and proving the acceptability of the proposed algorithm, we have tested our system on two benchmarked databases, i.e., Corel-1K database [17], which contain 1000 image of 10 different categories so that every category has 100 images each. Another one is GHIM-10K database [18], which contains 10,000 images of 20 different categories so that each category has 500 images. Every user selected query image goes through the proposed system and forms its query feature vector and similarity the respective database go through the same process and formed feature vector database. Then query feature vector is compared using Euclidean distance with each feature vector of the feature database to find the similar results. We have taken top 20 images for each passed query images. To check, whether our system is adequate to retrieve better results or not, we have computed precision and recall based on the Eqs. (1) and (2) respectively. The process of constructing intermediate color image construction has been depicted in Fig. 3. In general, object-based image categories are given based on acceptable image retrieval performance (Tables 1 and 2).

$$Precision = \frac{No.\, of\, relevant\, images\, retrieved}{Total\, no.\, of\, images\, retrieved} \tag{1}$$

$$Recall = \frac{No.\, of\, relevant\, images\, retrieved}{Total\, no.\, of\, relevant\, images\, available\, in\, the\, database} \tag{2}$$

Table 1 Precision comparison based on Corel-1K

Category	Yu et al. [19]	Elalami et al. [20]	Kekre et al. [21]	Proposed
People	55	70	65	50
Beaches	47	56	**60**	40
Buildings	56	57	**62**	45
Buses	**91**	87	85	60
Dinosaurs	94	98	93	**100**
Elephants	49	67	65	**80**
Flowers	85	91	**94**	60
Horses	52	**83**	77	50
Mountains	37	53	**73**	50
Foods	55	74	**81**	35

Table 2 Precision of GHIM-10K

Category	Precision	Category	Precision
Fireworks	75	Airplanes	75
Buildings	45	Butterflies	45
Walls	40	Forts	35
Cars	45	Sunsets	65
Flies	40	Bikes	100
Mountains	35	Boats	45
Flowers	55	Ships	45
Trees	50	Chicken	60
Grounds	35	Insects	45
Beaches	35	Horses	75

4   Conclusion

In this work, we have proposed a CBIR technique using object-based image features. The foreground image is obtained using active contour process on the Y-component and GLCM is used to exploit the shape feature of the image. As the background region of the image does not persist large significant information, we have evaluated some global statistical parameters. Then a new image is reconstructed by combining chrominance parts, i.e., Cb, Cr, and foreground image. Color autocorrelograms is deployed to extract the information. The proposed technique is tested over two benchmark databases and adequate results are retrieved for the object-based images. The technique is suitable for object retrieval-based applications.

References

1. Dharani, T., Aroquiaraj, I.L.: A survey on content based image retrieval. In: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), pp. 485–490. IEEE (2013)
2. Zhong, Y., Jain, A.K.: Object localization using color, texture and shape. Pattern Recogn. **33**(4), 671–684 (2000)
3. Varish, N., Pal, A.K.: Content based image retrieval using statistical features of color histogram. In: 3rd International Conference on Signal Processing, Communication and Networking (ICSCN) (2015)
4. Xu, Z., Ling, H., Zou, F., Lu, Z., Li, P.: Robust image copy detection using multi-resolution histogram. In: Proceedings of the International Conference on Multimedia Information Retrieval, pp. 129–136. ACM (2010)
5. Min, R., Cheng, H.D.: Effective image retrieval using dominant color descriptor and fuzzy support vector machine. Pattern Recogn. **42**(1), 147–157 (2009)
6. Yu, C.C., Jou, F.D., Lee, C.C., Fan, K.C., Chuang, T.C.: Efficient multi-resolution histogram matching for fast image/video retrieval. Pattern Recogn. Lett. **29**(13), 1858–1867 (2008)
7. Lu, T.C., Chang, C.C.: Color image retrieval technique based on color features and image bitmap. Inf. Process. Manage. **43**(2), 461–472 (2007)

8. Singha, M., Hemachandran, K.: Content based image retrieval using color and texture. Signal Image Process.: Int. J. (SIPIJ) **3**(1), 39–57 (2012)
9. Gagaudakis, G., Rosin, P.L.: Incorporating shape into histograms for CBIR. Pattern Recogn. **35**(1), 81–91 (2002)
10. Zhang, D., Lu, G.: Generic fourier descriptor for shape-based image retrieval. In: 2002 IEEE International Conference on Multimedia and Expo, 2002. ICME'02. Proceedings, vol. 1, pp. 425–428. IEEE (2002)
11. Wang, X.Y., Yu, Y.J., Yang, H.Y.: An effective image retrieval technique using color, texture and shape features. Comput. Stand. Interfaces **33**(1), 59–68 (2011)
12. Kumar, S., Pal, A.K.: A CBIR scheme using active contour and edge histogram descriptor in YCbCr color space. IJCTA **9**(41), 889–898 (2016)
13. Yue, J., Li, Z., Liu, L., Fu, Z.: Content-based image retrieval using color and texture fused features. Math. Comput. Model. **54**(3), 1121–1127 (2011)
14. Varish, N., Pradhan, J., Pal, A.K.: Image retrieval based on non-uniform bins of color histogram and dual tree complex wavelet transform. Multimed. Tools Appl. 1–37 (2016)
15. Xu, H., Jiang, G., Yu, M., Luo, T.: A global and local active contour model based on dual algorithm for image segmentation. Comput. Math. Appl. (2017)
16. Mahmoudi, F., Shanbehzadeh, J., Eftekhari-Moghadam, A.M., Soltanian-Zadeh, H.: Image retrieval based on shape similarity by edge orientation autocorrelogram. Pattern Recogn. **36**(8), 1725–1736 (2003)
17. Tang, J., Lewis, P.H.: A study of quality issues for image auto-annotation with the corel dataset. IEEE Trans. Circuits Syst. Video Technol. **17**(3), 384–389 (2007)
18. Li, J., Wang, J.Z.: Real-time computerized annotation of pictures. IEEE Trans. Pattern Anal. Mach. Intell. **30**(6), 985–1002 (2008)
19. Yu, J., Qin, Z., Wan, T., Zhang, X.: Feature integration analysis of bag-of-features model for image retrieval. Neurocomputing **120**, 355–364 (2013)
20. ElAlami, M.E.: A novel image retrieval model based on the most relevant features. Knowl.-Based Syst. **24**(1), 23–32 (2011)
21. Kekre, H.B., Thepade, S.D., Sarode, T.K., Suryawanshi, V.: Image retrieval using texture features extracted from GLCM, LBG and KPE. Int. J. Comput. Theory Eng. **2**(5), 695 (2010)

Trajectory Forecasting of Entities Using Advanced Deep Learning Techniques

**K. H. Apoorva, Raghu Dhanya, Anil Kumar Anjana
and S. Natarajan**

Abstract Recent growth in depth camera technology has significantly enhanced human motion tracking. Human future behaviour and intention forecasting become a challenging task due to high-dimensional interactions with the physical world. Prognostic methods that estimate ambiguity are, therefore, critical for supporting appropriate robotic responses to the numerous ambiguities posed within the human–robot interaction environment. Beyond autonomous agents, we will also see our surroundings—buildings, cities—becoming equipped with ambient intelligence which can sense and respond to human behaviour. In this paper, we present different deep learning models that can forecast the navigational behaviour of multiple classes (i.e. pedestrian, car, cycle) by considering influencing factors such as the neighbouring dynamic subjects and social behaviour of the classes under investigation. The results show that our approaches outperform the existing state-of-the-art forecasting models.

Keywords Recurrent Neural Networks · Autonomous agents · Predictive models

K. H. Apoorva (✉) · R. Dhanya · A. K. Anjana
PESIT, Bengaluru, India
e-mail: apoorva.khp@gmail.com

R. Dhanya
e-mail: dhanyaraghu52@gmail.com

A. K. Anjana
e-mail: anjana1909@gmail.com

S. Natarajan
PESU, Bengaluru, India
e-mail: natarajan@pes.edu

1 Introduction

The need for reliable models of navigation of collaborating agents is seen in multiple fields including robotics, computer graphics and behavioural science. A growing requirement for co-robotic applications that place robots as partners with humans play a role in cooperative and tightly interactive tasks [1]. Models of pedestrian navigation behaviour to identify the trajectories of nearby pedestrians are needed by mobile robots to navigate amid pedestrians in a safe, efficient and socially acceptable way [2].

When people navigate in a crowded public area such as a railway station, an airport terminal, a sidewalk or a shopping mall, they follow a number of common sense rules and conform to social conventions by considering right-of-way and private space. Machine learning methods usually consider fully observable environments with features that determine the motion and thus allow for predictions. Also, collision avoidance behaviours of humans (social aspect) pose a challenging aspect for developing optimal models. We thus investigate and evaluate various deep learning methods to predict trajectories of different moving entities in crowded spaces which is the main objective of the paper.

The deep learning models based on Recurrent Neural Networks (RNNs) are considered due to the relevance of their architecture in solving the problem at hand [3]. The RNNs' inherent ability to maintain temporal context appropriately fits our need of processing trajectories of entities, which is essentially sequential data [4]. As an extension to the RNN model, we propose three novel methods namely: multitarget pooling, attention network and Hierarchical RNN network which are seen to give better performance than existing state of the art methodologies.

2 Related Work

Deep learning models have been applied extensively to many sequence-to-sequence modelling problems such as the problem under consideration. In this section, we highlight the recent advancement in forecasting the future behaviour of people to improve intelligent autonomous systems. Alahi et al. [5] addresses the problem of forecasting pedestrians' destinations, a main problem in understanding crowd mobility. This paper proposes a new descriptor namely Social Affinity Maps (SAM) to link and capture the broken or unobserved trajectories of individuals in the crowd. Robicquet et al. [6] involves in predicting a trajectory and future intent of a human, similar to the objective of our research work. They address the high dimensionality and uncertainty problem by using predictive inverse optimal control methods for estimating probabilistic model of human motion trajectories. Our models include many other output classes other than humans such as bikers, cars, skaters etc., thus making it a multiclass regression model. Alahi et al. [7] proposes an LSTM model, which can learn general human movement and also further predict their future trajectories,

similar to the interests of our research. Their work fails to consider the speed with which the neighbouring entities approach each other and also considers only pedestrian interaction. Our approaches handle interactions between entities of multiple classes, not limiting to that of pedestrian alone. The attention mechanism approach we propose, tackles the speed aspect and thus handles collision avoidance and social force. In [8], the authors use GRU to predict object's future position. They consider the influence of neighbouring objects for the prediction, however, they assume no influence of the proximity or the speed of the objects. In the subsequent sections, we present our initiative to cope with above limitations.

3 Methodology

In this section, we describe the chosen deep learning models used for predicting an occupant's trajectory. Every occupant's trajectory is modelled as a sequence or time series. Each target entity belonging to a given class (e.g. pedestrian, car, etc.) in a real-world scene has a given characteristic navigation style which can be defined by a set of unique parameters associated with a forecasting model. If the parameters are considered as the learned navigation rules, this translates to all member occupants of a class. All entities or occupants belonging to the same class have same prediction model parameters which are essentially the weight matrices of the neural networks. The individual characteristics and context [9] of an entity are captured by the model's hidden layer. This helps to customise our model for each occupant and help better extract its moving trajectory considering its individual and collective class-wise behavioural pattern [10]. In addition, we define baseline models for comparison of performance with our approaches.

3.1 Baseline

We define a baseline model based on vanilla RNN and GRU without any customizations. In this model, the intention is to effectively predict the future trajectories of various entities, taking into consideration, only their past trajectories. The performance of the proposed approaches is evaluated by comparing it to that of the baseline model. Recently, Recurrent Neural Networks (RNN) and their variants including Gated Recurrent Units have proven to be very successful for sequence prediction tasks. We observe the entire path of the entity from T_0 to some $T_{observed}$ and predict the position for the next time step, $T_{observed+1}$.

Vanilla RNN:

The input to the neural network is the current location of the selected entity in the video frame. The RNN is trained to learn a sequence of video frames to capture the traversing behaviour of that entity, as depicted in Fig. 1a. The future path is then

(a) Internal Structure of RNN   (b) Pooling of hidden states in Multi-Target Approach
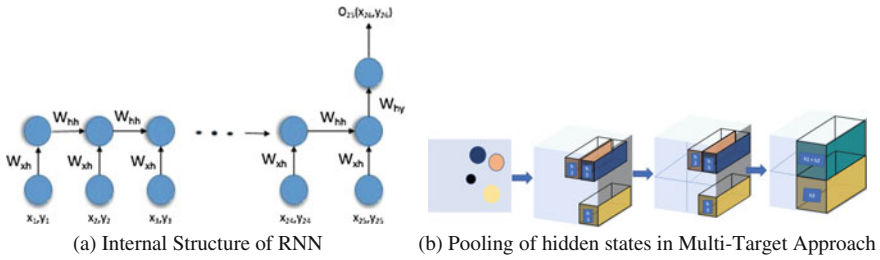
Fig. 1 Baseline and pooling models

predicted using the learned context captured by the hidden states of the RNN. The hyperparameters including the hidden state size, training data size and learning rate are varied and tested for different values.

3.2 Multitarget Trajectory Prediction

Behaviour of people in crowded scenes is dependent on the motion of other people and obstacles in their vicinity. For example, a person can deviate his/her path to avoid collision with another group of people moving towards him. A person cannot be observed in isolation to predict trajectories as we cannot sideline the influence from other entities. In this approach, we consider the positions of neighbouring entities of the selected entity along with its past trajectory for future path learning.

3.2.1 Pooling RNN Model

We introduce a pooling [8] technique based on RNNs as elaborated below. Every entity is modelled as a single RNN instance and the influence by surrounding entity instances is captured through this pooling technique. This is achieved by adding the hidden states of the neighbours in the 'pooling' window represented by a 'p × q' pixel square centred upon the entity as depicted in Fig. 1b. Pretraining of weight matrices is done to capture the behaviour of various output classes and these are used for the entities in calculating the context at the hidden layer. For each entity, we find its neighbours such that the Euclidean distance between the neighbour in the pooling window and the entity is less than or equal to a threshold value t.

$$H_t^e = H_t(i^m, j^m) + h_t^m \tag{1}$$

where, for each neighbour 'm' at time 't', we locate the appropriate grid coordinates i, j in the pooled matrix ($H_t(i^m, j^m)$) and summate it with the hidden state of the neighbour 'm' at time 't' (h_t^m) around the selected entity's location on the p × q grid.

$$h^e_{t,final} = W^s \cdot H^e_t \tag{2}$$

where, W^s is a high-density tensor which is learnt as an additional parameter during the backpropagation of weights, which is used to calculate the effective hidden state of the entity 'e' ($h^e_{t,final}$) and thus aids in finetuning the navigational skills of our trained model.

3.2.2 Hierarchical RNN Model

While we regard the influence of the neighbouring entities, we must take into account the prominence given to each neighbour. Intuitively, neighbours which are nearer to the considered entity have more influence on the entity as compared to a far off neighbour. We put forth a novel multi-scale RNN model, which can learn the hierarchical multi-scale structure from temporal data. The hierarchical approach uses a weighted method to append the hidden states. The closer entities have higher weights, thus, having more hand at hidden state variation. A virtual hierarchy of weights is given emerging from the entity under consideration. Hierarchy is applied on the pooled tensor to obtain the final hidden state, as seen in Fig. 2a. The p × q pooled tensor has x levels of hierarchy based on proximity. The weights provided in this approach for the neighbours are constant for a specific radius around the selected entity.

$$H^e_t = H_t(i^m, j^m) + w^m_{i,j} \cdot h^m_t \tag{3}$$

where, for each neighbour 'm' at time 't', we locate the appropriate grid coordinates i, j in the pooled matrix ($H_t(i^m, j^m)$) and summate it with the product of the constant weight assigned ($w^m_{i,j}$) and the hidden state of the neighbour 'm' at time 't' (h^m_t) around the selected entity's location on the p × q grid.

3.2.3 Attention Mechanism Model

The Attention Memory Network is modelled to give importance to the most relevant and highly affecting components of the input. In this approach, we consider the speed with which the neighbouring entities approach the selected entity to handle collision avoidance and social sense of humans. The weightage provided by the hierarchical approach is considered, to rank the initial importance of all of the entity's neighbours. As the neighbours with higher speeds have a greater chance of colliding, they are rated in increasing order of their speeds by providing corresponding weights to prioritise them as shown in Fig. 2b. We induce this ranked feature into our RNN, which are further involved in prediction of the trajectories.

$$H^e_t = \frac{\Delta d}{\Delta t} \cdot h^m_t \tag{4}$$
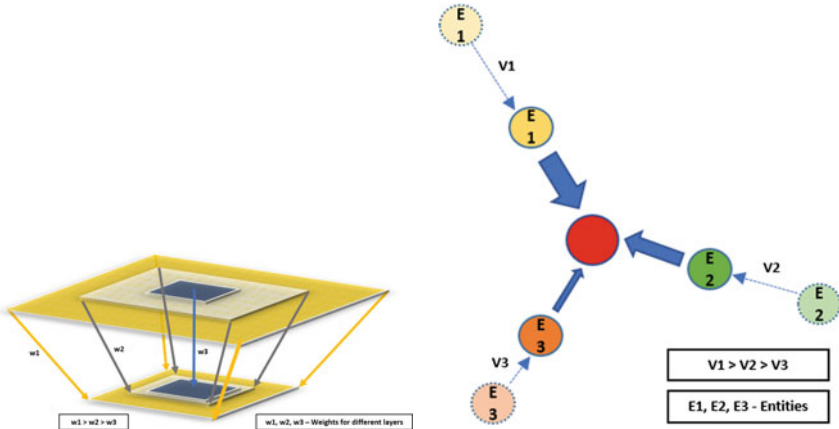
(a) Hierarchical Scaling of the Pooled tensor

(b) Applying attention on the hidden states of prominent neighbouring entities. E1 contributes more than E2, and E2 contributes more than E3 to the hidden state of the entity.

Fig. 2 Hierarchical and attention mechanism models

where, m is each neighbour within the boundaries of the restricted space radius around the selected entity's location on the $p \times q$ grid. The total rate of change of distance($\Delta d / \Delta t$) for each neighbouring occupant coupled with hidden state (h_t^m) plays a key role in updation of the pooled hidden state of the entity (H_t^e).

Algorithm:

Calculating hidden state for entity, e:

    N = neighbours of e

    for each n in N:

        d = euclidean dist between n and e

        If d < (100×100 pooling window):

            assign weight, w to n based on its speed towards e:

            $H_t^e = H_t(i^n, j^n) + h_t^n * w$

    Reduce dimension of H_t^e to 20×20 from 100×100.

    Final pooled hidden state of e, $h_{t,final}^e = W^s \cdot H_t^e$

4 Experiments and Results

Dataset

For our data, we use the Stanford Drone Dataset, taken from and pre-annotated by Stanford's Computer Vision and Graphics Lab. This new large-scale dataset contains
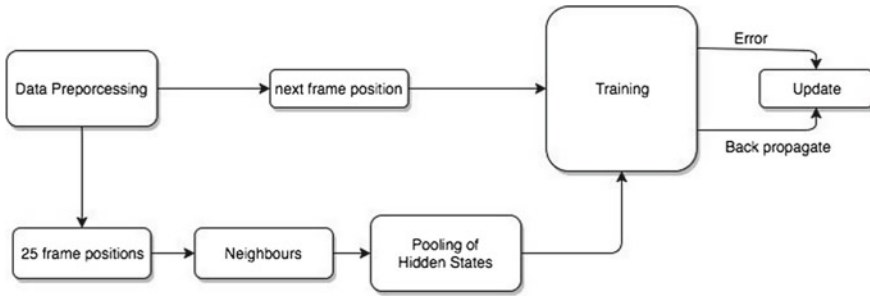
Fig. 3 System design

videos of various types of targets (pedestrians, bikes, skateboarders, cars, buses, golf carts) that navigate in a real-world outdoor environment such as a university campus. It comprises of more than 100 different top-view scenes for a total of 20,000 targets engaged in various types of interactions (Fig. 3).

4.1 Experimental Setup

As part of the implementation of our models, we have specified the dimension for both, the input and output layer, to be two, representing the x and y coordinates of the entity in the video frame. A hidden layer dimension of 64 is found to give the most accurate predictions. The models use *tanh* as the activation function. The model parameters are updated using Stochastic Gradient Descent as our backpropagation algorithm. The initial learning rate was set to 0.01 and annealed upon increased loss. The models are trained for 70 epochs. Each video frame represents a single increment in time step. To train the model, we allow the RNN to observe 25 video frames of an occupant's trajectory and predict its location in the next video frame (time step). For our Multi-target Pooling approach, the entity neighbourhood was set to 100×100 and used 5×5 pooling windows as they yield most optimal results.

4.2 Results and Analysis

The accuracies for all models are calculated using the path displacement metric, i.e. the displacement between predicted and actual position. The prediction is considered accurate if the path displacement between the predicted output paths and the actual paths of entities falls within a threshold of 36 pixels. We also evaluate each of our models based on the error rate metric. This is effectively the displacement between the predicted trajectory point and the actual target trajectory point of the entity along the path.

(a) Error Rate   (b) Accuracy
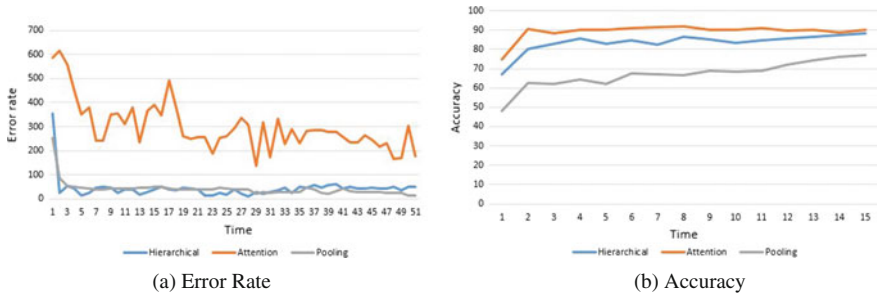
Fig. 4 The error rate and accuracy versus time is plotted for the three models

Baseline approach using RNN: Our single-track RNN model achieved an accuracy of 85% with learning rate as 0.3, 40 epochs and 64 as the hidden state dimension. This model does not handle collision avoidance, thus not incorporating social sensitivity specification.

Multitarget tracking approach: As seen in Fig. 4a, the error rate of attention mechanism approach is the least. The losses for all the models decrease consistently over time, conforming to the theoretical expectations. Attention method has considerably higher loss when compared to the other two models because of the dynamically changing preference given to neighbouring occupants. The fluctuation of the loss increases from pooling to hierarchical to attention method because of the seemingly unpredictable velocities of the surrounding entities.

According to Fig. 4b, the best accuracy is obtained for attention mechanism, followed by hierarchical approach and lastly, the pooling method. This is because attention mechanism takes the neighbouring entities' velocities and direction of approach into consideration. The accuracies for all the three models increase across time, gradually learning the features of the trajectories of various entities in each frame.

Table 1 describe the output of classwise accuracies for each approach mentioned in the paper. The classwise accuracies provide an insight into the ability to predict the path, given the characteristic behaviour of each class. The skater and cart class gave a good accuracy considering their motion to be linear most of the time. The pedestrian and biker class gave lesser accuracy as account of the number of entities being high and the added nonlinearity.

Table 2 represent accuracies based on various scenes in the dataset. Each scene has its own characteristic road paths with some having curves and some straight roads. Generally, the accuracy increases with each succeeding model for each scene. Accuracies for the 'Coupa' and 'Gates' scenes are high because the entities here mostly have linear paths. Our models can successfully predict non-linear paths of moving entities, as seen with the accuracies of the 'Death circle' scene which has a circular roundabout in the centre. Pre-training of weights, however improved the accuracies by a small factor.

Table 1 Class-wise comparison of accuracies for all three approaches

	Pedestrian	Car	Skater	Biker	Cart	Overall
Pooling	82.09	73.62	90.9	84.56	86.77	86.71
Hierarchical	86.36	90.5	93.6	85.82	87.36	87.5
Attention	91.4	74.55	87.6	88.52	92	91.75

Table 2 Class-wise comparison of accuracies for all three approaches

	Book-store	Death circle	Coupa	Gates
Pooling	82.52	83.6	88	82.5
Hierarchical	86.7	89.16	90.6	87.78
Attention	91.3	90.37	89.2	88.92

5 Conclusion

We have presented RNN-based modes that can reason across multiple entities to predict multi-class trajectories in various scenes. We use three approaches to perform trajectory forecasting in a multi-class setting, including a social pooling layer, hierarchical RNN approach and attention mechanism. On comparison of all our proposed methods, attention mechanism with RNN approach showed higher performance. We also effectively show that our approaches successfully predict various nonlinear behaviours arising from social interactions. However, our current methodologies of pooling cannot comprehensively predict perfect trajectories. In future, we will extend our models to consider more influencing factors and implement reinforcement learning for further qualitative comprehension of the results.

References

1. Kretzschmar, H., Kuderer, M., Burgard, W.: Learning to predict trajectories of cooperatively navigating agents. In: IEEE International Conference on Robotics and Automation-ICRA. IEEE (2014)
2. Antonini, G., Bierlaire, M., Weber, M.: Discrete choice models of pedestrian walking behavior. Transp. Res. Part B-Methodol. **40**(8), 667–687 (2006)
3. RNN. In Wikipedia, from https://en.wikipedia.org/wiki/Recurrent_neural_network
4. Bartoli, F., et al.: Context-aware trajectory prediction (2017). arXiv:1705.02503
5. Alahi, A., Ramanathan, V., Fei-Fei, L.: Socially-aware large-scale crowd forecasting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2014)
6. Robicquet, A., et al.: Forecasting social navigation in crowded complex scenes (2016). arXiv:1601.00998
7. Alahi, A., et al.: Social LSTM: human trajectory prediction in crowded spaces. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016)
8. Raju, K., Chang, M.: Predicting Human Trajectories in Multi-Class Settings (2016)

9. Varshneya, D., Srinivasaraghavan, G.: Human Trajectory Prediction using Spatially aware Deep Attention Models (2017). arXiv:1705.09436

10. Yi, S., Li, H., Wang, X.: Pedestrian behavior understanding and prediction with deep neural networks. In: European Conference on Computer Vision. Springer International Publishing (2016)

An Automatic Face Attractiveness Improvement Using the Golden Ratio

Hiranmoy Roy, Soumyadip Dhar, Kaushik Dey, Swaroop Acharjee and Debanjana Ghosh

Abstract Charles Darwin once wrote: 'It is certainly not true that there are in the mind of man any universal standards of beauty with respect to the human body'. The relation between facial beauty and the golden ratio is a known fact. In this paper, we have tried to establish the relation between face beauty and the golden ratio. Finally, we try to improve facial beauty using the golden ratio-based geometric transformation and some filtering operations. The work is divided into two parts: 1. verification of the relation between face beauty and golden ratio, 2. application of golden ratio for face beautification. The first part of the paper is based on the verification of a neoclassical theorem of beauty and the golden ratio based on the symmetry of the face, using various machine learning tools. Verification of the ratings is done using SCUT-FBP dataset. We used 450 images for the training purpose out of the 500 images and the rest 50 images are used for testing the data. The second part of the work is to beautify a face based on mathematical calculations and improve the skin texture, removes blemishes, and change the facial features according to the golden ratio. Test results show the significant improvement in facial beauty due to the application of the golden ratio.

Keywords Face beauty · Golden ratio · Pearson correlation · Random forest

1 Introduction

Attractiveness is a general and important topic to each and every society. Although the recognition of attractive people are universal, beauty varies from culture to culture. There is a measurable standard for beauty and what is generally found most desirable [1].

H. Roy (✉) · S. Dhar · K. Dey · S. Acharjee · D. Ghosh
RCC Institute of Information Technology, Kolkata 700015, India
e-mail: hiru.roy@gmail.com

Fig. 1 Pictorial representation of the golden ratio calculation
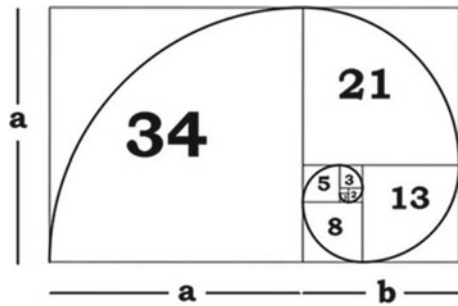
Faces with a high degree of symmetry are typically considered more attractive [2, 3]. In the literature, many different ways are used to measure the effects of symmetry and average on attractiveness, such as by generating composite faces, by altering the faces using positional changes of facial components.

Golden ratio is a special decimal number found when a line is divided into two parts, then is the ratio of the longer part to smaller part is also equal to the ratio of the whole line to the longer part and it is shown in Fig. 1. Golden ratio is often symbolized using the twenty-first letter of the Greek alphabet 'Phi' (ϕ). Let us assume that a line is divided into 'a' and 'b' where $(a > b)$, so golden ratio will be represented as follows:

$$\frac{a}{b} = \frac{(a + b)}{a} = 1.6180339887498948420\ldots \tag{1}$$

Leonardo da Vinci's 'Vitruvian Man', 'Mona Lisa' is said to illustrate the golden ratio. It is said that every geometrical structure with the golden ratio are said to have a perfect geometric structure. This paper is organized as follows: in Sect. 2, the proposed method is described in detail. Experimental results and comparisons are presented in Sect. 3, and finally, the paper concludes with Sect. 4.

2 Proposed Method

In the Fig. 2a we can see facial points, these points are the landmarks we would be using to calculate golden ratio of a given image and thus beautify it. Our work is divided into two parts: verification and application of golden ratio. SCUT-FBP is that database we have used over here. It consists of images which are pre-rated by a group of personnel. This rating will serve as the ground truth of our project which will be verified against the predicted truth of our proposed method.
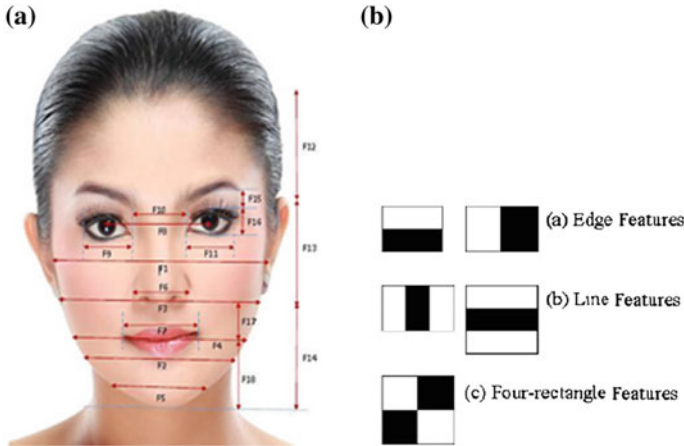
Fig. 2 **a** Example showing 18 geometric features, **b** three different Haar features

2.1 Facial Beauty Prediction

In this subsection, we evaluate the prediction performance of different algorithms on the basis of several criteria such as Pearson correlation (PC) [4], mean absolute error (MAE) [5] and root mean squared error (RMSE) [5]. The machine learning methods we used include SVM, linear regression, random forest regression and Gaussian regression.

We randomly selected 450 images from our SCUT-FBP dataset for training, and the remaining 50 images were used for testing. This testing resulted in the verification of our machine learning. Based on this machine learning, we can now take an image and apply this training set to get the facial landmarks of the image and append it with the existing database.

This testing can be done using any of the regression methods as follows:

- **Linear Regression**: In statistic, linear regression is the approach of modelling the relationship between the scalar dependent variable Y to independent variables denoted by X. In case of multiple independent variable the process is called multiple linear regression. To model the relationships, a linear predictor function and its unknown parameters are estimated from the training data. If the regression is based on straight line graph, then it is known as linear regression. Linear regression of 'y' on 'x' for 'n' number of variables can be written as follows:

$$y = a + bx, \quad where \quad a = \frac{\sum y_i - b \sum x_i}{n} \quad and \quad b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} \quad (2)$$

- **Support Vector Machine**: It is used both as a classification and regression purposes to analyse data in machine learning. In the training phase, SVM tries to generate a plane to separate the different training points plotted in a space. The plane is placed in such a way, so that it has almost equal distance from each and every different category of classes. Then, in testing phase, all the testing points are predicted in the same space by its belonging to the side of the plane.

- **Random Forest Regression**: It is a well-known ensemble learning mechanism, which is mainly used to boost the classification accuracy of decision tree. A set of decision tree classifiers are ensemble to generate a random forest regression or random decision forests [6–8]. At the time of training, a decision tree was generated depending on the training samples. The branches are the conditions and the leaf nodes are the decisions. From the tree some rules are generated. At testing phase, when a test sample is coming, according to the rules the output decision is taken. In random forest regression, a set of such decisions are ensembles.

- **Gaussian Regression**: It is also a learning mechanism, but a little bit lazy in the sense that it delayed the learning process until a query is given. It measures the similarity between different training points present in a local kernel. Then, from the similarity value, it predicts the unseen training points. The predicted value has both the estimated information and the uncertainty information regarding the training points. Gaussian regression provides a marginal distribution in a one-dimensional Gaussian distribution form.

Finally, using different statistical methods we selected random forest as our best learning strategy.

2.2 Golden Ratio-Based Facial Beautification

The second step of our project is based on beautification of a face using the golden ratio. Beautification of the face is done at first by detecting the face then applying Cascade Classifier to detect the eyes and face and finally applying bilateral filter and Gaussian Blur to provide an even skin tone and remove blemishes.

- **Haar Cascade**: At first, Haar cascade object detection method is used to detect the face from the image. Haar cascade is an ensembles-based machine learning approach, where a cascade function is used to train the system for a lot of face and non-face image samples. Then, the trained system is used to detect faces in other test image samples. Figure 2b shows three different Haar features used here for face detection.

- **Bilateral Filter**: Bilateral filter provides an even skin tone to the face. Bilateral filter changes the pixel intensity value of to a pixel by average weighted value of a nearby pixel. Generally, the weight value follows some Gaussian distribution. This Gaussian distribution-based weight has the sharp edge preservation property.

- **Gaussian Blur**: It is a widely accepted filtering technique used for image blurring or noise removal. The basic idea is to use a Gaussian function (in statistics it is called normal distribution) for measuring the transformation by applying it as a convolution operation for each image pixel. After applying this filter, the image becomes smooth and high-frequency noises are removed.

2.2.1 Algorithm: Facial Landmark Verification and Learning

Step 1: Collection of a dataset of images. (In this case, we are using SCUT-FB database of 500 pictures of Asian Females and their corresponding ratings).
Step 2: Detection of a face from each image.
Step 3: Detection of the facial landmarks from the detected face in each image.
Step 4: Save all the landmarks points in a text file.
Step 5: Calculate the facial landmark ratio of each face and store the data in another dataset.
Step 6: The dataset is divided into 450:50 ratio and 450 images are used training purpose and 50 are used for testing purpose.
Step 7: Machine learning algorithm tools like SVM, linear regression, random forest and Gaussian process are used to predict the facial attractive index of each face out of 5.
Step 8: Predicted score and ground reality of the 50 images are tested and Pearson correlation and residual is found out from the data.

2.2.2 Algorithm: Facial Beautification

Step 1: The image to be beautified according to the algorithm is uploaded.
Step 2: Face is detected in the image.
Step 3: Facial landmark is extracted.
Step 4: Then ratio values are calculated and collected.
Step 5: The ratio is perfected according to the defined algorithm
Step 6: The image ratio of the face is changed accordingly.
Step 7: Skin gradient, blemishes, grey hair, skin spots are reduced and removed from the image.
Step 8: Image is saved with a different name.
Step 9: Facial landmarks of the given new image is stored as to increase the accuracy of the dataset.

3 Experimental Results

We collected data to build a standard dataset that provides unified data for evaluating the performance of different algorithms. To reduce the effects of irrelevant factors

Fig. 3 Different levels of beauty in the SCUT-FBP dataset

such as age, gender and facial expression, the SCUT-FBP dataset [9] is confined to a unified form, i.e. it contains high-resolution, front on face portraits of Asian female subjects with neutral expressions, simple backgrounds, no accessories, and minimal occlusion. A previous study [2] has shown that beautiful individuals constitute a small percentage of the population. The SCUT-FBP dataset contains a higher proportion of beautiful faces than that in the general population in order to facilitate effective learning of facial beauty. Specifically, it contains 500 portraits, some of which we captured ourselves; others were licensed from different sources or downloaded from the Internet. All the images were rated by numerous raters. Figure 3 shows some examples of face portraits from the dataset. SCUT-FBP dataset publishers have already developed a web-based tool, namely, the facial beauty assessment system, to collect ratings. Images in the SCUT-FBP dataset were rated by 75 raters; the average number of raters per image was 70. Because the evaluation ground truth varied among individuals, we obtained raters opinions regarding the beauty of the portraits by asking them for answers to certain questions. The portraits were randomly shown to the raters. The raters could change their ratings if they accidentally selected an incorrect option. Although facial beauty has been shown to be a universal concept, it is subjective to some extent. The procedure described above aims to eliminate unnecessary effects.

Facial attractiveness rating can be regarded as a regression problem. Therefore, we compared the ground truth results provided by the SCUT-FBP web tool with
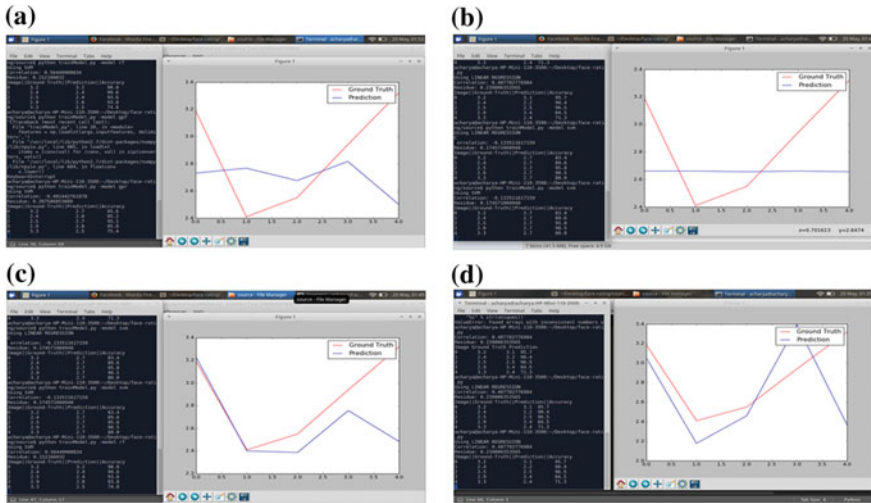
Fig. 4 Comparison results of ground truth provided by SCUT-FBP web tool and prediction given by proposed method using: **a** Linear progression learning, **b** SVM learning, **c** random forest learning and **d** Gaussian regression

our prediction. Different comparison results are provided in the Fig. 4. Comparison results show that random forest and Gaussian regression methods provide more accurate results. Another testing is done using Pearson correlation and graph residual [8] methods to choose the best learning method. Pearson Correlation (r) shows the linear relationship between two sets of data. For two datasets 'x' and 'y' with population size 'n', the value of 'r' is measured as follows:

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{[n\sum x^2 - \left(\sum x\right)^2][n\sum y^2 - \left(\sum y\right)^2]}} \tag{3}$$

The closer the value of 'r' gets to one, the greater the correlation between the data.

Again, the difference between the observed value of the dependent variable (y) and the predicted value ($\hat{y}$) is called the residual (e) $e = y - \hat{y}$. The sum of e, i.e. $\sum e$ and the mean of e i.e. $\frac{1}{n}\sum e$ both are equal to zero. A graph having the independent variable on the horizontal axis and the residuals on the vertical axis gives a residual plot. A linear regression model is appropriate for the data, when the points in a residual plot are randomly dispersed around the horizontal axis; otherwise, a nonlinear model is more appropriate.

Figure 5 shows the comparison results of these two methods on different learning techniques. From the results, again it is proved that random forest and Gaussian regression methods provide more accurate results. We choose random forest as our best learning strategy. Finally, in the second part, we applied the proposed algorithm
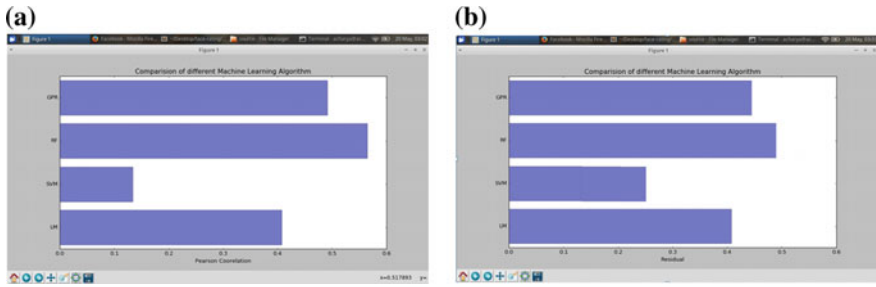
Fig. 5 Comparison of different learning methods using: **a** Pearson correlation, **b** graph residual methods
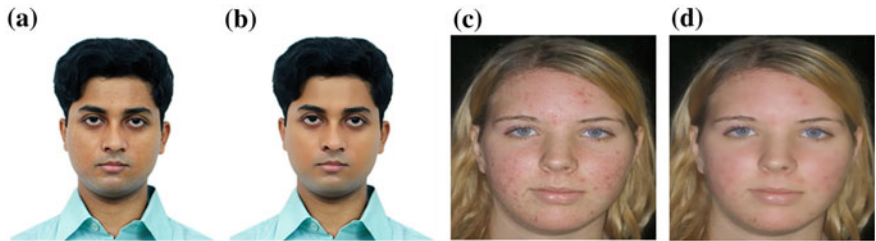
Fig. 6 One sample example: **a** Original face image, **b** face image after applying the proposed method, **c** original face image, **d** face image after applying proposed method

(Facial beautification) for face beautification. Figure 6 shows some sample examples, where the proposed method is applied to improve face beauty.

4 Conclusion

It can be concluded from our experiment that facial symmetry plays an important role in attractiveness of a face. This can be noted from the minimal difference between the predicted ratings and actual ratings out of 5 of the facial attractiveness. While beauty lies in the eye of the beholder, but in general it can be quantitatively said that a symmetric face is attractive in general. However, symmetry is not the sole parameter upon which facial attractiveness can be measured. Other factors like skin textures, skin gradient and facial hair play an important role.

We have not taken the skin gradient, skin textures into consideration. Similarly, our method is only able to predict the attractiveness index based on a single ethnicity due to the smaller size of the training dataset. It will fail considerably in case of males with facial hair. Proposed method fails when it comes to an image of a person wearing spectacles and angle of the face plays an important role in the efficient working of the program. In future, a modified version of the proposed method can be thought.

References

1. Otoole, A.J., Price, T., Vetter, T., Bartlett, J.C., Blanz, V.: 3D shape and 2D surface textures of human faces: the role of averages in attractiveness and age. Image Vis. Comput. **18**, 9–19 (1999)
2. Atiyeh, B.S., Hayek, S.N.: Numeric expression of aesthetics and beauty. Aesthet. Plast. Surg. **32**(2), 23–35 (2001)
3. Farkas, L.G.: Anthropometry of the Head and Face, 2nd edn. Raven Press, New York (1994)
4. Barnett, T.P., Preisendorfer, P.: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. Mon, Weather Rev (1997)
5. Ho, T.K.: Random decision forests. In: Proceedings of the 3rd International Conference on Document Analysis and Recognition, pp. 278–282 (1995)
6. Ho, T.K.: The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Int. **20**(8), 832–844 (1995)
7. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning, 2nd Ed. Springer (2008)
8. Farkas, L.G., Munro, I.R.: Anthropometric Facial Proportions in Medicine. Illinois (1987)
9. SCUT-FBP dataset. http://www.hcii-lab.net/data/SCUT-FBP/

A Trust-Based Intrusion Detection System for Mitigating Blackhole Attacks in MANET

Biswaraj Sen, Moirangthem Goldie Meitei, Kalpana Sharma, Mrinal Kanti Ghose and Sanku Sinha

Abstract MANETs (Mobile Ad hoc Networks) are wireless networks that are deployed for a particular purpose or short-term use. Because of the lack of central coordination, MANETs share an inherent trust relationship among the nodes forming the network. Each node implicitly trusts its neighbour to forward packets in the network till the packets reach their destination. Further, each node in a MANET can monitor its neighbours by keeping track of the packets passing through the neighbours. This characteristic ability of a node in MANET makes it possible to develop a trust model that can correlate with the innate trust shared among the nodes. This paper looks at developing such a trust model which is applied to all the nodes in the network. The trust model works like an Intrusion Detection System (IDS), which seeks to detect blackhole attacks in the system, and then identify and mitigate the malicious attacker.

Keywords Blackhole attack · MANET · Intrusion detection system
Threshold · Trust model

B. Sen (✉) · M. G. Meitei · K. Sharma
Computer Science and Engineering, Sikkim Manipal University, Rangpo, India
e-mail: biswaraj.sen@gmail.com

M. G. Meitei
e-mail: mgmeitei@gmail.com

K. Sharma
e-mail: kalpanaiitkgp@yahoo.com

M. K. Ghose
Computer Application, Sikkim University, Gangtok, India
e-mail: mkghose2000@yahoo.com

S. Sinha
Information Technology, DDE, Sikkim Manipal University, Rangpo, India
e-mail: sinhasanku@yahoo.com

1   Introduction

Ad hoc networks are wireless networks that are constructed for a particular purpose or an immediate need. Such networks differ from traditional networks in that they do not require a centralized coordinator or prior infrastructure to be in place. Thus, ad hoc networks are also called *infrastructureless networks* [1]. A Mobile ad hoc Network(MANET) refers to a network in which the nodes forming the ad hoc network are mobile [2]. Because of its characteristic properties, MANETs have been used in areas such as emergency search and rescue operations, military battlefields and in academic and commercial sectors [3, 4].

In a MANET, nodes cooperate with each other to share information. A node wanting to send information transmits the information to its neighbour which in turn propagates it to its neighbours until it reaches the required destination. This system places an inherent trust among other nodes in the network for information propagation. An attacker can take advantage of this trust relationship among the nodes, thereby compromising the network. Also, due to the mobility of the nodes and the dynamically changing network topology, it is hard to determine if a packet is dropped because of the intrinsic network characteristics or the presence of an attacker.

This paper briefly discusses the latent trust relationship among nodes in a MANET and seeks to develop a trust model to handle threats. The rest of the paper is organized as follows: Sect. 2 discusses the trust and intrusion detection systems in MANETs. Section 3 gives a brief explanation of the blackhole attack. Section 4 highlights some of the work done in this area. Section 5 describes the proposed trust model. Section 6 discusses the results. Section 7 provides the conclusion.

2   Trust and Intrusion Detection System (IDS) in MANETs

MANETs face vulnerabilities because of the shared wireless medium, lack of physical protection for the mobile nodes and complete trust among nodes because of lack of centralized decision-making entity [5]. MANETs are susceptible to DoS attacks as they do not have a clear line of defence [6, 7]. Ad hoc networks operate by establishing an intrinsic trust relationship among its participating nodes. Hence, each node in a MANET is able to function as a router. Each node in a MANET completely trusts its neighbours to carry out network activities such as packet forwarding and packet delivery until each packet reaches the intended destination. Often, attackers try to take advantage of this particular trait present in the nodes in a MANET. Thus, managing trust also becomes an important issue [8, 9].

Intrusion detection can be defined as a process of monitoring activities in a system, which can be a computer or network system [10]. An IDS monitors and collects network activity information and then analyses it to check for any anomalous behaviour in the network. Intrusion detection can be categorized into
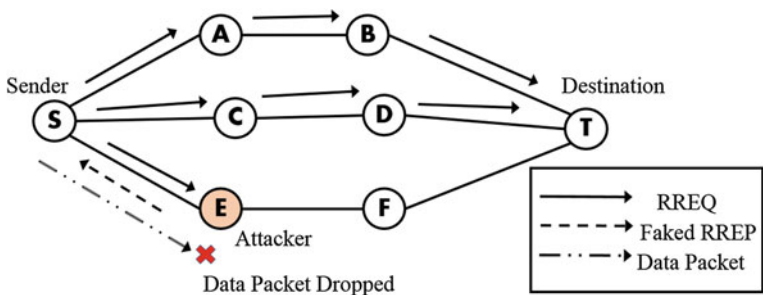
Fig. 1 Blackhole attack

two methods: anomaly detection and misuse detection. Anomaly detection is the method of monitoring the network for deviations from normal behaviour while misuse detection (also called signature-based detection) uses databases that contain signatures or patterns of known attacks [11] (Fig. 1).

3 Blackhole Attack

A blackhole attack is a DoS attack in which a malicious node falsely claims that it has the shortest path to the destination node. This attack is carried out by an attacker sending fake routing information [12]. When an attacker node receives a Route Request message from a sender node, it replies to the Route Request message with a Route Reply having a very high destination sequence number, hence ensuring that the attacker gets included in the route from the sender to the destination. On receiving the subsequent data packet from the sender, the attacker will not forward the data packets but instead drop them, thus preventing them from reaching the intended destination.

4 Related Work

Huang and Lee [13] proposed an intrusion detection system based on their previous work on anomaly detection, which used cross-feature analysis to detect intrusions [14]. Their work focused on detecting anomalies by implementing IDS on every node, and anomaly detection by implementing IDS for a cluster-based system. Trivedi et al. [15] proposed a detection mechanism based on reputation to deal with intrusions in MANETs. Their proposed mechanism has been termed as Reputation-based Intrusion detection System for Mobile ad hoc networks (RISM), which is a modification of the CONFIDANT protocol [16]. Nadeem and Howarth [17] proposed an IDS mechanism called Intrusion Detection and Adaptive

Response (IDAR), which uses both anomaly detection and knowledge-based intrusion detection. Hu et al. [18] proposed Rushing Attack Prevention (RAP), which is a generic route discovery mechanism for handling rushing attacks. Prathapani et al. [19] proposed the use of mobile honeypot agents to detect blackhole attacks in Wireless Mesh Networks (WMNs).

5 Proposed Strategy

The overview of the proposed strategy is shown in the following diagram (Fig. 2).
The proposed methodology consists of the following phases:

Phase 1: **Initial Network Deployment**

Initially, the network is made to function on its own using a standard routing protocol. This helps in calculating a baseline observation of the network under normal circumstances. Once we know how the network behaves normally, we can differentiate it from the conditions arising due to malicious behaviour in the network.

Phase 2: **Attack on Network**

The network is now subjected to a routing attack, viz. blackhole attack. The attack is carried out in four scenarios as follows:

(a) One blackhole attacker.
(b) Ten percent of total nodes as blackhole attackers.
(c) Twenty percent of total nodes as blackhole attackers.
(d) Thirty percent of total nodes as blackhole attackers.
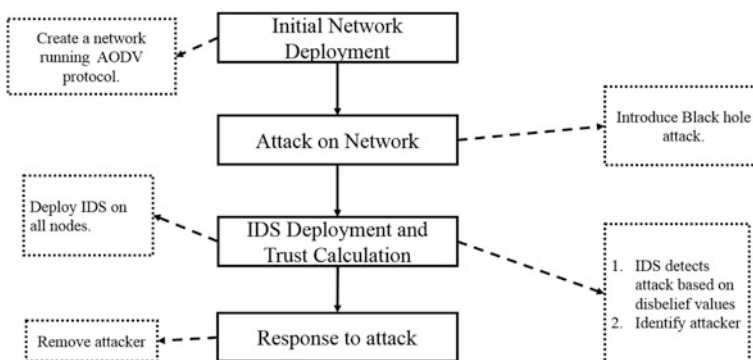
Fig. 2 Workflow diagram of proposed methodology

Phase 3: **IDS Deployment and Trust Calculation**

In this phase, IDS is deployed in all the nodes of the network. Each IDS node keeps track of the trust levels of its immediate neighbours. This tracking is done at runtime. The trust values at are calculated based on three parameters:

(a) The belief that a node has for its neighbour (b): The belief factor is calculated by taking into account positive events that occur during a transmission. Here, positive events are the events that signify a successful transmission in the network. The positive events chosen are the following:

   (i) Successful packet reception
   (ii) Successful packet forwarding

The belief factor is calculated as follows:

$$b = \frac{p}{p+n+2},$$ (1)

where p = number of positive events and n = number of negative events.

(b) The disbelief that a node has for its neighbour (d): Contrarily, the disbelief factor is calculated by taking into account negative events that occur during a transmission. Negative events are the events that signify an unsuccessful transmission in the network. The negative events chosen are the following:

   (i) Unsuccessful packet reception
   (ii) Unsuccessful packet forwarding

The disbelief factor is calculated as follows:

$$d = \frac{n}{p+n+2},$$ (2)

where p = number of positive events and n = number of negative events

(c) The uncertainty that a node has for its neighbour (u): Uncertainty factor is initially set to 1 before any transmission begins after a node has just discovered its neighbours.

The uncertainty factor is calculated as follows:

$$u = \frac{2}{p+n+2},$$ (3)

where p = number of positive events and n = number of negative events
Hence, these three parameters are taken in such a way that

$$b + d + u = 1$$ (4)

at all times.

So initially, a node will have uncertainty value of 1, belief value of 0 and disbelief value of 0 for its neighbour before transmission. As communication begins in the network, these values get updated based on positive and negative events. This trust value calculation is done periodically.

An anomaly in the system is detected when the disbelief factor rises above a certain threshold. In this case, the anomaly is first verified as an attack by applying identification rules for recognizing attacks such as a blackhole attack. This step ensures that network congestion factor is taken into consideration. Once positively identified as an attack, the node under question will be treated as malicious and it will not be allowed participate in the network.

Threshold calculation

The threshold value is set based on experimental values. This is done by calculating the average value of Packet Data Fraction (PDF) of several simulations of the network in its initial phase. This provides a measure of how the network performs normally in the absence of any malicious attacker. This average PDF value thus obtained serves as a threshold for discovering anomalous behaviour in the network.

Phase 4: **Response to attack**

When an attack such as a blackhole attack occurs, the intrusion is detected by the periodic update of the trust values. Once the attacker node is identified, it is not allowed to participate in routing and will be removed from the network. The node will then seek alternate routes to reach the destination after removing the attacker.

This process is explained as follows:

(a) Since each node will be running IDS, each node can monitor its neighbour's activities. Hence, each node keeps track of the belief, disbelief and uncertain factors of its neighbours.

(b) If the disbelief factor of a certain neighbour node rises above the calculated threshold value (as explained above), then appropriate action is taken by

- Identifying the attack and the attacker: This is necessary to differentiate between network congestion and a routing attack, viz. blackhole attack in this case. For this, the following formula for identifying blackhole attack is used:

$$PFP = \frac{pr(n)}{ps(nn)}, \tag{5}$$

where

PFP   packet forward percentage
pr(n)   no. of packets received by a node n
ps(nn)   no. of packets sent by n's neighbours and not destined for n

If n keeps dropping packets for a sufficiently long period, or more precisely, if the denominator is not zero and PFP = 1, then a blackhole is detected and n is identified as the attacking node.

- Removing the attacker from the routing process: Once the attacker is identified, the node that detected the attacker removes the malicious node from the network.

6 Results

Simulations were carried out in ns-2.35 using AODV protocol. The simulation parameters are as follows (Table 1).

The results of the simulations can be viewed in the following graphs (Figs. 3, 4, 5, 6, 7 and 8).

For 30 and 60 nodes, the proposed solution gives better performance in terms of throughput and PDF as compared to the native AODV protocol when both are exposed to blackhole attack. This is because the proposed method is able to detect, identify and remove the attacker, thereby increasing the throughput and the PDF.

But in terms of delay, for 30 nodes, the proposed solution suffers in performance as compared to the native AODV protocol. This is because the proposed method has to find a new path after detecting and removing the attackers from the network, thereby increasing the delay. However, for 60 nodes, the proposed solution experiences lesser delays as compared to the native AODV protocol because the network is denser as compared to 30 nodes and alternate routes to the destination can be found faster.

Table 1 MANET simulation parameters

Method	Type
Channel type	Channel/wireless channel
Radio propagation	Propagation/two-ray ground
Network interface	Phy/WirelessPhy
Mac type	Mac/802_11
Interface queue	Queue/Drop Tail/PriQueue
Link layer type	LL
No. of nodes	30, 60
Routing protocol	AODV
Area	1000 * 1000 sq. m.
Simulation Time	1200 s
Mobility	Random waypoint
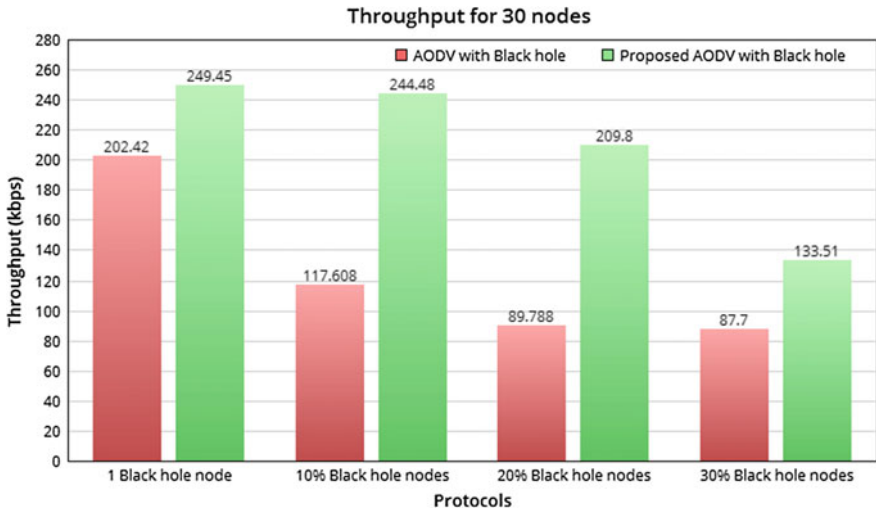
Fig. 3 Throughput of 30 nodes
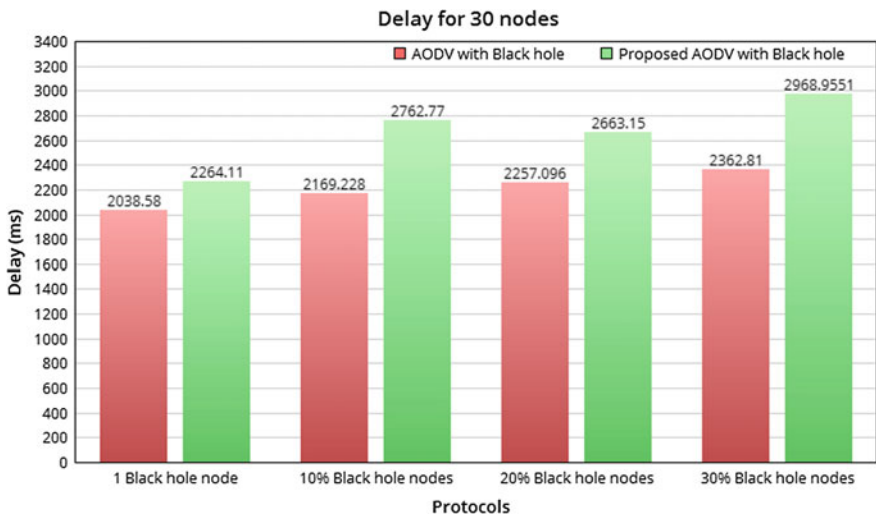
Fig. 4 Delay of 30 nodes
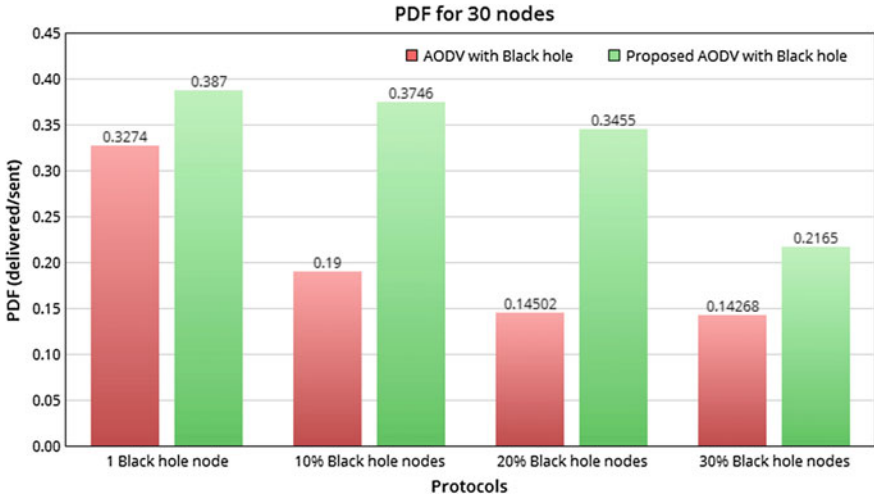
Fig. 5 PDF of 30 nodes
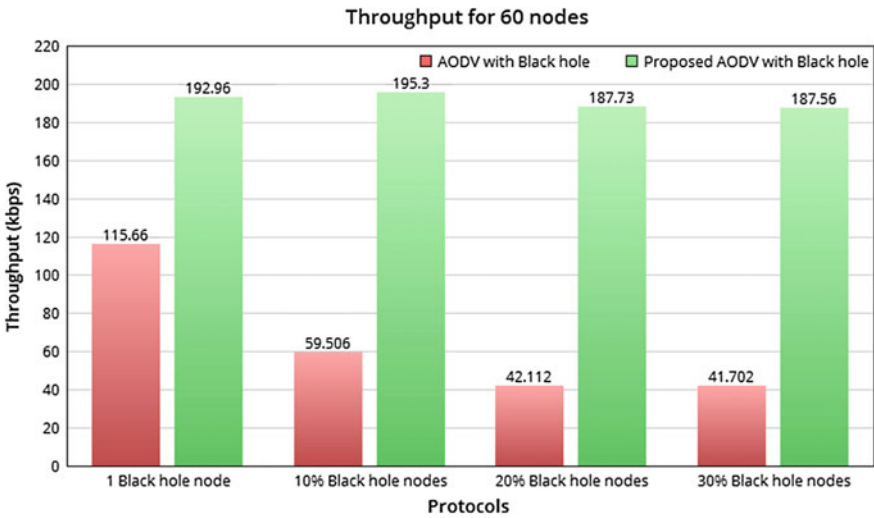
Fig. 6 Throughput of 60 nodes
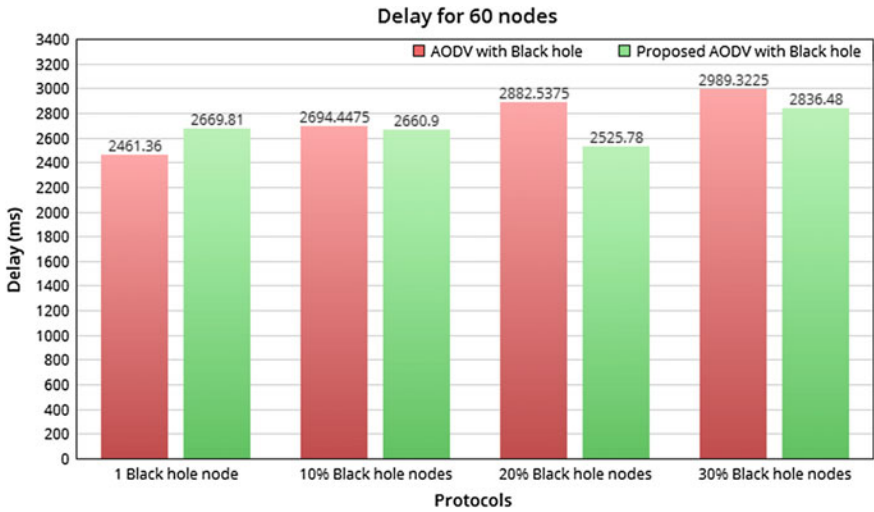
Fig. 7 Delay of 60 nodes
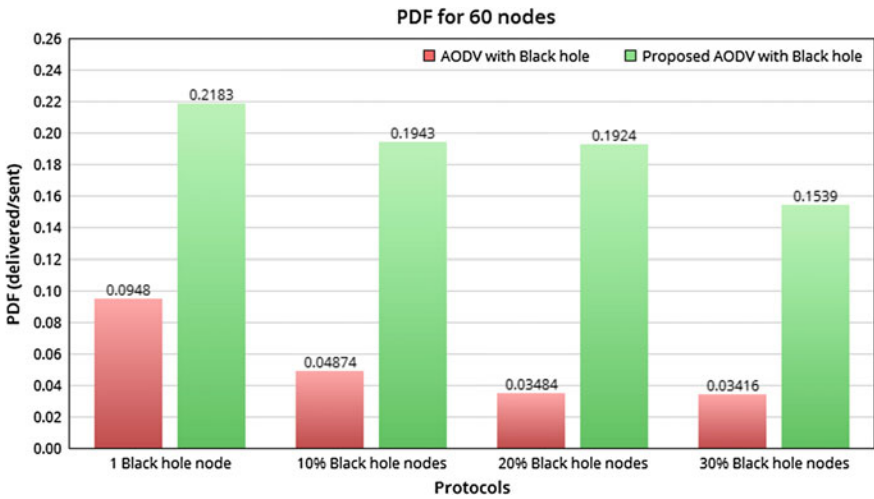
Fig. 8 PDF of 60 nodes

7 Conclusion

MANETs are susceptible to different kinds of attacks and threats because of their characteristic properties such as trust-based relationship and lack of central coordination. It is possible to use this characteristic feature of MANET and devise a trust model to monitor network activity. Therefore, this document has looked at

security measures in MANETs and devised a trust-based IDS system against blackhole attacks. The proposed mechanism is able to provide a substantial improvement in the affected network in terms of throughput and PDF, although it experiences higher end-to-end delays.

References

1. Deng, H., Li, W., Agrawal, D.P.: Routing security in wireless ad hoc networks. IEEE Commun. Mag. **40**(10), 70–75 (2002)
2. Chandra, P.: Bulletproof Wireless Security: GSM, UMTS, 802.11, and Ad Hoc Security. Elsevier (2011)
3. Murthy, C.S.R., Manoj, B.S.: Ad hoc Wireless Networks: Architectures and Protocols, Portable Documents. Pearson education (2004)
4. Güneş, M., Juraschek, F., Blywis, B., Mushtaq, Q., Schiller, J.: A testbed for next generation wireless network research. PIK-Praxis der Informationsverarbeitung und Kommunikation **32**(4), 208–212 (2009)
5. Zhang, Y., Lee, W., Huang, Y.A.: Intrusion detection techniques for mobile wireless networks. Wirel. Netw. **9**(5), 545–556 (2003)
6. Yang, H., Luo, H., Ye, F., Lu, S., Zhang, L.: Security in mobile ad hoc networks: challenges and solutions. IEEE Wirel. Commun. **11**(1), 38–47 (2004)
7. Wu, B., Chen, J., Wu, J., and Cardei, M.: A survey of attacks and countermeasures in mobile ad hoc networks. In: Wireless network security (pp. 103–135). Springer US (2007)
8. Li, W., Parker, J., Joshi, A.: Security through collaboration and trust in manets. Mob. Netw. Appl. **17**(3), 342–352 (2012)
9. Cho, J.H., Swami, A., Chen, R.: A survey on trust management for mobile ad hoc networks. IEEE Commun. Surv. Tutor. **13**(4), 562–583 (2011)
10. Anantvalee, T., Wu, J.: A survey on intrusion detection in mobile ad hoc networks. In: Wireless Network Security, pp. 159–180. Springer US (2007)
11. Nishani, L., Biba, M.: Machine learning for intrusion detection in MANET: a state-of-the-art survey. J. Intell. Inf. Syst. **46**(2), 391–407 (2016)
12. Kannhavong, B., Nakayama, H., Nemoto, Y., Kato, N., Jamalipour, A.: A survey of routing attacks in mobile ad hoc networks. IEEE Wirel. Commun. **14**(5), 85–91 (2007)
13. Huang, Y.A., Lee, W.: A cooperative intrusion detection system for ad hoc networks. In: Proceedings of the 1st ACM Workshop on Security of Ad hoc and Sensor Networks, pp. 135–147. ACM (2003)
14. Huang, Y.A., Fan, W., Lee, W., Yu, P.S.: Cross-feature analysis for detecting ad-hoc routing anomalies. In: Proceedings of the 23rd International Conference on Distributed Computing Systems, 2003, pp. 478–487. IEEE (2003)
15. Trivedi, A.K., Kapoor, R., Arora, R., Sanyal, S., Sanyal, S.: RISM–Reputation Based Intrusion Detection System for Mobile Ad hoc Networks (2013). arXiv:1307.7833
16. Buchegger, S., Le Boudec, J.Y.: Performance analysis of the CONFIDANT protocol. In: Proceedings of the 3rd ACM International Symposium on Mobile Ad hoc Networking & Computing, pp. 226–236. ACM (2002)
17. Nadeem, A., Howarth, M.P.: An intrusion detection & adaptive response mechanism for MANETs. Ad Hoc Netw. **13**, 368–380 (2014)
18. Hu, Y.C., Perrig, A., Johnson, D.B.: Rushing attacks and defense in wireless ad hoc network routing protocols. In: Proceedings of the 2nd ACM Workshop on Wireless Security, pp. 30–40. ACM (2003)
19. Prathapani, A., Santhanam, L., Agrawal, D.P.: Detection of blackhole attack in a Wireless Mesh Network using intelligent honeypot agents. J. Supercomput. **64**(3), 777–804 (2013)

Good-Quality Question Generation for Academic Support

Manisha Divate and Ambuja Salgaonkar

Abstract The paper presents a metric to automatically compute a score for machine-generated questions and transforms the questions which are having a lower score value, unacceptable, and ungrammatical into a human appealing form. Questions are unacceptable due to the flaws like incorrect grammar, selection of wrong wh-phrase, partial selection of answer phrase, negation, etc. Identifying such infirmities in the question is a challenge. Here, our attempt is to automatically detect and correct the flaws present in the question. We named this system as the Automatic Question Quality Enhancer (AQQE). By employing a multiple linear regression model, AQQE first computes the score (in range of 1–rejected to 5–accepted) for 174 questions. Higher score value tells the acceptance and lower score value shows the rejection of the question. AQQE's challenge is to enhance the quality of questions having a lower score. Out of 174, human evaluator had identified 84 questions as acceptable and 90 (51.72%) as an unacceptable. Performance of AQQE is judged with precision and recall and it is found well acceptable. AQQE enhanced 79(87.77%) questions are accepted by the human evaluator and 11 (6%) questions can be accepted with further modifications.

M. Divate (✉) · A. Salgaonkar
Department of Computer Science, University of Mumbai, Kalina, Santacruz,
Mumbai, India
e-mail: divate.manisha.79@gmail.com

A. Salgaonkar
e-mail: ambujas@gmail.com

1 Introduction

Question generation is a challenging job. The time required to generate a good-quality question will vary from tutor to tutor. Tutor asks questions to a learner to generate a feedback about the topic understanding. In the classroom, more than 90% of the questions asked are fact-based (or shallow learning) questions [1].

Question Generation (QG) system assists the learner in understanding the topic, to enrich the vocabulary, improvise the grammar. Table 1 shows the list of Automatic Question Generation (AQG) systems used for various academic purposes.

Literature survey reveals that the acceptance rate of various fact-based AQG system is nearly 60%. Table 2 shows the study of AQG's from the past decades with their acceptance rate. Productivity and question quality enhancement is the

Table 1 AQG in academic use

AQG system	Academic use
[2–4]	Vocabulary assistance
[5]	Reading comprehension
[6]	Comprehension assessment
[7]	English-language learning
[8, 9]	English grammar test, langauge Test
[10]	To test the english proficiency of nonnative speakers
[11]	Literature review writing support

Table 2 AQG with acceptance rate

AQG system	Types of AQG	Acceptance rate
[3]	Gapfill question	52.86%
[7]	Gapfill questions	93%
[11]	Factoid using template approach	60%
[12]	Factoid using syntactic approach	58%
[13]	Factoid using semantic–syntactic approach	46.5% for top 15% rank question 40% for top 30% rank question
[14]	Factoid using syntactic approach	43.3%
[15]	Factoid using template- semantic-based approach	23% questions with learning value 34% cross-validation result for precision −10
[16]	Factoid using syntactic approach	68%
[17]	Factoid using template–semantic approach	52%
[18]	Factoid using semantic–syntactic approach	46%
[19]	Factoid using template approach	53%

solution for the present AQGs and here we have proposed a system, AQQE, which enhances the quality of the unacceptable question.

Organization of the paper is as follows: Sect. 2 discussed the proposed methodology of question enhancement. Results are discussed in Sect. 3. The challenges encountered in defining the system is presented in Sect. 4 followed by the summary in Sect. 5.

2 Proposed Methodology

A question is *Ill-Formed* (IFQ), i.e., unacceptable, if one or more flaws of following types are observed in its articulation: incorrect grammar, semantic inadequacy, vagueness with respect to the answer, inadequate data, a wrong choice of wh-form while presenting a question, or some editing is needed [20, 21]. A question is *Well-Formed* (WFQ), hence acceptable, if it is not ill-formed, i.e., it does not have any of the 6 flaws mentioned above.

The objective of AQQE is to formulate a well-formed question (WFQ) automatically from the Ill-formed questions (IFQ).

Illustration of AQQE process:
Base Sentence (BS): As is the case in a parliamentary system, the government is formed by the party, alliance, or group of assembly members who command the majority.
Answer Phrase (AP): by the party, alliance, or group of assembly members who command the majority.
Q_{HSAQG}: What is the government formed by?
Q_{HSAQG} is grammatically correct, but it does not answer the same AP. Wh-phrase of a question should be who because AP (*the group of persons* who forms the government) is a person entity. Here, AQQE suggests the correct wh-phrase as who and frames Q_{AQQE} as: Who forms the government? The flaw detected is the wrong choice of wh-phrase [22].

Here, we state the hypothesis of the present study that the mean score value of a question generated through AQQE (i.e., score of Q_{AQQE}) is greater than the score value of a question before enhancing it (score of Q_{HSAQG}) with a significant level of 0.05.

μ_{HSAQG}: score value of HSAQG-generated question
μ_{AQQE}: score value of question enhanced by AQQE
μ_d = difference between the two score values μ_{AQQE} and μ_{HSAQG}
Null hypothesis is $H_0 = \mu_d = 0$, i.e., $[\mu_{AQQE} - \mu_{HSAQG}] = 0$
$H_1 = \mu_d > 0$, i.e., $[\mu_{AQQE} - \mu_{HSAQG}] > 0$

2.1 System Architecture

The basic framework of AQQE is shown in Fig. 1. The input set consists of BS, AP, and the questions generated by HSAQG. The experiment takes such 174 datasets [23].

Human judges the acceptance or rejection of questions based on the infirmities present in it. Similarly, the proposed score module judges the infirmities present in the questions basis on the feature set. AQQE defines features such as answer_length, wh-phrase_as_per_ans, leadingPP, ansPronoun, PPInsideNP, etc. With these feature values, AQQE's score module computes the score for the questions. The features value for the question illustrated above are calculated as answer_length = 0, wh-phrase_as_per_ans = 0, leadingPP = 1, ansPronoun = 1, PPInsideNP = 1.

The binary value of the parameters 1, here, indicates that the Q_{HSAQG} is grammatically correct while the parameters with value zero, indicating both, abnormal length of the AP and inconsistency of wh-phrase with respect to the AP, that models the unacceptability of Q_{HSAQG} because it does not answer the same as AP.

Human evaluator had rated the question on a 5-point scale (1–rejected, 5–accepted). To predict the scale of a question automatically, AQQE uses multiple linear regression model. Questions with a score between 2 and 4 are termed as IFQ and need to modify whereas those with having score above 4 are accepted and termed as WFQ.

AQQE's Question Quality Enhancement (QQE) module transforms those rejected questions to accepted one with the help of a type of infirmities which are
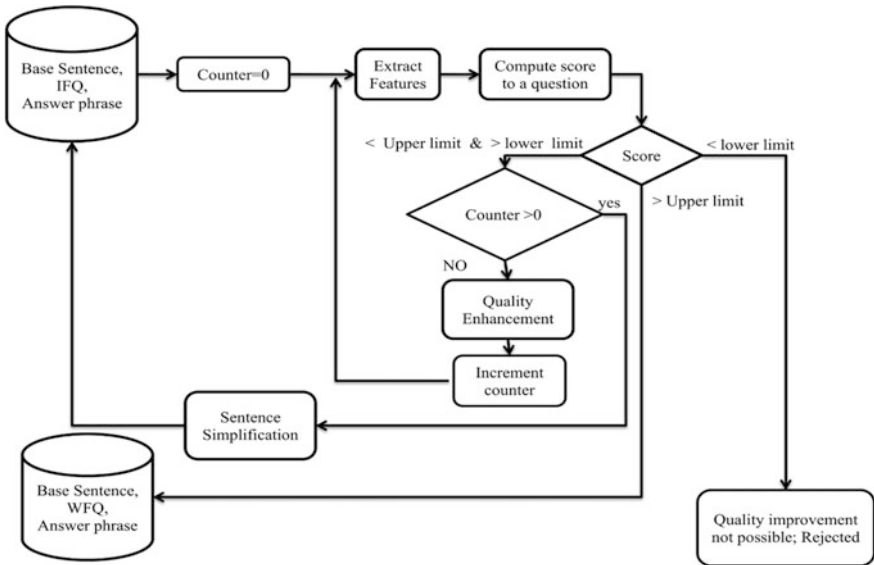
Fig. 1 Schematic of AQQE system

recorded in the form of feature value. For example, consider IFQ: What is the government formed by?

wh-phrase_as_per_ans feature tells about the wrong selection of *wh*-phrase in the question. Score module assigns value 0 to wh-phrase_as_per_ans and QQE module replaces the wh-phrase to who.

Various tasks performed by QQE module are as follows:

1. To check the voice of sentence, and convert passive voice to active voice
2. To remove the consecutive PP phrases present in a question
3. To suggest the complete answer phrase, if it is partially selected then
4. To check the grammar, and
5. To simplify the sentence

Figure 2 shows the working of QQE module.

Enhancement of IFQ to WFQ is a challenge. QQE uses Stanford parser to parse IFQ [24], Part Of Speech (POS) tagging, Tregex which helps to identify and retrieve the nodes, and Tsurgeon tool to remove or modify the IFQ nodes [25]. These NLP tools help in modification and restructuring of IFQ to WFQ. Table 3 shows the type of flaws identified in IFQ and modification suggested for transforming IFQ to WFQ.

To measure the quality enhancement in a question, AQQE recomputes the score of a modified question. QQE termed the question as WFQ which has a score above 4, otherwise, it termed the question as IFQ and performs the base sentence simplification. These newly simplified sentences, IFQ, and answer phrase are appended to an existing data set. Table 4 shows some questions enhanced by QQE module with the new score values.
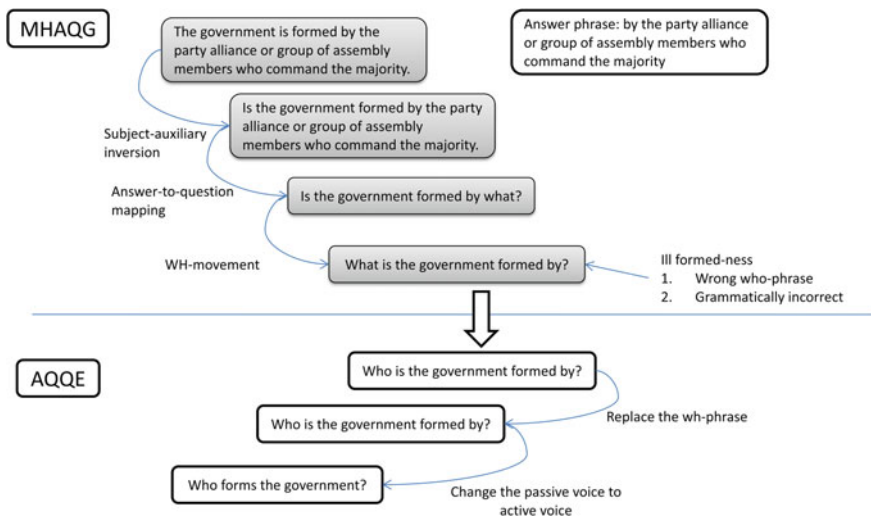
Fig. 2 Illustration of AQQE's QQE module with example

Table 3 AQQE suggested IFQ to WFQ transformation of a question

BS	AP	Parse tree of IFQ	Modification in the parse tree of IFQ
Wrong wh-phrase: The car was named after Edsel Ford, the son of the company's founder	After Edsel Ford		

(continued)

Table 3 (continued)
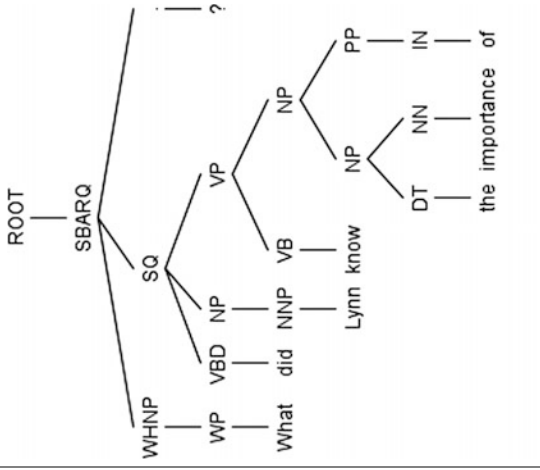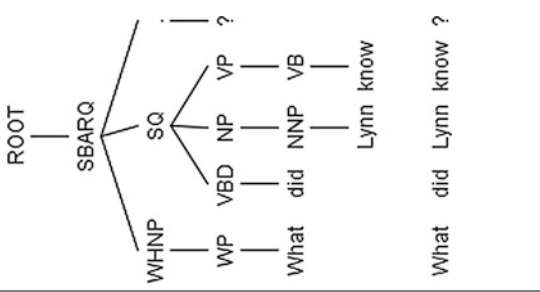
BS	AP	Parse tree of IFQ	Modification in the parse tree of IFQ
Negation: Neither Italy nor France got to the quarterfinals last year	Neither Italy nor France		

(continued)

Table 3 (continued)

BS	AP	Parse tree of IFQ	Modification in the parse tree of IFQ
Answer length: Sneezing exhausts Steve, who requires 8 tissues and 27 Gesundheits before he is done	8 tissues and 27 Gesundheits		

Table 3 (continued)

BS	AP	Parse tree of IFQ	Modification in the parse tree of IFQ
PP inside NP phrase: Having been a gymnast, Lynn knew the importance of exercise	Of exercise	ROOT — SBARQ — SQ — WHNP (WP — What), VBD — did, NP (NNP — Lynn know), VP (VB — know, NP (NP (DT — the, NN — importance), PP (IN — of))) — ? ... What did Lynn know the importance of ?	ROOT — SBARQ — SQ — WHNP (WP — What), VBD — did, NP (NNP — Lynn know), VP (VB — know) — ? ... What did Lynn know ?

Table 4 Question's score before and after the enhancement by QQE

BS	AP	IFQ	Score	WFQ	Improved score
As is the case in a parliamentary system, the government is formed by the party alliance or group of assembly members who command the majority	By the party alliance or group of assembly members who command the majority	What is the government formed by?	3.091	Who forms the government?	4.8736
China does not report its real level of military spending	Of military spending	What does China not report its real level of?	2.712	Which spending does China not report?	4.3889
Tilak was born in a Chitpavan Brahmin family in Ratnagiri headquarters of the eponymous district of present-day Maharashtra (then British India) on July 23, 1856	In a Chitpavan Brahmin family in Ratnagiri, Maharashtra	What was Tilak born in on July 23, 1856?	3.304	Where was Tilak born in July 1856?	4.7318
During Jotiraos father's time, the power and glory of the Peshwas had ebbed considerably	During Jotiraos father's	Who had the power and glory of the Peshwas ebbed considerably during?	3.41	When had the power and glory of the Peshwas ebbed considerably during?	4.8376
Tilak obtained his Bachelor of Arts in first class in Mathematics from Deccan College of Pune in 1877	Of Arts	What did Tilak obtain his Bachelor of in first class in Mathematics from Deccan College of Pune in 1877?	4.458	What did Tilak obtain from Deccan College of Pune in 1877?	4.8394

3 Result and Discussion

Human annotator had classified HSAQG-generated 174 questions as WFQ or IFQ. Score module computes the score for the same questions. To predict the score value of a question, multiple linear regression model with 11 feature is used [22]. Human annotator and AQQE independently had classified the HSAQG-generated questions into IFQ and WFQ on the basis of score value computed by AQQE's score module is shown in Table 5.

Performance of AQQE is measured using precision, recall.

Precision is the fraction of relevant instances among the retrieved instances, while *Recall* is the fraction of relevant instances that have been retrieved over total relevant instances.

Let us consider,

WFQ_h: human classified questions as WFQ

WFQ_{AQQE}: AQQE classified the question as WFQ

Precision and recall for AQQE is computed as follows:

$$Precision = \frac{WFQh \cap WFQAQQE}{WFQAQQE} = 78/89 = 87.64\%,$$

$$Recall = \frac{WFQh \cap WFQAQQE}{(WFQh \cap WFQAQQE) + (WFQh \cap \sim WFQAQQE)} = 78/(78+6) = 92.85\%.$$

87% precision shows that AQQE-computed score is satisfying the human annotators. The questions those are termed as IFQ are enhanced by QQE module, whereas WFQ are remaining unchanged. Once again the performance of AQQE's QQE module is computed by recomputing the score of all questions. The human annotator and AQQE's judgment for the enhanced questions are shown in Table 6.

$$Precision = 151/172 = 87.79\%,$$

$$Recall = 151/(1+151) = 99.34\%.$$

The precision of AQQE before and after the enhancement of questions is nearly same which indicate the exactness of AQQE to classify the questions as IFQ or WFQ is the same. The new recall value is impressive which indicates that 99% questions which are retrieved are WFQ. There are 11 questions which are still

Table 5 Evaluation of questions by human and AQQE

	Human	Human ∩ AQQE	AQQE
IFQ (before QQE)	90	79	85
WFQ (before QQE)	84	78	89

Table 6 Evaluation of the questions after enhancement

	Human	Human ∩ AQQE	AQQE
IFQ (after QQE)	22	1	2
WFQ (after QQE)	152	151	172

ungrammatical but acceptable by the human evaluator with little modification. The cost of that grammar correction in a question by AQQE is more and therefore manual corrections are suggested.

4 Challenges in Question Enhancement

AQG-generated questions may possess multiple flaws and, therefore, correcting those questions is a challenge. During the process of question enhancement, many challenges are an encounter which is discussed in this section.

4.1 Partial Selection of Answer Phrase

HSAQG selects either a Noun Phrase (NP) or a Prepositional Phrase (PP) as an answer phrase. For example consider BS: Having been a gymnast, Lynn knew the importance of exercise; Answer phrase: of exercise; IFQ: What *did Lynn know the importance of*?

Selected AP is PP which is a part of NP, therefore the IFQ produced is not human appealing. Here, AQQE suggest the new AP as *the importance of exercise* and WFQ as: What did Lynn know?

4.2 Enclosure of Answer Phrase in a Question

Presence of AP in question makes it ill formed. For example BS: Maharashtra has a bicameral legislature, i.e. it consists of two houses—Vidhan Sabha (legislative assembly) and Vidhan Parishad (legislative council). HSAQG resolves the pronoun phrase occurrence in BS by replacing it with NP *Maharashtra*. This leads to the formation of new BS: Maharashtra has a bicameral legislature, i.e. Maharashtra consists of two houses Vidhan Sabha and Vidhan Parishad.

With this new BS, HSAQG generates two wh-questions

1. What consists of two houses Vidhan Sabha and Vidhan Parishad Maharashtra has a bicameral legislature, i.e.?

2. What does Maharashtra consist of two houses Vidhan Sabha and Vidhan Parishad has a bicameral legislature. i.e.?

Here HSAQG fails to simplify the BS, which is a compound sentence; therefore, the above questions are IFQ. Here, AQQE-computed score for above IFQ is 3.58 and QQE suggested WFQ is: Which state does consist of two houses?

4.3 Conjoined Nouns in Answer Phrase

Though HSAQG selects the correct wh-phrase and AP, the generated question may be an IFQ. For example BS: Sneezing exhausts Steve who requires 8 tissues and 27Gesundheit before he is done; AP: 8 tissues and 27 Gesundheit and IFQ: How many tissues and 27 Gesundheit does Steve require?

Here, AP is conjoined with NP and with every NP a numeral adjective is attached. While formulating the question with wh-phrase How many, AQG removes the numeral adjective associated with NP. Here HSAQG removed the numeral adjective associated with first NP, but failed to remove another one leading to the formulation of an IFQ. QQE module identifies the said flaw and successfully formulates the WFQ as How many tissues and Gesundheit does Steve require?

5 Conclusion

This study reveals that researching AQGs is worth for their enhancement. A linear regression model with 11 features used for computing a score of questions outperforms. AQQE shows the enhancement in 79 (87%) questions out of 90 IFQ. WFQ generated by the QQE module is well accepted by the human expertise.

References

1. Rus, V., Graesser, A.C.: The question generation shared task and evaluation challenge. Natl. Sci. Found. **2009**, 1–48 (2009)
2. Aist, G.: Towards automatic glossarization: automatically constructing and administering vocabulary assistance factoids and multiple-choice assessment. Int. J. Artif. Intell. Educ. **12**, 212–231 (2001)
3. Brown, J.C., Frishkoff, G.A., Eskenazi, M.: Automatic question generation for vocabulary assessment. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, Oct 2005, p. 826
4. Susanti, Y., Iida, R., Tokunaga, T.: Automatic generation of english vocabulary tests. In: Proceedings of the 7th International Conference on Computer Supported Education, CSEDU 2015, vol. 1, pp. 77–87 (2015)

5. Mostow, J., Beck, I.L., Bey, J., Cuneo, A., Sison, J., Tobin, B., Valeri, J.: Using automated questions to assess reading comprehension, vocabulary, and effects of tutorial interventions. Technol. Instr. Cogn. Learn. **2**, 103–140 (2004)
6. Huang, Y., He, L.: Automatic generation of short answer questions for reading comprehension assessment. Nat. Lang. Eng. 1–33 (2016)
7. Kunichika, H., Katayama, T., Hirashima, T., Takeuchi, A.: Automated question generation methods for intelligent english learning systems and its evaluation. In: International Conference on Computers in Education (2004)
8. Chen, W., Aist, G., Mostow, J.: Generating questions automatically from informational text. In: AIED 2009, 14th International Conference on Artificial Intelligence in Education Workshops Proceedings, no. 1, pp. 17–24 (2009)
9. Hoshino, A., Nakagawa, H.: A real-time multiple-choice question generation for language testing—a preliminary study. In: Proceedings of the second workshop on Building Educational Applications Using NLP, June 2005, pp. 17–20
10. Mitkov, R., An Ha, L., Karamanis, N.: A computer-aided environment for generating multiple-choice test items. Nat. Lang. Eng. **12**(2), 177–194 (2006)
11. Liu, M., Calvo, R.A., Rus, V.: Automatic question generation for literature review writing support. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2010)
12. Ali, H., Chali, Y., Hasan, S.A.: Automatic question generation from sentences. In: Proceedings of the TALN 2010, pp. 19–23 (2010)
13. Chali, Y., Hasan, S.: Towards automatic topical question generation. In: COLING, Dec 2012, pp. 475–492 (2012)
14. Heilman, M., Smith, N.A.: Ranking automatically generated questions as a shared task. In: The 2nd Workshop on Question Generation (2009)
15. Lindberg, D.: Automatic Question Generatin from Text for Self-Directed Learning (2013)
16. Wolfe, J.H.: Automatic question generation from text—an aid to independent study. In: ACM SIGCUE Outlook (1976)
17. Wang, B., Liu, B., Sun, C., Wang, X., Zhang, D.: Generating questions from web community contents. In: Proceedings of COLING 2012: Demonstration Papers, Dec 2012, vol. 2, pp. 467–474
18. Yao, X.: Question generation with minimal recursion semantics. Communication 1–92 (2010)
19. Zhao, S., Wang, H., Li, C., Liu, T., Guan, Y.: Automatically generating questions from queries for community-based question answering. In: Proceedings of 5th International Joint Conference on Natural Language Processing, 2011, pp. 929–937 (2008)
20. Heilman, M., Smith, N.A.: Question generation via overgenerating transformations and ranking. Framework (2009)
21. Mannem, P., Prasad, R., Joshi, A.: Question generation from paragraphs at UPenn: QGSTEC system description. In: Proceedings of the QG 2010 Third Workshop on Question Generation Pittsburg, PA (2010)
22. Divate, M., Salgaonkar, A.: Ranking model with a reduced feature set for an automated question generation system. In: International Conference on Natural Language Processing, pp. 221–230 (2015)
23. Heilman, M.: Automatic Factual Question Generation from Text. Dissertation (2011)
24. Schuster, S., Manning, C.D.: Enhanced english universal dependencies: an improved representation for natural language understanding tasks. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), pp. 2371–2378 (2016)
25. Levy, R., Andrew, G.: Tregex and Tsurgeon: tools for querying and manipulating tree data structures. In: Proceedings of the 5th International Conference on Language Resources and Evaluation. (LREC 2006), pp. 2231–2234 (2006)

Facial Representation Using Linear Barcode

Sanjoy Ghatak

Abstract This paper suggests an innovative technique for quality-type linear barcodes from the face image. This procedure calculates the distinction in gradients of image shine and then it requires finding the average of the gradients into a finite number of intervals using normalization. After this, the result of quantization is converted into the limits of decimal digits from 0 to 9 and table is translated into an ultimate linear barcode. A theoretical analysis shows that the upper part of the physiognomy is not affected by a remark (like a change in physiognomy utterance, change of face range size, change in eye range (gaping and occlude eyes, after mirror rotation of input image, changes with human age)) ensuring the stability of representing its features. So, this method generates the standard-type linear EAN-8 barcode of the top part of face image (upper 70% and 75% of face image). However, in this work, a technique which is suggested is established based on the idea that the distances between attributes of a human face by comparing the feature edges of human faces that have been done by determining its image gradient using window technique. The determined gradient data is further reduced into a smaller data set for the representation of a histogram of each face. From this histogram, the barcode is generated for each face image. This method is tested by using "Face94", "YaleB Face Database", "Face database FERET", FG-NET dataset. A generated barcode holds knowledge about human's face and can be used for recording, finding, recognition, and search for peoples.

Keywords Facial image · Barcode · Image gradient · Real-time system

S. Ghatak (✉)
Department of CSE, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India
e-mail: sanjoy1cs@yahoo.co.in

1   Introduction

Barcoding technology entered into our everyday life widely and irrevocably for identification of various goods, finance documents, payments, advertising materials, and services. The approach of quality barcodes for people finding was submitted in the patent [1] in 1999. In that approach, the conclusion was that the finding of human is performed at the time of her/his e-payment, that is, in actual time, and the eccentric barcode printed on her/his hand is read using a remarkable gadget. Still, this procedure for human identification with barcode has limited use. Barcodes stamp on human body does not include any biometrical features of person [2]. Yet, the expectation is that if a barcode carries few biometric attributes, it will be taken up not only as a "stylish token", but also as a feasible plan of human recognition. Nevertheless, the creation of barcode in actual time and straight from a human's countenance or speak is an ideal solution. For this purpose, identification of human can be fulfilled at a gap, subtle for the human, in absent of back-breaking any action from the human or any barcode produced on the human's body. Among the two methods, based on voices or faces of people [3, 4], the new method will be developed based on new solution presented in [3]. New methods are, "A simple method for generating facial barcodes" [5] and "Barcode generation for face images" [6]. Both methods are proposed by Yuri Matvee, Georgy Kukharev, and Nadezhda Shchegoleva. By these methods, they generated the barcode in actual time and straight from human's face. However, the biggest problem of facial biometrics is the changeability of actual world face representations of today's practice, mainly lighting and posture variations, affecting the face expression and the brightness of the portraits. Clarification of this issue would make basic human finding and redesign the accomplishment and unwavering quality of comparing acknowledgment frameworks with the goal that standardized identification per users and decoders can satisfactorily utilize it in various functional applications for a drawn-out stretch of time. Consequently, there has been a colossal enthusiasm for the issue of solid facial scanner tag age, beginning from the season of the presence of first PC frameworks perceiving individuals by their appearances. In this work, a proposition is proposed for presenting facial representations as direct standardized tags.

Association of this paper is as per the following: in Sect. 2, a writing survey on standardized tag age has been shown and unique commitments are enrolled. In Sect. 3, the proposed scanner tag age technique is portrayed in points of interest. Observational outcomes and discourses are introduced in Sect. 4, and finally, the paper is finished up in Sect. 5.

2 Related Work

After the distribution of the patent [1], creators of [7] discovered that facial personality is essentially transmitted by even picture design, for example, eyes, eyebrows, and lips lines are put away as illustrative of this data as blend of two fold strips or a purported "organic scanner tag of human face". An example of this face area presentation in the form of "biological barcode" explained in [7] is shown in Fig. 1. A case of this face territory introduction as "natural standardized identification" clarified in [7] appears in Fig. 1.

As mentioned in [8] and other works by these authors, no formal algorithm for generating predetermined barcodes is given, but it is worth to take note of that the portrayal of a facial picture with just wide, straight dark and white stripes will never be an exact model of a human's face. In these types of model, while representing a precise "biological tag" of a face, it is a harsh estimate of the particular individual's face, which is shown in Fig. 1a and c. Behind this guess is the morphology of a man's face. Notwithstanding, more exact portrayal requires the more slender and more intricate arrangement of lines. Barcodes of the corresponding human face could look like introduce in Fig. 1d as was proposed in [9].

The fundamental idea of "biological code" is that it utilizes the calculation for separating two facial pictures presented in [10]. The thought behind this calculation is to decide the splendor angles of two commonly found strips which synchronously move along a face picture through and through, and after that assess the distinction over the present and mean estimations of the slopes and these distinctions are then encoded. On the off chance that that is the circumstance, the estimations of the fluctuation measure up to and underneath zero are coded as "0" or more zero as "1" and a two fold code is produced speaking to a face picture that is near a "biological code" [7].

However, the drawback of this approach [10] is that it is unable to produce a similar double code for confront pictures of a similar individual during the variation
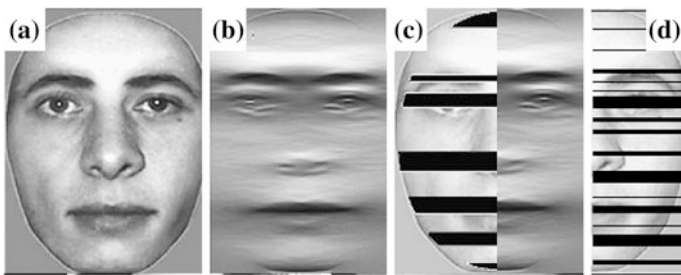
Fig. 1 Representing face region in the form of "biological barcode" [5]. **a** The actual face image; **b** level data contained in the face picture; **c** the "organic standardized tag" of a face picture; **d** the subsequent direct standardized identification [9]

of the face images. All these changes will not be with significant (but able to seen) variations in illuminations, scale and stance of facial pictures, outward appearance, and so forth. For these disadvantages, it is difficult to utilize this strategy when the variations are made in these image parameters and thus it is not material in a framework which requires high precision of change of face pictures into suitable codes.

For rectification of this entire problem [10], a new approach [5] is proposed. In this approach, the summed up structure of the framework for confront scanner tag age grasps of four fundamental squares are: (1) image preprocessing; (2) feature removing; (3) feature coding; (4) barcode creator. The technique that is proposed for creating standard sort direct scanner tags from the facial picture is expanded for the utilization of the distinction in inclinations of picture brightnesses. It includes averaging the gradients into a limited number of intervals, quantization of the outcomes in the scope of decimal numbers which ranges from 0 to 9; lastly, the standardized tag is produced from this table. However, this technique is built on producing a standard standardized identification specifically from the face picture, so it carries the individual data about a human's face. Utilizing this technique [5], for examining face pictures, the sliding windows are utilized and it starts at the "hair/brow" limit and completes at the lower limit of the nose. The weakness of this approach [5] is its incompetence to produce the barcode for the face image of the same human face during the expressions of lip lines; mouth, etc., are slightly different. Hence, in this paper, we set forward an approach for speaking to human face as EAN-8 standardized identification (straight scanner tag), in view of the approach which is examined in the paper [5, 6] with little change of a few parameters. In papers [5, 6], feature extraction is completed using only the three-fourths of the upper piece of a face picture, barring the part lower the center of the nose/mouth. If this parameter is changed, then, in the most of the cases, the better result is produced. For this reason, in our proposed method, feature extraction is completed using (i) 70% (considering bring down fringe of the nose territory) and (ii) 75% (considering between the nose and lips) of the upper piece of the face picture, barring the part bring down the center of the nose/mouth and the strength of standardized tag for both the cases are checked. The main features of the proposed approach are as follows:

(a) The features are input image analysis based on parameters estimation in terms of pose, expression, and size.
(b) Generate barcodes which are illumination invariant and checking the stability of barcode.
(c) To check the soundness of the standardized tag from the different facial pictures with various ages are likewise considered as another component.

3 Proposed Method

Following are the four steps to representing the proposed algorithm which is shown in Fig. 2 with block diagram. (a) Image preprocessing (b) Feature Extraction (c) Feature coding and (d) Barcode creator

3.1 Image Preprocessing

Input image preprocessing performs the following basic tasks. The input image is analyzed based on parameters estimation—pose, expression, and size. In bimodal Access control (AC) systems, frequently input data include voice and image. Then during conversion or uttering some control word, sentence; face expression is changed. Figure 3 shows an example of images in the mentioned situations.
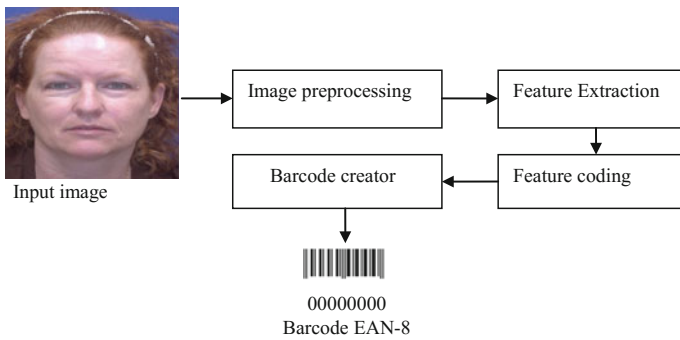
Fig. 2 Steps of barcode generation (with block diagram)
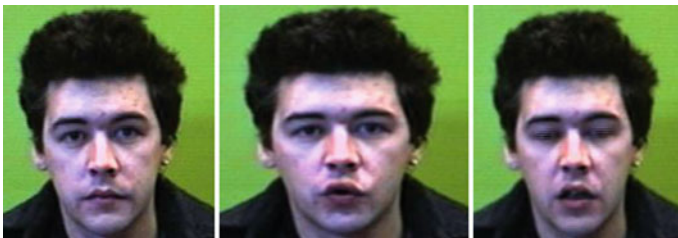
Fig. 3 Possible change of face image

3.2 Feature Extraction

Feature separation is implemented using the variance in gradients of image brightness. For this situation, introductory extractions of highlights from bonafide picture is expanded on ascertaining the change of splendor angles in two specularly found windows with the stature "h ≥ 1" pixels and width equivalent to the width of the genuine picture. With a step "S ≥ 1", the window synchronously moves (slide) along the facial image. They slide just in the vertical direction from upper to lower. At every step (in algorithm symbolized with i); the distance d(t) (gradient(i)) is calculated across the sub pictures in the sliding windows. These subspaces are the required changes in inclinations. Slipping starts from the "forehead/hair" boundary and ends at the bottom part of the nose. Along these lines, the distinction in inclinations changed into remove, features the drops in splendor on the fringe of the hair/brow, the line of eyebrows, the line of eyes and the lines of nose/lips, that is, on the lines of the "biological code" of the face picture. The integral characteristics of the least changeable part of the face are represented by the calculated values of distance if there should be an occurrence of perceptible changes of facial picture parameters. The arrangement of the strategy of angle contrast figuring and encoding utilizing a sliding window is shown in Fig. 4. In this Figure, (a) demonstrates the face picture with introductory position of the two rectangular windows (h-window tallness); (b) uncover a similar picture with the last position of the two rectangular windows; (c) the photo constituting the separation esteems is shown; d(t) over the tantamount territories of pictures "submerged by the windows" in light of the aggregate number of steps t = i=1, 2, …, T.

In this method, the actual face images have the size of "r × c" pixels and these face images are transformed into an EAN-8 barcode. The two windows "u" (upper window) and "d" (bring down window) comprising of "h" rows each, are orchestrated specularly separate to the present esteems I on the hub Y. There are T = j * m sliding advances, where 'j' is the code length and 'm' is the smoothing module. To make EAN-8 scanner tag from confront picture, the parameter j = 7. For the most part, i ≥ 8 is chosen from the state of limit end.
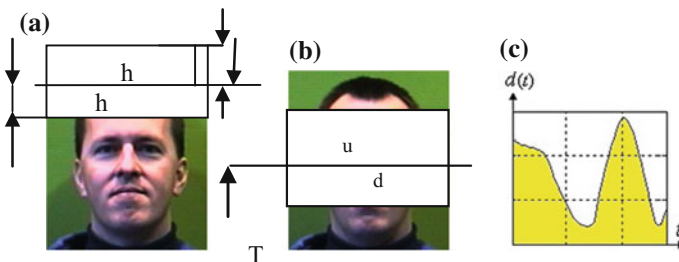
Fig. 4 Clarification of the possibility of the technique for gradients

$$T = j * m \leq (r - h). \tag{1}$$

Give us a chance to think about the estimation of 'T' around on the lower outskirt of the nose territory (if there should be an occurrence of 70% upper segment of the face) or across the nose and lips (if there should arise an occurrence of 75% upper segment of the face), which would bar from thought the base piece of the face and in this manner defeat the effect of feeling on the robustness of standardized identification creation. Furthermore, if the value of

$$T > (r - h), \tag{2}$$

then the input image size should be increased in order to fulfill the condition (1). Expanding the picture estimate is accomplished in stages 1, at the phase of picture measure redress, as clarified previously. The whole gang up and coming (current for i = 1, 2, 3, …, T) position of the hub between the windows u(i), d(i) is chosen from the condition.

$$h + s\ (i - 1) \leq i \leq (r - h) \tag{3}$$

where 'S' is sliding step.

The distance d(t) between the windows is defined by initial gradient,

$$d(t) = gradient(i) = \|u(i) - d(i)\| \text{ for all } i = 1, 2, 3, \ldots, T \tag{4}$$

Yield which is gotten in the wake of utilizing the condition (4) appears in Fig. 4 'c'. This yield is the contribution from stages 2 to stages 3, where the normalization and coding of estimations of a separation vector are taken out.

3.3 Feature Coding

In these approaches, step 3 remarks the encoding task, which is performed by the necessary number of decimal digits. For features coding, the following operations are performed. Elements of the gradient (vectors) are normalized to a maximum value of gradient (max_gradient).

$$gradient(i) = gradient(i) / max\_gradient(i),\ t = 1 \text{ to } T \tag{5}$$

The outcome from using Eq. (5) is averaged within the interim 'm' and using the scaling factor scale. This out is quantized in the range of decimal digits from 0 to 9.

$$gradient(i) = floor(gradient(i) * 10) \tag{6}$$

where i = 1 to T

$$num = num + gradient((i - 1) * m + j) \tag{7}$$

where i = 1, 2, …, 7 and j = 1 to m; initialize the num with value 0 and m is calculated with the Eq. 8.

$$m = (T/7) \tag{8}$$

$$barcode(i) = round((scale * num)/m) \tag{9}$$

where scale is the scale factor, which lies between 9 and 10 and it is not an integer. In this experiment, the scale factor is considered as 9.5 and also the output is observed after taking the scale factor between the given scales. After observation, it is clear that in this scale, the barcode of given face image is stable.

3.4   Barcode Creator

The output of steps 3 goes to steps 4, where the last 8 digits of EAN-8 barcode is produced. Figure 5 shows the final output of input image which is represented as an EAN-8 linear barcode. Total numbers of digits are 8 in this barcode. From the output of 8 digits, the eighth digit is considered as checksum digit for the first 7 digits. To generate final barcode, standard algorithm is used and this standard algorithm is described in the paper [11]. The formula for checksum calculation is

$$Checksum = [10 - [(3 * sum\_off\_odd\_position's\ digits + sum\_of\_even\_positions\ digits) \\ mod10]]\ mod10.$$

In Fig. 5 checksum digit is 2. The purpose of checksum digit is error correction and detection.

11001102
Barcode EAN-8
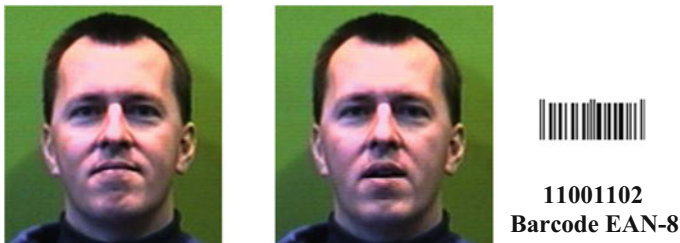
Fig. 5   Linear barcode (EAN-8)

Table 1 Description of database

Database name	Total no. of acquired images	File format	Variations (Expression, pose, age, brightness, illumination, etc.)	Stable barcode generated (Yes/No)
Face94	100 classes with 11 images	JPEG	Variations in expression and pose available	Yes
FERET database	14051 images with 10465 different subject	JPEG	Variations in expression pose, lighting condition, and illumination available	Yes but in case of illumination invariants and major directional change, it is difficult to generate stable barcode
YaleB face database	585 images with 10 different subjects	JPEG, BMP	9 pose X 64 illumination conditions. For every last one subject in a specific representation, a picture with surrounding (background) brightening additionally caught	For illumination invariant face, it is difficult to getting stable barcodes
FG-NET Aging database	FG-NET database contains 1002 face images of 82 subjects	JPEG	Different ages available (age 0(<12) to 69)	Yes possible but in case of major reflection, barcode varies
Database of composite faces at different ages	20 composite faces	JPEG, PNG	Representing changes in human age	Yes

4 Experimental Result and Discussions

A few analyses is performed to assess the soundness of the produced standardized tag of the face picture in different face acknowledgment situations, including outward appearance change, change in the splendor of test picture, reflect impression of the first picture, maturing of face and if there should be an occurrence of outward appearance varieties and the nearness of shadows on confront pictures from nearby lighting. In this technique, the test is done on a few databases: "Face94" [12], "YaleB Face Database" [13], "Face database FERET" [14], "FG-NET aging Database" [15], and "a database of composite faces at different ages". Table 1 enlists a complete database used in the present work and status of the result. Barcode is created in light of 70% and 75% of the upper piece of the face picture barring the part bring down the center of the nose/mouth for check the strength of barcode. Tables 2 and 3 demonstrate the exactness rate of creating scanner tags for facial pictures with various outward appearances on the changed database and checking the rate of the vigor of the produced standardized tags to maturing of

Table 2 Precision rate of producing standardized tags for facial pictures with various outward appearances on changed database

Database name	Result	
	For 75%	For 70%
Face94 database	78	80
Feret database	65	70
YaleB database	10	15

Table 3 Checking rate of vigor of the created scanner tags to maturing of appearances

Database name	Result	
	For 75%	For 70%
FG-NET aging dataset	75	80
Composite face dataset	80	83

countenances separately. From this result, it is clear that if the facial expression is changed then barcodes of the same face images are stable and if the age is also changed then there is no change in generated barcode of the same face. But the barcode is unstable in case of illumination invariant face image.

5 Conclusions

The paper presents a novel algorithm for face recognition using linear EAN-8 barcodes. This proposed procedure of creating linear EAN-8 barcodes from face images is built using the window technique, gradients calculation, gradient directions, normalization, and quantization. It requires computing the picture angle on limited quantities of interval, quantization of the yield in the size of decimal numbers from 0 to 9, and believers this table into the last standardized identification. Finally, input faces are represented with the help of EAN-8 linear barcodes. There is no requirement of specialized software for image processing (MATLAB 15 is used) and the proposed technique is computationally low cost. The test has been performed on the face94 database, Yale B dataset, FERET face dataset, 20 composite face dataset, and FG-NET aging dataset. The test comes about have demonstrated that the calculation which is depicted guarantee the quality of the created standardized tags on account of slight reflecting of the first picture, when scale, stance, and outward appearance changed. Moreover, the strategy depends on confront acknowledgment utilizing scanner tag specifically from the individual's face, in this way contains the inside data about the individual human's face. In this technique, the fundamental issue is creating standardized tags from enlightenment invariant face pictures. Next analysis will attempt to take care of this issue. The created barcode can be utilized for recording, distinguishing proof, acknowledgment, and look for the person.

References

1. Method for verifying human identity during electronic sale transactions: Patents US5878155A. http://www.google.com/patents/US78155A (1999). Accessed 15 Mar 2014
2. Barcode/Tattoos: http://www.barcodeart.com/store/wearable/tattoos/ Accessed 15 Mar 2014
3. Forczmanski, P., Kukharev, G., Shchegoleva, N.: An algorithm of face recognition under difficulty lighting conditions. Electr. Rev. (10 b), 201–204 (2012)
4. Matveev, Yu.N.: Technologies of Biometric identification of a person by voice and other Modalities. Vestnik MGTU. Priborostroenie, Special Issue Biometric Technologies, pp. 46–61 (2012) (in Russian)
5. Matveev, Y., Kukharev, G., Shchegoleva, N.: A simple method for generating facial Barcodes. In: WSCG2014 Conference on Computer Graphics, Visualization and Computer Vision in Co-operation with EUROGRAPHICS Association Exchange Anisotropy, pp. 213–220. Academic, Czech Republic (2014)
6. Kukharev, G., Matveev, Y., Shchegoleva, N.: Barcode generation for face images. Data Anal. Intellect. Syst. Bus. Inf. **3**(29), 201
7. Dakin, S.C., Watt, R.J.: "Biological barcodes" in human faces. J. Vis. **9**(4), 1–10 (2009)
8. UPC/EAN Barcode FAQ & Tutorial: http://www.idautomation.com/barcode-faq/upcean
9. Facial barcodes help us identify people: http://www.barcodesinc.com/news/?P=92. Accessed 15 Mar 2014
10. Hitrov, M.V. (ed.): Methods of facial images processing and recognition in biometrics. Politechnika, Saint Petersburg (2013) (in Russian)
11. Barcode generation algorithm for EAN-8 and EAN-13: http://www.cherry-notes.spb.ru/barcode_ean8.html
12. Face94 database: http://cswww.essex.ac.uk/mvallfaces/face94.html
13. http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html
14. CHUK face Sketch FERET database: http://mmlab.ie.cuhk.edu.hk/cufsf
15. FG-NET Aging database: http://www.fgnet.rsunit.com (2010)
16. Burt, M.D., Perrett, D.I.: Perception of age in adult caucasian male faces: computer graphic manipulation of shape and colour. Proc. R. Soc. Lond. **259**(1355), 13–143 (1995)
17. Barcode generation algorithm for EAN-8and EAN13: http://www.cherry-notes.spb.ru/barcode_ean8.htm. Accessed 15 Mar 2014

Bit-Reversal Encryption Towards Secured Storage of Digital Image in Cloud Deployment

Soumitra Sasmal and Indrajit Pan

Abstract Privacy preservation of cloud data is gaining importance since majority of digital records are transferred to cloud-based storage. Familiarization of cloud storage is rapidly growing and people are choosing cloud storage services as suitable alternatives for personal setup. Potential increase in the cloud usage creates a concern for data security. Cloud records demand robust encryption standards. Protection is needed both from external hackers and internal intruders. This article reports a bit-reversal encryption mechanism for privacy preservation of digital image data deployed on cloud. This encryption mechanism deals with different types of digital images toward secure storage in cloud infrastructure. Experimental results show that the method is quite simple and easy to implement to preserve the privacy and security of user data either off premises or on premises cloud storage.

Keywords Bit-reversal encryption · Cloud auditing · Cloud computing Digital image · Privacy preservation

1 Introduction

Cloud computing is a hostile framework that provides mobile and scalable computing resources and services at a large-scale. It relinquishes required resources and services to its clients within a nominal time through minimum managerial configuration [1]. This can be envisioned as on-demand software and hardware computational resources accessed through an Internet-based computer network infrastructure. There are three popular service models of cloud computing in the

S. Sasmal (✉)
Techno India, Salt Lake, Kolkata, India
e-mail: soumitrasasmal@gmail.com

I. Pan
RCC Institute of Information Technology, Kolkata, India
e-mail: indrajit.pan@rcciit.org

forms of Software as Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). Among these three services, cloud storage (also known as Data as a Service (DaaS)) is widely used by the community [1].

Cloud deployment frameworks are basically of two types: (a) public cloud and (b) private cloud. Private cloud is mainly set up within enterprises, whereas public cloud is broadly used by the common people, small and medium enterprises as a storage framework. Public cloud implements a networked enterprise storage architecture, where virtual pools of storage are contributed by different hosts. These hosts are basically called third-party cloud vendors like IBM, Amazon, etc. Private cloud storage users avail this service quite often to store their data, which are available and accessible on demand. A survey predicts that total data volume in cloud will touch 40 trillion gigabytes by 2020 [1].

The basic problem with this access is the security and privacy of data. Mostly, these records are stored in a plain format. Hence with the increasing use of data storage service, user gradually compromises security and privacy of data. Often these data owners are not fully guaranteed that the data will remain unaltered and only will be shared among an intended group of audience. These data need to be protected against unauthorized users and cloud system administrators.

Present work is going to address privacy and authenticity issues of digital image in cloud environment. Nowadays, people mostly use multimedia records, which are attractive and voluminous. One of the common forms of these multimedia records is image. Image data possesses a major share of records in cloud. However, in literature, very few works are reported on privacy preservation of digital image data in cloud. This work will illustrate a bit-reversal mechanism for encrypting digital images in cloud.

An image is normally received in three formats: RGB, gray scale, and black and white (b/w). Proposed method generates an indexed color map of the concerned image. Indexed image has a mapped representation of pixel values to color map values. Index matrix of the indexed image is then encrypted using the system key. Index image matrix is divided into some segments or blocks. Each of these blocks are applied with bit-reversal permutation method to generate cipher block. However, the color map matrix is left unchanged. The encrypted indexed image is again re-permuted to generate the original form.

Next section of this article presents some recent researches on privacy preservation of cloud data, which is followed by a basic concept of bit-reversal permutation. The proposed method is discussed in Sect. 2 followed by experimental results in Sect. 3 and conclusion is drawn in Sect. 4.

1.1 Recent Works

In this section, some very recent and diversified research references are discussed to present recent trends of research in this domain. Authors in [2, 3] have discussed

about different homomorphic cryptographic mechanisms. They have shown that homomorphic cryptography can be used to execute a script or a program in an encrypted space. Different types of security challenges for cloud data and some state-of-the-art solutions are discussed in [4].

The integrity of data storage and its endurance are discussed in [5]. The work proposal contains a public verifiability mechanism, which reduces user side computational overhead. Here, the work proposal comprises of different architectural constructs for cloud storage and cloud servers. Two cloud servers are deployed under this scheme, where the first one serves as a storage server and the second fulfils the requirement for auditing. Cloud audit server performs the task of preprocessing before uploading data into cloud storage server.

The ability of homomorphic encryption to assess cipher content without knowing the plaintext is discussed in [6]. This becomes very helpful in diverse situations for preserving the confidentiality of the document. However, in some situations, homomorphic encryption fails to meet the requisite standard. Thus, the article proposes some hybridization model to plug the fallacies of existing simple homomorphic model and to attain the required standard of operations.

A secondary scalar invariant feature transform mechanism to protect the privacy of image data is proposed in [7]. Authors have described that the volume of image data is rapidly growing and is being utilized for different computational purposes. Image-related computation consumes heavy computational resources and that can only be met if the resource is outsourced from cloud. A homomorphic encryption with garbled circuit is discussed in [8] for privacy preservation in face recognition process. This work is not based on cloud computing scenario but it provides a deep insight on secured encryption for images. Different image encryption techniques using chaotic scheme are described in [10, 11].

Dynamic data auditing and verification with the help of third-party auditor is discussed in [12]. Third-party auditing claims to eliminate the overhead of user to verify the integrity of data. This work attempts to attain both public verifiability and dynamic operations on data. The work also claimed to improve the proof of retrievability model.

This section establishes a requirement of privacy preservation, security auditing, and encrypted storage of data on cloud infrastructure. Recent research trend encompasses a versatile approach toward cryptographic modeling of data, however, the majority is focused on text-based records. Nowadays, data is not only confined to textual form, but also the volume of image and multimedia record are rapidly growing. Often, high-end resources are required to process these records. Cloud is a most convenient option for storing these records and cloud infrastructure is mostly suitable for processing these records. Thus, there remains a need for a simple yet powerful security mechanism to protect image data on cloud.

1.2   Bit-Reversal Permutation

Bit-reversal permutation concept belongs to applied mathematics. A sequence of k different items is permutated and rotated using this mechanism. Number of items (k) are usually represented in power of 2 (where $k = 2^n$).

If $n = 2$, then $k = 4$ and the numbers will be 00, 01, 10, and 11.

After applying bit-reversal technique, the above series will become 00, 10, 01, and 11. Here, each number will be read in reverse, which means earlier number 01 will be converted into 10. Similarly, if any number is 001 then its bit-reversal form will be 100.

Bit-reversal mechanism follows involution, which means if bit-reversal technique is reapplied on the data then it will produce back the old record.

2   Proposed Method

2.1   Design Background

Digital image can exist in different modes like RGB or true-color (though a true-color image is little different from RGB), grayscale or gray-level, black and white, etc., each of these modes contains different color pallets. RGB or true-color mode exhibits highest degree of color variation. In case of RGB or true-color images, its information storage requires three-dimensional matrix structure. This three-dimensional information is difficult to process. Operations on grayscale images are very common in image processing applications because it is less complex in computational requirements, and structural relation in color images are complex. In addition to that, grayscale image can hold almost all major properties of an image, which is not possible in black and white image [10].

Conversion of gray-level image to RGB is difficult and it often loses major properties of RGB image. Any encryption technique for digital images should maintain the quality and information of the original image after decryption. This requirement imparts a challenge for RGB mode digital image to be encrypted through gray-level operations.

In this present work, a concept of indexed color map image is applied for application of encryption mechanism. Indexed color approach helps to store digital RGB image in a restricted fashion. In this mechanism, an index is created to store pixel-wise information on index value of the color map. Here, a separate color map or palette is created to color information. Both index matrix and color map are two-dimensional. All modes of digital images can be decomposed in indexed image format, which develops an indexed matrix and a color map matrix. These two matrices together hold complete information of an image. These two matrices can be combined together to reconstruct the main image [9].
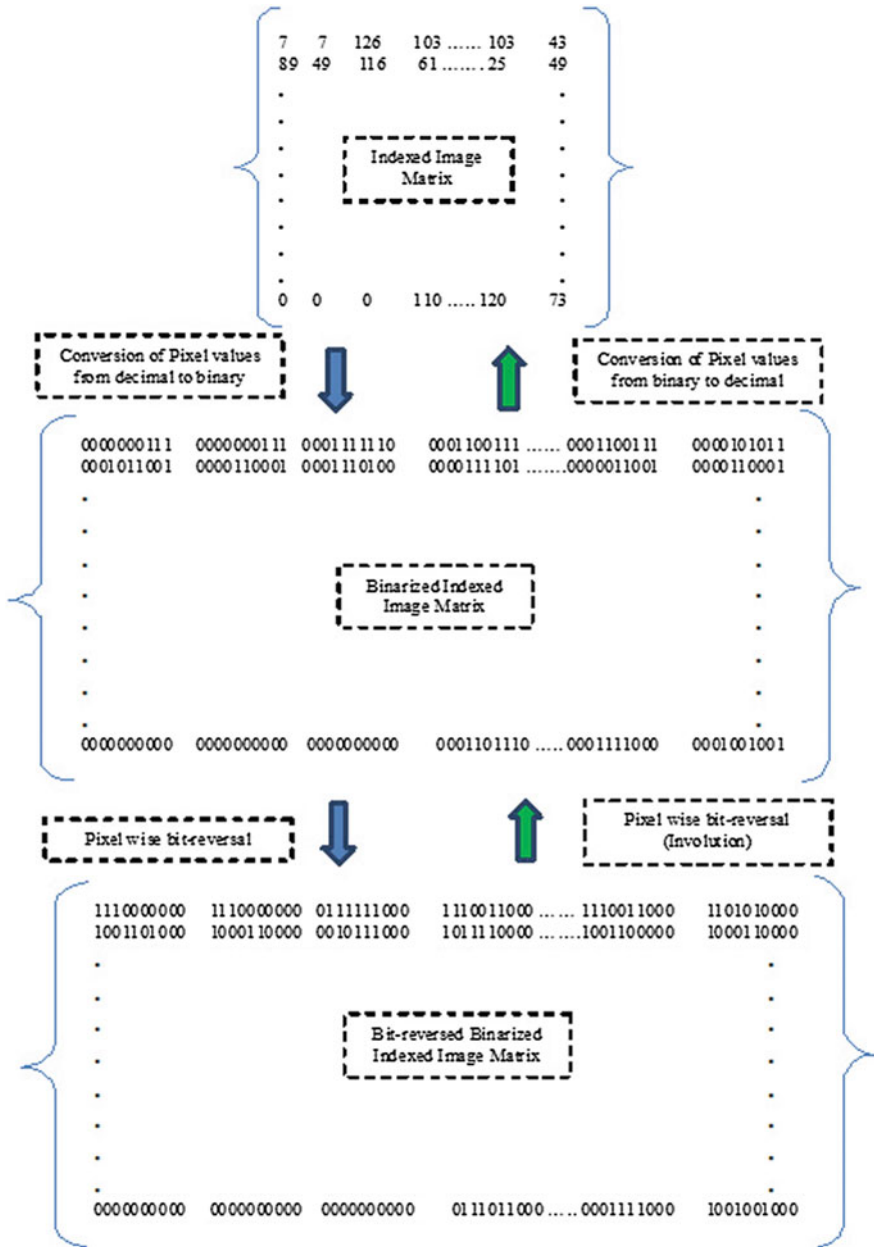
Fig. 1 Flow of encryption and decryption

Algorithm *(Bit-reversal Encryption)*

Input: *image in matrix (I)*
Output: *Encrypted index matrix (E$_{ind}$) and color-map (map)*

Method:

Read image in **I**
Convert I to Index matrix - (**ind**) and color-map (**map**)
Determine dimension (row (**r**) and column (**c**)) of **ind**
external_loop: 1 to r
        inner_loop: 1 to c
                binarize(element of **ind**)
                        //binary equivalent of decimal value
        end inner_loop
end external_loop

external_loop: 1 to r
        inner_loop: 1 to c
                bit-reversal(element of **ind**)
                        //bit-reversed equivalent of binary values
        end inner_loop
end external_loop
End Method

Algorithm *(Bit-reversal Decryption)*

Input: *encrypted index matrix (E$_{ind}$)*
Output: *Decrypted index matrix (ind) and color-map (map)*

Method:

Read index matrix in **E$_{ind}$**
Determine dimension (row (**r**) and column (**c**)) of **E$_{ind}$**
external_loop: 1 to r
        inner_loop: 1 to c
                bit-reversal (element of **E$_{ind}$**)
        end inner_loop
end external_loop

external_loop: 1 to r
        inner_loop: 1 to c
                decimalization(element of **E$_{ind}$**)
        end inner_loop
end external_loop
End Method

Fig. 2 Pseudocode of proposed method

Proposed work performs encryption and decryption operation on indexed matrix as described in Sect. 2.2.

2.2 Proposed Algorithm

Encryption

Bit-reversal encryption approach is applied on index matrix of an indexed image. A digital image is first converted into an indexed image during upload. This process segregates image information into two matrices (i.e., index matrix and color map matrix) as discussed in the previous section.

Index matrix is then processed for encryption as illustrated in Fig. 1. Index matrix contains decimal values indicative to the index location of color map matrix for each pixel of the image. Initially, the dimension (i.e., span of row and column) of index matrix is checked. Each pixel entry in decimal is then converted into its equivalent binary representation. Considering 512 different color entries for color map, binarization is represented in 10-bit representation. Conversion of decimal index values to its binary equivalent is shown in Fig. 1.

Once the binary equivalent of each pixel index is generated, then it is processed through bit-reversal modification as discussed in Sect. 1.2. An encrypted index matrix is generated after bit-reversal.

Decryption

Decryption process is also shown in Fig. 1. Bit-reversal technique supports involution as explained in Sect. 1.2. Same method is reapplied in involution to get back the earlier state. Hence, the bit-reversal technique is applied again on the encrypted index matrix to generate the main matrix. After applying bit-reversal on the encrypted matrix, pixel-wise binary values are converted into decimal equivalents. This generates the original index matrix of the corresponding image.

In Fig. 1, encryption flow is represented using blue arrow and decryption is represented using green arrow.

A formal pseudocode of this method is given in Fig. 2.

3 Experimental Results

The proposed method was implemented through GNU Octave in open GL cloud environment. In this work, the *Peppers image* and the *Cameraman image* were processed for encryption and decryption. Experimental results were tested through histogram analysis as shown in Figs. 3 and 4.

In the analysis, it is observed in the histogram analysis that the proposed cryptographic approach returns the original image without any distortion.
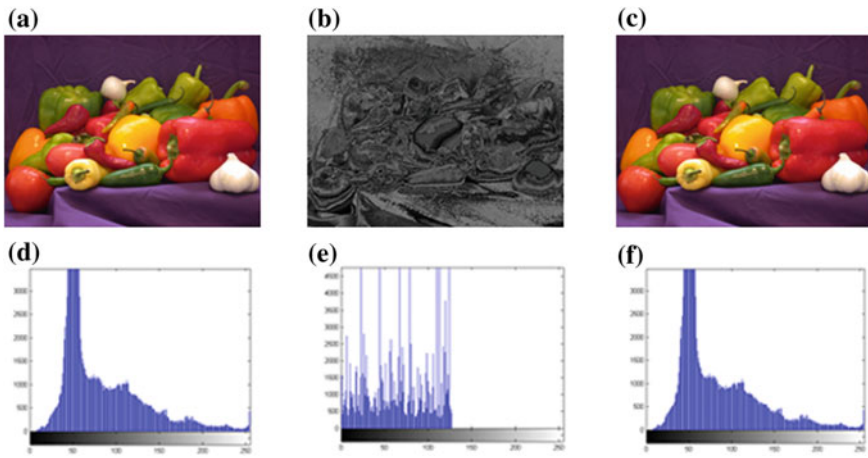
Fig. 3 **a** Original Pepper image. **b** Encrypted Pepper image. **c** Decrypted Pepper image. **d** Histogram of 3a. **e** Histogram of 3b. **f** Histogram of 3c
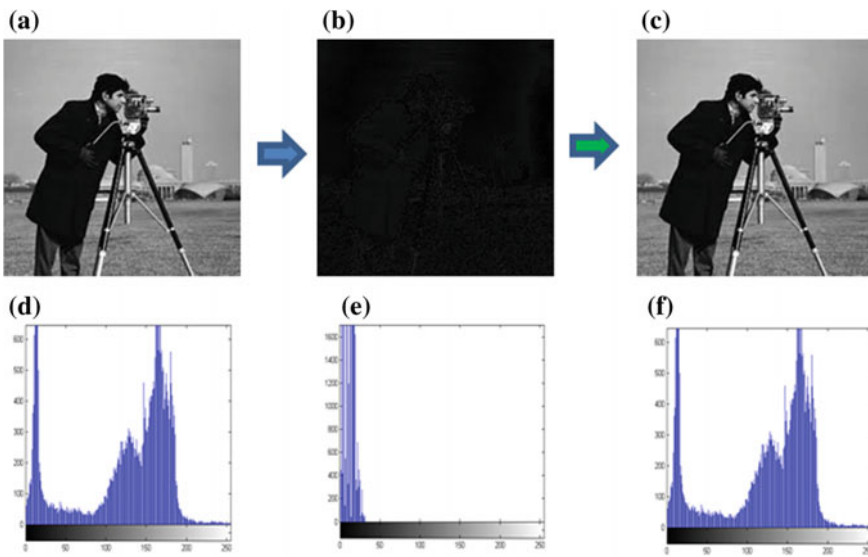
Fig. 4 **a** Original Cameraman image. **b** Encrypted Cameraman image. **c** Decrypted Cameraman image. **d** Histogram of 4a. **e** Histogram of 4b. **f** Histogram of 4c

Similarly, the histogram of the encrypted image shows no significant information within it. The results are not compared with other researches due to insufficient availability of literature.

Table 1 Performance analysis of encrypted and decrypted images

Sample image	Reference image	PSNR (dB)	MSE
Encrypted Pepper image (Fig. 3b)	Pepper image (Fig. 3a)	8.81	1.31
Decrypted Pepper image (Fig. 3c)	Pepper image (Fig. 3a)	55.72	0.006
Encrypted Cameraman image (Fig. 4b)	Cameraman image (Fig. 4a)	6.29	1.32
Decrypted Cameraman image (Fig. 4b)	Cameraman image (Fig. 4a)	46.93	0.003

3.1 Empirical Analysis

Data on cloud is vulnerable to many attacks. A common style of attack is brute-force approach. In this empirical analysis section, some standard statistical measurement is performed on both encrypted and the decrypted ones to understand the strength of the proposed encryption scheme.

Both encrypted and decrypted images are analyzed here with respect to Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) [10]. MSE determines pixel-based averaged squared difference between any two images. In a good encryption scheme, MSE of the encrypted image becomes high in comparison to original image and MSE between original and decrypted images is usually low. PSNR is represented in decibel (dB) unit. A high PSNR value means poor encryption and a low PSNR value denotes good encryption. PSNR of decrypted image with respect to original one is always high. Table 1 represents performance metric of proposed algorithms on the images of Figs. 3 and 4. The result shows efficacy and strength of the proposed algorithm.

4 Conclusion

Algorithm reported on bit-reversible approach for digital image encryption is implemented on different types of images. The method is mainly applied on index matrix of an image. It is observed that the encrypted index image does not contain any potential information, whereas, the decryption process returns the original image without any distortion.

This present method also has been tested upon its robustness against diverse statistical and differential attacks. PSNR and MSE analyses in Table 1 proves the robustness of the method.

In future, some hybrid intelligent techniques will be used to convert the present work into a key-based encryption and decryption.

References

1. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Knowinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M.: A view of cloud computing. Commun. ACM **53**(4), 50–58 (2010)
2. Brenner, M., Perl, H., Smith, M.: How practical is homomorphically encrypted program execution? An implementation and performance evaluation. In: 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, pp. 375–382 (2012)
3. Brenner, M., Perl, H., Smith, M.: Practical applications of homomorphic encryption. In: Proceedings of 7th International Conference on Security and Cryptography, pp. 5–14 (2012)
4. Datta, J., Pan, I., Bhattacharyya S.: Security challenges for cloud data: a survey. In: Proceedings of International Conference on Tele Communication, Power Analysis and Computing Techniques, Article 258 (2017)
5. Li, J., Tan, X., Chen, X., Wong, D.: An efficient proof of retrievability with public auditing in cloud computing. In: Proceedings of 5th International Conference on Intelligent Network Collaborating System, pp. 93–98 (2013)
6. Naehrig, M., Lauter, K., Vaikuntanathan, V.: Can homomorphic encryption be practical. In: Proceedings of 3rd ACM Workshop on Cloud Computing Security Workshop, pp. 113–124 (2011)
7. Qin, Z., Yan, J., Ren. K., Chen, C., Wang, C.: Towards efficient privacy preserving image feature extraction in cloud computing. In: Proceedings of ACM International Conference on Multimedia, pp. 497–506 (2014)
8. Sadeghi, A., Schneider, T., Wehrenberg, I.: Efficient privacy-preserving face recognition. In: Proceedings of 12th International Conference on Information Security and Cryptography, pp. 229–244 (2010)
9. Sanchez, J., Canton, M.P.: The PC Graphics Handbook. CRC Press (2003). ISBN 0-8493-1678-2
10. Sharma, M., Kowar, M.K.: Image encryption techniques using chaotic schemes: a review. Int. J. Eng. Sci. Technol. **2**(6), 2359–2363 (2010)
11. Srivastava, A.: A survey report on different techniques of image encryption. Int. J. Emerg. Technol. Adv. Eng. ISSN **2**(6), 2250–2459 (2012)
12. Wang, Q., Wang, C., Li, J., Ren, K., Lou, W.: Enabling public verifiability and data dynamics for storage security in cloud computing. In: Proceedings of Computer Security, pp. 355–370 (2009)

Author Index