

Daya K. Lobiyal
Durga Prasad Mohapatra
Atulya Nagar
Manmath N. Sahoo
Editors

Proceedings of the International Conference on Signal, Networks, Computing, and Systems

ICSNCS 2016, Volume 1

Lecture Notes in Electrical Engineering

Volume 395

Board of Series editors

Leopoldo Angrisani, Napoli, Italy
Marco Arteaga, Coyoacán, México
Samarjit Chakraborty, München, Germany
Jiming Chen, Hangzhou, P.R. China
Tan Kay Chen, Singapore, Singapore
Rüdiger Dillmann, Karlsruhe, Germany
Haibin Duan, Beijing, China
Gianluigi Ferrari, Parma, Italy
Manuel Ferre, Madrid, Spain
Sandra Hirche, München, Germany
Faryar Jabbari, Irvine, USA
Janusz Kacprzyk, Warsaw, Poland
Alaa Khamis, New Cairo City, Egypt
Torsten Kroeger, Stanford, USA
Tan Cher Ming, Singapore, Singapore
Wolfgang Minker, Ulm, Germany
Pradeep Misra, Dayton, USA
Sebastian Möller, Berlin, Germany
Subhas Mukhopadhyay, Palmerston, New Zealand
Cun-Zheng Ning, Tempe, USA
Toyoaki Nishida, Sakyo-ku, Japan
Bijaya Ketan Panigrahi, New Delhi, India
Federica Pascucci, Roma, Italy
Tariq Samad, Minneapolis, USA
Gan Woon Seng, Nanyang Avenue, Singapore
Germano Veiga, Porto, Portugal
Haitao Wu, Beijing, China
Junjie James Zhang, Charlotte, USA

About this Series

“Lecture Notes in Electrical Engineering (LNEE)” is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer’s high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer’s other Lecture Notes series, LNEE will be distributed through Springer’s print and electronic publishing channels.

More information about this series at <http://www.springer.com/series/7818>

Daya K. Lobiyal · Durga Prasad Mohapatra
Atulya Nagar · Manmath N. Sahoo
Editors

Proceedings of the International Conference on Signal, Networks, Computing, and Systems

ICSNCS 2016, Volume 1

 Springer

Editors

Daya K. Lobiyal
School of Computer and Systems Sciences
Jawaharlal Nehru University
New Delhi, Delhi
India

Atulya Nagar
Faculty of Science
Liverpool Hope University
Liverpool
UK

Durga Prasad Mohapatra
Department of Computer Science
and Engineering
National Institute of Technology
Rourkela, Odisha
India

Manmath N. Sahoo
Department of Computer Science
and Engineering
National Institute of Technology
Rourkela, Odisha
India

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-81-322-3590-3

ISBN 978-81-322-3592-7 (eBook)

DOI 10.1007/978-81-322-3592-7

Library of Congress Control Number: 2016942038

© Springer India 2017

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

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

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

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer (India) Pvt. Ltd.

Preface

International Conference on Signal, Networks, Computing, and Systems (ICSNCS 2016), organized by School of Computer and Systems Sciences, Jawaharlal Nehru University, India, during February 25–27, 2016, certainly marks a success toward bringing researchers, academicians, and practitioners in the same platform. It is indeed a pleasure to receive overwhelming response from researchers of premier institutes of the country and abroad for participating in ICSNCS 2016, which makes our endeavor successful. Being the first conference of its series, it was challenging for us to broadcast the conference among researchers and scientists and to receive their valuable works for review. A very systematic workflow by the committee has made it possible. We have received 296 articles and have selected 73 articles of the highest quality among them for presentation and publication through peer review done by at least two experts for each article. We are unable to accommodate many promising works as we restricted our selection to limited articles which can be elaborately presented in a three-day conference. We are thankful to have the advice of dedicated academicians and experts from industry to organize the conference. We thank all researchers participating and submitting their valued works in our conference. The articles presented in the proceedings discuss the cutting-edge technologies and recent advances in the domain of the conference. We conclude with our heartiest thanks to everyone associated with the conference and seek their support to organize the next editions of the conference in subsequent years.

New Delhi, India
Rourkela, India
Liverpool, UK
Rourkela, India

Daya K. Lobiyal
Durga Prasad Mohapatra
Atulya Nagar
Manmath N. Sahoo

Conference Organization

General Chair

Daya K. Lobiyal, Jawaharlal Nehru University, India

Organizing Chairs

Ram Shringar Rao, Ambedkar Institute of Advanced Communication Technologies and Research, India

Sushil Kumar, Jawaharlal Nehru University, India

Buddha Singh, Jawaharlal Nehru University, India

Program Chairs

Manmath N. Sahoo, National Institute of Technology Rourkela, India

Zaheeruddin, Jamia Millia Islamia University, India

Yulei Wu, University of Exeter, Exeter

Program Co-chairs

Sandip Rakshit, Kaziranga University, Assam, India

Syed Rizvi, Pennsylvania State University, USA

Yogesh H. Dandawate, SMIEEE, Vishwakarma Institute of Information Technology, India

Publication Chairs

Soubhagya Sankar Barpanda, National Institute of Technology Rourkela, India
Sambit Bakshi, National Institute of Technology Rourkela, India

Area Chairs

Asia: Omprakash Kaiwartya, Faculty of Computing Universiti Teknologi, Malaysia
Europe: Atilla Elci, Aksaray University, Turkey
USA: Adam Schmidt, Poznan University of Technology, Poland

Technical Track Chairs

Signal: Binod K. Kanaujia, AIACTR, India
Networking: Sanjay K. Soni, Delhi Technological University, Delhi, India
Computing: Nanhay Singh, AIACTR, India
Systems: Naveen Kumar, Indira Gandhi National Open University, India

Web Chairs

Sanjoy Das, Galgotias University, India
Rahul Raman, National Institute of Technology Rourkela, India

Technical Program Committee

Anand Paul, SMIEEEE, Kyungpook National University, Republic of Korea
Andrey V. Savchenko, National Research University Higher School of Economics, Russia
Ch Aswani Kumar, Vellore Institute of Technology, India
Dilip Singh Sisodia, National Institute of Technology Raipur, India
Ediz Saykol, Beykent University, Turkey
Flavio Lombardi, Roma Tre University of Rome, Italy
Jamuna Kanta Sing, SMIEEEE, Jadavpur University, India
Jaya Sil, Bengal Engineering and Science University, India

Krishnan Nallaperumal, SMIEEE, Manonmaniam Sundaranar University, India
Lopamudra Chowdhury, Jadavpur University, India
Narayan C. Debnath, Winona State University, USA
Nidul Sinha, SMIEEE, National Institute of Technology Silchar, India
Paulo Quaresma, University of Evora, Portugal
Patrick Siarry, SMIEEE, Université de Paris, France
Pradeep Singh, National Institute of Technology Raipur, India
Raghvendra Mall, University of Leuven, Belgium
Rajarshi Pal, Institute for Development and Research in Banking Technology, India
Sotiris Kotsiantis, University of Patras, Greece
Yogesh H. Dandawate, SMIEEE, Vishwakarma Institute of Information
Technology, Pune, India
Zhiyuan (Thomas) Tan, University of Twente, The Netherlands

Organizing Committee

Adesh Kumar, Shri Lal Bahadur Shastri Rashtriya Sanskrit Vidyapeetha, India
Ajay Sikandar, Jawaharlal Nehru University, India
Anil Kumar Sagar, Galgotias University, India
Arvind Kumar, Ambedkar Institute of Advanced Communication Technologies
and Research, India
Ashok Kumar Yadav, Amity School of Engineering and Technology, India
Indrani Das, Jawaharlal Nehru University, India
Kamlesh Kumar Rana, Galgotias College of Engineering and Technology (GCET),
India
Karan Singh, Jawaharlal Nehru University, India
Mahendra Ram, Jawaharlal Nehru University, India
Meenakshi Sihag, Guru Tegh Bahadur Institute of Technology, India
Prashant Singh, Northern India Engineering College, India
Rajesh Kumar Yadav, Delhi Technological University, India
Rameshwar Lal Ujjwal, Guru Gobind Singh Indraprastha University, India
Sanjeev Kumar, Ambedkar Institute of Advanced Communication Technologies
and Research, India
Shailender Kumar, Ambedkar Institute of Advanced Communication Technologies
and Research, India
Sunil Kumar, Jawaharlal Nehru University, India
Suresh Kumar, Ambedkar Institute of Advanced Communication Technologies
and Research, India

External Reviewers

Ajay Shankar Shukla, Central Council for Research in Ayurvedic Sciences, India

Amar Jeet Singh, Himachal Pradesh University, India

R. Kingsy Grace, Anna University, India

Shiv Prakash, Indian Institute of Technology Delhi, India

Snehasis Banerjee, Tata Consultancy Services Research, India

Taymaz Farshi, Gazi University, Turkey

Omprakash Kaiwartya, Jawaharlal Nehru University, India

Xavier Bellekens, University of Strathclyde, Glasgow

Contents

Part I Signal Processing Systems and Applications

DFT-DCT Combination Based Novel Feature Extraction Method for Enhanced Iris Recognition	3
Anunita Raghu, Meghana Gundlapalli and K. Manikantan	
Novel Digital Image Watermarking in SWT+SVD Domain.	13
Nikhil Purohit, M. Chennakrishna and K. Manikantan	
An Improved Histogram Bin Shifting Based Reversible Data Hiding of Color Images.	25
Smita Agrawal and Manoj Kumar	
Face Recognition Using Background Removal Based on Eccentricity and Area Using YCbCr and HSV Color Models.	33
Amith Lawrence, N.V. Manoj Ashwin and K. Manikantan	
An Efficient Multi-focus Image Fusion Approach Based on DWT	45
Sonam and Manoj Kumar	
A Novel Fuzzy Filter for Mixed Impulse Gaussian Noise from Color Images	53
M. Jayasree and N.K. Narayanan	
Face Recognition Using Snakes Algorithm and Skin Detection Based Face Localization	61
Rakshit Ramesh, Anoop C. Kulkarni, N.R. Prasad and K. Manikantan	
Quantifiable Image Nearness Approach Using Descriptive Neighbourhood	73
M. Sajwan and K.S. Patnaik	
Robust Speaker Verification Using GFCC Based <i>i</i>-Vectors	85
Medikonda Jeevan, Atul Dhingra, M. Hanmandlu and B.K. Panigrahi	

Enhanced Automatic Speech Recognition with Non-acoustic Parameters	93
N.S. Sreekanth and N.K. Narayanan	
Dynamic Gesture Recognition—A Machine Vision Based Approach . . .	105
N.S. Sreekanth and N.K. Narayanan	
Medical Image Security with Cheater Identification Using Secret Sharing Scheme	117
Arun Krishnan and Manik Lal Das	
The Role of Fractal Dimension, Lacunarity and Multifractal Dimension for Texture Analysis in SAR Image—A Comparison Based Analysis	127
Triloki Pant	
Efficient Storage and Processing of Video Data for Moving Object Detection Using Hadoop/MapReduce	137
Jyoti Parsola, Durgaprasad Gangodkar and Ankush Mittal	
Performance Evaluation of Digital Color Image Watermarking Using Column Walsh Wavelet Transform	149
Hemant B. Kekre, Shachi Natu and Tanuja Sarode	
Structural (Shape) Feature Extraction for Ear Biometric System	161
P. Ramesh Kumar and S.S. Dhenakaran	
 Part II Networking Theory and Distributed Systems	
DHAC Based Routing in Wireless Sensor Network with Asymmetric Links	171
Laxita Vyas, C.P. Gupta and Md Arqam	
Automatization of AAOCC to Find Trust Score of Websites	183
Manish Kumar Verma, Sarowar Kumar, Kumar Abhishek and M.P. Singh	
A Multi-level Weight Based Routing Algorithm for Prolonging Network Lifetime in Cluster Based Sensor Networks	193
Priyanka Pukhrambam, Sanghita Bhattacharjee and Himanish Shekhar Das	
An Approach to Optimize Unnecessary Energy Consumption During Dynamic Routing in Wireless Sensor Networks.	205
Narasimha Kamath and U.K. Anirudh	
Game Theoretic Modeling of Gray Hole Attacks in Wireless Ad Hoc Networks	217
Chintan Ketankumar Doshi, Sreecharan Sankaranarayanan, Vidyashankar B. Lakshman and K. Chandrasekaran	

Chi-Square Based Mobile Radio Propagation Model Analysis and Validation 227
 Lavanya Vadda, G. Sasibhushana Rao and L. Ganesh

Collision Theory Based Sentiment Detection of Twitter Using Discourse Relations 235
 Anuta Mukherjee and Saswati Mukherjee

Malicious Account Detection Based on Short URLs in Twitter 243
 Rasula Venkatesh, Jitendra Kumar Rout and S.K. Jena

Distance, Energy and Link Quality Based Routing Protocol for Internet of Things 253
 Kirshna Kumar, Sushil Kumar and Omprakash Kaiwartya

Effect and Suppression of Noise in 2D PC/OOC Scheme for Optical CDMA Systems 261
 Manisha Bharti, Ajay K. Sharma and Manoj Kumar

On-the-Fly Segment Density (OFSD) in Adaptive Beaconing System (ABS) Based Connectivity-Aware Geocast Routing (CAGR) in VANETs 269
 Durga Prasada Dora, Sushil Kumar and Puspanjali Mallik

Investigation and Analysis of Energy Efficiency in Distributed Antenna System: Technology Towards Green Communications 277
 Seetaiah Kilaru, S. Padmaja and K. Venugopal Reddy

A Novel Trust Based Access Control Model for Cloud Environment 285
 Pratap Kumar Behera and Pabitra Mohan Khilar

Live News Streams Extraction for Visualization of Stock Market Trends 297
 Vaishali Ingle and Sachin Deshmukh

Categorization of Cloud Workload Types with Clustering 303
 Piotr Orzechowski, Jerzy Proficz, Henryk Krawczyk and Julian Szymański

Development of a General Search Based Path Follower in Real Time Environment 315
 B.B.V.L. Deepak, G. Raviteja, Upasana Behera and Ravi Prakash

SDN Architecture on Fog Devices for Realtime Traffic Management: A Case Study 323
 Kshira Sagar Sahoo and Bibhudatta Sahoo

Maximizing Network Lifetime of Wireless Sensor Networks: An Energy Harvesting Approach 331
 Srikanth Jannu and Prasanta K. Jana

**Hybrid Network Intrusion Detection Systems:
A Decade's Perspective 341**
Asish Kumar Dalai and Sanjay Kumar Jena

Author Index 351

About the Editors

Dr. Daya K. Lobiya is currently serving as a professor in School of Computer and Systems Sciences in Jawaharlal Nehru University, India. His research works have been published in many journals and conference proceedings. He is a fellow of Institution of Electronics and Telecommunication Engineers, India.

Prof. Durga Prasad Mohapatra received his Ph.D. from the Indian Institute of Technology Kharagpur and is presently serving as an associate professor in NIT Rourkela, Odisha. His research interests include software engineering, real-time systems, discrete mathematics, and distributed computing. He has published more than 30 research papers in these fields in various international journals and conference proceedings. He has received several project grants from DST and UGC, Government of India. He has received the Young Scientist Award for the year 2006 by Orissa Bigyan Academy. He has also received the Prof. K. Arumugam National Award and the Maharashtra State National Award for outstanding research work in software engineering for the years 2009 and 2010, respectively, from the Indian Society for Technical Education (ISTE), New Delhi. He is going to receive the Bharat Shiksha Ratan Award for significant contribution in academics awarded by the Global Society for Health and Educational Growth, Delhi.

Prof. Atulya Nagar holds the foundation chair as a professor of mathematical sciences at Liverpool Hope University where he is the dean of Faculty of Science. He has been the head of the Department of Mathematics and Computer Science since December 2007. A mathematician by training, he is an internationally recognized scholar working at the cutting edge of applied nonlinear mathematical analysis, theoretical computer science, operations research, and systems engineering, and his work is underpinned by strong complexity–theoretic foundations. He has an extensive background and experience of working in the universities in the UK and India. He has edited volumes on intelligent systems and applied mathematics; he is the editor in chief of the International Journal of Artificial Intelligence and Soft Computing (IJAISSC) and serves on editorial boards for a number of prestigious journals such as the Journal of Universal Computer Science (JUCS). Professor Nagar received a prestigious Commonwealth Fellowship for pursuing his

doctorate (D.Phil.) in applied nonlinear mathematics, which he earned from the University of York in 1996. He holds B.Sc. (Hons.), M.Sc., and M.Phil. (with distinction) from the MDS University of Ajmer, India.

Dr. Manmath N. Sahoo received his M.Tech and Ph.D. degrees in computer science in the year 2009 and 2014, respectively, from the National Institute of Technology (NIT) Rourkela, India. He is an assistant professor in the Department of Computer Science and Engineering, NIT Rourkela, India. He has served as reviewer, guest editor, track chair, and program chair in many reputed journals and conferences. His research interests include mobile ad hoc networks, fault tolerance, and sensor networks. He is a professional member of prestigious societies such as IEEE, CSI, and IEL.

Part I
Signal Processing Systems
and Applications

DFT-DCT Combination Based Novel Feature Extraction Method for Enhanced Iris Recognition

Anunita Raghu, Meghana Gundlapalli and K. Manikantan

Abstract Iris Recognition (IR) using conventional methods is a challenging domain, and incorporating a combination of two transforms along with proposed novel extraction technique possesses the efficacy to address the problem at hand. This paper throws light upon the proposed unique *Combined DFT-DCT* feature extraction along with the inclusion of a *disc shaped morphological structuring element* in the preprocessing stage. Two novel methods, namely *astroid* and *astroid ring shaped extraction* techniques are proposed, and Binary Particle Swarm Optimization (BPSO) based algorithm for feature selection has been employed to procure the optimal subset of features from the feature space. Experimental results that have been obtained by implementing the proposed technique on two standard iris databases, IITD and MMU, lucidly outline the promising performance of the astroid shaped feature extraction resulting in a significant increase in rate of recognition accompanied by considerably lower number of features for iris recognition.

Keywords Iris recognition · Discrete cosine transform · Discrete fourier transform · Binary particle swarm optimization · Feature selection · Feature extraction

1 Introduction

Biometric Systems involve recognition of human physiological features namely face, retina, palm print, voice, fingerprint and iris, which are unique to an individual. Apart from its inherent accuracy, an added advantage of biometric systems is that existing

A. Raghu · M. Gundlapalli · K. Manikantan (✉)
Department of Electronics and Communication Engineering,
M S Ramaiah Institute of Technology, Bangalore 560054, India
e-mail: kmanikantan2009@gmail.com

A. Raghu
e-mail: anunitaraghu06@gmail.com

M. Gundlapalli
e-mail: gundlapallimeghana@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_1

technology, such as hand-held devices and electronic gadgets, can incorporate it with ease (Ref. [1]).

Iris Recognition has gained immense importance in recent times due to its age invariant features (unlike fingerprints which are eventually smoothed) and fast recognition. References [2, 3] proposes circular sector and triangular feature extraction using DCT while Ref. [4] introduces the concept of Binary Particle Swarm Optimization (BPSO) applied for Face Recognition. Reference [5] presents the role of the Golden Ratio in BPSO, assigning values to the cognitive and social factors. Reference [6] is a path-breaking work regarding the introduction of iris recognition in the field of biometrics and Ref. [7] proposes an algorithm to identify the iris using circular detection operator method. References [8, 9] are recent surveys carried out in the field of biometrics.

2 Problem Definition and Contributions

Reference [2] proposes a DCT based circular sector and triangular feature extraction. However, with the usage of DCT alone, recognition rate cannot be increased and the maximum number of features extracted for both the shapes is high, which in turn results in increased processing time.

To combat these limitations, our proposal consists of a *Combination of DFT and DCT* embodied feature extraction which results in a significantly higher recognition rate. Two novel approaches to considerably reduce the computation time are the *astroid* and *astroid ring shaped feature extraction*. The introduction of a *morphological structuring element* in the preprocessing stage enhances the features of the iris, ameliorating the recognition rate. The preprocessing stage also consists of Gaussian Blurring, Gamma Intensity Correction (GIC) along with Histogram Equalization (HE). Applying *Binary Particle Swarm Optimization (BPSO)* with *Euclidean Classifier* results in further diminution of the number of features extracted.

The subsequent portion of the paper is structured in the following manner: Sect. 3 addresses image preprocessing based on illumination and Sect. 4 introduces the proposed DFT-DCT Combination. Section 5 throws light upon the proposed shape of feature extraction while Sect. 6 deals with BPSO based feature selection. The proposed IR system with its experimental outcomes are illustrated in Sects. 7, and 8 summarizes the results obtained.

3 Image Preprocessing

Preprocessing an image is a fundamental step used to reduce information loss in the image in order to improve its suitability for subsequent procedures. The grayscale Iris image is not well suited for feature extraction due to non-uniform illumination, improper focus or insufficient lighting. Figure 1 represents the block diagram of our

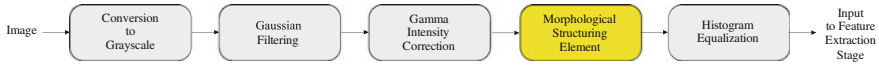


Fig. 1 Block diagram of the preprocessing system

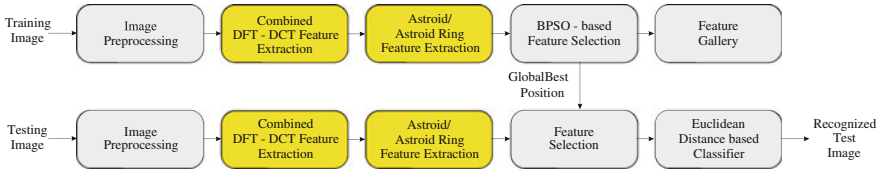


Fig. 2 Block diagram of the proposed IR system

suggested preprocessing system while Fig. 2 depicts the block diagram of the proposed IR system.

Gaussian Filtering: Gaussian blurring is used to obtain smooth edges using a filter of optimum dimensions 4×4 along with a standard deviation value of $\sigma = 1.2$.

Gamma Intensity Correction (GIC): According to the Power Law, the relation between the original image and gamma intensity corrected image is given by Eq. 1.

$$I(x, y) = I(x, y)^{1/\gamma} \quad (1)$$

where the exponent is gamma (γ), which determines the degree of brightness of the corrected image (Ref. [10]). Applying GIC with an optimum value of $\gamma = 0.78$ aided the enhancement of the iris' features.

Histogram Equalization (HE): HE is a graphical representation of the normalized number of pixels versus various values of intensity. It is a non-linear transformation that balances brightness of the image, giving a high contrast image which is visually distinguishable from the original image by the human eye.

3.1 Proposed Morphological Structuring Element

Constructed by employing a family of techniques referred to as structuring element decomposition, structuring elements are those wherein morphological operations by sizable structuring elements can be computed rapidly with a smaller sequence of structuring elements. The shape of the structuring element can be well defined, such as a polygon, or arbitrary. It contributes to obtaining sequiturs on how this shape misses or fits the shapes in the image. Since the iris profile is circular, the proposed concept involves the utilization of a flat, disk-shaped structuring element of radius $R = 23$. It is to be noted that the disk and ball shaped structuring elements are obtained by approximations, while the other structuring elements are exact.

4 Proposed Combination of DFT-DCT for Feature Extraction

Discrete Fourier Transform (DFT) is obtained by decomposing a sequence of values into components of various frequencies. Fast Fourier Transform (FFT), as the name suggests, is a fast and more efficient algorithm to compute the DFT of an image. DFT is mathematically represented as shown in Eq. 2 where $W_N = e^{-j2\pi/N}$.

$$X(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) W_N^{mk} W_N^{nl} \quad , \quad 0 \leq k, l \leq N - 1 \quad (2)$$

On applying DFT, the *low frequency components*, constituting majority of the distinguishable features in the image, get *amassed at the four corners* of the DFT spectrum. Since DCT coefficients are real, we consider the absolute value of the DFT terms.

Discrete Cosine Transform (DCT) enunciates a finite sequence of multiple points of data in terms of a sum of cosine functions that oscillate at diversified frequencies. DCT is mathematically expressed as Eq. 3. The application of DCT to the absolute value of the DFT coefficients obtained previously results in the *accumulation of the low frequency components at the top left corner* of the DCT spectrum. The high frequency components that correspond to the minutiae are inconsequential in recognition and hence, are discarded.

$$F(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos \left[\frac{\pi(2x+1)}{2M} \right] \cos \left[\frac{\pi(2y+1)}{2N} \right] \quad (3)$$

where

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{M}}, u = 0 \\ \sqrt{\frac{2}{M}}, u = 1, 2, \dots, M - 1 \end{cases} \quad , \quad \alpha(v) = \begin{cases} \frac{1}{\sqrt{N}}, v = 0 \\ \sqrt{\frac{2}{N}}, v = 1, 2, \dots, N - 1 \end{cases}$$

5 Proposed Shape of Feature Extraction

Any geometrical shape can be used to extract the top left corner features of the DCT spectrum. Conventionally, square extraction yields N^2 features where N is the number of pixels on each side of the square measured from the origin of the DCT spectrum. However, this has a relatively low recognition rate and extremely large number of features which leads to substantial computation time. As stated in Sect. 2, an alternative method involves the usage of circular sector and triangular shaped extraction, which consider $\pi N^2/4$ and $N^2/2$ features respectively. While recognition rate increases when compared to square extraction, it is found that the same can

be significantly increased along with abatement in the number of features extracted by applying the proposed extraction techniques. This paper presents two novel techniques for the shape of the region of extraction, namely the astroid and astroid ring. The equation of the astroid is given by Eq. 4.

$$x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}} \quad (4)$$

where 'a' is a constant.

5.1 Astroid Feature Extraction

An astroid is a hypocycloid with 4 cusps. We propose to extract the top left corner features of the DCT spectrum using one quadrant of the astroid. If N is the number of pixels measured from the origin and is equal to the constant 'a', then the number of features extracted, A, is given by Eq. 5.

$$A = \frac{3}{32}\pi a^2 = \frac{3}{32}\pi N^2 \approx 0.29452N^2 \quad (5)$$

5.2 Astroid Ring Feature Extraction

Astroid ring feature extraction involves extracting the pixels enclosed in the non-overlapping regions between two concurrent astroids of different constants, say 'a' and 'b'. If N_1 and N_2 are the number of pixels measured from the origin for each astroid and are equal to the constants 'a' and 'b' respectively such that $b > a$, then the number of features extracted, A, is given by Eq. 6.

$$A = \frac{3}{32}\pi (b^2 - a^2) = \frac{3}{32}\pi (N_2^2 - N_1^2) \approx 0.29452 (N_2^2 - N_1^2) \quad (6)$$

Table 1 presents a theoretical comparison between the maximum number of features extracted for different shapes, where N is the number of pixels measured from

Table 1 Comparison between different shapes for feature extraction

Shape of extraction	Square	Circular sector	Triangular	Astroid	Ring of astroid
Max. No. of features	N^2	$0.79N^2$	$0.5N^2$	$0.29N^2$	$0.29(N_2^2 - N_1^2)$

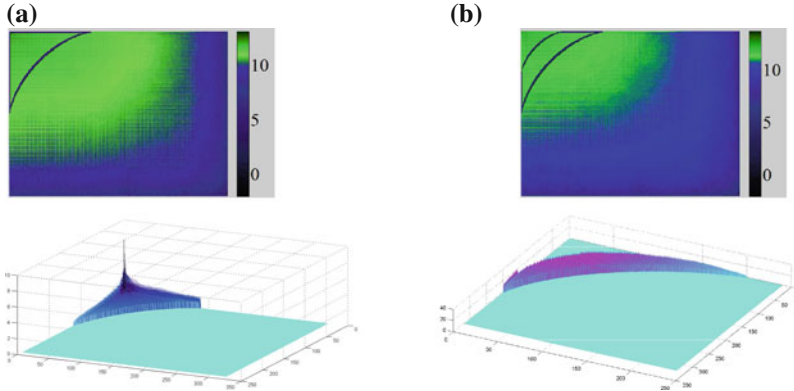


Fig. 3 DCT-DFT spectrum and surface plot of **a** Proposed astroid **b** Proposed astroid ring

the origin. From Fig. 3, it is observed that the proposed shapes of extraction provide the least number of features, providing satisfactory reduction in computation time.

6 Feature Selection Using Binary Particle Swarm Optimization

Introduced by Eberhart and Kennedy in the year 1995, Particle Swarm Optimization (PSO) is based on the behavioral pattern of birds, bees and fishes flocking to search for food. Velocity (represented as v_t^i) and position (represented as x_t^i) of the particles are the two variables under consideration and these are regularly updated until optimum convergence is achieved. Binary Particle Swarm Optimization (BPSO) maps the continuously varying position into binary bits. This process of mapping uses the binary sigmoidal function given by Eq. 7. With a swarm size of 35, inertial damping factor $\omega = 0.9$ is set. c_1 is the cognitive factor and c_2 is the social factor. From Ref. [5], c_1 is assigned the value of the golden ratio 1.618 and c_2 is assigned the inverse golden ratio 0.618. Velocity is updated as shown in Eq. 8.

$$f(x) = \frac{1}{1 + e^{-v_i^{t+1}}} \quad , \quad x_{id} = \begin{cases} 1 & \text{if } \text{rand}_3 < f(x) \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$v_i^{t+1} = \omega \times v_i^t + c_1 \times \text{rand}_1 \times (p_{best}^i - x_i^t) + c_2 \times \text{rand}_2 \times (g_{best} - x_i^t) \quad (8)$$

The Fitness Function characterizes the rate of misclassified patterns. Let rand_1 , rand_2 , rand_3 , and v_i^t lie in the range (0,1). p_{best}^i and g_{best} indicate the previous best position visited and global best position visited by any particle, respectively. They

are assigned on the basis of the highest value of the Fitness Function, represented by Eq. 9 where W_j and N_i symbolize the number of subjects, and the number of image samples for each subject. M_i and M_0 correspond to the means of the respective subjects and overall mean in the feature space. In Eq. 9, E denotes the Euclidean Classifier which calculates the N-dimensional distance vector, where N is the number of features extracted, q_i is the test image feature vector and p_i is the feature vector under consideration.

$$F = \sqrt{\sum_{i=1}^L (M_i - M_o)^t (M_i - M_o)} \quad , \quad E = \sqrt{\sum_{i=1}^N (p_i - q_i)^2} \quad (9)$$

where,

$$M_i = \frac{1}{N_i} \sum_{j=1}^L W_j^{(i)} \quad , \quad M_0 = \frac{1}{N} \sum_{i=1}^L N_i M_i$$

7 Discussion of Proposed IR System and Experimental Results

The experiments have been performed on 2 standard iris databases, Multi Media University (MMU) and Indian Institute of Technology—Delhi (IITD). Tabulated results for various values of N for both the proposed shapes of extraction are as indicated by Tables 2 and 3 for astroid and astroid ring respectively using (Ref. [11]). Any system having the ratio of training to testing images less than 1, is said to be intelligent and is more preferable. It was found that results with preprocessing were much better than those without. Both databases are in the Bitmap (BMP) format.

The MMU database (Ref. [12]) consists of 45 subjects with 5 images each and the images contain the iris as well as the eyelashes. The image size is 240×320 .

Table 2 Recognition rate (RR) and number of features extracted for different values of constant 'a' of proposed astroid IR system

Constant a	No. of features	MMU RR (%)	IITD RR (%)
9	33	51.41	77.26
19	106	66.15	87.10
44	429	75.93	91.27
59	804	80.15	93.25
74	1130	82.00	93.55
99	2020	84.67	93.73
119	2961	85.19	93.88
139	3716	84.37	94.75

Table 3 Rate of recognition (RR) and number of features extracted for distinct values of constants 'a' and 'b' of proposed astroid ring IR system

Constant a	Constant b	No. of features	MMU RR (%)	IITD RR (%)
9	44	424	79.11	92.28
9	59	790	81.70	93.75
14	74	1107	81.93	92.98
14	84	1527	84.07	93.08
19	99	1974	82.22	93.31
19	119	2906	81.70	93.16

**Fig. 4** Samples images of **a** MMU database **b** IITD database

Specimen images are as shown in Fig. 4a and Tables 2, 3 indicate the results for various value of constants for training to testing ratio of 2:3.

The IITD database (Ref. [13]) contains 224 subjects with 10 images each. The images contain the iris and the eyelashes. The image size is 240×320 . Specimen images are as shown in Fig. 4b and Tables 2, 3 include the results obtained for training to testing ratio of 4:6.

As observed from Table 2, the number of features extracted increases along with increase in constant 'a'. Best results were obtained for 'a' = 119 for MMU database and 'a' = 139 for IITD database.

Considering Table 3, the values of 'a' and 'b' must be chosen with utmost care to obtain good recognition rate. Best results were obtained for 'a' = 14, 'b' = 84 for MMU database and 'a' = 9, 'b' = 59 for IITD database.

With increase in the value of the constant(s), the performance of the system improves and eventually saturates for high values. However, near saturation, simply increasing the value of the constant(s) results in higher computation time and more number of features extracted with a negligibly marginal increase in recognition rate. Thus, it is a trade-off between computation time (number of features extracted) and rate of recognition.

Figure 5a compares the rate of recognition and number of features extracted for distinct shapes. The proposed IR System is found to produce optimum results in terms of both, the rate of recognition and the number of features extracted when compared to Ref. [2] and conventional shape of feature extraction. Figure 5b, c indicate the results for dissimilar training to testing ratios for MMU and IITD databases respectively. In both cases, the recognition rate saturates beyond a certain training to testing ratio and increase in recognition rate is minimal. It can also be observed that increasing the number of images used for training leads to increase in recognition rate.

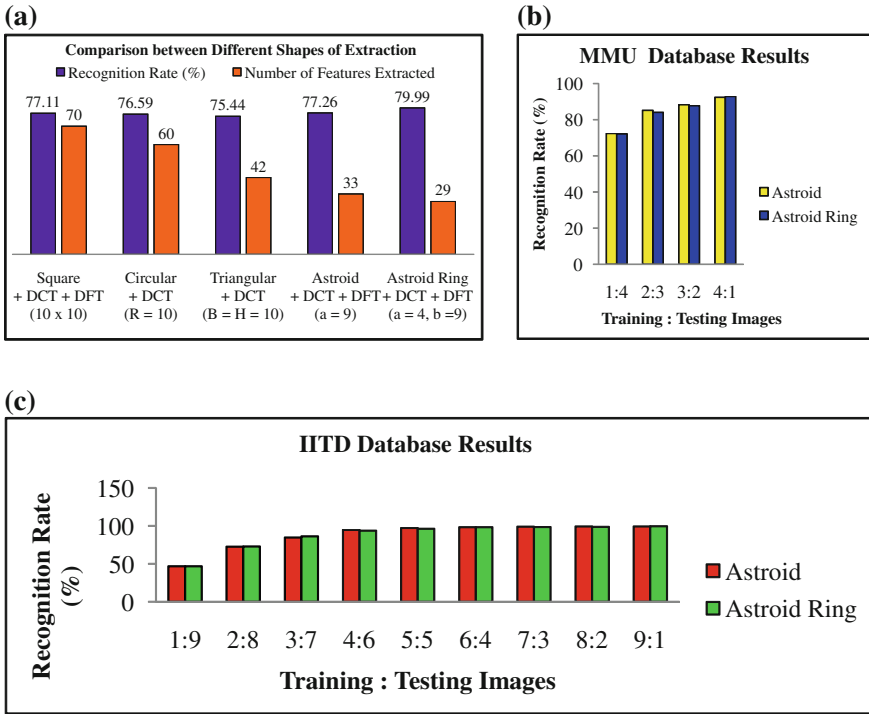


Fig. 5 Illustration of performance of the proposed IR system

8 Conclusions

A novel Iris Recognition System which incorporates a unique Combination of DFT-DCT, along with the application of a disc shaped morphological structuring element in the preprocessing stage, in addition to astroid and astroid ring shaped feature extraction has been proposed.

The application of this system in conjugation with Binary Particle Swarm Optimization (BPSO) based feature selection and the Euclidean Classifier was found to significantly decrease the number of features selected from the optimum feature subset along with a noteworthy increase in the recognition rate. The experimental results were found to be in accordance with the expected outcomes for two prominent iris databases, namely the MMU and IITD databases.

Integrating classifiers, such as the Support Vector Machine (SVM) and Random Forest, is expected to improve the existing results. This work is currently in progress and seems promising for creating improved IR systems in the future.

References

1. Kevin W. Bowyer, Karen Hollingsworth, Patrick J. Flynn.: Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281–307 (2008)
2. Abhiram M.H., Chetan Sadhu, Manikantan K, S. Ramachandran.: Novel DCT Based Feature Extraction for Enhanced Iris Recognition. *International Conference on Communication, Information & Computing Technology (ICCICT)* (2012)
3. Neal S. Latman, Emily Herb.: A field study of the accuracy and reliability of a biometric iris recognition system. *Science & Justice*, vol. 53, no. 2, pp. 98–102 (2013)
4. Rabab M. Ramadan, Rehab F. Abdel - Kader.: Face Recognition Using Particle Swarm Optimization-Based Selected Features. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 2, no. 2, pp. 57–59 (2009)
5. Arun B Vishwanath, Darshan Kumar S Yaradoni, Manikantan K.: Optimal Multilevel Thresholds based on Tsallis Entropy Method using Golden Ratio Particle Swarm Optimization for Improved Image Segmentation. *International Conference on Communication Technology and System Design 2011*, vol. 30, pp. 364–371 (2012)
6. John Daugman.: New Methods in Iris Recognition. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 37, no. 5, pp. 1167–1175 (2007)
7. Jinghui Li, Bairui Tao, Yanchun Wang, Xibing Li.: Research and Implementation of Iris Recognition Algorithm. *Procedia Engineering*, vol. 29, pp. 3353–3358 (2012)
8. Guodong Guo, Alice Lai.: A survey on still image based human action recognition. *Pattern Recognition*, vol. 46, Issue 12 (2013)
9. Ishan Nigam, Mayank Vatsa, Richa Singh.: Ocular biometrics: A Survey of Modalities and Fusion Approaches. *Information Fusion*, vol. 26 (2015)
10. Rafael C. Gonzalez, Richard E. Woods.: *Digital Image Processing*. Prentice Hall, Third edition (2008)
11. Matlab, <http://www.mathworks.in/>
12. MMU Database, <http://pesona.mmu.edu.my/ccteo/>
13. IIT Delhi database, <http://web.iitd.ac.in/biometrics/DatabaseIris.htm>
14. Ching-Han Chen, Chia-Te Chu.: High performance iris recognition based on 1-D circular feature extraction and PSO-PNN classifier. *Expert Systems with Applications*, vol. 36, no. 7, pp. 10351–10356 (2009)
15. Donald M. Monro, Soumyadip Rakshit, Dexin Zhang.: DCT-Based Iris Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 586–595 (2007)
16. Ajay Kumar, Arun Passi.: Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognition*, vol. 43, no. 3, pp. 1016–1026 (2010)
17. Izem Hamouchene, Saliha Aouat.: A New Texture Analysis Approach for Iris Recognition. *AASRI Procedia*, vol. 9, 2014 AASRI Conference on Circuit and Signal Processing (CSP) (2014)

Novel Digital Image Watermarking in SWT+SVD Domain

Nikhil Purohit, M. Chennakrishna and K. Manikantan

Abstract In this paper, Digital Watermarking is carried out in the frequency domain and the technique proposed uses single level Stationary Wavelet Transform (SWT) along with Singular Value Decomposition (SVD). SWT is used over other transformations because of its non-decimation and shift invariance property. The singular values of SWT transformed watermark image is embedded into the singularly decomposed HH sub-band (sub-image) of R, G or B channel of a Host color image. The experimental results of watermarked images shows increase in Peak Signal to Noise Ratio (PSNR) and the extracted watermark image is highly correlated with the original watermark for various types attacks.

Keywords Digital watermarking · Stationary wavelet transform · Singular value decomposition · Peak signal to noise ratio

1 Introduction

Digital watermarking is used extensively in this digital era wherein the sensitive information is merged with a carrier which acts like a camouflage and only the proper decoding algorithm is able to extract the information which is hidden inside the carrier. It is also used for various applications like authentication, copyright protection [1], information hiding [2], broadcast monitoring, content identification [3] and filtering i.e. for instance if a person is watching a movie scene, pop ups like promotions and advertisements are triggered by identifying the content being watched.

N. Purohit · M. Chennakrishna · K. Manikantan (✉)
Department of Electronics and Communication Engineering,
M S Ramaiah Institute of Technology, Bangalore 560 054, India
e-mail: kmanikantan2009@gmail.com

N. Purohit
e-mail: ramesh.nikhil89@gmail.com

M. Chennakrishna
e-mail: chennakrishna150@gmail.com

Similarly, it is also used for content blocking by recognising a specific part of the content. Hence watermarking is widely used for security purposes. The information can be in the form an image, video or audio and corresponding to it the carrier is chosen [4–6]. The multimedia industry are still facing a lot of challenges and trying find algorithms that are more robust and the watermarks which cannot be removed or altered, thereby eliminating piracy. Watermark can be visible or invisible [7] depending on the application.

2 Problem Definition and Contribution

Watermarking in the frequency domain [8, 9] is preferred since it alters the coefficients obtained after using different transformations rather than pixels which is done in spatial domain [10]. Many techniques and combinations have been implemented using DWT, DFT and SVD based [11–14]. But the PSNR values obtained are very low even though the correlation between the original and the extracted watermark is better. To improve the PSNR values this paper proposes.

SWT+SVD domain watermarking using HH sub-band:

Combination of SWT and SVD technique is used to embed the information into the host image and the use of HH sub-band instead of LL, HL, LH sub-bands provides better PSNR values for R, G and B channels.

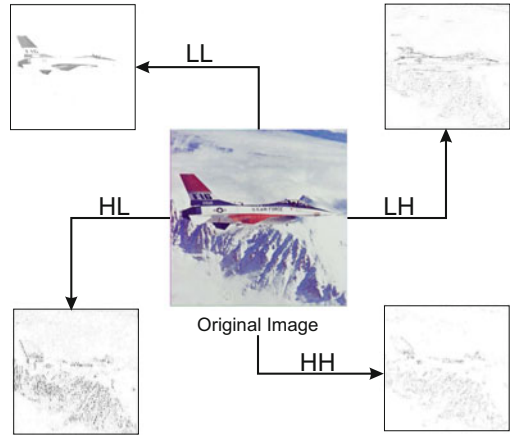
3 Fundamental Concepts

This section provides the basic ideas about the mathematical functions used in the proposed algorithm, they are described as follows

3.1 Stationary Wavelet Transform (SWT)

The single level two dimensional SWT [15] disintegrates the image into four sub-bands (sub-images) each having size same as that of the original image as shown in Fig. 1, i.e. LL, LH, HL, HH. It is also known as undecimated transform since there is no downsampling of the images. LL is the approximate image of input image it is low frequency sub-band. LH sub-band, HL sub-band and HH sub-band gives the horizontal, vertical and diagonal features of original image respectively. These sub-bands are used for integrating the watermark.

Fig. 1 Decomposition of the image into four components (LH, HL and HH sub-images are in negative form for better clarity) when SWT is applied



3.2 Singular Value Decomposition (SVD)

Singular value decomposition (SVD) [16] uses a rectangular matrix of an image ($m \times n$) where m is the number rows and n is the number of columns of the image. The mathematical equation for SVD is given below

$$A = U_{n \times n} \times S_{m \times n} \times V_{m \times m}^T \tag{1}$$

where U is the left singular vector, V is the right singular vector and S is diagonal matrix and it is singular. The U and V matrix are orthogonal to each other. The mathematical equation is given below

$$U \times U^T = I_{n \times n} \tag{2}$$

$$V \times V^T = I_{m \times m} \tag{3}$$

The S matrix is in descending order and it is always real. If A matrix is real, then U and V are also real. SVD can be calculated by finding eigenvector and eigenvalue of $A^T A$ and AA^T . The eigenvector $A^T A$ is given by columns of U and eigenvalue is given by columns of V .

4 Proposed SWT+SVD Domain Watermarking Using HH Sub-band

The combination of SWT+SVD along with their properties serve as a good alternative over other transforms. The host is separated into R, G and B channels and to one of the channels 2D-SWT (single level) is applied. The HH sub-band is chosen since it

gives better results compared to other sub-bands. Then it is decomposed using SVD. Similarly the HH sub-band of the watermark is also obtained by applying SWT and then the singular matrix is obtained by applying SVD. The singular values of the host images and the watermark images are combined and then inserted back into the selected host channel using ISWT and all the separated channels are combined to obtain the final watermarked image.

4.1 Watermark Embedding

Figure 2 shows the block diagram of the embedding process. The steps involved are as follows

1. The color host image is split into the individual channels (Red, Green and Blue).
2. To any one of the channels the 2D-SWT (single level) of SYM4 wavelet family is applied.
3. Among the different sub-bands obtained from SWT, the HH sub-band is selected.
4. SVD is applied to the HH sub-image to decompose it into U, S and V matrices

$$A = U \times S \times V^T \tag{4}$$

where A is the HH sub-image.

5. Similarly 2D-SWT (single level) of SYM4 wavelet family is applied to grey scale watermark image and SVD is applied to the HH sub-image to decompose it into U_1, S_1 and V_1 matrices

$$B = U_1 \times S_1 \times V_1^T \tag{5}$$

6. Insert the S_1 matrix into the S matrix using a scalar quantity α i.e. indirectly the insertion of the watermark is done.

$$S_2 = S + \alpha S_1 \tag{6}$$

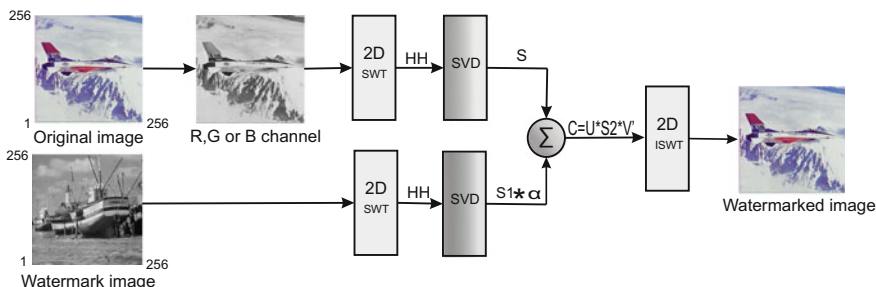


Fig. 2 Block diagram for the embedding process

where α is scaling factor which determines the strength of the watermark over the host image.

- Combine the newly obtained S_2 with U and V matrices obtained from Eq. 4 as shown in Eq. 7

$$C = U \times S_2 \times V \tag{7}$$

- The transformed HH sub-image C is combined with unaltered sub-images (LL, LH and HL) of the host image by applying ISWT to obtain the complete image of the selected channel.
- The resultant channel obtained is concatenated with the other two channels and the resulting image obtained is the required watermarked image.

4.2 Watermark Extraction

Figure 3 shows the block diagram of the extraction process. The steps involved are as follows

- The watermarked image is split into R, G or B channel.
- Level one 2D-SWT (SYM4) is applied to the channel into which the watermark was embedded.
- SVD is applied to the HH sub-band since the watermark is present in this sub-band and corresponding U, S and V matrices are obtained.

$$A^* = U^* \times S^* \times V^{*T} \tag{8}$$

where A^* is the HH sub-image.

- The difference of the singular matrix S^* of the watermarked image and the singular matrix S of original RGB image will fetch back the singular matrix of the

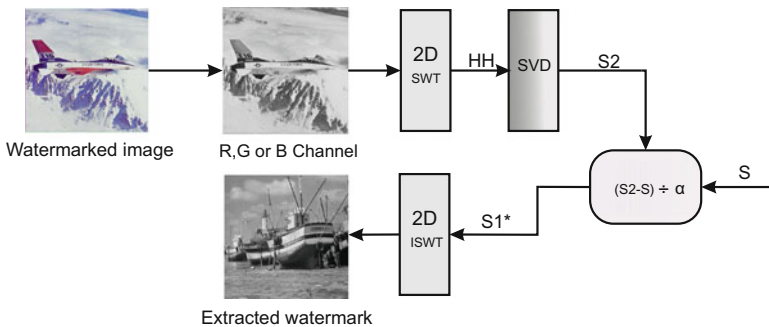


Fig. 3 Block diagram for the extraction process

watermark i.e. the separation of the watermark is done from the watermarked image. It is given by Eq. 9

$$S_1^* = \frac{(S^* - S)}{\alpha} \quad (9)$$

5. The new sub-image B^* is formed by multiplying S_1 with the U_1 and V_1 matrices obtained from Eq. 5

$$B^* = U_1 \times S_1^* \times V_1^T \quad (10)$$

where B^* is the HH sub-image of the extracted watermark.

6. This sub-image is combined with the unaltered LL, LH and HL sub-images of the watermark image to obtain the complete image by applying ISWT.
7. The obtained image almost matches the watermark image which was embedded during the embedding process. Correlation between the two images can be found to determine their resemblance.

5 Experimental Results and Analysis

The evaluation of the proposed algorithm is done by calculating the PSNR and also by finding correlation between watermark image before embedding and the watermark image obtained after the extraction. The MATLAB tool [17] is used for the evaluation of the algorithm. Table 1 shows the PSNR values for different combinations of host and the watermark images each of size 256×256 pixels. The PSNR is calculated using Eq. 11.

Table 1 PSNR values obtained for the watermarked image, after embedding the watermark image into R, G and B channels in dB

Host	Watermark	PSNR_R	PSNR_G	PSNR_B
Lena	Cameraman	97.52	95.96	96.80
Pepper	Airplane	99.75	97.58	99.45
Airplane	Boat	97.96	95.95	99.56
Baboon	Hunter	90.79	88.51	88.57
Tulips	House	103.34	99.14	101.98
Barbara	Livingroom	91.27	92.01	92.13
Plane	Barbara	96.27	96.17	96.31
Map	Butterfly	90.67	90.87	90.26
Lena	Mysail	97.58	96.03	96.87

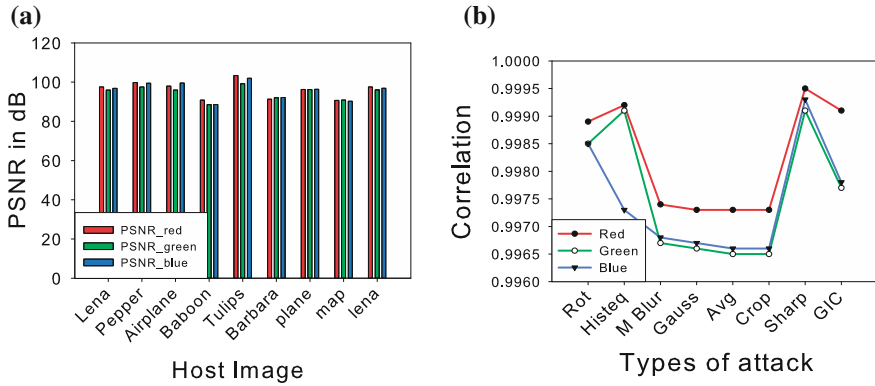


Fig. 4 a PSNR values for all three channels b Correlation coefficients for different types of attacks

$$PSNR = 20 \log_{10} \left(\frac{255}{RMSE} \right) \quad (11)$$

where RMSE is the root mean square error, $I(i, j)$ is the original image and the $\bar{I}(i, j)$ is the watermarked image.

$$RMSE = \sqrt{\frac{1}{MN} \left(\sum_{i=1}^M \sum_{j=1}^N (I(i, j) - \bar{I}(i, j))^2 \right)} \quad (12)$$

The correlation of the original watermark with the extracted watermark gives the information of how much the extracted watermark resembles the original watermark before embedding. Figure 4 shows the plot of PSNR values for different host images and correlation values obtained for different attacks. The correlation is calculated using Eq. 13.

$$corr(X, Y) = \frac{\sum_{i=1}^M \sum_{j=1}^N (X_{ij} - \bar{X})(Y_{ij} - \bar{Y})}{\sqrt{\left(\sum_{i=1}^M \sum_{j=1}^N (X_{ij} - \bar{X})^2 \sum_{i=1}^M \sum_{j=1}^N (Y_{ij} - \bar{Y})^2 \right)}} \quad (13)$$

To test whether the watermarking technique is reliable, various processing methods such as Cropping (Crop), Rotation (Rot), Gaussian Filter (Gauss), Histogram Equalization (Hist), Sharpening (Sharp), Average filter (Avg), Motion Blur (M Blur), Gamma Intensity Correction (GIC) have been applied. Tables 2, 3, 4 show the correlation values obtained for different channels after applying various attacks and Fig. 5 shows the corresponding watermarked images.

Table 2 Correlation coefficients obtained for watermark extracted from red channel after the application of various kinds of attacks

Host	Watermark	Rotate	Histeq	M blur	Gauss	Avg	Crop	Sharp	GIC
Lena	Cameraman	0.9993	0.9989	0.9982	0.9981	0.9981	0.9981	0.9990	0.9984
Pepper	Airplane	0.9998	0.9995	0.9990	0.9999	0.9989	0.9989	0.9995	0.9991
Airplane	Boat	0.9998	0.9999	0.9989	0.9989	0.9989	0.9989	0.9996	0.9991
Baboon	Hunter	0.9989	0.9992	0.9974	0.9973	0.9973	0.9973	0.9995	0.9981
Tulips	House	1.0000	0.9998	0.9996	0.9996	0.9996	0.9996	0.9998	0.9997
Barbara	Livingroom	0.9991	0.9992	0.9977	0.9977	0.9976	0.9976	0.9994	0.9984
Plane	Barbara	0.9982	0.9993	0.9969	0.9969	0.9968	0.9969	0.9984	0.9976
Map	Butterfly	0.9989	0.9989	0.9968	0.9967	0.9966	0.9967	0.9984	0.9972
Lena	Mysail	0.9999	0.9998	0.9991	0.9990	0.9989	0.9990	0.9999	0.9993

Table 3 Correlation coefficients obtained for watermark extracted from blue channel after the application of various kinds of attacks

Host	Watermark	Rotate	Histeq	M blur	Gauss	Avg	Crop	Sharp	GIC
Lena	Cameraman	0.9989	0.9988	0.9980	0.9980	0.9979	0.9979	0.9989	0.9985
Pepper	Airplane	0.9997	0.9994	0.9988	0.9988	0.9987	0.9987	0.9994	0.9994
Airplane	Boat	0.9998	0.9999	0.9988	0.9987	0.9987	0.9987	0.9995	0.9991
Baboon	Hunter	0.9985	0.9991	0.9967	0.9966	0.9965	0.9965	0.9991	0.9977
Tulips	House	0.9999	0.9998	0.9995	0.9995	0.9994	0.9994	0.9998	0.9997
Barbara	Livingroom	0.9991	0.9993	0.9979	0.9979	0.9978	0.9978	0.9993	0.9985
Plane	Barbara	0.9985	0.9993	0.9969	0.9969	0.9968	0.9969	0.9982	0.9975
Map	Butterfly	0.9988	0.9994	0.9968	0.9968	0.9967	0.9967	0.9983	0.9973
Lena	Mysail	0.9998	0.9998	0.9989	0.9988	0.9987	0.9987	0.9998	0.9995

Table 4 Correlation coefficients obtained for watermark extracted from blue channel after the application of various kinds of attacks

Host	Watermark	Rotate	Histeq	M blur	Gauss	Avg	Crop	Sharp	GIC
Lena	Cameraman	0.9988	0.9992	0.9981	0.9981	0.9980	0.9980	0.9980	0.9980
Pepper	Airplane	0.9997	0.9996	0.9989	0.9989	0.9989	0.9989	0.9989	0.9989
Airplane	Boat	0.9999	0.9999	0.9990	0.9990	0.9990	0.9990	0.9991	0.9990
Baboon	Hunter	0.9985	0.9973	0.9968	0.9967	0.9966	0.9966	0.9993	0.9978
Tulips	House	0.9998	1.0000	0.9996	0.9996	0.9996	0.9996	0.9998	0.9999
Barbara	Livingroom	0.9991	0.9991	0.9979	0.9979	0.9978	0.9977	0.9978	0.9978
Plane	Barbara	0.9988	0.9994	0.9969	0.9969	0.9969	0.9969	0.9984	0.9975
Map	Butterfly	0.9987	0.9998	0.9967	0.9966	0.9965	0.9966	0.9984	0.9973
Lena	Mysail	0.9998	1.0000	0.9990	0.9989	0.9988	0.9989	0.9999	0.9993

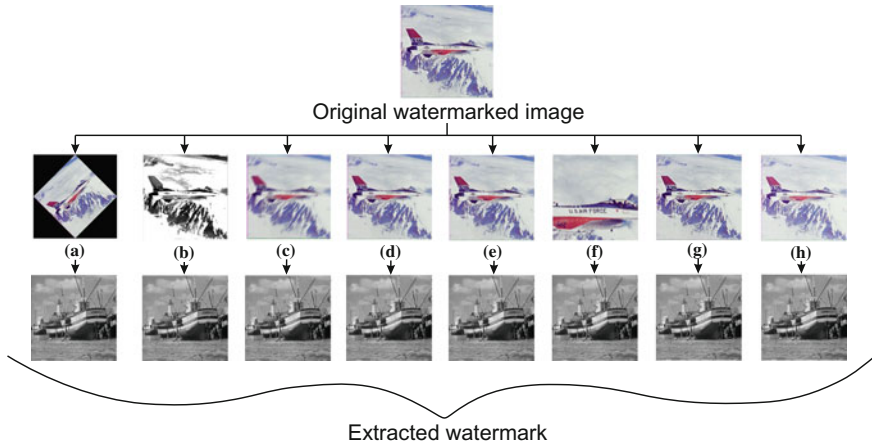


Fig. 5 **a** Rotated, **b** Histogram Equalised, **c** Motion Blurred, **d** Guassian filtered, **e** Average filtered, **f** Cropped, **g** Sharpened and **h** Gamma Intensity Corrected watermarked images

Table 5 Comparison of PSNR and correlation coefficients by proposed method with the DWT-DFT-SVD method by [14] for Airplane (Host) and Boat (watermark) image

Method	Channel	PSNR (dB)	Correlation coefficients		
			Rot	Hist	Sharp
As in [14]	Red	36.21	0.9905	0.9963	0.9889
	Green	33.06	0.9887	0.9970	0.9901
	Blue	37.33	0.9916	0.9924	0.9884
Proposed	Red	97.96	0.9998	0.9999	0.9996
	Green	95.95	0.9998	0.9999	0.9995
	Blue	99.56	0.9999	0.9999	0.9991

The PSNR and the correlation values obtained by proposed method is compared with DWT-DFT-SVD for watermarking in three different channels. Table 5 shows the comparison of PSNR and the correlation values with the method implemented by Ref. [14] and with the proposed method. The values obtained are better since the singular values of HH sub-band obtained from SWT is used instead of using singular values of HL sub-band obtained from DWT-DFT.

6 Conclusion

In this paper, the digital watermarking technique proposed uses combination of SWT+SVD which has been applied for nine different combinations of host and watermark images. The proposed technique was compared to the technique used

in Ref. [14] and the PSNR values obtained are superior. The main advantage of the SWT is that no downsampling is done after filtration which is not in case of DWT and the lack of shift invariance due to DWT is overcome by using SWT. Therefore increased PSNR values provide visually better quality watermarked images. Also in comparison with Ref. [14], the correlation coefficients obtained after applying for various kinds of attacks like rotate, crop, sharp etc are better. Therefore it is more robust and has more capability to withstand attacks that can cause information loss which is hidden inside the carrier and hence providing better security. Further this technique can be implemented for video watermarking and for embedding color watermark images.

References

1. Manjunatha Prasad, R., Shivaprakash Kollwad.: A Comprehensive Survey of Contemporary Researches in Watermarking for Copyright Protection of Digital Images. *International Journal of Computer Science and Network Security*, vol. 9, no. 4 (2009)
2. Shivani Khurana.: Watermarking and Information-Hiding. *International Journal of Computer and Information Technology*, vol. 2, pp. 1679–1681 (2011)
3. Zunera Jalil, Anwar M. Mirza and Maria Sabir.: Content based Zero-Watermarking Algorithm for Authentication of Text Documents. *International Journal of Computer and Information Technology*, ISSN 1947-5500, vol. 7, no. 2 (2010)
4. Mohammad Ali Nematollahi, S.A.R. Al-Haddad, Faraneh Zarafshan.: Blind digital speech watermarking based on Eigen-value quantization in DWT. *Journal of King University-Computer and Information Science*, vol. 27, 58–67 (2015)
5. Gaurav Bhatnagar, Balasubramanian Raman.: Wavelet packet transform-based robust video watermarking technique. *Indian Academy of science*, vol. 37, pp. 371–388 (2012)
6. Nidaa A. Abbas.: Image watermark detection techniques using quadrees. *Applied Computing and Informatics*, vol. 11, 102–115 (2015)
7. D.Vaishnavi, T.S. Subashini.: Robust and Invisible Image Watermarking in RGB Color space using SVD. *International Conference on Information and Communication Technologies (ICICT)*, pp. 1770–1777 (2014)
8. Srdjan Stankovic, Igor Djurovic, Ioannis Pitas.: Watermarking in the Space/Spatial Frequency Domain using Two-Dimensional Radon-Wigner Distribution. *IEEE transactions on Image Processing*, vol. 10, no. 4 (2001)
9. Rafael Gonzalez, Richard Woods.: *Digital Image processing*. Prentice Hall, Third edition (2008)
10. N. Nikolaidis, I. Pitas.: Robust image watermarking in spatial domain. *Signal processing*. vol. 66, pp. 385–403 (1998)
11. Bhupendra Ram.: Digital Image Watermarking Technique Using Discrete Wavelet Transform And Discrete Cosine transform. *International Journal of Advancements in Research and Technology*, ISSN 2278-7763, vol. 2, Issue 4 (2013)
12. Navnath S. Narwade, Narendra P. Deshmane, Pankaj Elchatwar, Pooja L.Pande.: Robust Watermarking for Geometric attack using DFT. *International Journal of Emerging Trends and Technology in Computer Science*, vol. 2, Issue 2 (2013)
13. O. Jane, E. Elbasi, H. G. ilk.: Hybrid Non-Blind Watermarking Based on DWT and SVD. *Journal of Applied research and Technology*, vol. 12, pp. 750–761 (2014)
14. Rahim Ansari, Murtyunjaya M, Devanalamath, K Manikantan, S Ramachandran.: Robust Digital Image Watermarking Algorithm in DWT-DFT-SVD Domain for Color Image. *International Conference on Communication, Information and Computing Technology* (2012)

15. B Siva Kumar, S Nagaraj.: Discrete and Stationary Wavelet Decomposition for IMAGE Resolution Enhancement. International Journal of Engineering Trends and Technology (IJETT), vol. 4, Issue 7 (2013)
16. Gaurav Bhatnagar, Q. M. Jonathan Wu, Balasubramanian Raman.: Biometric Template Security based on Watermarking. Procedia Computer Science, vol. 2, pp. 227–235 (2010)
17. Matlab, <http://www.mathworks.com>

An Improved Histogram Bin Shifting Based Reversible Data Hiding of Color Images

Smita Agrawal and Manoj Kumar

Abstract Reversible data hiding technique recovers the original image bit by bit after extracting the watermark bits from the watermarked image. In this paper, a novel reversible data hiding technique is proposed for color images. The proposed technique is based on histogram bin shifting technique, which is an efficient and widely used method for embedding the watermark in gray scale images in reversible manner. The proposed scheme extends the concept of histogram modification on RGB color images. However, as compared to basic histogram modification technique (used for reversible data hiding) which utilizes zero point and peak point, proposed scheme utilizes first peak point and second peak point so that the distortion generated due to the pixels shifting can be minimized. The successful application of proposed scheme on several standard color test images and experimental results in terms of higher PSNR values demonstrate the effectiveness of the proposed scheme.

Keywords Reversible data hiding · Color image · Histogram bin shifting · PSNR · First peak point · Second peak point

1 Introduction

In today's scenario of growing technologies, there is always a risk of copying, tampering, illegal access and modification of multimedia data such as images, video and audio. Digital Watermarking provides a very effective way to deal with these type of risks. Digital watermarking is a way of hiding some secret information i.e. watermark in multimedia data in such a way that distortion in watermarked media after embedding of watermark is perceptually negligible. In simple watermarking, some

S. Agrawal (✉) · M. Kumar
Babasaheb Bhimrao Ambedkar University, Vidya Vihar, Raibareli Road,
Lucknow, Uttar Pradesh, India
e-mail: smita.bbau@gmail.com

M. Kumar
e-mail: mkjnuiitr@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_3

25

information is always lost as distortion is caused after embedding of secret information. However, there are some fields, that are sensitive to information loss such as defence, artwork etc., where even a single bit information loss is not acceptable. Reversible data hiding is a special subset of fragile watermarking which not only extracts the embedded data from the watermarked content but also recovers the original cover media bit by bit. In reversible data hiding, there is no loss of information therefore, it is also called “lossless data hiding”. It is a subset of fragile watermarking where even a slightest modification or alteration in watermarked media alters or destroys the watermark and hence, the original media and watermark bits cannot be extracted. Hence, reversible watermarking is generally used in the applications where data authentication is important. Reversible watermarking techniques cannot be tested against various signal processing attacks [1].

During last few years, reversible data hiding has emerged as an interesting research area among researchers. Various researchers have proposed several techniques [1–12] for reversible data hiding after Barton [2] introduced the concept in 1997. In reversible watermarking algorithms, the main emphasis is on improving the embedding capacity. In the process of doing so, it is also important to maintain or improve the visual quality of the watermarked content. Existing reversible data hiding schemes can be classified generally into three group [13]: techniques based on histogram shifting [4], difference expansion [5, 8] and compression [7]. Amid all the schemes proposed so far, schemes using histogram modification belongs to a easy but efficient group of techniques. The principal concept in using the histogram modification based scheme is the shifting of pixels between peak point (highest occurring grayscale pixel value) and zero point (no pixel value corresponds to this value) so that the space can be created for embedding the data just adjacent to the peak point.

Techniques using histogram modification are enormously effective and computationally simple in comparison of other reversible watermarking schemes proposed for grayscale images and therefore, it is easily extendible for RGB images. RGB color image consists of three color components Red, Green and Blue and separately, these color components can be considered as grayscale images. Instead of applying it directly on color (RGB) images, proposed scheme has utilized first and second peak point in comparison of zero point and peak point used in previous existing versions [14]. Second peak point is used to minimize distortion created by pixels shifting between zero point and peak point, as there are invariably less pixels between first and second peak point in comparison to peak point and zero point [14]. Therefore, less distortion is caused in the watermarked content due to watermark embedding and the visual quality of watermarked content is improved.

In this paper, proposed modified variant of basic histogram modification scheme is applied on all three color portions i.e. R, G and B of color image separately, as an individual grayscale image and finally all watermarked color parts (R, G and B) are used for generating color watermarked image. Through this, watermarked image’s quality is improved in the terms of perceptibility at the same hiding capacity.

The organization of the paper is as follows: A brief idea of basic histogram modification scheme for gray scale images is given in Sect. 2 using algorithm given by

Ni et al. [4]. Section 3 explains the proposed scheme in detail. Experimental results are discussed on various test images in Sect. 4 and finally, conclusions are drawn in Sect. 5.

2 Related Work

In the last decade, various reversible watermarking algorithms are proposed for grayscale images. Some reversible data hiding techniques have been proposed for color images also. For example, Alattar [11] proposed a reversible data hiding technique for color images using expansion of color triplets. Another scheme is proposed in [12], in which reversible data hiding for color image is done by expanding pixel arrays which are adjacent.

First, Ni et al. [4] proposed the reversible data hiding technique utilizing the histogram of grayscale cover image for embedding the secret information. In this section, we describe basic histogram bin shifting technique for gray scale images given by Ni et al. [4].

This technique considers the histogram and utilizes the peak point and zero point for making the space so that the watermark can be embedded in the vacated space. The space for embedding the watermark is created by shifting all the pixel values between peak and zero point by 1. First, the original cover image is used to generate the histogram and then the zero and peak point of the histogram are calculated. Assuming that the zero point is always greater than the peak point, original image is scanned in a certain predefined order and the space is vacated adjacent to the peak point by incrementing the values of all pixels between peak and zero point by 1. Original image is scanned once again. If peak point is equal to the scanned value and if the watermark bit that has to be embedded is found to be 1, “1” is added to the scanned pixel otherwise the value of scanned pixel is left without any change. Resultant image is the final watermarked image.

In Fig. 1a–c, the steps of basic histogram modification technique [4] are shown for the embedding procedure. Histogram of original Lena image is shown in Fig. 1a. In Fig. 1b, histogram of Lena after shifting is shown. The space adjacent to the peak

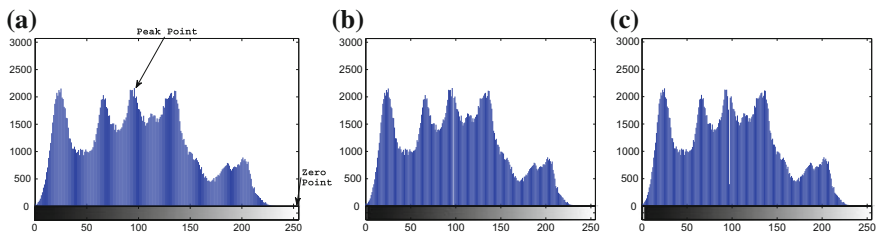


Fig. 1 Histogram modification scheme given by Ni et al. [4]. **a** Original Lena image histogram. **b** Shifted histogram of Lena image. **c** Lena image histogram after embedding of watermark

point is clearly visible in Fig. 1b. Space created by shifting the histogram (Fig. 1b) is used to embed the watermark and histogram of Lena after watermarking is shown in Fig. 1c.

Watermarked image is scanned in the similar sequence as above to extract the watermark and recover the original image. If scanned value is 1 greater than the value of peak point, extract “1” as embedded data otherwise if scanned value is found to be equal to the value of peak point, extract “0” as embedded data. Using this procedure, all watermark bits are extracted. Once again, image is scanned and 1 is subtracted from all scanned values z , where $z \in (\text{peak point}, \text{zero point}]$. At the end, recovered image is the original one.

3 Proposed Scheme

Proposed reversible watermarking technique for color images is explained in this section, which uses the modified and improved version of basic histogram bin shifting technique [4] for grayscale images.

3.1 Embedding Procedure

First, given RGB color image is divided into three color component R, G and B. Histogram of all these components are generated and first and second peak points (value just less than peak point) of all three histograms are recorded. The main aim behind using the peak point is to find the space to embed as much data as possible. Second peak point is used here so that the distortion caused due to the pixels shifting can be minimized to make space for embedding the watermark. Here, total capacity is equivalent to the total pixels that correspond to the highest pixel values in all three color components. Following embedding procedure is applied to all three color components separately:

The original image is scanned in a specific order. If the second peak point value is greater than the value of first peak point, all pixel values z , such that $z \in (\text{first peak point}, \text{second peak point})$, are incremented by “1” otherwise decremented by “1”, vacating the space just adjacent to the peak point. Again, the whole image is scanned and if scanned value is found equal to “*first peak point*”, check the watermark bit which has to be embedded. If watermark bit is equal to “1”, corresponding pixel value is incremented (if second peak point is greater than first peak point) or decremented (if first peak point is greater than second peak point) by “1”, otherwise there is no change in pixel value. Embedding is finished in this manner and generated image is called the watermarked image. This procedure is applied to all three individual grayscale images generated by dividing the RGB color image into three color components. Finally, all three watermarked color components are combined to obtain final color watermarked image.

3.2 Extraction Procedure

Extraction procedure is just the reversal of the embedding procedure. First, color watermarked image is divided into three color components R, G and B and then extraction procedure is applied to all three individual grayscale watermarked images obtained after dividing the color image into three components. Watermarked image is scanned in the similar order as done at the time of embedding. If the value of first peak point is less than the value of second peak point and scanned pixel is found having grayscale value “*first peak point + 1*”, “1” is extracted as embedded data otherwise if it is equal to the “*first peak point*”, “0” is extracted as embedded data. Image is scanned once again and if pixel value $z \in (\textit{first peak point}, \textit{second peak point})$, z is decremented by 1.

If the value of first peak point is greater than second peak point and scanned value is found to be “*first peak point - 1*”, “1” is extracted as embedded data otherwise if it is equal to the “*first peak point*”, “0” is extracted as embedded data. Whole image is scanned once again and pixel value z is incremented by 1 if $z \in (\textit{first peak point}, \textit{second peak point})$. Embedded bits are extracted in this manner and original image is also recovered. Finally all three recovered color components are concatenated to obtain the original color image.

4 Experimental Results

The proposed scheme is implemented on several color test images displayed in Fig. 2a–f. Proposed scheme has also been compared with an existing technique [4] (when applied to color images). Peak-Signal-to-Noise-Ratio (PSNR) is generally used as a metric for calculating distortion generated in watermarked content due to watermark embedding. High values of PSNR indicate better quality in terms of perceptibility. PSNR can be calculated by using the following formula:

$$PSNR = 10 * \log_{10} \left(\frac{255 \times 255}{MSE} \right)$$

where, Mean Square Error (MSE) is defined as-

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [I'(i,j) - I''(i,j)]^2.$$

where, I' is the original cover image of size $M \times N$ and I'' is the image obtained after watermark embedding.

Table 1 shows the PSNR values of various standard color test images (Fig. 2a–f) for proposed scheme and the existing scheme. Proposed scheme provides higher PSNR values as compared to the existing scheme [4] and thus indicates that the

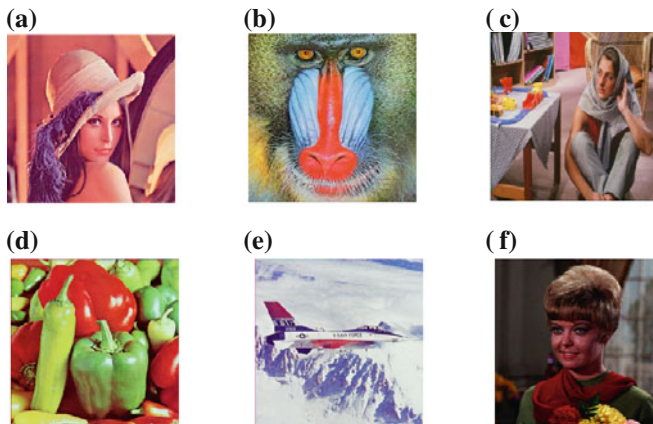


Fig. 2 Test images. **a** Lena. **b** Mandrill. **c** Barbara. **d** Peppers. **e** Airplane. **f** Girl

Table 1 PSNR values for several standard color test images (Fig. 2a–f) using proposed and existing technique

S.no.	Images	Capacity (no. of bits)	PSNR (Proposed scheme) (db)	PSNR (Existing scheme [4]) (db)
1.	Lena	9890	54.87	53.05
2.	Mandrill	6570	55.12	52.36
3.	Barbara	8353	57.21	50.43
4.	Peppers	33707	50.01	48.15
5.	Airplane	26217	60.79	59.28
6.	Girl	22501	56.23	54.10

proposed scheme can achieve better quality in terms of perceptibility. This demonstrates the strength of the proposed scheme. Higher PSNR values are achieved for images where less shifting is required between first and second peak point. Therefore distortion caused is perceptually negligible in the images after watermark embedding and the quality of the watermarked content is better at the same hiding capacity.

5 Conclusions

Histogram shifting is a widely exploited technique for grayscale images to embed the watermark in reversible way. An improved variant of existing histogram modification based reversible data hiding technique is proposed in this paper for color images. In the proposed scheme, first and second peak point are employed instead of zero and peak point. All the three color parts of the color images are used to embed the watermark separately. Rather than using zero point, the proposed scheme uses second

peak point, which helps in minimizing distortion as generally there are less pixel values between first and second peak point as compared to the pixels shifted between zero and peak point. Experimental results demonstrate that the proposed scheme achieves better visual quality as PSNR values for proposed scheme are higher in comparison to the basic histogram modification technique (when applied to RGB color images).

References

1. Leung, H., Cheng, L.M., Liu, F., Fu, Q.K.: Adaptive Reversible Data Hiding based on Block Median Preservation and Modification of Prediction Errors. *The Journal of Systems and Software*. 86, 2204–2219, (2013)
2. Barton, J. M.: Method and Apparatus for Embedding Authentication Information Within digital Data. U.S. Patent No. 5646997. (1997)
3. Honsinger, C.W., Jones, P.W., Rabbani, M., Stoffel, J.C.: Lossless recovery of an original image containing embedded data. U.S. Patent No.6278791. (2001)
4. Ni, Z., Shi, Y., Ansari, N., Su, W.: Reversible Data Hiding. *IEEE Transactions on Circuits and Systems for Video Technology*. 16(3), 354–362 (2006)
5. Tian, J.: Reversible Data Embedding Using a Difference Expansion. *IEEE Transactions on Circuits and Systems for Video Technology*. 13, 890–896 (2003)
6. Yang, B., Schucker, M., Funk, W., Busch, C., Sun, S.: Integer - DCT based Reversible Watermarking Technique for Images Using Companding Technique. In: *Proceedings SPIE*, 5306, 405–415, (2007)
7. Sharma, C., Tekalp, Saber: Lossless Generalized-LSB Data Embedding. *IEEE Transactions on Image Processing*. 14(2), (2005)
8. Thodi, D.M., Rodriguez, J.J.: Expansion Embedding Techniques for Reversible Watermarking. *IEEE Transactions on Image Processing*. 16(3), 721–730, (2007)
9. Kim, K. S., Lee, M. J., Lee, H. Y., Lee, H. K.: Reversible Data Hiding Exploiting Spatial Correlation between Sub-Sampled Images. *Pattern Recognition*. 42(11), 3083–3096, (2009)
10. Luo, L., Chen, Z., Chen, M., Xeng, X., Xiong, Z.: Reversible Image Watermarking Using Interpolation technique. *IEEE Transactions on Information Forensics and Security*. 5(1), 187–193, (2010)
11. Alattar, A.M.: Reversible watermark using difference expansion of triplets. In *Proceedings of IEEE International Conference on Image Processing (ICIP 2003)*, Spain. 501–504, (2003)
12. Alattar, A. M.: Reversible watermark using the difference expansion of a generalized integer transform. *IEEE Transaction of Image Processing*. 13(8), 1147–1156, (2004)
13. Feng, J.B., Lin, I.C., Tsai, C.S., Chu, Y.P.: Reversible Watermarking: Current Status and Key Issues. *International Journal of Network Security*. 2(3), 161–171, (2006)
14. Agrawal, S., Kumar, M.: An Improved Reversible Data Hiding Technique based on Histogram Bin Shifting. *Smart Innovation, Systems and Technologies (Springer Verlag)*. 43, 239–248, (2015)

Face Recognition Using Background Removal Based on Eccentricity and Area Using YCbCr and HSV Color Models

Amith Lawrence, N.V. Manoj Ashwin and K. Manikantan

Abstract The process of Face Recognition is complicated due to the background, pose variations in the images. Using the Pre-processing techniques proposed in this paper the essential invariant features in an image have been made available for extraction. *Background Removal based on Eccentricity* is implemented by incorporating both *YCbCr and HSV* color models to eliminate unnecessary features in the background. Multi-scaled fusion is included for nullifying the variation in pose. Next, the images are subjected to feature extraction using two-dimensional Discrete Wavelet Transform (DWT) and feature selection algorithm. Experimental results show the effectiveness of the above mentioned techniques for face recognition on two benchmark face databases, namely, CMU-PIE and Caltech.

Keywords Face recognition • Image pre-processing • Feature extraction • Feature selection

1 Introduction

The concept of Face Recognition has been at the centre of Image Processing more specifically in the field of biometrics for many years. The applications span over a variety of different fields i.e. Surveillance and Tracking, Criminal Identification, Authentication of Identity. It involves identifying the distinguishing features of a face and classifying them accordingly [1]. The success of this FR system is affected by various factors such as Background, Pose and Illumination which can be

A. Lawrence · N.V. Manoj Ashwin · K. Manikantan (✉)
Department of Electronics and Communication Engineering,
M S Ramaiah Institute of Technology, Bangalore 560 054, India
e-mail: kmanikantan2009@gmail.com

A. Lawrence
e-mail: amithlawrence94@gmail.com

N.V. Manoj Ashwin
e-mail: manojashwin.ediga@gmail.com

effectively nullified by Pre-processing techniques. The abilities of the human brain have been recreated using these techniques [2]. Images with a plain predefined static background removing the background will always give you the face boundaries. For color images, the image is subjected to skin segmentation [3] and contour of the face is obtained. The following sections deal with the techniques already available for executing the FR system, proposed pre-processing technique to improve our recognition rates, Experimentation results for Datasets (CMU-PIE and Caltech) and lastly conclusion and future work.

2 Problem Definition and Contribution

The images used for testing the proposed techniques may include unco-operative subjects with varying background. The background tends to affect the efficiency of the system. The background features need to be removed from the images in order to relieve the system of any unnecessary computation.

The face recognition process is optimized isolating only the face to be recognized while removing the background which is non-essential. This Background removal is a combination of skin segmentation, morphological operations, eccentricity—range based region selection. This method can be enhanced by introducing another color model namely HSV [4] during the skin segmentation stage, for extracting the face region.

In addition to the background issue there is the matter of pose variance which can also adversely affect the recognition ability of the system. The differences in pose can be neutralised by fusing the left and right poses of the face into a single image.

3 Prior Art

This paper has improved the performance of the FR system by incorporating the techniques mentioned below :

3.1 *Multi-scaled Image Fusion to Combat Varying Pose*

In this step the face is normalised in terms of pose by combining it with its mirrored image and compressing it as shown in Fig. 1. Thereby establishing symmetry along the vertical axis. Lengthwise compression and fusion of images with different ratios reduces the redundancy in images (due to vertical symmetry) [5]. This technique reduces the effect of pose variance in images and improves the correlation between images of the same subject in different profiles.

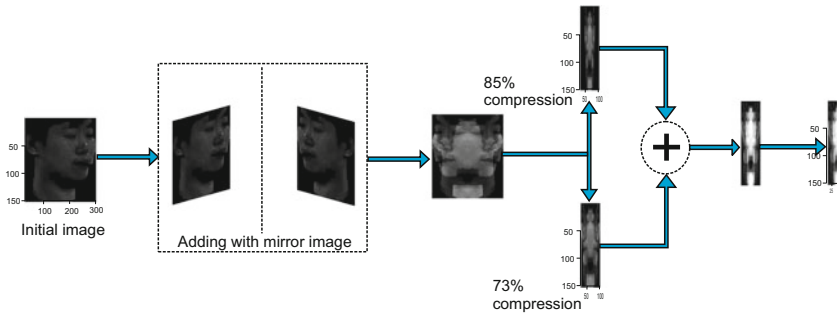


Fig. 1 Multi-scaled fusion process flow

3.2 Discrete Wavelet Transform for Feature Extraction

The Discrete Wavelet Transform (DWT) [6] is used in feature extraction because it can produce both frequency and spatial representations of a signal simultaneously. Here 2D DWT for feature extraction. Thus, the image is sampled into subbands approximation (cA), horizontal (cH), vertical (cV) and diagonal (cD). The information in low spatial frequency bands play a dominant role in face recognition [7]. The cA sub-bands facial features are least sensitive to external parameter variations.

3.3 Binary Particle Swarm Optimization for Feature Selection

This is a simple end efficient population-based optimization method proposed by Kennedy and Eberhart. Particle Swarm Optimization is inspired by the social foraging behavior of some animals such as flocking behavior of birds and the schooling behavior of fish, all examples of swarming [8]. In PSO, potential solutions are called particles. In PSO, the global best particle found among the swarm is the only information shared among particles. It is a one-way information sharing mechanism.

$$X_i^t = [x_{i1}, x_{i2}, x_{i3}, \dots, x_{iD}] \quad (1)$$

$$V_i^{(t+1)} = w \cdot V_i^t + C_1 \cdot \text{rand1} \cdot (P_{\text{best}_i} - X_i^t) + C_2 \cdot \text{rand2} \cdot (G_{\text{best}_i} - X_i^t) \quad (2)$$

Where $i = (1, 2, \dots, N)$ and N is the size of the swarm; P_{i_best} is the particle best reached solution and G_{best} is the global best solution in the swarm. C_1 and C_2 are cognitive and social parameters that are bounded between 0 and 2. rand_1 and rand_2 are two random numbers, with uniform distribution $U(0,1)$. (V_{max} is the maximum velocity).

The Binary Particle Swarm Optimization (BPSO) [8], is an evolutionary optimizer which maximizes the class separation by maximizing the fitness function and in turn selecting the most distinctive features. A bit value of 1 indicates selection and a value of 0 indicates rejection of the feature. The particle position is updated using

$$\text{If } rand_3 < 1/(1 + \exp(-v_i^{t+1})), X_i^{t+1} = 1 \text{ else } X_i^{t+1} = 0 \quad (3)$$

4 Proposed Methodology

The FR system is based on the system being divided into 2 stages i.e., training and testing stages, as shown in Fig. 2. The training stage deals with training the FR system using a certain amount of images from the database. Time taken to train and select the features (Training time) signifies the complexity and speed of the algorithm. After training, the system is tested using the remainder of the images from the database. The training and testing stages each consists of four general component blocks, as shown in Fig. 2.

4.1 Eccentricity and Area-Range Based Varying Background Removal for Facial Region Extraction

The images are subjected to skin segmentation in the YCbCr and HSV color space and are subjected to thresholding such that only skin colored regions are included. The HSV color model is more akin to human color perception [9]. The H and S channel provide information about skin color. Channel S is used to identify Asian and Caucasian ethnics [10]. The obtained binary image is subjected to morphological

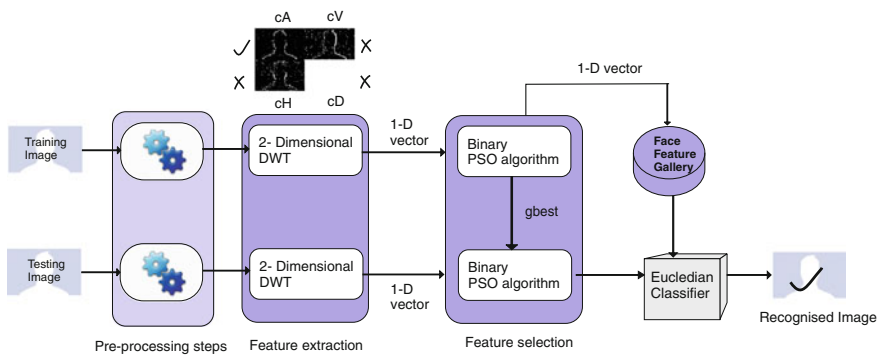


Fig. 2 Block diagram of face recognition system

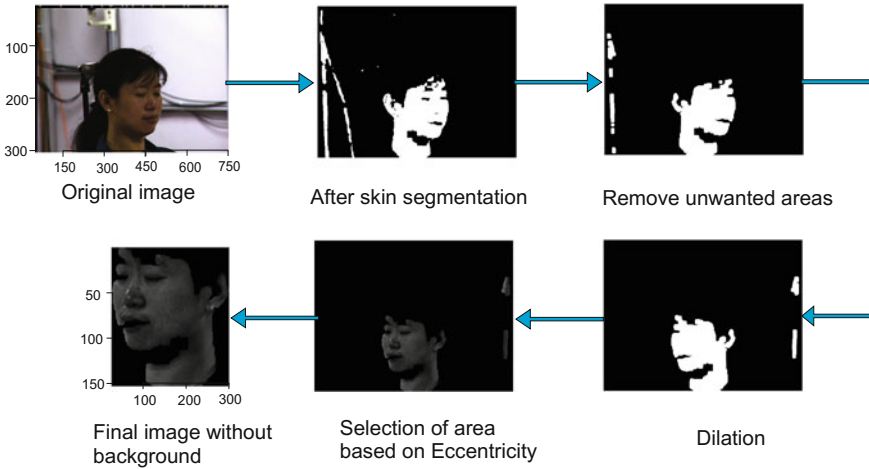


Fig. 3 Background removal process flow

operations such as opening by using a structuring element in order to remove the noise spots, which is followed by dilation and erosion. This binary image includes facial regions and other skin colored regions. Eccentricity is a scalar property which specifies the ratio of the distance between the foci of the ellipse and its major axis length [11]. The value is between 0 and 1. Considering that a general human face is ovoid in shape and that the face is in the foreground, the facial regions are extracted based on the eccentricity and the area occupied by that region as shown in Fig. 3. These are the steps involved:

- (1) The color images of the subject in RGB color space is converted to YCbCr and HSV color space as it is relatively easier to threshold skin color regions using this. The YCbCr color space is relatively immune to illumination changes and thus provides better separation between illumination component Y and the color components *Cb* and *Cr*. The skin in channel *H* (Hue) is characterized by values between 0.01 and 0.25, in the channel *S* (Saturation) from 0.9 to 0.103. The skin pixels are assigned a 1 and non-skin pixels are assigned a 0 using (5). Thus the image is converted to a binary image with white portions being the skin colored regions within the threshold value.

$$\begin{aligned}
 77 &\leq Cb \leq 123 \\
 133 &\leq Cr \leq 173 \\
 0.01 &\leq H \leq 0.25 \\
 0.9 &< S < 0.103
 \end{aligned}
 \tag{4}$$

- (2) The skin color segmentation is succeeded by a series of area opening operations, erosion and dilation to remove the unwanted sputtering and to round the edges of the segmented regions [12].

- (3) Regions having higher values of eccentricities are removed from the image. The regions with higher values of eccentricities usually correspond to rectangular long areas. The facial areas correspond to areas with low values of eccentricities (usually between 0.4 and 0.6). In the event that multiple regions are identified then they are selected based on their area.

5 Feature Extraction Stage

The extraction stage deals with reducing the dimensions of the target thereby reducing and eliminating the unimportant aspects of the image. In this paper, we have used one dimensional DWT to achieve this. For which the cA component is used for further extraction as shown in Fig. 2. 2D DWT is applied to this and the cA sub-band is extracted. This is done up to 2 levels using the sym4 wavelet. DWT filters out the number of features required to effectively represent an image and thus used as an extractor in this system.

6 Feature Selection Stage

The BPSO algorithm is used for feature selection. This optimises the set of features by acting as funnel to reduce the quantity of features [13], such that the class separation is maximized. This minimizes the number of redundant features to give a compact representation of the image.

7 Euclidean Distance Classifier

The Euclidean classifier is used to ascertain the similarities between the testing image and training images and thus, in the process recognize the subject in the testing image [13]. The Euclidean distance termed as the straight line distance between two images is a measure of this similarity. If A and B are feature vectors corresponding to a training image and a testing image respectively, and L is the length of the feature vector, then the Euclidean distance is calculated.

8 Experimental Results and Discussions

The background removal technique proposed in the paper, is applied to a CMU-PIE database, the results for a particular subject are as shown in Fig. 4.



Fig. 4 13 images of a subject in CMU-PIE database (*left*) and the corresponding images after background removal (*right*)

8.1 Database Details

In order to verify the effectiveness and robustness of the proposed technique, the following experiments were conducted using standard face databases, namely, CMU-PIE and Caltech.

8.1.1 CMU-PIE Database

The CMU Pose, Illumination and Expression (PIE) database [14] contains more than 40,000 images of 68 subjects taken between October 2000 and December 2000. It contains images with highly complex backgrounds and non-uniform lighting conditions these images adopt different poses as well. For the experiments related to CMU-PIE, a database which contains images of 30 subjects with pose variations captured before complex backgrounds has been utilised. In order to simulate the pose variance factor, a total of 13 images from each subject with a size of 640×486 (RGB) have been considered. The images differ from one another in pose and background.

8.1.2 Caltech Database

The Caltech Frontal face dataset [15] contains a total 450 face images of 27 unique subjects each of size 896×592 , in JPEG format. It consists of images with different lighting conditions and non-uniform backgrounds. All the images are frontal with a constant pose in all the images. This database has been customised by considering a total of 16 subjects having 15 images each with varying expressions. The backgrounds for each of these images is complex and the background removal process has been adjusted accordingly. This database has been created for experimental purposes from the initial set of images. The creation of this database takes a total time of 62s.

Table 1 Comparison of recognition rate for different Tr:Te ratio for the proposed datasets

Dataset	Tr:te	RR (%)	Avg. No. of features	Training time (s)	Testing time (ms)
CMU-PIE	5 : 8	32.04	395	75.36	118.34
	7 : 6	34.97	393	64.56	121.45
	3 : 10	26.23	387	70.37	103.23
	1 : 12	21.65	402	76.97	100.64
Caltech faces 1999	5 : 10	50.20	386	46.62	32.45
	6 : 9	54.16	391	55.36	35.62
	7 : 8	58.73	345	64.79	36.80
	8 : 7	58.63	358	74.37	38.67

8.2 Experimentation

8.2.1 Experiment 1

In real world applications, the training to testing ratio varies depending on the situation, we try to simulate this scenario by varying the training to testing ratio. The experiment is carried out for different ratios as shown in Table 1. For a given ratio, the RR varies slightly with different iterations as the training and testing images are chosen in a pseudo-random nature [16]. Thus, the results for each of the databases mentioned in this paper, are computed by averaging over 10 iterations. In this experiment, as the ratio of training to testing images increases, the RR and the average training time increases. In addition, testing time also increases because the image has to be compared with more number of training images.

In the Caltech database, it can be observed that the RR seems to vary from 58 to 50%. In this case, the best and worst subjects. The results associated with the CMU-PIE dataset is similar. The system performs well under most constraints and the average RR varies between 34 to 28%. The CMU-PIE dataset consists of images that are not at face level also and therefore getting high RR is a challenge.

8.2.2 Experiment 2

In this experiment as in Table 2, the results by implementing the proposed background removal method using both color models in conjunction (1B & 2B) against background removal in YCbCr (1A & 2A). The method to cancel pose variation is also applied to the system in each of these scenarios and the results are tabulated in Table 3. In the CMU-PIE database the RR is 32.04 which is 4% greater than the RR when the dataset is subjected to background removal in YCbCr in combination with multi-scaled fusion technique. Whereas the Caltech database, the RR improves nearly 8% when the background is removed using both color models against only

Table 2 IDs for each scenario

Database	ID	Pre-processing technique
CMU-PIE	1A	Background removal in YCbCr with scaled-fusion technique
	1B	Background removal in YCbCr and HSV with scaled-fusion technique
	1C	Only background removal in YCbCr and HSV
Caltech	2A	Background removal in YCbCr with scaled-fusion technique
	2B	Background removal in YCbCr and HSV with scaled-fusion technique
	2C	Only background removal in YCbCr and HSV

Table 3 Comparison of recognition rate for different scenarios

ID	RR (%)	Avg. no.of features	Training time (s)	Testing time (ms)
1A	28.29	349	73.02	112.66
1B	32.04	395	75.36	118.34
1C	25.09	381	63.25	102.58
2A	42.36	342	45.58	31.56
2B	50.20	386	46.42	32.45
2C	47.52	344	38.29	27.31

YCbCr method. The images are pose invariant and therefore multi-scaled fusion technique does not effect the RR significantly. Thus the proposed method enhances the recognition efficiency by implementing the proposed pre-processing technique.

9 Conclusions

The proposed pre-processing method utilises two color models in order to zero in on the face, and reduce the image to obtain the features which improves the chances of recognition. The processed images are in coupled with DWT based feature extraction, BPSO based feature selection and Euclidean Classifier, using MATLAB, to achieve improved face recognition. The accuracy of this system is realised by the results obtained by applying the proposed FR system on two different databases, which have variations in pose (CMU-PIE) and background (Caltech) as shown in Fig. 5. The results for CMU-PIE database indicate that the improved background removal along with the pose remedial method has given a recognition rate of 32.04 % with training-to-testing ratio of 5 : 8. Considering that the Caltech database has varying background but no pose variation, the results are mainly influenced by background removal method with a recognition rate of 50.2 % with training-to-testing ratio of 5 : 10. Therefore the improved background removal technique has effectively increased the recognition efficiency.

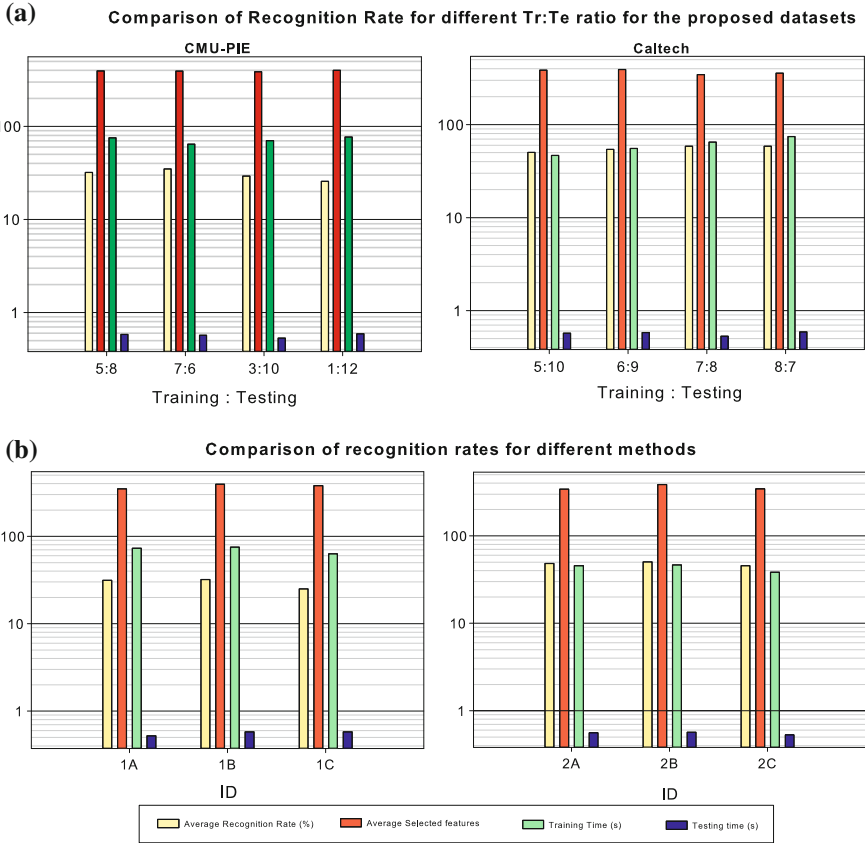


Fig. 5 Graphical plots describing Experiment 1(a) and Experiment 2(b). **a** Comparison of recognition rate for different Tr:Te ratio for the proposed datasets. **b** Comparison of recognition rates for different methods

References

1. John D Woodward Jr, Biometrics: The Ultimate Reference, 2009 edition John Wiley and Sons, 2009.
2. G Gordon, T Darrell, M Harville, J Woodfill, Background Estimation and Removal Based on Range and Color,” International Conference on Computer Vision and Pattern Recognition, 1999.
3. Son Lam Phung, Abdesslam Bouzerdoum, Douglas Chai, “Skin Segmentation Using Color Pixel Classification: Analysis and Comparison,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, 2005.
4. Daithankar M V, Karande K J, Harale A D, “Analysis of skin color models for face detection,” *International Conference on Communications and Signal Processing*, 2014.
5. Nitish S Prabhu, Thejas N Kesari, K Manikantan, S Ramachandran, “Face Recognition using Eccentricity-Range based Background Removal and Multi-Scaled Fusion as Pre-processing Techniques Subsequences,” *Nirma International University Conference on Engineering, NUiCONE*, 2013.

6. R C Gonzalez, R E Woods, Digital Image Processing, 3rd Edition, NJ Prentice-Hall, 2008.
7. Roohi S, Sachin D, K Manikantan, S Ramachandran, "Feature Accentuation using Uniform Morphological Correction as Pre-processing Technique for DWT based Face Recognition," *International Conference on Computational Intelligence: Modeling, Techniques and Applications*, 2013.
8. J Kennedy, R Eberhart, "Particle Swarm optimization," *IEEE International Conference on Neural Networks*, 1995.
9. Face Detection Techniques, <http://www.facedetection.com/techniques/>
10. V A Oliveira, A Conci, "Skin Detection using HSV color space," *International Journal of Computer Trends and Technology*, 2009.
11. Matlab, <http://in.mathworks.com/help/images/ref/regionprops.html>
12. R M Ramadan, R F Abdel-Kader, "Face Recognition Using Particle Swarm Algorithm," *IEEE International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2009.
13. Li Bai, Yihui Liu, "A Novel Face recognition method," *Artificial Intelligence, Automated Reasoning, and Symbolic Computation*, 2002.
14. Robotics Institute, Carnegie Mellon University - Pie Database. www.ri.cmu.edu/research_project_detail.html?project_id&menu_id
15. Computational Vision, Caltech Faces 1999. www.vision.caltech.edu/Image_Datasets/faces/faces.tar
16. R Brunelli, T Poggio, "Face Recognition: Features versus Templates," *Pattern Analysis and Machine Intelligence*, vol.15, no.10, 1993.

An Efficient Multi-focus Image Fusion Approach Based on DWT

Sonam and Manoj Kumar

Abstract In this paper, a Discrete Wavelet Transform (DWT) based approach for multi-focus image fusion using a novel coefficients selection scheme is proposed. The proposed method enables the decomposition of source images into low and high frequency sub-bands. To combine the low and high frequency sub-band coefficients of the transformed images, pixel averaging method and gradient based fusion rule are applied to the various frequency sub-bands. Finally, we applied inverse DWT and obtained an enhanced fused image. The performance of the proposed scheme is evaluated on various multi-focus images and experimental results are also compared with existing methods. Experimental results demonstrate that the proposed scheme is better and effective in terms of Peak-Signal-to-Noise-Ratio (PSNR).

Keywords Multi-focus image fusion · Image sharpness · DWT · PSNR

1 Introduction

In current scenario, image fusion has become a vital and useful technique for image analysis and computer vision [1]. The purpose of image fusion is to integrate the information of two or more images of same view or scene into a single composite image and to obtain more informative fused image. There are various application of image fusion in the field of medical imaging, security, remote sensing, military, robotics, etc. Generally, single image of a complex scene does not contain enough information to observe, means if one image is focused in object and another one is in out-of-focus. This problem can be solved with multi-focus image fusion. In which, the relevant information from multiple source images of same scene with

Sonam (✉) · M. Kumar

Department of Computer Science, Babasaheb Bhimrao Ambedkar University,
Lucknow, India

e-mail: sonam870115@gmail.com

M. Kumar

e-mail: mkjnuiitr@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_5

45

different focuses are combined into a single focusing sharp output image. This output image contain more information and having better visual appearance in comparison to any of the source image. The resultant fused image is also useful for the perception of human/machine. Fusion techniques can be performed into spatial and transform domain [2]. In spatial domain techniques, fusion process is performed on all source images pixel by pixel and a single fused image is produced. Some spatial domain techniques of fusion are maximum method, minimum method, average method, Principal Component Analysis (PCA) method [3, 4], Intensity-Hue-Saturation (IHS) method, Brovey Transform and High Pass Filter (HPF) [5, 6]. Generally, the spatial domain methods produce distorted fused images. To overcome this problem a transform domain approach is introduced. Transform domain approaches include discrete wavelet transform [7–9], Lifting Stationary Wavelet Transform (LSWT) [10], Contourlet Transform [11] and Nonsubsampled Contourlet Transform (NSCT) [12], etc. The image fusion process can be broadly categorized into pixel level, decision level and feature level. Among these, the pixel level fusion method is easy to implement and time efficient. The maximum method [13] is widely used technique for fusion. This method is sensitive in terms of noise and artifacts. The simplest fusion method is average method, in which pixel by pixel average values are calculated. But this method, reduce the contrast [14]. For the problem of contrast, a new fusion method had been introduced based on directive contrast [15]. The high frequency sub-bands had been fused by directive contrast, while low frequency sub-bands are fused using median method. The multiresolution fusion techniques based on wavelet transform, provides a better spatial and spectral localization information of image than other multiresolution techniques [16, 17]. A new fusion method based on DWT [18] has been introduced for multi-focus images. The low frequency sub-bands are fused using a maximum sharpness focus measure method, while high frequency sub-bands are combined by a maximum neighboring energy method. The sharp details of images are represented in high frequency coefficients. Here, we propose a fusion method in which sharp details are obtained from the high frequency coefficients with the aim to increase the sharpness of image. This paper is based on pixel level image fusion. In which, the blurring is evaluated by selecting the important information from the sharp details. The objective of this paper is to improve the quality and visual appearance of image by combining all important details of individual source images. In the proposed work, DWT based image fusion method is applied using a novel selection rule for wavelet coefficients. After performing DWT, low and high frequency coefficients are obtained and combined using different fusion methods, separately. Over the obtained new fused wavelet coefficients Inverse Discrete Wavelet Transform (IDWT) is performed to reconstruct the image. The performance of the proposed scheme is evaluated and compared with few existing fusion methods.

The remainder of this paper is organized as follows. Proposed scheme is explained in Sect. 2. Experimental results are discussed in Sect. 3. The quality of fused image evaluated in Sect. 4. Finally, concluding remarks are given in Sect. 5.

2 Proposed Scheme

To improve the quality of multi-focus images a scheme is proposed where images are of same view but with different focuses. In our proposed scheme, two multi-focus images are taken. Over the each source image DWT is performed, by which images are decomposed into low-pass (L) and high-pass (H) frequency sub-bands. The main idea of our proposed scheme is to improve the sharpness of image and reconstruct the enhanced image, by selecting the sharp details from high frequency sub-bands. For that, high frequency coefficients of both image are fused using gradient based method. The gradient coefficients of both of the transformed images are calculated using high frequency sub-bands and compared, respectively. If gradient coefficients of one image is greater than other image then pixel wise averaging method is performed otherwise maximum method is opted. Similarly, on low frequency sub-bands, simple pixel averaging method is performed. Finally, the resultant fused image is reconstructed by performing IDWT on new fused wavelet coefficients. The block diagram of proposed scheme is shown in Fig. 1. The sharpness focus measure is defined as [18].

$$\nabla G(p) = [\nabla G_x(p)^2 + \nabla G_y(p)^2]^{1/2} \tag{1}$$

where $\nabla G_x(p), \nabla G_y(p)$ can be defined as:

$$\nabla G_x(p) = \left\{ \begin{aligned} & -D(s-1, t-1, u, v) - 2D(s-1, t, u, v) - D(s-1, t+1, u, v) \\ & + D(s+1, t-1, u, v) + 2D(s+1, t, u, v) + D(s+1, t+1, u, v) \end{aligned} \right\}$$

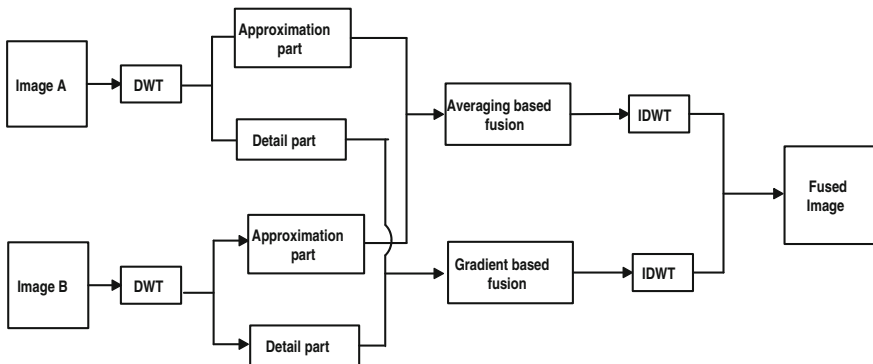


Fig. 1 Block diagram of proposed scheme

$$\nabla G_y(p) = \left\{ \begin{aligned} &D(s-1, t-1, u, v) + 2D(s, t-1, u, v) + D(s+1, t-1, u, v) \\ &- D(s-1, t+1, u, v) - 2D(s, t+1, u, v) - D(s+1, t+1, u, v) \end{aligned} \right\}$$

In this method, let $p = (s, t, u, v)$ be represent the index of a particular multiscale decomposition coefficient, where s and t represent spatial position, u and v denote decomposition level and frequency band of multi-scale decomposition representation. The proposed method can be summarized using following steps:

1. Take two multi-focus source images A and B .
2. Perform DWT to decompose the source images into low and high frequency sub-bands.
3. Apply pixel averaging fusion method on low frequency coefficients of both transformed images in approximation part.
4. Apply gradient based fusion rule on high frequency coefficients of both transformed images in detail parts.
5. Perform an inverse DWT on fused wavelet coefficients to achieve the final fused image (I_f).

3 Experimental Results and Discussion

The proposed scheme is tested on some standard images such as mandrill, lena and barbara, each of size 512×512 . These image are considered as the ideal reference image (I_r) shown in Figs. 2a, 3a and 4a. In the proposed work, DWT based method with Haar as the wavelet basis is used upto 2-level decomposition. But before applying DWT, reference images are blurred with Gaussian filter of 13×13 window and with standard deviation (σ) = 5. Figures 2b, 3b and 4b are upper side blurred images, while other Figs. 2c, 3c and 4c are lower side blurred images. The results obtained by some existing method such as minimum method [13], maximum method [19] and directive contrast method [15] are shown in Figs. 2d-f, 3d-f and 4d-f, respectively. The results obtained by our proposed method are shown in Figs. 2g, 3g and 4g. For the efficiency of our proposed work, we also compared the results of our method by above existing methods. The obtained results from these methods and proposed method are shown in Table. 1. From the experimental results, it can be easily observed that the proposed scheme is better than others.

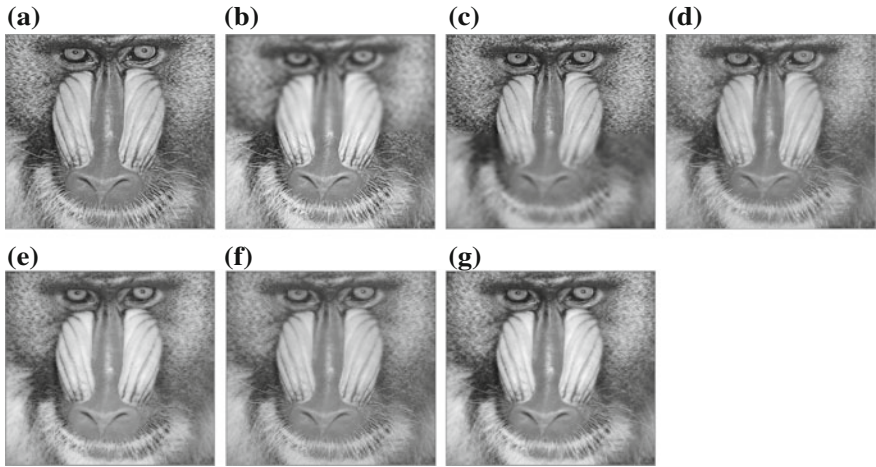


Fig. 2 Fusion results for mandrill image. **a** Reference image; **b** upper side blurred image; **c** lower side blurred image; **d** fused image by maximum method; **e** fused image by minimum method; **f** fused image by directive contrast method; **g** fused image by proposed method



Fig. 3 Fusion results for lena image. **a** Reference image; **b** upper side blurred image; **c** lower side blurred image; **d** fused image by maximum method; **e** fused image by minimum method; **f** fused image by directive contrast method; **g** fused image by proposed method



Fig. 4 Fusion results for barbara image. **a** Reference image; **b** upper side blurred image; **c** lower side blurred image; **d** fused image by maximum method; **e** fused image by minimum method; **f** fused image by directive contrast method; **g** fused image by proposed method

Table 1 PSNR of fused images

Input images	Fusion rules	PSNR w.r.t reference image
Mandrill	Minimum method	28.3171
	Maximum method	29.3383
	Directive contrast	28.6002
	Proposed method	29.8569
Lena	Minimum method	32.1539
	Maximum method	34.09539
	Directive contrast	34.2217
	Proposed method	35.7386
Barbara	Minimum method	29.9535
	Maximum method	31.3625
	Directive contrast	31.9447
	Proposed method	33.1251

4 Evaluation Factor of Fused Image

It is difficult, to discriminate the proposed scheme with others by visual inspection. Therefore, PSNR metric is used to evaluate the performance of the methods. PSNR indicates the visual quality of the fused image, while Mean Square Error (MSE) indicates how much error conveys by the fused image. The lower value of MSE and higher value of PSNR indicate better fused result.

4.1 Peak Signal to Noise Ratio (PSNR)

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (2)$$

where MSE is given as

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I_r(i,j) - I_f(i,j)]^2$$

where I_r and I_f represent the reference and fused image.

5 Conclusions

In this paper, a simple and effective DWT based fusion scheme, for constructing a single image with enhanced quality using multi-focus image fusion is proposed. These multi-focus images are of same scene but with different focusing. The main contribution of this work is to introduce a novel fusion scheme for selecting the coefficients in DWT domain to increase the sharpness. In this method, fusion of low frequency coefficients is performed using pixel averaging method, while a gradient based fusion method is opted for high frequency coefficients. The proposed method is also compared with few existing fusion methods. To evaluate the performance of our approach, PSNR values are compared with other methods. We can observed that the quality of the result from our proposed scheme is good in comparison of other fusion techniques.

References

1. Hall, D.L., Llinas, J., An introduction to multisensor data fusion. Proc. IEEE, 85 (1), 6–23, 1997.
2. Goshtasby, A.A., Nikolov S., Image fusion: Advances in the state of the art. Inf. Fusion, 8 (2), 114–118, 2007.
3. Naidu, V.P.S., Raol, J.R., Pixel-level image fusion using wavelets and principal component analysis. Defence Science Journal, 58 (3), 338–352, 2008.
4. Desale, R.P., Verma, S.V., Study and analysis of PCA, DCT and DWT based image fusion techniques. International Conference on Signal Processing Image Processing and Pattern Recognition, IEEE, 66–69, 2013.
5. Wang, Z., Ziou, D., Armenakis, C., Li, D., A comparative analysis of Image fusion methods. IEEE transactions on Geoscience and Remote Sensing, 43 (6), 1391–1402, 2005.

6. Chavez, S.J., Sides, S.C., Anderson, J.A., Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic. *Photogrammetric Engineering and Remote Sensing*, 57 (3), 295–303, 1991.
7. Pajares, G., de la Cruz, J.M., A wavelet-based image fusion tutorial. *Pattern Recogn*, 37 (9), 1855–1872, 2004.
8. Yang, Y., Park, D.S., Huang, S., Rao, N., Medical image fusion via an effective wavelet-based approach. *EURASIP Journal on Advances in Signal Processing*, 10, 1–13, 2010.
9. Li, H., Manjunath, B.S., Mitra, S.K., Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 57 (3), 235–245, 1995.
10. Chai, Y., Li, H., Li, Z., Multifocus image fusion scheme using focused region detection and multiresolution. *Optics Commun*, 284 (19), 4376–4389, 2011.
11. Do, M.N., Vetterli, M., The contourlet transform: an efficient directional multiresolution image representation. *IEEE Trans. Image Process*, 14 (12), 2091–2106, 2005.
12. da Cunha, A.L., Jianping, Z., Do, M.N., The nonsubsampling contourlet transform: theory, design and applications. *IEEE Trans. Image Process*, 15 (10), 3089–3101, 2006.
13. Wang, Z., Ma, Y., Gu, J., Multi-focus image fusion using PCNN. *Pattern Recognition*, 43 (6), 2003–2016, 2010.
14. Li, S., Yang, B., Multifocus image fusion using region segmentation and spatial frequency. *Image Vis. and Comput*, 26 (7), 971–979, 2008.
15. Bhatnagar, G., Raman, B., A new image fusion technique based on directive contrast. *Electronics letters on computer vision and image analysis*, 8 (2), 18–38, 2009.
16. Amolins, K., Zhang, Y., Dare, P., Wavelet based image fusion techniques - An introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62 (4), 249–263, 2007.
17. Chen, Y., Chen, L., Gu, H., Wang, K., Technology for multi-focus image fusion based on wavelet transform. *IEEE, Third International Workshop on Advanced Computational Intelligence*, 405–408, 2010.
18. Yong, Y., A novel DWT based multi-focus image fusion method. *International conference on advances in engineering, Procedia engineering*, 24, 177–181, 2011.
19. Tamilselvan, K.S., Murugesan, G., Survey and analysis of various image fusion techniques for clinical CT and MRI images. *Wiley Periodicals, Inc.*, 24, 193–202, 2014.

A Novel Fuzzy Filter for Mixed Impulse Gaussian Noise from Color Images

M. Jayasree and N.K. Narayanan

Abstract Removal of Mixed noise from digital color images is a challenging task because it requires processing of different types of noise. Also, noise need to be distinguished from the original image structures such as edges and details. Fuzzy theory is an effective solution to this problem. In this paper, a new fuzzy method is proposed to reduce impulse and Gaussian noise from color images. A weighted averaging filtering operation is used for this purpose. The weights in the averaging process are assigned using a fuzzy rule system based on a new certainty function, so as to reduce both noise types and to preserve image structures. Experimental results show that the method outperforms the state-of-the-art filters.

Keywords Mixed noise • Fuzzy theory • Gaussian noise • Certainty function

1 Introduction

Mixed noises degrade digital image quality to a large extend. Gaussian noise can corrupt digital color images at the image acquisition phase that may be further ruined by impulse noise due to transmission errors or storage faults. Noise can be reduced separately using various existing techniques [1–3], but only very few methods have been published to process mixed noise, and most of them have been developed for gray level processing. Consecutively applying specific filtering methods, one for each type of noise would reduce mixed noise in a simple way. But, this process could decrease the computational efficiency which implies that this

M. Jayasree (✉)

Department of Computer Science and Engineering, Government Engineering College,
Thrissur, India
e-mail: Jayasree_mm@yahoo.com

N.K. Narayanan

Department of Information Technology, Kannur University, Kannur, India
e-mail: nknarayanan@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_6

53

solution is not practical for real applications. Therefore, we need to devise a specific filter for mixed noise.

Some of the methods in the literature survey have addressed this problem efficiently. A Peer group averaging (PGA) technique detailed in [4–7] was extended to the fuzzy method in [8] which combines a statistical method with an averaging operation to smoothen Gaussian noise. Studies in [4, 5, 7] make use of Fisher linear discriminant. Region analysis method [6], and fuzzy method [8] has been studied. The trilateral filter (TF) [4] which is based on the popular bilateral filter [5, 6] is used to remove Gaussian noise. The study in [7] discusses an impulse noise removal technique which uses a switching mechanism in the bilateral filter. A weighted averaging technique is used in the adaptive nearest neighbor filter (ANNF) [8, 9] which reduces impulses that receive lower weights. The partition based filters [10, 11] classify each intended pixel to be processed into various signal activity groups. A hybrid approach of Bayesian classification with kernel regression is used in [12].

Joan-Gerard Camarena et al. proposed a simple fuzzy method called Simple Fuzzy Rule Filter (SFRF) to remove mixed Gaussian-impulse noise from color images [13]. A simple weighted average operation was carried out over the pixels in a filtering window, which filters each image pixel only once. A fuzzy rule-based system is used to compute the weights adaptively. Even though the method showed good results, the filtering performance and noise suppression capability can be further enhanced by introducing an improved certainty degree. This paper presents an enhanced model for color image filtering using a novel empirical formula for calculating fuzzy certainty degree.

The organization of this paper is as follows: Sect. 2 details the simple fuzzy rule filtering (SFRF) method [13]. Section 2.3 discusses the modified certainty degree for the medium similarity index. Experimental results are provided in Sect. 3, and finally conclusions are drawn in Sect. 4.

2 Simple Fuzzy Rule Filtering (SFRF) Method

Let a color image to be processed be represented as F , and let W be a sliding filtering window, of size $n \times n$ (n equal to an odd number starting from 3), which is centered at the pixel F_0 . The vectors in F denoted as $F_i = (F_i^R, F_i^G, F_i^B)$, are used in the RGB color space. The simple fuzzy rule filter [13], replaces a pixel F_i by a pixel \hat{F}_i which is a weighted average on certain selected pixels in W , denoted by F_0, F_1, \dots, F_m and is given by

$$(\hat{F}_i = \sum_{i=0}^m w_i F^i) / (\sum_{i=0}^m w_i) \quad (1)$$

The weight w_i lies between 0 and 1, and are calculated through a defuzzification method using fuzzy logic inference.

The process of obtaining the weights for each filtering window is explained in the following sections.

2.1 Noisiness of Pixels

Initially, evaluate how much noisy each image pixel is. So, assign a certainty degree $\delta(\mathbf{F}_i)$ to each \mathbf{F}_i as follows. Order the pixels \mathbf{F}_j in the window W centered at \mathbf{F}_i which is of size $n \times n$ in the way $\mathbf{F}(0), \mathbf{F}(1), \dots, \mathbf{F}(n^2 - 1)$ in the order of a distance measure ρ so that $\rho(\mathbf{F}_i, \mathbf{F}(0)) \leq \rho(\mathbf{F}_i, \mathbf{F}(1)) \leq \dots \leq \rho(\mathbf{F}_i, \mathbf{F}(n^2 - 1))$, where $\mathbf{F}(0) = \mathbf{F}_i$. As the distance measure ρ , we use the metric L_∞ [14] given by

$$L_\infty(\mathbf{F}_i, \mathbf{F}_j) = \max \left\{ \left| F_i^R - F_j^R \right|, \left| F_i^G - F_j^G \right|, \left| F_i^B - F_j^B \right| \right\} \tag{2}$$

for the impulse noise detection. Consider the first $s + 1$ pixels $\mathbf{F}(0), \mathbf{F}(1), \dots, \mathbf{F}(s)$. Calculate the *RODs statistic* for \mathbf{F}_i , [4]

$$RODs(\mathbf{F}_i) = \sum_{j=0}^s L_\infty(\mathbf{F}_i, \mathbf{F}_j) \tag{3}$$

As $\mathbf{F}_i = \mathbf{F}(0)$, $L(\mathbf{F}_i, \mathbf{F}(0)) = 0$, *RODs* takes integer values in the interval $[0, 255]$. A small range of *RODs* (\mathbf{F}_i) indicates that the $s + 1$ pixels in W lie close to \mathbf{F}_i which means that \mathbf{F}_i is probably noise free.

If $x = RODs(\mathbf{F}_i)$, the certainty degree $\delta(\mathbf{F}_i)$ is defined by

$$\delta(\mathbf{F}_i) = f(x) = \begin{cases} 0, & x \leq k_1 \\ \frac{x - k_1}{k_2 - k_1}, & k_1 < x < k_2 \\ 1, & k_2 \leq x \end{cases} \tag{4}$$

where the parameters k_1 and k_2 will be detailed in Sect. 3.

2.2 Similarities Between the Processed Pixel and the Remaining Pixels in the Window

To measure the similarity between the two pixels, the metric L_1 is used.

$$L_1(\mathbf{F}_i, \mathbf{F}_j) = \left| F_i^R - F_j^R \right| + \left| F_i^G - F_j^G \right| + \left| F_i^B - F_j^B \right| \tag{5}$$

For assigning the certainty degrees to the three vague propositions above, the following step is performed. Let $x = L_1(\mathbf{F}_0, \mathbf{F}_i)$, and define $\mu_H(\mathbf{F}_0, \mathbf{F}_i) = g_H(x)$ by

$$g_H(x) = \begin{cases} 1, & x \leq a \\ -\frac{x}{(3a)} + \frac{4}{3}, & a < x < 4a \\ 0, & 4a \leq x \end{cases} \tag{6}$$

2.3 Certainty Degree for the Fuzzy Filter

$$g_M(x) = \begin{cases} \frac{(x-a)}{a}, & a < x < 2a \\ 1, & 2a \leq x \leq 3a \\ \frac{(4a-x)}{a}, & 3a < x < 4a \\ 0, & \text{elsewhere} \end{cases} \quad (7)$$

For improving the result of filtering, the certainty degree corresponding to medium similarity given in (7) has been modified by studying the effect of various empirical formulas. The empirical formula which showed best noise removal capability is introduced as (8) given below.

Applying fuzzy negation, assign $\mu_L(\mathbf{F}_0, \mathbf{F}_i) = 1 - \mu_H(\mathbf{F}_0, \mathbf{F}_i)$. The certainty for $\mu_M(\mathbf{F}_0, \mathbf{F}_i) = g_M(x)$ is

$$g_M(x) = \begin{cases} \frac{(2x-a)}{a}, & a < x < 2a \\ 1, & 2a \leq x \leq 3a \\ \frac{4a-2x}{a}, & 3a < x < 4a \\ 0, & \text{elsewhere} \end{cases} \quad (8)$$

2.4 Fuzzy System for Computation of Weights

The fuzzy system rules can be summarized in two main ideas: (1) Noisy pixels should be assigned a small weight; and (2) Noise free pixels can be associated with a larger weight if either they are similar to the central pixel or if the central pixel is noisy.

Before performing the fuzzy inference process, we need to define the fuzzy sets corresponding to the consequents of the fuzzy rules. A certainty degree is associated with each weight $w_i \in [0, 1]$ in the vague statements “ w_i is a large weight,” “ w_i is a medium weight,” and “ w_i is a small weight,” which are denoted by $\nu_L(w_i)$, $\nu_M(w_i)$, and $\nu_S(w_i)$, respectively. The fuzzy sets ν_L , ν_M , and ν_S are represented by a triangular-shape fuzzy membership functions for simplicity of the defuzzification process, as follows:

$$\nu_M(\omega_i) = \begin{cases} \frac{(2\omega_i-1)}{2b-1} + 1, & 1-b < \omega_i \leq 0.5 \\ \frac{(1-2\omega_i)}{2b-1} + 1 & 0.5 < \omega_i < b \\ 0, & \text{elsewhere} \end{cases} \quad (9)$$

$$\nu_L(w_i) = (w_i - 1)/(1 - b) + 1, \text{ if } b < w_i \leq 1 \text{ and } 0 \text{ elsewhere} \quad (10)$$

$$\nu_S(w_i) = w_i/(b - 1) + 1, \text{ if } 0 \leq w_i \leq 1 - b \text{ and } 0 \text{ elsewhere} \quad (11)$$

The certainties corresponding to the antecedents are assigned to their consequents, and finally, by defuzzification, weight w_i of the pixel F_i is obtained. The center of gravity (COG) technique is used, which is one of the popular defuzzification method [15, 16] and is given below.

3 Experimental Results

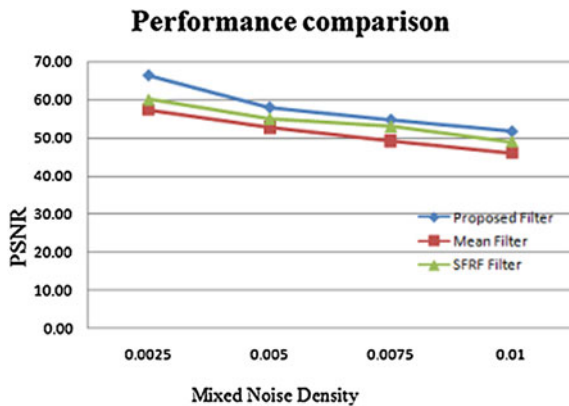
Numeous images from Berkeley Segmentation Dataset [17] and Kodak Dataset [18] have been considered for evaluating the performance of the proposed filter. Table 1 shows the average PSNR for the parrot image that is added with various densities of the two types of mixed noise. Experimentally m has been set to 7. From Table 1, it is clear that the proposed fuzzy filter outperforms mean and Simple Fuzzy Rule Filter (SFRF).

Figure 1 shows graphically the PSNR values of the proposed fuzzy method in comparison with mean and Simple Fuzzy Rule Filter (SFRF). Figure 2a shows the original images from the standard data sets, Fig. 2b shows images contaminated by mixed Gaussian and impulse noise and Fig. 2c–e shows the restored image after applying mean filter, Simple Fuzzy Rule Filter and the proposed filter.

Table 1 Performance in terms of average PSNR value when filtering the parrot image contaminated with different densities of the two types of mixed noise

Noise		PSNR		
Gaussian noise variance (σ)	Impulse noise density (p)	Proposed fuzzy filter	Mean filter	SFRF
0.01	0.1	51.6954	46.0689	48.9632
0.0075	0.075	54.7845	49.1278	52.9658
0.005	0.05	58.0215	52.6389	54.9601
0.0025	0.025	66.4206	57.2541	60.0047

Fig. 1 PSNR value comparison for the proposed method, mean and SFRF filter



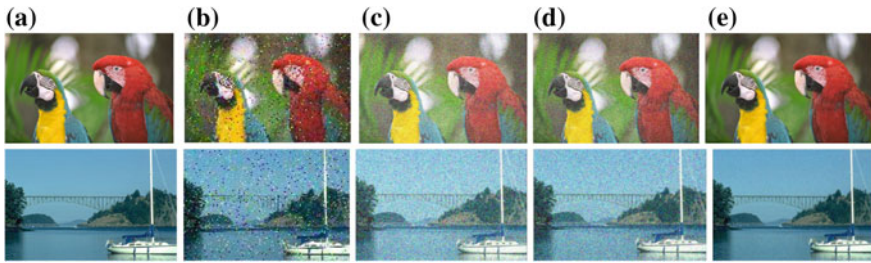


Fig. 2 Outputs for visual comparison. **a** Original images from Berkeley Segmentation dataset and Kodak Dataset, **b** images corrupted with $\sigma = 0.01$ Gaussian and $p = 0.1$ impulse noise, **c** restored image using mean filter, **d** restored image using SFRF filter, **e** restored image using proposed filter

4 Conclusion

A simple and effective fuzzy method is implemented to reduce Gaussian and impulse noise from color images. Only a single filtering operation is used, which is a weighted averaging operation, in which a set of weights are computed by a fuzzy rule system. Experimental results prove that the method reduces noise and also preserve image details, providing competitive results.

Acknowledgment The authors are thankful to the graduate student Ms. Razana K M for her help in the simulation experiment.

References

1. K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*. Berlin, Germany: Springer-Verlag, 2000.
2. R. Lukac, B. Smolka, K. Martin, K. N. Plataniotis, and A. N. Venetsanopoulos, "Vector filtering for color imaging," *IEEE Signal Process. Mag. Spec. Issue Color Imag. Process.*, vol. 22, no. 1, pp. 74–86, Jan. 2005.
3. R. Lukac and K. N. Plataniotis, "A taxonomy of color image filtering and enhancement solutions," *Adv. Imag. Electron Phys.*, vol. 140, pp. 187–264, 2006.
4. R. Garnett, T. Huegerich, C. Chui, and W. He, "A universal noise removal algorithm with an impulse detector," *IEEE Trans. Imag. Process.*, vol. 14, no. 11, pp. 1747–1754, Nov. 2005.
5. C. Tomasi and R. Manduchi, "Bilateral filter for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.
6. M. Elad, "On the origin of bilateral filter and ways to improve it," *IEEE Trans. Imag. Process.*, vol. 11, no. 10, pp. 1141–1151, Oct. 2002.
7. C. H. Lin, J. S. Tsai, and C. T. Chiu, "Switching bilateral filter with texture/ noise detector for universal noise removal," *IEEE Trans. Imag. Process.*, vol. 19, no. 8, pp. 2307–2320, Sep. 2010.
8. K. N. Plataniotis, D. Androutsos, and A. N. Venetsanopoulos, "Multichannel filters for image processing," *Signal Process. Imag. Commun.*, vol. 9, no. 2, pp. 143–158, Jan. 1997.
9. K. N. Plataniotis, D. Androutsos, and A. N. Venetsanopoulos, "Adaptive fuzzy systems for multichannel signal processing," *Proc. IEEE*, vol. 87, no. 9, pp. 1601–1622, Sep. 1999.

10. Z. Ma, H. R. Wu, and D. Feng, "Partition based vector filtering technique for suppression of noise in digital color images," *IEEE Trans. Imag. Process.*, vol. 15, no. 8, pp. 2324–2342, Aug. 2006.
11. Z. Ma, H. R. Wu, and D. Feng, "Fuzzy vector partition filtering technique for color image restoration," *Comput. Vis. Imag. Understanding*, vol. 107, no. 1–2, pp. 26–37, Jul./Aug. 2007.
12. E. L´opez-Rubio, "Restoration of images corrupted by Gaussian and uniform impulsive noise," *Pattern Recog.*, vol. 43, no. 5, pp. 1835–1846, May 2010.
13. Joan-Gerard Camarena, Valent´ın Gregori, Samuel Morillas, and Almanzor Sapena, "A Simple Fuzzy Method to Remove Mixed Gaussian-Impulsive Noise From Color Images," *IEEE Trans. Imag. Process.*, vol. 21, no. 5, October 2013.
14. J. G. Camarena, V. Gregori, S. Morillas, and A. Sapena, "Two-step fuzzy logic-based method for impulse noise detection in colour images," *Pattern Recog. Lett.*, vol. 31, no. 13, pp. 1842–1849, Oct. 2010.
15. K. Passino and S. Yurkovich, *Fuzzy Control*. Menlo Park, CA: Addison Wesley, 1998.
16. E. Cox, *The Fuzzy Systems Handbook*, 2nd ed. New York: Academic, 1999.
17. Eecs.berkeley.edu,(2015).[online]Available at: <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images.html>.
18. KODAK Test Images Database [Online]. Available: <http://r0k.us/graphics/kodak/>.

Face Recognition Using Snakes Algorithm and Skin Detection Based Face Localization

Rakshit Ramesh, Anoop C. Kulkarni, N.R. Prasad and K. Manikantan

Abstract Skin detection is an important part of face localization since the most exposed part of human skin is the face. This paper proposes a novel algorithm for face localization via skin detection. The algorithm utilizes a kernel iterative procedure to check for the region of interest in an image where it is likely that the face exists. The algorithm utilizes a swarm of particles in a kernel that randomly check the fitness of the corresponding pixel BGR value which is determined from a skin BGR dataset. Consequently, the kernel which has the best fitness value is chosen as the region of the image where it is likely that the face exists. Following this, we employ an active contour model called *Snakes algorithm* which further converges on our region of interest and finally a contour of the face region is extracted from the original image.

Keywords Face localization · Swarm intelligence · Face recognition · Binary particle swarm optimization · Snakes algorithm

1 Introduction

Face Recognition (FR) is a topic of major research in recent years as it has proved to be quite successful as a biometric. Advancement in image analysis has improved the performance of FR systems however varying backgrounds, pose and illumination still are a major challenge. A comprehensive survey of the techniques can be

R. Ramesh · A.C. Kulkarni · N.R. Prasad · K. Manikantan (✉)
Department of Electronics and Communication Engineering,
M S Ramaiah Institute of Technology, Bangalore 560054, India
e-mail: kmanikantan2009@gmail.com

R. Ramesh
e-mail: rakshitadmar@gmail.com

A.C. Kulkarni
e-mail: rnianoop@gmail.com

N.R. Prasad
e-mail: prasadr606@yahoo.in

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_7

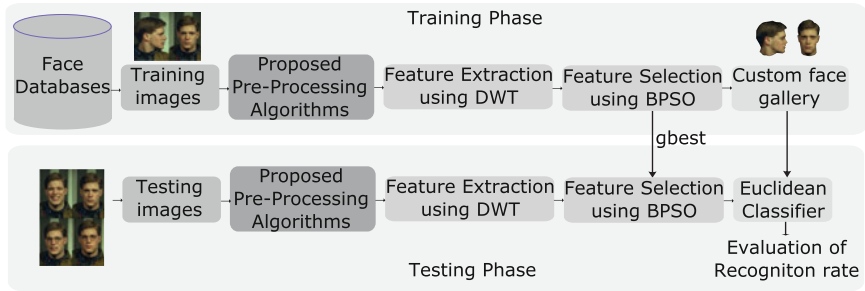


Fig. 1 Proposed face recognition system

found in [1]. In this paper, we describe a simple technique to eliminate the issue of varying background. The proposed algorithm is capable of face localization in order to obtain the exact region of interest on which further face recognition algorithms may be applied. It is found that without a face localization algorithm, face recognition becomes erroneous since the background becomes a part of the features to be recognized. A localization algorithm such as ours thus becomes necessary in resolving the issue. It is also found that due to such a process, recognition rate increases and false positive detection is averted as is explained in the course of this paper.

The proposed system is used as a pre-processing step in the overall face recognition system as shown in Fig. 1.

2 Problem Statement and Contributions

As experimentally proved in [2] if the training face image contains significant background, hair and shoulders, it will mislead the recognition process, therefore it is necessary to use the 'correct' face-only database and remove the irrelevant portions. In order to overcome this problem, we have proposed a simple technique wherein face region of interest can easily be localized thereby improving the recognition rate and removing false positive detection. The contributions made in this paper are:

- i Implemented Skin Detection algorithm as a means to detect face. The algorithm has proven to be very reliable under all circumstances and provides no false positive cases under uniform illumination. Proposed skin detection algorithm uses a novel new method for determining the fitness of a region and its likelihood of being skin, called Particle Swarm Search.
- ii Used an Active Contour model called Gradient Vector Flow Snake to further obtain face boundary region of interest. The algorithm proves to reliably converge on to face like contour and therefore provide us the best possible regions of interest.

3 Proposed Preprocessing Methodology

We propose a combination of two techniques to achieve exact facial region of interest from an image.

- i A novel Skin Detection Algorithm which localizes the image to an area with a high probability of containing the face.
- ii Snakes algorithm to further converge on to the face from the localized area of interest.

3.1 Proposed Skin Detection Algorithm for Face Localization

The proposed algorithm is a very simple approach towards face localization by skin detection, detecting that part of an image which has the best fitness value obtained by an ellipsoidal fitness function enclosing the points defined by a dataset containing randomly sampled skin pixel’s BGR values [3]. The ellipsoidal function and the dataset from which it’s obtained is as shown in Fig. 2.

The ellipsoidal function is obtained by fitting an ellipsoid to the BGR dataset by the LMS Curve fitting method in 3 dimensions. The equation of the ellipsoid for the dataset is obtained as in equation Eq. 1.

$$\frac{(B - 166)^2}{166^2} + \frac{(G - 100)^2}{37^2} + \frac{(R - 71)^2}{182^2} - 1 = 0 \tag{1}$$

We define a kernel (region of interest(ROI)) of fixed size and iterate it throughout the image as shown in Fig. 3, this kernel is the input to our face detection function. A swarm of randomly positioned particles test randomly sampled pixels for their BGR values. Our fitness function then defines the fitness for these individual particles by testing whether the particles pixel coordinates lie inside the fitness ellipsoid or not. A positive fitness value is obtained if the said particles BGR value lies inside the ellipsoid and negative fitness value if it lies outside.

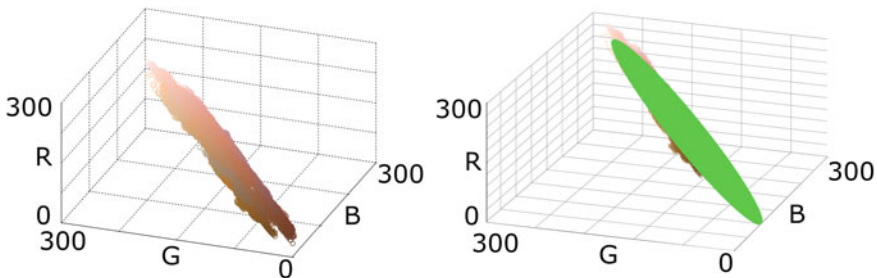


Fig. 2 Ellipsoidal fitness function fit on a skin dataset

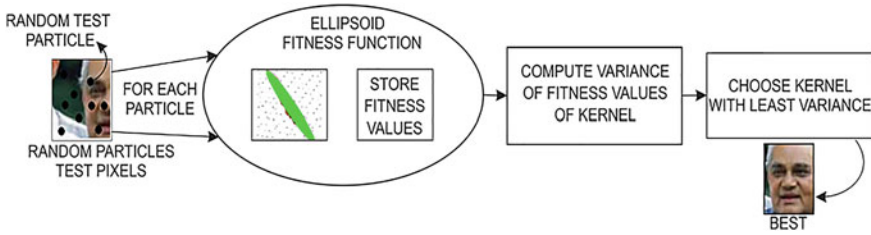


Fig. 3 Graphical representation of algorithm

The above process is repeated for each ROI kernel which is iterated pixel wise or block wise (using an arithmetic progression). A fitness value for each of these ROI kernels is also defined as the variance of the fitness value of each of the particles comprised within the ROI kernel. The ROI kernel with the least variance and most positive mean is chosen as the region of the image with a high probability of being the face.

Algorithm 1 Face Region of Interest Selection Algorithm

1: **procedure** FACEDETECT

2: **for** Each kernel $i \in N$ **do**

3: Initialize k particles with random position

$$p(k) = rand(x_k, y_k)$$

4: Obtain fitness value of each particle with pixel value (Blue, Green and Red)

$$v(k) = \frac{(B - 166)^2}{166^2} + \frac{(G - 100)^2}{37^2} + \frac{(R - 71)^2}{182^2} - 1$$

5: Obtain the fitness value of the kernel

$$V(i) = \text{variance}(v(k))$$

6: **end for**

7: Choose best kernel K as face

$$K = \min(V)_{index}$$

8: **end procedure**

3.2 Snake Algorithm for Head Boundary Extraction

Snakes [4, 5], or active contours are parametric curves guided by internal and external forces to converge onto an object boundary. Forces are defined in a such a way that the contour corresponding to the object boundary Eq. 2 has the minimum energy Eq. 3.

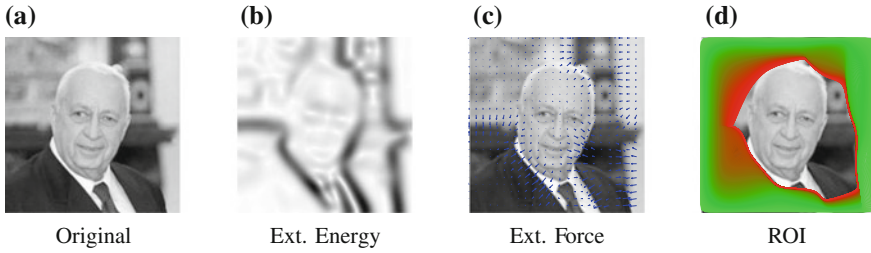


Fig. 4 Snake convergence on region of interest

$$E_{snake} = E_{internal} + E_{external} + E_{constraint} \tag{2}$$

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term} \tag{3}$$

$$E_{snake} = \int_s \frac{1}{2}(\alpha(s)|v_s|^2 + \beta(s)|v_{ss}|^2) + E_{image}(v(s))ds \tag{4}$$

where α is the elastic parameter β is the bending parameter. Hence in order to find the contour the above energy functional in Eq. 4 is to be minimized.

Gradient Vector Flow [6] was proposed as a external force to overcome shortcomings of traditional snake such as convergence in boundary concavities and large capture range [7].

$$V(x, y) = (u(x, y), v(x, y)) \tag{5}$$

$$E = \oint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 dxdy \tag{6}$$

$V(x, y)$ in Eq. 5 is defined such that it minimizes the above energy functional in Eq. 6, where $f(x, y)$ is the edge map of the image. An algorithm [8] for GVF snake using the above, is able to settle on the required region of interest, the face in our case as shown in Fig. 4.

4 Feature Extraction Using DWT

A family of wavelets called symlets which increased symmetry in comparison to Haar, hence respond well to facial features are applied as a Discrete Wavelet Transform on the testing and training image. Therefore DWT acts as a feature extractor by retaining only the useful components [9].

5 Binary Particle Swarm Optimization Based Feature Selection

We use the Binary Particle Swarm Optimization [10–13] technique for optimized feature selection. The BPSO then selects N distinct elements of this feature vector as a binary chromosome which defines the features of that image. Since the algorithm is binary, the chromosome either contains 1 for selected feature or 0 for non-selected feature. A swarm of particles test each of the elements to produce the Binary feature chromosome according to the equations Eqs. 7 and 8.

$$V_i^{t+1} = \beta \times V_i^t + c_1 \times rand_1 \times (p_{best}^i - X_i^t) + c_2 \times rand_2 \times (g_{best} - X_i^t) \quad (7)$$

$$X_i^{t+1} = \begin{cases} 1, & rand_3 < \frac{1}{1+e^{-V_i^{t+1}}} \\ 0, & otherwise \end{cases} \quad (8)$$

6 Euclidean Classifier

An Euclidean distance classifier is then used to find out the difference in Euclidean distance between the training feature vector and the testing feature vector chromosome, which is a measure of disparity between the training and testing images, to recognize the subject.

7 Experimental Results and Discussion

We use *Recognition Rate* (RR) as a metric for face localization. A localized face image provides higher recognition rate. Localized faces also provide the ideal training image feature vectors hence false positive detection is averted.

The proposed techniques are implemented in MATLAB® [14] and various experiments are conducted on standard face databases to verify the performance. Experiment 1 evaluates the performance on databases with significant head pose variation whereas Experiment 2 evaluates on databases with significant background variation. Experiment 3 is done using Viola-Jones algorithm as the face extractor for comparison with the proposed technique.

Table 1 describes the details of the 4 databases used, and Table 2 is results of the experiments.

Table 1 Database information

Parameter	Color FERET	LFW	FEI	Caltech
Total images	11388	13233	2800	450
Total subjects	856	5,749	10	27
Selected subjects	10	5	10	10
Images/subject	20	20	30	21
Type	PPM	JPG	JPG	JPG
Size (pixels)	384 × 256	250 × 250	288 × 384	896 × 592
Variation	P, E, S	P, E, S, B	P	E, B, I

P = Pose, E = Expression, S = Scale, I = Illumination, B = Background

Table 2 Experimental results

Database	Proposed algorithm	Viola Jones	Tr:Te	Kernel	Snake iteration	RR (%)
Caltech	×	×	4:8	–	–	30.75
	✓	×	4:8	350 × 350	50	73.25
	✓	×	4:8	350 × 350	300	58.13
	×	✓	4:8	–	–	63.75
LFW	×	×	–	–	–	22.53
	✓	×	4:8	150 × 150	200	36.13
	✓	×	4:8	150 × 150	600	42.25
	✓	×	4:8	200 × 200	300	44.50
	✓	×	4:8	200 × 200	600	43.25
	×	✓	4:8	–	–	30.25
FERET	×	×	–	–	–	39.33
	✓	×	8:12	300 × 300	300	34.91
	✓	×	8:12	300 × 300	600	64.64
	×	✓	8:12	–	–	55.13
FEI	×	×	4:8	–	–	59.25
	✓	×	4:8	350 × 350	600	77.59
	×	✓	4:8	–	–	57.56

VJ = Viola Jones Face Extractor, Tr:Te = Training testing ratio

7.1 Experiment 1 : Color FERET and FEI

The original Color FERET [15] database has a total of 11338 images, with varying poses and facial expressions. A subset of this database, using 10 subjects, with 20 images per subject has been used for the experiments.

FEI [16] database has a total of 2800 images of 200 subjects.

Pose variant algorithms pose a slight issue to the feature selector and feature classifier part of the face recognition system. Hence, despite background being removed,

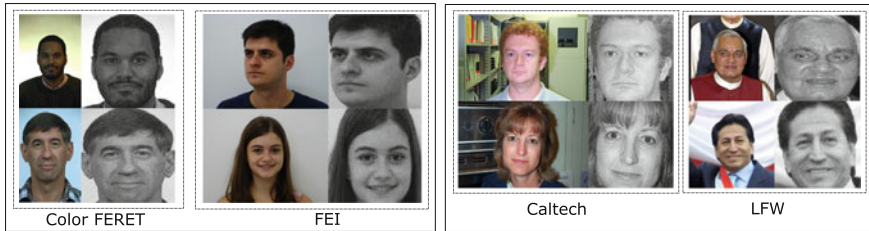


Fig. 5 Backgrounds being removed using proposed methodology

we don't obtain higher percentages of recognition rate. Despite this, recognition rates when proposed algorithm is used is higher in comparison to that when it is not used.

7.2 *Experiment 2 : LFW and Caltech 1999*

The LFW database [17] consist of 13,233 images faces of 5749 different persons, obtained from news images by means of a face detector. These images have a very large degree of variability in the face expression, age, race, background and illumination conditions. Images are taken in very noisy background such as public places. From the results, we can conclude that for images with neutral pose and varying backgrounds, proposed algorithm is capable of exact localization as shown in Fig. 5, therefore we can observe high percentage of recognition rate. In the case of LFW, we obtain 44.5 % with our algorithm as compared to 22.5 % without our algorithm. This clearly indicates that background removal has positively affected recognition rate and therefore validates the requirement of our algorithm.

Caltech 1999 [18] is a frontal pose database with 450 face images, 27 unique people under with different lighting, expressions and backgrounds. The images are taken in noisy non-uniform environments such as classrooms, libraries, etc. We obtain 73.3 % recognition rate with our algorithm whereas without our algorithm, we obtain 30.8 %. This again validates the use of our algorithm as noted in Table 2.

7.3 *Experiment 3 : Using Viola-Jones Face Extraction*

The standard alternative to our algorithm has been the Viola-Jones Face Extractor [19] which uses the Haar Wavelets and a Cascade object classifier to classify region of image as face. However, it both yields lesser recognition rate and plenty of false positives as shown in Fig. 6. The results are shown in Fig. 7.

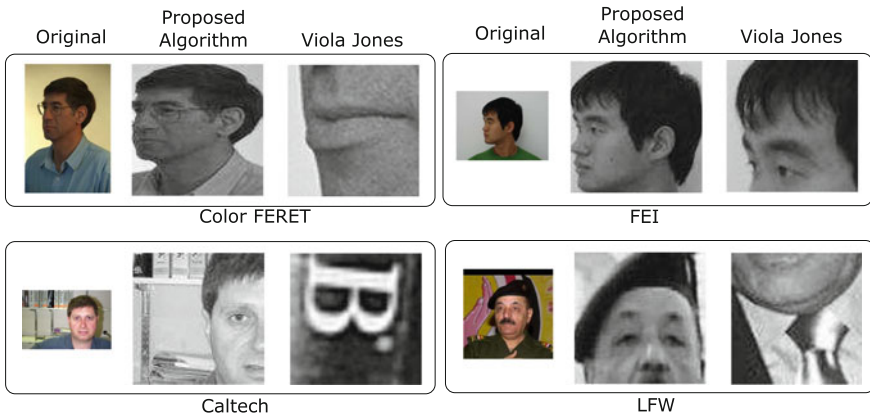


Fig. 6 Successful detection by the proposed algorithm versus failed detection by Viola-Jones method

Recognition rates with different databases

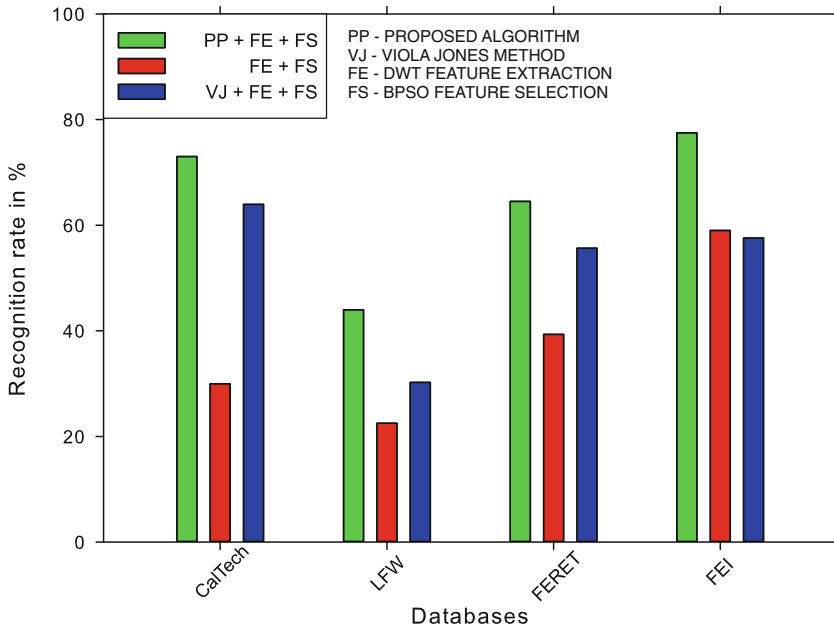


Fig. 7 Plot of experimental results

8 Conclusions and Future Work

Two novel algorithms were proposed which address the issue of varying background and distracting elements in the face image. The extensive experimentation with the proposed techniques on FERET, FEI, Caltech and LFW have shown significant improvement in recognition rates. The face only region extracted by the proposed algorithm successfully mask the background and hence eliminate the possibility of FR system incorrectly learning elements like hair, shoulder and background.

Therefore the significant improvement in FR rate without any additional pre-processing can be attributed to the FR algorithm correctly learning the face features only. Some algorithms artificially boost RR by exploiting image characteristics such as correlation between faces and backgrounds. We can confidently claim the RR obtained correlates the pure face only region only.

Scale invariant face detection and multiple face localization can be evaluated as a future extension to the Particle Swarm Search algorithm. Snakes algorithm convergence is limited to spatial gray scale domain, further convergence in other domains such as wavelet domain or contourlet domain in a HSV color representation can be pursued.

A face validation [20] technique needs to be implemented to ensure that detected face is actually face. Faster algorithms for testing and hence real time algorithms can be developed in the future.

References

1. W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld.: Face recognition: A literature survey, *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458 (2003)
2. Li-Fen Chen, Hong-Yuan Mark Liao, Ja-Chen Lin, Chin-Chuan Han.: Why recognition in a statistics-based face recognition system should be based on the pure face portion a probabilistic decision-based proof, vol. 34, no. 7, pp. 1393–1403 (2001)
3. UCI Machine Learning Repository <https://archive.ics.uci.edu/ml/datasets/Skin+Segmentation>
4. S. R. Gunn and M. S. Nixon.: A dual active contour for head and boundary extraction, in *IEEE Colloquium on Image Processing for Biometric Measurement*, London, pp. 6/1–6/4 (1994)
5. Nixon, M., Aguado, A.: *Feature Extraction and Image Processing*, 3rd edition, Elsevier (2012)
6. C. Xu and J. L. Prince.: Snakes, shapes, and gradient vector flow. In: *IEEE Transactions Image Processing*, vol 7 issue 3, The Johns Hopkins University, (1998)
7. J. L. Prince and C. Xu.: A new external force model for snakes - 1996 *Image and Multidimensional Signal Processing Workshop*, pp. 30–31 (1996)
8. Dirk-Jan Kroon, University of Twente.: Snake Active Contour, Mathworks, Matlab Central www.mathworks.com/matlabcentral/fileexchange/28149-snake---active-contour
9. Rafael Gonzalez, Richard Woods.: *Digital Image processing*. Prentice Hall, Third edition (2006)
10. Rabab M. Ramadan and Rehab F. Abdel Kader.: Face Recognition Using Particle Swarm Optimization-Based Selected Features, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 2, no. 2 (2009)
11. J. Kennedy, R. Eberhart.: A New Optimizer Using Particles Swarm Theory. *Proceedings of 6th International Symposium Micro Machine and Human Science*, pp. 39–43 (1995)

12. J. Kennedy, R. Eberhart.: A Discrete Binary Version of the Particle Swarm Algorithm. In: Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp. 4104–4108 (1997)
13. J. Kennedy, R. Eberhart.: Particle swarm optimization. In: IEEE international conference on neural network, vol. 4, issue 4, pp. 1942–1948 (1995)
14. Mathworks (MATLAB) <http://www.mathworks.in>
15. Color FERET Database <http://www.nist.gov/itl/iad/ig/colorferet.cfm>
16. FEI Database <http://fei.edu.br/~cet/facedatabase.html>
17. LFW database <http://vis-www.cs.umass.edu/lfw/>
18. CalTech Face (Front) Database www.vision.caltech.edu/Image_Datasets/faces/faces.tar
19. Paul Viola, Michael J. Jones.: Robust real-time face detection, International journal of computer vision, vol. 57, no. 2, pp. 137–154 (2004)
20. Sanket N J, K Manikantan and S Ramachandran.: Recursive Binary Particle Swarm Optimization based Face Localisation, Computer Vision, Pattern Recognition, Fourth National Conference, Image Processing and Graphics, NCVPRIPG (2013)

Quantifiable Image Nearness Approach Using Descriptive Neighbourhood

M. Sajwan and K.S. Patnaik

Abstract Similarity metrics plays an important role in Content-Based Image Retrieval (CBIR). This article introduces a new technique called *Descriptive Proximal Coverage (DPC)* to measure the quantifiable similarity index between images. This work is based on Near Neighbourhood approach in which perceptually relevant information is extracted from group of objects based on their description. Two images are considered as sets of perceptual objects and affinities between objects is defined by a tolerance relation. Two visual objects are similar if the difference between their descriptions is smaller than a tolerable level of error. Existing Notion of nearness stems from the observation that in nature it is rare to find exact objects but similar objects are often seen. It is imperative to provide a numeric value which will quantify similarity and nearness between images.

Keywords Tolerance relation • Hausdorff distance • Perceptual objects • Granularity • Near sets

1 Introduction

Quantifying the visual similarity between pairs of digital images is of prime importance in CBIR [1–3]. This is achieved by extracting some visual descriptions (i.e. features) from images and then finding the distance between feature vector values in order to determine how much they resemble each other. A common methodology to compare features is to calculate some form of distance like Hausdorff [4, 5], Minkowski distance [2, 6, 7] or to use Histogram Intersection and L_p Norm distance between histograms [8, 9]. There are various other distance metrics which can be found in [4, 5, 10–13].

M. Sajwan (✉) • K.S. Patnaik
Department of Computer Science & Engineering,
Birla Institute of Technology, Mesra, Ranchi 835215, Jharkhand, India
e-mail: mohit.sajwan@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_8

73

The above metrics come under the family of geometric approaches to image similarity. In 2007, James F. Peters introduced Near Set theory [14, 15] in image processing and opened a new dimension for research. The following section describes the concept of near set theory and tolerance relation, which are foundation stones of DPC.

1.1 Near Set Theory

Disjoint sets containing objects with similar descriptions are near sets [16–18]. Objects that have some commonality in their features are called perceptually near objects. Features of objects are given by Probe functions [15] represented as ψ . Probe functions take perceptual object as input and returns its characteristic property as output [19]. A set of describable/perceptual object in conjunction with a set of probe functions forms the perceptual system $\langle P_O, P_F \rangle$ where P_O denotes a set made up of describable objects and P_F represents set constituting real-valued functions $\psi \in P_F$ [3, 18].

Let α be a perceptual or describable object and $\psi(\alpha) = (\psi_1(\alpha), \psi_2(\alpha), \psi_3(\alpha), \psi_4(\alpha), \dots, \psi_k(\alpha), \dots, \psi(\alpha))$ where $\psi(\alpha)$ is a feature vector and $\psi_i(\alpha)$ is a probe function value of the k th probe function ψ_k and for example, $\psi_{\text{red}}(\text{image}_x)$ represents the red color pixel value of a particular image X .

Tolerance space has been widely referred in [16, 20–22]. Let P_O be a collection of perceivable objects, and let \mathbb{Y} be a tolerance relation (i.e. a binary relation) on I ($\mathbb{Y} \subset I \times I$) be reflexive ($\forall i \in I, i\mathbb{Y}i$) and symmetric ($\forall i_1, i_2 \in I$, if $i_1\mathbb{Y}i_2$, then $i_2\mathbb{Y}i_1$) but is not necessarily transitive [21].

Now, a tolerance space can be described as $\langle I, \mathbb{Y} \rangle$ and tolerance relation $\cong_{p,t}$ is defined by: $\cong_{p,t} = \{(o_1, o_2) \in P_O \times P_O: \|\psi(o_1) - \psi(o_2)\|_2 \leq t\}$, where $\|\cdot\|_2$ is L^2 Norm, pre-determined threshold $t \in R_0^+$ and probe function $p: p \subseteq P_F$.

2 Perceptual Objects (Visual Element) in an Image

A perceptual object ($x \in O$) is something that can be presented to the senses and recognized by the human mind [18, 23]. For example, a group of pixels in an image can be described as a perceptual object that can be seen and perceived. Image is made up of indivisible units called as pixels. In our case a Perceptual object can be a single pixel or a group of neighboring pixels. Any pixel (or a cluster of neighboring pixels) at any position can be designated as visual element. Generally a patch of image (pixel with its adjacent pixels) is selected as perceptual object. From

computation point of view, program will run faster and less information will be needed to represent a sub image rather than just a pixel. This can be explained by an example, for a granule size of 1 pixel; and image size of 512×512 , where each pixel has 10 features, requires a matrix of $512 \times 512 \times 10$ (for a single image) to compute the results. But if we choose a granule size of 32×32 pixel then it would lead to only 256 blocks requiring a 256×10 matrix size for computation. Hence, the storage space can be reduced by 1024 times. As the human eye does not perceive images in pixel based resolution (i.e. human eyes are not sensitive enough to observe pixel level details), choosing coarse grain of pixels is a justifiable explanation. Hence for simplicity, and above mentioned arguments, we consider each describable object to be a small square consisting of ' n ' pixels. The value of ' n ' is optional and it must be carefully chosen, ensuring it neither very small ($n = 2$ or 4) nor very high ($n = 512$).

3 Perceptual Object Granularity

To preserve the descriptive and spatial relationship among visual objects of an image, recursive decomposition of an image into ' m ' equal size blocks (or quadrants) is adopted. This process can be stopped at any iteration depending upon the granularity level i.e. from coarse to fine grain. Hence it stops earlier for coarse matching, where an entire image can be considered as a grain as compared to fine matching where each single pixel is considered as a grain. Thus, this process can be described as an m -ary tree formation of divide and conquer approach.

In Fig. 1 the initial undivided image is shown as numeric 1. Then after first recursive decomposition formation of four sub-blocks is represented as numeric 2, 3, 4 and 5. Then each sub-block is further divided into four subsections. For simplicity, division of only block 5 is shown and the finest grains are 6, 7, 8 and 9. Same process is applied for another image and corresponding blocks of corresponding images are mapped for similarity check. This can be depicted in form of 4-ary tree as

In Fig. 2, L1 represents initial undivided image. Iteratively, L2 represents 4 sub-blocks or quadrants and L3 represents 16 sub blocks of the image. From moving top to bottom in the complete 4-ary tree, coarse size decreases. The finest level will be a single pixel. In this example $m = 4$, indicates at each iteration each block will be divided into 4 sub-blocks recursively. This m is called as *division factor*.

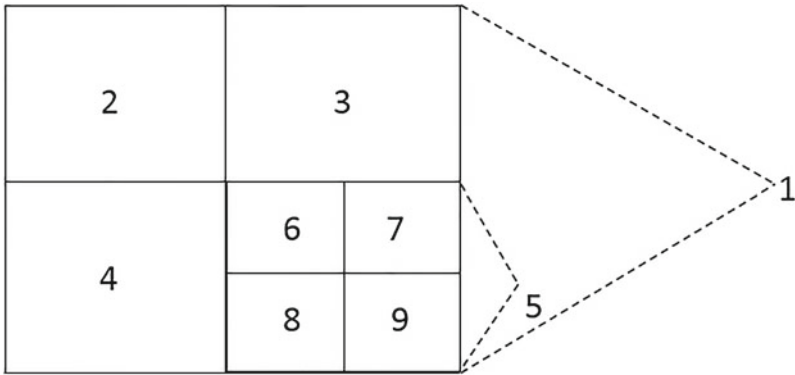


Fig. 1 Image decomposition

It must be noted that $m = 2^i$, $\{i \in \mathbb{Z} : i > 1\}$ for sub-block division. Value of $i = 0$, indicates no recursive decomposition will occur and entire image would be considered as perceptual object and value of $i = 1$ indicates that, on each iteration image will be divided into two equal sized strips and the finest granularity that can be achieved is either a single row or column (of pixels and not a single pixel). If *division factor* is an even number then at each iteration equal sized blocks will be generated and if division factor is in odd number equal sized strips will be generated at each iteration.

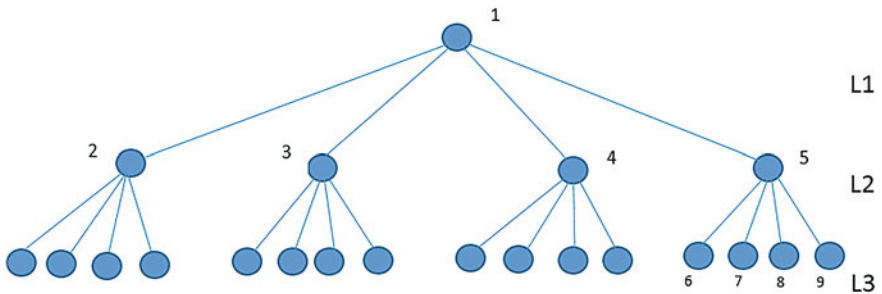


Fig. 2 Decomposition tree

4 Descriptive Proximal Coverage (DPC) Algorithm

1. Input : $I_A, I_B, \varepsilon, \eta$ (η is size of each block/granule/perceptual object)
2. Output: $I_A I_B$
3. for each block \mathfrak{b} of size $\eta \times \eta$
4. Find $\mathcal{F}_\eta(I_A)$ (The unique features in sub-image $I_{A\eta \times \eta}$);
5. Find $\mathcal{F}_\eta(I_B)$ (The unique features in sub-image $I_{B\eta \times \eta}$);
6. $\mathcal{F}_\eta(I_C) \leftarrow \mathcal{F}(I_B)$;
7. $\mathcal{F}_\eta(I_A \cap I_B) \leftarrow \text{NULL}$;
8. for $\psi(i_a) \in \mathcal{F}_\eta(I_A)$
9. for $\psi(i_b) \in \mathcal{F}_\eta(I_B)$
10. if $\|\psi(i_a) - \psi(i_b)\|_2 \leq \varepsilon$
11. $\mathcal{F}_\eta(I_A \cap I_B) \leftarrow \mathcal{F}_\eta(I_A \cap I_B) \cup \psi(i_a)$;
12. $\mathcal{F}_\eta(I_A \cap I_B) \leftarrow \mathcal{F}_\eta(I_A \cap I_B) \cup \psi(i_b)$;
13. $\mathcal{F}_\eta(I_C) \leftarrow \text{SetDifference}(\mathcal{F}_\eta(I_C), \psi(i_b))$;
14. if $|\mathcal{F}_\eta(I_C)| > 0$
15. for $\psi(i_c) \in \mathcal{F}_\eta(I_C)$
16. if $\|\psi(i_a) - \psi(i_c)\|_2 \leq \varepsilon$
17. $\mathcal{F}_\eta(I_A \cap I_B) \leftarrow \mathcal{F}_\eta(I_A \cap I_B) \cup \psi(i_c)$;
18. $\mathcal{F}_\eta(I_C) \leftarrow \text{SetDifference}(\mathcal{F}_\eta(I_C), \psi(i_c))$;
19. break;
20. $I_A I_{B\eta \times \eta} \leftarrow \{i_x \in I_A \cup I_B : \psi(i_x) \in \mathcal{F}_\eta(I_A \cap I_B)\}$;
21. $\mathcal{R}(i_b) \leftarrow \text{PercentCoverage}(I_A I_{B\eta \times \eta}, I_A, I_B)$;
22. $\text{DPC} \leftarrow \text{Average}(\mathcal{R})$;

4.1 Percentage Coverage

1. **PercentCoverage**(I_A $I_{B\eta \times \eta}$ I_A , I_B) // function call from DPC Algorithm
2. Find $\mathcal{F}(I_A I_B)$ (Features in $I_A I_{B\eta \times \eta}$)
3. for each granule $g \in \mathcal{F}(I_A I_{B\eta \times \eta})$
4. for each occurrence g in I_A
5. Increment *CountPixelA*;
6. end
7. for each occurrence g in I_B
8. Increment *CountPixelB*;
9. end
10. *CommonPixelA*(f) \leftarrow *CountPixelA*;
11. *CommonPixelB*(f) \leftarrow *CountPixelB*;
12. end
13. Evaluate T_A (Total of all pixels in *CommonPixelA*);
14. Evaluate T_B (Total of all pixels in *CommonPixelB*);
15. $Perc_A \leftarrow T_A / \text{PixelCount}(I_A) * 100$ ($Perc_A$ is percent coverage of I_A);
16. $Perc_B \leftarrow T_B / \text{PixelCount}(I_B) * 100$ ($Perc_B$ is percent coverage of I_B);
17. $\mathbf{P} \leftarrow \min(\mathbf{P}_A, \mathbf{P}_B)$ (Similarity between two images in percentage);

DPC algorithm is based on intersection algorithm given by Henry [23]. In DPC, the descriptive intersection is calculated on corresponding sub-images of A and B (of size $\eta \times \eta$) by comparing feature vector $F_\eta(I_A)$ to $F_\eta(I_C)$ (which is a copy of feature vector $F_\eta(I_B)$). During the execution of DPC algorithm, perceptual object characteristic $\psi(i_a) \in F_\eta(I_A)$ is compared to each perceptual object characteristic $\psi(i_b) \in F_\eta(I_B)$. If a match is found, then characteristic from $F_\eta(I_C)$ is removed, i.e. $F_\eta(I_C) \leftarrow F_\eta(I_C) - \psi(i_b)$, leaving the original feature/characteristic set $F_\eta(I_B)$ unchanged. After this step, $\psi(i_a)$ is compared with remaining perceptual object characteristics in $F_\eta(I_C)$, beginning from b position. If further match occurs, it is

also deleted from $F_{\eta}(I_C)$. This process is repeated by comparing elements in $F_{\eta}(I_A)$ with $F_{\eta}(I_B)$, and converges to the set $F_{\eta}(I_C)$ by reducing the number of comparisons.

In next section of the algorithm percentage similarity is calculated as follows. Vector $I_A I_B$ contains the intersected pixels. Frequency of each pixel in both the images is computed (i.e. how many times each intersected pixel is occurring in both the images). The minimum of two values gives the descriptive proximal coverage. Since in this approach we are dividing the images into blocks/perceptual objects, object_i of first image will be mapped to object_i of second image. If image is divided into N equal blocks, N results will be obtained, and average of them will give the DPC value for the test images.

The term descriptive has its own significance. If user decides to consider an entire image as a descriptive object (or a single grain), then spatial relationship between pixels wanes because pixel at (1, 1) location in first image can be compared with pixel at (255, 255) location in second image as entire image is a block itself. This dilutes the spatial proximity because the top left region of first image is mapped to the right end, which is not allowed under spatial proximity constraints. If finer grains are chosen then the image is further divided into equal sized strips or blocks (or quadrants) depending on the division factor. Since we divided the image into sub-blocks, corresponding blocks of corresponding images are mapped to each other and descriptive similarity is computed. In this context not only descriptive properties but to an extent spatial proximity also materializes.

5 Results

Table 1 represents percentage similarity between different images of buses. DPC algorithm was implemented with Bus dataset as input, and results are tabulated in matrix form. Since each matrix corresponds to images of similar objects, their percentage similarity must also be high. Values in the matrix support this argument. Bus matrix consist of images of similar buses.

In this article, the proposed algorithm (DPC) is compared with the Hausdorff Metric [4, 5] only, because the Hausdorff metric allows a portion of one image to be compared with another, since it is a point (pixel) to set distance or set to set distance while other distances are point(pixel) to point(pixel) distance, which are sensitive to small perturbations. DPC also resembles Hausdorff, since it is also a set to set comparison.

For further discussions, it must be noted that the test image is the first image of dataset. This is applicable for all cases. In Figs. 3, 4, 5, 6, 7 and 8, graphs depict the similarity comparison of test image with rest of the images in the same dataset.

Figures 3 and 4 represent the comparative analysis of DPC and Hausdorff distance for an image set consisting of images of same person (his face) with different moods and different face orientations. In both cases, it can be visualized that test image has high similarity quotient with the given image set (it is obvious, as images are of same person). From Fig. 3 it can be claimed that DPC values (in percentage

Table 1 Bus matrix

DPC matrix						
	100 %	80.1229 %	76.1247%	86.9551%	69.6221%	80.7387%
	80.1229 %	100%	73.2395%	78.8605%	71.9215%	78.1708%
	76.1247%	73.2395%	100%	82.0560%	65.3171%	72.5759%
	86.9551%	78.8605%	82.0560%	100%	72.5759%	80.5712%
	69.6221%	71.9215%	65.3171%	72.5759%	100%	75.3133%
	80.7387%	78.1708%	72.5759%	80.5712%	75.3133%	100%

Fig. 3 Similarity comparison on dataset1

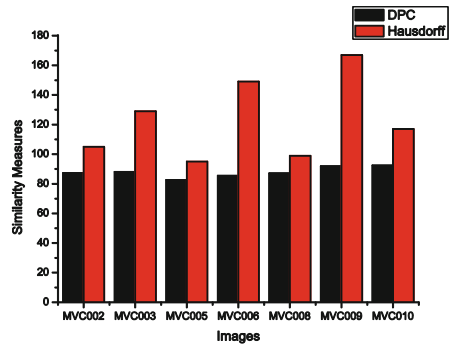


Fig. 4 Similarity comparison on dataset2

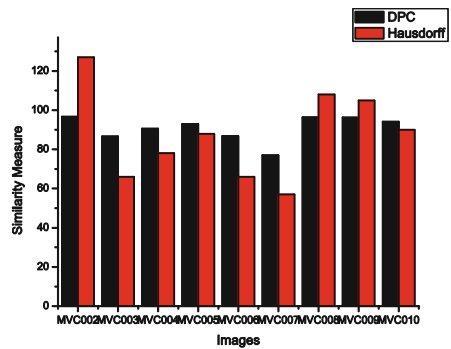


Fig. 5 Similarity comparison on dataset3

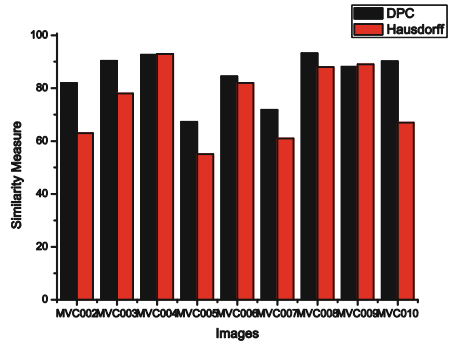


Fig. 6 Similarity comparison on dataset4

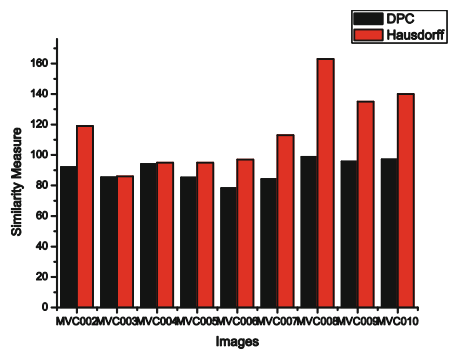
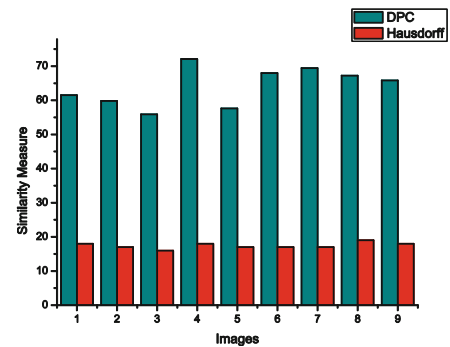


Fig. 7 Similarity comparison on dataset5

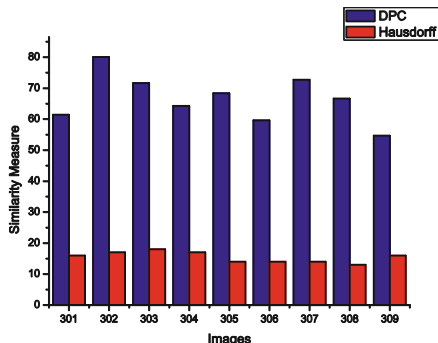


similarity) is approximately following the same trend whereas Hausdorff trend is uneven (Refer Dataset 1 and 2).

Figures 5 and 6 are again showing similarity pattern between images of similar object (in this case same female face, different orientation and moods). DPC is in accordance with Hausdorff (Refer Dataset 3 and 4).

Continuing the same pattern similarity index was calculated for Elephants and buses and is represented graphically in Figs. 7 and 8 It can be observed that DPC

Fig. 8 Similarity comparison on dataset6



provide more visual information than Hausdorff, according to DPC, image 0 (Dataset5) is 75 % similar to image 4 (Dataset5), and according to Hausdorff metric, distance between the two is nearly 15. It is evident that former result provide better insight to image similarity than latter one (Refer Dataset 5 and Dataset 6).

In all scenarios, performance of DPC is at par with the well-established Hausdorff Distance (HD) as it gives the precise percentage similarity between the two images, whereas all other distance metric just conveys the relative nearness or farness among different images. For example, for cases considered, the HD between images A and B is 140 and the corresponding DPC value is 90, whereas HD between A and C is 40 while the DPC value is 25. HD signifies that image A is more similar to image B than C and DPC value signifies that A and B are 90 % similar to each other, whereas images A and C are 25 % similar. Thus, DPC is a sound approach, because if we say two images are 80 % similar, for any two images if it holds, it will conclude the same semantic meaning.

6 Conclusion

The DPC algorithm, for the study of resemblance between perceptual objects is introduced in this article. Particularly, perceptual object resemblance is quantified by measuring percentage similarity between descriptions of sets of objects. Adopting DPC as image similarity approach, we can provide explicit control to user, to decide the granularity of image mapping. Also this approach is more semantically justified, as it gives better insight to the user by providing degree of similarity in terms of percentage. Experimental results with the DPC also suggest that it is a better approach than the existing well established Hausdorff metric.

Dataset [24, 25]

Dataset1.



Dataset2.



Dataset3



Dataset4



Dataset5.



Dataset6.



References

1. R. Datta, D. Joshi, J. Li, J. Z. Wang, Image retrieval: Ideas, influences, and trends of the new age, *ACM Computing Survey*, 40(2), 2008, pp. 1–60, 1348248.
2. N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique, Content based image retrieval using motif cooccurrence matrix, *Image and Vision Computing*, 22(14), 2004, pp. 1211–1220.
3. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 2000, pp. 1349–1380.
4. D. Huttenlocher, G. Klanderman, W. Rucklidge, Comparing Images Using the Hausdorff Distance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9), 1993, pp. 850–863.

5. M. Shapiro, M. Blaschko, Stability of Hausdorff-based Distance Measures, Proceedings of IASTED Visualization, Imaging, and Image Processing, Marbella, Spain, 2004.
6. C.H. Lin, R.T. Chen, Y.K. Chan, "A smart content-based image retrieval system based on color and texture feature, Image and Vision Computing", 27(6), 2009, pp. 658–665.
7. G.H. Liu, L. Zhang, Y.K. Hou, Z.Y. Li, J.Y. Yang, Image retrieval based on multi-texton histogram, Pattern Recognition, 43(7), 2010, pp. 2380–2389,.
8. C. P. Lee, Robust image segmentation using active contours: Level set approaches, Ph.D. dissertation, North Carolina State University, 2005.
9. M. A. Stricker, M. Oren, Similarity of color images, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2420, 1995, pp. 381–392.
10. T. Gevers, A.W.M Smeulders, Image Search Engines, An Overview, The International Society for Optical Engineering, 2003.
11. K. Yang, J. Trewn, Multivariate Statistical Methods in Quality Management, McGraw-Hill Professional, 1st Edition 2004.
12. P. N. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison-Wesley, 2005.
13. C. Spearman, The proof and measurement of association between two things, The American Journal of Psychology, 15(72–101), 1904, pp.72–101.
14. J.F Peters, General theory about nearness of objects, Applied Mathematical Sciences 1(53), 2007, pp. 2609–2629.
15. J.F Peters, Special theory about nearness of objects, Fundamenta Informaticae, 75(1–4), 2007, pp.407–433.
16. C.J Henry, S. Ramanna, Maximal clique enumeration in finding near neighbourhoods, Transactions on Rough Sets XVI, Springer Berlin Heidelberg, 2013, pp. 103–124.
17. C. Henry, J. F. Peters, Image pattern recognition using approximation spaces and near sets, Proceedings of the Eleventh International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computer, Joint Rough Set Symposium (JRS07), Lecture Notes in Artificial Intelligence, 4482, 2007.
18. C. Henry, J. F. Peters, Perception based image classification, Computational Intelligence Laboratory, University of Manitoba, UMCI Laboratory Technical Report No. TR-2009-016, 2009.
19. J. F. Peters, S. Ramana, Affinities between perceptual granules: Foundations and perspectives, Human-Centric Information Processing Through Granular Modelling, Springer-Verlag, 2008, pp. 409–436.
20. J.F. Peters, P. Wasilewski, Tolerance spaces: Origins, theoretical aspects and applications. Information Sciences 195, 2012, pp. 211–225.
21. A.B. Sossinsky, Tolerance space theory and some applications, Acta Applicandae Mathematicae: An International Survey Journal on Applying Mathematics and Mathematical Applications 5(2), 1986, pp. 137–167.
22. E.C Zeeman, The topology of the brain and the visual perception, Topology of 3-manifolds and selected topics, Prentice Hall, New Jersey, 1965, pp. 240–256.
23. C. Henry, G. Smith, Proximity Systems, UM CI Labs Technical Report Number TR-2012-021, University of Manitoba, Canada, 2014.
24. V. Jain, A. Mukherjee, The Indian Face Database, <http://viswww.cs.umass.edu/~vidit/IndianFaceDatabase>.
25. J. Z. Wang, J. Li, and G. Wiederhold, Simplicity: Semantics-sensitive integrated matching for picture libraries, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9), 2001, pp. 947–963, Database URL:<http://wang.ist.psu.edu/docs/related/>.

Robust Speaker Verification Using GFCC Based *i*-Vectors

Medikonda Jeevan, Atul Dhingra, M. Hanmandlu and B.K. Panigrahi

Abstract This paper presents to ameliorate the performance of text-independent speaker recognition system in a noisy environment and cross-channel recordings of the utterances. In this paper presents the combination of Gammatone Frequency Cepstral Coefficients (GFCC) to handle noisy environment with *i-vectors* to handle the session variability. Experiments are evaluated on NIST-2003 database.

Keywords Mel-frequency cepstral coefficient (MFCC) • Gammatone frequency cepstral coefficient (GFCC) • *i-vector*

1 Introduction

Human are able to recognize a speaker and differentiate it from other individuals independent of the text spoken. With the rise of Artificial Intelligence applications, we have tried to emulate these innate human abilities to our advantage. Intuition behind speaker recognition is to characterize the unique signature of the vocal tract. The major factors that inhibit accurate recognition in a non-ideal environment are Noise and Cross-Channel utterances. Over the last few years a lot of research has ensued in this area, of which three important methods have emerged.

Mel-Frequency Cepstral Coefficient (MFCC) features have been long used as standard in speaker recognition. In this method, a short segment of time-windowed speech segment is used to extract spectral information [1–3]. These features are efficient in speaker recognition system in clean speech environment and are sensitive in noisy environment.

To more accurately emulate a human auditory system, and in depth psychophysical observation of the auditory periphery was carried out, resulting in a standard model of cochlear filtering [4], known as Gammatone Filters. This filter bank models the basilar membrane as a set of band-pass filters. Gammatone Fre-

M. Jeevan (✉) · A. Dhingra · M. Hanmandlu · B.K. Panigrahi
Indian Institute of Technology, New Delhi 110016, India
e-mail: jeevanmedi@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_9

85

quency Cepstral Coefficients (GFCC) is used to alleviate the poor performance of MFCC in noisy or mismatched conditions.

Apart from noisy conditions, cross-channel recording or session variability of speaker utterances also a challenging research in robust speaker recognition systems. Session variability was handled efficiently by *i-vector* is being a state-of-the-art method was initially developed by Dehak et al. [5], and a lot of improvements have been made since. *i-vector* calculates a total variability statistics of an utterance, taking into account both user dependent and user independent characteristics. In this research we have studied and compared MFCC, GFCC and *i-vector* methods and their fusion. Most of the experiments in the literature [5] have used MFCC features while calculating *i-vector*. We have therefore extrapolated it to GFCC features while calculating *i-vector* because robustness of GFCC as compared to MFCC in noisy environment.

The rest of the paper is organized as follows. Section 2, describes the methodology. Section 3 will provide experimental results. Section 4 presents conclusions.

2 Methodology

The basic frame work of the methodology is as shown in Fig. 1.

2.1 Mel Frequency Cepstral Coefficients

For analyzing and processing a speech waveform, it needs to be converted into parametric representation. Basic pre-processing methods such as Voice Activity Detection (VAD) [6] are generally used before delving into feature extraction. It is

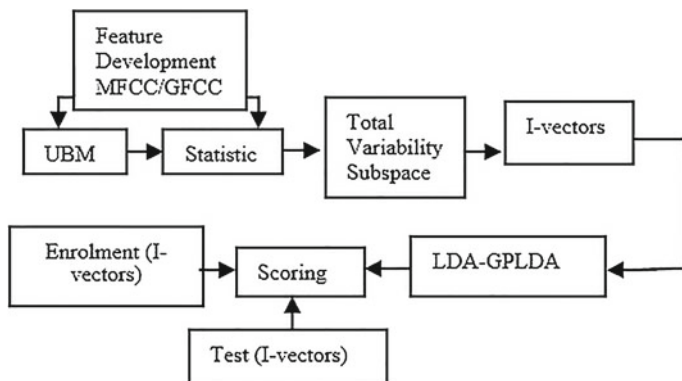


Fig. 1 Detailed frame work of the methodology

used to eliminate the unvoiced segments in speech signal. For our research we have used a simple energy based VAD.

A speech waveform has fairly stationary characteristics when examined over a sufficiently short period of time. MFCC is one of the most widely used methods for parametric representation of these characteristics. Cepstrum is the inverse Fourier transform of the logarithm of magnitude spectrum. Using cepstrum enables us to compute the similarity between two cepstral feature vectors to be computed by a simple distance metric such as Euclidean distance. MFCC are based on variation of human ear's critical bandwidth with linearly spaced frequency filters at low frequencies and logarithmically at high frequencies. This has been used to capture phonetically important characteristics of speech. The following MFCC processor is used for feature extraction.

A short 20 ms blocks of frames with 10 ms overlapping are obtained from entire speech signal and coefficients are calculated using feature extraction. To minimize the signal discontinuities and spectral distortion, hamming window $w(n)$ operation is used. Consequently Fast Fourier transforms (FFT) is applied to obtain the spectrum. The spectrum is wrapped in Mel-Frequency filter bank to obtain the mel-spectrum. Mel Filter banks are triangular filter banks whose center frequencies are logarithmically spaced given by

$$Mel(f) = 2595 * \log_{10}(1 + f/700) \quad (1)$$

Logarithm of the spectrum is calculated, followed by DCT operation, which results in cepstrum. The 0th coefficient of DCT is ignored as it contains no relevant information. Finally, cepstral liftering is used to perform low-pass filtering operation in cepstrum domain. This enables us to extract the vocal tract information, and at the same time suppressing the excitation information of the speaker utterance.

2.2 Gammatone Frequency Cepstral Coefficients

After the wide use of Mel Frequency Filters due to its similarity to ear model, a more comprehensive model was developed to mimic the characteristics of the ear based on psychophysical observations of auditory periphery [4]. Gammatone Frequency Cepstral Coefficient (GFCC) was developed by Patterson and Smith [7]. The Gammatone filter bank models the human auditory systems as a series of overlapping band-pass filters. The impulse response of each filter is given by the following equation,

$$g(t) = at^{n-1} e^{-2\pi bt} \cos(2\pi f_c t + \phi) \quad (2)$$

where, a is a constant (usually equal to 1), n is the order of the filter, Φ is the phase shift, f_c is the central frequency and b is the bandwidth. The central frequency and

bandwidth can be derived from filter’s Equivalent Rectangular Bandwidth (ERB), given by Eqs. 3 and 4.

$$ERB(f_c) = 24.7 \left(4.37 \frac{f_c}{1000} + 1 \right) \tag{3}$$

$$b = 1.019ERB = 1.019 \left(ERB(f_c) = 24.7 \left(4.37 \frac{f_c}{1000} + 1 \right) \right) \tag{4}$$

For a GFCC feature extraction process, speech signal is multiplied to the Gammatone filter bank in the frequency domain. Consequently this signal is again reverted back to the time-domain using Inverse Fourier transform. The signal is then decimated to 100 Hz to reduce the effects of noise. Absolute value of the signal is taken and rectified using a non-linear process. In GFCC we use cubic root operation instead of logarithmic operation as in the case of MFCC. Finally, DCT is applied to obtain GFCC features. Please note that, we cannot ignore the 0th term of DCT in this case, as we have used cubic root rectification which cannot split multiplicative terms into addition as in the case of a logarithmic operation (Figs. 2 and 3).

Fig. 2 Block diagram of MFCC feature extraction

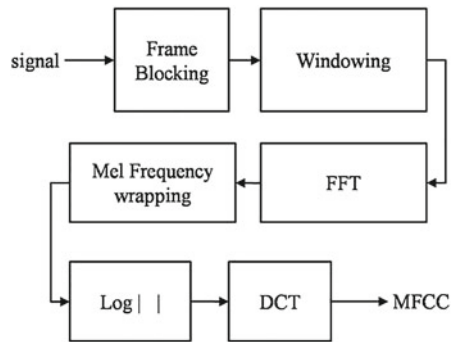
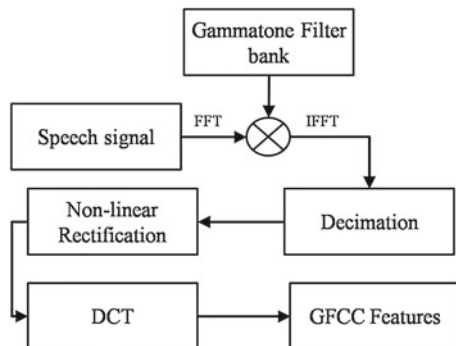


Fig. 3 Block diagram of GFCC feature extraction



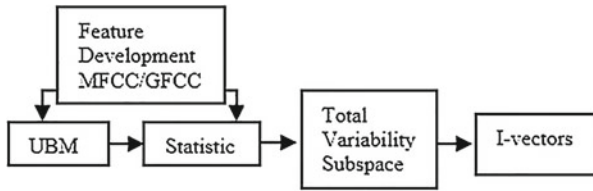


Fig. 4 Block diagram of I-vector extraction

2.3 *i*-Vector

i-vector approach makes use of a set of low-dimensional total variability factors (w) to represent each conversation side. To model the total variability statistics, each factor controls an Eigen-dimension of total variability matrix (T). For calculating universal Background Model (UBM) [8], first-order Baum-Welch statistics are used. Figure 4 gives a high level description of *i*-vector feature extraction.

$$M = m + Tw + \epsilon \quad (5)$$

where m , is event-independent GMM, w is the low dimensional *i*-vector, and ϵ is the residual not represented by the earlier two terms. The model and test segments are represented as *i*-vectors for classification. Linear discriminant analysis (LDA) is used for dimensionality reduction of *i*-vectors, so as to annihilate the non-speaker related subspace, and thus increasing the discrimination between speaker subspaces. Before modeling the dimensionality reduced *i*-vectors, they are mean and length normalized via a generative factor analysis approach called the probabilistic LDA (PLDA). In our work we have leveraged on the fact that GFCC is a better feature model as compared to MFCC. We therefore combined it with *i*-vectors to achieve a better verification rate with GFCC features.

3 Experiment and Results

We have evaluated the performance of the proposed models on the fixed set of 100 speakers randomly chosen from the total 356 speakers of NIST-2003 [9] database involving different mismatch conditions. Zhao et al. [10] have categorically compared the performance of GFCC over MFCC in identification. In our research we extrapolated the same effects of GFCC over MFCC, in attempts to prove that GFCC features out performs MFCC features in speaker verification as well using *i*-vectors.

Table 1 shows the Equal Error Rate (EER) for MFCC based *i*-vectors (*i*-mfcc) and GFCC based *i*-vectors (*i*-gfcc) under different noise environments with SNRs from -10 db to 20 db. At low SNR, *i*-gfcc performs comparable with *i*-mfcc which

Table 1 k-fold average EER (%)

Noise	SNR(db)	EER (%)	
		i-mfcc	i-gfcc
White	-10	45.25	46.88
	-5	40.17	38.50
	0	35.17	28.64
	5	26.31	17.26
	10	22.13	9.99
	15	17.52	8.06
	20	17.48	7.12
Factory	-10	47.65	47.14
	-5	44.53	42.68
	0	36.92	32.53
	5	31.08	18.84
	10	26.33	10.02
	15	21.18	8.94
	20	19.90	6.71

Fig. 5 EER graph for -5 db white noise

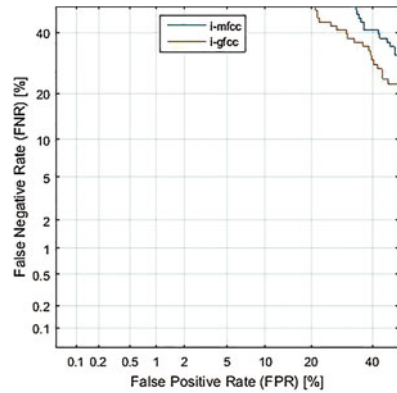
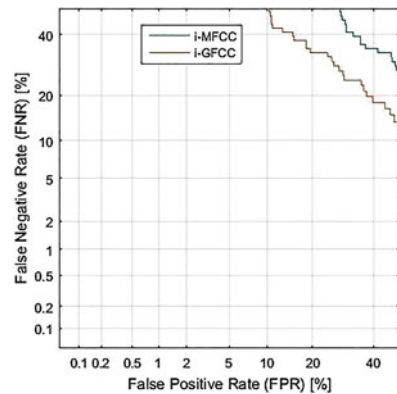


Fig. 6 EER graph for 5 db white noise



can be observed from Fig. 5. As the SNR is increased it is observed that *i*-gfcc outperforms *i*-mfcc, as evident in Figs. 5 and 6.

4 Conclusions

As GFCC is a far better approximation of the human auditory systems as compared to MFCC, it is evident in literature that GFCC outperforms MFCC in speaker identification. This paper shows that GFCC features using state-of-the-art *i*-vectors produces better speaker verification results as compared to the *i*-vector based on MFCC features.

Acknowledgment The present study is a part of ongoing project on “Personal Authentication using Multimodal Behavioural Biometrics: Voice and Gait” and the authors express their gratitude to the Department of Science & Technology, Govt. Of India for funding the project.

References

1. D. A. Reynolds, “Experimental Evaluation of Features for Robust Speaker Identification,” *IEEE Trans. on Acoustic Speech and Audio Processing*, vol. 2, no. 4, 1994.
2. Davis, S. Mermelstein, P.,” Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences”, In *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 28 No. 4, pp. 357–366, 1980.
3. H. Seddik, A. Rahmouni and M. Sayadi, “Text independent speaker recognition using the mel frequency cepstral coefficients and a neural network classifier”, *Ecole Nationale Des Sciences Informatiques*, 2010, Manouba, Tunisia.
4. R.D Patterson, I. Nimmo-Smith, J. Holdsworth, and P. Rice, “An efficient auditory filterbank based on Gammatone function,” in Paper presented at a meeting of the IOC Speech Group on Auditory Modelling at RSRE, December 14–15, 1987.
5. N. Dehak, R. Dehak, P. Kenny, N. Brummer, P. Ouellet, and P. Dumouchel, “Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification,” in *Proceedings of Inter speech*, Brighton, UK, 2009.
6. T. Kinnunen and P. Rajan. “A practical, self-adaptive voice activity detector for speaker verification with noisy telephone and microphone data,” in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, 2013, pp. 7229–7233.
7. R. Patterson and I. N. Smith, “An efficient auditory filter bank based on the gammatone function,” *Speech-Group meeting of the Institute of Acoustics on Auditory Modelling*, vol. 54, Apr 1987/FLEXChip Signal Processor (MC68175/D), Motorola, 1996.
8. D. Povey, S.M. Chu, B. Varadarajan, Universal background model based speech recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.
9. M. Przybocki, A. Martin, and A. Le, “NIST speaker recognition evaluations utilizing the mixer corpora 2004, 2005, 2006,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 7, pp. 1951–1959, Sep. 2007.
10. Zhao, Xiaojia et Wang, DeLiang. Analyzing noise robustness of MFCC and GFCC features in speaker identification. In: *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 *IEEE International Conference on*. IEEE, 2013. p. 7204–7208.

Enhanced Automatic Speech Recognition with Non-acoustic Parameters

N.S. Sreekanth and N.K. Narayanan

Abstract A novel method for improving the accuracy of automatic speech recognition system by adding non-acoustic parameters are discussed in this paper. The gestural features which are commonly co-expressive with speech is considered for improving the accuracy of ASR system in noisy environment. Both dynamic and static gestures are integrated with speech recognition system and tested in various environmental conditions, i.e., noise levels. The accuracy of continuous speech recognition system and isolated word recognition system are tested with and without gestures under various noise conditions. The addition of visual features provides stable recognition accuracy under different environmental noise conditions for acoustic signals.

Keywords Audio-visual speech recognition • Non-acoustic features for speech recognition

1 Introduction

Studies on automatic recognition of speech by machine is a topic of interest for researchers for many decades. During the initial period, the speech recognition was considered as a mere signal processing problem where the acoustic features are analyzed to recognize the speech. Latter stages rather than considering the speech as mere acoustic signals, researchers started investigating the correlation of speech signals with underlying language. This results, the building automatic speech recognition systems by combining both acoustic and language features. Stochastic based N-gram models are built to compute the next probable word which user

N.S. Sreekanth (✉)

C-DAC Bangalore, #68 Electronics City, Bangalore 560100, Karnataka, India

e-mail: nss@cdac.in

N.K. Narayanan

Department of Information Technology, Kannur University, Kannur, Kerala, India

e-mail: nkrarayanan@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_10

93

might have spoken. The feature vector comparison will be done against the most probable words, and if it matches system return the result. If the variance is not within the acceptable limit, system will go for next higher one. Automatic speech recognition system build by combining acoustic and language models work well in the lab-conditions or indoor environment, i.e. either noise free or less noisy environment. The performance of these systems are not acceptable in the noisy environment (noisy for acoustic signals) so deploying the practical solutions at public places may not be really reliable. The efforts for improving the speech recognition accuracy by machine have mainly two different directions. The first one is implement anti-noise cancellation techniques, as part of signal acquisition and feature extraction module i.e., FrontEnd module of the ASR systems [1-3]. The second method is to extract acoustic independent features, but depends on speech; model such parameters as part of feature vector. When we communicate each other we use gestures, co-expressed with. This motivates the researchers to investigate the use of acoustic independent parameters to improve the accuracy of speech recognition system. Context dependent visual parameters are extracted from the source and model it suitably for recognition application. Understanding and modeling the lips shapes and its movement corresponding to speech sound has been widely investigated by researches and fair results are reported for recognizing the speech in noisy environment and also recognizing the distorted speech. Lips modeling is being a popular acoustic independent parameter used for speech recognition and even for speaker recognition [4]. The approach for integrating the visual features for improving the accuracy of speech recognition system is depends on the type of application or task that user would like to perform. The recognized gestures are generally substitute or interpolate a missing word or it can be a redundant information in the communicated speech. Some time it can even be a anaphora equivalent in the recognized speech. For example if we says "Please take that", where the instruction will be complete only if the user point to the object to be taken. So here the pointing gesture is expected to replace the word "that" in the speech based communication. The approaches for integration of visual features with ASR system for isolated word recognition differ from the continuous speech recognition system. In both the context the gestures are equivalent to a word or a phrase in speech. In word recognition system it can have a bijective relationship between speech and gestures, it means we expect one-to-one correspondence for each word in speech and gestures. Digit recognition systems, system used for controlling devices, robotic control environment are the best example for this. Whereas for the continuous speech recognition system a bijective relationship between the individual words in the communicated speech with gesture is not guaranteed always. In continuous speech recognition system we use gestures for certain word or phrase in continuous speech. For example when we call someone we says "*could you please come here*" and shows the gesture to call him with hand. In this gesture is equivalent to the phrase "*come here*". The acoustic features extracted from the speech will be used for recognizing the word or phones as a primary step and this will be validated with the gestural features if present and confirms the recognized string in case of isolated word recognition system. For continuous speech

recognition system, the primary recognition of acoustic signals will be done using the acoustic features combined with language models. The resulted word will be validated with the string i.e., word returned by the gesture recognition system. In this study we have implemented the speech recognition system with conventional MFCC based parameter along with the time domain based phase space point distribution features [5]. The details of gesture recognition are discussed in the latter part of this paper. In this paper we discuss about the performance improvement of automatic speech recognition system by adding visual features, i.e., hand gestures which are co-expressed with speech. The gestures are used as an augmenting or supporting modality for recognizing speech. Hand based gestures are recognized and fused with speech recognition system for improving the accuracy. Both static and dynamic gestures are experimented separately. This paper is organized as follows. Review of previous work is discussed in section two. Section three discuss about the visual features used for improving the speech recognition accuracy. Algorithm for isolated word recognition and continuous speech recognition with visual features are discussed in section four and five respectively. The experimental settings and results are discussed in section six.

2 Review of Previous Work

Adding acoustic independent features for improving the speech recognition accuracy is introduced three decades ago. The work done by [6, 7] introduced visual information by the use of lip information as an important aid for speech recognition [6, 7]. Kittler et al. [8] presented a study using geometric features of the lip shapes from model based lip boundary tracking confirming the importance of lip information in identity recognition [8]. Yamamoto et al. [9] proposed visual information semi automatically mapped to lip movements through the aid of sensors put around the mouth to highlight the lips. The experimental results showed that significant performance could be achieved even by only using visual information [9]. Another interesting study done by Heracleous P. et al, developed HMM based model for cued speech recognition. Lip and mouth movement based geometrical feature extracted through image processing techniques provided and effective method for recognizing speech in noisy environment. The extracted visual information will act as a augmenting feature for recognizing speech. If we look at human-human way of communication, we use gestures or other forms of visual artifacts along with speech as an augmented or support system. Especially in noisy environment e.g. inside factory we, extensively use the gestures in addition to speech for communicating with others. In such context it is found that gestures play a very important role for interpolating the missing part of speech. Providing gestures via hand, head, eyes or with other external object is very popular in human-human interaction. The visual parameters form these gestures can be extracted through computer vision (image processing) techniques or sensor based techniques become popular among the researchers in human-computer interaction for building interface technologies.

Considering these gestures i.e. visual input as complement to missing or a noisy speech signals to recognize by machine is also an interesting topic to research for improving the accuracy of speech recognition system. Work done by Neti C et al. at IBM labs reports with a effective audio visual fusion techniques for building large vocabulary continuous speech recognition system. Audio-only recognizer, introducing the visual modality reduced ASR word error rate by 7 % relative in clean speech, and by 27 % relative at an 8.5 dB SNR audio condition [10]. The extraction of vocal tract construction gestures fused with acoustic parameters experimented and reliable results were reported [11]. Hand gesture and speech based interaction techniques are experimented in the field of control system and robotic environment for controlling machines [12]. Along with gestures, online-handwriting based techniques are also experimented and hopeful results are reported across literatures [13]. The visual patterns also used for repairing the speech signal, form video information [14]. The increased performances of recognition accuracy of above systems are at the cost of increase in computational complexity. In this paper we report a computationally simple architecture for fusion of visual features for improving the accuracy of speech recognition for isolated and continuous speech recognition system.

3 Visual Features Used for Speech Recognition

As discussed earlier, the gestures are co-expressive with speech when human communicate each other. We have used the hand gestures as part of this experiment. The gestures produced by hand during the communication may be either static or dynamic. Static gestures are the gestures where the movement of body part does not involve. For example when we say “three” and also show three fingers while speaking is treated as static gesture. For dynamic gestures, the body part also moves and the orientation and direction of movement of the object take part in the gesture production determine the gesture produced. User says “*go to next page*” and shows gesture of moving the hand from left to right is a good example for dynamic gesture. Similarly user can say a number or a letter and can write in the air with finger. The movement of finger in the air can be tracked and feature vector can be generated. This types of gestures also considered as dynamic gesture. Static gestures are recognized using the orientation and geometrical properties of hand and fingers. Skin colour based segmentation is used to segment the hand form the complex image [15]. The convex hull algorithm returns the enclosed polygon of the hand and along with the fingers orientation. Various geometrical features like area, perimeter, centroid, solidity along with histogram distribution is used for recognizing the static gestures [16]. For recognizing dynamic hand gestures we have used the skin color based segmentation algorithm for segmenting the hand form the complex image frames. The finger tips will be identified form the segmented frames using the convex hull algorithm. The movement of finger will be tracked and the feature vector will be generated using Freeman’s eight directional codes (popular algorithm used for online handwriting recognition) [17]. A dynamic

time wrapping based Leventian Minimum Edit distance algorithm is used for classification of incoming gestural features.

4 Isolated Word Recognition System Using Visual Features

We have enhanced the conventional speech recognition system by incorporating the visual features for improving the accuracy. The input signal from the microphone will be given to the speech recognition module. The acoustic feature vector i.e., 12 MFCC coefficients with one energy parameter and whose first and second derivatives are extracted along with 20 sized phase space point distribution parameter from the speech frames. If user shows gestures corresponding to spoken word with hand, that will also be captured and relevant features will be extracted. Both the speech and gesture recognition systems are independently recognized the word corresponds to speech and gesture respectively. The conclusion of results will be done by the system as follows. Let us assume that user spoke a word S which is a part of the vocabulary W where $S \in W$ and corresponding gesture also shown. The feature vector corresponds to the input speech signal will be extracted and the distortion vector with all word in the vocabulary with respect to the incoming signal will be calculated. Similarly the distortion vector for gesture also will calculated. The distortion vector return by both speech and gesture are normalizing to a scale of 0 to 1. Let $d_{s1}, d_{s2}, d_{s3} \dots d_{sn}$ are the respective distortion vectors for $W_1, W_2, \dots W_n$. Against the input speech signal. So the list of words which will be returned by speech recognition module will be $\langle w_1, d_{s1} \rangle, \langle w_2, d_{s2} \rangle, \dots \langle w_n, d_{sn} \rangle$, Similarly the gesture also will return the distortion vector corresponding to the elements in the data base if user issues corresponding gestures. $\langle W_1, d_{g1} \rangle \langle W_2, d_{g2} \rangle \dots \dots \dots \langle W_n, d_{gn} \rangle$

The result return by speech recognition engine W_i

$$W_i = \forall_i \{ \text{return}(i) \text{ with } \{ \min \langle W_i, d_{si} \rangle \} \}$$

The result return by gesture recognition engine is W_k

$$W_k = \forall_k \{ \text{return}(k) \text{ with } \{ \min \langle W_k, d_{gk} \rangle \} \}$$

If $i = k$ i.e. both the recognizer unit return the same word that result will be more authenticated. If $i \neq k$ i.e., system returns two different word index then the decision will be taken by finding the difference between the speech distortion vector and gesture distortion vector so that the difference is more than predefined threshold value T then select the minimum form resultant vector as follows

$S = \forall_i (\text{Min} \{ \langle W_i, d_{si} \rangle - \langle W_k, d_{gk} \rangle \})$ if $|d_{si} - d_{gk}| > T$, else suggest both words as probable words.

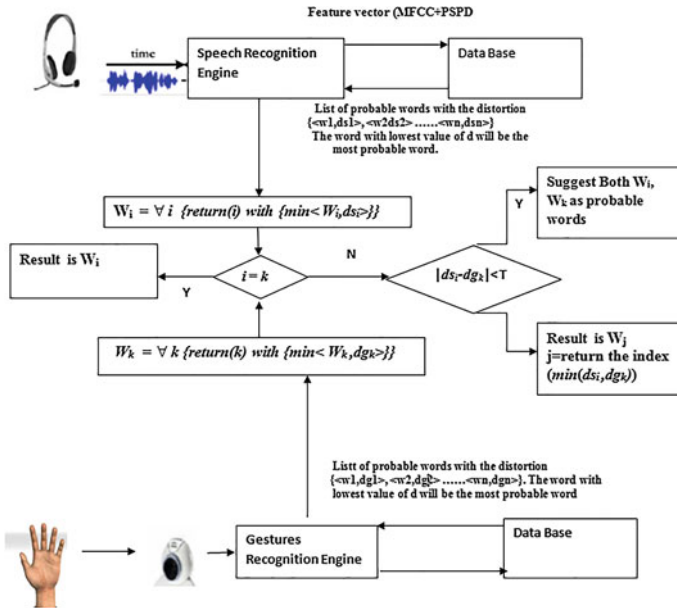


Fig. 1 Isolated word recognition system with visual features

System diagram for audio visual based isolated word recognition is shown in Fig. 1.

5 Continuous Speech Recognition System Using Visual Features

The method for recognizing speech by adding additional features is more or less same for isolated word recognition and continuous speech recognition system. The major difference is the probability choosing a word corresponding to input sequence form microphone (for speech) or form camera or sensor (for gestures), have a strong correlation with language models in case of continuous speech recognition system. whereas the probability of occurrence of any word in the given vocabulary are equally likely for an isolated word recognition. In continuous speech recognition system stochastic based N-gram models are used for predicting the next probable word given that certain words are already recognized. N-gram models are popular language modals widely used for continuous speech recognition system [18]. So here for recognizing the gestures corresponding to a word can also be correlated to the language model build as part of the system.

Let O_s be the acoustic observation sequence generated by FrontEnd module for speech based input and O_g be the visual observation sequence generated by the FrontEnd module corresponds to gesture based system, For a given context the probability of recognizing a given observation as w_j by speech recognizer, w_k by gesture based module is calculated as follows, assume that input via both methods are present [18].

$$P(w_j|O_s) = \left[\arg \max_{w_j \in L} \right] \frac{P(O_s|w_j)P(w_j)}{P(O_s)}$$

Similarly for gesture based interaction

$$P(w_k|O_g) = \left[\arg \max_{w_k \in L} \right] \frac{P(O_g|w_k)P(w_k)}{P(O_g)}$$

The word return by language model can be calculated as, the probability of choosing a word W_i given that $W_{i-1} W_{i-2} \dots W_{i-N+1}$ is already occurred where N is number of words looking backward to predict the next word W_i based on N-gram calculated from the training corpus. So W_i will be the word return by language model.

So The resulted word W_R can be decided from the individual results returned by different recognizers and language models. $W_R \in (W_i, W_j, W_k)$

The choosing of the resulted word as W_i or W_j or W_k depends on the probability value returned by individual recognizer units and language model. If $i = j = k$, the system returns one of the index as word are same. If $i \neq j \neq k$ an additional rule can be imposed so that the probability value for every word index returned by recognizers are less than a predefined threshold τ , can be avoided and system will return the null string. An algorithm for continuous speech recognition system with visual features is shown in Fig. 2.

6 Experimental Setting and Results

The evaluation of isolated word recognition system and continuous speech recognition system are carried out separately. The performance of the system is evaluated with and without gestural features under different environmental conditions. The equal number of both dynamic and static gestures are tested for digits and isolated word. For continuous speech recognition system, the static and dynamic gestures are used alternatively with equal numbers. For the isolated word recognition system the overall recognition accuracy of the system under different environmental conditions with gestures and without gestures are tabulated. The standard evaluation metric for a continuous speech recognition system is the **word error rate**. The word error rate (WER) for a continuous speech recognition system is generally calculated using the minimum edit distance between the hypothesis word and correct string. The word error rate with and without gestural features are tested separately under

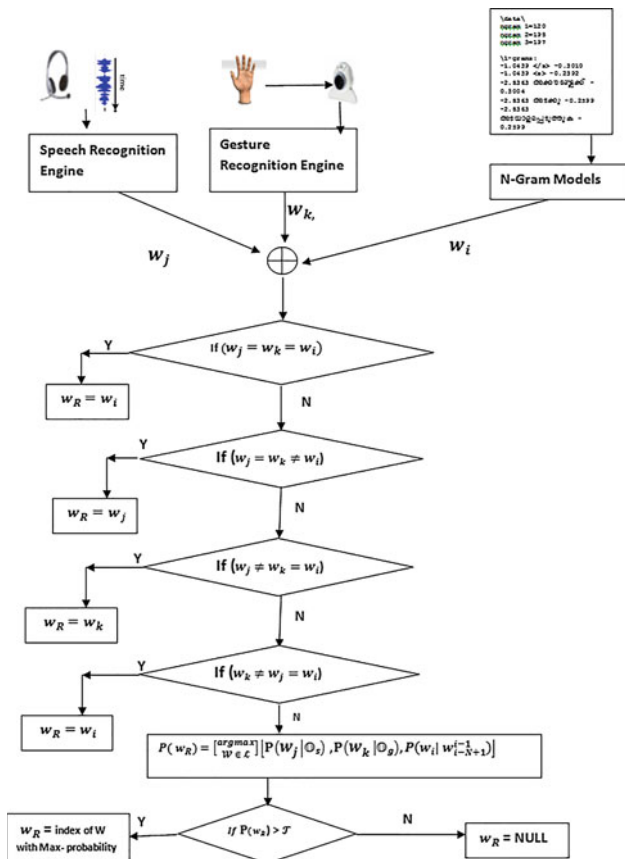


Fig. 2 Continuous speech recognition system with visual features

different noise conditions. The environmental noise is measured using the standard software called sound meter. The training data base is created with speech and gesture data accrued form 20 different people (12 male and 8 female) within the age group of 30–40 with 100 samples per each word for isolated word recognition system and 50 sentences for continuous speech reorganization system. A total of 30 isolated words are the vocabulary size for isolated word recognition system which includes the ten digits and 20 operational keywords. Both static and dynamic gestures are integrated separately with speech recognition system and results are evaluated. The accuracy of isolated word recognition system is tested under various noise level, sound proof rooms (20–30 dB), Lab and Quiet rooms without fan 30–40 dB, 40–50 dB at Office, 50–60 dB at open park, 60–70 dB street and outside Office surroundings during day time, 70–80 dB on Busy traffic 80 dB and above factory environment are tested for measuring the performance. The recognition accuracy for isolated word recognition system is tabulated in Table 1.

Table 1 Average recognition accuracy of the isolated word recognition system under different noise conditions without gestures and with gestures

Environmental noise		20–30 dB (%)	30–40 dB (%)	40–50 dB (%)	50–60 dB (%)	60–70 dB (%)	70–80 dB (%)	80 dB and above (%)
Recognition accuracy	Without gesture	88	83	80.2	72.2	62	43	18
	With gesture	89.2	88	88.1	87	87.3	86	85

Table 2 Word error rate of continuous speech recognition system under different noise conditions without gestures and with gestures

Environmental noise		20–30 dB (%)	30–40 dB (%)	40–50 dB (%)	50–60 dB (%)	60–70 dB (%)	70–80 dB (%)	80 dB and above (%)
Recognition accuracy	Without gesture	11.4	12.1	22.3	24.8	28.7	33	70.1
	With gesture	9.2	9.4	10	10.3	14.5	17	21

For continuous speech recognition system N-gram language models are created for the sentences. The vocabulary is restricted to the command related to basic computer operations. Some example sentences are “*Open this file*”, “*Close this picture*”, “*copy this file to that folder*”, “*paste this*”, “*move this file from this folder to my document*” etc. The training set is also created for static and dynamic gestures corresponding to various operational keyword/commands. The performance of continuous speech recognition system is tested with gestures and without gestures under various noise conditions same as that discussed for isolated word recognition. The word error rate under different conditions for 40 random sentences selected from the data base with gesture and without gestural features are tabulated in Table 2.

From the above results it is found that addition of visual features provides a more or less stable recognition accuracy under different environmental noise conditions for acoustic signals. For continuous speech recognition system, the visual feature may help to recognize the prominent word under high noise conditions, and less relevant part of the communication can be neglected. For example user says “*can you please delete this file*” and user shows the gesture corresponding to delete and the file icon to be deleted, the system can still recognize the command issued under noisy conditions. The words like “*can you please*” are not prominent at that context.

7 Conclusions

Enhanced versions of conventional speech recognition system by fusing the non-acoustic feature vectors are discussed in this paper. The hand gesture features are extracted which is co-expressive with speech are used for improving the accuracy of the speech recognition. The approach and method used in this paper for improving the recognition accuracy is to be seen objectively, i.e., no matter the communicated message is recognized from speech recognizer or gesture recognizer but the purpose of communication need to be met. Hence the proposed method provides a natural and effective interaction mechanism for interacting with machines.

References

1. Dong Yu, Li Deng; Droppo, J.; Jian Wu; Gong, Yifan; Acero, A. “A minimum-mean-square-error noise reduction algorithm on Mel-frequency cepstra for robust speech recognition” Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on DOI: [10.1109/ICASSP.2008.4518541](https://doi.org/10.1109/ICASSP.2008.4518541). pp. 4041–4044.
2. Wouters, Jan; Vanden Berghe, Jeff “Speech Recognition in Noise for Cochlear Implantees with a Two-Microphone Monaural Adaptive Noise Reduction System”- Ear & Hearing: Journal of American Auditory society. October 2001 - Volume 22 - Issue 5 - pp 420–430.

3. Willie Walker, Paul Lamere, Philip Kwok, Bhiksha Raj, Rita Singh, Evandro Gouvea, Peter Wolf, Joe Woelfel, "Sphinx-4: A Flexible Open Source Framework for Speech Recognition" White paper -SMLI TR2004-0811 c2004 SUN MICROSYSTEMS INC.
4. Maycel Isaac Faraj, Josef Bigun, "Lip Motion Features for Biometric Person Recognition" Book chapter of Medical Information Science Reference, IGI Global, Chapter XVII, pp. 495–532. Year 2009.
5. P.Prajith, "Investigations on the applications of dynamical instabilities and deterministic chaos for speech signal processing", Ph.D Thesis, University of Calicut 2008.
6. Petajan, E. (1984). Automatic lipreading to enhance speech recognition. Global Telecommunications Conference. (pp. 265–272).
7. Mase, K., & Pentland, A. (1991). Automatic lip-reading by opticalflow analysis. *Systems and Computers in Japan*, 22(6), 67–76.
8. Kittler, J., Li, Y., Matas, J., & Sanchez, M. (1997). Combining evidence in multimodal personal identity recognition systems. *Proceedings of the First 48 International Conference on Audio- and Video-Based Biometric Person Authentication*, LNCS 1206, (pp. 327–334).
9. Yamamoto, E., Nakamura, S., & Shikano, K. (1998). Lip movement synthesis from speech based on hidden markov models. *Journal of Speech Communication*, 26(1), 105–115.
10. Neti, C Potamianos, G.; Luetin, J.; Matthews, I.; Glotin, H.; Vergyri, D." Large-vocabulary audio-visual speech recognition: a summary of the Johns Hopkins Summer 2000 Workshop ", IEEE Fourth Workshop on Multimedia Signal Processing, 2001, pp. 619–624.
11. Mitra, V; Hosung Nam; Espy-Wilson, C.Y.; Saltzman, E.; Goldstein, L"Gesture-based Dynamic Bayesian Network for noise robust speech recognition", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011. pp. 5172–5175, IEEE-DOI [10.1109/ICASSP.2011.5947522](https://doi.org/10.1109/ICASSP.2011.5947522).
12. Ze Lei; Zhao Hui Gan; Min Jiang; Ke Dong "Artificial robot navigation based on gesture and speech recognition", International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), 2014, pp. 323–327, IEEE DOI [10.1109/SPAC.2014.6982708](https://doi.org/10.1109/SPAC.2014.6982708).
13. Wu-chun Feng "An integrated multimedia environment for speech recognition using handwriting and written gestures", *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, 2003. IEEE DOI: [10.1109/HICSS.2003.1174293](https://doi.org/10.1109/HICSS.2003.1174293).
14. Lei Chen; Harper, M.; Quek, F. "Gesture patterns during speech repairs", *Proceedings of Fourth IEEE International Conference on Multimodal Interfaces*, 2002. pp. 155–160, DOI [10.1109/ICMI.2002.1166985](https://doi.org/10.1109/ICMI.2002.1166985).
15. Lei Yang, Hui Li, Xiaoyu Wu, Dewei Zhao, Jun Zhai. — An algorithm of skin detection based on texture|. *IEEE Image and Signal Processing (CSIP)*, 2011.
16. Noor Adnan Ibraheem, RafiqulZaman Khan "Survey on Various Gesture Recognition Technologies and Techniques", *International Journal of Computer Applications* (0975–8887), Volume 50 – No.7, July 2012, pp. 38–44.
17. B.J Manikandan, Gowri Shankar, V Anoop, A Datta, V S Chakravarthy: LEKHAK: A System for Online Recognition of Handwritten Tamil Characters. *Proceeding of the International Conference on Natural Language Processing (ICON-2002)* Vikas Publishing House Pvt. Ltd. pp. 285–291.
18. Daniel Jurafsky and James H. Martin "Speech and Language Processing", Prentice Hall, Englewood Cliffs, New Jersey 07632, 2000.

Dynamic Gesture Recognition—A Machine Vision Based Approach

N.S. Sreekanth and N.K. Narayanan

Abstract Computationally simple method for dynamic hand gesture recognition is presented in this paper. The segmentation of hand which take parts in gesture production is being addressed tow different ways using color band based segmentation algorithms. The first one uses the special color stickers as part of finger and the second one does segmentation based on normal skin color. The movement of hand is tracked and Freeman's eight directional code is generated corresponds to each gestures. A dynamic time wrapping based Levenshtein minimum edit distance algorithm is used for classification. The results of dynamic hand gestures with special color approach and without special color are discussed separately. The accuracy of the system is found to be more for special colour based segmentation than skin colour based segmentation techniques.

Keywords Gestures recognition • Dynamic gestures

1 Introduction

Gestures are considered as natural and intuitive method for communication which generally co-exist with speech. The gestures can also have independently express especially when a person is unable to talk because of disability or situational impairments, for example in noisy environment people depend on gesture for communication rather than speech. For past few decades, gestures are evolved as an alternative input mode to the keyboard and mouse interactions for most of the application domains. Gesture are defined as nonverbal mode of communication and it's symbol set are quiet large compared to symbol set of formal language like

N.S. Sreekanth (✉)

C-DAC Bangalore, #68 Electronics City, Bangalore, Karnataka, India

e-mail: nss@cdac.in

N.K. Narayanan

Department of Information Technology, Kannur University, Kannur, Kerala, India

e-mail: nknarayanan@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_11

105

English, Hindi, Malayalam, etc. When human communicate each other we use various gestures which are produced with hand or head (face, eyes, lips, or head itself) and will have specific meaning during the context. Automatic gesture recognition addressed in various way starting from hardware based solution to machine vision based soft solution. If we look at the evolution of gesture based interaction, initially there were glove based devices, but they lacked the naturalness factor as they had introduced an additional hardware constraints on the user [1–3]. The main advantage of these approach are high accuracy and fast reaction speed but this approach can be quite expensive. The modern techniques employed for gesture processing are either 3D models or image based processing. The former lacks the computational efficiency and the simplicity compared to other. In the image based processing method there are several techniques based on color, contour and correlation for identifying gestures [1]. In this paper we have used the image based processing and follows the algorithms and procedures of Machine vision a.k.a computer vision.

There are various ways to classify the gestures. One of the important aspects of classifications of gestures is based on the movement of the object take part in gesture production. In that basis there are two types of gestures, static gestures and dynamic gestures. Static gestures communicated via some symbols or fixed orientation of human body/body parts (hand, head) in particular manner. Generally sign languages like ASL (American Sign Language) are the subset of the static gestures [4]. The dynamic gestures as its name implies it involves a moving body parts, i.e., hand, head, leg etc. Example To call someone we may use the hand/fingers will keep in horizontal position and wave it in the air. In this paper we manly focuses on dynamic hand gesture recognition. The recognition or segmentation of hand from a complex background and track the movement of hand and generate the appropriate feature vector are two complex task as far as computer vision based gesture recognition is concerned. The color band based segmentation algorithms are used for segmenting the hand from the complex background image. Two different approaches are studied and simulated. The first method uses special color stickers which are affixed on fingers. The assumption is the special color stickers will not be a part of background. BLOB coloring algorithm is used for segmenting the finger i.e., special color. The movements of fingers are tracked and corresponding feature vector will be generated. In another method which does not uses any special color but that also uses the color band technique to segment the hand from the complex background. The segmentation will be performed using the BLOB coloring method using RGB equivalent of skin color band. Once the hand is segmented the movement of hand will be tracked and feature vector will be generated. The Freemans chain code features are extracted which is one of the popular method used for online handwriting recognition techniques. The classification of gestures is done using the Levenshtein Minimum edit distance algorithms. This paper is organized as follows. Section two discuss the review of previous work in the area of dynamic gesture recognition. The overall architecture of dynamic gesture recognition system and the details about dynamic gestures are discussed in section three and four. The hand segmentation and Feature vector extraction are

discussed in section five and six respectively. Section seven describes the gesture analysis and recognition with classification algorithm. Section eight discuss the experimental details and results.

2 Review of Work

Hand gestures are most power full gestures for human-human interaction. Computer vision based hand tracking and gesture recognition is extremely challenging problem due to the complexity of hand gestures [5]. Various studies are reported across literature for recognizing the dynamic gestures extracted from human communications. The recognition of dynamic gestures by machine generally fall into two categories, they are hardware/sensor based gesture recognition and soft computing based techniques, it is also known as machine vision based method. In this paper we are more focused on the machine vision based techniques for extracting the gestures. Quite a few studies and results reported in this area really motivate researchers to integrate the gesture based applications as part of human computer interaction. The studies carried out by Hairong Jiang et al., reports the experimental experience of using a robot to perform the chemistry laboratory experiments which is controlled through gesture based interactions. For gesture encoding, dynamic motion models were created employing the dynamic time warping method [6]. Support vector machine and HMM models are also used for dynamic hand gesture recognition and reliable results are reported [6]. Highly reliable practical systems are developed through color band models were the standard colors are used on the finger tips or hands to make the object detection easy from the complex scene for gesture recognition. The main disadvantage of this method is the designated color used for recognition should not be present as a background or as part of the scene [7]. The use of bare hand or finger is preferred over color band based recognition techniques so as to ensure a generic method for recognizing the gestures and interaction. Skin color based segmentation algorithms are popular among the researches who work in the area of machine vision based static gesture recognition technique [8, 9].

The extraction of appropriate features vector with suitable classification algorithm also plays major role in dynamic hand gesture recognition. Mixtures of spatio-temporal features are highly recommended by the literature. The work done by Nishikawa et al. reports a technique that uses optical flow from monocular motion image for recognition of human gestures based on rate of change of gesture motion and direction [10]. Another interesting work reported by Feng-Sheng Chen et al., where they use the motion analysis to characterize the temporal features and spatial features are modeled using Fourier descriptor (FD) [11]. In this method the frame differences are calculated and moving region is segmented form the consecutive video frames. This method is computationally intensive as it needs to process every incoming frame to capture the movements of hand at every context. In this paper we have used a computationally simple technique which is widely

applied in online handwriting techniques. Freeman's Eight Directional Chain code algorithm is used for generating the feature vector. The movement of hand will be converted to directional vector based on the direction of movement from the current point [12].

Once the feature are extracted it is found that researchers uses different classification algorithms for recognizing the gestures, they are mainly Hidden Markov Models (HMMs), Dynamic Time Warping (DTW), Neural Networks (NNs), data-driven template matching and statistical feature-based classifiers [13]. DTW based methods are quiet effective for dynamic gesture recognition [14]. We have used the computationally simple DTW based Leventian Minimum Edit distance algorithm for classifying the gestures.

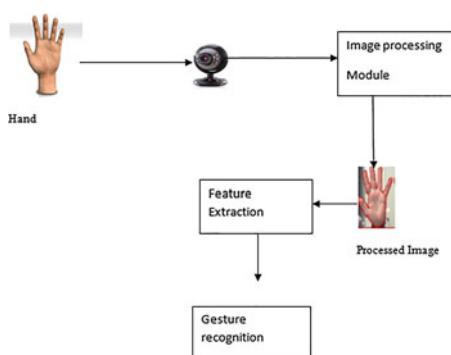
3 Dynamic Gesture Recognition System

A typical block diagram of a dynamic hand gesture recognition system is shown in Fig. 1. The hand gestures captured via web cam will be processed. From the processed image the position orientation and movement pattern of hand will be extracted, which is called feature vector. The extracted parameters will be used for gesture recognition. As discussed in the introduction section the dynamic gestures are produced by the movement of hand/finger in a peculiar manner. The orientation and direction of movements of hands are to be tracked.

The video frames captured via camera will be processed for segmenting the desired object. Special colour method, the segmentation of object of interest will be relatively easy. All other color can be eliminated from the frames other than a special color. (In this experiment Blue colour). The presence of special colours in the background can be eliminated by removing the non moving pixels i.e. part of back ground.

Similarly in the other method the skin color segmentation algorithms are used to segment the skin colors (hand regions) which takes part in the gesture production. The other skin region (face) can be eliminated by removing the non-moving or less

Fig. 1 Block diagram of hand gesture recognition system



moving pixels same as that of used in the special colour based segmentation. In the both the cases the object of interest will be identified and whose movement and orientation in the consecutive frames have to be recorded. From the recorded locus of points of the object of interest in 2D image plane, the motion vector (directional vectors) can be generated for further analysis. This feature vector will be used for gestures recognition.

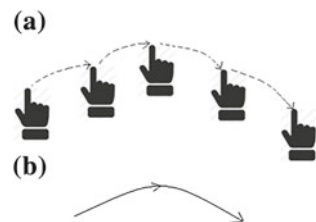
4 Dynamic Gestures

A dynamic hand gesture comprises a sequence of hand shapes with associated movements in the space. For recognizing dynamic hand gestures both orientation and instantaneous hand position need to be captured. The movement of hand or the object of interest will be tracked and it will be modeled as feature vector for recognizing the dynamic gestures. For example the gesture corresponding to a command *rotate right* or we expect user to move their hand from left to right as shown in Fig. 2. In above Fig. 2a represent the instantaneous hand positions when hand moves from left to right. The trajectory of the hand is represented as shown in Fig. 2b. Such trajectories are mathematically modeled and converted as feature vector for recognition of gestures.

5 Segmentation

The gesture recognition process using a colour band method is discussed here. One of the method uses designated special colors or stickers in the finger or hand to recognize the object of interest (finger) among the complex background. The other method uses the skin color based segmentation algorithms. During the segmentation process anything other than the designated color will be removed, and the object with a pre-defined colour will be segmented and tracked. The following section discusses the segmentation of object of interest who takes parts in gesture production.

Fig. 2 The dynamic gesture equivalent to rotate right
a Instantaneous hand position
b trajectory of hand



5.1 Segmentation—Hand with Colour Stickers in Fingers

In this method we have used the blue color as designated colour for discriminating the object of interest from the complex scene. The process of segmentation is nothing but isolating the object of interest from the background in each image frame. In this case the occurrence of blue colour has to be identified and isolated. The movement and orientation has to be recorded for gesture recognition. To identify the region of pixels (blue colour region), BLOB coloring algorithm is used [15]. Let m be the current pixel and z the one of the neighbor pixel under consideration the Euclidean distance between z and m is calculated as

$$\begin{aligned} D(z, m) &= \|z - m\| \\ &= \left[(z - m)^T (z - m) \right]^{1/2} \\ &= \left[(z_R - m_R)^2 + (z_G - m_G)^2 + (z_B - m_B)^2 \right]^{1/2} \end{aligned} \quad (1)$$

where $\| \cdot \|$ represent the norms of the argument and subscripts R, G, B denotes the RGB components of vectors m and z . The locus of points such that $D(z, m) < \text{Threshold}$, it can be counted as a part of the pixel [16]. In this method it is assumed that the special colour attached with the object of interest will be absent in the background. The presence of special colors in the background can be eliminated by removing the non moving pixels i.e., part of back ground. However if the chosen colors are part of moving object then elimination will become difficult. This is one of the major drawbacks of this method.

5.2 Segmentation of Hand Using Skin Color

Hand segmentation from the complex image is done using the skin colour based segmentation. Skin colors are modeled in RGB space with decision rule given below.

(R, G, B) is classified as skin if:

$$\begin{aligned} R &> 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and} \\ \max\{R, G, B\} - \min\{R, G, B\} &> 15 \\ |R - G| &> 15 \text{ and } R > G \text{ and } R > B. \end{aligned}$$

The BLOB coloring algorithm discussed above will be used to segment the object of interest, i.e., hand. The pixels fall in the region as mentioned in the above decision rules will be segmented [17]. Once the object of interest is segmented whose movements can be tracked and feature vector will be generated. The following section discusses how feature vector is generated.

6 Feature Extraction—Freeman’s Chain Code

After the object of interest is isolated, whose movements are to be tracked for feature extraction. The locus of points of object of interest, moving in the plane can be extracted from the consecutive video frames. The feature vector will be generated based on the movement with respect to current position. We have used Freemans Chain Code which is widely used for online handwriting recognition [18]. Freemans Chain code is a trajectory representation technique of movement of an object in a plane with respect to the current position in Cartesian plane. There are two directions of chain code, namely 4-directional and 8-directional as shown in Fig. 3. In 4-directional code only 4 directions are possible whereas for 8 directional codes there are 8 possible directions from the current position. For the recognition purpose 8 Directional code are more accurate than 4-Directional code.

Let us discuss how directional codes are generated. The segmented object can make move from the current location to any of 8 directions as shown in Fig. 3. Depending on the direction of movement from the current position corresponding number will be generated. This sequence of number will generate an unique pattern (string) which can be used for recognition purpose. Group of 10×10 pixels are grouped together and considered as single unit. The directional codes are generated based on the movement of this grouped pixel form one position to other position based on the directional code mentioned in Fig. 3. The directional vector code generated will have lots of noise values. For example dynamic gesture corresponding to a digit “1” or Move down should ideally be $S = \text{“77777777777777777777..”}$ but the actual feature vector generated at a given context might be $S1 = \text{“777687777715777...”}$. From S1 we can derive S because the noise values like 68 and 1,5 can be deleted from S1 as it is not repeating for sufficient amount of time, So this values can be glitch in the actual signal and this will be removed at the generation stage itself. The feature vector generated for few dynamic gestures gesture corresponding to certain command are listed below with ideal expected string is shown in Table 1.

In addition to the gestures for system command the gesture based digit recognition also experimented as part of this study. Here user have to move his finger in air as if he write the digit from 1 to 9 as shown in Fig. 4. The directional codes corresponds to the digit will be generated, and this will be used for classifying the gestures.

Fig. 3 Chain codes
a 4-Directional and **b** 8 directional

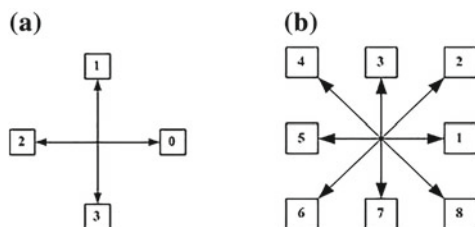


Table 1 Gesture and its freeman’s directional codes

Gesture	Ideal string	Generated string
Move right	1,1,1,1,1,1,1,1,	1,1,2,1,1,2,8,1,1,1
Move up	3,3,3,3,3,3,3,3,	3,3,4,3,3,3,2,3,3,4,3
Move left	5,5,5,5,5,5,5,	5,5,3,7,5,5,5,4,5,5,6
Move down	7,7,7,7,7,7,7,7,	7,7,7,1,5,7,7,7,6,5,7

Fig. 4 Dynamic gesture corresponds to digit 2



7 Gesture Analysis and Recognition

Various calcification algorithms are used to classify an incoming gesture. A dynamic programming based Levenshtein Minimum edit distance [19] algorithm is used for the directional code string comparison. The method is similar to the Dynamic time wrapping algorithm which measures the similarity between two time series data. Levenshtein distance (LD) is a measure of the similarity between two strings. The distance is calculated as the number of insertions, deletions, or substitutions required for converting the source string to target string. The best match string in the presorted template will have minimum edit distance. If two strings are same, the minimum edit distance between them will be zero. If minimum edit distance values for two strings found to be equal, the probability of occurrence of the string on that particular context will be analyzed. For example the probability of issuing a page up or page down command will be high if user opened a word processor.

8 Experimental Details

The efficiency of dynamic gesture recognition algorithm is tested in two different case implementations. The first one is with special colour band sticker as part of fingers and this will be used as prominent information for segmentation. The other method does not use any special colour, where the segmentation will be done using the natural human skin colour identification method. Ten commands, ten digits are the vocabulary size for this experiment. One hundred samples per gesture are given to system for training. Gaussian models are build corresponding to each feature vector, i.e. corresponds to each gesture. The incoming feature vector will be tested against the Gaussian model generated values for a given gesture and its similarity can be measured using the Levenshtein distance algorithm. The system is tested

Table 2 The dynamic gesture recognition results with special color based segmentation and with skin color based segmentation techniques for operational keyword











		Recognition accuracy in percentage (%)	
Operational command	Corresponding gesture	Segmentation with special color	Segmentation with skin color
Volume up		81	77
Volume down		79	72
Move down		95	93
Move up		93	90
Move back		93	89
Move next		90	89
Rotate right		94	90
Rotate-left		91	87
Close		93	94
Select		86	82

Table 3 The dynamic gesture recognition results with special color based segmentation and with skin color based segmentation techniques for digits

		Recognition accuracy in percentage (%)	
Gesture	Corresponding gesture	Segmentation with special color	Segmentation with skin color
One	1	98	97
Two	2	89	85
Three	3	90	83
Four	4	88	82
Five	5	92	91
Six	6	90	87
Seven	7	96	95
Eight	8	86	83
Nine	9	93	91
Zero	0	92	90

with five different users. The average recognition accuracy of the dynamic gesture recognition system for special color based method and skin color based method are listed in Tables 2 and 3 respectively. It is found recognition accuracy of special color method is more over the skin color based segmentation techniques.

9 Conclusion

A study on dynamic gesture recognition using computationally simple, Freeman chain code based feature vector with simulation results are discussed. Hand based dynamic gestures are experimented in this paper. Ten digits and ten operational commands used for interacting with computers are simulated as part of this experiment. The segmentation of hand from the background is being addresses using special color band and skin color segmentation methods. The accuracy of the system is found to be more with special colour based segmentation than skin colour based segmentation techniques. This system can be extended with more number of operational gestures so that it can be used as primary mode of interaction for interacting with computer systems. The dynamic gesture recognition system can also jointly work with speech recognition system, so that it can improve the accuracy of the speech recognition system especially in noisy environment.

References

1. Harshith.C, Karthik.R.Shastry, Manoj Ravindran, M.V.V.N.S Srikanth, Navee Lakshmikhanth “Survey on various Gesture Recognition Techniques for Interfacing Machines based on Ambient Intelligence” (IJCSSES) Vol.1, No.2, November 2010.
2. PragatiGarg, Naveen Aggarwal and SanjeevSofat, 2009. Vision Based Hand Gesture Recognition, World Academy of Science, Engineering and Technology 49, pp. 972–977.
3. Laura Dipietro, Angelo M. Sabatini, and Paolo Dario, 2008. Survey of Glove-Based Systems and their applications, IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 38(4), pp. 461–482, doi:10.1109/TSMCC.2008.923862.
4. Jayshree R. Pansare, Kirti S. Rampurkar, Pritam L. Mahamane, Reshma J. Baravkar, Sneha V. Lanjewar” Real-Time Static Devnagri Sign Language Translation using Histogram “International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013, pp. 1455–1459.
5. Francis Quek, David McNeilly, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E. McCulloughy, and Rashid Ansari: Gesture and Speech multimodal conversational interaction. ACM Transactions on Computer-Human Interaction, Vol. 9, No. 3, September 2002, pp. 171–193.
6. Hairong Jiang Duerstock, B.S.; Wachs, J.P., “A Machine Vision-Based Gestural Interface for People With Upper Extremity Physical Impairments”, Systems, Man, and Cybernetics: Systems, IEEE Transactions on (Volume:44, Issue: 5) May 2014, pp. 630–641.
7. N.S Sreekanth, Gopinath, Supriya Pal, N.K Narayanan. “GESTURE BASED DESKTOP INTERACTION.” International Journal of Machine Intelligence 3.4 (2011): pp. 268–271. <http://dx.doi.org/10.9735/0975-2927.3.4.268-271>.
8. Fabio Dominio, Mauro Donadeo, Pietro Zanuttigh. —Combining multiple depth-based descriptors for hand gesture recognition|| .Elsevier Publications, Pattern recognitions 2013.
9. Lei Yang, Hui Li, Xiaoyu Wu, Dewei Zhao, Jun Zhai. — An algorithm of skin detection based on texture||. IEEE Image and Signal Processing(CSIP), 2011.
10. Ohknishi, A. Nishikawa, Curvature-based segmentation and recognition of hand gestures, Proceedings Annual Conference On Robotics Society of Japan, 1997, pp. 401–407.

11. Feng-Sheng Chen, Chih-Ming Fu, Chung-Lin Huang “Hand gesture recognition using a real-time tracking method and hidden Markov models”, *Image and Vision Computing* 21 (2003)-Elsewhere pp. 745–758.
12. B.J Manikandan, Gowri Shankar, V Anoop, A Datta, V S Chakravarthy: LEKHAK: A System for Online Recognition of Handwritten Tamil Characters. Proceeding of (ICON-2002) Vikas Publishing House Pvt.Ltd. pp. 285–291.
13. Mingyu Chen, Ghassan AlRegib, Biing-Hwang Juang, “Feature Processing and Modeling for 6D Motion Gesture Recognition” *IEEE TRANSACTIONS ON MULTIMEDIA*, VOL. 15, NO. 3, APRIL 2013, pp. 561–571.
14. Kaustubh S. Patwardhan Sumantra Dutta Roy, “Dynamic Hand Gesture Recognition using Predictive EigenTracker”, Proceedings of. Indian Conference on Computer Vision, Graphics and Image Processing, 2004. https://www.ee.iitb.ac.in/~sumantra/publications/icvgip04_gesture.pdf.
15. Wöllert, Thomas (Dipl.-Inf. (FH)): About Portable Keyboards with Design and Implementation of a Prototype Using Computer Vision, Semester Thesis, Master of Science - Computer Graphics and Image Processing, Munich University of Applied Sciences, Germany, June-2006.
16. Rafael. C Gonzalez, Richard. E Woods, Steven .L Eddin: *Digital Image Processing 2/e* Pearson Education –Third Indian Reprint -2005, pp. 251.
17. PEER, P., KOVAC, J., AND SOLINA, F. 2003. Human skin colour cluster-ing for face detection. Proceeding of EUROCON 2003.
18. B.J Manikandan, Gowri Shankar, V Anoop, A Datta, V S Chakravarthy: LEKHAK: A System for Online Recognition of Handwritten Tamil Characters. Proceeding of the International Conference on Natural Language Processing (ICON-2002) Vikas Publishing House Pvt. Ltd. pp. 285–291.
19. Algorithm Implementation/String/Levenshtein distance wikibooks http://en.wikibooks.org/wiki/Algorithm_Implementation/Strings/Levenshtein_distance.

Medical Image Security with Cheater Identification Using Secret Sharing Scheme

Arun Krishnan and Manik Lal Das

Abstract Security of medical images and relevant patient information is a matter of important concern while using public networks for transfer of medical images between patients and clinicians. Clinicians require to confirm the legitimacy of patient medical images for applications such as telediagnosis and teleconsultation. Furthermore, medical image should not be perceivable to unauthorized parties with malicious intentions on the patient's health. As a result, medical images must be protected with suitable primitives while transferring them over public channel. In this paper, we present a scheme for protecting medical images using a threshold secret sharing scheme. The proposed scheme protects images from unauthorized access and intermediate tampering, thus, ensuring confidentiality and integrity of the shared images and associated patient records. The scheme takes into consideration the possibility of malevolence from any of the participating clinicians and detects and identifies cheating among the clinicians, if any. The proposed scheme is analyzed and simulated with electronic patient records and the experimental results satisfy all the properties of the scheme.

Keywords Medical image · Cheater identification · Secret sharing

1 Introduction

With rapid advancement of Information and Communication technology in health-care applications security and privacy of medical data has become a crucial factor in protecting patients records as well as clinicians information. A simple data infringement by an (un)intentional attempt can pose serious problems in the diagnosis process, even can lead to loss of patients lives. For example, consider the case of

A. Krishnan · M.L. Das (✉)
DA-IICT, Gandhinagar Gujarat, India
e-mail: maniklal_das@daiict.ac.in

A. Krishnan
e-mail: arun_krishnan@daiict.ac.in

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_12

117

teleconsultation, wherein a clinician at one hospital may call a senior colleague of him located at another hospital to get his opinion regarding a complicated diagnosis case. The former may have to sent some medical images and patient records to the latter to make some opinion regarding the case. Now, if the data are transmitted in open networks without any form of security, an adversary (could be a professional rival) may tamper the data and thus, prevent the senior clinician in getting a correct opinion from his colleague. Therefore, protecting medical images through public networks is of extreme significance, as a corrupted image can tend to a wrong diagnosis and may cost a patient's life. The main challenges faced while transferring a medical image are whether the medical image received is same as that sent and whether the image is discernible to malicious parties. In this paper, we discuss the security aspects of electronic patient records (EPR) with respect to different primitives used in protecting the EPR, and then present a scheme that addresses both confidentiality and integrity of EPR while transmitting them in public channel. The proposed scheme uses a (k, n) threshold secret sharing scheme [1] for transferring medical image, where the image and patient records are broken into n share images such that at least k of them must share their individual images to reconstruct the original medical image and associated patient records. The proposed scheme offers mechanism to detect the presence of cheating by any internal or external participant in the image construction phase. The features of the proposed scheme are summarized as follows:

- Shared medical images are hidden in natural cover images to thwart the attention of malevolent attackers.
- Sharing image among a team of clinicians reduces the risk of loss of secret image or unprecedented modifications.
- Integrity of the shared images is ensured which helps the receivers to detect if any tampering done on the images.
- Clinicians providing a wrong share image to force a wrong diagnosis can be identified uniquely.

The organization of the paper is as follows. Section 2 gives a brief overview of the various techniques used for image security. Section 3 presents an improved scheme. Section 4 analyses the security of the scheme and gives the experimental results. We conclude the work with Sect. 5.

2 Background and Motivation

Secret sharing scheme. Secret sharing scheme introduced by Shamir [1] is an interesting primitive used for sharing a secret among a number of participants n such that at least k of them must collect their individual shares to recover the secret. Typically, the secret sharing method uses a $(k - 1)$ degree polynomial to break the secret down into shares such that each participant gets a share of the secret. The secret to be shared will be the constant value of the polynomial and the polynomial is evaluated for unique x values (known only to the secret dealer and the corresponding

participant) so that each participant gets a unique share. To reconstruct the secret, k ($< n$) of the participants can come together and provide their shares and respective x values, and then using the Lagrange's interpolation they can obtain the actual secret.

Cheating Detection and Identification. Shamir's secret sharing scheme can work efficiently and reconstruct the secret accurately only when the involved participants provide correct shares. Cheating detection, that is detecting a false share, and more importantly cheater identification [2, 3], identifying the participant providing the false share, is essential for the scenarios of medical applications. Wu and Wu [4] proposed a scheme to share a secret among a set of participants using Shamir's (k, n) threshold sharing scheme with the objective that any false share during image reconstruction can be detected and the respective participant can be uniquely identified.

Sudoku-based secret image sharing scheme with reversibility. Chang et al. [5] proposed a secure image sharing scheme which uses the concept of Sudoku number grid. The scheme generates share images from the secret image using Shamir's secret sharing scheme and uses a 16×16 Sudoku grid to hide the share images within natural looking cover images to avoid unnecessary attention to the shares. The scheme ensures the confidentiality of transmitted image through generation of random looking share images, where the pixels of the share images are considered as Sudoku digits and their location is encoded as the pixels of the cover image. At the receiver side using the same Sudoku grid used in the sender side the share images are retrieved from the cover images and then, Lagrange's interpolation is carried out to reconstruct the secret image.

(k, n) secret sharing scheme from Quadratic Residues for grayscale images. Abdelsatir et al. [6] proposed a (k, n) image sharing scheme which uses a QR (Quadratic Residues) approach for the image security. In their scheme the pixels of the secret image are first permuted using a secret key and then are encrypted using quadratic residues. The encrypted image is then input to Shamir's (k, n) threshold scheme which produces n share images. At the receiver side at least k of the n share images and respective keys are required to retrieve the encrypted image. Once the encrypted image is decrypted using the QR technique and depermuted using the secret key, the secret image is reconstructed correctly.

Medical image security and EPR hiding. Ulutas et al. [7] proposed a scheme for the secure transfer of medical images over a public network to a set of participants. They use Shamir's secret sharing scheme to generate n share images from the original medical image and EPR (Electronic Patient Record) characters using a set of unique x values known to the respective participants. In their scheme the secret sharing acts as an encryption step, as the share images look like noise, thus ensuring the confidentiality of transmitted shares to other participants of the system. Furthermore, to thwart attackers's attention Steganography technique is used to hide the noisy looking share images into natural looking cover images. The scheme uses certificates for each share image for achieving the integrity of received share images. At the receiver side k out of the n participants can come together to reconstruct the original medical image and associated EPR using Lagrange's interpolation technique. We

observe some limitation of [7] and propose an improved scheme with additional features in next section. We refer to [7] for details of the scheme. The limitations of the scheme [7] are as follows.

- Secure exchange of x value suffers from man-in-the-middle attack between dealer and the participants, through which the attacker, posing as a participant, can establish an x value with the dealer and it can, posing as the dealer, establish a different x value with the participant. Therefore, any communication from the dealer to the participant can be compromised by the attacker and modified without knowledge of either the dealer or the participant.
- Once the Verification phase is completed successfully, that is, at least k stego images have been verified to be authentic and untampered, then the reconstruction phase begins where the medical image and EPR is reconstructed. However, there is no mechanism to check the authenticity of the share images used for reconstruction, that is, there is no way to ensure that the share images retrieved from untampered stego image by respective clinicians are same as the share images provided for reconstruction. Any of the participants can modify the received share image or provide a different share image altogether, thus causing a entirely different image to be reconstructed. The scheme does have any mechanism to detect any such malevolence or to identify the respective malevolent participants.
- The Initialization phase is prone to man-in-the-middle(MITM) attack. An attacker posing as a participant can mislead the dealer to sharing all the x values with him, thus enabling him to reconstruct the medical image himself by collecting all the share images and executing the procedure discussed above. Furthermore, the attacker can then pose as the Dealer and perform the entire secret sharing scheme and can share the medical image among the actual participants. Neither the dealer nor the participants will be aware of the attacker's presence in this scenario. This will affect the confidentiality of the transmitted medical image and associated patient records. It also can affect the integrity of the scheme as the attacker, who has all the x values, can generate valid certificates for any image, therefore, causing any participant to accept any image sent by the attacker as a valid share image from the dealer.

3 An Improved Scheme

The proposed scheme is based on the scheme [7], but aims to address the limitations, as observed in the previous section, of the scheme [7]. In addition, the proposed scheme provides one important feature, that is, the cheating detection and cheater identification. The proposed scheme includes an additional entity called TTP (Trusted Third Party) who performs the reconstruction of the medical image from the shares. The Fig. 1 depicts the architecture of the improved scheme. The improved scheme has the same two procedures as in the scheme [7]—Partitioning and Retrieving procedures. In order to resist the MITM attacks in the Initialization phase, the modified Initialization phase works as follows.

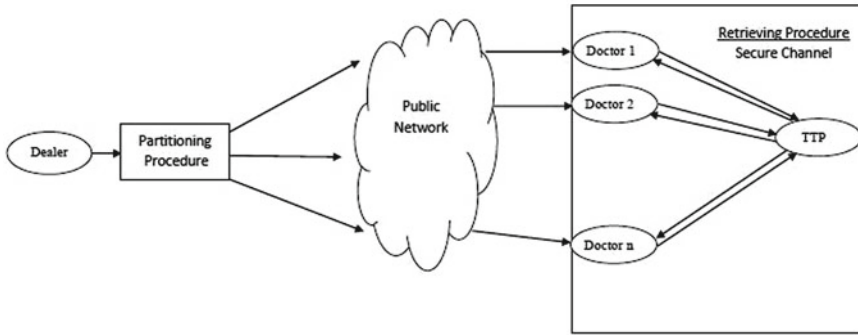


Fig. 1 Architecture of the improved scheme

Initially a secret session key is established between the Dealer and the TTP using the password authenticated key exchange protocol [8]. This key exchange protocol in [8] establishes a secure session key between the TTP and the Dealer with mutual authentication of the two parties. Once the session key, K , is established between the Dealer and the TTP, the TTP sends the secret key K to the authorized participant clinicians through a secure channel. Now, while establishing unique x values using the steps followed in the Initialization phase, the public parameters R_i and R_0 are being encrypted using the key K before their transfer through the network. Since the values R_i and R_0 are now encrypted using K which is only known to the Dealer, TTP and authorized participants, an adversary will not be able to generate false public values and mount a MITM attack. Once each participant has unique x_i values, share images are generated using the Sharing phase. The Sharing phase, however, includes some additional computations which will be used later for Cheating detection and identification. A value T is calculated by the Dealer which is used in the cheater detection phase. The value T is calculated as follows:

$$T = \sum_{i=1}^n h(ST_i || x_i) * p^{2(i-1)} + \sum_{i=1}^{n-1} c * p^{2i-1}$$

where $h(ST_i || x_i)$ is the authentication code of the share image for the i th participant ST_i . The Sharing phase is followed by the Embedding phase as in [7] which performs Steganography of the share images and a Protection phase generating Certificates for each image.

The Verification phase of the Retrieving procedure will follow the same steps as in [7]. The Reconstruction phase is performed by TTP, to whom the share images and corresponding x values are sent by at least k participants. The TTP, on reception of the shares, carries out a Cheating detection phase to check for any false shares. Let G be the group of k participants who provide their shares and their x values for reconstruction of the original medical image and patient records. TTP performs the following computations:

1. All the k participants provide their share images ST_i and x_i values. TTP calculates a value \tilde{T} as $\tilde{T} = \sum_{j \in G} h(ST_j || x_j) * p^{2(j-1)}$, where j is the index of each participant.
2. For each participant check whether $\lfloor \frac{T-\tilde{T}}{p^{2(j-1)}} \rfloor \pmod{p} = 0$.

If the equation holds, the j th participant is an honest participant; else, he is a cheater. Therefore, TTP can detect and identify cheating by any of the participants. Also, since integrity of the received stego images are tested using Certificates, no participant will be falsely accused of tampering caused by adverse parties in transit. If all the participants have provided honest shares then the TTP can reconstruct the original image and EPR from the share images and the respective x values. If not, then the TTP can detect it and will blacklist the participant, who provided false shares, and restart the Reconstruction phase using the share images of other honest participants involved.

4 Analysis of the Improved Scheme

Security Analysis. We consider an active attacker who has full control over the communication such as he can view, intercept and also modify messages as per his attack goal. The following security goals have been achieved by the improved scheme.

Confidentiality: Shamir's secret sharing technique acts as an encryption step which converts the image to illegible noisy looking share images, where the corresponding x values act as the encryption key. An attacker can intercept the stego images and retrieve the share images, but to reconstruct the original image from the share images he requires the corresponding x values which are unknown to him. Therefore, the attacker cannot obtain the original image.

Authentication: The improved scheme employs a Protection phase which generates a certificate, $Cert_k$, for each stego image. The $Cert_k$ transmitted along with the stego images allows the recipient to check for the authenticity of received stego images. Since the share images are extracted from the stego images, which verifies the authenticity of the received share images and the original image reconstructed using the share images. Furthermore, the input to the hash function is the stego image concatenated with the corresponding x value, known only to the dealer and the respective participant, which prevents the generation of counterfeit hash values. Therefore, the improved scheme provides authentication of the reconstructed medical image.

Data Hiding: Patient records are hidden in the share images in the Sharing phase wherein EPR characters are also input to the polynomial along with the medical pixel values. The generated share images are random and do not provide any information about the hidden data, which can only be retrieved only by authorized participants, holding the respective x value.

Cheating detection and identification of the source: The improved scheme identifies cheating from any of the participants and uniquely identifies the source of it. The technique works on the basis of the two values, T and \tilde{T} , calculated in the Sharing

phase and Reconstruction phase. Using these two values the cheating detection and identification occurs as follows:

Claim *If $\lfloor \frac{T-\tilde{T}}{p^{2(j-1)}} \rfloor \pmod p = 0$ then the j th participant, with x value = x_j , has provided the correct share for reconstruction and he is an honest participant; else, he is a cheater.*

The proof can be explained based on the following theorem in [4].

Theorem *If $T = \sum_{i=1}^n a_i * p^{2(i-1)} + \sum_{i=1}^{n-1} c * p^{2i-1}$ where $-(p - 1) \leq a_i \leq (p - 1)$ and $1 \leq c \leq (p - 1)$, then $\lfloor \frac{T}{p^{2(j-1)}} \rfloor \pmod p = a_j \pmod p$, where p is a prime number $-p < a_i < p$ and $1 \leq c < p$. The theorem is modified to produce the cheater detection and identification technique used in the proposed improved scheme. The value a_i is replaced by the hash of the share image and corresponding x value $H(ST_i || x_i)$. Now, $\lfloor \frac{T-\tilde{T}}{p^{2(j-1)}} \rfloor \pmod p$ can be simplified as:*

$$\lfloor \frac{T - \tilde{T}}{p^{2(j-1)}} \rfloor \pmod p = (H(ST_j || x_j) - H(\tilde{S}T_j || x_j)) \pmod p$$

The equality $H(ST_j || x_j) - H(\tilde{S}T_j || x_j) = 0$ is satisfied, if and only if $H(ST_j || x_j) = H(\tilde{S}T_j || x_j)$, in turn, means that $ST_j = \tilde{S}T_j$.

In other words, the share image transmitted to the j th participant and the share image provided by him for reconstruction of the image, ST_j and $\tilde{S}T_j$ are same, which proves that the participant is honest. If they do not match, then the participant has provided a wrong share for reconstruction and is a deceiver. The strengths of the improved scheme are compared with the Sudoku scheme [5] and QR scheme [6], and results are shown in Table 1.

Experimental Results. We implement the improved scheme using MATLAB R2013a in an Intel i3 processor with 4GB RAM. The test image used was a 256×256 MR (Magnetic Resonance) image of brain having 8-bit pixel depth as shown

Table 1 Comparison with other image sharing schemes

Properties	Proposed scheme	Sudoku scheme [5]	QR scheme [6]
Confidentiality	Yes	Yes	Yes
Authentication	Yes	No	No
Steganography	Yes	Yes	No
Data hiding	Yes	No	No
Cheating detection and identification	Yes	No	No
Recovery of original image	Yes	Yes	No

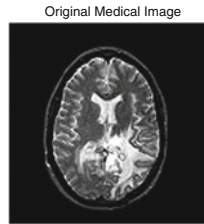


Fig. 2 MRI image of brain

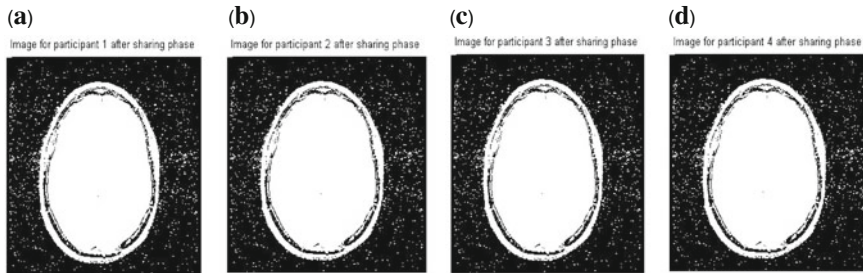


Fig. 3 Share images generated after sharing phase. **a** Share image 1. **b** Share image 2. **c** Share image 3. **d** Share image 4

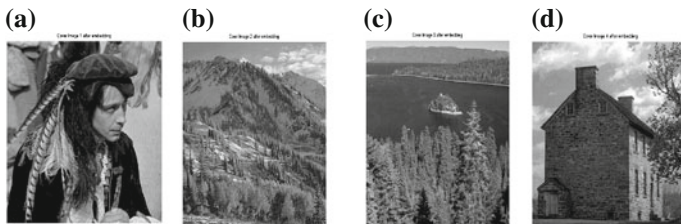


Fig. 4 Stego images distributed to clinicians. **a** Stego image 1. **b** Stego image 2. **c** Stego image 3. **d** Stego image 4

in Fig. 2. The EPR used was the text “disease” which is later converted to a set of corresponding ASCII values for use in the polynomial.

The improved scheme is then applied to the test image and EPR. A $(3, 4)$ threshold secret sharing scheme is used to reconstruct the original medical image and associated EPR. The first step was to generate the x_i values for each participant. Then using the four generated x_i values four share images were generated, one for each participant, as shown in Fig. 3a–d. The stego images generated after the embedding phase, wherein each share image is embedded in a different cover image as shown in Fig. 4a–d. Each of the stego image is received by the respective clinician, which then goes through the Verification phase, checking the integrity of the received stego image. Once the Verification phase is successful, the Reconstruction phase is exe-

Fig. 5 Reconstructed medical image

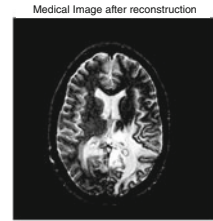
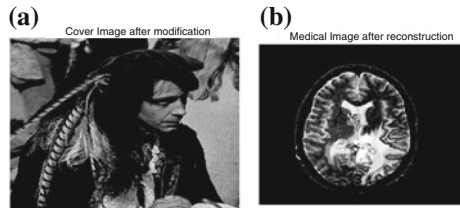


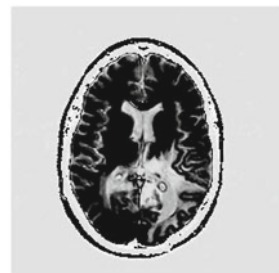
Fig. 6 Tampered image.
a Tampered cover image 1.
b Reconstructed corrupted image



cuted by the TTP generating the medical image and EPR which is then distributed to the respective clinicians through secure channels. The reconstructed medical image is shown in Fig. 5.

If any of the stego images have been corrupted (accidentally or maliciously) then it will be detected in the Verification phase as it uses a hash function for verification which detects even a 1-bit change. The result of tampering of a cover image and the resulting corruption in the medical image are shown in Fig. 6a, b. It is noted that the reconstructed medical image has some distortions compared to the original image. The Figures depict the case of a normal scenario when all the participants provide honest shares and TTP, after checking for any cheating by any of the participants, reconstruct the original medical image and EPR. If one of the participants provides a false share, it can result in the reconstruction of a false medical image as shown in Fig. 7. The improved scheme takes care of this using the cheating detection phase, performed by the TTP before actual reconstruction.

Fig. 7 Medical image reconstructed using a false share



5 Conclusion

We discussed about the secure image transfer aiming at protecting EPR while transmitting them over an insecure public channel. We reviewed techniques on which images can be shared among multiple participants over an insecure channel. We then reviewed a scheme [7] on medical image protection using secret sharing scheme. We found some weaknesses on the scheme [7], and proposed an improved scheme. In addition to image confidentiality and authenticity, the improved scheme prevents any malevolent clinician from providing a false share, and thus causing a wrong diagnosis, by detecting and identifying any cheating involved. The experimental results of the improved scheme show that the scheme satisfies all the properties claimed in the analysis section.

References

1. A. Shamir. How to share a secret. In: *Communications of the ACM*, 22(11):612–613, 1979.
2. M. Tompa and H. Woll. How to share a secret with cheaters. In: *Journal of Cryptology*, 1(2):133–138, 1988.
3. I. C. Lin. A (t, n) Threshold Secret Sharing System With Efficient Identification Of Cheaters. In: *Computing and Informatics*, 24:529–541, 2005.
4. T. C. Wu and T. S. Wu. Cheating detection and cheater identification in secret sharing schemes. In: *IEEE Proceedings of Computers and Digital Techniques*, 142(5):367–369, 1995.
5. C. C. Chang, P. Y. Lina, Z. H. Wangb, and M. C. Lib. A Sudoku-based Secret Image Sharing Scheme with Reversibility. In: *Journal of Communications*, 5(1):5–12, 2010.
6. E. B. Abdelsatir, S. Salahaldeen, H. Omar, and A. Hashim. A Novel (k, n) Secret Sharing Scheme from Quadratic Residues for Grayscale Images. In: *International Journal of Network Security and Its Applications*, 6(4), 2014.
7. M. Ulutas, G. Ulutas, and V. Nabyev. Medical image security and EPR hiding using Shamir’s secret sharing scheme. In: *Journal of Systems and Software*, 84:341–353, 2011.
8. M. Saeedand H. S. Shahhoseini, and A. Mackvandi. An improved two-party Password Authenticated Key Exchange protocol without server’s public key. In *Proceedings of the IEEE International Conference on Communication Software and Networks*, pp. 90–95, 2011.

The Role of Fractal Dimension, Lacunarity and Multifractal Dimension for Texture Analysis in SAR Image—A Comparison Based Analysis

Triloki Pant

Abstract In present paper, a fractal approach to study the texture in SAR images has been explored and the utility and problems of fractals for texture analysis are discussed. Since satellite images are rich in texture, they have to be studied in details for texture analysis. In present study, an ERS2 SAR image has been used for estimation of fractal dimension, lacunarity and multifractal dimension where the texture has been studied on the basis of these parameters and compared. A conclusion regarding the applicability of these three parameters has been drawn in the study.

Keywords Fractal dimension • Image analysis • Lacunarity • Local fractal dimension • Multifractal dimension • Texture

1 Introduction

Texture is a property of neighbourhood in digital images which can be defined as a variation of pixel values at a scale smaller than the scale of interest [1–8]. In digital images, texture is defined and studied with the help of a local window which is a predefined neighbourhood of pixels. The peculiar thing about the texture is that it varies with the neighbourhood and can disappear if the pixel neighbourhood changes to certain extent. For example, in a window of 3×3 , texture may be defined which can vary if the window size is increased and the same texture may not be called as texture if the window size is increased to say, 15×15 . Similarly a pixel pattern may define a texture in the window size 15×15 , which may disappear in a lower window size, say 5×5 and hence can't be said to be a texture. Thus, texture is a context dependent property of digital images which also changes according to the measuring scale. The study of texture is important in image analysis as it can define various image objects on the basis of their context and thus becomes a topic of prime interest in numerous applications including computer

T. Pant (✉)
Doon University, Dehradun, Uttarakhand, India
e-mail: trylucky@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_13

127

vision. There are various types of textures defined as per the variation of image pattern, e.g., homogeneous and heterogeneous [1, 3]. Thus, measurement of texture is an important issue in image analysis because of context and scale dependency of texture. As per the human perception, texture is called to be rough, highly rough, smooth, very smooth and so on which becomes a fuzzy terminology and hence difficult to measure. In order to measure the texture for analysis purpose and other applications like computer vision, texture must be measured in terms of numeric values. As a result, numerous measures to map the texture exist in practice and are used in various applications, e.g., co-occurrence matrix, autocorrelation index, local binary pattern, entropy value, fractal dimension [1–8].

Since context plays a major role in identification of objects and in the absence of enough local information it is difficult to identify the objects accurately [9], it is necessary to define the context or local neighbourhood appropriately. As a consequence, it is necessary to define the context, for which the fractal features are to be identified, i.e., the size of local window is to be defined properly. The exercise regarding selection of local window has been done in [10]. In present study, the local neighbourhood has been fixed as 5×5 pixels and the fractal features have been estimated using this moving window. It is very important to define the neighbourhood prior to estimation of fractal features because the fractal dimension varies as the local window size changes [10].

2 Fractal Based Texture Study

Although, a number of texture measures are available for study, fractals are preferred in present study due to the fact that fractals can best model natural scenes and natural objects [10–12]. As described by Pentland [11], natural scenes can be modeled with the fractional Brownian motion (fBm) and thus the digital image $I(x, y)$ satisfies following pdf of an fBm

$$\Pr\left(\frac{I(x + \Delta x, y) - I(x, y)}{\|\Delta x\|^H} < z\right) = F(z) \quad (1)$$

where H is the Hurst parameter, $0 < H < 1$ and $F(z)$ is the cumulative distribution function. The two properties of a fractals, i.e., self-similarity and fractal dimension are fundamental to any fractal object, however there exist other parameters to define the fractals, e.g., self-affinity, lacunarity, multifractal dimension. Since the study is based on fractal dimension, lacunarity and multifractal dimension, these measures are discussed now.

2.1 Fractal Dimension (D)

As per the Eq. (1), digital images are to be considered as fractal objects with some value of fractal dimension defined as [1, 2, 10–13]

$$D = \frac{\log(N_r)}{\log(\frac{1}{r})} \quad (2)$$

where r is the scale and N_r is the number of self-similar objects at scale r . The value of fractal dimension is independent of scale, i.e., second property of fractals. It is customary to note that the fractal dimension of digital images lies between 2.0 and 3.0 which represents the roughness of images in terms of numbers.

Theoretically these two facts are perfectly valid for digital images, however, the scale independence has to be compromised for digital images due to digitization of pixel values. Further, Eq. (1) is used to model natural objects with fractals which follow a statistical self-similarity instead of strictly self-similarity and thus, natural scenes will be considered to follow the properties of fractals. As a matter of fact, the fractal dimension is independent of the scale, i.e., at various scales if fractal dimension is measured, it would remain same, however for natural scenes it would not follow this fact. Since fractal features are identified by the fractal dimension, the major concern regarding estimation of fractal features would be estimation of fractal dimension of the images. The study is not dealing with the scale because the images are digitized at once and used for the study with no further modification in terms of sampling or resizing. The only fact to be kept at hand is the fact that the fractal dimension can be estimated either for the full image at once called global fractal dimension or for the local context defined by the neighbourhood of pixels, called local fractal dimension.

It has been reported in [10] that the local fractal dimension is of much importance in the study of texture as compared to the global fractal dimension. Since texture is a property of neighbourhood and so the fractal dimension, it has been proposed that fractal dimension can be used to map the texture in terms of numbers. The fractal dimension of images ranges from 2.0 to 3.0 and thus image texture can be mapped in a range between 2.0 and 3.0. The analogy of fractal dimension and texture comes from the fact that for an equal distribution of pixels in a given neighbourhood, the fractal dimension comes to be lower (near to 2.0) and the object is said to be smooth in nature and equally the texture is also called a smooth texture. On the other hand, for roughly distributed pixels in the neighbourhood, the fractal dimension comes to be higher (near to 3.0) and the corresponding texture is also called to be rough texture. Thus, fractal dimension is an appropriate measure of texture which maps the texture in the range 2.0–3.0.

The issue of non-uniqueness of fractal dimension was resolved by introducing the parameter lacunarity by Mandelbrot [2], which refers to the gaps occurring in the fractal objects. The concept says that if the fractals with different geometry have same value of fractal dimension, they can be identified with gaps among the

constituent objects, i.e., the lacunarity of the fractals. This fact has been well established in practice and utilized in present study to differentiate the fractals with same fractal dimension. Thus, lacunarity comes to be a rescue to fractal dimension in order to defy the problem of non-uniqueness. The idea here is that if the fractal dimension of different objects is same, their lacunarity will vary otherwise these objects are same. Thus, these two measures being complementary in nature are sufficient to identify the objects uniquely.

2.2 Lacunarity

The additional measure of texture, based on fractal property is lacunarity (L), which has been used in present study. Lacunarity is the measure that deals with the gaps in fractal objects, by definition and it takes care of the distribution of pixels in the image [1, 2, 14, 15]. If the pixels are distributed uniformly, the gaps between them become uniform and these pixels in the neighbourhood can be considered as similar. On the other hand, the pixels roughly distributed show different gaps and can be considered as dissimilar. Thus, similar and dissimilar pixels can be identified to form different clusters and this fact is used as an aid to the fractal dimension. The pixels having a fixed range of D and uniform distribution, i.e., fixed L can be grouped as one and this process can be repeated for a number of D ranges.

Lacunarity is defined as the ratio of the variance over the mean value and defined for an $N \times M$ image having the pixel values $I(n, m)$ by any of following [1]

$$L = \frac{\frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M I(n, m)^2}{\left(\frac{1}{NM} \sum_{k=1}^N \sum_{l=1}^M I(k, l) \right)^2} - 1 \quad (3)$$

or

$$L = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M \left| \frac{I(n, m)}{\frac{1}{NM} \sum_{k=1}^N \sum_{l=1}^M I(k, l)} - 1 \right| \quad (4)$$

Both the above definitions are equivalent and hence any can be used for defining the lacunarity. In our study, the first definition has been used. It is obvious from the definition that the value of L will be 0 for uniform image, i.e., an image with all the pixel values to be same. Further, higher value of L represents more inhomogeneous image, i.e., highly rough surface [1, 3]. The reason for using L as a measure for labeling is the fact that lacunarity can be used to fill the gap created by fractals. Since the fractal dimension can be same for two different objects and it becomes difficult to differentiate them, however they may be differentiated on the basis of lacunarity.

Non-uniqueness of fractal dimension has raised another issue in digital images called multifractal analysis. Multifractal is a fractal object for which the value of fractal dimension is different at different parts of the object which is contrary to the property of fractals. This feature is desirable in real life objects such as digital images where the mass is not uniformly distributed along the whole object and hence the fractal dimension changes at different parts. The multifractal property can be explained by considering a digital image, e.g., satellite image of agriculture land for which the pixels are not all same, however the image object is same. In such kind of object the pixels representing same object are varying in values and hence not uniformly distributed throughout the image and thus lead to different value of fractal dimension at different subimages. However, if the image has various objects distributed among the space available in the image, they can be identified by using local fractal features only, which include both fractal dimension and lacunarity. Thus in present study the guideline that local fractal features can be used over multifractal features when various objects are distributed in the image in place of single object has been followed. In other words, when various types of texture is present in the image instead of single or two types of textures, local fractal features are preferred over multifractal features in order to reduce the calculation of features and hence reduce the processing time.

2.3 *Multifractal Dimension*

A fractal has a basic property, called self-similarity which is defined for a fractal object such that each part of the object is a reduced copy of the whole object [2, 16–19]. However, satellite images do not exactly follow self-similarity due to discretization of image pixels and the self-similarity property changes at different parts of the images. Such an object can better be modeled with the multifractals instead of fractals [2, 19]. The multifractal property denotes the distribution of mass within the whole object and satellite images follow the multifractal behavior. The multifractal measures not only look into spatial relation but also into distribution of points [19], i.e., the distribution of pixel values with respect to images. The multifractal behavior can be measured by the multifractal dimension denoted by D_q and defined as [13, 18]

$$\begin{aligned}
 D_q &= \frac{1}{q-1} \lim_{r \rightarrow 0} \frac{\log X(q,r)}{\log r} \quad (q \neq 1) \\
 D_1 &= \lim_{r \rightarrow 0} \frac{\sum_i P_i \log P_i}{\log r}
 \end{aligned}
 \tag{5}$$

where $-\infty \leq q \leq \infty$ and $X(q,r) = \sum_i P_i^q P_i$ is the normalized measure of i th box and it can be defined for integer values of q , i.e., $q = 0, 1, 2$ etc. For a fractal object, D_q is same for all values of q , i.e., D_0 which is actually the fractal dimension.

The multifractal behaviour depends on the factor q , for $q = 0$, the value of D_q , i.e., D_0 represents the fractal dimension, for $q = 1$, D_q represents the information measure or entropy, for $q = 2$, D_q represents correlation dimension [18, 19]. Similarly for higher values of q , D_q represents higher order correlations [18].

Multifractals are useful for image classification purpose due to the fact that for different values of q , different values of D_q are obtained. The values of D_q can identify the objects better and hence become the multifractal signature of the objects. The concept of signature is similar to the fractal signature, where an image object has a fractal signature representing its unique features. For various values of q , i.e., $q = 0, 1, 2, \dots$ different local multifractal images can be generated and each multifractal image can be analyzed and classified. The classification results can be enhanced using multifractals as compared to those of using fractal dimension alone because fractal features alone are less efficient to identify various land types [13].

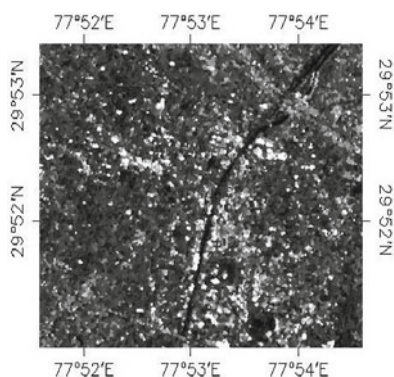
3 Methodology

The methodology follows to generate the texture maps of the SAR image by applying a moving window approach. The size of moving window has been fixed to be 5×5 and for this fixed size window texture maps estimating fractal dimension, lacunarity and multifractal dimension (D_0) are generated. Before performing the texture analysis, the data used in the study has been discussed and subsequently the methodology has been discussed.

3.1 Data Used

For the current study ERS-2 (European Remote sensing Satellite) SAR image generated in C Band (5.3 GHz) and having the spatial resolution 12.5 m has been used. The image lies between longitudes 77.807° E and 77.901° E and latitudes 29.890° N and 29.850° N and shown in Fig. 1.

Fig. 1 Data used (ERS2 SAR Image)



3.2 Methodology Applied

Since local fractal dimension is the measure of texture, it is necessary to estimate the value of local fractal dimension appropriately. The local fractal dimension is defined for a predefined local window which is of odd size, e.g., 3×3 , 5×5 etc. The fact that fractal dimension is a property of context is supported by the fact that local features give an identity to the objects due to their contextual information. The objects which are presented in a familiar context are faster to recognize [9] and thus it is obvious to estimate the local features accurately. In order to estimate the fractal dimension in the local context, an odd sized square local window is selected and it is moved over the image starting from the top-left corner in a row wise movement. The process is similar to the masking of digital image with the difference that the local window provides the value of fractal dimension instead of masked value. Inside the local window, the value can be estimated by following any of the available methods of fractal dimension estimation, e.g., Box counting method, Triangular prism surface area method (TPSAM), Variogram method, 2D Variation method and so on. In present study, TPSAM has been followed for estimation of the fractal dimension due to the ease and familiarity of the method [10, 20–22]. Using the same local window approach lacunarity and multifractal dimension are estimated as per the Eq. (3) and (5) respectively.

4 Results and Discussions

The fractal map, lacunarity map and multifractal dimension map obtained from the SAR image are given in Fig. 2, 3, and 4 respectively.

The fractal dimension of different land classes has been estimated and on the basis of this value, the image is classified. The process has been repeated for both lacunarity and multifractal dimension. In each case, 4 classes have been identified

Fig. 2 *D* Map with $w = 5 \times 5$

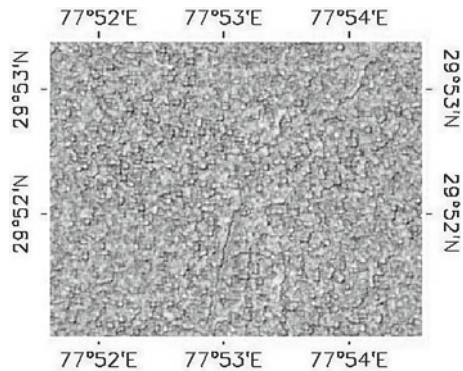


Fig. 3 L Map with $w = 5 \times 5$

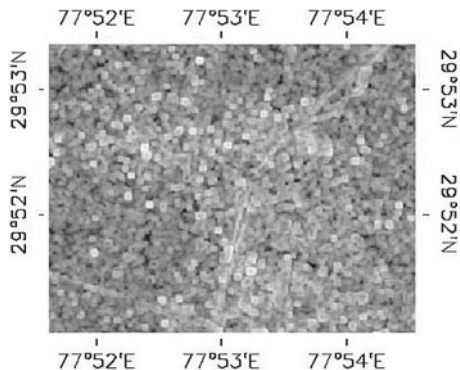
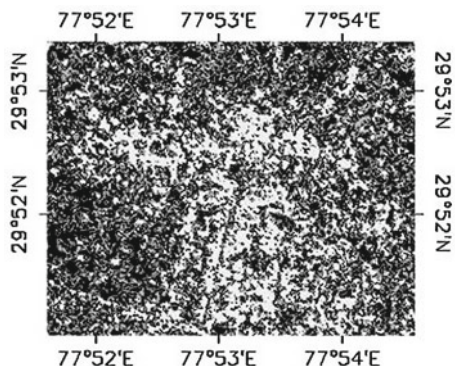


Fig. 4 D_0 Map with $w = 5 \times 5$



using an unsupervised classification scheme and the classification accuracy is estimated. The results of classification are listed in Table 1.

In present study it has been proposed that multifractal analysis can be done if the fractal objects in the image are uniform whereas lacunarity estimation can be done if different classes are mixed instead of uniform distribution of single class. The data of Table 1 shows that lacunarity based classification of texture is better as compared to the classification on the basis of fractal dimension and multifractal dimension and verifies the proposal.

The issues associated with texture identification have been discussed now very briefly. It is noteworthy to state at this juncture that both texture and fractal dimension are not unique in nature, i.e., for two different objects the same texture can be defined or obtained and equally two different objects may possess same value of

Table 1 Classification accuracy of SAR image using fractal measures

Texture measuring parameter	Classification accuracy (%)
D	51.05
L	72.50
D_0	65.34

fractal dimension. Various examples of such fractals can be found in [2] and the fact is also highlighted in [10]. The reasons are obvious as differently oriented pixels can generate same pattern for which fractal dimension comes to be same.

The novelty of present study lies in the comparison of three fractal based texture measures, especially for SAR images. The basis of comparison is classification accuracy for 4 major classes in SAR image and it has been observed that lacunarity is preferred over multifractal dimension when the image has more than one class.

5 Conclusions

The present study deals with fractal based texture analysis of SAR images which are known to possess rich texture. The texture is studied in a fractal based approach where fractal dimension, lacunarity and multifractal dimension are used as the measures of texture. Since texture is a property of the context, a 5×5 pixel window has been used to map the local texture content for each of these fractal parameters. The fractal dimension represents the roughness and hence texture in the range 2.0–3.0 whereas lacunarity is a measure of gaps between the fractal objects. Multifractal dimension, on the other hand represents the dimension at different parts of the image. The study shows that multifractal dimension is useful when the texture is uniform as compared to fractal dimension which is preferred when the texture is complex and possesses mixed textures.

References

1. Petrou, M. and Sevilla, P.G.: Image Processing Dealing with Texture. John Wiley and Sons, Ltd., England (2006)
2. Mandelbrot, B.B.: The Fractal Geometry of Nature. W.H. Freeman and Co., New York (1982)
3. Turner, M.J., Blackledge, J.M. and Andrews, P.R.: Fractal Geometry in Digital Imaging, Academic Press (1998)
4. Pant, T.: Implementation of Fractal Dimension for Finding 3D Objects: A Texture Segmentation and Evaluation Approach. Second International Conference, IITM 2013, Allahabad, India, March 9–11. (2013) 284–296
5. Myint, S.W.: Fractal Approaches in Texture Analysis and Classification of Remotely Sensed Data: Comparisons with Spatial Autocorrelation Techniques and Simple Descriptive Statistics. *Int. J. Remote Sens.* 24(9) (2003) 1925–1947
6. Rajesh, K., Jawahar, C.V., Sengupta, S. and Sinha, S.: Performance Analysis of Textural Features for Characterization and Classification of SAR Images. *Int. J. Remote Sens.* 22(8) (2001) 1555–1569
7. Satpathy, A., Jiang, X., and Eng, H.: LBP-Based Edge-Texture Features for Object Recognition. *IEEE Trans. Image Proc.* 23(5) (2014) 1953–1964
8. Chaudhuri, B.B. and Sarkar, N.: Texture Segmentation using Fractal Dimension. *IEEE Trans. Pattern Anal. Mach. Intell.* 17(1) (1995) 72–77
9. Oliva, A. and Torralba, A.: The Role of Context in Object Recognition. *Trends in Cognitive Sciences.* 11(2) (2007) 520–527

10. Pant, T., Singh, D. and Srivastava, T.: Advanced Fractal Approach for Unsupervised Classification of SAR Images. *Advances in Space Research*, 45(11) (2010) 1338–1349
11. Pentland, A.P.: Fractal-based Description of Natural Scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-6(6) (1984) 661–674
12. Riccio, D. and Ruello, G.: Synthesis of Fractal Surfaces for Remote-Sensing Applications. *IEEE Trans. Geosci. Remote Sens.* 53(7) (2015) 3803–3814
13. Pant, T., Singh, D., and Srivastava, T.: Multifractal Analysis of SAR Images for Unsupervised Classification. *International Conference on Recent Advances in Microwave Theory and Applications, Microwave-2008, Jaipur, India, Nov. 21–24.* (2008) 427–430
14. Chen, S.S., Keller, J.M. and Crownover, R.M.: On the Calculation of Fractal Features from Images. *IEEE Trans. Pattern Anal. Mach. Intell.* 15(10) (1993) 1087–1090
15. Plotnick, R.E., Gardner, R.H., Hargrove, W.W., Prestegard, K. and Perlmutter, M.: Lacunarity Analysis: A General Technique for the Analysis of Spatial Patterns. *Physical Review E.* 53(5) (1996) 5461–5468
16. Teng, H.T., Ewe, H.T. and Tan, S.L.: Multifractal Dimension and its Geometrical Terrain Properties for Classification of Multi-band Multi-polarized SAR Image. *Progress in Electromagnetics Research.* 104 (2010) 221–237
17. Cheng, Q.: Multifractality and Spatial Statistics. *Computers and Geosciences.* 25 (1999) 949–961
18. Sarkar, N. and Chaudhuri, B.B.: Multifractal and Generalized Dimensions of Gray-tone Digital Images. *Signal Processing.* 42 (1995) 181–190
19. Parrinello, T. and Vaughan, R. A.: Multifractal Analysis and Feature Extraction in Satellite Imagery. *Int. J. Remote Sens.* 23(9) (2002) 1799–1825
20. Clarke, K.C.: Computation of the Fractal Dimension of Topographic Surfaces using the Triangular Prism Surface Area Method. *Computers and Geosciences.* 12(5) (1986) 713–722
21. Sun, W., Xu, G., Gong, P. and Liang, S.: Fractal Analysis of Remotely Sensed Images: A Review of Methods and Applications. *Int. J. Remote Sens.* 27(21–22) (2006) 4963–4990
22. Ju, W. and Lam, N.S.N.: An Improved Algorithm for Computing Local Fractal Dimension using the Triangular Prism Method. *Computers and Geosciences.* 35 (2009) 1224–1233

Efficient Storage and Processing of Video Data for Moving Object Detection Using Hadoop/MapReduce

Jyoti Parsola, Durgaprasad Gangodkar and Ankush Mittal

Abstract Technological advances and easy availability of low cost video camera has further encouraged users for deploying network of surveillance systems. These systems generate massive data. Thus, storage and processing of the huge video data for application such as video forensics, post event investigation etc., has emerged as a challenging problem to the research community. In this paper we propose a powerful approach that makes use of Hadoop Distributed File System (HDFS) for efficient storage of video data and programming model called MapReduce for data intensive computing. The proposed approach detects moving objects and provides their coordinates which can be used for localizing post event investigation. We have analyzed the storage and processing of video data of varying resolution and size to assess the performance of proposed approach.

Keywords Motion detection · Hadoop distributed file system · MapReduce

1 Introduction

Video surveillance has emerged as an active area of research, due to the increasing appeal for security and safety in public and private areas. Data is stored as evidence. If any event is triggered data can be used for investigation or identification. Reduction in the cost of cameras deployed for video surveillance makes it an effective tool for monitoring system. Consequently data generated by such appli-

J. Parsola (✉)

Department of Computer Applications, Graphic Era University, Dehradun, India
e-mail: jyotee.negi@gmail.com

D. Gangodkar · A. Mittal

Department of Computer Science and Engineering, Graphic Era University,
Dehradun, India
e-mail: dgangodkar@yahoo.com

A. Mittal

e-mail: dr.ankush.mittal@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_14

137

cations is massive as, data keeps on growing in seconds, and therefore storing such a huge data is a challenge. The second problem is, round the clock monitoring with human supervision is cumbersome and tiring. Processing large data efficiently becomes a complex problem as it consumes a lot of computation time. This calls for a measure which can not only process the data efficiently but also provide means to manage and store massive data.

Video processing task like motion detection and classification is an essential aspect of video surveillance system. To accelerate the video processing task like motion detection and resolving the storage problem, Hadoop MapReduce framework has been utilized in this paper.

Remaining paper is organized as follows. Section 2 surveys the relevant work done using Hadoop MapReduce. Section 2.1 introduces Hadoop framework and describes how motion is detected from videos using HDFS and MapReduce. Section 3 presents results performance evaluation and Sect. 4 discusses conclusions and future work.

2 Related Work

Ground work of data processing task with MapReduce was established by [1]. Hadoop is best suited for processing text data. Computational time of applications proposed by researchers like word count, sort etc. is reduced using MapReduce cluster. Some researchers have exploited video data processing with Hadoop video transcoding with MapReduce framework as discussed by [2]. Image processing techniques are applied to process video utilizing MapReduce cluster as proposed in [3]. [4] Presents an approach of parallel video processing and analysis using MapReduce cluster. Parallel processing of distributed database is performed by [5]. It transforms a colored image into grayscale image and features are extracted in parallel. The framework proposed in [6] demonstrates extensible processing videos with MapReduce. A face tracking system was implemented on top of the framework. A general cloud based platform is proposed in [7], which provides analytics and storage of video data. A novel metadata extraction and correction system is proposed in [8], to handle large amount of data produced due to surveillance system, data is distributed with Hadoop. Hadoop MapReduce is used for processing large resolution images and features are extracted in [9]. MapReduce framework is used for application like image retrieval on the basis of content as discussed in [10], with the purpose of combining an image analysis algorithm with text based image search engine with downgrading their response time. An image filtering technique using MapReduce programming model is proposed in [11] which requires images to be streamed only once as compared to the other distributed file system which requires a whole image or portion of image is streamed every time a filter is applied.

2.1 Hadoop Framework

Hadoop is an apache software framework used for distributed processing [12, 13, 14] and distributed storage of large dataset across cluster of nodes build on low cost computers. It consists of two main components. One is distributed file system named computer. It consists of two main components. One is distributed file system named as Hadoop Distributed File System (HDFS), second is execution engine or data processing framework MapReduce as shown in Fig. 1.

HDFS is designed to store huge files in a large number of clusters. It is a master slave architecture where master contains a single name node and slave contains multiple data node. NameNode contains metadata which manages the file system. DataNode stores data. Input files or data are split into block (default size 64 MB) and these are stored in data node. Data blocks are replicated (default value is 3) for fault tolerance and fast access.

MapReduce is a structure for parallel distributed processing large volume of data inspired by Google File System (GFS). It is master slave architecture, master runs JobTracker and slave executes TaskTracker as shown in Fig. 1. Objective of JobTracker is to divide jobs into task and decide where to run each task. It continuously communicates with the TaskTracker. A TaskTracker executes tasks, monitors execution of each task and continuously sends feedback to the JobTracker. MapReduce computing model consist of three components.

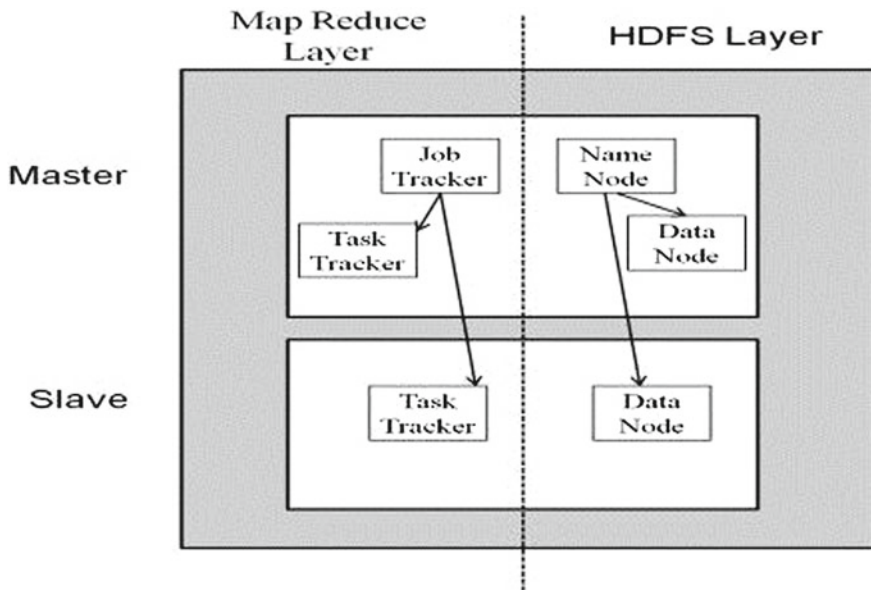


Fig. 1 Hadoop architecture [13]

- Map phase—Input data is split into data blocks and every block is processed by each map tasks in a parallel manner. It takes input as a key value pair as $\langle k, v \rangle$ and generates one or multiple $\langle K', V' \rangle$ pair as intermediate output.
- Shuffle phase—map output is shuffled and grouped by the key in construction for reduce phase.
- Reduce phase—it takes input as $\langle K', \text{LIST } V' \rangle$ where LIST V' is the list of all V' that are correlated with a key K' , an output is produced as a key-value pair.

2.2 *Processing Video Files and Detecting Motion with MapReduce*

Video file comprises of sequence of frames (images) therefore to process videos, individual frames are extracted. Hadoop works as its best for text files as discussed in [10], hence reading images/frames with Hadoop is a challenging task. Further size of a single image frame is relatively very small. Large set of video files result in huge amount of frame/images which are small in size whereas Hadoop efficiently reads files bigger in size and larger in number by distributing it to the various nodes. Therefore, processing small size files is again a complex problem. To resolve these problems extracted images are converted into sequence file. Sequence files are a Hadoop specific archive file format similar to tar and zip. Sequence file merges the file set using a key and a value pair where file name is used as the key and the file content is used as value. The size of sequence files generated is half the size of the original data. It occupies less memory space in HDFS thus making it storage efficient. These files can be split and processed in parallel.

Motion detection is the fundamental task of computer vision application. It segments region analogous to the moving object from the remaining frame established on approaches like subtracting background with the foreground [15], optical flow segmentation, segments moving regions by availing attributes of flow vector of moving regions measured with time [16], frame differencing [17]. Frame differencing is an efficient and easiest method for finding motion. Frame differencing, subtracts two frames separated at an interval of frame to segment moving objects.

Block matching approach is one of the effective method of frame differencing. It involves dividing the current frame and previous frame into blocks and block is searched with corresponding blocks as well as the neighboring blocks based on some criteria. Matching result provides a vector displaying movement of objects. Many approaches have been proposed for matching criteria for blocks like Sum of Absolute Difference (SAD), Sum of Squared difference (SSD), mean of absolute difference etc. SAD is the effective and efficient method for matching blocks. We followed approach proposed in [18] for finding motion detection in videos. As discussed in [18] previous frame is divided into square block of size $n \times n$ pixels called as motion block. Motion blocks are non-overlapping blocks and current frame is divided into overlapping search window of size $2n \times 2n$ pixels.

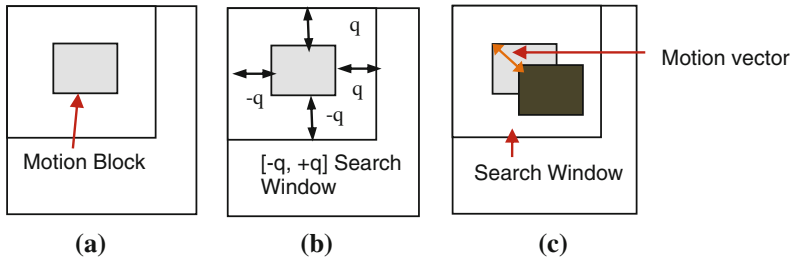


Fig. 2 Motion detection view (a) motion block in previous frame (b) overlapping block in current frame and (c) motion vector within search window [18]

Block matching strategy as shown in Fig. 2, every motion block is inspected into respective search window using (1). In (1) and (2) $I_{per}(\cdot)$ and $I_{cur}(\cdot)$ indicates the pixel intensity in the previous and current frame respectively. SAD (a, b) is the sum of absolute difference at the pixel location a, b. $[-q, q]$ is the search region in the search window. MV refers to the motion vector at minimum value of SAD calculated between frames separated by time interval t_1 and t_2 . The size of motion block is chosen for 16×16 and search window size is of 32×32 pixels.

$$SAD(a, b) = \sum_{i=0}^{j=N-1} \sum_{i=0}^{j=N-1} |I_{per}(i, j, t_1) - I_{cur}(i + a, j + b, t_2)| \quad (1)$$

where $(a, b) = \{-q \leq a, b \leq q\}$

$$\text{Motion vector (MV)} = (a, b) | \min \text{ sad } (a, b) \quad (2)$$

The task of calculating motion vector in the range between $[-31, 32]$ into frames of a video is computationally expensive so search is initialized when motion block is on the same location of the reference frame i.e. block and search window is overlapping in the center. If value of SAD is zero then there is no motion of block and if blocks involve motion then block display the highest value of absolute difference SAD_0 is computed as.

$$SAD_0 = \sum_{i=0}^{j=N-1} \sum_{i=0}^{j=N-1} |I_{ref}(i, j, t_1) - I_{cur}(i, j, t_2)| \quad (3)$$

where i, j refers to the pixel location in previous and current frame.

In addition a threshold is applied to SAD_0 to reduce the computation time. It decides to initialize the search or not as based on (4).

$$\text{Search Decision} = \begin{cases} 1 & \text{If } SAD_0 < \text{threshold} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

For every block moderate threshold can be fixed as portion of $256 \times 15 = 3840$ where, 256 is 16×16 block pixel value. Portion can vary from 0.4 to 0.1.

2.3 Reducing SAD Computation Time with Hadoop

Calculating sum of absolute difference for large number of videos is computationally expensive task. Apache Hadoop provides a framework to store as well as compute up to terabyte of data efficiently. Processing flow of motion detection in video is shown in Fig. 3. Frames are extracted from video for motion detection. In order to reduce the unnecessary storage space in HDFS, frames with a certain time interval are considered for motion detection as, consequent frames do not differ much. Frames are converted into sequence file. In Sequence file image files are merged in a key value pair where Key is the name of the frame and value is the contents of the frame. Sequence or input file is split into blocks (default value of each block is 64 MB).

Hadoop processes data in a batch processing manner and moreover rather than sending the data to the computation it sends computation where data is residing. This overcomes data transfer latency resulting into reduction in execution time. Frames are passed to Mapper. We provide two alternate fifth frame to a single Mapper which converts each frame into a overlapping 32×32 blocks. Intermediate output generated by Mapper is, *Key* with an integer value and *Value* contains 32×32 pixel blocks as well as the name of frame containing it. *Value* possessing the same *Key* is moved to the single reducer. Each Reducer performs SAD calculation on 32×32 blocks with the unique *key* value. The final output of Reducer

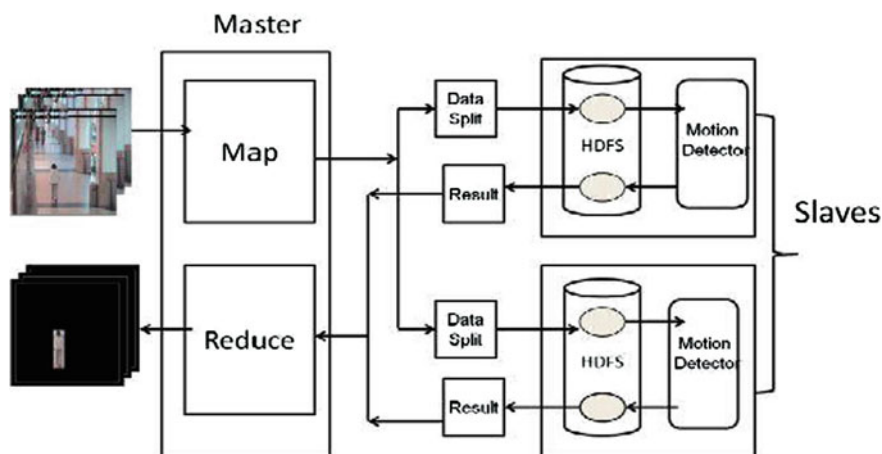


Fig. 3 MapReduce flow for motion detection

is the collection of motion vectors. By processing all the image files reducer accumulates results and stores in a text file with *key* as the frame number and *Value* as list of motion vectors with respect to the frame.

3 Results Along with Performance Evaluation

The implementation of proposed framework is performed on Ubuntu 12.04 as an operating system Hadoop version is 1.2.1. The processor of the device is Intel core i5 1.70 GHz with 4 GB of memory. To check how file size and data size affects the computation time in Hadoop MapReduce cluster, motion detection is determined on grayscale video frames having pixel size 256×256 and 512×512 . In case of colored frames, firstly frames are converted into grayscale. Figures 4, 5, 6, 7 shows the motion detection obtained on different video sequences. Motion vectors obtained are plotted with the help of JavaCV which is wrapper for OpenCV library [19]. Results have been calculated on different dataset archives available publically. Performance Evaluation of Tracking and Surveillance (PETS) [20], LITIV Datasets Ecole.Polytechnique Montreal [21] and OCTBVS Benchmark Dataset collection, Ohio State University [22]. Context Aware Vision using Vision Based Recognition [23].

Execution time for 256×256 and 512×512 pixel size video frames in sequential and Hadoop MapReduce single node cluster is shown in Figs. 8 and 9 respectively. X axis is the data size of videos and y axis is processing time in seconds. Processing time for small data set does not differ much whereas speed up execution time is obtained for big data size. The comparison on processing of both pixel size video frames in Fig. 10 clearly shows execution time decreases for higher size of data as well as for big file size. Reduction in computation time in comparison to sequential using Hadoop MapReduce single node cluster is 34 % in case of 256×256 pixel video frames and 40 % for 512×512 pixel video frames.

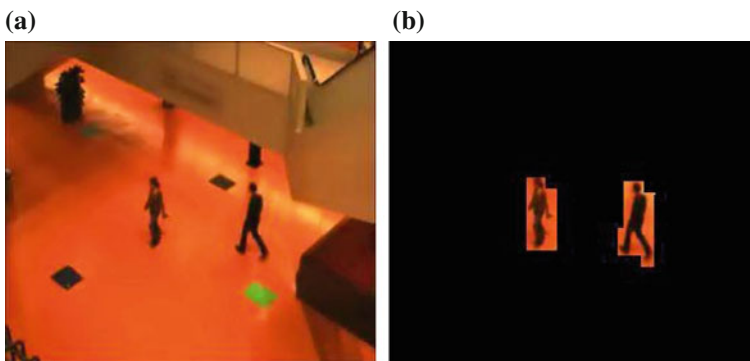


Fig. 4 LITIV Sequence4 [21]. **a** Actual frame. **b** Identified moving objects

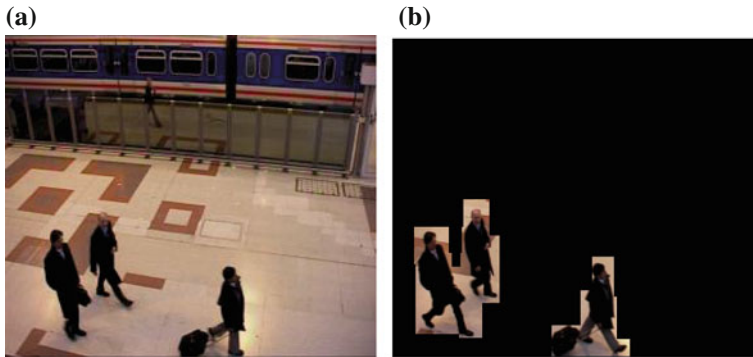


Fig. 5 PETS S1-T1 sequence [20]. **a** Actual frame and, **b** Identified moving objects

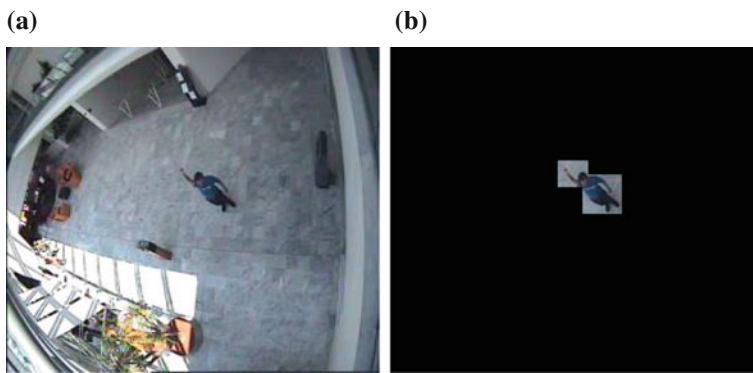


Fig. 6 CAVIAR Walk [23]. **a** Actual frame. **b** Identified moving objects

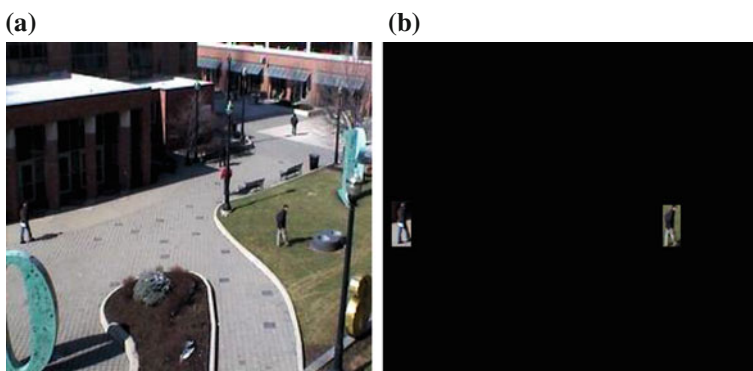


Fig. 7 OTCBVS, Sequence 4b [22]. **a** Actual frame. **b** Identified moving objects

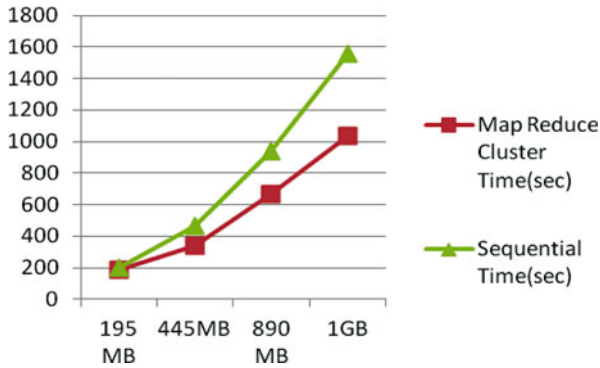


Fig. 8 Comparison of processing time for 256×256 images in sequential and MapReduce cluster

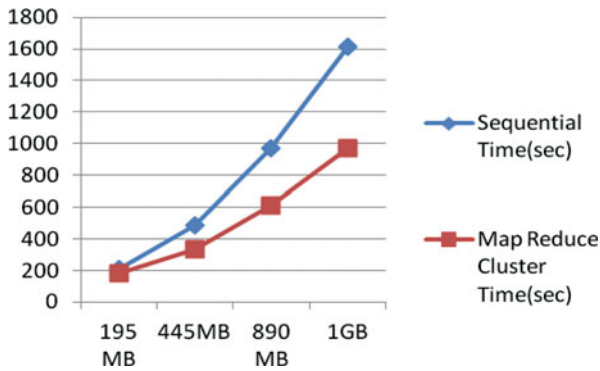


Fig. 9 Comparison of processing time for 512×512 images in sequential and MapReduce cluster

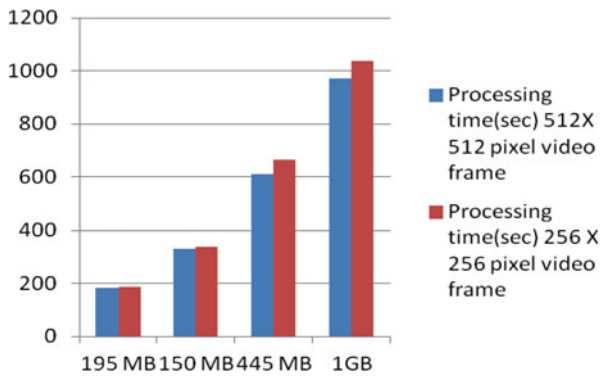


Fig. 10 Comparison of processing time for 256×256 and 512×512 pixel size video frame in MapReduce single node cluster

4 Conclusion and Future Work

An efficient approach has been implemented for finding motion detection in large data size using Hadoop MapReduce. We have shown that the computational time is reduced as pixel resolution of images is increased. We could not check it for higher resolutions images as dataset available is small in size. Computation time can be reduced further by increasing nodes in a cluster. Execution time in a cluster is also affected by network latency, resolving this problem execution time can be more optimized. MapReduce framework works best for application which require similar job to be executed in different data size. The proposed framework can be utilized for post event investigation application like theft detection, criminal activities, abandoned object detection, suspicious behavior detection etc.

References

1. Dean, J. and Ghemawat, S.: Map Reduce: Simplified Data Processing on Large Cluster. Operating Systems Design and Implementation (OSDI). San Francisco, United States (2004) 137–150.
2. Pereira, R. and Breitman, K.: A Cloud Based Architecture for Improving Video Compression Time Efficiency: The Split & Merge Approach. 3rd IEEE Int. Conf. on Cloud Computing (CLOUD). Miami, United States July (2010) 482–489.
3. Schmidt, R. and Rella, M.: An approach for processing large and non-uniform media objects on MapReduce-based clusters. 13th Int. Conf. on Asia-Pacific Digital Libraries, (ICADL). Beijing, China (2011).
4. Tan, H. and Chen, L.: An approach for fast and parallel video processing on Apache Hadoop clusters. IEEE Int. Conf on Multimedia and Expo (ICME). Chengdu, China (2014) 1–6.
5. Yamamoto, M. and Kaneko, K.: Parallel image database processing with MapReduce and performance evaluation in pseudo distributed mode. Int. J. of Electronic Commerce Studies, Vol. 3, No. 2 (2012) 211–228.
6. Ryu, C., Lee, C. M. Jhang, Kim, C., Seo, E.: Extensible video processing framework in apache hadoop. in Proc. IEEE 5th Int. Conf on Cloud Computing Technology and Science (CloudCom), Vol. 2, Bristol, England, Dec (2013) 305–310.
7. Zhang, W., Xu, L., Duan, P., Gong, W., Liu, X. and Q. Lu.: Towards a High Speed Video Cloud Based on Batch Processing Integrated with Fast Processing. in Proc. of Int. Conf of Identification, Information and Knowledge in the Internet of Things (IIKI), Beijing, China (2014) 28–33.
8. Zhao, X., H. Ma, H., Zhang, H., Tang, Y. and Fu, G.: Metadata extraction and correction for large-scale traffic surveillance videos. in Proc. on Big Data (Big Data), Washington, DC, United States (2014) 412–420.
9. Zhu, H., Shen, Z., Shang, L. and Pang, X.: Parallel Image Texture Feature Extraction under Hadoop Cloud Platform. Intelligent Computing Theory. Springer International Publishing (2014) 459–465.
10. Premchaiswadi, W., Tungkatsathan, A., Intarasema, S. and Premchaiswadi, A.: Improving performance of content-based image retrieval schemes using Hadoop MapReduce. In Int. Conf. of High Performance Computing and Simulation (HPCS). Helsinki, Finland (2013) 615–620.

11. Gamage, T.D, Samarvikrama, J.G., Rodrigo, R.and Pasqual, A. A.: Image filtering with MapReduce in pseudo-distribution mode. in IEEE Conf. of Moratuwa Engineering Research Conference (MERCCon). Moratuwa, Sri Lanka (2015) 160–164.
12. APACHE, 2010. Hadoop MapReduce framework. Available <http://hadoop.apache.org/mapreduce/>.
13. T. White. Hadoop: “The Definitive Guide”. Yahoo Press, 2010.
14. Holmes, Alex, “Hadoop in practice”, Manning Publications Co., 2012.
15. Grimson, W.E.L. and Stauffer, C.: Adaptive background mixture models for real-time tracking, in Proc. Of IEEE Conf. Computer Vision and Pattern Recognition, Vol. 2, Fort Collins, United States (1999) 22–29.
16. Lim, S., Apostolopoulos, J.G. and Gamal, A.E.: Optical flow estimation using temporally oversampled video, IEEE Trans. of Image Processing, Vol. 14, No. 8(2005) 1074–1087.
17. Yu, Y. and Chen, Y.: A real-time motion detection algorithm for traffic monitoring systems based on consecutive temporal difference. 7th Conf of Asian Control Conference (ACC). Hong Kong, China (2009) 1594–1599.
18. Gangodkar, D., Kumar, P. and Mittal, A.: Segmentation of moving objects in visible and thermal videos. In Int. Conf of Computer Communication and Informatics (ICCCI), Coimbatore, India (2012) 1–5.
19. Open Source Computer Vision (OpenCV) [Online]. Available: <http://opencv.willowgarage.com/wiki/>.
20. Performance Evaluation of Tracking and Surveillance, (PETS).Available: <http://www.cvg.cs.rdg.ac.uk/slides/pets.html>.
21. Ecole Polytechnique, Montreal, LITIV Datasets Available: <http://www.polymtl.ca/litiv/en/vid/index.php>.
22. OTCBVS Benchmark Dataset Collection, Ohio State University. Available: <http://www.cse.ohio-state.edu/otcbvs-bench/>.
23. Context Aware Vision Using Image-Based Active Recognition (CAVIAR).Available: <http://groups.inf.ed.ac.uk/vision/CAVIAR/html>.

Performance Evaluation of Digital Color Image Watermarking Using Column Walsh Wavelet Transform

Hemant B. Kekre, Shachi Natu and Tanuja Sarode

Abstract This paper proposes a robust watermarking technique using wavelet transform generated from well-known orthogonal transform Walsh. Watermark embedding is done in middle frequency band of column wavelet transformed host image. Performance of proposed technique is evaluated against image processing attacks like compression, cropping, addition of run length noise with binary and Gaussian distribution and image resizing. Comparison of performance of these transforms is done on the basis of robustness to attacks using Mean Absolute Error (MAE) as a metric of robustness. Column wavelet is found preferable over full wavelet. Also column Walsh wavelet is preferable over column DCT wavelet for robustness.

Keywords Watermarking • Walsh wavelet transform • Column wavelet • Run length noise

1 Introduction

Growth of internet has led to transmission of data in digital form. Unauthorized copies of these digital data can be created and it becomes necessary to prove ownership of digital data to avoid copyright abuse. Digital watermarking is a popular method which provides copyright protection. Visible and invisible watermark can be used to protect the copyright. Visible is usually not preferred as it

H.B. Kekre
MPSTME, NMIMS University, Mumbai 400056, India
e-mail: hbkekre@yahoo.com

S. Natu (✉)
MPSTME, Vileparle, Mumbai 400056, India
e-mail: shachi_natu@yahoo.com

T. Sarode
Thadomal Shahani Engineering College, Mumbai 400050, India
e-mail: tanuja_0123@yahoo.com

spoils the beauty of digital host and also position of watermark gets disclosed. This makes tampering of watermark easy. Popular domains for performing watermarking in digital images are spatial domain where pixel values of image are altered to insert watermark. Another is frequency domain in which image is transformed into frequency domain and then frequency coefficients of image are altered to insert the watermark. Depending on the application context of the watermarking, requirements of watermark differ. To identify the origin of contents, a single watermark should be embedded at the source of distribution [1]. For access control, watermark needs to be checked at every recipient. For this semi blind or blind watermarking scheme is required [1]. Usually there is a trade-off between robustness and invisibility of watermark which are the most desirable characteristics of good watermarking mechanism. Robustness can be made high by hiding watermark into perceptually significant regions of host. However this leads to reduced invisibility of watermark. To achieve high invisibility of watermark, it can be embedded in perceptually insignificant regions of host. Such regions are the desired candidates for elimination or alteration during majority of geometric and signal processing attacks. In this paper a frequency domain watermarking technique is proposed using wavelet transforms generated from Walsh transform. Using suitable combination of Walsh transform size, Walsh wavelet transform is generated. Column Walsh wavelets transform is tried in the experimental work. Performance of column Walsh wavelet is evaluated against various attacks such as compression, cropping, noise addition and image resizing. Each of these attacks is performed using some variations in it. Organization of paper is as follows: Sect. 2 presents the related work in the field of digital image watermarking. Section 3 explains the proposed method. Discussion of Results and discussion for performance of proposed technique against various attacks is given in Sect. 4. Comparison of column Walsh wavelet with full Walsh wavelet and Column DCT wavelet is given in Sect. 5. Conclusion of paper is presented in Sect. 6.

2 Related Work

Deb, Al-Seraj, Hoque and Sarkar have proposed a wavelet and DCT based watermarking technique [2] in which watermark is embedded as a bit stream in low frequency DCT coefficients of HL frequency components. The technique is proved to be robust for JPEG compression, contrast adjustment, cropping and noise attacks. In [3], Lee et al. proposed a multiple digital watermarking scheme. It has two phases, Grayscale watermark phase and binary watermark phase. In first phase, grayscale watermark is embedded in original image by dividing it into blocks and calculating average of each block. In second phase, binary watermark is embedded by generating polarity matrices from original image and grayscale watermark. Reverse procedure is followed to extract the watermarks. Wang et al. [4] proposed DWT-SVD based watermarking scheme. Single level wavelet decomposition is applied to image. SVD is applied to LL sub-band. Bit stream is generated from LL

band and U, V matrices. Chaotic sequence is generated and then converted into binary and this binary sequence is then hashed to generate a key. These secret key and initial value i , are registered in a third party intellectual property for copyright protection. Song et al. [5], proposed a semi-fragile watermarking scheme based on wavelets. Watermark is generated from 3rd level wavelet coefficients of LL frequency band by condition judgment based on the mean value. Watermark is then embedded into predetermined bit plane of LH2, HL2 and HH3 frequency bands. In [6], pixelwise image watermarking has been proposed in spatial domain. Watermark is embedded by using Human Visual System characteristics to improve robustness and imperceptibility. Watermark strength is modulated according to local image characteristics. In [7], Keerthi et al. proposed wavelet domain watermarking technique with varying embedding strength. With varying strength value, better robustness and imperceptibility is observed than watermarking techniques proposed by Barni in [6] and by Naforita in [8]. Jiansheng et al. [9] proposed a robust and invisible watermarking based on DCT and DWT in which host image is subjected to three level wavelet transform and Discrete Cosine transformed watermark is inserted into it. Singh, Rawat and Agrawal also proposed a watermarking technique based on DCT and DWT in which scrambled watermark is inserted into HH3 band of host image using Arnold transform [10]. Li, Yap and Lei proposed a watermarking technique based on DCT and SVD in [11]. In this technique, SVD is applied to cover image. First singular values are selected to form macro block and DCT is applied on it. Particular relationship is imposed between pairs of the DCT coefficients that are selected from SVD-DCT block pseudo randomly and then watermark is hidden in high frequency band of the block. Kekre et al. presented a DCT wavelet transform based watermarking technique [12, 13]. Watermark is compressed before embedding with acceptable loss of information from image. Embedding compressed image also reduces the payload of information embedded in host image and thus causes good imperceptibility of watermarked image.

3 Proposed Method

In the proposed method, Walsh wavelet transform matrix is generated from Walsh transform using algorithm of wavelet generation given by Kekre in [14]. Since our experimental work consists of fifteen host images each of size $256 \times 256 \times 3$ bytes and five watermarks of size $128 \times 128 \times 3$ bytes, wavelet transform matrices of size 256×256 and 128×128 are required. According to algorithm of wavelet generation in [14], these matrices can be generated using different size combinations of orthogonal transforms like (64, 4), (32, 8), (8, 32), (4, 64) and (2,128). Similarly 128×128 size transform matrix can be generated using pairs (64, 2), (32, 4), (16, 8), (8, 16), (4, 32) and (2, 64). Proposed watermark embedding process is illustrated in Fig. 1.

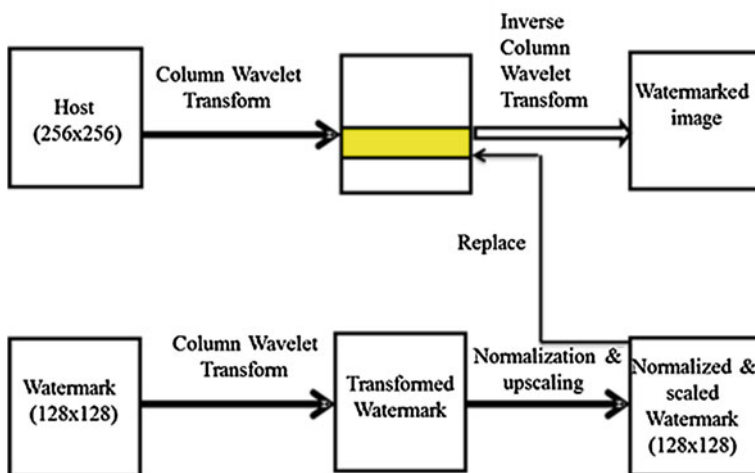


Fig. 1 Watermark embedding process

First, wavelet transform matrices of size 256×256 and 128×128 are generated using suitable combination of Walsh transform. Column wavelet transform of host image and watermark image are obtained using wavelet transform matrix. Middle frequency region of transformed host is chosen for embedding watermark. For column transform, it is middle rows due to energy concentration taking place at upper region of transformed image. Rows 108–171 of transformed host are chosen as middle frequency region to embed the watermark. Wavelet transform coefficients of watermark are normalized to bring them in a range $[0, 1]$. Energy gap between middle frequency region of transformed host and normalized watermark is filled by scaling up the normalized transform coefficients of watermark using suitable scaling factor. From one host image to other this scaling factor varies. Scaled coefficients of watermarks are replaced in middle frequency region of host. Inverse column transform of host gives us the watermarked image. Another aspect of our experimental work is to observe the performance of proposed method when the energy gap between middle frequency region of host and watermark is zero and $\pm 40\%$. Thus energy of watermark is maintained same as that of energy of middle frequency region where it is embedded or it is 60% of the energy of middle frequency region or it is 140% of the energy of middle frequency region. Figure 1 shows the flow of embedding procedure. Watermark extraction process is reverse of embedding. To extract the watermark, column Walsh wavelet transform of watermarked image is obtained. Middle frequency band is extracted out of it. This middle frequency band is scaled down using same scaling factor used in embedding process to scale up followed by denormalization. This process results into extraction of transform coefficients of watermark. Inverse column wavelet transform of these coefficients gives us extracted watermark. Quality of extracted watermark is compared with that

of embedded watermark by calculating Mean Absolute Error (MAE) between them. Smaller the error value better is the robustness.

4 Results and Discussions

As mentioned earlier the best size combination for generation of column Walsh wavelet is found to be (64, 4) to get 256×256 matrix and (32, 4) to generate 128×128 transform matrix.

4.1 Compression Attack

The most obvious attack that can be performed on images before transmitting them over a network is compression as we always strive for optimal usage of available bandwidth. In the proposed technique, compression of watermarked images is done with three variations. First is using transformed based compression. In this type various transforms like DCT, DST, Walsh, Haar (with compression ratio 1.142) and DCT wavelet with compression ratio 1.954 is used. Second, using JPEG compression with quality factor 100 and third using Vector Quantization obtained using Kekre’s Fast Codebook Generation algorithm (KFCG) [15] with codebook size 256. Watermarked images after compression using DCT, JPEG and VQ with watermark recovered from them are shown in Fig. 2. As a representative case of results where the case of embedding watermarks with zero energy difference is considered. Watermarked image Bird and recovered watermark austral are shown for DCT compression, JPEG compression and VQ compression. Below watermarked image MAE between original and watermarked image is mentioned while below extracted watermark image MAE between embedded and extracted watermark is mentioned.

Table 1 summarizes Average of MAE values between embedded and extracted watermark against various compression attack. These MAE values are summarized over fifteen host and five watermarks separately embedded in them. As can be seen from Table 1, proposed method shows excellent robustness against compression







					
0.862	3.897	1.642	46.637	4.214	38.895
DCT compression		JPEG compression		VQ compression	

Fig. 2 Result images for compression attack

Table 1 Average MAE values against compression attack

Compression type	DCT	DST	Walsh	Haar	DCT Wavelet	JPEG	VQ
Average MAE	8.217	8.487	0	0	59.930	31.029	40.551

using Walsh and Haar followed by compression using DCT and DST. Comparatively it gives poor robustness against JPEG, VQ and DCT wavelet compression.

4.2 Cropping Attack

Image cropping is done by removing image portion of size 16×16 and 32×32 at four corners and 32×32 at center. Figure 3 shows sample result images for cropping.

Figure 3 shows that column Walsh wavelet shows excellent robustness against cropping at corners and at center as cropping done in spatial domain does not cause loss of watermark information in transform domain. That is transform coefficients corresponding to cropped pixel values do not contain significant information about watermark. Table 2 summarizes the performance of column Walsh wavelet against cropping attack. From Table 2 it can be concluded that when number of cropped pixels are same but their location is different, MAE between embedded and extracted watermark is varying. MAE is less when cropping is done at center and more for cropping at corners.

4.3 Noise Addition Attack

Two variations of noise addition attack are performed. In first type, binary distributed run length noise with discrete magnitude $+1$ or -1 noise is added and in second type, Gaussian distributed run length noise with discrete magnitude between -2 and $+2$ is added to watermarked images. Binary distributed run length noise is increased in runlength from the range of $1-10$ to $5-50$ and then $10-100$. Figure 4

Fig. 3 Result images for cropping attack





			
3.135	0.341	2.513	0.173
16x16 cropping at corners		32x32 cropping at center	

Table 2 Average MAE values against cropping attack

Cropping type	16 × 16 at corners	32 × 32 at corners	32 × 32 at center
Average MAE	0.396	4.244	0.158

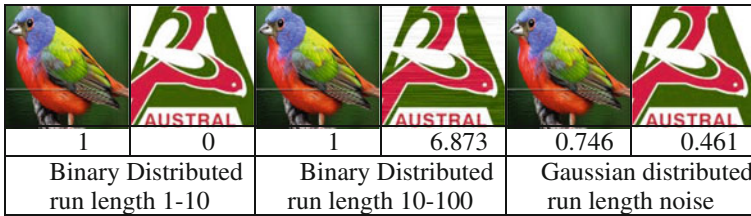


Fig. 4 Result images for cropping attack

Table 3 Average MAE values against noise addition attack (BRLN = Binary Distributed Run Length Noise, GRLN = Gaussian Distributed Run Length Noise)

Noise type	BRLN (1–10)	BRLN (5–50)	BRLN (10–100)	GRLN
Average MAE	0	10.543	10.453	0.721

shows sample result image against noise addition attack which indicate strong robustness.

Table 3 gives MAE values between embedded and extracted watermark averaged over fifteen hosts and five watermarks. Values in bracket in each column heading of table indicate run length of noise.

Table 3 shows that proposed method using column Walsh wavelet shows high robustness against both the types of noises, especially Gaussian distributed run length noise and smaller run length of binary distributed run length noise.

4.4 Resizing Attack

In resizing attack, watermarked image is first increased in size two times of its original size and then reduced back to original size. This is done using bicubic interpolation, grid interpolation [16] and various transforms like DCT, DST, DFT, Hartley, and Real Fourier transform as proposed in [17]. Sample images for this attack are shown in Fig. 5.

Figure 5 shows that proposed method using column Walsh wavelet is highly robust against resizing using DCT. Comparatively it shows less robustness against resizing using bicubic interpolation and grid interpolation. Table 4 shows summarized MAE value between embedded and extracted watermark for resizing attack.







					
1.487	18.662	0.531	44.578	0	0
Resizing using Bicubic interpolation		Resizing using Grid interpolation		Resizing using DCT	

Fig. 5 Result images for resizing attack

Table 4 Average MAE values against resizing attack

Resizing Type	Bicubic interpolation	Grid interpolation	DCT/DST/DFT/RFT/Hartley
Average MAE	20.451	22.285	0

Table 4 shows that column Walsh wavelet yields highest robustness against resizing using transforms. For resizing using bicubic and grid interpolation it is not very good in robustness.

5 Comparison of Column Walsh Wavelet, Full Walsh Wavelet and Column DCT Wavelet Against Various Attacks

Performance of column Walsh wavelet when compared to full Walsh wavelet, it is observed that column transform gives better robustness than majority of the attacks. Also these results when compared to column DCT wavelet, column Walsh wavelet proves to be better than column DCT wavelet. This comparison for each of the attack is shown in Figs. 6, 7, 8 and 9.

From Fig. 6 statistical improvement in robustness is as follows. Column Walsh wavelet is observed to be 100 % more robust than full Walsh wavelet against compression using Walsh and Haar, 69 and 71 % more robust against DST and DCT compression and 7 % better in robustness against JPEG compression. Column Walsh wavelet is also observed to be 100 % more robust than column DCT wavelet against compression using Walsh and Haar, 49 and 51 % better in robustness against compression using DST and DCT and 19 % more robust against JPEG compression respectively. Figure 7 shows performance comparison of column Walsh wavelet with full Walsh wavelet and column DCT wavelet against cropping attack.

As can be seen from Fig. 7, column Walsh wavelet shows significant improvement in robustness by reducing the MAE values by 90–98 % for different types of cropping attacks. For column DCT wavelet this improvement is 15 % for

Fig. 6 Comparison of Column Walsh wavelet, full Walsh wavelet and column DCT wavelet against compression attack

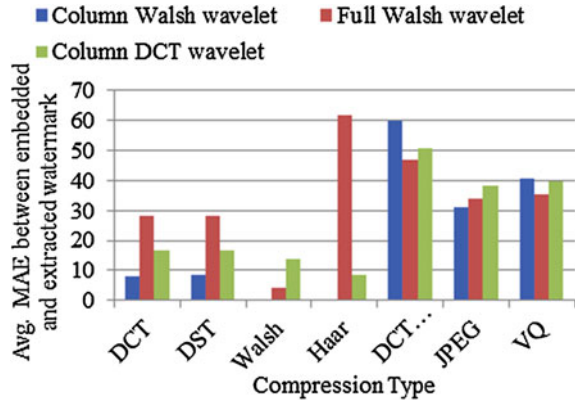


Fig. 7 Comparison of Column Walsh wavelet, full Walsh wavelet and column DCT wavelet against cropping attack

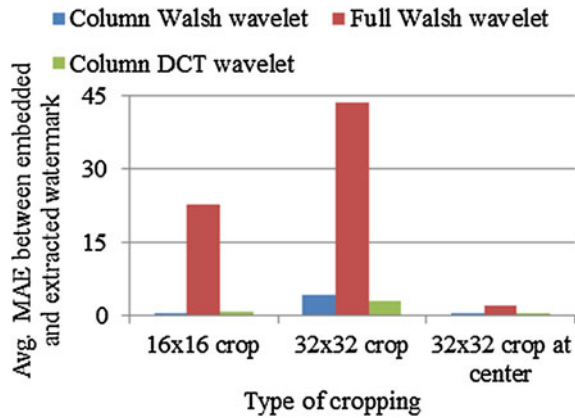


Fig. 8 Comparison of Column Walsh wavelet, full Walsh wavelet and column DCT wavelet against noise addition attack

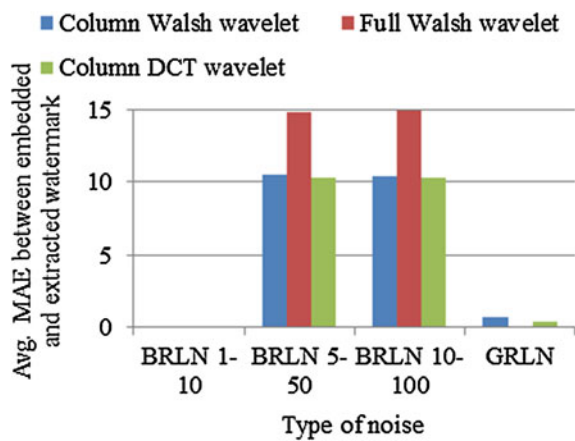
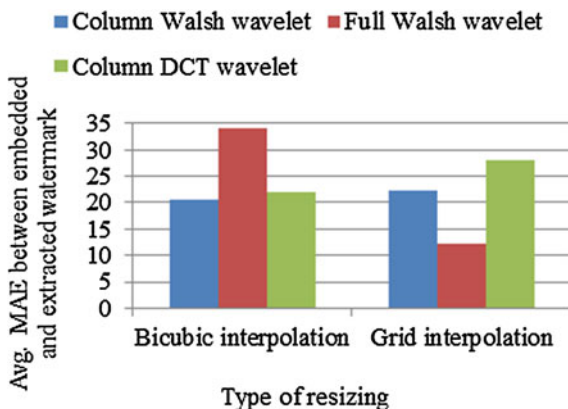


Fig. 9 Comparison of Column Walsh wavelet, full Walsh wavelet and column DCT wavelet against resizing attack



cropping 32×32 at center and 47 % for cropping 16×16 at corners. Figure 8 shows the comparison of MAE values for noise addition attack obtained when column Walsh wavelet, full Walsh wavelet and column DCT wavelet is used to embed the watermark.

Figure 8 shows that performance of column Walsh wavelet and column DCT wavelet give 28–30 % better robustness than full Walsh wavelets transform against noise addition attack. Figure 9 shows comparison of three transforms against resizing using bicubic interpolation and grid interpolation. Since for resizing using various transforms, MAE between embedded and extracted watermark is zero, they are not shown in graph.

Column Walsh wavelet is 40 % more robust in resizing using bicubic interpolation than full Walsh wavelet. Column Walsh wavelet is also more robust than column DCT wavelet in both the types of resizing by 6 and 20 % respectively.

6 Conclusions

From various results presented in the paper it can be concluded that column Walsh transform is more robust than full Walsh wavelet against many attacks like compression in the range of 69–100 %, against cropping in the range of 90–98 %, against noise addition from 28 to 30 % and against resizing attack by 40 %. Column Walsh wavelet is also found more robust than column DCT wavelet for attacks like compression by 19–51 %, against cropping by 15–47 %, cropping and against resizing attack by 6–20 %.

References

1. Emir Ganic, Ahmet Eskicioglu: Robust embedding of visual watermarks using DWT-SVD. *Journal of Electronic Imaging*, Volume 14(4), (2005).
2. Kaushik Deb, Md. Sajib Al-Seraj, Md. Moshiul Hoque and Md. Iqbal Hasan Sarkar: Combined DWT-DCT Based Digital Image Watermarking Technique for Copyright Protection. In *IEEE proc. of 7th International Conference on Electrical and Computer Engineering*, (2012) 458–461.
3. Gil-Je Lee, Eun-Jun Yoon, Kee-Young Yoo: A Novel Multiple Digital Watermarking Scheme for the Copyright Protection of Image. In *Proc. of 4th International Conference on Innovative Computing, Information and Control*, (2009) 756–759.
4. Wei Wang, Aidong Men, Bo Yang and Xiaobo Chen: A novel robust zero watermarking scheme based on DWT and SVD. In *proc. of IEEE 4th International Congress on Image and Signal Processing*, (2011) 1012–1015.
5. Haohao Song, Zihua Qiu, Jian Gu: A Novel Semi-fragile Image Watermarking Scheme Based on Wavelet. In *Proc. of IEEE, ICALIP*, (2010) 1504–1510.
6. Mauro Barni, Franco Bartolini, and Alessandro Piva: Improved wavelet-based watermarking through pixel-wise masking. *IEEE Transactions on Image Processing*, Vol. 10, No. 5 (2001) 783–791.
7. M. S. Keerthi, Nair Prithi, M. Jaymohan: An improved wavelet domain image watermarking with varying embedding strength. *World Congress on Information and Communication Technologies* (2012) 404–407.
8. C. Nafornita: Contributions to digital watermarking of still images in the wavelet transform. Ph. D Thesis, Technical University of Cluj-Napoca, Romania (2008).
9. Mei Jiansheng, Li Sukang, Tan Xiomeri: A digital watermarking algorithm based on DCT and DWT. *International symposium on web information systems and applications* (2009) 104–107.
10. Surya pratap Singh, Paresh Rawat, Sudhir Agrawal: A robust watermarking approach using DCT-DWT. *International journal of emerging technology and advanced engineering*, vol. 2, issue 8 (2012) 300–305.
11. Zhen Li, Kim-Hui Yap and Bai-Ying Li: A new blind robust image watermarking scheme in SVD-DCT composite domain. *18th IEEE international conference on image processing* (2011) 2757–2760.
12. H. B. Kekre, Tanuja Sarode, Shachi Natu: Performance of watermarking system using wavelet column transform under various attacks. *International Journal of Computer Science and Information Security*, Vol. 12, No. 2 (2014) 30–35.
13. H. B. Kekre, Tanuja Sarode, Shachi Natu: Robust watermarking scheme using column DCT wavelet transform under various attacks. *International Journal on Computer Science and Engineering*, Vol. 6, No. 1 (2014) 31–41.
14. Kekre H. B., Archana Athawale, Dipali Sadavarti: Algorithm to Generate Wavelet Transform from an Orthogonal Transform. *International Journal of Image Processing (IJIP)*, 4.4 (2010) 444–455.
15. Kekre H. B., Tanuja K. Sarode: Fast Codebook Generation Algorithm for Colour Images using Vector Quantization. *International Journal of Computer Science and Information Technology* 1.1 (2009) 7–12.
16. H. B. Kekre, Tanuja Sarode, Sudeep Thepade: Grid based image scaling technique. *International Journal of Computer Science and Applications*, Volume 1, No. 2 (2008) 95–98.
17. H. B. Kekre, Tanuja Sarode, Shachi Natu: Image Zooming using Sinusoidal Transforms like Hartley, DFT, DCT, DST and Real Fourier Transform. *International journal of computer science and information security*, vol. 12, no. 7 (2014) 11–16.

Structural (Shape) Feature Extraction for Ear Biometric System

P. Ramesh Kumar and S.S. Dhenakaran

Abstract The Ear Biometrics is an emerging modern human identification system, which is a passive biometrics system where the recognition of an individual is done without the knowledge of the human subject. This kind of passive biometrics system can help the image analyst to extract useful information feature for public surveillance system. The proposed article extracts structure (shape) features such as Area, Perimeter, Eccentricity, Elongation, Compactness, Horizontal Height, Vertical Height, Major Axis, Minor Axis, Circularity and Rectangularity of the human ear to construct an effective ear biometrics system.

Keywords Structural • Shape • Feature extraction • Surveillance

1 Introduction

The Biometrics is the process of identifying the authorization of an individual human based on physical parts such as finger, palm, Iris etc., In general, the biometric system can be divided based on active and passive involvement of the subject into the system. The problem with the active biometrics system is subject cooperation, not hygiene, lost feature. To overcome the issues in the active biometrics system, the passive biometrics system is emerging where subject involvement is not required. The surveillance officer gets the required biometric physical image and extract the feature to authenticate the individual human.

P.R. Kumar (✉)

Department of Computer Science and Engineering, VR Siddhartha Engineering College, Vijayawada 520 007, Andhra Pradesh, India
e-mail: send2rameshkumar@gmail.com

S.S. Dhenakaran

Department of Computer Science and Engineering, Alagappa University, Karaikudi, India
e-mail: ssdarvind@yahoo.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_16

161

The Ear biometrics [1–7] is one type of passive biometric system where the Left and Right ear image of a subject can be captured; the interesting unique features can be extracted to verify the individual. The future of the ear biometrics can be coupled with video surveillance system, where ear image can be easily cropped and computed for verification.

The researches in automatic ear recognition [8] are emerging to validate human in a public place to stop intruders. The stability of the ear is also constant; the local features are more and more stable. The physical ear is a hearing sensor of the human, so it cannot be hidden from a surveillance camera.

The structural features are based on the shape of the input ear image and this information on the ear shape can be analyzed for understanding unique feature present in each ear image in the verification process. The input image can be processed for its structural related features and its stability.

2 Methodology

In this section, we Proposed a structural (shape) based feature extraction for ear biometrics system to verify the authentication. Here, for experimentation, we take input as RGB images from the AMI ear dataset [9]. Pre-processing the images, i.e. convert the RGB image folder into the gray scale image folder and resize the images into the fixed size then crop the images into a particular shape.

Extract the structure or shape features such as Area, Perimeter, Eccentricity, Elongation, Compactness, Horizontal height, vertical height, major axis, minor axis, circularity, rectangularity. Export these parameter features into the local feature database. In verification, take tests random sample image and extract the features. Then verify these features with the local feature, whether the sample image feature present in the database or not. Finally, using two methods false match rate and the false mismatch rate, we check the performance of the system.

The structural (shape) feature extraction methods for ear biometric system consists of the following phases

- i. Enrollment
- ii. Verification

The enrollment is the process of registering the individual's ear image features into the local database for future verification. The enrollment process consists of the following stages

1. Ear image Acquisition
2. Preprocessing (resizing, Cropping)
3. Convert to Gray level Image
4. Structural (Shape) Feature Extraction
5. Create the Local Feature DB (Fig. 1)

Fig. 1 Gray scale image of the AMI Ear database [9]



The verification process authenticates the individual ear image feature with a collection of local feature database. The verification process consists of the following steps

1. Capture the probe Ear Image
2. Preprocessing (resizing, Cropping)
3. Convert to gray level image
4. Structural (Shape) features extraction
5. Compare with the Local Feature DB

Feature Extraction

To extract the features, i.e., area, perimeter, major axis, minor axis, eccentricity, horizontal height, vertical height, rectangularity, circularity, compactness, elongation (Fig. 2).

- (a) **Area:** Defined as, the total number of pixels in the image. It is calculated as,

$$\text{Area} = \text{length}(\text{find}(\text{pixel})).$$

- (b) **Perimeter:** The distance around the boundary of the region.
The formula is,

Square	$4S$
Rectangle	$2(l + w)$
Triangle	$l + s_2 + s_3$

Polygon adds length of all sides

Where, s = square, l = length, W = width

Fig. 2 Cropped image of the AMI Ear Database [9]



- (c) **Major axis:** It is the longest diameter of the image. Major axis length is used to find the X and Y directions in clockwise directions horizontally.
The Major axis length is calculated as,
Major axis = a + b, a and b are distanced from each focus to any point
- (d) **Minor axis:** It is the shortest diameter of the image. Minor axis length is used to find the X and Y directions in anti clock wise.
Minor axis length is calculated as,

$$\text{Minor axis length} = \sqrt{(a + b)^2 - f^2}$$

Where f is the distance between focus points

- (e) **Eccentricity:** Eccentricity is the ratio of the length of the major axis to the minor axis [10].
Eccentricity is calculated as,

$$\text{Eccentricity} = 2 * \sqrt{((\text{Major Axis Length}/2)^2 - (\text{Minor Axis Length}/2)^2)}/\text{Major Axis Length}$$

- (f) **Horizontal height:** It is used to calculate the height of the given image horizontally.

$$\text{Horizontal height} = \text{width}$$

- (g) **Vertical height:** It is used to calculate the height of the given image vertically.

$$\text{Vertical height} = \text{height}$$

- (h) **Rectangularity:** Rectangularity represents how rectangular a shape is, i.e. how much it fills bounding rectangle [10]

$$\text{Rectangularity} = A_S/A_R,$$

where A_S is the area of shape

A_R is the minimum bounding rectangle

- (i) **Circularity:** Circularity is defined as to find roundness of the objects

$$\text{Circularity} = (\text{perimeter}^2)/(4*\pi*\text{area})$$

- (j) **Compactness:** Compactness is the property that generalizes the notion of a subset of Euclidean space being closed and bounded.

$$\text{Compactness} = (\text{Perimeter}^2)/\text{Area}$$

- (k) **Elongation:** The process of increasing in length in which the image is longer than what is being x-rayed.
Elongation is calculated as, $T = 2\pi w$,
where w = relative angular velocity

A Local feature, database created using the structure with the following features Area, Perimeter, Eccentricity, Elongation, Compactness, Horizontal height, major axis, minor axis, circularity, rectangularity, vertical height as shown in Table 1.

Performance evaluation defines the performance of the ear image features present in the database or not. Here, Genuine accept rate is defined as, if we check the true image features present in the database or not, if the result is true it is genuinely accept rate, and if the result is false it is known as false reject rate.

The above Fig. 3 explains that we take hundred images of the performance evaluation of the given database. After evaluating the performance with the database, it shows the result like in 100 images 75 images accepted by the database known as genuine accept rate and the remaining 25 images doesn't accepted and those are rejected with the database called as false reject rate shown in Fig. 3. We take genuine accept rate in x-axis and false reject rate in the y-axis.

Here, we check whether the false images feature is present in the database or not. The false images present in the database are known as false accept rate and the false images not present in the database are known as false reject rate.

Fig. 4 explains that we take twenty images of the performance evaluation from the database. After evaluating the performance with the database, it shows the result like in 20 images, 15 images accepted by the database known as false accept rate and the remaining 5 images doesn't accepted those are rejected by the database called as false reject rate. We take false accept rate in x-axis and false reject rate in the y-axis.

Table 1 Extracted structural or shape feature DB

S. No	Image name	Area	Perimeter	Major axis length	Minor axis length	Eccentricity	Horizontal height	Vertical height	Rectangularity	Circularity	Compactness	Elongation
1.	082backear	11138	412	148.7421241	96.64570331	0.760145052	180	256	0.241710069	0.824142709	15.24007901	1.421206796
2.	082downear	9935	402	150.9814792	85.52830179	0.824074061	180	256	0.215603299	0.77215663	16.26612984	0.573773774
3.	082frontear	10172	399	146.7727749	88.90096086	0.795689516	180	256	0.220746528	0.802509532	15.65090444	0.601355514
4.	082leftear	9492	2	147.9155151	82.61580095	0.829481825	180	256	0.205989583	0.298048814	15.30421408	0.574665009
5.	082rightear	11053	419	155.8791528	91.42945491	0.809919812	180	256	0.239865451	0.790754666	15.88356102	0.532226403
6.	082upear	10233	0	148.2676477	89.01289871	0.799735272	180	256	0.222070313	0.652903882	15.97288912	0.614602588
7.	082zoomear	14204	475	177.6867878	102.9531887	0.81503748	180	256	0.308246528	0.790702449	15.88460997	0.585727245
8.	083backear	9496	395	150.0106665	82.21152331	0.836453534	180	256	0.206076389	0.764427239	16.43060236	1.855167693
9.	083downear	8900	381	149.5115008	77.20888587	0.856342874	180	256	0.193142361	0.770069096	16.31022472	1.966101695
10.	083frontear	9298	394	153.978412	78.40983532	0.860636202	180	256	0.201779514	0.752292509	16.69563347	0.505170631
11.	083leftear	6688	353	146.2459593	59.03424065	0.914907269	180	256	0.145138889	0.674118884	18.63172847	0.40312516
12.	083rightear	103141	412	156.3267898	84.86221295	0.829828782	180	256	0.224414063	0.765169667	16.41466009	0.542724098
13.	083_up_ear	9745	0	146.023259	85.53211792	0.810496767	180	256	0.211480035	0.782358031	17.19215655	0.361531532
14.	083zoomear	12932	464	181.1048626	92.13428639	0.86092348	180	256	0.280642361	0.754430737	16.64831426	0.506225925
15.	084backear	9741	403	153.4195569	92.13428639	0.845556998	180	256	0.211393229	0.753326232	16.67272354	1.88315565
16.	084downear	9888	398	154.2217368	82.70996569	0.844023939	180	256	0.214583333	0.784028686	16.01982201	0.396883593
17.	084frontear	10345	408	155.3551347	85.52870512	0.834810954	180	256	0.224500868	0.780548347	16.09125181	0.543586896
18.	084leftear	9159	392	153.0769217	77.58486231	0.862042623	180	256	0.198763021	0.748627135	16.77737744	0.676704742
19.	084rightear	10411	412	158.3360766	84.27264542	0.846594104	180	256	0.22593316	0.770349232	16.30429354	0.525810324
20.	084upear	10132	406	155.7958275	83.46221069	0.844398419	180	256	0.219878472	0.77202747	16.26885116	0.527586207

Fig. 3 True images feature present in the database or not

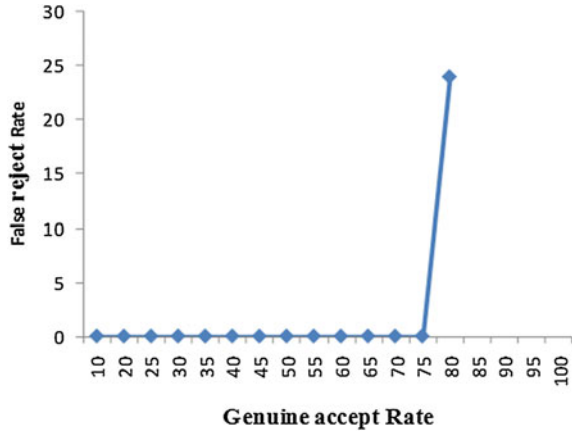
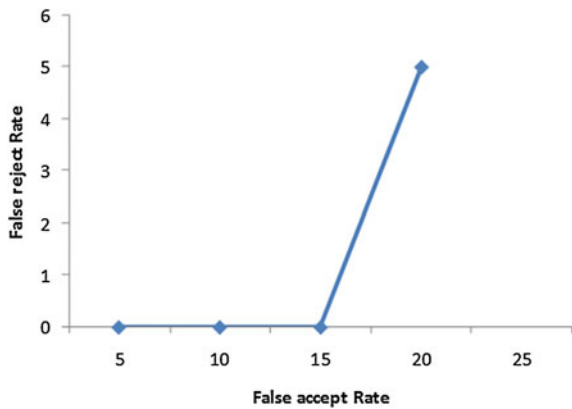


Fig. 4 False images feature present in the database or not



3 Conclusion

This article mainly focused on the shape to extract the features, which can be used to recognize the unauthorized persons in security systems. The AMI ear database is used for experimentation and the database consists of 200 images. The proposed approach uses structural or shape properties of the ear image to extract the features, which can be used to recognize the unauthorized persons in the security systems. The features area, perimeter, major axis length, minor axis length, eccentricity, horizontal height, vertical height, rectangularity, circularity, compactness and elongation for the given database are verified to identify whether the unauthorized person is present in the database or not. The performance is measured by using two techniques, false match rate and false non-match rate to generate graphs.

References

1. Devesh Narayan, Sipi Dubey 'A Survey Paper on Human Identification using Ear Biometrics', *International Journal of Innovative Science and Modern Engineering (IJISME)*, ISSN: 2319-6386, Volume-2 Issue-10, September 2014.
2. Anika Pflug, Christoph Busch, 'Ear Biometrics-A Survey of Detection, Feature Extraction and Recognition Methods', *IET Biometrics*, Volume 1, Number 2, pages 114–129, June 2012.
3. Jain, Anil K., Arun Ross, 'An Introduction to Biometric Recognition'. *IEEE TRANSACTIONS on circuits and systems for video technology* 14.1 (2004): 4–20. *IEEE Explore. Web.* 5 Dec. 2011.
4. Abaza, A., Ross, A., Herbert, C., Harrison, M. A. F., and Nixon, M. S. 2011. 'A Survey on Ear Biometrics'. *ACM Trans. Embedded. Comput. Syst.* 9, 4, 33 pages, Article 39, March 2010.
5. D. J. Hurley, B. Arbab-Zavar, and M. S. Nixon, 'The Ear as a Biometric', In A. Jain, P. Flynn, and A. Ross, *Handbook of Biometrics*, Chapter 7, Springer US, 131–150, 2007.
6. Michal Choras, 'Image Feature Extraction Methods for Ear Biometrics'. University of Technology & Life Sciences, Bydgoszcz. *Computer Information Systems and Industrial Management Applications*, 2007. *IEEE Explore. Web.* 28–30 June 2007.
7. Hanna-Kaisa Lammi, 'Ear Biometrics', Department of Information Technology, Lappeenranta University of Technology, Laboratory Information Processing, Lappeenranta, Finland, 2004.
8. Hurley, David J., Mark S. Nixon, and John N. Carter. "Automatic ear recognition by force field transformations." *Visual Biometrics (Ref. No. 2000/018)*, IEE Colloquium on. IET, 2000.
9. AMI ear database', Esther Gonzalez, Luis Alvarez and Luis Mazorra, PhD, Department of computer science and technology, under a Creative Commons Reconocimiento- No commercial- SinObraDerivada 3.0.
10. Yang, Mingqiang, Kidiyo Kpalma, and Joseph Ronsin. "A survey of shape feature extraction techniques." *Pattern recognition* (2008): 43–90.

Part II
Networking Theory and Distributed
Systems

DHAC Based Routing in Wireless Sensor Network with Asymmetric Links

Laxita Vyas, C.P. Gupta and Md Arquam

Abstract In Wireless Sensor Network (WSN), various routing strategies have been adopted to prolong the network lifetime. Clustering is the important technique in comparison to all the other existing routing techniques. Proposed algorithm adopts the hierarchical structure for cluster formation and selecting cluster head in distributed approach. In this paper, clustering of nodes is carried out by considering the asymmetric communication links between nodes. Energy consumption in proposed technique is reduced and hence lifetime of the network is improved. The Simulation is carried out in MATLAB7.9 and the obtained result verifies that proposed algorithm increased network as compared to the existing routing protocols.

Keywords Wireless sensor network • Clustering • Routing • Cluster head • Asymmetric links

1 Introduction

Wireless Sensor Network is tiny low power transceivers, which gathers information from surrounding or environments and sends the to the Base Station for processing. WSN consists of large number of sensors or nodes that are widely distributed in a network. WSN shows its significance in various application areas [1, 2]. To improve the performance of wireless sensor network several constraints of WSN is considered for research. The major constraints in WSN is energy utilization as due

L. Vyas (✉) · C.P. Gupta
Department of Computer Engineering, Rajasthan Technical University,
Kota, Rajasthan, India
e-mail: Laxita.mca@gmail.com

C.P. Gupta
e-mail: Gupta.cp2@rediffmail.com

M. Arquam
Department of Computer Engineering, NIT Delhi, Delhi, India
e-mail: mdarquam@gmail.com

to small size nodes, the energy depletion is fast in WSN which require the communication are done in organized way that reduce the energy depletion rate of the sensors.

The asymmetric communication links [3] between the sensors is also the one of the most important challenges in WSN. In asymmetric communication pattern, the cost of transferring packets from node A to node B would not be equivalent of transferring packets from node B to node A.

This research motivates a clustering algorithm that considers the asymmetric parameters in communication while clustering and selecting the cluster head in a distributed approach (by eliminating the need of global knowledge).

In this paper hierarchical structure is used for clustering of node while considering their asymmetric links. Fewer efforts have been taken place to consider asymmetric links while creating clusters of the node. Most of the existing routing protocol technique considers symmetric parameters such as euclidean distance between two nodes in which the cost of moving from node A to B is same as cost of moving from node B to A. In proposed algorithm communication cost will be a function of Euclidean distance between nodes and their respective energy index [4], which transforms the symmetric, cost into asymmetric, as energy level of each node are different in network.

2 Related Work

Intensive research in wireless sensor has been carried out to organise communication that leads to development of many clustering algorithm. Each algorithm has some pros and cons. In this course LEACH [5] is the first protocol, in which the cluster head randomly selected among all the nodes in the network to balance the energy consumption. Initially some probability value is assigned to every node and the node with maximum probability would be selected as the cluster head for that round. LEACH comprises two stages for cluster head selection, i.e., setup phase and steady phase. In first phase, cluster organization takes place. In steady phase the sensor node transmits data to cluster head. The issue with LEACH protocol is the random selection of node to become cluster head that causes the quickly energy consumption of the node.

LEACH-C [6] is an enhanced version of LEACH protocol. In LEACH-C, cluster head selection done by considering the location information rather than randomly selecting the cluster head. The LEACH-C requires the global knowledge and the location of sink for selecting the cluster head, which make it entirely dependent on the Base Station location.

PEGASIS [7] protocol is an improvement over the existing routing protocols in which each node trans receive data to the close neighbors and rotates the role of being leader or cluster head for transmitting data to the base station. The key idea of

PEGASIS is to form a chain of node in which on cluster head gathering of data would be done which send the data to the Base Station. If there are N number of nodes in the network then the node $i \bmod N$ would be selected as the cluster head for the round i .

HEED [8] protocol considers the energy parameters for selecting the cluster head among all the nodes in the network. Node with more residual energy would have more of selecting the cluster head. HEED selects the leader or cluster head by calculating the probability value of each node by:

$$CH_{prob} = C_{prob} * E_{residual} / E_{initial}$$

Node i would be selected as the cluster head if its probability value, i.e., CH_{prob} would be either less than or equal to 1.

EECH [9] is a distributed and randomized algorithm in which every sensor node advertises itself as a CH with a probability p to the adjacent nodes within its communication range. These CHs are called as volunteer CHs. All nodes receive this advertisement either by direct communication or by forwarding within k hops range of a CH. Any node that receives that advertisement and is not itself a CH becomes the member of the closest cluster. Forced CHs are those node that are neither CHs nor belong to any cluster directly communicates with sink. The limitation of EECH is that it requires the global knowledge about the distances between the cluster-heads and the base station.

RRCH [10] have fixed cluster and apply round robin method to select the node to be cluster head within cluster. But without periodic re-clustering, RRCH cannot handle cluster with bad quality that results in either too small or big size cluster.

V-LEACH [11] is an improvement over LEACH protocol in which vice-CH selected when selected CH dies while communication. V-LEACH does not perform well when the CH is not presented in the network for routing packets from sensor nodes to sink.

CH Lung proposes a three protocol namely HAC, DHAC and H-DHAC. HAC (Hierarchical Agglomerative Clustering) [12] is a theoretically and mathematically simple clustering model. It provides efficient clustering strategy by using the clustering approach bottom-up rather than top down. In HAC the bottom-up approach is follows in which nodes is associated and clusters are formed before CHs selection. In this way, the bottom-up strategy can be a better way to deploy self-organization, scalability and flexibility.

DHAC [12] provides more innovation plan to make cluster and maintain cluster as compared to HAC. DHAC overcomes the HAC limitations of having global knowledge as it require only 1-hop neighbour information. DHAC is the distributed algorithm, which assumes all nodes in the network have same capabilities, processing, communication and initial energy. DHAC consider only the symmetric parameters, i.e., distance measures form one node to another for creating the cluster and selecting the Cluster Head. H-DHAC [13] works on both qualitative data (such

as RSSI) and quantitative data (such as location). H-DHAC designed to overcome the limitation of GPS [14–17]. H-DHAC reduces the randomness while clustering.

All the above described routing protocols consider the symmetric parameters i.e. either it take Euclidean distance, RSSI or residual energy for creating cluster. Little work has been carried in taking more than one parameters simultaneously for cluster formation. Various factors such as residual energy, Euclidean distance and Received Signal Strength are involved in creating a cluster which makes proximity between two nodes asymmetric. Here in this research we propose a protocol that considers the asymmetric proximity between the nodes in which we combine two parameters such as residual energy and Euclidean distance into single parameters that identify the energy distance proximity between nodes.

3 Proposed Work

In this paper an algorithm is proposed for cluster formation and selecting the cluster head among the generated cluster in distributed environments without requiring the global knowledge and by considering the asymmetric links between nodes. The following assumptions are made for an algorithm

- The sink is static.
- Each node has the information of their neighbouring node that is at its one-hop distances.
- Each node has the equal initial energy.
- Processing capabilities of individual nodes vary as the energy level of each node varies.

Optimal Number of Cluster

According to the energy model given by [5]:

The energy consumption during transmission of k -bit message over a distance d using first order radio model is:

$$E_{Tx} = E_{Tx-elec}(m) + E_{Tx-amp}(m, d) = \left\{ \begin{array}{l} mE_{elec} + m\epsilon_{fs} d^2, (d < d_0) \\ mE_{elec} + m\epsilon_{amp} d^4, (d \geq d_0) \end{array} \right\}$$

- m is the number of bit transmitting on the distance d .
- E_{elec} is the energy dissipation per bit of transmitter circuitry.
- ϵ is the transmit amplifier dissipation per bit.

The receiving cost is:

$$\begin{aligned}
 E_{Rx} &= E_{Rx-elec}(m) - mE_{elec} \\
 E_{CH} &= L \cdot E_{elec} \frac{N}{k} + LE_{DA} \frac{N}{k} + L\epsilon_{fs} d_{toBS}^2 \\
 E_{NonCH} &= L \cdot E_{elec} + L\epsilon_{fs} d_{toCH}^2 \\
 E_{Total} &= k \cdot E_{CH} + nE_{Non-CH} \\
 E_{Total} &= nLE_{elec} + nLE_{DA} + k \cdot L\epsilon_{fs} d_{toBS}^2 + nLE_{elec} + nL\epsilon_{fs} d_{toCH}^2 \\
 E[d_{toCH}^2] &= \frac{M^2}{2\pi k} \\
 nE_{elec} + nE_{DA} + k\epsilon_{fs} d_{toBS}^2 + nE_{elec} + n\epsilon_{fs} \frac{M^2}{2\pi k} &= 0
 \end{aligned}$$

Differentiating above eq. w.r.t k -

$$\begin{aligned}
 \epsilon_{fs} d_{toBS}^2 - n\epsilon_{fs} \frac{M^2}{2\pi k^2} &= 0 \\
 k &= \frac{\sqrt{n}M}{\sqrt{2\pi} d_{toBS}^2}
 \end{aligned}$$

According to [9]

$$d_{toBS} = 0.3826 * a$$

Substituting value of d_{toBS} in above eq.-

$$k = \frac{\sqrt{n}}{0.9587}$$

Since there are on an average $n * p$ cluster heads so value of k equal to $n * p$ where p is the probability of becoming CH-

$$\begin{aligned}
 k &= n * p \\
 p &= \frac{1}{\sqrt{n} * 0.9587}
 \end{aligned}$$

The proposed algorithm comprised of the following steps for communication in WSN

- I. Initialization: Node presented in a network identified their neighbouring nodes that are suited at their one hop distances.
- II. Each node starts computing its proximity value from another node that respond to their Hello message by below formulae,

$$C_{A,B} = (1/n_A * n_B)^* \sum_{x \in A, y \in B} proximity(x, y)$$

III. Cluster Head Selection

```

for r = 1 to r_max
    for i = 1:N
        if r == 1
            t(i) = random number generation from 0 to 1
        else
            t(i) = p * (Er/Ei)
        end of if else statement
    end of for loop
[C, ind] = max(t)
Node i with index ind would be selected as cluster Head
End of for loop

```

Cluster Maintenance

The cluster Head selection has below illustrated 3 phases:-

1. The Cluster Head election: It contains of three phases mention following:
 - (a) Cluster Head Advertisement: All nodes in the cluster transmit advertisement to all the sensor nodes with their probability of becoming cluster head.
 - (b) Joining Request to Cluster: Sensor Nodes, on hearing advertisement request, added to the closest cluster head according to the maximum probability of becoming CH. Along with the connection message, node position and residual energy of the sensor nodes also transmitted to their cluster head.
 - (c) Cluster Head Acknowledgement: An acknowledgement has been sent by cluster head to the sensor node confirming that the sensor node is connected to its cluster on receiving a cluster join request.
2. Steady State Phase: In this phase, data collected by the sensor nodes from their surroundings environment and then transmitted to their respective cluster heads. Cluster Head create the TDMA schedule to avoid collision among nodes.
3. Aggregation and Forwarding of Data: Local Cluster heads, after receiving data from all their member nodes, collect the data and send the collected data to the Base Station.

4 Simulation and Result Analysis

To analysis the performance of proposed algorithm MATLAB 7.9 is used for simulation. For simulation model, network of 100 stationary sensors and one sink is used. The nodes are assumed to be randomly arranged within the field of square area of $(a * a)$ (Table 1).

Validating Optimal Number of Cluster

According to [14] silhouette coefficient, cophenetic and spearmen correlation coefficient are used to validate the number of optimal cluster.

From the Tables 2 and 3 of silhouette coefficient, cophenetic correlation and spearmen correlation coefficient we conclude that maximum coefficient value for all model are obtained for average linkage therefore in this paper it is adopted that average linkage method for updating the resemblance matrix (Figs. 1, 2 and 3).

Table 1 Simulation parameters

Parameter	Value
Node number	100
Sensing field range	(0, 0)–(100, 100)
Channel bandwidth	1 Mbps
Threshold distance(d_0)	87.7 m
E_{elec}	50 nJ/bit
ϵ_{fs}	10 pJ/bit/m ²
ϵ_{amp}	0.0013 pJ/bit/m ⁴
$E_{initial}$	0.5 J
Data packet size	2000 bytes
E_{DA}	5 nJ/bit

Table 2 Silhouette Coefficient Value at Different Cluster at Different Linkage Method

No of cluster	Silhouette coefficient			
	Single	Complete	Average	Ward
5	-0.20	0.30	0.32	0.38
10	-0.10	0.28	0.40	0.35
20	0.15	0.32	0.36	0.35

Table 3 Cophenetic And Spearmen Correlation Coefficient Value For Different Linkage Method

Linkage Method	Coefficient	
	Cophenetic	Spearmen
Single	0.55	0.53
Complete	0.67	0.63
Average	0.69	0.67
Ward	0.65	0.65

Fig. 1 Graph showing coefficient value for different number of cluster for various linkage methods

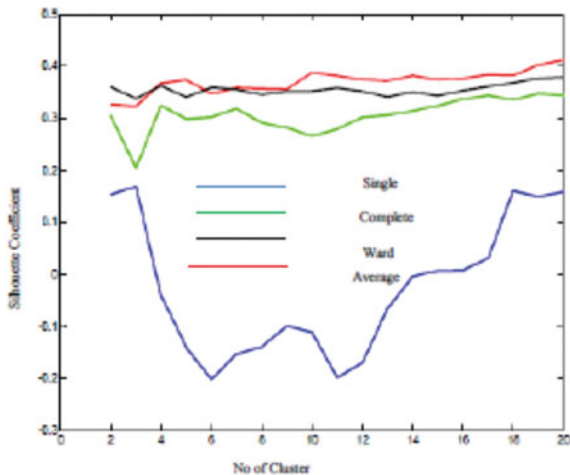
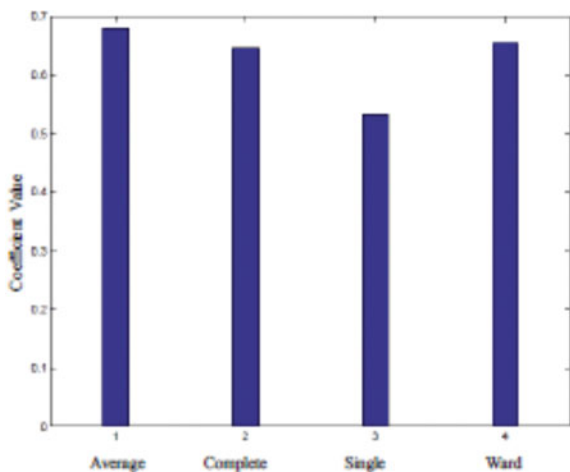


Fig. 2 Graph showing different cophenetic coefficient value at for different linkage method



A comparison of proposed work performance to that of LEACH, LEACH-C and PEGASIS shown in Figs. 4 and 5. The proposed algorithm provides improvements of up to 60 % in lifetime. LEACH has the worst performance. It is due to those CHs in LEACH send data to base station directly which costs large amount of energy and thus the energy consumption is not uniform. LEACH-C give a better performance than LEACH because it modified LEACH by using global information and centralized clustering algorithm for cluster formation in order to realize uniform distribution of cluster heads throughout the network. PEGASIS give better performance than both the LEACH and LEACH-C as each node communicates only with a close neighbour and takes turns transmitting to the base station, thus reducing the amount of energy spent per round. The performance of our proposed algorithm is the best. It divides the network into k clusters, which can guarantee the

Fig. 3 Graph showing different sparmen coefficient value at for different linkage method

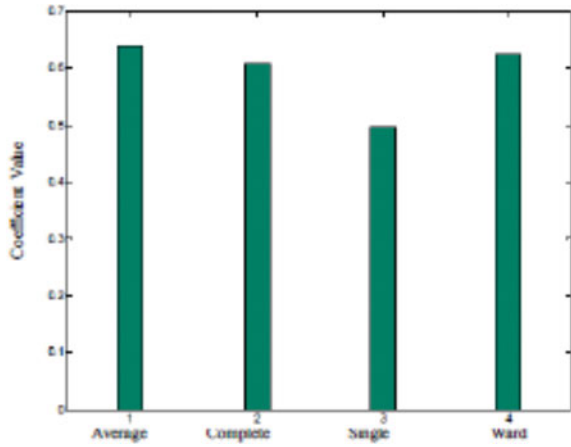


Fig. 4 Energy utilization for different routing protocols

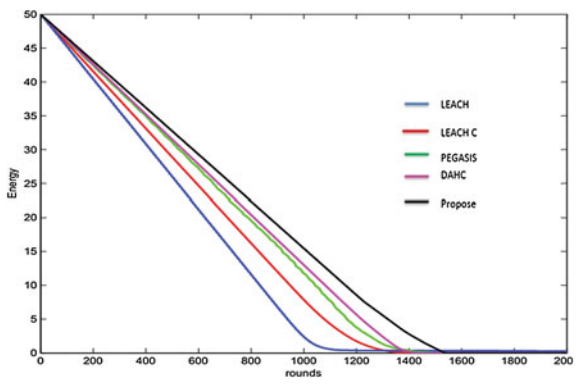
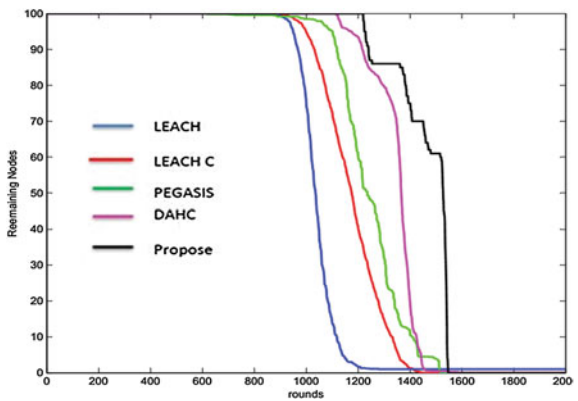


Fig. 5 Network lifetime of different routing protocols



CHs uniformly distribute in the whole network. CHs conduct data aggregation in cluster and relay cluster heads are responsible for forwarding data between clusters through a routing tree.

5 Conclusion

In this research we undertake more than one factor that affect the clustering of a node where as in most of existing technique only one factor would be used. Proposed protocol is able to handle asymmetrical proximity value between sensor nodes, which is an important characteristic of a dynamic sensor networks. It selects the cluster head by locally available information by eliminating the requisite of sink's global knowledge. Proposed protocol is more effective since it is designed to exploit the application of sensor networks; producing high level information about the environment the nodes are monitoring in an energy-efficient way.

References

1. Akyildiz I.F, Wu S., Sankarasubramaniam Y, Cayirci E., "Wireless Network Survey", *Computer Networks Journal (Elsevier)*, vol 40, no. 8, 2002, pp 393–422.
2. Gupta C.P., Kumar A., "Wireless Sensor Networks: Review", *International Journal of Sensors, Wireless Communication and Control*, vol., no., 2013, pp 25–36.
3. Takumi S., & Miyamoto, S., "Agglomerative Clustering Using Asymmetric Similarities" in *Proc. Of 2006 3rd IEEE international Conference on information Technology New Generation*, 2006, pp 34–39.
4. Matthew H, et al. "Optimizing physical-layer parameters for wireless sensor networks." *ACM Transactions on Sensor Networks (TOSN)* 7.4 (2011): 28.
5. Heinzelman, W.B., Chandrakasan, A., Balakrishnan, H., "Energy efficient communication protocol for wireless microsensor networks", in *Proc of 33rd Hawaii International Conference on System Sciences (HICSS)*, Wailea Maui, Hawaii, USA, January 2000.
6. Heinzelman, W.B., Chandrakasan, A., Balakrishnan, H., "An application specific protocol architecture for wireless microsensor networks", *IEEE Transactions on Wireless Communications*, vol.1, no.4, 2002, pp 660–670.
7. Lindsey, S.; Raghavendra, C.; Krishna Sivalingam, "Data gathering in sensor networks using the energy*delay metric," *Parallel and Distributed Processing Symposium.*, in *Proc of 15th International*, vol. 23, no. 27, April 2000, pp. 001– 2008, doi:10.1109/IPDPS.2001.925196.
8. Younis, O.; Fahmy, Sonia, "Distributed clustering in ad-hoc sensor networks: a hybrid, energy-efficient approach," *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, vol.1, no.7, March 2004, pp., 640, doi:10.1109/INFCOM.2004.1354534.
9. Bandyopadhyay, S.; Coyle, E.J., "An energy efficient hierarchical clustering algorithm for wireless sensor networks," *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, vol.3, no., 2003, pp 1713–1723, doi:10.1109/INFCOM.2003.1209194.
10. Nam D.H. & Min H.K., "An efficient ad-hoc routing using a hybrid clustering method in a wireless sensor network," in *Proc. of 3rd IEEE Int'l. Conf. WiMOB.*, Oct. 2007, pp. 60.

11. Young M.B., Al-zou'bi A., Khamayesh Y& Mardini W.“Improvement on Leach Protocol of Wireless Sensor Network (VLEACH),”*International Journal of Digital Contents Technology and Its Applications*, Vol. 3, No. 2, 2009, pp. 132–136.
12. Lung,C.H., &, Zhou,C.,”Using Hierarchical Agglomerative Clustering in Wireless Sensor Network: An Energy Efficient and Flexible Approach”,*Adhoc Networks*, vol.8,no.3,2010, pp 328–344, doi:[10.1109/GLOCOM.2008.ECP.55](https://doi.org/10.1109/GLOCOM.2008.ECP.55).
13. Lung C.H, Zhu J.,Srivastava V.,”H-DHAC: A Hybrid Clustering Protocol for Wireless Sensor Networks”, *in Proc. Of IWCMC*,2013,pp.183–188,doi:[10.1109/IWCMC.2013.6583556](https://doi.org/10.1109/IWCMC.2013.6583556).
14. Gupta, C P, Kumar A,”Optimal Number of Cluster in Wireless Sensor Network with Mobile Sink ”, *International Journal of Scientific and Engineering Research*, vol 4,no. 8, 2013,pp 1706–1710.
15. Privat L. (2007, March 29th).TI announced new GPS chip,cost Under \$5. Available:http://www.gpsbusinessnews.com/Tiannounced_new-GPSchipcostunder-5_a102.html
16. Gakstatter E. (2010, June 1st). The Dawn of a New ERA in GPS Accuracy. Available:<http://www.gpsworld.com/gis/gssweekly/thedawn-a-new-era-gps-accuracy-10016>.
17. Bhatti,U.I. & Ochieng,W.Y., “Failure Modes and Models for Integrated GPS/INS Systems,”*The Journal of Navigation*,vol.60, pp. 327–348, 2007.

Automatization of AAOCC to Find Trust Score of Websites

Manish Kumar Verma, Sarowar Kumar, Kumar Abhishek
and M.P. Singh

Abstract Today World Wide Web has emerged as a second world. Everything we can think of is now available on this digital world. In the real world there exists many kinds of people. Some are good, some are bad, and some are trustworthy while the rest are liars. Similarly, in the digital world, there exists many websites out of which some are good while the rest are useless. So the big question is how can one know which content on www is trustworthy as every information can be changed with a few keystrokes. A lot of algorithms have been developed to identify the trust rank of a website, but none of them are up to the mark. So through this paper, we propose a simple mechanism by which we can test a website and automatically calculate the trust score of the website on the basis of not a single parameter but a cumulative combination of five parameters. Various test results have been included in this paper which prove that the method proposed in this paper is much better than the conventional methods which test websites manually.

Keywords AAOCC · PHP · XAMPP · MySQL · PageRank · TrustRank

M.K. Verma (✉) · S. Kumar · K. Abhishek · M.P. Singh
Department of Computer Science, National Institute of Technology, Patna, India
e-mail: manish123977@nitp.ac.in

S. Kumar
e-mail: sarowarkumar@gmail.com

K. Abhishek
e-mail: kumar.abhishek@nitp.ac.in

M.P. Singh
e-mail: mps@nitp.ac.in

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_18

183

1 Introduction

On the World Wide Web there exist a variety of websites. Out of these a lot of websites can be called web spam/spam websites. The purpose of these spam websites is to spread viruses, Trojan, worm etc. Some of them are used to display useless ads and information and increase the user visit count so that they can earn more revenue.

The spam websites contain web pages with a lot of hidden common keywords that a normal user searches for. Normally a search engine indexes these websites at the top because they contain more percentage of matching keywords. Now when the user clicks on these sites he/she is redirected to chain of bogus sites which include pornographic materials, illegal contents, useless ads etc.

These websites are a threat to user's privacy and security and waste a lot of user's time. So different algorithms like Trust Rank Algorithm, Page Rank Algorithm etc. were developed to calculate trust rank of websites and to filter out good websites from spam pages. These algorithms were later followed by different search engines to optimize their search results and display pages with higher trust rank at the top.

But a major disadvantage of these algorithms is that they take a few thousands of websites as seed and check their trust rank manually. The algorithm gives higher trust rank to those websites who hyperlink to these good seed pages and the chain continues. So if a spam page refers to these seed pages, then they get higher trust rank. Hence a lot of false results occur.

Through this paper, we developed a technique to automate the calculation of trust score. The method treats every website individually and calculate the trust rank on the basis of five parameters. So even if a spam webpage refers to a good website it does not mean they will get higher rank. Although this paper is inspired by previous works yet it provides a simple and efficient technique to calculate trust score.

2 Related Works

There have been a lot of research in this field, but all proposed theoretical approaches [1–3]. Practical implementation of trust rank had a lot of flaws. Through this paper, we provide a better practical method to calculate trust score automatically [4–8].

Gyongyi et al. [9] proposed a technique to semi automatically separate good web pages from spam web pages using a set of seed pages which have been evaluated manually.

Chandratre and Kulkarni [10] proposed a web mining framework for e-commerce websites and used trust rank algorithm and page rank algorithm for web structure mining analysis.

Bianchini et al. [11] performed a detail analysis of page rank algorithm and found out the fundamental properties that affect the stability and complexity of computation of trust rank.

Haveliwala [12] improved Page Rank by introducing a set of topic biased Page Vectors which helped to obtain a more accurate result in case of search keywords which are related to a particular context or topic.

3 Proposed Methodology

The traditional algorithms use seed pages and their connectivity to good seed pages as a prime parameter. But the technique proposed in this paper follows a new set of parameters known as AAOCC. Here AAOCC stands for

A-Accuracy
 A-Authority
 O-Objectivity
 C-Currency
 C-Coverage

To calculate the entropy of keyword i.e. occurrences of a particular term in a website, we use the formula

W: Total word count in webpage

K: Total match of keyword in webpage

$G = K/W$

$H = -1 * \text{Log}(G)$

$$\text{Hence Entropy of keyword} = G * H \quad (1)$$

Using these parameters we calculate the total trust score. Each parameter contributes evenly in the total score. On the basis of total score obtained we decide the final ranking or indexing of web pages.

Accuracy

- Listing the author and institution that uploaded the webpage and find methods to contact him or her.
- It also includes finding the purpose of writing the web content and checking whether the author is qualified enough to write such content.

Authority

- Checking the domain of webpage and finding whether the webmaster is different from the author or not.

Objectivity

- We check whether the website meets its goal/objective or not.
- We also find out extent of details up to which the topic is explained in the web page.

- If the website is used as cover page to display ads then it might contain biased information about a certain group of people/company who are paying for the ads. Such websites are given lower trust score in case of objectivity.

Currency

- It mainly converse the current status of web page.
- It includes information about recent update/creation date of web page.
- We also check the total number of active and dead links on a website. Also, we find out how whether the site redirects to any misleading links or not.

Coverage

- We find out whether the site has enough text and multimedia content to properly explain the topic.
- Also, we check if the site requires certain special software to view the information.
- We determine the user-friendly elements present in the web page. User-friendly sites get higher coverage score.

4 System Implementation

The algorithm used in the paper is

- Step 1 Search string taken as input using html form
- Step 2 stopwords were removed from the search string
- Step 3 the keywords were searched in the database and resultset obtained
- Step 4 loop start if resultset is not null
- Step 5 the URL of websites was fetched
- Step 6 total word count in body of websites calculated
- Step 7 keyword was broken into individual terms
- Step 8 loop start
- Step 9 total match count for each term in website body calculated
- Step 10 using the formula entropy was calculated
- Step 11 loop end
- Step 12 the URL and total entropy was put in an array
- Step 13 loop end
- Step 14 the array was sorted as per total entropy using bubble sort
- Step 15 Loop start
- Step 16 The URL of website to evaluate was obtained from array.
- Step 17 CheckAuthority function was called with URL as parameter.
- Step 18 CheckCurrency function was called with URL as parameter.
- Step 19 Total trust score was calculated using the return values of CheckAuthority function and CheckCurrency function

- Step 20 The title was displayed as hyperlink for the website.
- Step 21 URL along with entropy and total trust score was displayed.
- Step 22 Loop end

The steps for CheckAuthority function

Input Parameters: URL

- Step 23 begin
- Step 24 Domain “.gov” was searched in URL. If match found score 100 was returned.
- Step 25 Domain “.edu” was searched in url.If match found score 80 was returned.
- Step 26 Domain “.org” was searched in url.If match found score 60 was returned.
- Step 27 Domain “.ac.” was searched in url.If match found score 40 was returned.
- Step 28 Domain “.nic.” was searched in url.If match found score 30 was returned.
- Step 29 Domain “.com” was searched in url.If match found score 20 was returned.
- Step 30 If no match found score 0 was returned.
- Step 31 end

The steps for CheckCurrency function

Input Parameters: URL

- Step 32: Using substr function the baseurl was extracted from the given URL.
- Step 33 The httpheader for the URL was obtained using get_header function
- Step 34 initialized total count = 1 and successcount = 0
- Step 35 If httpstatus is 200 or 302
- Step 36 URL printed along with last-modified information if present. successcount incremented by 1.
- Step 37 DOM Document Object was created and html data was loaded into it.
- Step 38 search for <a> tags and store it in array
- Step 39 Loop start
- Step 40 Get attribute “href” value for each <a> tag
- Step 41 If value not absolute
- Step 42 Change its value to absolute
- Step 43 If the URL follows https then exception thrown and loop control sent to top
- Step 44 else the httpheader for the URL was obtained using get_header function
- Step 45 totalcount incremented by 1.
- Step 46 If httpstatus is 200 or 302
- Step 47 URL printed along with last-modified information if present successcount incremented by 1.
- Step 48 Loop end
- Step 49 Percentage of success was calculated and returned using the variables Totalcount and successcount.
- Step 50 Success percentage was summed with authority score and total trust score was printed.

5 Test Cases and Results

In the test results the entropy value is calculated using the entropy formula (1).

The total score has been calculated as summation of Authority score and Currency score using the CheckAuthority function and CheckCurrency function.

The results use entropy and AAOCC to decide whether a website is trusted or not. The conventional methods use only PageRank to decide trust status of web-pages (Figs. 1, 2, 3 and 4).

Note:- OLD RANKING is as per <http://www.seocompany.ca/pagerank/page-rank-10-sites.php> on 14th May 2015.

**AAOCC Score has been calculated as summation of Authority score and Currency score using the CheckAuthority function and CheckCurrency function.

The top 6 urls in above table have almost same ranking in AAOCC as compared to PageRank. The AAOCC method adopted in this paper is highly automated and dynamic in nature as the scoring functions run in real-time whereas PageRank is calculated manually.

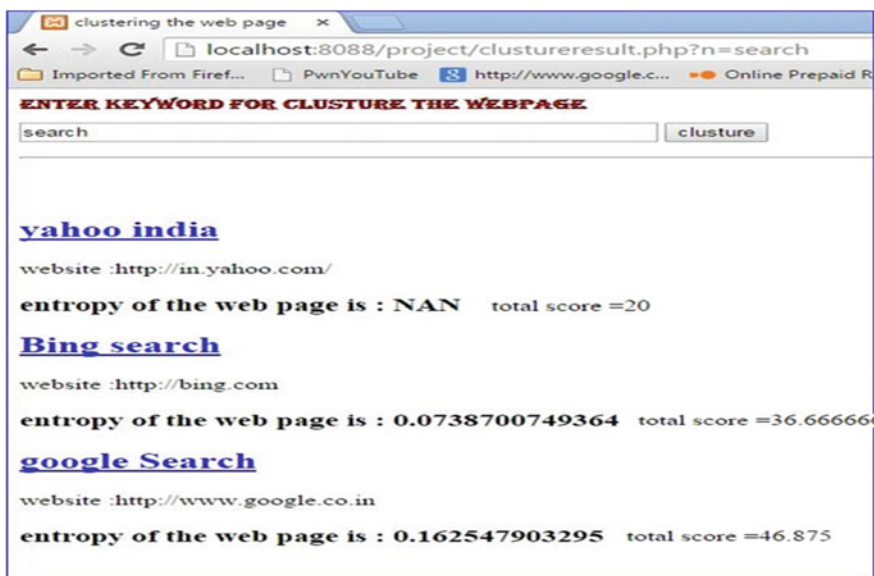


Fig. 1 Calculation of entropy and trust score for keyword "search"

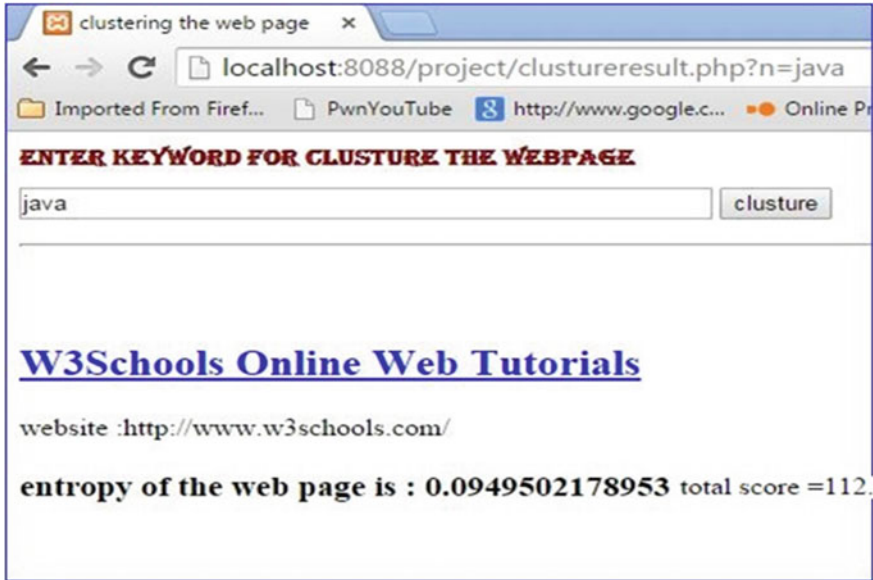


Fig. 2 Calculation of entropy and trust score for keyword “java”

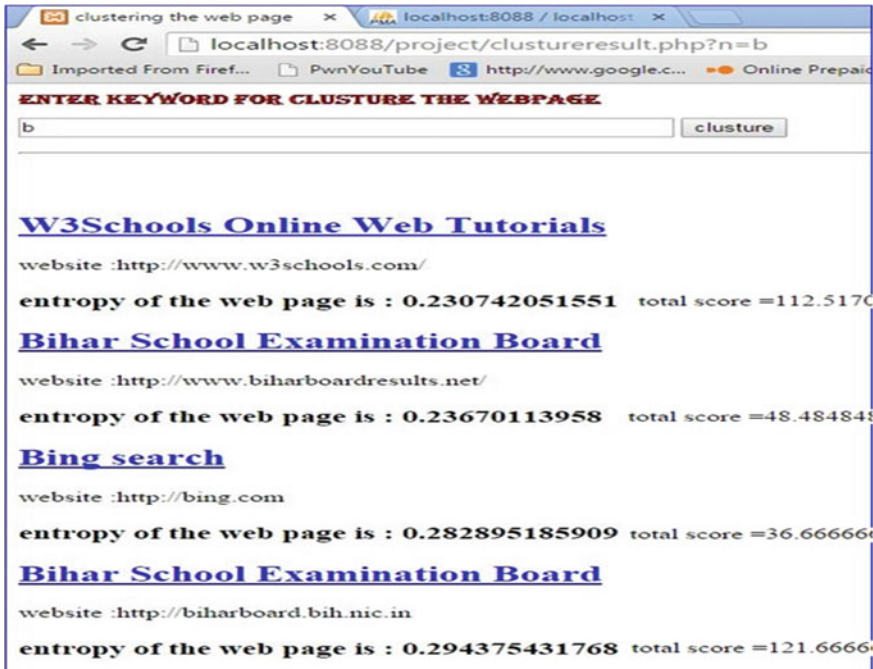


Fig. 3 Calculation of entropy and trust score for keyword “b”

RANK	OLD RANKING	RANKING after AAOCC	AAOCC score
1	http://www.adobe.com/	http://www.energy.gov/	196
2	http://www.apple.com/	http://www.keio.ac.jp/	140
3	http://www.energy.gov/	http://www.apple.com/	117
4	http://www.ercim.org/	http://www.adobe.com/	109
5	http://www.firstgov.gov/	http://www.firstgov.gov/	100
6	http://froogle.google.com/	http://www.google.com/	93
7	http://www.google.com/	http://www.ercim.org/	60
8	http://www.google-store.com/	http://www.google-store.com/	20
9	http://www.intel.com/	http://www.intel.com/	20
10	http://www.keio.ac.jp/	http://froogle.google.com/	20

Fig. 4 Comparison of PageRank and AAOCC rankings

6 Conclusion

Since the number of spam and fake websites are increasing day by day. We need website evaluation methods like AAOCC using which we can examine each website individually and determine whether they can be trusted or not. The total trust score calculated using AAOCC parameters can be used to optimize the search results of search engines and index trusted websites at the top of search results. The ranking and score will also help the users to get good and authentic information from World Wide Web. It will save a lot of time as the user will get what he/she wants at the top of search result with much higher accuracy and coverage to searched keyword.

7 Future Work

The parameter accuracy is a very broad topic and details vary from point of view of different persons. Hence, further research is needed to generalize accuracy and develop algorithm to calculate trust score.

The parameter coverage includes information about combination of text and multimedia that are enough to explain the topic. The combination may vary person to person and with different topics. So we need to have a parameterized algorithm which can sufficiently estimate the final combination of text and multimedia.

The method is not efficient enough in case of websites which follow https protocol. Effort is being made to extract information from https header.

References

1. Jim Kapoun “Teaching undergrads WEB evaluation: A guide for library instruction” C&RL News (July/August 1998)
2. Hiteshwar Kumar Azad, Kumar Abhishek “Entropy measurement and Algorithm for Semantic-Synaptic Web Mining” INTERNATIONAL CONFERENCE ON DATA MINING AND INTELLIGENT COMPUTING (ICDMIC-2014) IGDTUW Delhi, (September 2014)
3. Hiteshwar Kumar Azad, Rahul Raj, Rahul Kumar, Harshit Ranjan, Kumar Abhishek, and M. P. Singh. “Removal of Noisy Information in Web Pages”. In Proceedings of the 2014 International Conference on Information and Communication Technology for Competitive Strategies (ICTCS 14). ACM, New York, NY, USA, Article 88, 5 pages
4. PHP Manual for DOM Document <http://php.net/manual/en/class.domdocument.php>
5. PHP Header function <http://php.net/manual/en/function.get-headers.php>
6. PHP Manual for string functions <https://php.net/ref.strings>
7. MySQL database and documentation <https://www.mysql.com/>
8. XAMPP Package <https://www.apachefriends.org/index.html>
9. Zoltan Gyongyi, Hector Garcia-Molina and Jan Pedersen “Combating Web Spam with Trust Rank” Proceedings of the 30th VLDB Conference, Toronto, Canada, 2004
10. Mrs.Pallavi Chandratre and Prof. Umesh Kulkarni “ Implementation of Trust Rank Algorithm on Web Pages “ International Journal of Science Technology & Management, Volume No.04, Special Issue No.01,February 2015
11. Monica Bianchini, Marco Gori and Franco Scarselli “Inside PageRank “ACM Transactions on Internet Technology, Vol. 5, No. 1, February 2005
12. Taher H. Haveliwala “Topic Sensitive PageRank “WWW 2002, May 7–11, 2002, Honolulu, Hawaii, USA

A Multi-level Weight Based Routing Algorithm for Prolonging Network Lifetime in Cluster Based Sensor Networks

Priyanka Pukhrambam, Sanghita Bhattacharjee
and Himanish Shekhar Das

Abstract Energy efficiency in routing is an important design issue in wireless sensor networks where nodes are battery operated which may or may not be rechargeable in many cases. Due to relaying high volume data packets, nodes closer to the base station consume more energy than other nodes in the network. In this paper, we propose a Multi-level Weight based Routing Algorithm (abbreviated shortly as MWRA) for cluster based sensor networks with an aim to minimize inter-cluster energy consumption and to balance the energy dissipation at nodes. In MWRA, clusterheads are assigned levels based on the distance between the clusterheads and the base station and then a 2-connected backbone network is formed to find energy efficient next hop nodes. Moreover, a weight function based routing technique is used based on which clusterhead chooses its relay node for forwarding the packets. In MWRA, the network reconfiguration is performed periodically to achieve balanced energy consumption among clusterheads. The experimental results show that our algorithm significantly improves the network lifetime and the energy efficiency of the network than the existing clustering algorithm.

Keywords Clustering • Level • Backbone network • Network lifetime • Backbone range

P. Pukhrambam (✉) • S. Bhattacharjee
Department of Computer Science and Engineering,
National Institute of Technology, Durgapur, India
e-mail: yuprina612@gmail.com

S. Bhattacharjee
e-mail: sanghita.b@gmail.com

H.S. Das
Department of Computer Science and Engineering,
National Institute of Technology, Silchar, India
e-mail: hsdas0815044@gmail.com

1 Introduction

A wireless sensor network (WSN) consists of a large number of tiny, resource constraint and unattended sensor nodes capable of monitoring ambient conditions such as weather, pressure, noise levels etc. Such nodes are typically empowered with limited energy supply. When a node runs out its energy, it becomes non-functional. Therefore, energy efficiency is an essential design criterion to keep the node networked. Energy efficient routing protocols not only ensure high message delivery and low energy consumption for data forwarding, but also balance the entire network energy consumption, and thereby improve the network connectivity significantly. Node clustering is one of the popular routing techniques for minimizing the energy consumption of sensor nodes. In such approach, clusters are organized by grouping the nodes, which increases scalability, robustness and reduces network traffic. Each cluster is supervised by a leader known as clusterhead (CH) and other ordinary nodes referred as cluster members. Moreover, CHs form a virtual backbone for forwarding the data packets to the base station when clusterhead and base station are not within the communication range of each other. Nevertheless, this technique causes imbalanced energy dissipation among CHs and as a result, CH nodes deplete their energy quickly and the network becomes partitioned. Although re-clustering improves energy balancing, repeated re-clustering incurs network overhead and high delay. Several clustering schemes have been developed in past couple of years to minimize the energy consumption at nodes. However, proper attention has not been given to cluster based multihop routing which plays a vital role in reliable data transmission.

In this paper, we present a Multi-level Weight based Routing Algorithm (shortly abbreviated as MWRA) for cluster based sensor networks to prolong the network lifetime. The proposed method is classified into three phases: *level assignment*, *backbone network formation* and *routing*. In level assignment phase, clusterhead nodes are assigned various levels depending on the distance between the clusterhead and the base station. After level assignment, a 2-connected backbone network is formed. Clusterheads utilize their topological information for constructing such network. We develop a weight function based routing technique using both distance and residual energy by which each clusterhead can choose its next hop node to route the data packet. Furthermore, in MWRA, set of forwarding nodes are updated periodically to balance the energy of over utilized nodes and to minimize the energy consumption of the entire network. Simulation experiments demonstrate that the proposed MWRA improves the lifetime of network significantly than the existing clustering algorithm EMRA [1].

The paper is organized as follows. Section 2 states the related work. In Sect. 3, system model is briefly described. The proposed Multi-level Weight based Routing Algorithm (MWRA) is detailed in Sect. 4. Performance of the proposed algorithm is evaluated and analyzed in Sect. 5. Finally, in Sect. 6, we conclude the paper.

2 Related Work

Many clustering methods [2] have been addressed in the literature to improve the energy efficiency of the network. [3–7] are single hop based clustering protocols for wireless sensor networks where clusterhead directly communicates to the base station. The benefit of this technique is that the nodes near to the base station are not overburdened and have less delay. However, in these methods, the nodes away from the base station suffer from excessive energy dissipation and thereby expire shortly. Authors in [1, 8–12] proposed various multihop based clustering protocols where clusterheads use multihop routing strategy for sending data packets. An energy aware clustering protocol in WSNs has been designed in [9]. In this approach, each CH selects the next hop clusterhead based on residual energy only. A three layer routing protocol based on LEACH (TL-LEACH) has been designed in [10]. DWEHC [8] uses multiple levels for minimizing intra-cluster energy consumption. Nevertheless, in DWEHC, clusterheads transmit the data directly to the base station. Various multilevel clustering algorithms have been addressed in [1, 11, 12]. In [1], the authors proposed a distributed energy aware multilevel routing algorithm (EMRA) for cluster based sensor networks. EMRA uses multiple paths for sending the packets to next hop nodes. However, EMRA minimizes energy consumption at nodes, the computation of multiple paths and management of paths has high overhead.

3 System Model

We use undirected graph $G = (V, E)$ to model wireless sensor network where V is set of static sensor nodes and E denotes set of links. The base station (BS) is static and placed outside the field. Initial energy of all nodes is assumed to be same and it is given by E_0 . The distance between two nodes u and v is denoted by d_{uv} , while the distance from node u to BS is d_{uBS} . Nodes are location aware. However, the base station knows information of all the nodes in the network. In addition to this, we assume that nodes are organized into different clusters. Clusterhead selection and cluster formation are done according to [6]. CHs collect the data packets from their respective clusters and then send them to the BS in multihop routing. Furthermore, we consider that node's transmission can be adjusted to communicate with neighbor nodes.

Here, we use *first order radio model* [3] for calculating energy consumption in communication. Both the free space (d^2 path loss) and the multipath fading (d^4 path loss) channel models are used depending on the distance between the transmitter and the receiver. Therefore, the energy required for transmitting a l -bit packet over distance d is:

$$E_{tx}(l, d) = \begin{cases} lE_{elc} + lE_{amp}d^2, & d < d_0 \\ lE_{elc} + lE_{amp}d^4, & d \geq d_0 \end{cases} \quad (1)$$

where E_{elc} is energy per bit spent by the circuitry electronics, E_{amp} is the energy dissipation for transmitter amplifier and d_0 is threshold distance. In this energy model, E_{amp} is written as ϵ_{fs} for free space model and to ϵ_{mp} for multipath model. The energy spent for receiving a l -bit packet is:

$$E_{rx}(l) = lE_{elc} \quad (2)$$

4 Proposed Multi-level Weight Based Routing Algorithm

The proposed MWRA is divided into three phases: level assignment mechanism, 2-connected backbone network formation and finally routing. In first phase, the base station assigns level to each sensor node in the network. In next phase, a 2-connected backbone network is formed using CHs. In third phase, nodes forward their packets to the lower level nodes based on the weight function as detailed subsequently.

4.1 Level Assignment Mechanism

In this phase, we assign level (L) to every node u in the network depending on its distance from the base station. Initially, level of all the nodes is zero including the base station i.e., $L(BS) = 0$. Maximum number of levels is k and the network size is given by $N \times N$. Algorithm 1 describes the level assignment mechanism in details. For a given value of k , Algorithm 1 checks whether d_{uBS} lies between 0 and N/K or not. If answer is yes, level of node u i.e., $L(u)$ is assigned to L_1 . If d_{uBS} is between $((N/K) \times j)$ and $((N/K) \times (j + 1))$, where $\forall j = 1$ to $(k - 1)$, then $L(u)$ is set to L_{j+1} .

Algorithm 1: Level Assignment to Deployed Nodes

```

for  $u := 1 : N_{CH}$  do      /*  $N_{CH}$  is number of clusterheads*/
  Get  $d_{uBS}$  ;
  for  $j := 1 : (k - 1)$  do
    if  $(0 < d_{uBS} \leq N/k)$ 
       $L(u) := L_1$  ;
    end if
    if  $((N/k) \times j < d_{uBS} \leq ((N/k) \times (j + 1)))$ 
       $L(u) := L_{j+1}$  ;
    end if
  end for
end for

```

4.2 2-Connected Backbone Network Formation

In this phase, the backbone network is formed with CHs in the network. Backbone range (BR) is used to provide backbone connectivity. Backbone range is the range between the connected CHs. Initially, BR is same for all the nodes $u \in N_{CH}$ and it is equal to R_{max} where R_{max} is maximum transmission range of sensor nodes.

Here, the clusterheads in level L_1 are directly connected to the BS. However, other CHs except nodes in L_1 utilize their levels and BR for selecting next hop CHs. The process is as follows. Each CH, say u , broadcasts a $\langle Hello \rangle$ message (HM) within the range of BR . If distance between two CHs u and v , d_{uv} , is less than BR , then node v sends a $\langle Reply \rangle$ message which contains its level and ID . If $L(v)$ is lower than or equal to $L(u)$ then a directed link (e_{uv}) is formed between them and node v becomes the backbone neighbor (BN_u) of node u . If there is no reply from any CH or number of backbone neighbor is less than 2, the backbone range BR is increased by $\frac{1}{2}BR$ and the process continues. The details of backbone formation method are summarized in Algorithm 2.

4.3 Weight Based Routing

In this phase, routing is performed where clusterhead in a level is only allowed to transmit the packet to a next hop node in next lower level based on a weight function W . Here, the weight function takes residual energy level and link distance into account. For two CH nodes u and v , the weight for the directed link e_{uv} is calculated as:

$$W_{uv} = \left[\frac{RE_v}{E_0} \right] \left[1 - \frac{d_{uv}}{d_{max}} \right] \quad (3)$$

where RE_v is residual energy level of CH node v and d_{max} is maximum distance between nodes. Formula in (3) emphasizes that the CH with less residual energy is less preferred than the CH with high residual energy in next hop selection process.

Algorithm 3 shows the weight based routing process. In this routing process, CH node u sends its packet to one of its backbone neighbors (BN_u). At first, node u sends $\langle Hello \rangle$ message to the nodes in $L(u) - j$ (where $j = 1, 2, \dots$) and waits for $\langle Reply \rangle$ message. Each node $v \in BN_u$ sends $\langle Reply \rangle$ message only if its energy RE_v is greater than *Maximum Threshold* ($MAXTH$). On receiving $\langle Reply \rangle$ messages from the backbone neighbor nodes, W_{uv} is calculated and the node with largest value is chosen as next hop. If there is no $\langle Reply \rangle$ received, variable j is incremented until the condition is satisfied ($j \leq 2$). It is to be noted that, when the node does not receive any $\langle Reply \rangle$ further, the node can send the packets directly to the base station using its maximum power level.

Algorithm 2: 2-Connected Backbone Network Formation

```

for each  $u \in N_{CH}$  in  $L_1$  do
    Add  $e_{uBS} := 1$  ;
end for
 $BR := R_{max}$ ;
for each  $u \in N_{CH} \notin L_1$  do
    Get level  $L(u)$  ;
    repeat
        Send  $HM$  in the range of  $BR$ ;
        if ( $Reply == yes$ )
            Get level  $L(v)$ 
            if ( $L(u) \geq L(v)$ )
                Add  $e_{uv} := 1$ ;
                 $BN_u ++$ ;
            end if
        end if
         $BR := BR + \frac{1}{2}BR$  ;
    until (No response |  $BN_u \leq 2$ )
end for

```

Algorithm 3: Weight based Routing

```

 $MAXTH := \frac{1}{|V|} \sum_{u=1}^V RE_u$  ;
 $j := 1$ ;
for each node  $u \in N_{CH}$  do
    repeat
        Send  $HM$  to nodes in  $L(u) - j$  ;
        Wait  $Reply$  ;
        if ( $RE_v > MAXTH$ ), Send  $Reply$  ;
        if ( $Reply == yes$ )
            Calculate  $W_{uv}$ ;
             $NextHop(u) := \max_{v \in BN_u} W_{uv}$ ;
        end if
        if (No Response),  $j ++$ ;
    until ( $j \leq 2$ )
    if (no neighbor exists)
        Send directly to  $BS$ ;
    end if
end for

```

4.4 Network Reconfiguration Stage

Network reconfiguration is important to achieve balanced energy consumption and also to reduce the unnecessary networking overhead due to frequent re-clustering. In this procedure, the average residual energy of CHs is compared with Maximum Threshold ($MAXTH$) as well as Minimum Threshold ($MINTH$) value as shown in Eq. (7). It has three possible cases. These are:

1. If the average residual energy of CHs is higher than $MAXTH$, same forwarder set is used for data forwarding and *Reclustering* flag is set to 0.
2. If the average energy of CHs is between $MAXTH$ and $MINTH$, then a new set of forwarder nodes is chosen from the existing CHs and *Reclustering* flag becomes -1.
3. If the average energy is less than $MINTH$, re-clustering process is invoked in order to select a new set of CHs and *Reclustering* flag is set to 1.

$$MAXTH = \frac{1}{|V|} \sum_{u=1}^V RE_u \quad (4)$$

$$MINTH = lE_{elc} + lE_{amp}d_{uBS}^4 \quad (5)$$

$$E_{avg-CH} = \frac{1}{|N_{CH}|} \sum_{u=1}^{N_{CH}} RE_u \quad (6)$$

$$Reclustering = \begin{cases} 0: & E_{avg-CH} > MAXTH \\ -1: & MINTH < E_{avg-CH} \leq MAXTH \\ 1: & E_{avg-CH} \leq MINTH \end{cases} \quad (7)$$

5 Performance Evaluation

In this section, the performance of our proposed MWRA is validated through simulations and results are compared with EMRA [1]. Simulation experiments are performed using C compiler and MATLAB. Network lifetime is considered as the primary performance metric and it is defined in terms of number of alive nodes. The second performance metric is average residual energy of the network. In this study, 100 nodes are randomly distributed in $100 \times 100 \text{ m}^2$ square area and the base station is placed outside the field. Other parameters used here are listed in Table 1.

1. *Network lifetime (Number of alive nodes)*: Fig. 1 shows the number of alive nodes as the round proceeds for MWRA and EMRA. The lifetime of both the algorithms gradually decrease as the round proceeds. It is clear from the figure

Table 1 Simulation Parameters

Parameter	Value
Area size	100 × 100 m ²
Number of nodes	100
Transmission range	30 m
Initial energy level	1 J
ϵ_{fs}	10 pJ/bit/m ²
ϵ_{mp}	0.0013 pJ/bit/m ²
E_{elc}	50 nJ/bit
Size of packet	500 bits
Size of control packets	300 bits

Fig. 1 Number of alive nodes versus number of rounds

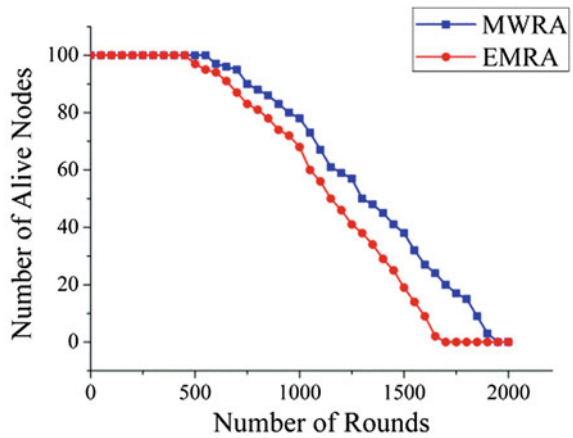
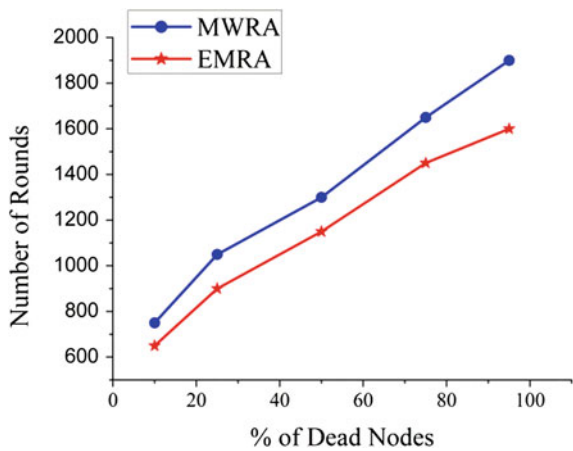


Fig. 2 Percentage of dead nodes versus number of rounds



that our MWRA improves the network lifetime than EMRA. Compared to EMRA, the lifetime under MWRA is increased by 25 % which proves that the proposed method has better network efficiency as well as network coverage. Figure 2 illustrates percentage of dead nodes for different routing algorithms. In all cases, the proposed MWRA improves the network connectivity than EMRA. In MWRA, 25 % dead nodes exist at 1050 rounds, while it is 900 rounds for EMRA. On the other hand, 95 % nodes become dead nodes at 1900 rounds for MWRA and 1600 rounds for EMRA.

- 2. *Average residual energy*: Fig. 3 displays average residual energy of the network for different rounds. In MWRA, average residual energy is much higher than EMRA. This implies that less amount of energy consumed for transmitting packets under MWRA. Moreover, rotating the role of forwarder set minimizes the energy consumption at nodes and thereby improves the energy efficiency of the network significantly.

Fig. 3 Average residual energy versus number of rounds

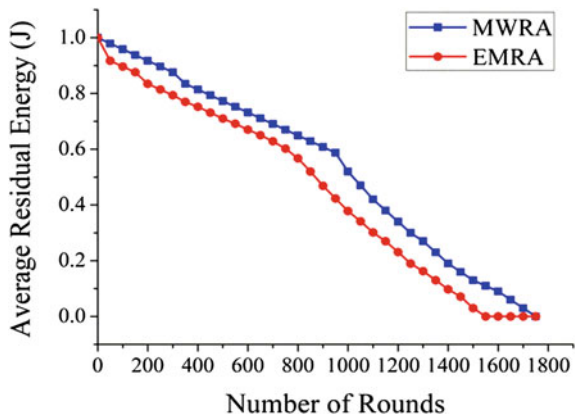


Fig. 4 Number of alive nodes versus number of rounds for various levels

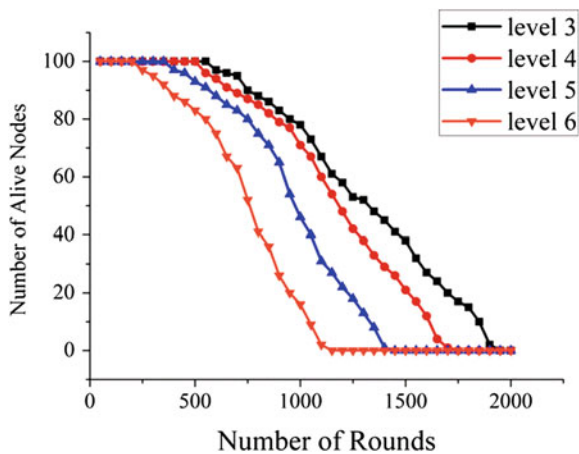
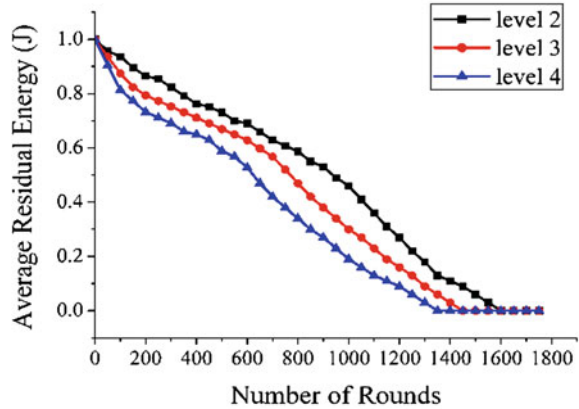


Fig. 5 Average residual energy versus number of rounds for various levels



3. *Energy efficiency at various levels:* Figs. 4 and 5 highlight the number of alive nodes and average residual energy of the network for various levels in MWRA. In this experiment, the number of alive nodes and average residual energy gradually decreases as the number of levels increases. The reason behind this phenomenon is that increase in number of levels reduces the possibility of appearing next hop nodes in next level and as a result, nodes increase their transmission power level to get their next hop.

6 Conclusion

In this paper, we presented a Multi-level Weight based Routing Algorithm (MWRA) for minimizing energy consumption cost in cluster based sensor network. The proposed method builds a 2-connected backbone network for selecting relay nodes. An energy aware weight function has been developed for choosing next hop node. Furthermore, MWRA uses network re-configuration to balance the energy level of nodes. Extensive simulations have been performed to evaluate the effectiveness of the proposed scheme. Simulation results reveal that the proposed MWRA improves the energy efficiency of the network than the clustering algorithm, EMRA.

References

1. Amgoth, T., Ghosh, N., Jana, P. K.: Energy Aware Multi-Level Routing Algorithm for Two-Tier Wireless Sensor Networks. In: International Conference on Distributed Computing and Internet Technology, Springer International Publishing, 111–121 (2014)
2. Abbasi, A.A., Younis, M.: A Survey on Clustering Algorithms for Wireless Sensor Networks. *Computer Communications* 30, 2826–2841 (2007)

3. Heinzelman, W.B., Chandrakasan, A., Balakrishnan, H.: Energy Efficient Communication Protocols for Wireless Micro Sensor Networks. In: Hawaii International Conference on System Sciences, pp.1–10 (2000)
4. Heinzelman, W.B., Chandrakasan, A., Balakrishnan, H.: An Application Specific Protocol Architecture for Wireless Micro Sensor Networks. *IEEE Transactions on Wireless Communications* 1, 660–670 (2002)
5. Dimokas, N., Katsaros, D., Manolopoulos, Y.: Energy Efficient Distributed Clustering in Wireless Sensor Networks. *Journal of Parallel and Distributed Computing* 70, 371–383 (2002)
6. Lee, H.S., Kim, K.T., Youn, H.Y.: A New Cluster Head Selection Scheme for Long Lifetime of Wireless Sensor Networks. In: International Conference on Computational Science and Its Applications, Springer Berlin Heidelberg, pp. 519–528 (2006)
7. Bhowmik, S., Sen, A., Bhattacharjee, S.: Balanced Neighborhood Aware Clustering Technique in Wireless Sensor Networks. In: IEEE International Conference on Power, Control and Embedded Systems, pp. 1–5 (2012)
8. Ding, P., Holliday, J., Celik, A.: Distributed Energy Efficient Hierarchical Clustering for Wireless Sensor Networks. In: Distributed Computing in Sensor Systems, Springer Berlin Heidelberg, pp. 322–339 (2005)
9. Yu, M., Kin, K.L., Ankit, M.: A Dynamic Clustering and Energy Efficient Routing Techniques for Sensor Networks. *IEEE Transaction on Wireless Communications* 6, 3069–3079 (2007)
10. Zhixiang D., Bensheng, Q.: Three Layered Routing Protocol for WSN Based on LEACH Algorithm. In: IET Conference on Wireless, Mobile and Sensor Networks, pp.72–75 (2007)
11. Liu, Y., et al.: Multi-Layer Clustering Routing Algorithm for Wireless Vehicular Sensor Networks. *IET Communication* 7, 810–816 (2009)
12. Kim, M., Kim, S., Seo, J., Choi, K., Han, S.: CAPNet: An Enhanced Load Balancing Clustering Algorithm for Prolonging Network Lifetime in WSNs. *International Journal of Distributed Sensor Networks*, pp. 1–8 (2014)

An Approach to Optimize Unnecessary Energy Consumption During Dynamic Routing in Wireless Sensor Networks

Narasimha Kamath and U.K. Anirudh

Abstract The paper proposes an approach to optimize unbalanced energy consumption during dynamic routing in WSN (wireless sensor network). The concept of intermediate nodes are generally used to route or transmit overloaded data. In a normal scenario, when gathered data is not overloaded and are to be transmitted to the required destination, presence of intermediate nodes might pose a communication delay causing the overall energy to reduce by the certain amount. We have analyzed the given problem to ensure optimality and thus reduce the delay due to intermediate nodes in WSN.

Keywords WSN • Unbalanced energy • Dynamic routing • Intermediate nodes • Delay • Optimization

1 Introduction

The primary definition of routing is a path or a route that leads the sensed data or information to the required destination. The two major types of routing include:

- Static Routing
- Dynamic Routing

Static routing uses a table that consists of connected nodes, shortest paths or routes from the desired node to the destination node. Updating the route table is not possible, due to which static routing is meant for small scale WSN having primitive functionalities and are majorly used for small data transfer [1]. Whereas, dynamic

N. Kamath (✉) · U.K. Anirudh
Department of Computer Science and Engineering,
PESIT-BSC, Bangalore 560100, India
e-mail: narasimha4593@gmail.com

U.K. Anirudh
e-mail: anirudh.anirudh93@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_20

205

routing provides several functionalities that support information updating and inclusion of other factors to accomplish routing in a large scale WSN.

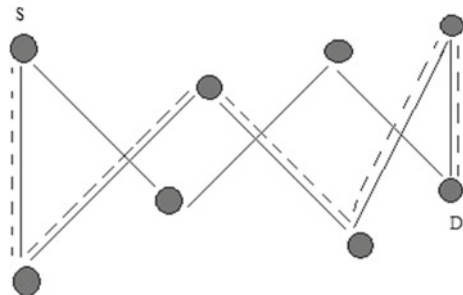
The major overheads in routing are mainly due to unbalanced energy consumption or delay in the delivery of data packets due to inherent ambiguity in the underlying WSN [2]. Security attacks in WSN is also a primary reason for routing overheads but the paper focuses on optimizing the unbalanced energy consumption during dynamic routing hence, security attacks in WSN is of least importance in the proposed scheme [2]. The predefined routing scheme in WSN may or may not lead to an unnecessary energy expense depending upon the type of data that is to be carried towards the destination node.

Major hurdles in dynamic routing scheme occur when the aggregated data from the CH (Cluster Head) is transmitted to the BS (Base Station) and the information is overloaded. Data overloading is an indirect reason for unbalanced energy consumption. Therefore, routing an overloaded data can be simplified by including the intermediate nodes and assigning the overloaded data to them.

Another cause for unbalanced energy consumption is solely due to the presence of duplicate nodes. A duplicate node contains dummy data that drains the overall energy in WSN. Presence of duplicate nodes cannot be identified in the route table and hence causes a severe damage. Once a network is formed, a message is flooded to the connected nodes from the source node. When a reply message is received along with the additional information like node ID, current energy level and the battery life, a route table is formed to minimize further upgrading thus reducing the count of duplicate nodes in WSN [3]. The data organization for dynamic routing will be in the form of packets. Data packets encapsulating the original messages are transmitted within the network for a given time slot to minimize network congestion.

Designing a suitable approach for minimizing the unwanted energy expense requires several networking parameters to be taken into account. The next section of the paper describes the aspects of energy optimization along with a proposed scheme to minimize routing overheads that leads to energy wastage in WSN (Figs. 1 and 2).

Fig. 1 Routing in WSN



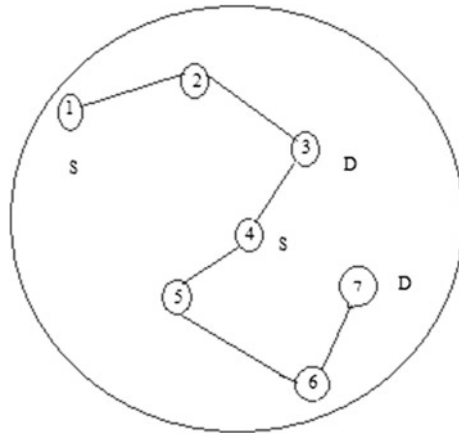


Fig. 2 WSN nodes connected through symmetric links

2 Related Works

The main aim of the paper is to optimize the unbalanced energy consumption during dynamic routing. The energy consumption becomes unbalanced depending upon the intermediate node count and the density of the gathered data. Work described in [4], provides some useful information regarding the routing protocol that provides application specific service guarantees. This paper presents a new cluster-based route optimization and load-balancing protocol, called ROL combining several application requirements. The approach proposed in [4] automatically selects the nodes to transfer the gathered data. Energy budget for data communication is known before routing to reduce unbalanced energy consumption (Tables 1, 2, 3, 4 and 5).

The approach proposed in [5] proposes an algorithm to reduce uneven, energy consumption in large scale wireless sensor network. By using greedy policy and dynamic programming, a heuristic topology control algorithm with time complexity $O(n(m + \log(n)))$ where n and m are the number of nodes and edges in the network.

Table 1 Tabular representation of connected nodes

Source nodes	Destination nodes	Intermediate nodes	Delay parameter
1	3	4	α
4	7	5, 6	$\alpha + \alpha = 2\alpha$

Table 2 Representation of connected nodes

Source nodes	Destination nodes	Intermediate nodes	Delay parameter
1	3	–	α
4	7	–	α

Table 3 Representation of connected nodes

Source nodes	Destination nodes	Intermediate nodes	Delay parameter
1	3	2	α
4	6	5	α
7	9	8	$\alpha + \alpha'$

Table 4 Tabular representation of connected nodes

Source nodes	Destination nodes	Intermediate nodes	Delay parameter
1	3	–	–
4	6	–	–
7	9	8	$\alpha + \alpha'$

Table 5 Experimental values

Parameters	Values (J)
E_{DSP}	3.8012
$E_{OPT'}$	2.5678
$E_{OPT''}$	1.2334
β	3.2340
μ	7.0093
$\Delta\mu$	3.6703
$E_{overall}$	8.3480
$E_{OPT(overall)}$	3.5190
$\Delta\beta$	1.3480

The proposed approach uses a mobile sink over a multi-hop communication path to maximize the network life time.

The reasons for unbalanced energy consumption are mainly due to lack of application of specific nature in WSN. [6] proposes a new application specific low routing protocol named ASLPR to elect the optimal cluster heads. In this work, a hybrid algorithm based on genetic algorithm and simulated annealing is applied to ASLPR.

3 Proposed Approach

3.1 Network Model

In this paper, we consider a WSN having K sensor nodes dispersed in a given region of a rectangular area $L \times W$. We consider some of the basic assumptions of the sensor nodes and the underlying network model.

- The nodes are uniformly distributed within a cluster C with a density ∂ .
- A BS is connected to the sensor nodes through symmetric links.
- Dynamic routing scheme is employed to transmit the data packets.
- Sensor nodes can recognize their geographical location and BS's via information exchange.
- Size of the data packets remains the same for the entire process.
- A unique ID is assigned to each sensor nodes operating with the initial energy E_i .

If a d -bit message is transmitted via the transmitting nodes through dynamic routing approach, the minimum amount of energy required to do so is given by:

$$E_{\min} = R_c * K_{\text{current}} * S_t \leq E_{\text{overall}} \tag{1}$$

The consumed energy of aggregating a message with d -bit is

$$E_A(m, d) = d * E_{DA} \tag{2}$$

where, E_{DA} is the energy dissipated per bit to aggregate message signal.

If (k_i, k_{i+1}) and (k_j, k_{j+1}) are the sensor nodes represented in the coordinate form, then the distance d is calculated as follows:

$$d = \sqrt{(k_i - k_j)^2 + (k_{i+1} - k_{j+1})^2} \tag{3}$$

The operating frequency f_o of the sensor nodes is affected by the processor frequency f . The performance of the sensor nodes is dependent on f_o and the variation in f_o is solely due to changes in f .

$$\Delta f_o = f_o - f/f_o \tag{4}$$

Let $K_{\text{current}} = \{k_1, k_2, k_3 \dots k_i\}$ be the set of nodes in the routing table T . The energy expense due to routing overheads can be classified into two different categories:

- Intracluster Routing Overheads
- Intercluster Routing Overheads

Consider a node $K \in K_{\text{current}} \in T, \exists$ a node k_i with an initial energy E_{\min} carrying a d -bit information towards the destination node D with an operating frequency f_o . Consider E_{DSP} as energy dissipated due to dynamic routing overheads and E_{OPT} as optimized energy that minimizes the unbalanced energy expense. Though the energy expenditure due to dynamic routing overhead is considered to be minor, preserving E_{overall} becomes a major task to accomplish. The problem is considered to be minor depending upon the sensor nodes included in the routing table. A very small expense will also add up during the optimization process to maximize E_{DSP} which can later become a critical issue to resolve.

The problem can be formulated as follows:

$$\max \sum_{k=1}^i (E_{OPT(overall)}) \geq \sum_{k=1}^i ((E_{min} + E_i) - E_{DSP}) \quad \text{s.t. } E_i > E_{min} \quad (5)$$

The above optimization problem can be solved as two sub problems as follows.

- a. Optimizing the overheads due to intracluster dynamic routing
- b. Optimizing the overheads due to intercluster dynamic routing.

3.2 Intra-cluster Dynamic Routing

In order to optimize the overheads due to intracluster dynamic routing, we consider a node separation approach to minimize the unbalanced energy expense.

Lemma 1 Consider a WSN represented using a graph $G = (V, E)$ having k_i sensor nodes. Say if node A is connected to node C via an intermediate node B within a cluster C having a radius R_c and α being the delay parameter. Prove that $E_{OPT} = \sum_{k=1}^i \alpha_i - \alpha_j$ where, $i > j$.

Proof If a WSN has k_i sensor nodes, considering the assumption of symmetric links between the nodes we can make an additional assumption of symmetric cluster size to ensure that the delay parameter becomes equivalent depending upon the intermediate nodes.

The major hurdle in dynamic routing is due to intermediate nodes k_j . When data packets are overloaded, we make use of intermediate nodes to accomplish dynamic routing. In a generic scenario, intermediate nodes might add an additional delay to data communication causing unbalanced energy consumption.

Consider a region having m symmetric clusters. The overall cluster radius can be calculated as shown below.

$$R_c = r_1 + r_2 + r_3 + \dots + r_m \cong \sum_{r=1}^m R_c \quad (6)$$

The tabular representation of the network is shown below.

By referring the table, we can infer that the delay parameter α can be represented in a general form when a WSN has k nodes.

$$E_{DSP} = \sum_{k=1}^i \alpha_i = i * \alpha \quad (7)$$

Delay can occur in many ways. In this paper, we consider the energy loss due to unwanted intermediate nodes in WSN. The approach used to minimize the delay parameter uses node separation technique. In this technique, instead of flooding the

message from source node to the connected nodes, we send a connectivity message directly to the destination node. Since WSN has a predefined network topology, we can know the desired node with which a connection has to be formed. When the data is not overloaded, the presence of intermediate nodes becomes unwanted and hence we use this approach to separate the intermediate nodes.

Removing the intermediate nodes minimize the energy expense from $i * \alpha$ to α . Thus, we can prove that if j unwanted intermediate nodes are separated from the existing network, the overall delay α_i becomes $\alpha_i - \alpha_j$. If j intermediate nodes having a delay parameter α_j , subtracted from k sensor nodes having a delay α_i . the optimized energy E_{OPT} is given by:

$$E_{OPT} = \sum_{k=1}^i \alpha_i - \alpha_j \cong \alpha \text{ provided, } i > j \tag{8}$$

3.3 Inter-cluster Dynamic Routing

Intercluster routing is evident, when the destination lies outside the existing cluster. When the sensor nodes within the cluster communicates with a node residing outside the specified boundary, preserving the signal strength (S_T) of the existing network becomes a major task to accomplish. The major hurdle in intercluster routing is to balance the delay parameter of the existing network and the remote network.

Lemma 2 Consider a WSN W represented using a graph $G = (V, E)$ and a remote network W' represented as $G' = (V', E')$. Say if node A of the existing network is connected to the node C of the remote network W' via an intermediate node B residing in the existing network. If α is the delay parameter of W and α' is the delay parameter of W' , prove that $E_{OPT} = (\sum_{k=1}^i \alpha_i - \alpha_j) + (\sum_{k'=1}^i \alpha')$ and also prove that the optimized energy due to intracluster routing is always greater than the optimized energy due to intercluster routing.

Proof For an intercluster routing, we consider an existing network W connected to the individual node of remote network W' . The intermediate node connecting the remote node, resides in the existing network. The diagrammatic representation of the above scenario is shown below.

By referring the above table, we can infer that the parameters α and α' can be represented in a general form as follows.

$$E_{DSP} = \sum_{k=1}^i \alpha_i + \sum_{k'=1}^i \alpha'_i = i * \alpha + i * \alpha' \tag{9}$$

Applying the node separation technique to the above network, we get:

By observing the table, we can conclude that the node connecting the existing and the remote network (node no 8) is not separated using node separation

technique. When node 8 is removed, connection between the networks is lost and there is no room for intercluster routing. Hence, the delay due to node 8 must be retained because when we choose one of the sensor nodes as an intermediate node for intercluster routing, we place several criteria to select an intermediate node and hence node separation technique is not applicable to node 8.

It is obvious that the optimized energy due to intercluster routing is lesser than the optimized energy due to intracluster routing. Since the former methodology has an additional node from the remote network connected to the existing network. We also consider the delay in remote network causing $E_{OPT'} > E_{OPT''}$.

If j nodes are separated from i sensor nodes and considering the delay due to the remote network, the equation for $E_{OPT''}$ can be written as follows

$$E_{OPT'} = \left(\sum_{k=1}^i \alpha_i - \alpha_j \right) + \left(\sum_{k'=1}^i \alpha' \right) \cong \alpha + \alpha' \quad (10)$$

3.4 Problem Optimization

To optimize the formulated problem, consider the Eqs. (8) and (10). $E_{OPT(overall)}$ is given by:

$$E_{OPT(overall)} = E_{OPT'} + E_{OPT''} \quad (11)$$

Thus,

$$E_{OPT(overall)} = \sum_{k=1}^i \alpha_i - \alpha_j + \left(\sum_{k=1}^i \alpha_i - \alpha_j \right) + \left(\sum_{k'=1}^i \alpha' \right) \quad (12)$$

Simplifying Eq. (13), we get:

$$E_{OPT(overall)} = \sum_{k=1}^i \alpha_i - \alpha_j \left(1 + \sum_{k'=1}^i \alpha' \right) \quad (13)$$

Since the remote network W' consists of a single node, we can replace $\sum_{k'=1}^i \alpha'$ by α' and $\sum_{k=1}^i \alpha_i - \alpha_j$ by α , we can re write Eq. (13) as follows.

$$E_{OPT(overall)} = \alpha(1 + \alpha') \quad (14)$$

E_{DSP} becomes minimal after applying the node separation technique. So, the value of E_{DSP} is nearly equal to $\alpha + \alpha'$. The minimum energy (E_{min}) required to accomplish dynamic routing is always greater than the individual node energy (E_i).

If the value of E_{\min} for the entire network is considered to be β and the value of E_i is μ , we can substitute the values of E_{\min} and E_i in Eq. (5).

$$\alpha(1 + \alpha') > ((\beta + \mu) - (\alpha + \alpha')) \quad (15)$$

The value of E_{\min} and E_i can vary depending upon the sensor nodes included in the routing table T. It is always known that the energy of WSN is never stable and can change invariably depending upon the internal factors like processor frequency, hardware components used, operating frequency and so forth.

If change in β is $\Delta\beta$ and change in μ is $\Delta\mu$, we can infer that the changed value $\Delta\beta < \beta$ and $\Delta\mu < \mu$. We are assuming the value of α to be consistent throughout the process because, the changes in the optimized value lead to further optimization and the process requires the initial value of β and μ which is practically not possible to achieve. Once the delay parameter is minimized, we should try minimizing α' . But in this paper, the existing network is connected to an individual node of the remote network. Thus, α' value is assumed to be negligible. We can rewrite Eq. (16) as follows:

$$\alpha > ((\Delta\beta + \Delta\mu) - \alpha) \quad (16)$$

$$\alpha / ((\Delta\beta + \Delta\mu) - \alpha) > 1 \quad (17)$$

Equation (17) clearly notifies that the formulated problem holds well irrespective of the networking parameters and the specified constraints. The value of $\Delta\beta$ and $\Delta\mu$ is considered after the routing phase is completed and their values decide the percentage of optimality in the formulated equation. Hence, $\Delta\beta$ and $\Delta\mu$ are called the deciding parameters of the formulated equation.

Thus, the energy expense or delay due to dynamic routing can be minimized if the supporting network parameters are valid and functional. The proposed scheme is applicable for a small scale WSN for time being and gives satisfactory results for the input parameters if specified correctly.

4 Simulation and Graphical Representation

We conduct simulation based studies to evaluate the performance of the proposed scheme. For this purpose, we have used ns2 simulator, to analyze the performance and interpret the simulation results. Figure 3 depicts a snapshot of WSN simulation.

We have also compared the delay due to intracluster and intercluster routing. A graph of time versus delay time depicts the overall lag caused due to the routing overheads (Figs. 4, 5, 6 and 7).

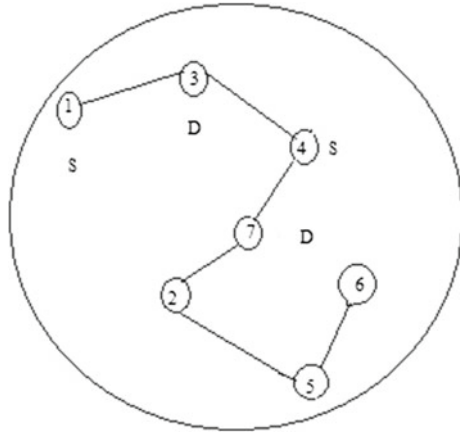


Fig. 3 Applying node separation technique to WSN

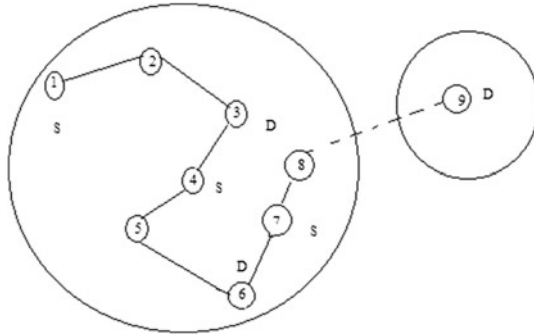


Fig. 4 Intercluster routing in WSN

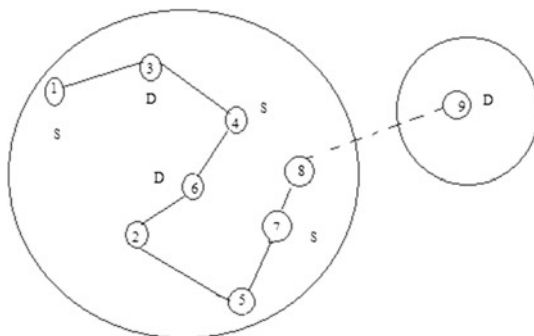


Fig. 5 Applying node separation technique to WSN

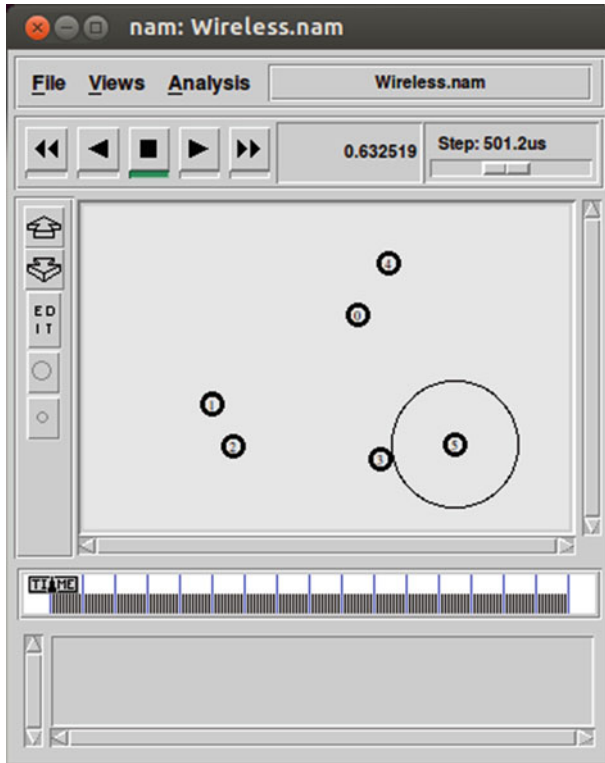


Fig. 6 WSN simulation snapshot

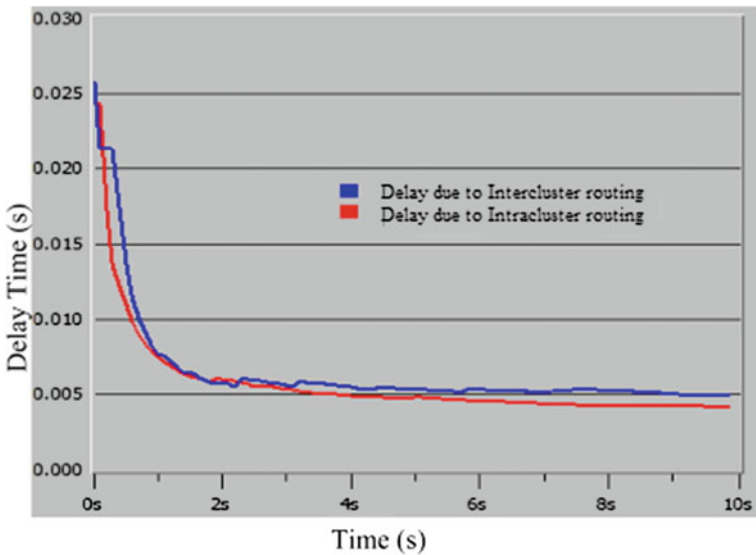


Fig. 7 Graphical representation of the routing overheads

Since the existing WSN is connected to an individual node of the remote network, the delay due to intercluster routing is slightly higher than the delay due to intracluster routing. If the existing network was connected to the entire remote network, the delay gap would have been much larger causing linear raise in the value of E_{DSP} .

5 Conclusion

In this paper, we have proposed an approach to simplify the overheads due to dynamic routing in WSN. Basically, the delay or unnecessary energy expense due to dynamic routing is considered as a minor issue. In this paper, we have tried to prove that the delay isn't minor and have also described the criticality posed especially in a large scale WSN having large number of functional sensor nodes. Though the outcome does not severely affect the functionalities of a WSN, we have to make sure that the energy dissipation are simplified using the existing solutions to achieve better results.

References

1. SM Jameii, K Fuez, M Deghan, "Multi Objective optimization for topology and coverage control in wireless sensor networks ", International Journal of Distributed Sensor Networks", Hindawi Publishing, 2015.
2. Jun Long, Anfeng Liu, Zhi Li, "An energy-efficient and sink-location privacy enhanced scheme for WSN's through ring based routing", Journal of parallel and Distributed computing, Vol.81, Elsevier, 2015.
3. Tongtong Li, Jian Ren, Jie Wu, "cost-Aware Secure Routing (laser) protocol Design For Wireless Sensor Networks", V0 1-26, issue-4, IEEE, 2015.
4. Mohammad Hammoudeh, Robert Newman, "Adaptive Routing in Wireless Sensor networks, QOS Optimization for enhanced application performance", Information Fusion, Elsevier, 2015.
5. Huan Zaho, Songtao Guo, Xiaojian Wang, Fei Wang, "Energy Efficient Topology control Algorithm For maximizing The Network Lifetime In Wireless Sensor Networks With Mobile-sink", applied soft computing, Elsevier, 2015.
6. Mohammad Sho Kouhifar, Ali Jalali, "A New Evolutionary Based Application Specific Routing Protocol For Clustered Wireless Sensor Networks", International journal of electronics and communication, Vol.69, Elsevier, 2015.

Game Theoretic Modeling of Gray Hole Attacks in Wireless Ad Hoc Networks

Chintan Ketankumar Doshi, Sreecharan Sankaranarayanan,
Vidyashankar B. Lakshman and K. Chandrasekaran

Abstract Wireless ad hoc networks rely on the cooperation of participating nodes to undertake activities such as routing. Malicious nodes participating in the network may refuse to forward packets and instead discard them to mount a denial-of-service attack called a packet drop or blackhole attack. Blackhole attacks can however be easily detected using common networking tools like trace route as all packets passing through the malicious node is dropped. A gray hole attack on the other hand accomplishes denial of service by selectively dropping packets thus escaping detection. In this paper, a novel two player incomplete information extensive form game is used to model the defender and the attacker both of whom are considered rational agents in an effort to determine their optimal (equilibrium) strategies under different values for the parameters true detection rate, false alarm rate, packet value, packets forwarded per unit time, probability of the node being a gray hole, cost of exposure of the attacker and cost of not using a node for the defender. The respective equilibrium strategies if followed guarantee maximum possible protection for the defender and maximal possible damage potential for the attacker.

Keywords Game theory for security · Wireless ad hoc networks · WANETs · Packet drop attack · Blackhole attack · Gray hole attack

C.K. Doshi (✉) · S. Sankaranarayanan · V.B. Lakshman · K. Chandrasekaran
Department of Computer Science and Engineering,
National Institute of Technology Karnataka, Surathkal, India
e-mail: chintandoshi94@gmail.com
URL: <http://cse.nitk.ac.in>

S. Sankaranarayanan
e-mail: sreecharan93@gmail.com

V.B. Lakshman
e-mail: vidyashankarbl@gmail.com

K. Chandrasekaran
e-mail: kchnitk@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International
Conference on Signal, Networks, Computing, and Systems*, Lecture Notes
in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_21

217

1 Introduction

Wireless ad hoc networks (WANETs) are decentralized wireless networks that do not rely on any pre-existing infrastructure. Instead, every node that is a part of the network agrees to forward or relay packets thereby ensuring the collective functioning of the network. For a selfish node however, the forwarding of packets not meant originally for it results in the wastage of precious bandwidth. A malicious node therefore, may choose to drop packets instead of relaying them to mount a denial of service attack called the packet drop attack.

A packet drop attack where the node drops all packets being relayed through it is called a blackhole attack [2]. Blackhole attacks however can be easily detected using common networking tools such as trace route. Moreover, once the node is compromised or revealed to be malicious, other nodes begin to remove the compromised node from their forwarding table and eventually no traffic flows through the compromised node [20].

More sophisticated packet drop attacks have therefore emerged called gray hole attacks where the malicious node tries to drop packets selectively at a rate that can be misconstrued to be due to loss or congestion in the network thereby proving much harder to detect [1].

In this paper, gray hole attacks in WANETs are modeled as an extensive form incomplete information two player game. Earlier approaches for modeling gray hole attacks are either not applicable broadly to WANETs or use weaker game models. Incomplete Information is a stronger notion that considers that every node knows its own type (malicious or non-malicious) but does not know the type of the other nodes. This interim equilibrium analysis therefore proves stronger than the earlier approaches which are discussed in more detail in the Sect. 2.

The rest of the paper is organized as follows. Section 2 describes the related work in this area while elucidating the research gap which our solution addresses. Section 3 is used for defining the game and Sect. 4 provides the equilibrium analysis for the game defined. Section 5 provides experimental results and discussions regarding the same. Section 6 summarizes the contents of the paper while Sect. 7 concludes with pointers to future work.

2 Related Work

The lack of a centralized architecture in WANETs exposes them to various security vulnerabilities most of which have been studied since the advent of WANETs. Denial-of-service (DoS) attacks [4] which WANETs are susceptible to have been surveyed by various authors at different time periods. Wu et al. [16] presented a survey of attacks on mobile ad hoc networks (MANETs) and possible countermeasures. Routing security attacks such as wormhole and sinkhole attacks have been discussed extensively.

Existing solutions to identify and tackle gray holes in the network often involve sweeping changes to the underlying routing protocol or demand additional implementation at the network level. Karlof et al. [7] first described a multi-path based prevention mechanism against gray hole attacks. Yu and Xiao [19] used a multi-hop acknowledgement scheme at intermediate nodes that reports abnormal packet losses to both the base station and the source node. They also provided an improvement to this multi-hop acknowledgment scheme called CHEMAS [17]. CHEMAS uses checkpoint nodes which are responsible for generating acknowledgements. Loss of acknowledgements indicate a suspect node. Kaplantzis et al. [6] proposed a support vector machine based approach for gray hole node detection in wireless sensor networks.

Methods tackling decentralized, infrastructure-less nature of MANETs emerged later. Kanthe et al. [5] proposed an algorithm for detecting gray hole nodes based on sequence number comparison between sent and received packets. The sequence number difference is essentially compared with the expected forwarding rate in the network. A significant drop in the same is used to flag malicious nodes. Xiaopeng and Wei [18] use a proof based method that uses an audit node to determine the current congestion rate, the uses that as the basis to judge other nodes. The Trace Gray algorithm [15] uses a timer to identify gray holes. A packet not returning to its home context before the timeout indicates the presence of a malicious node. All of these methods have significant overhead on nodes that are already selfish in terms of resource utilization. This shortcoming is therefore addressed by our game theoretic approach that aims to give every node an optimal strategy to pursue in order to maximize defense and minimize the damage potential without additional overhead at the network level.

Game Theory has been extensively applied to security and privacy scenarios. As networks start to play an increasingly important role in modern society, security scenarios which involve the direct participation of network agents have emerged. These agents often act on self behalf and can be cooperative, selfish, malicious or any combination of these. Consequently, decision making approaches in general and game theory in particular lend well to security scenarios [12]. Liao et al. [9] provide a classification of attacks on WANETs that can be tackled by game theoretic approaches.

Game theoretic approaches have been used already to tackle gray hole attacks. Mahmoud et al. [11] proposed a stimulation and punishment mechanism to evict irrational packet droppers from the network. This in the game theoretic sense can be considered akin to payoff. Reddy et al. [13] use a zero-sum game approach coupled with selective acknowledgements to model the detection of malicious nodes [8]. Shila et al. [14] use a non-cooperative Markov game for detection of malicious nodes in wireless mesh networks. All of these approaches lack the stronger notion of incomplete information which considers that every node knows only its own type (malicious or non-malicious) and not the type of any of the other nodes.

Our approach therefore aims to provide an optimal strategy for every node without additional network level overhead while considering the stronger incomplete information notion.

3 Game Definition

The two players in the game are

$$N = \{1: \text{The forwarder (Malicious with probability } \mu), 2: \text{The sender}\}$$

The type set for both players can be given by

$$\theta_1 = \theta_2 = \{0: \text{Malicious}, 1: \text{Non-Malicious}\}$$

The type of player 2 is 1 Non-malicious. This is common information i.e. every node is aware of it. The type of player 1 however, is private to the player without any deterministic way of revealing the same. The probability distribution on the type of player 1 is known. He can be malicious with probability μ and therefore non-malicious with probability $1 - \mu$. This game model has been inspired by the intrusion detection system presented by Liu et al. [10]

The strategy set for each player is given below

$$S_1 = \{1: \text{Drop}, 2: \text{Forward}\}$$

$$S_2 = \{1: \text{Do Not Use}, 2: \text{Use}\}$$

Player 1 can choose to either drop or forward the packer that he receives. Player 2 can choose to whether to use the node that Player 1 represents.

Figure 1 describes the sequential play of the game. The utility values for each strategy profile is given by the leaf nodes.

Table 1 provides a summary of all the parameters and their meaning. If a sender does not use a particular node, he will pay some cost due to the absence of that node. This cost represents the overhead incurred because of removal of that node from the forwarding table as well as establishing a new route by initiating rerouting. This cost

Fig. 1 Extensive form game tree

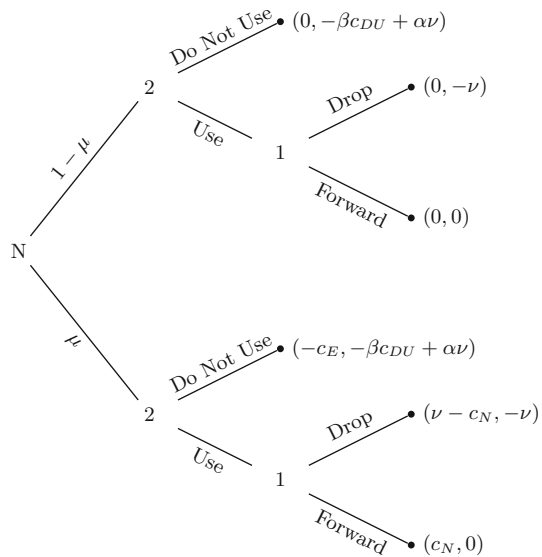


Table 1 Summary of parameters

α	True detection rate
β	False alarm rate
v	Value of a packet
μ	Probability of a node being malicious
c_E	Cost of exposure
c_N	Cost incurred in the next game
c_{DU}	Cost of not using a node
I_i	Information set for player i
Θ_i	Type set for player i
N	The set of players
S_i	The strategy set for player i

is given by c_{DU} . Therefore, if the sender chooses to avoid a particular node, the payoff he receives by doing so is given as $(-\beta c_{DU} + \alpha v)$ where α is the true detection rate and β is the false alarm rate [12]. The values of α and β are inherent properties of the node. v is the value of the data packet. Therefore, if the forwarder drops the packet, the sender receives a payoff of $(-v)$.

Non-malicious nodes may also drop packets because of congestion in the network. If a node is malicious and it is not used by the sender, then it receives $(-c_E)$ payoff which is the cost of exposure for that malicious node. When a malicious node is used and it drops the packet, it receives a payoff of $(v - c_N)$ where c_N is the cost the malicious node pays in the next round. c_N captures history in the repeated play of the game. If a node has dropped packets maliciously in the past, it risks detection in the future. That risk is quantified by c_N . The sender keeps track of packets dropped by every node it is connected with and updates the value of α and β accordingly. This will result in an equilibrium not in favor of the malicious node thus proving successful in thwarting the gray hole attack. Assumptions regarding the parameters and their value is given below.

$$\begin{aligned}
 c_N, c_E &< v \\
 v &> 0 \\
 \alpha, \beta &\in [0, 1]
 \end{aligned}$$

4 Interim Equilibrium Analysis

The extensive form game is useful for providing an intuitive explanation of the game. For the purpose of analysis however, the game is converted into a Bayesian normal form game. It is assumed that Player 1 is malicious and only he knows his type. We can therefore calculate ex-interim utility values for both the players. Table 2 shows the ex-interim utility value for each player.

Table 2 Expected utility value matrix for normal form game

1, 2	Use	Do not use
Drop	$v - c_N, 1 - 2\mu v$	$-c_E, -\beta c_{DU} + \alpha v$
Forward	$c_N, 0$	$-c_E, -\beta c_{DU} + \alpha v$

4.1 Dominant Strategies

It is clear from Table 2 that for Player 1, if Drop has to be the dominant strategy then

$$v - c_N \geq c_N \implies v \geq 2 * c_N$$

For Player 2, if Use has to be the dominant strategy then

$$1 - 2\mu v \geq -\beta c_{DU} + \alpha v \text{ and } 0 \geq -\beta c_{DU} + \alpha v$$

$$\implies v \leq \frac{\beta c_{DU}}{\alpha} \text{ and } v \leq \frac{1 + \beta c_{DU}}{2\mu + \alpha}$$

4.2 Nash Equilibrium Analysis

The strategy (Drop, Use) becomes a Nash equilibrium for the game if

$$v \geq 2c_N \text{ and } v \leq \frac{1 + \beta c_{DU}}{2\mu + \alpha}$$

The strategy (Forward, Use) becomes a Nash equilibrium for the game if

$$v \leq 2c_N \text{ and } v \leq \frac{1 + \beta c_{DU}}{2\mu + \alpha}$$

The strategy (Drop, Do Not Use) becomes a Nash equilibrium for the game if

$$v \geq \frac{1 + \beta c_{DU}}{2\mu + \alpha}$$

The strategy (Forward, Do Not Use) becomes a Nash equilibrium for the game if

$$v \geq \frac{\beta c_{DU}}{\alpha}$$

5 Experimental Results and Discussions

The game that has been analyzed mathematically in the previous section is played exactly once. In the real life scenario however, the sender and the forwarder repeatedly play this game. In the repeated play of the extensive form game, current game play is affected by every player’s actions in previous plays of the game. Consequently, some parameter values are carried over or updated to capture this. In our game model, α and c_N are such parameters.

5.1 α Value Determination

As defined previously, α is the True Detection Rate. The value of α however is dynamic and changes after each game play. α therefore, represents the current assumption that the sender has about the malicious nature of the forwarder. The updates to the value of α are done according the following equations.

$$\alpha_i = \begin{cases} f_i(\alpha_{i-1}), & \text{if the packet is dropped by the forwarder.} \\ f_d(\alpha_{i-1}), & \text{if the packet is relayed by the forwarder.} \end{cases}$$

If the packet is dropped, then the value of α is incremented by the function f_i . If the packet is forwarded, the value of α is decremented by the function f_d . The updated value of α is then used in the next round of game play. Since the value of $\alpha \in [0, 1]$, the functions should be such that the increment and decrement does not produce an α outside this range. The functions f_i and f_d can thus be selected in the following ways.

- Additive Increase, Additive Decrease (AIAD)

$$f_i(\alpha_{i-1}) = \alpha_{i-1} + IF$$

$$f_d(\alpha_{i-1}) = \alpha_{i-1} - DF$$

Here IF and DF are the linear increase and decrease factors respectively.

- Multiplicative Increase, Multiplicative Decrease (MIMD)

$$f_i(\alpha_{i-1}) = IF * \alpha_{i-1}$$

$$f_d(\alpha_{i-1}) = \alpha_{i-1} / DF$$

Here IF and DF are the multiplicative increase and decrease factors respectively.

- Multiplicative Increase, Additive Decrease (MIAD)

$$f_i(\alpha_{i-1}) = IF * \alpha_{i-1}$$

$$f_d(\alpha_{i-1}) = \alpha_{i-1} - DF$$

Here IF is the multiplicative increase and DF is the linear decrease factor.

The strongest possible assumption is made in this case i.e. the scheme chosen for increment and decrement by the sender is known to every other player as well. Chiu et al. [3] in their seminal paper proved that AIAD and MIMD do not converge. AIMD they proved, is the best choice for TCP congestion avoidance. AIMD in their case corresponds to MIAD in our case because the factor α is detrimental when increased and beneficial when decreased. MIAD is therefore, the best choice for increment-decrement schemes. This fact is experimentally verified as shown below.

The graph of utility of every player at their Nash equilibrium against the total number of games played considering the following values of each parameter is shown in Fig. 2a, b, c.

$$v = 1.5, c_E = 1, c_{DU} = 2, \mu = 0.5, \beta = 0.5, \alpha = 0.01$$

For AIAD, IF = 0.02 and DF = 0.01

For MIAD, IF = 1.5 and DF = 0.01

For MIMD, IF = 1.5 and DF = 1.2

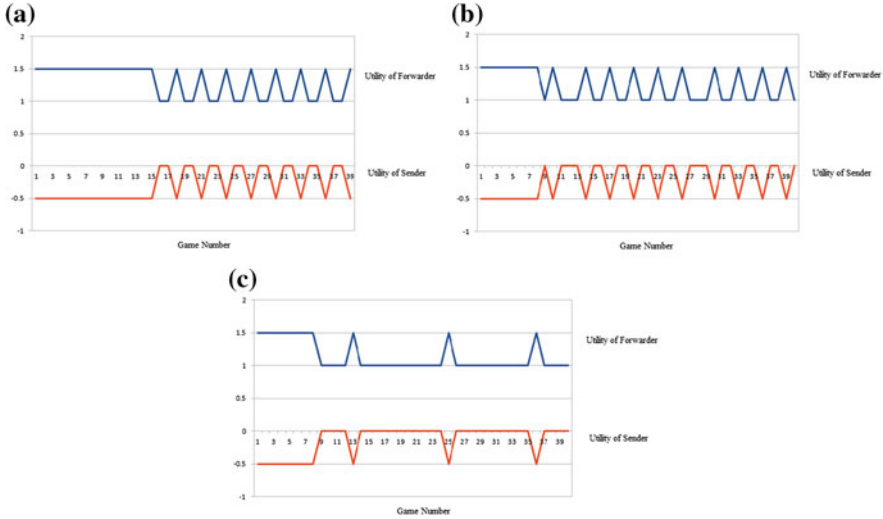


Fig. 2 Graph of utility versus game number for forwarder and sender. **a** AIAD strategy, **b** MIMD strategy, **c** MIAD strategy

The utility value of the forwarder (malicious) node reaches peak value when the packet is dropped. Number of packet drops can therefore be identified as the number of peaks in the forwarders graph. As can be seen, MIAD produces the least number of packet drops over repeated plays of the game. It is therefore experimentally verified that the α value update should follow the MIAD strategy.

5.2 c_N Value Determination

The value of c_N represents the benefit that the malicious node receives from forwarding the packet and faking normal behavior. The benefit is based on the change in α at the sender node. If the value of v is greater than $(1 + \beta * c_{DU}) / (2 * \mu + \alpha)$ computed using the updated alpha value, then $c_N = c_E$. This is because, the best response for the sender becomes not sending a packet through the forwarder and hence the forwarder is exposed. In all other cases, $c_N = 0$.

6 Summary

This paper presents a game theoretic approach to the thwarting of gray hole attacks in WANETs. Unlike other approaches, there is no additional network level overhead as every node has only to follow the suggested strategies to counter the attack. An

extensive form game model is used for the game definition and interim equilibrium analysis is done after converting it into a Bayesian normal form game. Repeated play of the game is simulated using the parameters α and c_N . α values in the repeated play of the game are updated using the MIAD strategy which is experimentally verified to be the best. The sender thus succeeds in thwarting the gray hole attack if he follows the suggested strategy under each condition.

7 Conclusion and Future Work

The game theoretic approach described in this paper succeeds in thwarting the gray hole attack under empirically defined parameter values. This approach is widely applicable to WANETs, uses the stronger two-player incomplete information extensive form game notion and induces no additional network level overhead. It is therefore superior to related work in the area. In subsequent work, we would like to inform parameter values over subsequent plays of the game by learning over past network metrics such as congestion and packet drop rate.

References

1. P. Agrawal, R. K. Ghosh, and S. K. Das. Cooperative black and gray hole attacks in mobile ad hoc networks. In *Proceedings of the 2nd international conference on Ubiquitous information management and communication*, pages 310–314. ACM, 2008.
2. M. Al-Shurman, S.-M. Yoo, and S. Park. Black hole attack in mobile ad hoc networks. In *Proceedings of the 42nd annual Southeast regional conference*, pages 96–97. ACM, 2004.
3. D.-M. Chiu and R. Jain. Analysis of the increase and decrease algorithms for congestion avoidance in computer networks. *Computer Networks and ISDN systems*, 17(1):1–14, 1989.
4. A. Desai. Review paper on detection and prevention techniques of gray-hole attack in manet. *International Journal of Computer Science and Mobile Computing*, 2:105–108, 2013.
5. A. M. Kanthe, D. Simunic, and R. Prasad. A mechanism for gray hole attack detection in mobile ad-hoc networks.
6. S. Kaplantzis, A. Shilton, N. Mani, and Y. Sekercioglu. Detecting selective forwarding attacks in wireless sensor networks using support vector machines. In *Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on*, pages 335–340, Dec 2007.
7. C. Karlof and D. Wagner. Secure routing in wireless sensor networks: Attacks and countermeasures. *Ad hoc networks*, 1(2):293–315, 2003.
8. W. Z. Khan, Y. Xiang, M. Y. Aalsalem, and Q. Arshad. The selective forwarding attack in sensor networks: Detections and countermeasures. *International Journal of Wireless and Microwave Technologies (IJWMT)*, 2(2):33, 2012.
9. X. Liao, D. Hao, and K. Sakurai. Classification on attacks in wireless ad hoc networks: A game theoretic view. In *Networked Computing and Advanced Information Management (NCM), 2011 7th International Conference on*, pages 144–149, June 2011.
10. Y. Liu, C. Comaniciu, and H. Man. A bayesian game approach for intrusion detection in wireless ad hoc networks. In *Proceeding from the 2006 workshop on Game theory for communications and networks*, page 4. ACM, 2006.

11. M. Mahmoud and X. Shen. An integrated stimulation and punishment mechanism for thwarting packet dropping attack in multihop wireless networks. *Vehicular Technology, IEEE Transactions on*, 60(8):3947–3962, Oct 2011.
12. M. H. Manshaei, Q. Zhu, T. Alpcan, T. Başçar, and J.-P. Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys (CSUR)*, 45(3):25, 2013.
13. Y. B. Reddy and S. Srivathsan. Game theory model for selective forward attacks in wireless sensor networks. In *Control and Automation, 2009. MED'09. 17th Mediterranean Conference on*, pages 458–463. IEEE, 2009.
14. D. M. Shila and T. Anjali. A game theoretic approach to gray hole attacks in wireless mesh networks. In *Military Communications Conference, 2008. MILCOM 2008. IEEE*, pages 1–7. IEEE, 2008.
15. A. Taggu and A. Taggu. Tracegray: An application-layer scheme for intrusion detection in manet using mobile agents. In *Communication Systems and Networks (COMSNETS), 2011 Third International Conference on*, pages 1–4, Jan 2011.
16. B. Wu, J. Chen, J. Wu, and M. Cardei. A survey of attacks and countermeasures in mobile ad hoc networks. In Y. Xiao, X. Shen, and D.-Z. Du, editors, *Wireless Network Security, Signals and Communication Technology*, pages 103–135. Springer US, 2007.
17. B. Xiao, B. Yu, and C. Gao. Chemas: Identify suspect nodes in selective forwarding attacks. *Journal of Parallel and Distributed Computing*, 67(11):1218–1230, 2007.
18. G. Xiaopeng and C. Wei. A novel gray hole attack detection scheme for mobile ad-hoc networks. In *Network and Parallel Computing Workshops, 2007. NPC Workshops. IFIP International Conference on*, pages 209–214. IEEE, 2007.
19. B. Yu and B. Xiao. Detecting selective forwarding attacks in wireless sensor networks. In *Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International*, pages 8 pp.–, April 2006.
20. X. Zhang, S. F. Wu, Z. Fu, and T.-L. Wu. Malicious packet dropping: how it might impact the tcp performance and how we can detect it. In *Network Protocols, 2000. Proceedings. 2000 International Conference on*, pages 263–272. IEEE, 2000.

Chi-Square Based Mobile Radio Propagation Model Analysis and Validation

Lavanya Vadda, G. Sasibhushana Rao and L. Ganesh

Abstract In urban and semi-urban areas, the ever increasing population is creating high raised structures and increasing tele-density. It is becoming difficult for the mobile network providers to offer quality service to the mobile user. One of the main reasons causing degradation in signal quality is multipath propagation. Because of this, the Received Signal Strength (RSS) may be either reduced or completely attenuated at the receiver. So modelling and characterisation of the channel is necessary. If there is no line-of-sight signal component from transmitting station to the receiver, then the envelop of the received signal can be statistically described by Rayleigh distribution. In this paper, real time mobile RSS in terms of power (in dBm) is recorded, analysed and its quality is tested using theoretical Rayleigh distribution and also validated using Chi-square fitness-of good test.

Keywords Multipath · Received signal strength · Rayleigh distribution · Chi-square goodness of fit test

1 Introduction

The radio signal travelling through the channel (communication path) can propagate in all possible paths and it can be affected by reflection, refraction, Doppler spread etc. The number of available propagation paths of the signal increases the complexity of the channel [1, 2].

A mobile signal travels from a transmitting station to the mobile station over a number of reflective paths and is known as multipath propagation in wireless communication system. The deviation of received signal strength due to variations

L. Vadda (✉) · G. Sasibhushana Rao
Department of ECE, Andhra University, Visakhapatnam, Andhra Pradesh, India
e-mail: lavanyavadda@gmail.com

L. Ganesh
Department of ECE, ANITS College of Engineering,
Visakhapatnam, Andhra Pradesh, India

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_22

227

in channel is called as fading, and multipath fading indicates changes in the received signal's characteristics because of multipath propagation. The error rate of the received data increases because of multipath fading [3].

In wireless communication, two types of fading may predominantly affect the mobile users. They are large scale fading and small scale fading. The fading rapidly changes according to the changes in distance between the transmitting station and receiving station.

In a multipath environment, the variations of a signal can be analysed by Rayleigh or Rician distribution [4]. When the Line Of Sight (LOS) between the transmitting and the receiving stations is blocked, then the signal variations are statistically analysed by Rayleigh distribution [5]. The Rayleigh probability distribution function is given by Eq. (1).

$$f(x) = \begin{cases} \frac{x}{\sigma^2} \exp(-\frac{x^2}{2\sigma^2}) & \text{for } x \geq 0 \\ = 0 & \text{otherwise} \end{cases} \quad (1)$$

where x is a voltage and σ is scale parameter.

The main intention of this paper is to characterize the real time mobile phone signal data and then validating the recorded data with the theoretical Rayleigh distribution using Chi-Square goodness of fit test.

2 Analysis of Mobile Radio Propagating Signal

The following steps are followed during the analysis of data:

1. The mobile radio propagating signal data is recorded under homogeneous multipath conditions for a period of 1 h 25 min and the same is analysed. The recorded data illustrates the signal variation at certain time intervals in multipath environment as shown in Fig. 1.
2. Received power is transformed from dBm to dB and then to model the voltage, a load resistance, R , of 50Ω is considered. Voltage values for the recorded data are calculated by using the following equation

$$v = \sqrt{2RP} \quad (2)$$

Recorded data is complex to analyse and is normalized by considering one of its parameters. The expression for the Probability Distribution Function (PDF), as a function of the mode is considered. Here for Rayleigh distribution, the mode is predicted from mean using the equality $\sigma\sqrt{\pi/2} = 1.25\sigma$.

Thus Fig. 2 shows the recorded data of voltage which is normalized with its predicted modal value $\sigma = 1$.

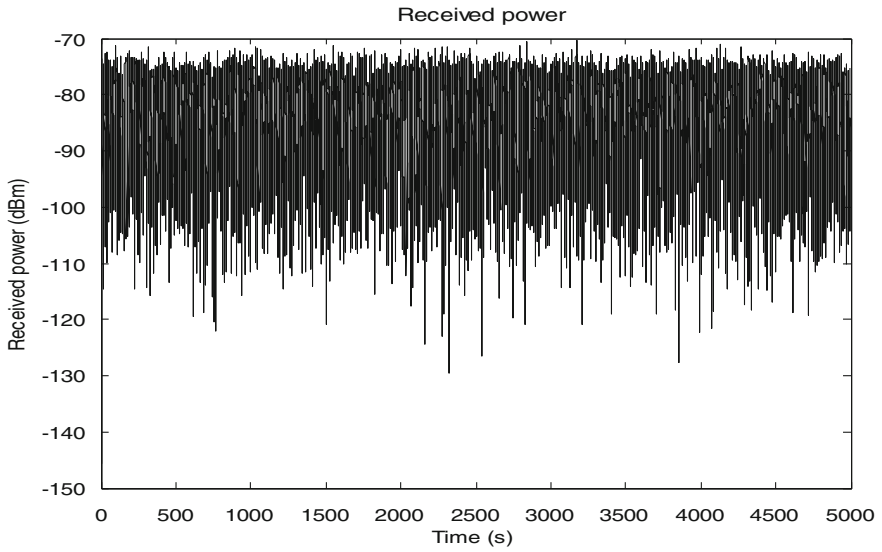


Fig. 1 Power received (dBm) versus time(s)

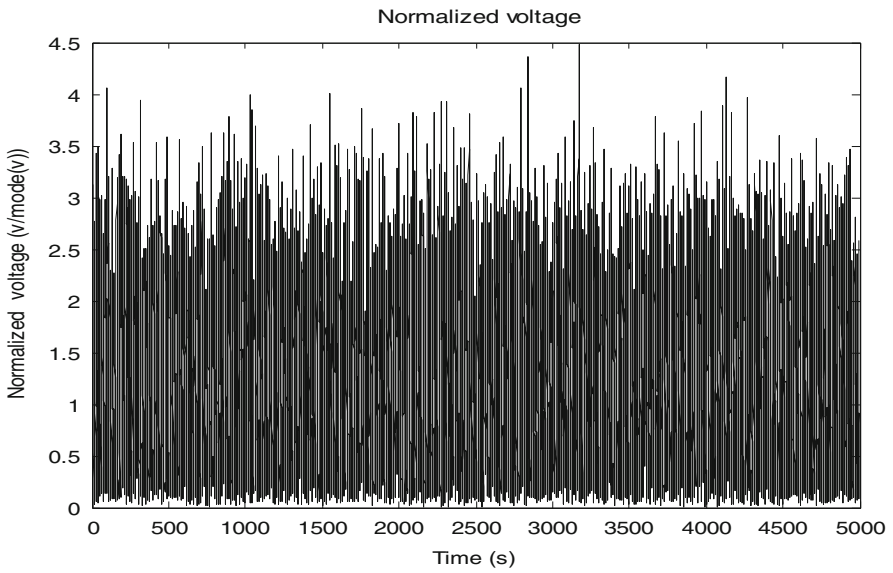


Fig. 2 Normalized voltage versus time

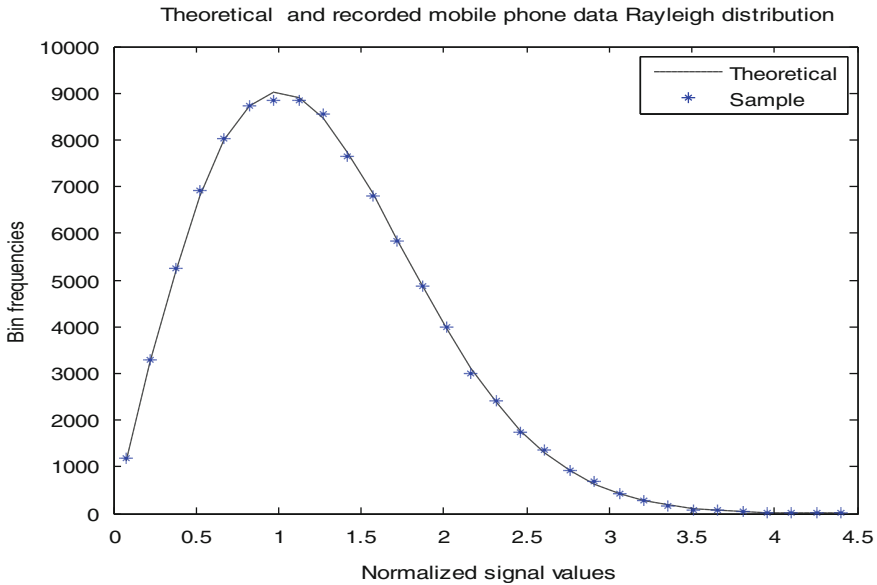


Fig. 3 Theoretical and recorded mobile phone data Rayleigh PDF's

3. Comparison of recorded mobile data PDF with theoretical Rayleigh PDF: The theoretical Rayleigh PDF and PDF for the observed data is plotted using Eq. (1) and shown in Fig. 3. The match between the observed and theoretical values is reasonably good. Figure 4 shows the histogram plot between the theoretical and observed data using the following equation

$$\int_{X_1}^{X_2} f(x)dx = \exp\left(-\frac{X_2^2}{2\sigma^2}\right) - \exp\left(-\frac{X_1^2}{2\sigma^2}\right) \tag{3}$$

where X1 and X2 are the bin limits.

4. The bin frequencies and centers are provided by the histogram of the measured data. These frequencies are converted into probabilities by dividing them by the total number of samples.
5. Validation of the recorded mobile data with the theoretical Rayleigh distribution using Chi-square goodness of fit test.

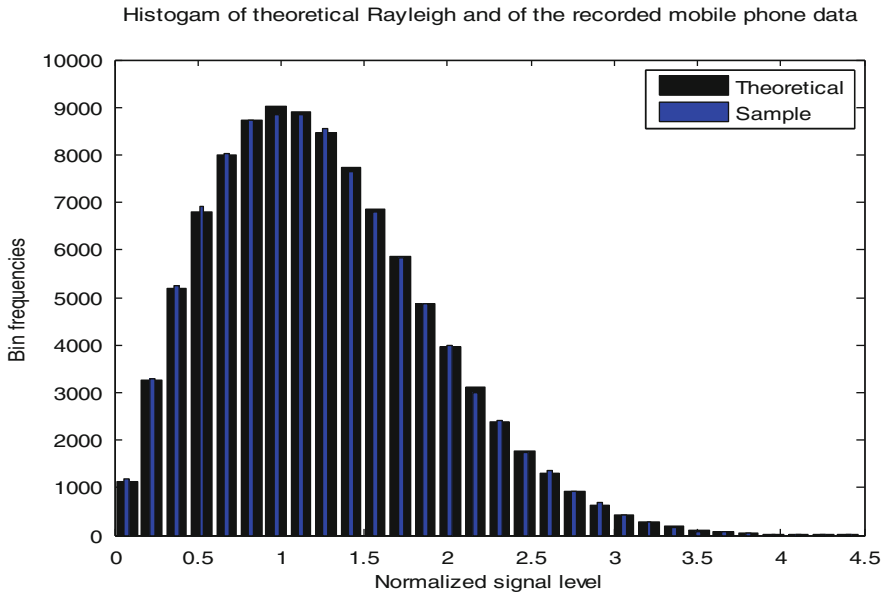


Fig. 4 Rayleigh histogram of the theoretical and measured data

3 Validation of Observed Values with True Values Using Chi-Square Test

From the visual comparison, it is evident that measured data follows theoretical PDF. But by using Chi-square goodness of fit test, we can quantify how good the fit is, between the theoretical and observed data [6].

To determine goodness of fit between the theoretical (expected values) and the observed values in one or more cases, chi-square test is used. The term ‘Chi-square’ is used because the Greek letter χ is used to define this distribution. The elements on which this distribution is based are squared, so that the symbol χ^2 is used to denote the distribution [6].

In the Chi-square test, the difference between the values observed experimentally and the values that would be expected are measured. Then, this measure has to be compared with critical values which are tabulated for significance level and degree of freedom of the Chi-square test. Generally this significance level is set to 1 % or 5 %. A significance level of 5 % means that it is the point at which there is 95 % confidence that the difference is NOT due to chance alone.

Degree of freedom is denoted by ‘v’. For any χ^2 distribution, a number of independent free choices can be made in allocating values to the expected frequencies. The degrees of freedom are defined as $M - 1$, the total number in the group minus one restriction. The critical χ^2 value can be read directly from the table. A critical value is the value corresponding to a degree of freedom for a

given significance level. First is to determine degrees of freedom and locate the value in the appropriate column. Then, locate the critical value for corresponding degrees of freedom and significance level. To read the chi-square table, we need degree of freedom as well as the significance level of the test.

The steps in using the chi-square test may be summarized as follows:

- I. The sample space S_x , is partitioned into a union of Y disjoint bins.
- II. In ‘ n ’ repetitions of the experiment, the expected number of outcomes that fall in the Y th interval is given by $E_Y = nb_Y$ where the probability that a outcome is in the Y th interval is indicated by b_Y .
- III. Chi-square measure and is denoted by χ^2 , it is the weighted difference between the observed number of outcomes, O_Y , and the expected number, E_M that fall in the Y th interval

$$\chi^2 = \sum_{n=1}^Y \frac{(O_n - E_n)^2}{E_n} \tag{3}$$

- IV. If $\chi^2 \geq t_\alpha$, then the fit is good otherwise the hypothesis is rejected. Where t_α is the critical value for significance level α and for degree of freedom.

The test is performed by splitting the recoded data into intervals or bins of equal length. The results for the theoretical and obtained for 30 bins are tabulated in Table 1, and in Table 2, the critical values for different degrees of freedom and significant levels are tabulated.

Table 1 Chi-square measure table for 30-bins

Bin Min.	Bin Max.	Measured freq (O)	Theo-freq (E)	(O-E) ² /E	Bin Min.	Bin Max.	Measured freq (O)	Theo-freq (E)	(O-E) ² /E
0	0.26	3442.0	3333.3	3.54	1.18	1.23	3470.0	3333.3	5.60
0.26	0.37	3364.0	3333.3	0.28	1.23	1.29	3397.0	3333.3	1.22
0.37	0.46	3353.0	3333.3	0.12	1.29	1.35	3285.0	3333.3	0.70
0.46	0.53	3496.0	3333.3	7.94	1.35	1.42	3261.0	3333.3	1.57
0.53	0.63	3311.0	3333.3	0.15	1.42	1.48	3308.0	3333.3	0.19
0.63	0.67	3372.0	3333.3	0.45	1.48	1.55	3288.0	3333.3	0.62
0.67	0.73	3325.0	3333.3	0.02	1.55	1.63	3356.0	3333.3	0.15
0.73	0.79	3370.0	3333.3	0.40	1.63	1.71	3263.0	3333.3	1.48
0.79	0.84	3307.0	3333.3	0.21	1.71	1.79	3351.0	3333.3	0.09
0.84	0.90	3313.0	3333.3	0.12	1.79	1.89	3316.0	3333.3	0.09
0.90	0.96	3245.0	3333.3	2.34	1.89	2.01	3365.0	3333.3	0.30
0.96	1.01	3224.0	3333.3	3.59	2.01	2.15	3264.0	3333.3	1.44
1.01	1.07	3313.0	3333.3	0.12	2.15	2.33	3329.0	3333.3	0.01
1.07	1.12	3296.0	3333.3	0.42	2.33	2.61	3351.0	3333.3	0.09
1.12	1.18	3311.0	3333.3	0.15	2.61	99.0	3354.0	3333.3	0.13

$\chi^2 = 33.5432$

Table 2 Critical values for significance levels 1 and 5 % for different degrees of freedom

DOF	5 %	1 %	DOF	5 %	1 %	DOF	5 %	1 %
1	3.84	6.63	8	15.51	20.09	22	33.92	40.29
2	5.99	9.21	9	16.92	21.67	24	36.42	42.98
3	7.81	11.35	10	18.31	23.21	26	38.89	45.64
4	9.49	13.28	12	21.03	26.22	28	41.34	48.28
5	11.07	15.09	14	23.68	29.14	30	43.77	50.89
6	12.59	16.81	16	26.30	32.00	32	46.19	53.48
7	14.07	18.48	18	28.87	34.80	34	48.60	56.06

Here the degree of freedom is the number of bins minus number of parameters (mean, mode) extracted from the data i.e., $Y - m - 1$.

The obtained χ^2 value from Eq. (3) should always be less than the tabulated critical value. The percentage of fitness is determined by the significance level. The χ^2 value after calculation is 33.57 and this value is compared with the critical value for different degrees of freedom and significance levels. For degree of freedom 'v' = 28 and for 1 % significance the critical value is 48.28 and for 5 % significance it is 42.34. In both the cases, threshold value exceeds the obtained value and hence, the test is passed.

4 Conclusions

In this paper, the PDF of recorded real-time data is compared with the theoretical Rayleigh distribution function and is validated with the Chi-square goodness of fit test. Fading pattern for real-time mobile phone data is characterised for a typical urban environment. Chi-Square goodness of fit test is used to validate the percentage of fitness between the theoretical and measured mobile phone data. The obtained goodness of fit parameter is compared with the predefined critical value with certain threshold level (α), i.e. corresponding to particular degrees of freedom. Finally it can be concluded that the recorded mobile phone data follows the Rayleigh distribution.

References

1. Liberti, J.C., and T.S. Rappaport, Smart Antennas for Wireless Communications: IS-95 and third Generation CDMA Applications, Prentice Hall, New York (1999).
2. Bertoni, H.L., Radio propagation for Modern Wireless Systems, Prentice Hall, New York (2000).
3. Nikolay Kostov, "Mobile Radio Channels Modeling in MATLAB" (2003).

4. Sklar, B. Rayleigh fading channels in mobile digital communications, part I: Characterization. IEEE Communications Magazine (1997) Vol.35, no. 7, p. 90–100.
5. Gottapu Sasibhushana Rao, “Mobile cellular communication”, New Delhi, Pearson Education (2013).
6. Chi- square test <http://www2.lv.psu.edu/jxm57/irp/chisquar.html>.

Collision Theory Based Sentiment Detection of Twitter Using Discourse Relations

Anuta Mukherjee and Saswati Mukherjee

Abstract Social networking sites such as Twitter are contributing to a large increase in the growth of data today, and are a rich source for sentiment detection or mining. This research employs collision theory to achieve a query based sentiment detection of Twitter data with discourse analysis. Hadoop has been exploited for speed. Our experiments show effective results.

Keywords Hadoop • Sentiment analysis • Collision theory • Discourse relations

1 Introduction

Twitter is a micro-blogging website that allows users to post short messages, called tweets. The popularity of this website and its continuous use en mass ensures that any information garnered from this site is directly coming from the users. While research on Twitter data is attractive, researchers are faced with the challenge of Big Data.

Perhaps the largest use of Twitter data is found in opinion mining or sentiment detection [1]. The availability of opinions of millions of users can be exploited for online applications to provide quick replies to all possible queries spanning all domains using online data. However for this we need a domain independent sentiment detection method suitable for online applications.

A. Mukherjee (✉)

Department of Computer Science and Engineering, CEG, Anna University,
Chennai 600025, India
e-mail: manuta94@gmail.com

S. Mukherjee

Department of Information Science and Technology, CEG, Anna University,
Chennai 600025, India
e-mail: msaswati@auist.net

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_23

235

In this paper, we propose a fast, online, domain independent Twitter sentiment detection using Hadoop [2], Apache's MapReduce framework, as an end-to-end solution. We follow a lexicon-based approach, and apply a modified version of collision theory [3] combined with discourse analysis [4] to re-assign polarity values to the sentiment-bearing terms in the proposed sentiment analysis mechanism. We propose to use a weighted approach, exploiting retweet feature of Twitter data.

In this paper our contributions are (1) Use of Hadoop as a filter to preprocess Twitter data for domain independent Twitter sentiment analysis, (2) Inclusion of discourse analysis in modified collision theory and (3) Use of retweet feature for sentiment analysis.

The rest of the paper is organized as follows: Sect. 2 presents a short related work. Section 3 presents a comprehensive view of the proposed mechanism. The validation of the proposed method is presented in Sect. 4. Section 5 concludes the paper.

2 Related Work

In lexicon based Twitter sentiment analysis, the polarity of a sentence is the summation of the polarity values of every sentiment bearing individual word in the text [5]. The central requirement of this approach is a dictionary of terms indicating the polarity. A pioneering lexicon based sentiment detection from tweets has been proposed by Taboada et al. [1].

Unfortunately, most of the sentiment analysis techniques for micro blogs do not scale well for big data. To this end some researchers combined machine learning and lexicon based analysis for Twitter data [6, 7] while other researchers applied big data based solutions like Hadoop for the purpose of Twitter sentiment analysis [8].

Use of Hadoop for sentiment analysis has the inherent need of a cluster having a reasonably large number of nodes and this restricts the use of such mechanisms for online applications. In the proposed method Hadoop is applied with this constraint in mind.

3 Proposed Work

In this research, we propose a two-phase approach; the Preprocessing phase and the Analysis phase. A query is given from which sentiment is to be extracted. We expand noun terms of the query to obtain the feature terms (FT) and use these for Hadoop based extraction of relevant tweets in the Preprocessing phase. In this phase, we use only the Map function of Hadoop to ensure reduced number of nodes, reduced processing time and enhanced speed. This allows a domain independent sentiment analysis using any dump of Twitter data. In the Analysis phase, actual sentiment analysis of each tweet is carried out using the proposed lexicon-based discourse oriented collision theory. This too has been implemented in

a Hadoop cluster to improve speed and accommodate large Twitter data. In the proposed method two important items, viz., the polarity term (PT) and the discourse terms (DT) are used.

In lexicon-based sentiment detection, the sentiment orientation of a tweet must be calculated by aggregating the semantic orientation of terms or phrases, mainly adjectives and adverbs, called the polarity terms (PT). In the proposed method Bing Liu sentiment lexicon containing around 6800 PT terms is used [9]. We obtain the initial semantic orientation (SO) values of all PT terms in the extracted tweets $Value_{PT}$, modifying it with the distance factor between PT and any FT terms that are found in the sentence as shown in Eq. (1).

$$Value_{PT} = P_Val_{PT} + \frac{P_Val_{PT}}{2^{*(Distance|feature - PT|)}} \quad (1)$$

where P_Val_{PT} is the signed value of polarity term PT in the Bing Liu list and $Distance_{feature_i - PT}$ is the number of terms between the nearest i th FT 'feature' found in the sentence of the tweet and the polarity term PT.

We propose a modified collision theory [3] approach to detect sentiment of tweets using transition of PT terms. We traverse a tweet from left to right looking for the presence of all PT terms and after the first one, each encounter of a PT term is designated as a collision. Each collision requires an adjustment of the SO value of the second PT term using one of the two types of transitions. The first transition is called the Positive-Follow transition as shown in (2) below. In this transition an occurrence of a positive PT term influences the polarity of the next PT term in the absence of any influencing intermediate DT term.

$$Value_{PT} = Value_{PT} + \frac{\sqrt{|P_Val_{PT_{prev}}|}}{Distance|feature - PT|} \quad (2)$$

The second transition called the Negative-Follow transition is shown in Eq. (3). In this transition an occurrence of a negative PT term influences the polarity of the following PT term in the absence of any influencing intermediate DT term.

$$Value_{PT} = Value_{PT} - \frac{\sqrt{(|P_Val_{PT_{prev}}|)}}{Distance|feature - PT|} \quad (3)$$

In a tweet, there may be multiple transitions and each of these transitions affect the SO value of the appropriate PT term.

Discourse relations have been harnessed in [4] using certain semantic operators found in the text. Table 1 below shows various discourse relations and the corresponding semantic operators identified to be important.

We further claim that in a sentence of a tweet, the influence of the Discourse Terms or DT terms on the PT terms have to be defined with respect to the strength of the polarity value of the PT term in question and hence an appropriate transition has to be defined for this purpose. In the proposed work, we have incorporated all

Table 1 Discourse relations and semantic operators

Relations	Semantic operators
Conj_Fol	But, however, yet, still, nevertheless, otherwise nonetheless
Conj_Prev	Till, until, despite, in spite, though, although
Conj_Infer	Therefore, furthermore, consequently, thus, as a result, eventually, hence
Modal	Might, could, can, would, may, should, ought to, need not, must
Conditional	If
Negation	Not, neither, never, no, nor

the semantic operators of Table 1 in the transitions of collision theory based sentiment analysis. This gives our collision based sentiment detection model an improved accuracy. To this end, we define our third transition called the PT_Previous transition as shown in Eq. (4). If during traversal from left to right, a DT belonging to Conj_prev is encountered, this transition is used to calculate the value of the current PT term.

$$Value_{PT_{Prev}} = Value_{PT_{Prev}} + \sqrt{(Value_{PT})} \quad (4)$$

The fourth transition, expressed in Eq. (5), is called the DT_Previous transition. If during traversal from left to right, a DT belonging to Modals or Conditionals is encountered, this transition is used to calculate the value of the current PT term.

$$Value_{PT} = \frac{1}{2} Value_{PT} \quad (5)$$

For negation, we've flipped the value of the PT term.

The sentiment value for each of the tweets is the summation of the calculated polarity values of all the PT-s present in the tweet. A tweet has a field labeled 'retweet' that indicates how many times a tweet was resent by users. This makes certain tweets more important than other and hence we propose to use this feature to decide the weightage of a certain tweet in the overall sentiment detection of a certain topic. We finally apply the retweet-based weighting for each of the tweet sentiment values and obtain the overall sentiment for the given query.

The final overall value of the query is the summation of the sentiment values of all the tweets extracted by the preprocessing system.

4 Experimental Section

To test the efficacy of Hadoop to run as a filter to Twitter data and to implement sentiment analysis, we have used a YARN version 2.4.1. To compare the performance of Hadoop-based sentiment analysis, we have run our experiments on a

normal Java platform in a single node Hadoop cluster and in a 4-node Hadoop cluster. In the 4-node cluster, one node is used as a Resource Manager and three nodes are used as Node Managers.

We have used two data sets to check the efficiency and accuracy of our approach. The first data set constitutes of tweets downloaded using Twitter API. We obtained 3000 tweets related to five keywords, viz., *Cricket*, *Narendra Modi*, *Earthquake*, *Facebook* and *Instagram*. Tweets from each group were manually classified as positive, negative, neutral or uncertain towards the specific set of expanded keywords. After removing duplicate and uncertain tweets from each group, we obtained five test sets of tweets. This data set contains 930 positive tweets, 468 negative tweets and 1,212 neutral tweets.

We used Sanders data set [10] as our second data set. There are 5513 tweets for four different categories, Apple, Google, Microsoft and Twitter.

4.1 Evaluation

We have a three-part evaluation: the first part is Map-only Hadoop as a filter, the second part is the sentiment analysis with respect to a query and the third part is the efficacy of running the sentiment analysis in standard Hadoop cluster environment.

To evaluate the Hadoop as a filter, we manually created different queries that are appropriate to each keyword for each of the two data set. The queries are POS tagged individually to identify nouns and are then semantically expanded using WordNet to obtain sets of expanded feature terms or FT. We then ran each data set using only the Map function in Hadoop to obtain separate groups of tweets. For data set 1, we compared the per group tweets extracted by Hadoop with the ones obtained using Twitter API for each keyword. Similarly, for the second data set, the extracted groups are compared with the four different groups of tweets provided in [10].

Here, Hadoop is used as a two-class classifier. Since recall is widely used for evaluating the performance of unstructured as well as semi-structured document categorization, we have adopted recall here for comparing the performance of Hadoop as a filter. The result is shown in Fig. 1.

Fig. 1 Recall of Hadoop filter

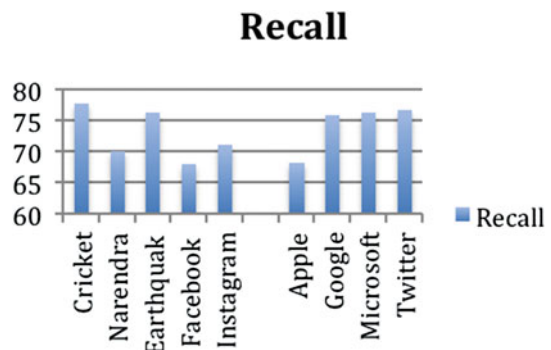
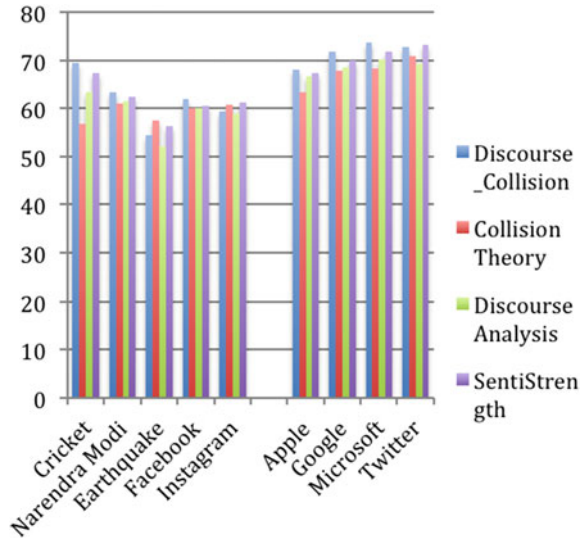


Fig. 2 Comparison of accuracy



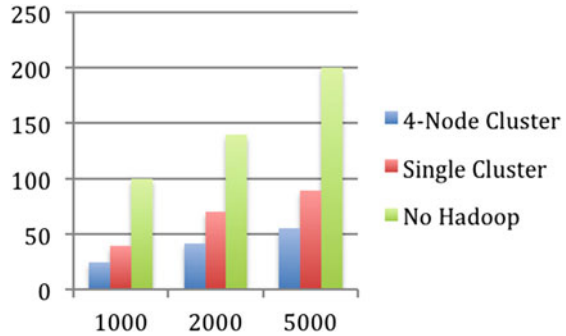
In the above figure, the worst recall values are for Facebook and Apple at 67.9 and 68.1 respectively, indicating that there were at least 67.9 % of correct tweets obtained among all the tweets extracted by Hadoop for the specific keyword.

The second part of our evaluation is for the sentiment analysis. We used the tweets extracted in the first part of evaluation in the second part. To compare performance of our analysis system, the accuracy of the proposed algorithm (Discourse_Collision) is compared with three algorithms, viz., *Collision Theory* [3], lexicon based *Discourse Analysis* [4] and *SentiStrength* [1] as shown in Fig. 2.

As seen above, our proposed method gives vastly better results than both the Collision theory and the Discourse Analysis methods and outperforms SentiStrength for many keywords in both the data sets.

We test the third part of our system, the efficacy of Hadoop, by running Discourse_Collision experiments in three different setups, viz., non-Hadoop environment, single-node Hadoop environment and a 4-node Hadoop cluster environment. We used the two data sets from earlier experiments and created three sets of data with 1000 tweets, 2000 tweets and 5000 tweets. Proposed sentiment analysis is run in the three setups for the all the sizes of data and the time spent is observed, as shown in Fig. 3.

The result reinforces the need of using Hadoop cluster for sentiment analysis of Twitter data.

Fig. 3 Efficiency of Hadoop

5 Conclusion

This paper proposes a collision theory based discourse analysis mechanism for sentiment analysis of Twitter data using Hadoop. We have compared the result with three existing related methods that show the efficacy of the proposed method. While there are a multitude of research using machine learning methods, lexicon based mechanism brings in advantages to Twitter sentiment analysis that are important to the research and cannot be ignored. This research explores and includes such a method, namely collision theory in Twitter sentiment analysis. As future work, we would like to explore the possibilities of combining machine learning methods with the proposed collision based discourse analysis to improve the results further.

References

1. Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M.: Lexicon-based Methods for Sentiment Analysis. In: Computational Linguistics 37, no 2, (2011), 267–307.
2. Apache Hadoop. Available at <http://hadoop.apache.org>.
3. M. S. Murugesan, S. Mukherjee: Novel Relevance Model for Sentiment Classification based on Collision Theory. In: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (LNICST), Springer, (2012).
4. Mukherjee, Subhabrata, and Pushpak Bhattacharyya: Sentiment Analysis in Twitter with Lightweight Discourse Analysis. In: COLING. (2012).
5. Ding, X., Liu, B. and Yu, P. S.: A Holistic Lexicon-based Approach to Opinion Mining. In: WSDM, (2008).
6. Zhang, L., Ghosh, R., Dekhil, M., Hsu, M. and Liu, B.: Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis. Technical report, HP Laboratories, (2011).
7. J. Y. Chang: Automatic Retrieval of SNS Opinion Document Using Machine Learning Technique. The Journal of The Institute of Internet, Broadcasting and Communication, vol. 13, no. 5, article 27, (2013).

8. V. N. Khuc, C. Shivade, R. Ramnath, and J. Ramanathan: Towards Building Large-Scale Distributed Systems for Twitter Sentiment Analysis. In: 27th Annual ACM Symposium on Applied Computing (SAC '12), pp. 459–464, March (2012).
9. Minqing Hu and Bing Liu: Mining and Summarizing Customer Reviews. In: ACM SIGKDD, (2004).
10. Sanders, Niek J.: Sanders-Twitter Sentiment Corpus. In: Sanders Analytics LLC, (2011).

Malicious Account Detection Based on Short URLs in Twitter

Rasula Venkatesh, Jitendra Kumar Rout and S.K. Jena

Abstract The popularity of Social Networks during the last several years have attracted attention of cyber-criminals for spreading of spam and malicious contents. In order to send spam messages to lured users, spammers creating fake profiles, leading to fraud and/or malware campaigns. Sometimes to send malicious messages, cyber-criminals use stolen accounts of legitimate users. Nowadays they are creating short URLs by the short URL service provider and post it on friend's board. Lured users unknowingly clicking on these links, are redirected to malicious websites. To control such type of activities over Twitter we have calculated a trust score for each user. Based on the trust score, one can decide whether a user is trustable or not. With usage of trust score, we have achieved accuracy of 92.6% and F-measure of 81% with our proposed approach.

Keywords Short URLs · Cyber crime · Spam messages · Trust score

1 Introduction

Social networking is a platform which provides to build a social relationship among people using the Internet. Over recent years, social networks are largest and fastest growing networks. Out of hundreds of online social networks present today, only few like Facebook, Twitter, LinkedIn etc. are the most popular based on the number of active users. In these networks the users are sharing their personal information. These sites can be used by the government to get opinion of public quickly. On Twitter users are communicating through tweets. Twitter is playing a crucial role for connecting

R. Venkatesh · J.K. Rout (✉) · S.K. Jena
National Institute of Technology, Rourkela, Sundargarh 769008, Odisha, India
e-mail: jitu2rout@gmail.com

R. Venkatesh
e-mail: Venky1233@gmail.com

S.K. Jena
e-mail: skjena@nitrkl.ac.in

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_24

243

people, and people can discuss on a particular topic like earthquake in Nepal. In Twitter, the user can send a message maximum up to 140 characters only. Twitter allows only unidirectional relationship among the users. User can add tags to the tweets (i.e. #tags) which provides combination of all the related information.

Twitter has a concept of following. Suppose, if a user A follows a user B signifies that all tweet posted by B would be posted on timeline of A. But user B cannot see the tweets posted by the user A. By this we can specify that whose tweets the user having an interest to see. These user could be friends, co-workers, celebrities, researchers etc. Twitter acting as news social media for spreading the breaking information over the globe. Twitter has trending topics on the left side of the user timeline. Trending topics contain top 10 hot topics to discuss. In order post a tweet related to trending topic user must include # followed by topic name. With millions of tweets generated per day, there is also an increasing concerns about the trustworthiness of information disseminated throughout the social networks and the privacy breaching threats of participant's private information.

Few years ago, the users are limited to viewing of information on the websites. Now online social networks are providing a platform for the users to actively participate over the websites. At the same time there are cyber-criminals attacks like stealing credentials, fake messages etc. Cybercrimes are serious threat for Internet users. Twitter is the one of social network attracted by the most of the malicious users. They are providing malicious links and fake information for advertising purpose or get the money from the lure users. Twitter having limitation that we can on send 140 characters, the user can not send whole URL in a tweet. There are some of the URL shortening service provider (goo.gl, bit.ly, t.co) present for shortening the long URL to short URL. Spammers are masquerading the actual URLs, i.e. user doesn't know the actual link behind the short URL.

In this work, we have focused on "trust score" of a user. In social network (like Twitter) user can participate in several social activities. How much trustable a person is in social networks? Based on the trust score, the user can decide tweets posted by the particular user is trustable or not. If the user is having higher trust score [6], the information posted by user is legitimate content. Lesser the trust score, the information posted by him is more vulnerable, i.e. containing malicious information. The trust score is numerical score with in the range of 0 to 1. For calculating trust we are considered many parameters like user activities, social connection, user profiles etc. In the past, several machine learning algorithms have been employed to extract and classify social network users, still most of them failed to accurately classify malicious users.

The remainder of the paper is organized as follows: in Sect. 2 some existing work on malicious account detection are discussed Sect. 3 deals with methodology for the data collection Sects. 4 and 5 describes the proposed algorithm and machine learning classification techniques respectively. Results and analysis discussed in Sect. 6. Section 7, concludes this paper and suggests future scope.

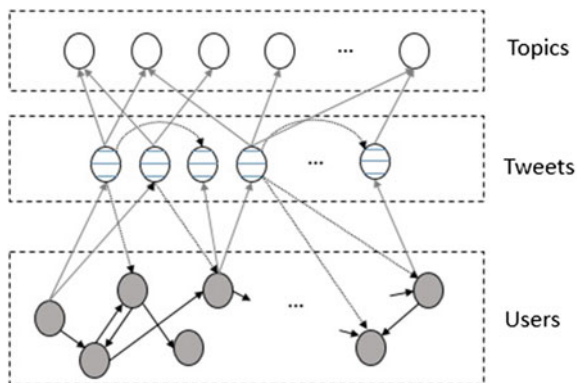
2 Related Work

In the recent years a lot of work done related to developing feature models for detecting suspiciousness. Lee et al. [8], have proposed a method for detection of suspicious URLs, these are redirecting chain of URLs. Caverlee et al. [7], have discussed profile based features for classification of Twitter users into malicious or legitimate. Pasquale et al. [3], have suggested classification of the malicious and fraudulent behavior by the global and local reputation of user. In addition to the spam detection [2, 4] in social network, rumors identification attain a much attraction. And rumors trust score [1, 6] is unverifiable because it definitely false [9]. Short URLs are used to reduce the size. Clicking on short URLs, so it is redirecting on to the long URLs. There is ambiguity with the short URLs [5].

In heterogeneous graph representation as shown in Fig. 1 there are three types of vertices, corresponds to three major entities in online social networks e.g., users, tweets, and hashtag topics.

Directed edges connecting vertices in the graph as shown in Fig. 1 represent different types of social activities. First, an edge from user u_i to user u_j means that u_i relates to u_j in the network (e.g., u_i is following u_j in Twitter). Second, an edge from user u_i to tweet t_j indicates that u_i is the author of t_j (e.g., u_i posts a tweet t_j in Twitter). Third, an edge from tweet t_i to topic h_j represents that h_j is one of the topics covered in t_i (e.g., h_j is a hashtag topic in a tweet t_i). In addition, there are two more types of directed edges in the graph. One edge starts from tweet t_i and points to another tweet t_j . This represents that t_j is a retweet of t_i . Another type of edges connects a tweet t_i and a user u_j . This specifically captures the mention function in Twitter.

Fig. 1 Heterogeneous social graph



3 Data Collection

For our experimental dataset, the first step for our analysis is to gather data by crawling Twitter. We have used a Twitter API to collect the data. We can collect only the information that are in the public domain. If a user is keeping his data secret cannot be accessed (i.e. does not allowing other to access his personal information). Initially we have collected the user profile information. We received a dataset of 4,230 user’s information and collected 380 suspicious profile data. By using the user profile names tweets are collected. Here, the tweets contains a hashtags. So we have collected hashtags from all the tweets. And further we have collected tweets which are related to hashtags. From each hashtag tweets the short URLs are extracted. The details of the data collection process is shown in Fig. 2.

Twitter quickly reacts to detected malicious profile, as well as deletes any malicious tweet found in order to get the Social Network clean from fraud. So if we want to get this malicious data for our analysis we should be quicker than Twitter and gather as much data as possible before it is deleted.

As shown in Fig. 3 each extracted short URLs from hashtag tweets, is queried to google safe browsing API to find whether the short URLs are malicious or not. Google safe browsing maintain a black listed URLs. When the request is sent, it searches against blacklisted URLs. If query returns “false” then the requested URL is malicious. If it returns “true” then the URL is legitimate. We are assigning trust score to “hashtags” based on number of legitimate URLs i.e.

$$Trust\ score\ of\ hashtag = \frac{\#Number\ of\ legitimate\ URLs}{\#Total\ number\ of\ URLs} \tag{1}$$

If the hashtag having the high trust score, then the information related to that is more trustable. If the trust score value is low all the information related to that is malicious. If the trust score is 0.5 then it is not decided (i.e. it may be either malicious/legitimate).

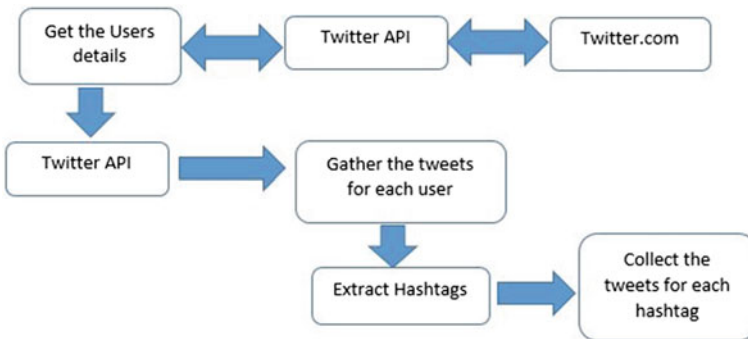


Fig. 2 Data collection

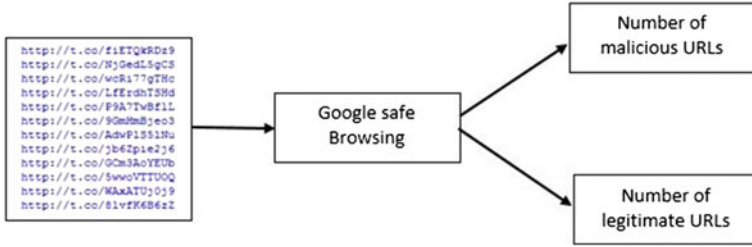


Fig. 3 Short URLs labeling

4 Proposed Algorithm

Data: a heterogeneous graph representation $G(V, E)$, a trust threshold θ ;

Result: a set of malicious activities Mal ;

- 1 Initialize a trustworthiness score of 0.5 to each node in G ;
- 2 Initialize a trust score to each T in G based on the formula given in Eq. 1
- 3 **repeat**
 - 4 $\forall v \forall u \text{ Trust score}(u) = \sum_{x \in B_u} \text{Trustscore}(x) / N_u$
 - 5 **until** all nodes are visited in U ;
 - 6 **repeat**
 - 7 $\forall v \text{ Trust score}(v) = \sum_{x \in B_v} \text{Trustscore}(x) / N_v$
 - 8 **until** all nodes are visited in V ;
 - 9 Repeat step 6 to 8 until reaching a stable status; each vertex v is calculated a trust score $T(v)$;
 - 10 initialize Mal to be \emptyset ;
 - 11 **for** every $v \in V$ **do**
 - 12 **if** $(T(v) \leq \theta)$ **then**
 - 13 $\text{let } Mal = Mal \cup v$;
 - 14 **return** Mal ;

Where

T, U, V is collection of topic node, tweets and user nodes

N_u, N_v is the out degree of the node U, V

B_u, B_v is the set of nodes pointed by node U, V

$T(v)$ is trust score of node v .

θ is threshold value for trust score

The most important step in the above algorithm is the calculating trust score for the user node in heterogeneous social graph. Trust score is calculated based on the PageRank algorithm. Initially Mal is empty and it store the information about the nodes which are less than. Here we are classifying based on the trust score. If the user having score less than the threshold value are classified as malicious.

4.1 Feature Selection

In this approach, we propose a new feature i.e. trust score for detecting malicious user. The following are feature used in our classification

- **User ID:** ID of a user assigned by Twitter
- **Followers Count:** Number of user following him
- **Friends Count:** To the number of user the user is following
- **Status Count:** Total number tweets posted
- **User location:** Geographic location of the user
- **Has URL:** Some users having URL in profile data
- **Spam URLs:** number of spam URLs present in tweets
- **Duplicate URLs:** Ratio of total URL and number of unique URLs
- **Trust score:** It is extracted from all the user information

Trust score is more important feature it is calculated based on tweets, hashtags etc.

5 Machine Learning Classification

Now, we describe the way of classification of malicious users. Initially, the dataset is divided into training dataset (80 %), testing dataset (20 %). In order to assess the most efficient mechanism to detect malicious accounts, we inspected various machine learning algorithms. For this classification, we have used the most popular Weka software package. In this most of the classification algorithms are implemented. Weka is open source collection of machine learning classifiers for data mining. The following classifiers are evaluated for our dataset.

5.1 Evaluation Metrics

Accuracy (A) and F-measure are the metrics which are used for the evaluation of the classifier performance. If evaluation metrics having higher value, then the classifier is best suitable for data set. The evaluation metrics described effectively by confusion matrix as shown in Table 1.

Table 1 Confusion matrix

	Malicious	Legitimate
Malicious	TP	FN
Legitimate	FP	TN

6 Results

The objective of current study is identifying aberrant behavior of users in Twitter. We have analyzed user suspiciousness based on the trust score. If the calculated trust score is greater than the threshold value θ then the user is legitimate user. We are taken a threshold value as 0.5. If the user score is less than 0.5 then the user no more trustable as shown in Fig. 4. Here we treat the obtained trust score as a feature along with the all obtained user profile features like followers count, following count, status count etc. In the above Table 2, it shows that decision tree works better compared with the other classifiers, Fig. 5 also describes that the accuracy of decision tree is better than the other two classifiers. In our dataset, decision tree correctly classifies 75 % malicious users. 25 % malicious users are misclassified as legitimate.

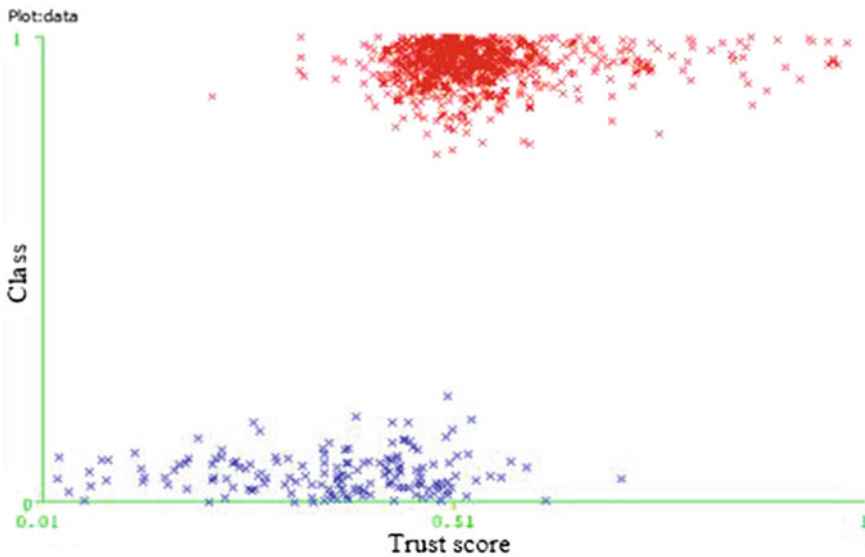


Fig. 4 Classification of users

Table 2 Comparison of classifiers

Evaluation metric	Decision tree (%)	Naive bayes (%)	Random forest (%)
Accuracy	92.6	89.9	90.4
F-measure (Malicious)	81.0	64.4	76.3
F-measure (Legitimate)	95.5	93.4	94.0
True positive rate	88.2	80.9	79.0
False positive rate	93.6	90.1	93.0
Positive predictive rate	74.9	53.5	73.7
Negative predictive rate	97.3	97.1	94.8

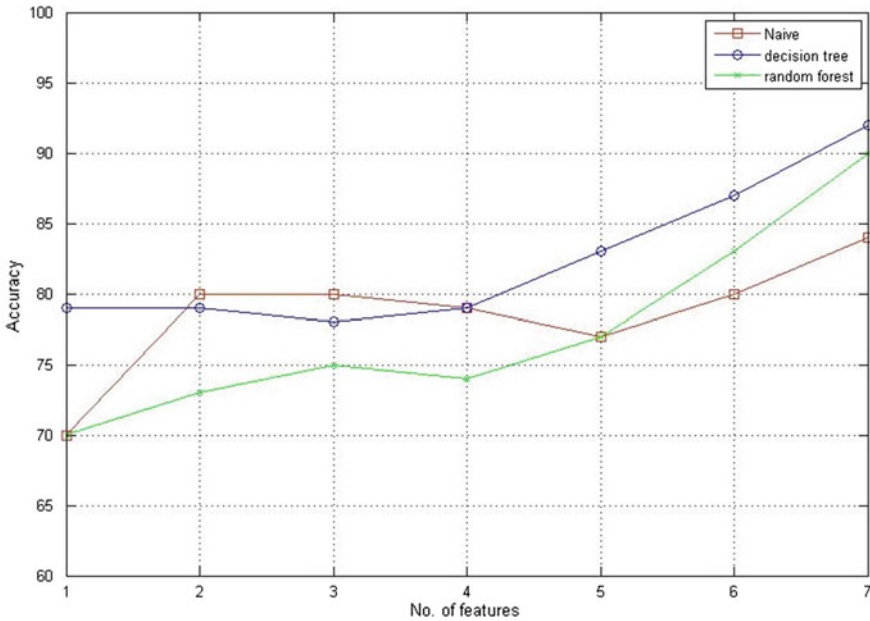


Fig. 5 Efficiency versus no of features in training data set

7 Conclusion and Future Work

In this paper, we have developed an algorithm for calculating trust score of each user in heterogeneous social graph for Twitter. The trust score is a special feature that can be used to detect malicious activities in Twitter with high accuracy. Our classifier attains an improved F-measure of 81 % and with an accuracy of 92.6 %. In this work, we have successfully detected malicious users. For calculating trust score we have considered only short URLs of trending topics. Based on the backward propagation, we assign trust score to tweets if trending topics present in that tweet and followed by the users. Future work deals with calculation of trust score by considering the short URLs present in the tweet.

References

1. Abdul-Rahman, A., Hailes, S.: A distributed trust model. In: Proceedings of the 1997 workshop on New security paradigms. pp. 48–60. ACM (1998)
2. Agarwal, M., Zhou, B.: Detecting malicious activities using backward propagation of trustworthiness over heterogeneous social graph. In: 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). pp. 290–291. IEEE (2013)

3. De Meo, P., Messina, F., Rosaci, D., Sarné, G.M.: Recommending users in social networks by integrating local and global reputation. In: *Internet and Distributed Computing Systems*, pp. 437–446. Springer (2014)
4. Fire, M., Kagan, D., Elyashar, A., Elovici, Y.: Friend or foe? fake profile identification in online social networks. *Social Network Analysis and Mining* 4(1), 1–23 (2014)
5. Gupta, N., Aggarwal, A., Kumaraguru, P.: bit.ly/malicious: Deep dive into short url based e-crime detection. In: *Electronic Crime Research (eCrime)*, 2014 APWG Symposium on. pp. 14–24. IEEE (2014)
6. Jiang, W., Wang, G., Wu, J.: Generating trusted graphs for trust evaluation in online social networks. *Future generation computer systems* 31, 48–58 (2014)
7. Lee, K., Caverlee, J., Webb, S.: The social honeypot project: protecting online communities from spammers. In: *Proceedings of the 19th international conference on World wide web*. pp. 1139–1140. ACM (2010)
8. Lee, S., Kim, J.: Warningbird: Detecting suspicious urls in twitter stream. In: *NDSS* (2012)
9. Qazvinian, V., Rosengren, E., Radev, D.R., Mei, Q.: Rumor has it: Identifying misinformation in microblogs. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. pp. 1589–1599. Association for Computational Linguistics (2011)

Distance, Energy and Link Quality Based Routing Protocol for Internet of Things

Kirshna Kumar, Sushil Kumar and Omprakash Kaiwartya

Abstract In future communication networks with IoT, each of the things will be able to communicate with other things ubiquitously throughout the time clock. Multipath distortion, noise and interference create problems for low power communication devices. These smart devices contain limited amount of battery, energy and processing power. In this context this dissertation proposes Distance, Energy and Link quality based Routing protocol (DELR) to enhance routing success probability and to minimize route setup delay. Finally, the performance of the proposed algorithm is evaluated with respect to the protocol: REL considering the metric such as routing success probability and route setup delay in the various rounds. The results demonstrate that the performance of proposed algorithm is better than the compared algorithm: REL in terms of routing success probability and route setup delay on the simulated network in IoT applications.

Keywords Internet of things · Route setup delay · Routing success probability · Energy efficiency · Link quality

K. Kumar (✉) · S. Kumar

Wireless Communication and Networking Research Lab, School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: kirshnakumar7@gmail.com

S. Kumar

e-mail: skdohare@yahoo.com

O. Kaiwartya

Faculty of Computing, Pervasive Computing Research Group,
Universiti Teknologi Malaysia, 81310 Skudua Johor, Malaysia
e-mail: omokop@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_25

1 Introduction

Internet of Things (IoT) is the process of creating, accumulating and communicating information among smart devices with or without human intervention. Things or Objects with communication capabilities and embedded intelligence are known as Smart Devices [1]. In Internet communication occurs between users, while in IoT communication occurs between devices automatically [2]. This paper proposes Distance, Energy and Link quality based Routing protocol (DELR) for IoT to enhance routing success probability and to minimize route setup delay. It considers link quality and energy efficiency. The performance of the proposed algorithm is evaluated with respect to the protocol: REL. The rest of the paper is organized as follows. In Sect. 2, related literatures are reviewed. In Sect. 3, DELR is presented. In Sect. 4, experimental results are discussed. Finally, Sect. 5 concludes the work presented in this article with future research direction in the area.

2 Related Work

Routing algorithms provide the forwarding methods on the basis of location of the current forwarding node, its neighbors, and the packet destination. Greedy routing can be based on progress, distance and direction. In the progress based scheme [3], the next forwarding node is selected based on the progress. Distance based [3] greedy routing scheme is based on the distance. Direction based greedy routing is based on direction. It considers the deviation (angle between current, next hop and destination node) from the line connecting active sender and destination [3]. GEDIR [4] is a position based greedy forwarding algorithm.

Energy-efficient probabilistic routing (EEPR) algorithm [5], which controls the transmission of the routing request packets stochastically in order to increase the network lifetime and decrease the packet loss under the flooding algorithm. In EEPR algorithm the routing setup delay is slightly increased and the routing success probability is slightly decreased. Routing protocol based on Energy and Link quality (REL) [6] can be used to optimize route selection mechanism by using link quality estimation, energy evaluation and providing load balancing mechanisms. REL algorithm provides less routing success probability and routing set up delay is also increased.

3 DELR

3.1 Enhanced Range Directional Routing (ER-DIR)

ER-DIR algorithm provides different possible routes as output and that output is used as input in RELD algorithm to find optimum path from source to destination.

Algorithm: ER-DIR

Input: (Current Node, neighbor Node, Destination Node)

Process:

1. N = number of neighbor nodes of current node inside Range-DIR region.
2. m_1 = the slope of current node (a_1, b_1) and neighbor node (a_2, b_2) .
3. m_2 = the slope of current node (a_1, b_1) and destination node (a_2, b_2) .
4. M = Angle between slope m_1 and m_2 .
5. $m_1 = (b_2 - b_1)/(a_2 - a_1); m_2 = (b_3 - b_1)/(a_3 - a_1);$
6. $M = \tan^{-1}((m_2 - m_1)/(1 + m_1 m_2));$
7. Find all neighbor nodes of current node inside Range-DIR region.
8. If $(N > 0)$ then
9. Select all neighbor nodes of current node in parallel manner inside Range-DIR region to find different route through them but in increasing order of angle (M) .
10. else
11. Select one nearest angle (M) node, out of this region.
12. Set current node = neighbor node
13. Repeat steps 1 to 13 until we find every route in range-DIR region from source node to destination node.

Output: A set of different routes (S_i) between source node and destination node.

3.2 Distance Calculation

Distance between two nodes with coordinates (a_1, b_1) and (a_2, b_2) is calculated using Eq. 1

$$Distance = \sqrt{(a_2 - a_1)^2 + (b_2 - b_1)^2}; \quad (1)$$

Distance of one route is calculated by adding distance between all intermediate nodes from source node to destination node. One value of distance is selected as threshold value that is called Distance threshold (D^{th}) , which is selected with respect to straight distance (d) from source node to destination node to increase routing success probability.

3.3 Link Quality Evaluation

Link quality is measured in terms of either Link Quality Indicator (LQI) or Receiver Signal Strength Indicator (RSSI). Optimal threshold value (LQI_{th}) is selected to provide optimal packet delivery ratio (PDR). Link is divided into good links and weak links on the basis of LQI_{th} value, if it is lesser than LQI value and more than LQI value respectively. Finding path sender sends Route Request (RREQ) message and receiver node sends Route Reply (RREP) message. Both messages contain LQI information of link and this information is updated at every hop. If it is weak link then updated if required. On the basis of average LQI value of present destination, routes are evaluation.

3.4 Energy Evaluation

The consumption of energy by each node is required in balanced way. For energy evaluation one threshold value, Energy threshold (E^{th}) is selected. In starting phase of execution in the network, percentage of residual energy has to be stored by every node and, the present energy level E_t is compared with previous energy level $E_{(t-1)}$, after each t time units by each node. An energy event of discharge is indicated if E^{th} is less than the difference between E_t and $E_{(t-1)}$ which is called ($Index_{RADV}$). Difference of energy consumption in the nodes is dependent on the Energy threshold (E^{th}) value (Fig. 1).

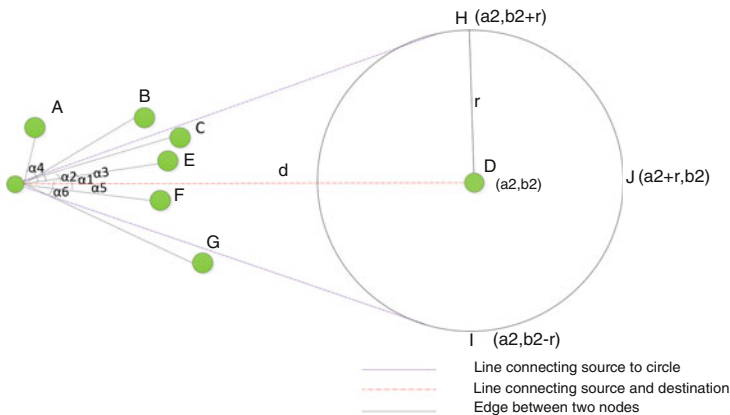


Fig. 1 Selection of neighbor nodes based on angle

3.5 Optimum Path Algorithm

Algorithm: RELD

Notations:

D^{th} : Distance threshold; E^{th} : Energy threshold; R_c : Current route; R_n : New route
 E_c : Energy of current route; E_n : Energy of new route; D_n : Distance of new route
 D_c : Distance of current route; Bl_c : Bad links of R_c ; Bl_n : Bad links of R_n
 GoToRoute(): function to switch between R_c and R_n .

Input: A set of different routes (S_i)

Process:

1. if $E_c = E_n$ then
2. if $D_c > D_n + D^{th}$ then
3. if $Bl_c \geq Bl_n$ then
4. GoToRoute(R_n)
5. end if
6. end if
7. else if $E_c < E_n$ then
8. if $D_c + D^{th} \geq D_n$ then
9. if $Bl_c \geq Bl_n$ then
10. GoToRoute(R_n)
11. end if
12. end if
13. else if $E_c > E_n$ and $E_c \leq E_n + E^{th}$ then
14. if $D_c > D_n + D^{th}$ and $Bl_c \geq Bl_n$ then
15. GoToRoute(R_n)
16. end if
17. end if

Output: optimized route

4 Experimental Results and Discussion

The tool used to simulate our work is Matlab 2009b.

Figure 2 shows PDR for various values of D^{th} which changes from 4 to 20 m. PDR is higher than desired 80 %. According to our results, the most appropriate values for D^{th} is 12 m for better PDR. Figure 3 shows the result of route setup delay. Route setup delay is described as the time difference between the time when a source node forwards the RREQ packets and the time when destination node receives the first RREQ packet. It is directly proportional to collision. The route setup delay under the DELR algorithm has approximately 1.106 ms lesser than that under the typical REL protocol (Table 1).

Figure 4 shows routing success probability of DELR algorithm is 93.233 % which is approximately 2.5 % higher than that of the typical REL protocol.

Fig. 2 Simulation result of distance threshold (D^{th})

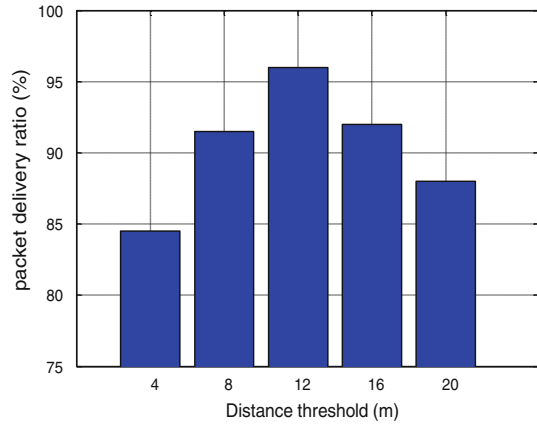


Fig. 3 Route setup delay

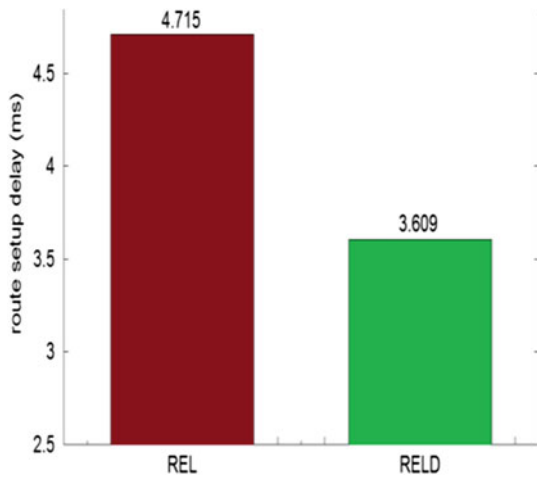
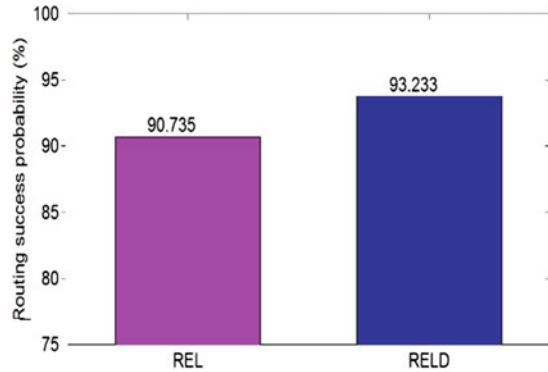


Table 1 Simulation parameters

Parameters	Values
Area	200 m × 200 m
Number of nodes	100
Simulation time	60 min
Base station location	(60, 60)
Initial energy	17565 J (2 AA batteries)
E^{th}	2
LQI_{th}	200
D^{th}	12 m

Fig. 4 Routing success probability



5 Conclusion and Future Work

The design of this protocol and its analysis theoretically prove that DELR protocol minimizes route setup delay and enhances routing success probability of the network. In particular, using DELR protocol routing success probability is enhanced from 90.735 to 93.233 % and route setup delay is reduced from 4.715 to 3.609 ms as compared to REL. Algorithm considered in this paper, works well in static environment. In future Adaptability in a dynamic network environment will be also explored.

References

1. Bello, O., Zeadally, S.: "Intelligent Device Communication in the Internet of Things," Systems Journal, IEEE, Vol.1, no.99, pp.1–11, 2014.
2. Evans, D.: "The Internet of things: How the next evolution of the Internet is changing everything," Cisco IBSG, San Francisco, CA, USA, 2011.
3. Frey, H., Rührup, S., Stojmenović, I.: "Routing in Wireless Sensor Networks," Springer-Verlag London Limited, pp.81–111, 2009.
4. Madani, S.A., Weber, D., Mahlknecht, S.: "Position-based Routing Protocol For Low Power Wireless Sensor Networks," Journal of Universal Computer Science, vol.16, no.9, pp.1215–1233, 2010.
5. Park, S., Cho, S., Lee, J.: "Energy-Efficient Probabilistic Routing Algorithm for Internet of Things," Journal of Applied Mathematics(213106), Available online: <http://www.ietf.org/rfc/rfc3561.txt> (accessed on 2014).
6. Machado, K., Rosário, D., Cerqueira, E., Loureiro, A.A.F., Neto, A., de Souza, J.N.: "A Routing Protocol Based on Energy and Link Quality for Internet of Thing Applications," Sensors, MDPI, Vol.13, pp.1942–1964, 2013.

Effect and Suppression of Noise in 2D PC/OOC Scheme for Optical CDMA Systems

Manisha Bharti, Ajay K. Sharma and Manoj Kumar

Abstract The performance of 2D wavelength/time optical code division multiple access (OCDMA) system is adversely affected by the presence of noise at its physical layer. It is present in the form of Multiple Access Interference (MAI) and beat noise in OCDMA systems and degrades its performance. In this paper an attempt has been made to determine as well as to mitigate the effect of noise using either Optical Hard Limiter (OHL) or Coherent detection technique for a system employing prime sequence permutations over time spreading optical orthogonal codes (OOCs). Investigations reveal that performance of system is severely affected by noise and it can be improved by using coherent detection as well as hard limiter. Specifically for this PC/OOC coding technique, it has been reported in the results that coherent detection for reducing the noise outperforms the use of optical hard limiter.

Keywords Optical code-division multiple access (OCDMA) • Multiple access interference (MAI) • Optical hard limiter (OHL) • PC/OOC

1 Introduction

Optical code division multiple access (OCDMA) system is based on the principle of electrical CDMA. The significant distinction between the two lies in the frequency and channel used [1]. In both the techniques, unique codes are mapped to the

M. Bharti (✉)

I. K. Gujral Punjab Technical University, Kapurthala, Jalandhar, India
e-mail: manishabharti@aiactr.ac.in

A.K. Sharma

National Institute of Technology, Delhi, India
e-mail: sharmaajayk@rediffmail.com

M. Kumar

DAV Institute of Engineering and Technology, Jalandhar, India
e-mail: drmanojkumarindia@gmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_26

261

identities of users using the principle of spread spectrum. OCDMA system was initially proposed for the implementation of ultrafast asynchronous broadcast LANs [1, 2]. It continues to receive attention due to its ability for enhance information security, simplified network control and improved spectral efficiency [2]. However, noise, on the other hand, is the ever present nuisance in any OCDMA system affecting its performance at physical layer. It restricts the maximum number of users due to the presence of multiple access interference (MAI) and beat noise. MAI results when many users simultaneously share the same transmission bandwidth and interfere with each other [3]. It is a serious issue to be addressed specially in asynchronous OCDMA systems. This is caused due to the non-ideal orthogonal property of the optical codes [4]. Due to random time offsets between the signals, the interference comes into existence that makes it impossible to design the code waveforms to be completely orthogonal [5]. Secondly, beat noise is caused by the interaction among electrical fields belonging to signal, spontaneous emission and local laser oscillator [6]. At the receiver side, photo detector generates a photocurrent along with several side components. These side components are due to the beating action of noise electric field with other optical noise components leading to the formation of beat noise [7].

Several methods have been proposed by researchers to improve the OCDMA system's performance in presence of noise. Both MAI and beat noise can be reduced to a great extent by using optical hard limiter (OHL) and heterodyne coherent detection receiver. OHL is a non linear device that is capable of limiting the received optical power to a fixed level when it is higher than or equal to the threshold [8]. According to Dang et al. [9], at a particular wavelength, only those pulses whose maximum power are longer or equal to the threshold power can pass through the OHL. Another method to combat with beat noise and MAI is heterodyne detection. Pham et al. [10], proposed the use of heterodyne receiver for optical systems to improve their performance. Heterodyne detection is a type of coherent detection technique that consists of a local oscillator (LO), whose frequency differ from the received signal carrier frequency. The frequency of local oscillator is coherently mixed with the received signal.

The purpose of this manuscript is to analyze the impact of beat noise and MAI on PC/OOC codes and to apply optical hard limiter (OHL) as well as heterodyne detection to combat with these effects. The results are reported on the basis of probability of error. The paper is organized in the following manner. Section 1, introduces the basic concepts of OCDMA and noise types that affect the performance of system along with the methods to combat with these effects. Section 2 describes the performance analysis of PC/OOC codes with and without the presence of noise and it also includes the enhancement in performance using OHL and coherent detection at receiver. Finally, Sect. 3 concludes the paper.

2 Performance Analysis of PC/OOCs

This family of codes support multiple wavelength hopping for a given time spreading code by using prime sequences for wavelength hopping and $(n, w, 1, 1)$ optical code for time spreading. This is represented as $(m \times n, w, \lambda_a = 1, \lambda_c = 1)$. The time slot may contain none or one pulse of the appropriate wavelength in this code [8]. This wavelength is decided with the help of permutations of wavelengths that are done by prime sequences over non-zero time slots of the pre decided time spreading orthogonal code. The code weight w of optical orthogonal codes can be as high as p (prime number) providing a total group $\Phi_{group} = p$, of prime sequences to generate the new matrices [10, 11]. The value of auto correlation and cross correlation functions of the PC/OOC codes is 1 to minimize the effect of multiple access interference MAI. The cardinality of optical orthogonal codes is described by [11]:

$$\emptyset_{ooc} \leq \frac{N_{ooc} - 1}{w(w - 1)}, \tag{1}$$

Here $N_{ooc} \geq w(w - 1)\emptyset_{ooc} + 1$ denotes the length of code and w denotes the weight of code. The cardinality of PC/OOC is described by [10]:

$$\emptyset_{pc/ooc} = \emptyset_{ooc} * P^2 \tag{2}$$

Let q^0 and q^i be the probability of having one hit among the code set from group 0 and group i (for $i = 1, 2, \dots, p-1$), respectively. Then we have from [11]:

$$q^0 = \frac{w^2(\emptyset_{ooc} - 1)}{2N_{ooc}(\emptyset_{ooc}P^2 - 1)} \tag{3}$$

$$q^i = \frac{w^2(\emptyset_{ooc}P^2 - 1) + (w - 1)^2}{2N_{ooc}(\emptyset_{ooc}P^3 - 1)}, \tag{4}$$

Taking Th and K as the decision threshold and number of users in the system respectively, the probability of error for PC/OOC is given by [6]:

$$P_e \leq \frac{1}{2} \sum_{i=0}^{Th} (-1)^i \binom{w}{i} \left[1 - \frac{q_i}{w} \right]^{k-1} \tag{5}$$

In general, for the optimum performance of the system the value of decision threshold is taken to be equal to the weight of code. The generic equation used for MAI is obtained as [12]:

$$P_e = \frac{1}{2} \sum_{i=Th}^{K-1} \binom{K-1}{i} q^i (1-q)^{K-1-i} \tag{6}$$

Here, K is the number of simultaneous users, Th is pre-determined threshold of the receiver the threshold (= code weight). The equation for q (average probability) is defined separately for each case. Similarly the generic equation for error probability in the presence of Beat Noise is obtained from [13] and is given as:

$$P_e = \sum_{i=1}^{K-1} \binom{K-1}{i} \frac{1}{2} \{Q(SNR_0) + Q(SNR_1)\} q^i (1-q)^{K-1-i} \tag{7}$$

$$SNR_0 = \frac{p_s P_d D - j P_c}{\sqrt{2 P_c P_c \binom{j}{2} \frac{1}{p_s}}}; \quad SNR_1 = \frac{p_s P_d + j P_c - p_s P_d D}{\sqrt{2 j P_d P_c + 2 P_c P_c \binom{j}{2} \frac{1}{p_s}}} \tag{8}$$

Here, p_s is the available wavelengths (= w), P_d and P_c are optical powers of data and interferer pulses respectively, D is the threshold (= $1/2$). In Eq. (7) it is assumed that out of K simultaneous users, i users out of possible $K-1$ are transmitting “1” (binomial distribution with probability $1/2$) and at the time of thresholding, j pulses among i are deposited to form the autocorrelation peak. The equation for error probability of PC/OOC under MAI condition is obtained by substituting the value of q_{PC} in Eq. (6).

In Fig. 1, the error probability for $m = 7, 11$ and 13 have been individually calculated using Eq. (7). In the case of $m = 7$, the wavelength-hopping sequence

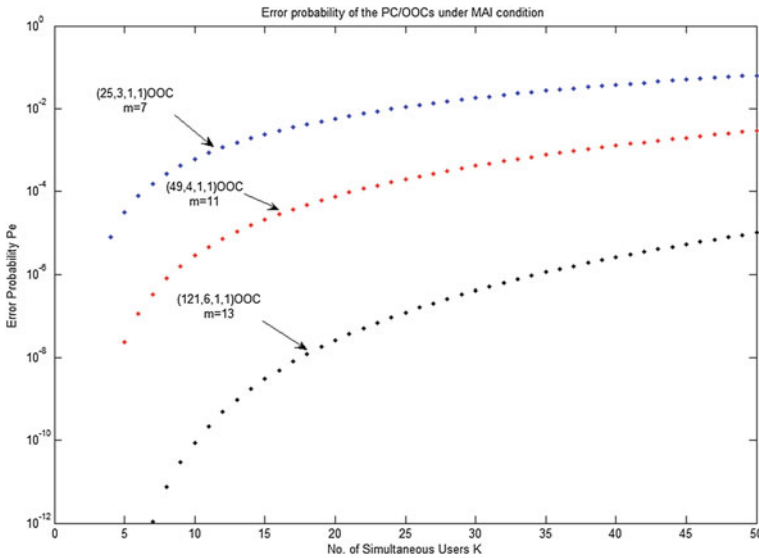


Fig. 1 Error probability of the PC/OOCs under MAI condition versus the number of simultaneous users for various m

has been generated over GF (7) and for time spreading (25, 3, 1, 1) OOC has been used. Similar procedure is used for $m = 11$ and 13. From the above figure, the effect of m on the error probability in the presence of MAI can be easily seen. The error probability decreases as the value of m is increased from 7 to 13. One may expect MAI to increase the error probability as m is increased, however, due to the code properties of (n, w, l, l) OOC, the code weight w increases by a very small value (from 3 to 6 as m is varied from 7 to 13) as compared to the increase in code length n (which increases from 25 to 121 as m is varied from 7 to 13). The larger code length along with a smaller code weight ensures lesser number of interfering terms. Hence, MAI decreases with the increasing value of m in this case.

Similarly the equation for PC/OOC in the presence of Beat Noise has been deduced by substituting the value of q_{PC} in Eq. (7). In Fig. 2, the error probability of PC/OOC is shown in the presence of Beat Noise. This is compared with the respective error probabilities when Beat Noise is absent. For all values of m , the error probability increases when Beat Noise is present. The higher the value of m , greater is the Beat Noise. From the above figure, when the code weight, w , is increased marginally from 3 to 6, difference in the error probabilities between the presence and absence of Beat Noise increases considerably. This is because; more beating components are produced in the receiver for heavier code weights.

For analyzing the effect of using hard limiter the hard limiting error probability (HEP) for PC/OOC as given by following equation:

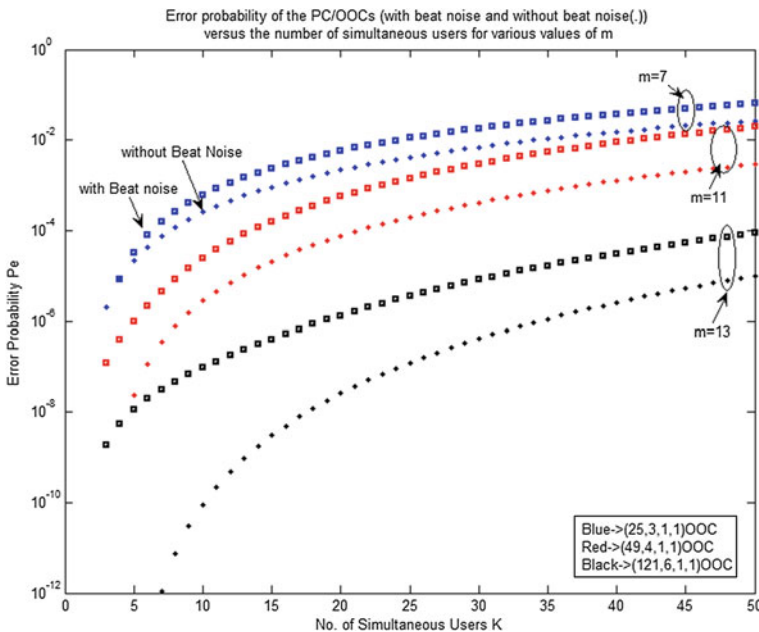


Fig. 2 Error probability of the PC/OOCs (with and without Beat Noise) versus the number of simultaneous users for various m

$$P_e \leq \frac{1}{2} \sum_{i=0}^{Th} (-1)^i \binom{w}{i} \left[1 - \frac{q_{PC} i}{w} \right]^{K-1} \tag{9}$$

where K denotes the number of simultaneous users and q_{pc} is obtained by (3).

From Fig. 3, the effect of OHL is readily seen. For all values of m , the use of OHL improves the performance considerably. The performance also improves with the increase in code length, code weight and number of available wavelengths. The performance for $m = 13$ is the best as seen from the figure. The generic equation for error probability when Coherent Detection is employed in the receiver is obtained from [11] and given as:

$$P_e = \sum_{i=1}^{K-1} \binom{K-1}{i} \frac{1}{2} \{Q(SNR_0) + Q(SNR_1)\} q^i (1-q)^{K-1-i} \tag{10}$$

where K is the number of simultaneous users, q is the average hit probability defined separately for each code. SNR_0 and SNR_1 can be obtained from equation as given below [11].

$$SNR_b = \frac{(i_b - i_D)^2}{i_{nb}^2} \tag{11}$$

i_b , where b may be bit “0” or bit “1” is the sum of data and MAI currents and is given as [12]. The equation for PC/OOC for Coherent Detection has been given by:

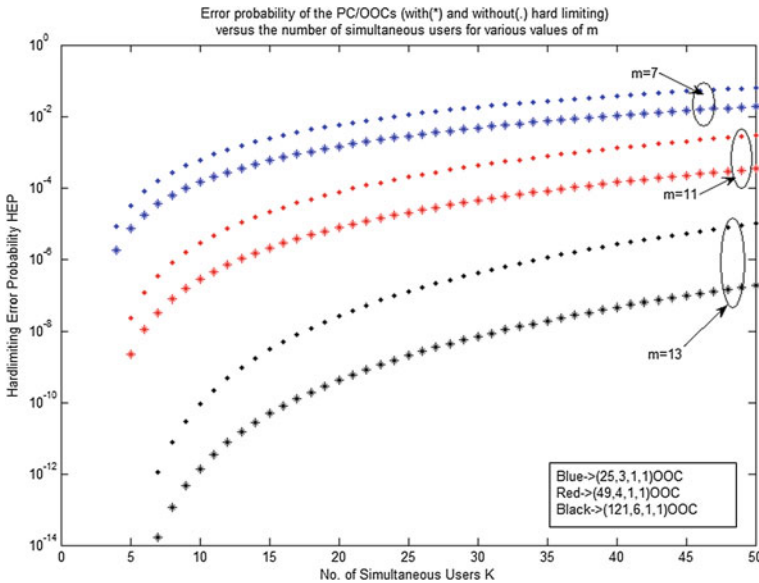


Fig. 3 HEP of the PC/OOCs versus the number of simultaneous users for OHL

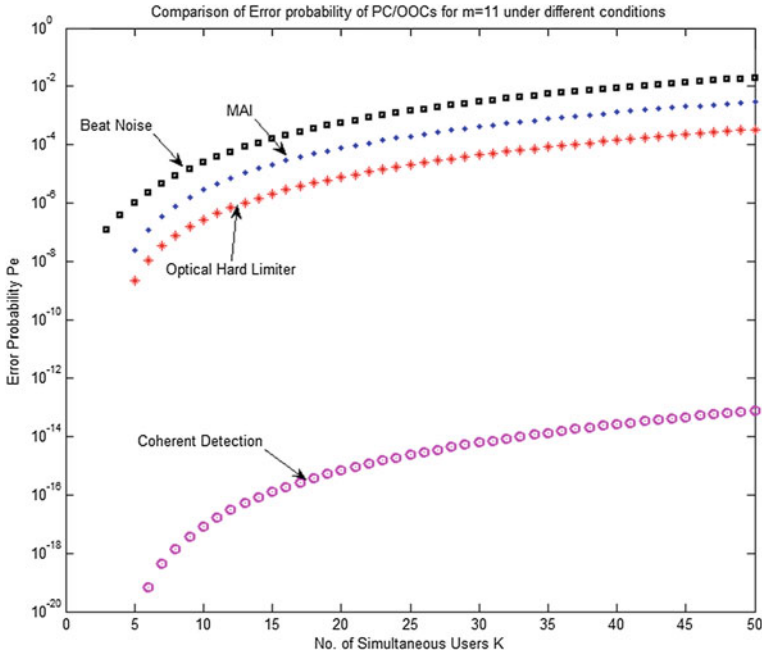


Fig. 4 Error probability of the PC/OOCs versus the number of simultaneous users for $m = 11$ for Coherent Detection

$$P_e = \sum_{i=1}^{K-1} \binom{K-1}{i} \frac{1}{2} \{Q(SNR_0) + Q(SNR_1)\} q_{PC}^i (1 - q_{PC})^{K-1-i} \quad (12)$$

Figure 4 shows the performance of Coherent Detector as compared to Optical Hard Limiter for PC/OOC. The curves for MAI and Beat Noise is also plotted. It is seen that OHL is capable of reducing both MAI and Beat Noise. The performance due to Coherent Detector, however far exceeds that of OHL.

3 Conclusion

In this paper, we have examined the effect of MAI and Beat Noise for PC/OOC 2D optical codes for different number of available wavelengths for OCDMA system. The various investigations reveal that effect of MAI and beat noise is more at higher number of wavelengths and where the code weight is larger. An attempt has also been made to improve the system performance by using OHL and coherent detection at the receiver. The results obtained provide significant conclusions. First, both OHL and Coherent Detection successfully reduce MAI and Beat Noise, with Coherent Detection outperforming OHL.

References

1. Prucnal, P. R., Santoro *et al.*, "Spread Spectrum Fiber-optic Local Area Network using Optical Processing", *IEEE/OSA Journal of Lightwave Technology*, vol.4, No. 5, May 1986, pp. 547–554.
2. Jawad. A. Salehi, "Code Division Multiple Access Techniques in Optical Fiber Networks-Part I: Fundamental Principles", *IEEE Transactions on Communications*, vol. 37, No. 8, August. 1989, pp. 824–833.
3. S.V.Mric *et al.*, "A New Family of Optical Code Sequences for use in Spread Spectrum Fiber-optic Local Area Networks", *IEEE Transactions on Communications*, vol. 41, No. 8, August. 1993, pp. 1217–1221.
4. Andrew Stok and Edward H. Sargent, "The Role of Optical CDMA in Access Networks", *IEEE Communications Magazine*, September 2002, pp. 83–87.
5. L.Bin, "One-coicidente sequences with specified distance between adjacent symbols of frequency hopping multiple-access", *IEEE Transactions on Communications*, vol. 45, No. 4, April, 1997, pp. 408–410.
6. B. K. Whitlock *et al.*, "Modeling and Simulation of the OETC optical bus", *IEEE J. Sel. Areas Commun.*, vol. 15, No. 4, May, 1997, pp. 717–730.
7. Tancevski *et al.*, "Hybrid wavelength hoping/time spreading schemes for use in massive optical networks with increased security", *IEEE/OSA Journal of Lightwave Technology*, vol.14, No. 12, Dec 1998, pp 2636–2647.
8. G.C.Yang and W.C. Kwong, "Prime Codes with Applications to CDMA Optical and Wireless Networks", Boston, MA: Artech House, 2002.
9. R. M. Gagliardi, "Design and performance analysis of wavelength/time (W/T) matrix codes for optical CDMA", *IEEE/OSA Journal of Lightwave Technology*, vol. 21, No. 11, Nov, 2003, pp. 2524–2533.
10. Sun Shurong, "A New Family of 2D Optical Orthogonal Codes for Optical CDMA Access Networks", Asia Pacific Conference on Communications, Perth, Western Australia, 3–5 October-2005.
11. W. C. Kwong, G.-C. Yang, V. Baby, C.-S. Brès, and P. R. Prucnal, "Multiple-wavelength optical orthogonal codes under prime-sequence permutations for optical CDMA," *IEEE Trans. Communications.*, vol. 53, no. 1, pp. 117–123, Jan. 2005.
12. Tancevski, A. Selvarajan, and T. Srinivas, "Two-dimensional optical orthogonal codes for fiber-optic CDMA networks," *J. Lightwave Technol.*, vol. 23, no. 2, pp. 647–654, Feb. 2005.
13. T. C. Wanget *et al.*, "Anew Family of 2D codes for fiber optic CDMA system with and without the chip synchronous assumption", *J. Lightwave Technology.*, vol. 27, no. 14, pp. 2612–2620, July 2009.

On-the-Fly Segment Density (OFSD) in Adaptive Beaconing System (ABS) Based Connectivity-Aware Geocast Routing (CAGR) in VANETs

Durga Prasada Dora, Sushil Kumar and Puspanjali Mallik

Abstract Vehicular Ad-hoc Networks (VANETs) is a Critical Communication Infrastructure (CII), a part of Intelligent Transportation System (ITS) that enable the system to minimize vehicular accident by reducing the traffic congestion. The embedment of connectivity analysis in designing of geocast routing is the in-separable phase in VANETs for fulfillment of the said objective. High mobility, link disconnection and uneven distribution of vehicular nodes makes analysis of connectivity in VANETs imperative. This paper proposes On-the-Fly Segment Density (OFSD) in Adaptive Beaconing System (ABS) based Connectivity-Aware Geocast Routing (CAGR) protocol which enhances the data delivery ratio by guaranteeing assured connectivity. Simulation result shows that the proposed scheme works better when the no of nodes is higher, as compared to greedy routing.

Keywords Connectivity-aware geocast routing • OFSD • ABS • VANETs

1 Introduction

VANETs, a subclass of Mobile Ad-hoc Networks (MANETs) is vehicular traffic networks used for communication and cooperation between vehicular nodes for safety message dissemination that minimizes vehicular congestion as well as

D.P. Dora (✉) · S. Kumar

Wireless Communication and Networking Research Lab, School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: doradurga@gmail.com

S. Kumar
e-mail: skdohore@yahoo.com

P. Mallik
Department of Computer Science, Ravenshaw University, Cuttack, Odisha, India
e-mail: puspanjalimallik.cs@ravenshawuniversity.ac.in

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_27

269

accident and makes the transportation smoother [1, 2]. Efficient coordination and communication between vehicular nodes is done in VANETs to disseminate critical information such as road accidents, traffic jams, multimedia services and other e-commerce application [3, 4].

This paper proposes On-the-Fly Segment Density (OFSD) in Adaptive Beaconsing System (ABS) based Connectivity-Aware Geocast Routing (CAGR) protocol. In the following sections, the subsequent divisions of this paper are presented. Section 2 reviews the relevant literatures and related work regarding connectivity-aware geocasting. Section 3 introduces the proposed routing scheme and its different phases. Section 4 provides information regarding simulation and analysis of results. Section 5 concludes this paper and gives potions for future work.

2 Related Work

Shortest path based traffic aware routing (STAR) [5] considers effect of traffic light on connectivity in designing of routing protocols. It assumes high density on green light segment. A connectivity-aware intersection-based routing in VANETs (CAIR) [6] chooses intersected connected routes having higher connectivity and lower transmission delay based on the fly real time traffic density collection scheme. Connectivity-aware routing (CAR) [7] uses adoptive beaconsing mechanism where beaconsing interval is related to the node density of the neighbors and network load due to generation of beacon packets reduced significantly. In [8], Adaptive Traffic Beacon (ATB) incorporates two parameters such as message utility and channel quality and shown that adoptive beaconsing enhances the penetration rate by broadening dissemination of message than flooding based approaches.

3 Connectivity-Aware Geocast Routing (CAGR)

Dis-connectivity, as shown in the Fig. 1, can be arise in VANETs due to three different circumstances such as directional changes of vehicular nodes, variation in vehicular speed and degradation of link quality due to interference and shadow fading.

3.1 *On-the-Fly Segment Density (OFSD) in Adaptive Beaconsing System (ABS)*

In the proposed on-the-fly-node density (OFSD) approach, the source or the relay node attaches extra information such as node density (number of neighbor nodes

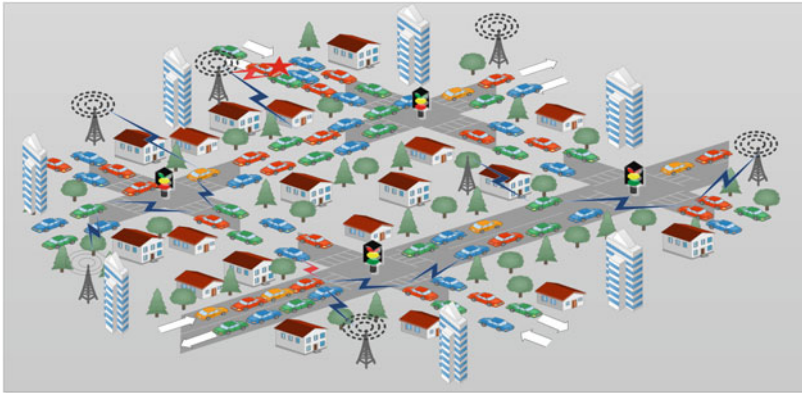


Fig. 1 Connectivity-aware geocasting in vehicular adhoc networks

around the sender) as well as segment density (number of nodes present in path segment of the network) in the beamed packet during beaconing. Proposed OFSD-ABS algorithm describes as follows. Node N stores the IP address information I_p of its neighbors from beamed information BI and stores in its IP address table T_{IP} and calculates node density N_d . When critical information is disseminated by the sender, it attaches N_d with it and send. In this way, when the data packets reaches to the receiver through routers by multi hopping, each router does the same thing and segment density S_d is available.

3.2 Impact of Traffic Signal on OFSD in ABS

When the traffic signal is Red, all the vehicles in the network segments starts accumulating and joined in a small area near traffic post. For that reason it is unnecessary to search the whole segment length SL to find S_d . The proposed algorithm TIM-OFSD-ABS searches only one-third of SL . Similarly, in case of Yellow traffic signal the nodes scattered relatively in larger area and search is done in one half of SL . Green traffic signal helps the accumulated vehicular nodes to move and the nodes scattered in whole path segment and for that reason search in SL .

Algorithm: On-the-Fly Segment Density (OFSD) in Adaptive Beaconsing System (ABS)

Notations: N_d : Node density; S_d : Segment density; T_S : Threshold Segment density
 T_r : Transmission range; I_p : IP address; BI : beaconsed Information;
 T_{IP} : IP address Table; SL : segment length; T : Traffic Signal

OFSD-ABS algorithm

```

1.  While (! Beacon ( $N_d, S_d, T_s$ ))
2.  {
3.    getDensity() // collect density information
4.    IF( $N_d=0$  &&  $S_d=0$ ) //no node within  $T_r$ 
5.      SAVE(CM); //store in cache memory
6.    ELSE IF( $N_d=0$  &&  $S_d!=0$ )
7.      WAIT(); //wait till node within  $T_r$ 
8.    ELSE
9.      {
10.     IF( $S_d>T_S$ ) //more nodes in segment
11.       BI="HIGH" //Beaconsing Interval is more. hence, less beaconsing
12.     ELSE
13.       BI="LOW" // Beaconsing Interval is less. hence, more beaconsing
14.     BEACON()//beaconsing of critical information
15.     }
16.  }
17.  getDensity() //collect density function
18.  {
19.     $N_d$ =getNdensity(); // collect node density function
20.     $S_d$ =getSdensity(); //collect segment density function
21.    RETURN  $N_d, S_d$ ; //both node and segment density are available
22.  }
23.  getNdensity() //collect node density function
24.  {
25.    WHILE(!BEACON(BI))
26.    {
27.       $I_p$ =getIP(BI); //get IP address of sender from beaconsed information
28.      STORE( $T_{IP}$ ); //store in  $T_{IP}$ 
29.    }
30.     $N_d$ =getTip()++; //increment node density
31.    RETURN  $N_d$ ;
32.  }
33.  getSdensity() //collect segment density function
34.  {
35.    WHILE( $SL!=END$ ) //search inside path segment
36.    {
37.      FLAG= SEARCHNODE(SL); //whether any node present in path segment
38.      IF(FLAG=="YES")
39.      {
40.         $I_p$ =getIP(BI); //get IP address of sender from beaconsed information
41.        STORE( $T_{IP}$ ); //store in  $T_{IP}$ 
42.         $N_d$ =getTip()++; //calculate node density
43.      }
44.      ++ $S_d$ ; //increment segment density
45.    }
46.    RETURN  $S_d$ ; //return segment density
47.  }

```

TIM-OFSD-ABS algorithm

```

1.  WHILE( $S_d!="NULL"$ ) //at least one node present in path segment
2.  {
3.    IF ( $T_s=="RED"$ ) //when traffic signal is RED
4.    {
5.       $SL=SL/3$ ; //segment length is now one-third of original length
6.      SEARCHNODE(SL); //find if any node present in path segment
7.      OFSD-ABS(); //beacon
8.    }
9.    ELSE IF( $T_s=="YELLOW"$ ) // when traffic signal is YELLOW
10.   {
11.      $SL=SL/2$ ; // segment length is now one-half of original length
12.     SEARCHNODE(SL); // find if any node present in path segment
13.     OFSD-ABS(); // beacon
14.   }
15.  ELSE
16.  {
17.    SEARCHNODE(SL); // find if any node present in path segment
18.    OFSD-ABS(); // beacon
19.  }
20. }

```

4 Simulation Results and Performance Analysis

Network Simulator NS-2 is used in Linux environment to evaluate the performance of our proposed protocol. A 2×3 double lane road with eight junction points is used as our proposed simulation area. The range of vehicles such as minimum range 100 and maximum range 500 are taken. The average speed of vehicular nodes is considered between the range 50–80 km/h. Transmission range is taken as 300 m and packer size is 512 bytes. CBR is the traffic type used in urban scenario. MAC protocol 802.11 DCF is used for our prosed vehicular network set up.

Figure 2 shows that the proposed scheme, OFSD-ABS based CAGR has greater successfully delivered data packets rate, when segment length is greater as compared to greedy routing. It shows that in our proposed scheme the rate of delivery ratio increases very fast as compared to greedy forwarding. Figure 3 shows that delay decreases very rapidly as compared to greedy forwarding when the no of nodes increases. In Fig. 4, the proposed scheme shows rapidly decreasing routing overheads, when the no of nodes increases the curve initially decreases very rapidly, but as the no of nodes increases, the rate of decrement is steady. In Fig. 5, the average hop count can be clearly gauged. As the no of nodes increases in greedy forwarding, it fluctuates very rapidly. But, in the proposed scheme, initially the rate of fluctuation is quite higher and as the no of nodes increases, the curve increase steadily.

Fig. 2 packet delivery ratio versus no. of nodes

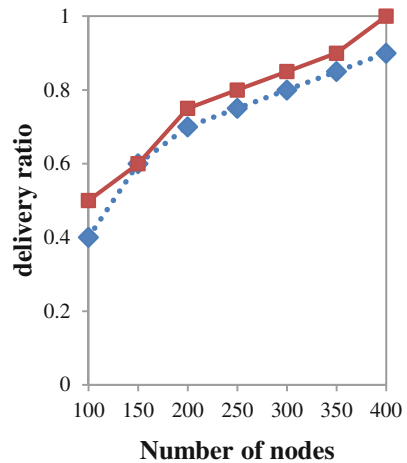


Fig. 3 Packet delay versus no. of nodes

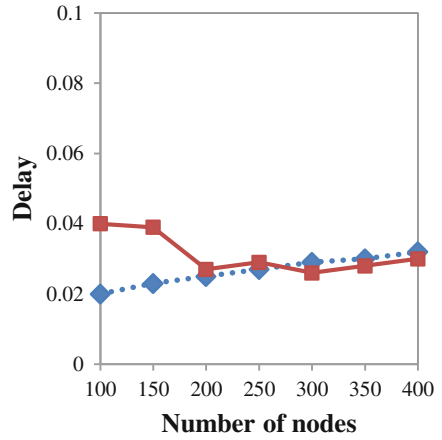


Fig. 4 Routing overhead versus no. of nodes

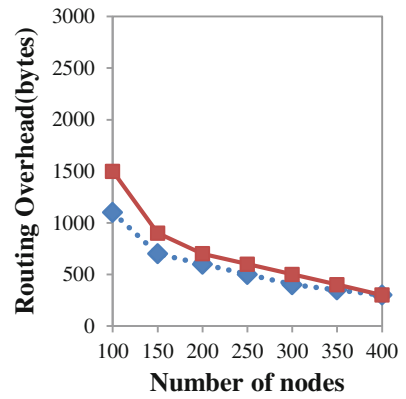
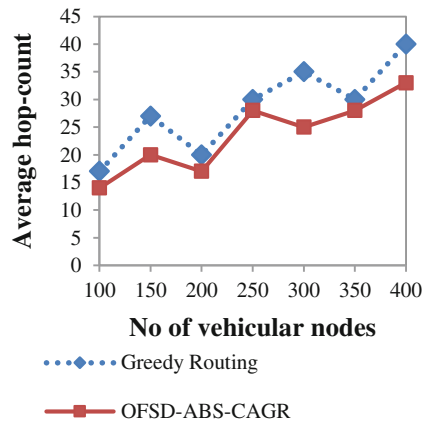


Fig. 5 Average hop count versus no. of nodes



5 Conclusion and Future Work

In this paper, On-the-Fly Segment Density (OFSD) in Adaptive Beaconsing System (ABS) based Connectivity-Aware Geocast Routing (CAGR) protocol is proposed to ensure connectivity during dissemination of critical information. The performance of the proposed scheme has been evaluated in terms of packet delay, average hop count, packet delivery ratio, routing overload and results compared with greedy forwarding. The performance analysis shows that the proposed scheme worked well in urban vehicular environment. In future, authors will explore the comparative study of the proposed scheme with some existing geocast routing approaches in the literature and develop mathematical models for result validation.

References

1. Fan Li, Yu Wang, "Routing in vehicular Ad Hoc Networks: A Survey", IEEE Vehicular Technology Magazine, Pages 12–22, June 2007.
2. S. Al-Sultan, M. Al-Doori, A. H. Al-Bayatti, & H. Zedan, "A comprehensive survey on vehicular Ad Hoc network" Journal of Network and Computer Applications, Elsevier, 37(1), pp. 380–392, 2014.
3. G.M.T. Abdalla, M. A. Abu-Rgheff, and S. M. Senouci, "Current Trends in Vehicular Ad-hoc Networks," IEEE Global Information Infrastructure Symposium, Morocco, July 2007.
4. M. Khatri, S. Malhotra, "An insight overview of issues and challenges in Vehicular Ad Hoc Networks", Journal of Global Research in Computer Science, Volume 2, No. 12, pp. 47–50, December 2011.
5. J. J. Chang, L. Yh, W. Liao, C. Ic, "Intersection-based routing for urban vehicular communications with traffic-light considerations", *J Wireless Commun IEEE* 19, pp.82–88, 2012.
6. C. Chen, Y. Jin, Q. Pei and N. Zhang, "A connectivity-aware intersection-based routing in VANETs," *EURASIP Journal on Wireless Communications and Networking*, Mar. 2014.
7. C. Sommer, O. K. Tonguz, F. Dressler, "Traffic information systems: efficient message dissemination via adaptive beaconsing," *Communications Magazine, IEEE*, vol.49, no.5, pp.173–179, May 2011.
8. R. Reinders, M. Eenennaam, G. Karagiannis, G. Heijenk, "Contention window analysis for beaconsing in VANETs," *Wireless Communications and Mobile Computing Conference (IWCMC), 7th International*, pp.1481–1487, July 2011.

Investigation and Analysis of Energy Efficiency in Distributed Antenna System: Technology Towards Green Communications

Seetaiah Kilaru, S. Padmaja and K. Venugopal Reddy

Abstract Distributed Antenna Systems has the potential to get higher Spectral and Energy efficiencies. This paper focused on implementation of DAS system in moderate and low load modes. It compares DAS with CAS and proved that Energy efficiency is excellent in DAS. We considered different load scenarios and investigated power consumption parameter and proposed a novel DAS implementation to reduce the overall power consumption per day. The results revealed that the proposed system efficiency is far ahead than CAS. We achieved 27 % energy savings too with the guarantee of high speed and capacity.

Keywords Distributed antenna structure • Energy efficiency • Spectral efficiency

1 Introduction

From the past decade, the cellular subscribers are increasing in a rapid way. Subscribers are using both voice and data services. Hence, the operators are increasing the number of Base stations (BS's) which leads to higher consumption of energy and carbon footprint of cellular networks. Investigation and improvisation of energy efficiency became challenging task for researchers with respect to reducing the operational cost and also to reduce the global carbon emissions. BS's alone are consuming much of the energy [4] (approximately 55–60 %) and hence there is a need to introduce green cellular network at base station. From [18], Internet Communication technology (ICT) is consuming more than 3 % of the world

S. Kilaru (✉) · S. Padmaja · K.V. Reddy
Jawaharlal Nehru Technological University, Hyderabad, India
e-mail: dr.seetaiah@gmail.com

S. Padmaja
e-mail: padmajasakkuru@gmail.com

K.V. Reddy
e-mail: gopalreddy0071@gmail.com

electricity consumption. A considerable research happened till now to address this problem by introducing the green cellular network. We can achieve this in two ways. As a first approach, we should introduce smart hardware system which can save power and as a second approach, we should create an intelligent mechanism to handle cellular traffic.

Wireless network efficiency can be decided by two factors, they are spectral efficiency (SE) and energy Efficiency (EE). In [10], the authors found that the considerable SE is achieved in 2015 when compared to 2010. The LTE and LTE-A technologies are playing vital role to increase the SE. Traditional cellular architecture follows the principle of Centralized Antenna System (CAS). In this system, there is a maximum possibility of wastage of power and poor network coverage. To avoid this, if we tried to increase the transmitting power, cell edges experience severe interference. To overcome the disadvantage of CAS, research is centered towards Distributed Antenna System (DAS). With this technology, we can increase the spectral efficiency throughout the region without deploying the extensive base stations.

In DAS architecture, central base station accommodates number of antennas and micro cell type antennas are spread over the regions which are also called as Radio Remote Units (RRU). These RRU's are under the controlled of BS but have their processor, Up/downlink converter, low noise amplifier and other required components [19]. BS and RRU's are connected with each other with the help of optical fiber [5]. This distribution of RRU's will reduce the distance between User Equipment (UE) and nearest serving point. Hence, the transmitter power is reduced which in turn reduces the interference. DAS will help to provide Uniform Coverage [11], to increase the capacity and will increase the SE [15, 16]. This technology will also reduce path loss [6]. In [12, 14], they proved that DAS causes low power consumption than CAS.

Multiple Input and Multiple Output is also one of the promising technology to improve SE [17]. The recent research work is focusing on application of MIMO in DAS [1, 8, 15, 20]. The EE of MIMO is not good than SISO in CAS, especially at smaller distances. This paper will explain the possibility of EE in MIMO and DAS. In cellular network, Power Amplifiers (PA) will consume most of the energy and hence if we reduce the number of antennas, power consumption will also reduce. We should make PA switch off randomly based on existed traffic load conditions. In [3], they proposed EE and SE of MIMO in CAS, in [10], they compared both CAS with DAS and concluded that DAS has good power efficiency. In [13], they investigated optimum location identification and required number of RRU's to achieve EE and SE. In [8], they proposed a method to increase EE of DAS. In [2, 7, 19], the contributions concluded that the performance of DAS was remarkable when compared to CAS.

DAS is the combination of central base station with several RRU's. We should consider the out-of-cell interference where RRU's are distributed over a region. This factor influence Signal-to-Interference plus Noise Ratio (SINR). To the best of my knowledge, the researches on the implementations of DAS under critical issues

like path loss, shadowing and fading are not addressed properly. This paper focused to address this problem using LTE-A standard simulator.

2 System Model

2.1 Downlink Model

The following Fig. 1 represents the DAS with I cells and each cell contains one BS. Each BS has central transmit antenna $M_{c,i}$ ($i = 1, 2, \dots, I$) and R_i RRU's, each RRU is equipped with $M_{r,i}$ antennas. The number of served users in cell is K_i and each user has $N_{k,i}$ received antennas (each user $k \in \{1, 2, \dots, K_i\}$). The I/O relation on kth user in cell i is given by

$$y_{k,i} = H_{k,i}x_{k,i} + H_{k,i} \sum_{u=1, u \neq k}^{K_i} x_{u,i} + \sum_{j=1, j \neq i}^I H_{k,i}^j X_j + n_{k,i}$$

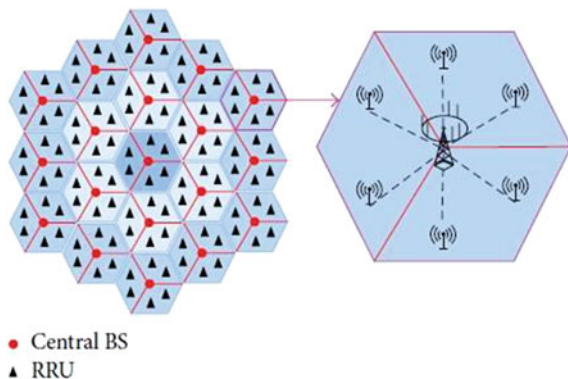
where, $y_{k,i}$ is the received signal and $H_{k,i}$ is the channel matrix between user k and all transmit antennas of cell j.

$x_{k,i}$ is the scheduled and pre-coded vector and is given as $f_{k,i}s_{k,i}$. Here, we used proportional fair scheduler and the users are selected based on greedy algorithm. Here, $s_{k,i}$ is the symbol vector and $f_{k,i}$ the pre-coded vector. If $p_{k,i}$ is the power vector [8], we can define pre-coded matrix as

$$F_i = \hat{H}_{k,i}^H (\hat{H}_{k,i}^H \hat{H}_{k,i})^{-1} \text{diag}(p_{k,i})^{\frac{1}{2}}$$

where, F_i is the sum of all pre-coded vectors. Here, to achieve high gain Multi User MIMO is used. In [21], authors recommended Zero-Forcing (ZF) [9] to get high gain. RRU's are spread throughout the cell and hence channel matrix could be decoded as explained in [8].

Fig. 1 DAS with 19 cells



The pass loss between antenna n in cell j and user k in cell i is given as

$$L_{k,i}^{n,j} = \max(70, 128.1 + 37.6 \log_{10}(\frac{d_{k,i}^{n,j}}{1000 \text{ m}}))$$

where $d_{k,i}^{n,j}$ is the distance between antenna n in cell j and user k in cell i .

3 Methodology

For simulation purpose, we defined following parameters. They are (Table 1).

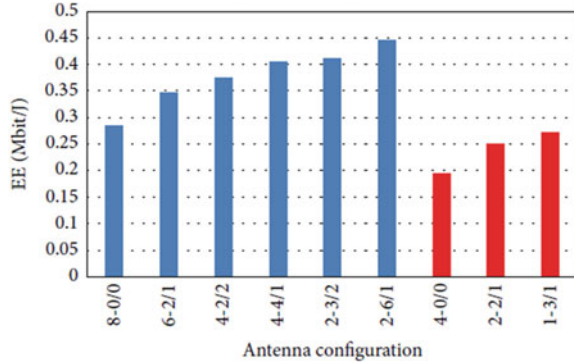
To reduce daily power consumption, we are considering two sleep modes; they are low-load side and moderate load side. The number of activated antennas will experience 3 loads (high, moderate and low). We have to find the optimal number of activated antennas to increase EE and power consumption. Antenna switch off decision can be determined from 2 considerations. The first one is handling high data rates with respect to load and secondly, number of antennas ready for switched off to obtain decrease in power consumption. Here, detailed discussions about these 2 considerations are explained in 2 steps.

Step 1 EE is achieved only when system throughput L is maximized [1] and transmit power should be reduced and L is maximum only when there are

Table 1 Simulation parameters

Parameter	Value/measurement
Carrier frequency	2.6 GHz
Number of UE in high load	36
Number of UE in low load	6
Cell radius	500 m
Channel model	Flat relay fading
Bandwidth	15 kHz
Maximum power	45 dBm
Noise spectral density	-164 dBm/Hz
Number of simulations positions	20
Number of UE in moderate load	18
CSI feedback	Moderate-perfect
UE speed	3 km/h
Standard deviation	4 dB
Receiver antenna combiner	MMSE
Number of UE	1
Transmission strategy	Zero-forcing
User scheduler	Greedy algorithm
Number of base stations	19

Fig. 2 EE with 8 and 4 antennas configuration



more activated antennas [6, 18]. In [5, 19], they suggested less distance between UE and BS. By increasing number of RRU's, we can reduce transmitted power [20]. By increasing number of antennas throughout the cell; we can achieve EE and is shown in following Fig. 2.

Step 2 For power saving one should use optimal number of antennas. In [11], they suggested to reduce number of active antenna to achieve EE. We calculated total power consumption as defined by [11] and set an EE degradation based on cell load. Power transmission is minimized by reducing number of active antennas. EE can be achieved by increasing number of active antennas. Achieving these two efficiencies are impossible with similar environment. Hence, trade-off is required between EE and power saving.

4 Results

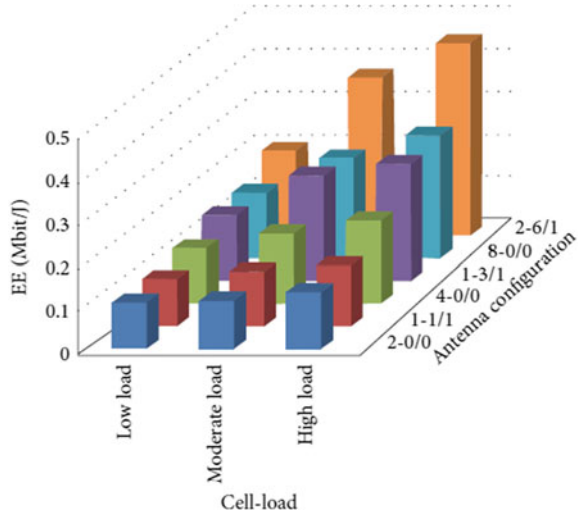
The following Fig. 3 represents EE for various load factors and with different antenna configurations.

These results explained the dependency of EE on both load constraints and antenna configurations. DAS with more number of antennas increased energy efficiency of the system. The load 2-6/1 configuration is the high load scenario with excellent EE. The EE of 2-6/1 is 56 % more when compared to 8-0/0.

With respect to energy saving, the following analysis computes daily power consumption.

1. During regular operation, i.e. without switch off process, let us consider 2 antennas are at BS and 6 antennas are distributed, therefore

Fig. 3 EE with different loads and antenna configurations



$$\begin{aligned}
 \text{Power consumed per day} &= (\text{number of antennas active}) \cdot (\text{power per each antenna}) \cdot 24 \\
 &= 126817 \times 24 \\
 &= 30.44 \text{ KW/day}
 \end{aligned}$$

- By proposed DAS scheme, in moderate load hours, only half of the antennas are activated for 5 h (1 in centre and 3 RRU at edges) and during low load scenario, only one antenna at centre and one RRU is used, then

$$\begin{aligned}
 \text{The resultant power consumed per day} &= 12 (126817) + 5 (743) + 7 (441.75) \\
 &= 22.03 \text{ KW/day}
 \end{aligned}$$

Hence, we observed that the proposed system reduced 8 KW/day when compared to existing system.

5 Conclusion

This paper focuses on EE of DAS network and performance of proposed system is analyzed. It also proved that, the distribution of antennas over region will increase the efficiency. Results showed that the dependency of load over DAS will considerably affect the performance of the system. We studied the tradeoff between EE and power saving and then we proposed DAS configuration in moderate and low load configurations. From the calculations, we achieved 8 KW/day power saving which is 26 % saving for proposed configuration when compared to traditional configuration.

References

1. X.-H. You, D.-M. Wang, B. Sheng, X.-Q. Gao, X.-S. Zhao, and M. Chen, "Cooperative distributed antenna systems for mobile communications," *IEEE Wireless Communications*, vol. 17, no. 3, pp. 35–43, 2010.
2. X. Ge, C. Cao, M. Jo, M. Chen, J. Hu, and I. Humar, "Energy efficiency modelling and analyzing based on multi-cell and multi-antenna cellular networks," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 4, pp. 560–574, 2010.
3. Kilaru, S., Prasad, Y. A., Kiran, K. S., & Chandra, N. S. (2014). Design and Analysis of Heterogenous Networks. *International Journal of Applied Engineering Research*, 9(17), 4201–4208.
4. Kilaru, S. (2014). Ability of OFDMA in Handling Interference of Femto Cells Under Random Access Process. *Journal of Engineering Science and Technology Review*, 7(2), 133–136.
5. Kilaru, S., Harikishore, K., Sravani, T., Anvesh, C. L., & Balaji, T. (2014, August). Review and analysis of promising technologies with respect to Fifth generation networks. In *Networks & Soft Computing (ICNSC), 2014 First International Conference on* (pp. 248–251). IEEE.
6. T. Zhang, C. Zhang, L. Cuthbert, and Y. Chen, "Energy efficient antenna deployment design scheme in distributed antenna systems," in *Proceedings of the IEEE 72nd Vehicular Technology Conference Fall (VTC '10)*, pp. 1–5, Ottawa, Canada, September 2010.
7. O. Onireti, F. H'eliot, and M. A. Imran, "On the energy efficiency-spectral efficiency trade-off of distributed MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3741–3753, 2013.
8. G.-H. Chen, C.-M. Yu, and C.-C. Huang, "A simulation study of a distributed antenna-based CDMA system," in *Proceedings of the 7th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*.
9. N. Jindal, "Antenna combining for the MIMO downlink channel," *IEEE Transactions on Wireless Communications*, vol. 7, no. 10, pp. 3834–3844, 2008.
10. S. Schwarz, R. W. Heath, and M. Rupp, "Multiuser MIMO in distributed antenna systems with limited feedback," in *Proceedings of the IEEE Globecom Workshops (GC Wkshps '12)*, pp. 546–551, IEEE, 2012.
11. C. He, G. Y. Li, B. Sheng, and X. You, "Energy and spectral efficiency of distributed antenna systems," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '13)*, pp. 3225–3229, IEEE, April 2013.
12. L. Dai, "Distributed antenna system: performance analysis in multi-user scenario," in *Proceedings of the 42nd Annual Conference on Information Sciences and Systems (CISS '08)*, pp. 85–89, Princeton, NJ, USA, March 2008.
13. S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089–1098, 2004.
14. M. Etoh, T. Ohya, and Y. Nakayama, "Energy consumption issues on mobile network systems," in *Proceedings of the International Symposium on Applications and the Internet (SAINT '08)*, pp. 365–368, August 2008.
15. H. Kim, S. R. Lee, C. Song, and I. Lee, "Optimal power allocation for energy efficiency maximization in distributed antenna systems," in *Proceedings of the IEEE International Conference on Communications (ICC '13)*, pp. 5769–5773, June 2013.
16. R. Heath, S. Peters, Y. Wang, and J. Zhang, "A current perspective on distributed antenna systems for the downlink of cellular systems," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 161–167, 2013.
17. M. Trivellato, F. Boccardi, and F. Tosato, "User selection schemes for MIMO broadcast channels with limited feedback," in *Proceedings of the IEEE 65th Vehicular Technology Conference (VTC '07)*, pp. 2089–2093, April 2007.

18. Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: a survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 4, pp. 524–540, 2011.
19. A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–217, 2004.
20. T. Chen, Y. Yang, H. Zhang, H. Kim, and K. Horneman, "Network energy saving technologies for green wireless access networks," *IEEE Wireless Communications*, vol. 18, no. 5, pp. 30–38, 2011.
21. G. Fettweis and E. Zimmermann, "ICT energy consumption trends and challenges," in *Proceedings of the 11th International Symposium on Wireless Personal Multimedia Communications*, pp. 1–6, 2008.

A Novel Trust Based Access Control Model for Cloud Environment

Pratap Kumar Behera and Pabitra Mohan Khilar

Abstract Cloud computing is a service oriented technology which offers the services (IaaS, PaaS, and SaaS) as a utility over the Internet. Since cloud computing is one of the most popular form of Internet application, the resources and services in cloud environment is more vulnerable to security threats and attacks. In order to protect the cloud environment from malicious users, we proposed a novel trust based access control model. The proposed model authorize the user based on user trust value before entering to cloud environment. The user must be trusted before accessing the resources and the resources must be trusted before providing the services to the user. In this paper, we evaluate the trust value of both user and cloud resources. The user trust value is evaluated based on the user behaviour parameter and the resource trust value is evaluated based on the Service Level Agreement (SLA) parameter. If the trust value of both users and cloud resources are more than their threshold value then they are considered as trusted. We implement the proposed model using java and oracle as database server. The implementation result shows the trust value of different type of users and CSP and compare with the QoS model. The proposed model performs better than QoS model in terms of Rate of Successful Transaction (RST).

Keywords Authorization · Access control · SLA parameter · Cloud computing

P.K. Behera (✉) · P.M. Khilar
Indian Institute of Technology Roorkee, Roorkee, India
e-mail: pratapbehera.bls@gmail.com

P.K. Behera · P.M. Khilar
National Institute of Technology Rourkela, Rourkela, India
e-mail: pmkhilar@nitrkl.ac.in

1 Introduction

Cloud computing is a distributed computing paradigm which provides the services to the users on pay-per use basis [1]. The main aim of designing the cloud computing system is to provide a scalable, on-demand services to the end users in a cost effective manner. Users do not need to be worry about the installation of high cost application on their system. In addition to this, the users do not maintain their own physical infrastructure for high performance computing and can obtain the required services on-demand from cloud server [2]. The services provided by the CSP is called cloud services and these services are infrastructure as a service (IaaS), platform as a service (PaaS), software as a service (SaaS). However, all of the above services can be deployed in four different cloud computing models. The four deployment models are private, public, hybrid and community cloud [3].

Since users outsourced their data and computation to the cloud servers, they may lose the physical control over their data and computation. Loss of physical control means the users are unable to resist the certain type of threats and attacks. In order to secure the users data and computation in cloud environment, the CSP should protect the cloud server from various threats and attacks. Before the users outsourced their data and computation, they must trust completely on the cloud service provider and the cloud service provider must trust on the users before providing the services to the users. So both the users and CSP must be trusted among themselves before their interaction. The user trust on the CSP based upon the security parameter and the quality of service. The CSP trust on the user based on the user behaviour parameter. If the user behaviour is malicious then the user is not considered as a trusted user.

Before we establish the trusted user and trusted CSP, first we must evaluate the trust value of both user and cloud resources. There are various parameter to be considered before we evaluate the trust value of user and cloud resources. From users perspective, we should consider those parameters which involves to identify the trusted users in order to achieve all security goals. The main security goals for cloud computing are: Confidentiality, Integrity, Availability, and Accountability.

The rest of the paper is organised as follows. Section 2 describe the existing work on various access control mechanism and trust based models for cloud environment. Section 3 describe the problem statement of our research work. Section 4 describe the proposed trust based model, algorithm for authorization, trust management module and trust evaluation strategy. Simulation results are presented in Sect. 5. Finally, Sect. 6 concludes the paper.

2 Related Work

Cloud computing is one of the most popular form of internet application, which faces a lot of security threat and attacks. There are several vulnerabilities present in the cloud environment by which the attacker and malicious user can get a chance to

introduce the attacks. Since the cloud computing provide the on demand and scalable service hence, the environment is highly dynamic. Therefore, the traditional security mechanism cannot satisfy the security requirements of cloud computing. One of the most fundamental and important key technique that can meet the security requirements in cloud environment is the access control technology [4].

Access control technology is one of the fundamental security requirements of cloud computing in order to protect cloud resources from unauthorized access and legal users malicious activity. An access control mechanism is a set of methods and policies which grants or deny the access right to the user [5]. The main goal of access control mechanism is to restrict the users from performing any unauthorized activity to protect the information. There are wide variety of models, methods and policies proposed for designing the access control system. Each access control system has its own attributes, methods, and functions based on set of policies.

We review all existing access control models that can be applied for specific environment and also cloud environment to achieve the users security requirement. Mandatory access control model (MAC) [6] is an access control policy in which a subject or request initiator can perform some sort of operation on a particular object or resource. When a subject or user attempts to access an object or resources an authorization rule is enforced to determine whether the access can take place by examining the security attributes. Discretionary access control model (DAC) [7] is an access control policy which determines the owner of an object. The owner decides who is allowed to access the service based on users identities. But the above two models are traditional models and can only applied in specific environment like operating system, data base management system. Since, cloud computing is highly dynamic and openness in nature the above model is not effective for achieving the users security requirements.

Some researchers analyze the dynamic requirements for cloud environment and introduce the Role based access control model (RBAC) [8] to the cloud environment. RBAC is an access control policy determined by the systems rather than by the owner itself. RBAC model can only applied within a closed network and it is based on identification. It only checks the user identity and roles assigned to users and based on this role it checks user is authorized or not. However, RBAC model fails to check the malicious activity done by the authorized users. To protect the legal user from doing malicious activities, trust mechanism concept is arised which evaluates the user trust value in order to authorize the user.

Grandison and Sloman [9] have surveyed several existing trust models and they explained the trust as the firm belief in the capability of an entity to act consistently, securely and reliably within a specified context. They also claim that the trust is composed of several attributes such as reliability, truthfulness, dependability, security parameter, timeliness, Quality of Service (QoS) and accountability in the context of an environment. User trust based access control model (UTBAC) [10] evaluates the trust value based on the user behaviour information in order to authorize the user. Guoyuan et al. [10] proposed a mutual trust based access control (MTBAC) model in which the user and cloud service node both trust to each other before their interaction. Paul Manuel [11] proposed a trust model for cloud computing environment

based on the quality of service. In this model, the trust value is evaluated for cloud resources based on quality of service parameter. But only the user can get to know whether the CSP is trusted or not, but there is no point to check the user performs any malicious activities or not. Because, if the user performs any malicious activity then the resources in CSP may be affected so that the performance of the CSP is degraded and less trusted.

In this paper, we proposed a novel trust based access control model which authorize all the users based on their trust value, before entering to the cloud environment for accessing the resources. The proposed model evaluates the trust value of both users and cloud resources. In this paper, we evaluate the trust value of user based upon the user behaviour parameter and the cloud resource trust value based upon SLA parameter.

3 Problem Description

The main aim is to evaluate the user trust value and cloud resource trust value based upon user behaviour parameter and SLA parameter respectively. Let us consider U_i where $i = 1, 2, \dots, n$ as n number of cloud users and R_j where $j = 1, 2, \dots, m$ as m set of cloud resources present. We have to identify trusted users $TU_k \subseteq U_i$ and the trusted resources $TR_l \subseteq R_j$ based on the user behaviour parameter and the quality of service respectively. TU_k and TR_l are set of trusted users and cloud resources respectively.

4 Proposed Model

We proposed a novel trust based access control model in which the user request is passed through various sub modules to complete the authorization process. Figure 1 shows the proposed model is deployed and running inside cloud service provider (CSP).

4.1 Authorization Process

Before accessing to any resources from the cloud provider, first the user submit the QoS requirements such as security, cost, computing power and networking speed to the cloud provider. The user and cloud provider may negotiate among themselves about the cost and quality of service for final agreement. This agreement is called as service level agreement (SLA). The following algorithm describes about the process of user authorization in order to access the resources.

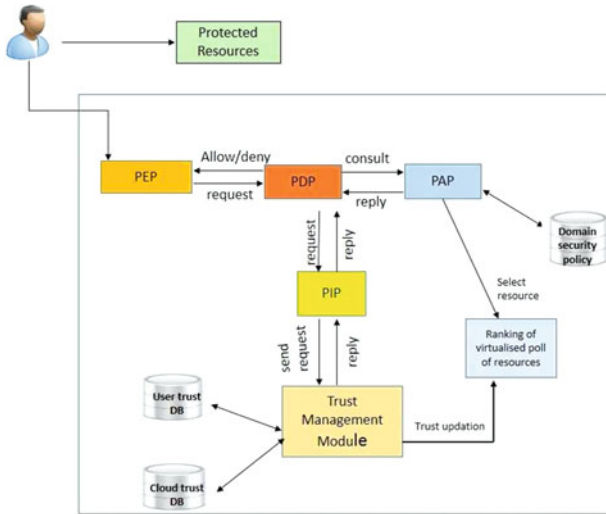


Fig. 1 Authorization model

Algorithm: The following algorithm describes about the process of user authorization in order to access the resource.

step-1: Policy enforcement point (PEP) accepts the user request and sends it to policy decision point (PDP).

step-2: PDP first checks all the user credentials and if it is correct then send it to policy information point (PIP).

step-3: PIP send the request to trust management module (TMM) for obtaining user trust value.

step-4: After getting required information from TMM, PIP send it to the PDP.

step-5: PDP consults with the policy administrative point (PAP) by sending all information coming from PIP and PEP.

step-6: Finally, PAP checks and compares all the collected information of user with domain security policy and resource database. If the user trust value is more than the user trust threshold value then the user is allowed to access otherwise rejected.

step-7: if user is allowed, then PAP identify the resources based on users request type, select the most trusted resource out of all available resource and send it to the PDP.

step-8: PDP sends resource ID to the PEP and if rejected then send deny message.

step-9: PEP forward it to the user with allow or deny message.

step-10: Finally user can able to access its service and executes their jobs or processes.

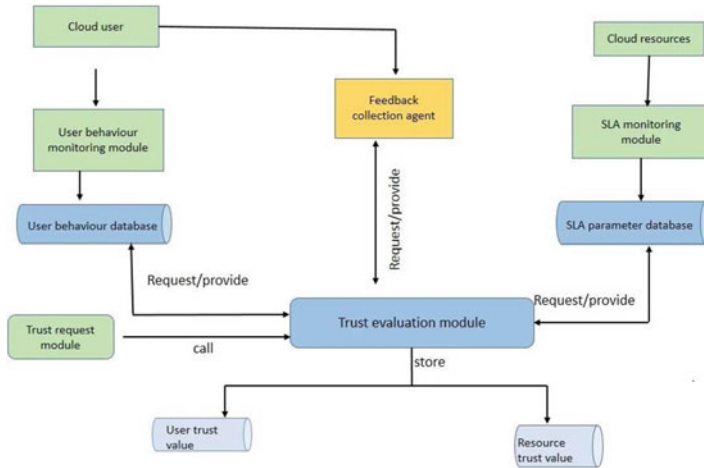


Fig. 2 Architecture of trust management module

4.2 Trust Management Module

This module is always running inside the CSP to monitor the user behaviour and the quality of service provided by the resources. Figure 2 shows the TMM is composed of several sub-modules which is involved for the evaluation of trust value for the user and the cloud resources.

4.3 Trust Evaluation Parameter

Trust evaluation parameter is considered for evaluation of trust values of the users and cloud resources. It consists of user behaviour parameter and SLA parameter.

User behaviour parameter: User behaviour parameter is used to evaluate the trust value of the user. The different user behaviour parameters are discussed as follows:

Bogus Request Rate (BRR): Bogus requests are dummy requests or huge amount of nonsensical request send to the cloud server for consuming the cloud resources intentionally for denial of service attack. Dummy request is used for introducing DOS attack by performing bandwidth starvation. So, this parameter is mainly used for achieving the availability of the cloud computing. Let R_k is the number of nonsensical or dummy requests and R_a is the total number of requests made by a user in a unit time interval.

$$BRR = \frac{R_k}{R_a} \tag{1}$$

Resource Affected Rate (RAR): This parameter is used to achieve the reliability of cloud computing. Resource affected rate is the percentage of resources affected out of the resources the user is accessing in a unit time interval. Resources may be any hardware or software resources such as system files, CPU registers, memory location, etc.

Let RA_k is the amount of resources affected and R_b is total number of accessible resources in a unit time interval.

$$RAR = \frac{RA_k}{R_b} \quad (2)$$

Unauthorized Operation Rate (UAR): Unauthorized operation is the illegal operation performed by the user for the purpose of stealing or modifying the data or computation by introducing some attacks. Illegal operation may be modify the content of system files, change the address of program memory etc.

Let UA_k is the number of unauthorized operation performed and R_c is total operation in a unit time interval.

$$UAR = \frac{UA_k}{R_c} \quad (3)$$

SLA parameter: SLA parameter is used to evaluate the trust value for cloud resources. If the CSP is unable to meet the requirements of any user according to the service level agreement (SLA), then the CSP would be less trusted. The parameter which is involved for evaluation of trust value for cloud resources is explained below:

Turnaround Efficiency (TE): Turnaround time is the exact time between the submission of a job by a user and delivery of the completed job to the user. It is promised by the cloud service provider to the user during the service level agreement. This actual turnaround time is normally different from the estimated turnaround time. Turnaround efficiency of a resource R_k (TE) is the average of turnaround efficiency over all the jobs submitted during the period T. Let TAT_{act} is the actual turnaround time and TAT_{est} is the estimated turnaround time promised by the cloud service provider. TE for a job by resource (R_k) =

$$\frac{TAT_{est}}{TAT_{act}} \quad (4)$$

Turnaround efficiency of a resource R_k (TE) is the average of the turnaround efficiency over all the jobs submitted during the time period T.

$$ATE(R_k) = \frac{\sum_{i=1}^n TE_i}{n} \quad (5)$$

Resource Availability (RAV): Availability is the degree to which the system must be functional or operational when it is accessible. If the resources are affected by the attackers or the malicious users which consumes the resources intentionally, then the

resources are unavailable and unable to process by the users request. Let us assume that R_1, R_2, \dots, R_k are the cloud resources. For each $k = 1, 2, \dots, m$. Let N_k denotes the number of jobs submitted to cloud resource R_k over a time period T . Out of N_k jobs submitted to R_k , let A_k denotes the number of jobs accepted by the resource R_k over a time period T .

$$RAV(R_k) = \frac{A_k}{N_k} \quad (6)$$

Rate of Successful Transaction (RST): RST is defined as the total number of jobs executed successfully by the resource R_k . It is also called as the success rate. RST of a cloud resource is a measure of successful completion of accepted jobs by the cloud resource [12]. Out of A_k jobs accepted by resource R_k , let C_k denotes the number of jobs completed successfully by the resource R_k over the period T .

$$RST(R_k) = \frac{C_k}{A_k} \quad (7)$$

Correctness of result (COR): Correctness of result is used to define the data integrity in cloud computing. Data may be modified or corrupted during the poor network latency or any hardware and software failure. Data precision loss might happen due to obsolete computing resources.

Let D_k is the number of jobs which preserved the data integrity or correctness of output of the C_k jobs.

$$COR(R_k) = \frac{D_k}{C_k} \quad (8)$$

4.4 Trust Evaluation Strategy

We evaluate the trust value of the users and cloud resources after finishing their interaction in a unit time interval. There may be several interaction in this time interval. First we evaluate the trust value in current time window and evaluate the average trust value by using previous average trust value and current time window trust values. Let t_n and t_{n-1} be the time interval for the current time window and previous time window respectively. The formulas for evaluating the user trust (UT) and average user trust (AUT) value are shown below:

$$UT = 1 - (W_1 * UAR + W_2 * BRR + W_3 * RAR) \quad (9)$$

$$AUT = \alpha * (UT)_{t_n} + (1 - \alpha) * (AUT)_{t_{n-1}} \quad (10)$$

where UT and AUT are the user trust value and average user trust value respectively. W_1, W_2, W_3 are the weight parameters of UAR, BRR and RAR respectively.

The formula for evaluating cloud resource trust (CT) value and average cloud resource trust value (ACT) are shown below:

$$CT = Q_1 * TE + Q_2 * RAV + Q_3 * RST + Q_4 * COR \tag{11}$$

$$ACT = \alpha * (CT)_{t_n} + (1 - \alpha) * (ACT)_{t_{n-1}} \tag{12}$$

where CT and ACT are the resource trust value and average resource trust value respectively. Where Q_1, Q_2, Q_3 and Q_4 are the weight parameters of TE, RAV, RST and COR respectively. α and $1 - \alpha$ are the weight values for t_n and t_{n-1} time intervals.

5 Simulation Experiment and Results

We implement the proposed model by using Java and Oracle as the database server. We assume the probability values for different weight parameters. According to the priority of the security requirement in our experiment, we consider the probability values of weight parameter for W1, W2, W3 as 0.5, 0.2 and 0.3 respectively. Similarly we assume the probability value of weight parameter for Q1, Q2, Q3, Q4 as 0.1, 0.2, 0.2 and 0.5 respectively. We show the trust value of different type of users and CSP such as good and malicious in Fig. 3a, b respectively. We compare our proposed model with the QoS model in terms of RST of cloud resources which is shown in Fig. 4a.

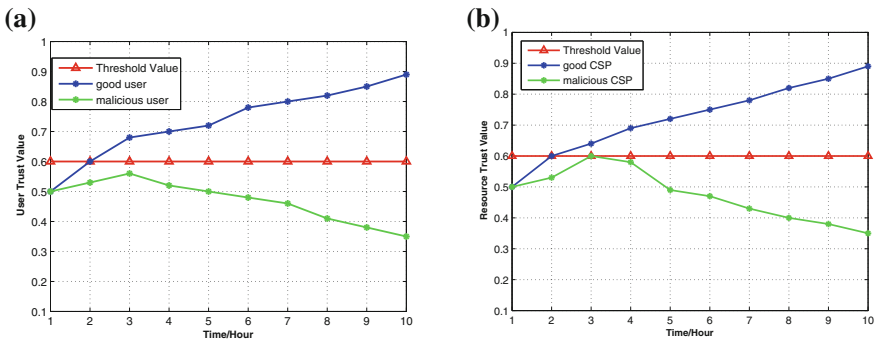
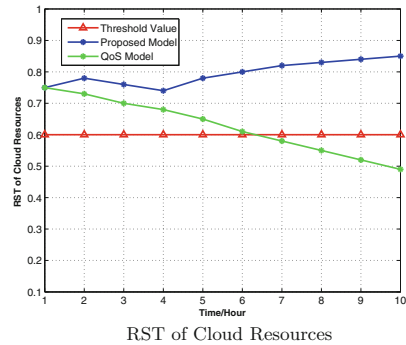


Fig. 3 Trust value of users and CSP. a Trust value of user. b Trust value of CSP

Fig. 4 Comparison of two model



6 Conclusion and Future Work

In this paper, we proposed a trust based access control model which authorize the user and protects the cloud resources from the malicious activities. It considers both user behaviour parameter and SLA parameter in order to evaluate the trust value of the users and cloud resources respectively. We evaluate the trust value of all the users in a unit time interval and identify the users according to their trust values. The users whose trust value is more than their threshold value is allow to access the required service. We evaluate the trust value of different CSP such as good and malicious CSP. The trust value in good CSP is always more than the malicious CSP. Finally, we compute the rate of successful transaction of the cloud resources of our proposed model and compare with the existing QoS model. The proposed model has more RST values than QoS model.

In future, we can extend this model and adding some new module, parameter to implement in multi domain cloud environment.

References

1. Xiao, Zhifeng, and Yang Xiao. Security and privacy in cloud computing. *Communications Surveys & Tutorials*, IEEE 15, no. 2 (2013): 843–859.
2. S. K. Panda, P. K. Jana. Efficient task scheduling algorithms for heterogeneous multi-cloud environment. *The Journal of Super Computing* 71(4): 1505–1533 (2015).
3. Mell, Peter, and Tim Grance. The NIST definition of cloud computing. (2011).
4. Samarati, Pierangela, and Sabrina Capitani de Vimercati. Access control: Policies, models, and mechanisms. In *Foundations of Security Analysis and Design*, pp.137–196. Springer Berlin Heidelberg, 2001.
5. Younis, Younis A., Kashif Kifayat, and Madjid Merabti. An access control model for cloud computing. *Journal of Information Security and Applications* 19, no. 1 (2014): 45–60.
6. <http://en.wikipedia.org/wiki/Mandatory-access-control>.
7. Samarati, Pierangela, and Sabrina De Capitani Di Vimercati. “Access control: Policies, models, and mechanisms.” *Lecture notes in computer science* (2001): 137–196.

8. Sandhu, Ravi S., Edward J. Coyne, Hal L. Feinstein, and Charles E. Youman. Role-based access control models. *Computer* 29, no. 2 (1996): 38–47.
9. Grandison, Tyrone, and Morris Sloman. A survey of trust in internet applications. *Communications Surveys & Tutorials, IEEE* 3, no. 4 (2000): 2–16.
10. Lin, Guoyuan, Danru Wang, Yuyu Bie, and Min Lei. MTBAC: A mutual trust based access control model in Cloud computing. *Communications, China* 11, no.4 (2014): 154–162.
11. Manuel, Paul. A trust model of cloud computing based on Quality of Service. *Annals of Operations Research* (2013): 1–12.
12. Gupta, P., Kumar Goyal, M., Kumar, P., & Aggarwal, A. (2013). Trust and reliability based scheduling algorithm for cloud IaaS. In *Lecture notes in electrical engineering: Vol. 150. Proceedings of the third international conference on trends in information, telecommunication and computing* (pp. 603607).

Live News Streams Extraction for Visualization of Stock Market Trends

Vaishali Ingle and Sachin Deshmukh

Abstract The live news data is vital role in the movement of stock prices. The real time unstructured data is generated through electronic and online news sources. Text mining is used for preprocessing of news stories from web sources. The visualization of stock trends can be correlated with actual market prices. The proposed analysis on current news stories helps to predict stock trends. Other techniques like tagging of stock related terms can be added for improvement in results. Stock market trends can be captured with help of this technique.

Keywords Text mining • Stock market • News streams • Word cloud

1 Introduction

The unprocessed documents are input to text mining [1] system to generate various types of output such as patterns, connections and trends. There is increasing growth in raw data by means of social media, blogs and images. The sources [2] of unstructured and semi structured data include the world wide web, online news and forums, chats, digital repositories, email and blogs. If the information is not having proper structure [3], it is called “Unstructured”. For example, even a company meeting email file has many elements embedded in it for example date and time, the message, links, author etc. To search and analyze with database query is easier than finding answers within unstructured data.

There is a strong yet complex relation between the market trades [4] and the finance news. Breaking news at every instant can change the response towards a company. The traders and investors have regular access to the online news and it

V. Ingle (✉) · S. Deshmukh

Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University,
Aurangabad, India

e-mail: ingle_123@rediffmail.com

S. Deshmukh

e-mail: sndeshmukh@hotmail.com

© Springer India 2017

D.K. Lobiyal et al. (eds.), *Proceedings of the International
Conference on Signal, Networks, Computing, and Systems*, Lecture Notes
in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_30

297

can change their decision for investing in a specific company. As the number of web sources for online news is growing, it is a tedious job for investors to filter out the required news from total news available. In this paper, real time news from news sources like Google News available in the web domain is collected to predict the stock market trends. In this we also have calculated significant correlation between the relevant news and the original stock price found at Stock Exchange.

In this paper Sect. 2 overviews literature of various machine learning techniques used for stock market predictions along with textual representations. Section 3 describes the data preparation. Section 4 investigate result as well as discuss their inference. Section 5 highlights conclusion and furthermore concise on future directions for the research.

2 Literature Review

Named Entities, Bag of Words and [5] Noun Phrases are various forms of Text illustration. As a result of this analysis, Artificial Neural Network [6] approach is a better technique in correlating the structure of [7] stock values and its dependent factors more exactly than many other statistical techniques.

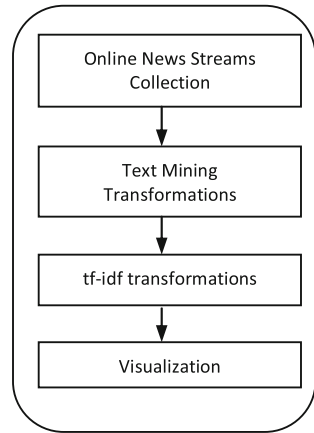
Stop-words like ‘at’, ‘by’ are removed from the text data for formation of Bag of Words [5]. Textual illustration is created by the left over text. Nouns are recognized with lexicon [3]. Later syntactic rules on adjacent parts of speech are applied to get Noun Phrases. Semantic lexical hierarchy on Noun Phrases presents Named Entities. Out of these three Named Entities provide a more abstract illustration. Past data for selected time period is input to the majority of the machine learning algorithms. The stock market movements are determined by carrying out regression analysis. Afterwards Stock prices are classified as up, down and neutral. Table 1 shows examples of classification with machine learning techniques.

It was found that the genetic algorithm and evolution strategies have performed almost evenly. Genetic algorithm [8] consistently outperformed regression model, backpropagation algorithm in predicting the stock market values.

Table 1 Classification of previous algorithmic research

Algorithm	Classification	Source material	Examples
Naive Bayesian	positive, negative, up, down, neutral	News articles	Thomas and Sycara, 2002
SVM	rise, drop, neutral	News articles	Fung 2002
Genetic algorithm	positive, negative	Chat room postings	Lavrenko 2001

Fig. 1 Data preparation stages



3 Data Preparation

There are stages for data preparation as shown in Fig. 1. At stage one we collect online news streams, then for second stage apply basic text mining transformations for data cleaning purpose and third stage find correlation between terms and finally visualization with word cloud.

3.1 Data Aggregation

Google News, Yahoo finance selects current news from about 25000 publishers worldwide which are registered. In this work online news from sources such as Yahoo Finance, Google News and Google Finance etc. are collected. We aggregated all the news [9] in a text corpus. Following table shows list of currently implemented websources (Table 2).

Table 2 Summary of web sources [9] with maximum number of items per feed

Source name	Items	URL	Auth	Format
GoogleBlogSearchSource	100	http://www.google.com/blogsearch	–	RSS
GoogleFinanceSource	20	http://www.google.com/finance	–	RSS
GoogleNewsSource	100	http://news.google.com	–	RSS
NYTimesSource	100	http://api.nytimes.com	x	JSON
ReuterNewsSource	20	http://www.reuters.com/tools/rss	–	ATOM
YahooFinanceSource	20	http://finance.yahoo.com	–	RSS
YahooInplaySource	100+	http://finance.yahoo.com/marketupdate/inplay	–	HTML
YahooNewsSource	20	http://news.search.yahoo.com/rss	–	RSS

3.2 Pre-processing of Data

Plain text is preprocessed by use of transformations like [10] Standard English stop words removal, numbers and punctuation elimination, stemming.

Further word frequency of total words in given corpus is calculated to change the count by the significance of that word. TF-IDF weights [11] computation is used for counting of word score. We can determine whether word is present or absent in a document with these positive scores.

4 Experimental Findings

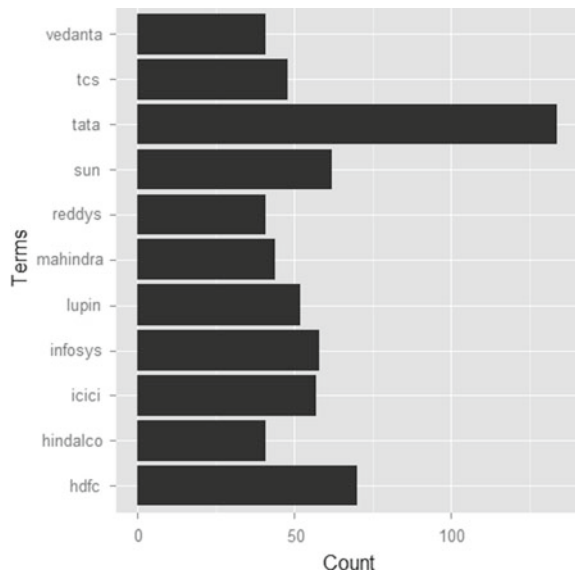
A word cloud is a diagram drawn using text mining method. It emphasizes the most frequent words in provided textual data. Each word height indicates frequency of word in whole document. Word cloud is one of the most powerful visualization tools. They are easy to share as well as understand. Usually tables are not that impactful and attractive as word cloud.

The above word cloud clearly shows that “tata”, “infosys”, “sun” and “icici” are the current market gainers. The results are correlated with the live stock market data. Our method found to have approximately 80 % accuracy. The stock market

Fig. 2 Word cloud related stock news streams



Fig. 3 Bar chart of extracted data



trends can be predictable; it is closely associated with real stock price progress (Figs. 2 and 3).

The prediction algorithm [12] can be associated to a model to predict stock trends as Up, No change and Down. The data can be stored in relational databases [13] to find out the business intelligence [14] out of it.

5 Conclusion and Future Directions

At initial stage extracted results and stock market trends are matching. The stock news data collected weekly is found useful in our technique but we need to improve data collection and analysis methods. Further research points towards use of dimension reduction techniques for feature extraction and Hidden Markov Models can be used for prediction of future trends.

References

1. Feldman, R., Sanger J.: The Text Mining Handbook :Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press New York (2007)
2. Unstructured Data. Proceedings of the 33rd international conference on Very large data bases (VLDB '07), Vienna, Austria, pp. 1045–1056, September 23–28, 2007.
3. Buneman P., Davidson S., Fernandez M., Suciu D.: Adding structure to unstructured data. Database Theory-ICDT'97, (1997) 336–350
4. Anandarajan, M., Anandarajan, A. ed. : e-Research collaboration: Theory, techniques and challenges. Springer Science & Business Media, 2010.
5. Schumaker R., Chen H.: Textual analysis of stock market prediction using financial news articles. AMCIS 2006 Proceedings (2006) 185
6. Majumder M., Hussian M.A.: Forecasting Of Indian Stock Market Index Using Artificial Neural Network. Information Science (2007) 98–105.
7. Vui C. S., Soon G. K., On C. K., Alfred R., Anthony P.: A review of stock market prediction with Artificial neural network (ANN). Control System, Computing and Engineering (ICCSCE), Proceedings -2013 IEEE International Conference (2013) 477 – 482
8. Venugopal K. R., Srinivasa K. G., Patnaik L. M. : Fuzzy Based Neuro - Genetic Algorithm for Stock Market Prediction. Soft Computing for Data Mining Applications .Volume 190 of the series Studies in Computational Intelligence. Springer Berlin Heidelberg (2009)139–166
9. Annau M.: Short Introduction to tm.plugin.webmining. <https://cran.r-project.org/web/packages/tm.plugin.webmining/vignettes/ShortIntro.pdf> (Accessed on 15 July, 2015)
10. Nagar A., Hahsler M.: Using Text and Data Mining Techniques to extract Stock Market Sentiment from Live News Streams. IACSIT Press, Singapore, Vol. 47 (2012) 91–96
11. Weiss S. M., Indurkha N., Zhang, T., Damerou, F.: Text mining: predictive methods for analyzing unstructured information. Springer Science & Business Media, (2010).
12. Chowdhury S. G., Routh S., Chakrabarti S.: News Analytics and Sentiment Analysis to Predict Stock Price Trends. International Journal of Computer Science and Information Technologies, Vol. 5, Issue 3 (2014) 3595–3604.
13. Mansuri I. R., Sarawagi S.: Integrating unstructured data into relational databases. Proceedings of the 22nd IEEE International Conference on Data Engineering (ICDE'06), (2006) 29–29
14. Rao R.: From unstructured data to actionable intelligence. IT professional. Vol. 5, Issue 6 (2003) 29–35

Categorization of Cloud Workload Types with Clustering

Piotr Orzechowski, Jerzy Proficz, Henryk Krawczyk
and Julian Szymański

Abstract The paper presents a new classification schema of IaaS cloud workloads types, based on the functional characteristics. We show the results of an experiment of automatic categorization performed with different benchmarks that represent particular workload types. Monitoring of resource utilization allowed us to construct workload models that can be processed with machine learning algorithms. The direct connection between the functional classes and the resource utilization was shown, using unsupervised categorization approach based on moving average for finding a class number, and k-means algorithm for clustering.

Keywords Workload categorization · IaaS · Cloud computing · Clustering · Cloud load prediction

1 Introduction

Growing expenses of purchase and maintenance of IT infrastructure cause organizations and companies to seek cheaper ways of handling their services. Recently, a very popular way is to deploy services on rented infrastructure, for example using the cloud computing model named Infrastructure as a Service (*IaaS*) [1].

P. Orzechowski (✉) · J. Proficz
Academic Computer Centre in Gdansk (CI TASK), Gdańsk, Poland
e-mail: p.orzechowski@task.gda.pl

J. Proficz
e-mail: j.proficz@task.gda.pl

H. Krawczyk · J. Szymański
Department of Computer Systems Architecture,
Gdańsk University of Technology, Gdańsk, Poland
e-mail: hkrawk@eti.pg.gda.pl

J. Szymański
e-mail: julian.szymanski@eti.pg.gda.pl

Popularity of IaaS in cloud computing is currently growing very fast. Compound Annual Growth Rate (CAGR) for 2010–2015 is equal to 26.2 % [2] and would be 42.91 % [3] in 2015–2019, which shows the speed of growth for this market. IaaS grows all over the world. At the moment, the cloud computing market is dominated by big companies, such as: Amazon EC2, Google Cloud, Microsoft Cloud, or Rackspace, but there is also a group of smaller providers.

In Poland, the growth of cloud computing market is forecasted to increase from \$63 million in 2012 to \$223.5 million at the end of 2017 [4]. This data shows the importance of the cloud computing market, and the IaaS model specifically. On the Polish market we can also find local IaaS providers, such as: Home Cloud, and Oktawave.

In the IaaS model, the providers offer their clients the possibility to rent server machines, either with predefined hardware parameters (ready to use flavors), or with the parameters customized to meet individual requirements. During the process of hardware configuration, the infrastructure providers give the clients the possibility to deploy some applications. The number of available applications depends on the infrastructure provider and varies from several (Oktawave) to hundreds (Amazon). The combination of prepared flavors, the possibility to define custom flavor, and hundreds of applications ready to be deployed, give the cloud's clients a whole set of ready-to-use environments in which they can start their own services.

The problem arises when a client has to choose which flavor is appropriate for him. Usually the client has some application usage forecast, e.g. the expected number of requests per second for a web application, or the size of data stored in a caching application like Redis, but frequently the user does not have any specific knowledge about the actual resource utilization of these applications. It is therefore more suitable to recommend him or her which resources will be necessary for performing a particular tasks in the cloud environment. To do this in an automatic way, categorization of workload types and modeling of their resource utilization is required. The thesis of this research is that construction of a representation of workloads based on monitoring of their resource utilization (cloud measured parameters) will allow us to perform automatic workload categorization which will agree with a manually created categorization. Workload identification is an important feature for cloud monitoring. It can be applied e.g. for load prediction in which a model of a particular workload allows us to allocate the required resources and optimize their utilization.

On the other hand, users want to declare resource requirements themselves, and in most cases the declared requirements are overestimated. This causes many servers in a cloud to be significantly underutilized. The currently accepted solution is to dynamically migrate virtual machines between physical servers, but it is difficult to precisely predict the resource utilization and choose an appropriate server to run a virtual machine. In such a case, the knowledge of workload type and the possibility to estimate workload's resource utilization could help decrease the number of working servers and, in consequence, reduce power consumption.

The next section provides the summary of related works, then we present our proposition of the workload classification based on the functional characteristics. Section 4 presents the experimental verification of the proposed categorization based

on observing workloads' resource utilization. The last section draws the conclusions and presents possible future works.

2 Related Works

The problem of analyzing resource utilization of different kinds of workloads in cloud computing has been already discussed in literature [5–7]. In most cases authors focused on proper workload allocation or changes in resource utilization during the workload's life cycle.

In [8] authors present a classification method named FBWC model (Feedback-Based Workload Classification) and a classification algorithm TSRSVM (Training Set Refresh Support Vector Machine). Research was based on 22 metrics from 16 synthetic benchmarks monitored during execution and five predefined classes: CPU-Intensive, Memory-Intensive, I/O-Intensive, Network-Intensive, Compound.

Rao et al. [9] describe the distributed iBalloon framework, which is able to forecast resource utilization generated by a 2-layers web application. The framework uses an algorithm utilizing a reinforcement learning method, and it employs four metrics as the input parameters: CPU and memory utilization, number of I/O operations and the used swap space. The authors claim iBalloon provides results close to the optimal ones (minimizing the server response time), and show their solution's superiority to the other methods, such as Adaptive Proportional Integral (PI) or Auto-Regressive-Moving-Average (ARMA).

Another framework designed for automatic configuration of VMs is VCONF [10]. The authors present the results of experiments for three types of workloads: e-commerce (TPC-E), online transaction processing (TPC-C) and an application server (SPECweb). The tests were performed for various hardware configurations and for different parameters of the workloads, e.g. different number of users. The framework is based on the measurements of three parameters: CPU utilization time, load of virtual CPUs and amount of used RAM memory. The algorithm is based on an artificial neural network and Markov chains. Its use for dynamic VMs reconfiguration led to an increase in the cloud performance.

Zhang and Figueiredo [11] presents a classifier based on PCA, which was used for optimizing resource utilization (reported gain of about 20 %) while maintaining comparable application performance. The above results were achieved by mixing different types of workloads on the same nodes, in a way which allowed the workloads to use their specific resources exclusively. Four types of the workloads were distinguished: CPU-intensive, I/O-intensive, memory-intensive and idle. For the learning purposes, the following parameters were used: CPU System/User, Network Bytes In/Out, Disk IO IN/OUT, Swap IN/OUT, which were collected using 14 different benchmarks.

In [12] the authors present an analysis of workload characteristics measured in the Google cloud. The goal of the paper was to propose a workload classification based

on the resource utilization. The analysis input was related to workload execution time, CPU and memory utilization.

The classification related to the software functionality was analyzed in [13]. In the research, benchmarks representing different application types, such as: application server, database, file server, etc., have been used. In the experiments, the following parameters were monitored: CPU and memory utilization, and I/O latency, related to the storage access. Two solutions were proposed for workload modeling: a neural network and a support vector machine. Both were tested against regression models, and provided more accurate results.

The paper [14] presents a new method for VM migration in a cloud. For the method analysis the authors proposed a following classification of the workloads (VMs): idle, OLTP, file server, science, application server. Using of these proposed classes showed that the migration requirements depend on the functionality realized in the migrated VM.

The related works presented either functional or automatic classifications of the workload, without any hybrid approach. We argue that merging of both of these methods provides better classification results. Moreover, we postulate using more than a few monitoring parameters, extending the standard set of CPU, memory, I/O and network utilization with more detailed measurements related to the hardware and system characteristics (e.g. cache references or memory bus utilization).

3 The Workload Categorization

In our research we propose a novel approach for computing cloud workload classification. We start with the functional approach, based on the state of the art in this field, with additional proof of using the results of unsupervised classification [15] based on the k-means clustering algorithm. The analysis of the measurements performed during execution of different benchmarks in an HPC cluster environment allows us to create groups of the most similar workloads. The whole procedure is presented in Fig. 1.

Based on the related works, as well as the review of the typical applications hosted by the cloud providers, we performed the functional analysis of the workloads. As a result, we proposed the following functional classification of tasks that are performed within our cloud:

1. Science,
2. Big Data,
3. OLTP,
4. Caching,
5. Streaming,
6. Webserving.

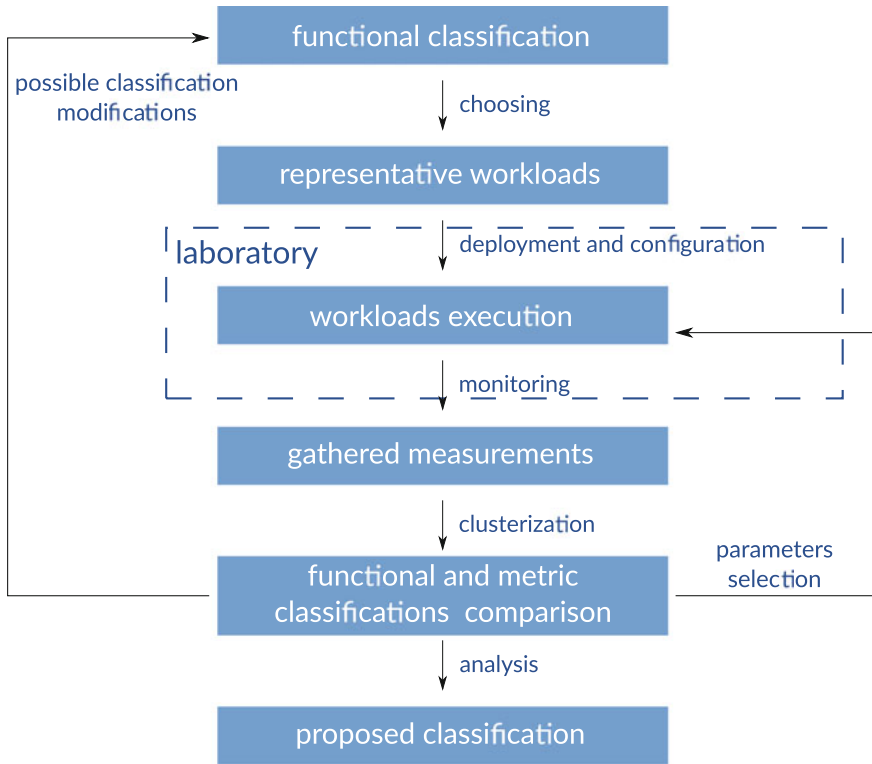


Fig. 1 Scenario of our experiment

To test the cloud load with applications of a particular type we use benchmarks that allow us to emulate the given workload. Details of the benchmarks used in the experiments have been shown in the Table 1.

The next step was preparing an appropriate laboratory, where the chosen benchmarks were deployed and configured. The experiments were performed in an HPC cluster environment: Galera+ supercomputer located at the Academic Computer Center of Gdansk University of Technology in Poland. They used six compute nodes, each equipped with 2 Intel Xeon L5640 processors (2.27 GHz, 12 MB cache) and 16 GB RAM memory interconnected by Gbit Ethernet.

For the experiments, the KVM virtualization environment was used with Linux Ubuntu Server v12.04 as host and v14.04 as guest operating systems. Only one guest VM was executed on each host node at any given time, thus all the node resources were available for the tested workloads, including 12 virtual CPUs (cores) and 12 GB RAM. The measurements were performed on the host OS, with our own tool based on the PAPI library,¹ registering 115 metrics of resource utilization related to various hardware and OS parameters, such as: processor load, L2 cache misses, used network

¹<http://icl.cs.utk.edu/papi/>.

Table 1 Set of the workloads emulators (benchmarks)

No.	Workload type	Benchmark name	No. of configurations
1	OLTP	OLTPBench [oltpbench] [16]	36
		CloudSuite—data serving [cloudsuite] [17]	
2	Science	HPCG [hpcg] [18]	9
		HPCC [hpcc] [19]	9
3	Streaming	Cloudsuite—streaming [cloudsuite] [17]	27
4	Web serving	Cloudsuite—web serving [cloudsuite] [17]	7
		WP Mark Twenty Twelve [wpmark] [20]	
5	Caching	Cloudsuite—caching [cloudsuite] [17]	27
		Mentier [mentier] [21]	27
6	Big data	PUMA Map Reduce [puma] [22]	9

bandwidth or even CPU temperature. The full list of the observable metrics can be found in [23]. The measurements were performed every five seconds and were stored in a NoSQL database (MongoDB).

4 Experiments and Results

Every measured workload was installed on a separate VM, and for some of them additional VMs with client software were prepared, e.g. for Webserving. Depending on the specificity of a workload class, there could be different parameters affecting its behavior, e.g. number of concurrent users for Webserving, types of queries for OLTP, length and bitrate of the streams for Streaming. Thus, apart from the workloads themselves, specific configurations of the parameters were defined. Table 1 presents the total numbers of such configurations used for every workload type.

For each of the configurations, dedicated tests were prepared and performed, whose results were further analyzed. During the experiments, over 136.000 samples were collected, each including measurements of the 115 hardware metrics. Furthermore, for each of the tested configurations, statistics were calculated, which included the average, minimum, maximum and standard deviation of every metric. For visualization, we used multidimensional scaling (MDS) [24], which allowed us to map the 115-dimensional feature space into a 2D chart, keeping the minimal distortion of the distances between the points.

Figure 2 presents the results of the experiments projected into the 2D space with targeted workload classes marked with particular colors. It should be noted that the functional classes are concentrated in separate groups. It suggests that in the original, higher dimensional space, the workloads also create separable groups, related to the proposed categorization. For the verification of the above assumption we

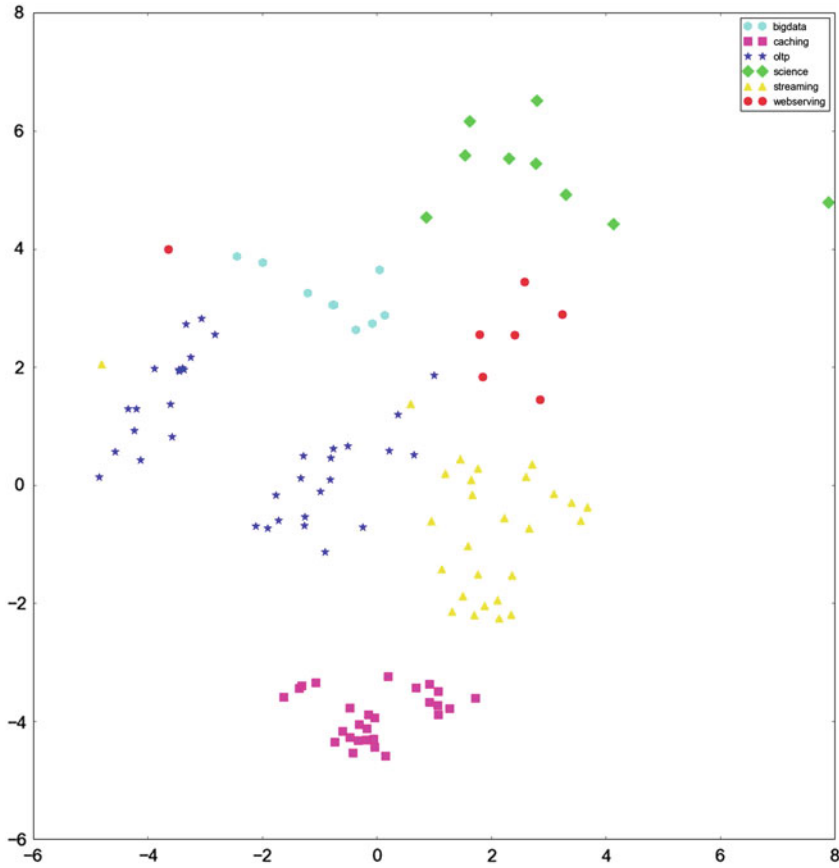


Fig. 2 Visualization of the data in 2D space, reduced with MDS

performed automatic clustering, using combination of two methods: moving average and k-means algorithms [25]. The former was used to determine the number of groups within the input data and the latter enabled partition of the workloads into the groups. The results of the analysis, mapped into 2D space (using MDS), are presented in Fig. 3.

The obtained results prove that the proposed functional classes have different characteristics related to hardware resource utilization. The automatically formed groups of workload types are consistent in 94 % of the cases, evaluated in 10-fold cross validation. However there is a number of workloads with characteristics significantly deviated from the others but belonging to the same class, what leads to misclassification. We argue that for more complex cases other approach for categorization, based on supervised learning, will be necessary. Detailed information on misclassification between particular workload types has been shown in Fig. 4.

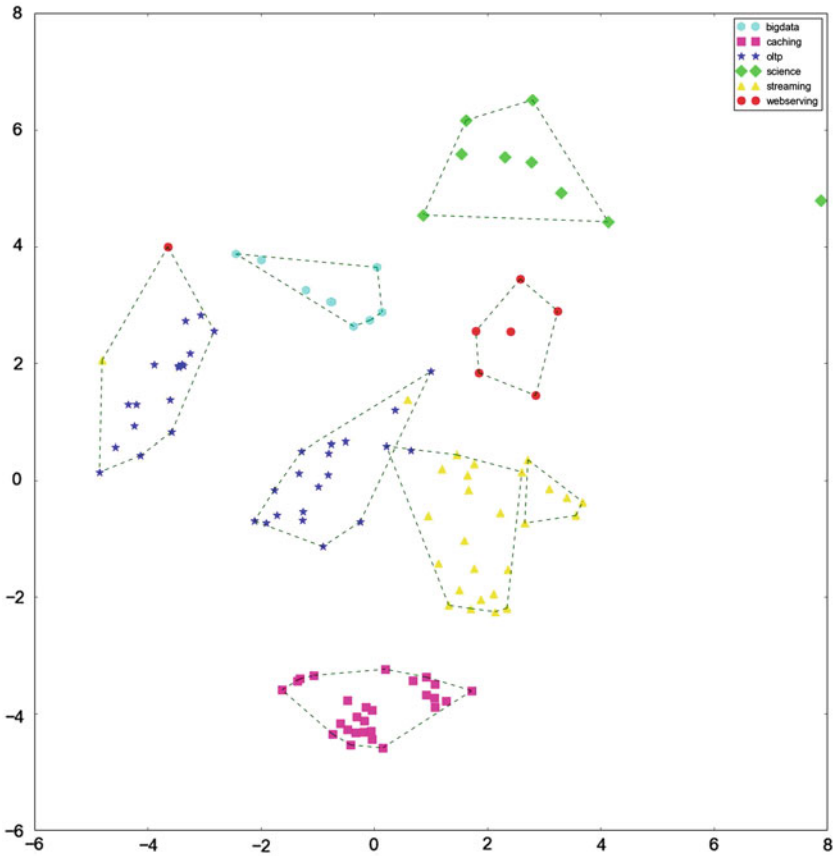


Fig. 3 Results of automatic workloads categorization with k-means, presented in 2D projected with MDS

The results of the performed experiments and their analysis proved that the proposed functional classification is consistent with the classification based on the resource utilization measurements. However we assume that some cases of the workloads could have been omitted and that additional method customization may be necessary. For example, the OLTP workloads seem to be partitioned into subclasses, and it may even be necessary to introduce completely new classes. To improve the results, additional workload-based resource utilization measurements or end user application supervised categorization [26] can be performed.

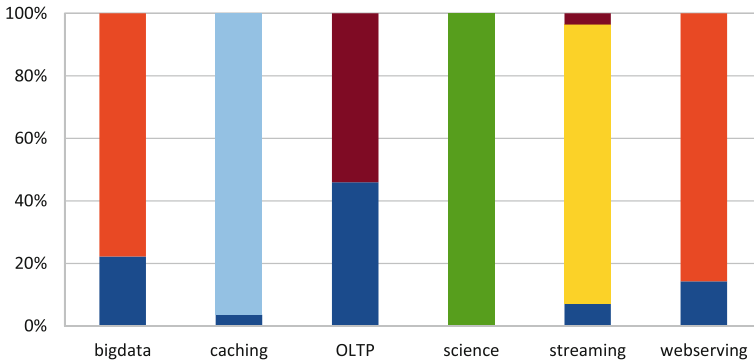


Fig. 4 Miss classifications within particular workload types

5 Conclusions and Future Works

In the paper we propose a categorization of workloads types used in our experiments and we perform their automatic functional categorization. Based on 105 resource utilization measurements we construct the parameter space, in which we performed k-means clustering. Some of the parameters may be redundant, thus in future we plan to perform a deeper analysis based on feature selection methods [27]. For the most of the cases the constructed groups of workload types are consistent with their functional assignment, which indicates that the proposed approach can be used for identification of workloads implemented in end user applications.

In the future, we plan to use categorization of workloads for cloud load prediction, where the allocation of particular resources for a given workload type can be optimized according to monitoring of the cloud resources utilization. Other future research area is related to performing the tests using real cloud environment including workloads, and integrating the proposed solution into some cloud management software (e.g. OpenStack) Additionally we plan to provide the method of performing both automatic workload categorization and prediction for a heterogeneous environment [28] where particular cloud servers can vary in available resource quantities and types.

Acknowledgments This work was carried out as a part of the research project “Recommendation Component for Intelligent Computing Clouds”, co-financed by the European Regional Development Fund, Innovative Economy Operational Programme 2007–2013, Priority Axis 1: “Research and development of state-of-the-art technologies”. The experiments were performed using high-performance computing infrastructure provided by the Academic Computer Centre in Gdansk (CI TASK).

References

1. Mell, P., Grance, T.: The nist definition of cloud computing. In: NIST Special Publication 800-145, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg (2011).
2. Markets and Markets: Cloud Computing Market: Global Forecast (2010 – 2015). <http://www.marketsandmarkets.com/Market-Reports/cloud-computing-234.html> (2010) [Online; accessed July-2015].
3. Infiniti Research Limited: Global IaaS Market 2015–2019. <http://www.technavio.com/report/global-iaas-market-2015-2019> (2014) [Online; accessed July-2015].
4. Zborowska, E.: Poland cloud services market 2012 analysis and 2013 – 2017 forecast. <http://idcpoland.pl/eng/research/published-reports/52486-poland-cloud-services-market-2012-analysis-and-2013-2017-forecast/> (2013) [Online; accessed July-2015].
5. Kundu, S., Rangaswami, R., Gulati, A., Zhao, M., Dutta, K.: Modeling virtualized applications using machine learning techniques. In: ACM SIGPLAN Notices. Volume 47., ACM (2012) 3–14.
6. Liu, H., Jin, H., Xu, C.Z., Liao, X.: Performance and energy modeling for live migration of virtual machines. *Cluster computing* **16** (2013) 249–264.
7. Rao, J., Bu, X., Xu, C.Z., Wang, L., Yin, G.: Vconf: a reinforcement learning approach to virtual machines auto-configuration. In: Proceedings of the 6th international conference on Autonomic computing, ACM (2009) 137–146.
8. Zhao, X., Yin, J., Chen, Z., He, S.: Workload classification model for specializing virtual machine operating system. In: Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. (2013) 343–350.
9. Rao, J., Bu, X., Wang, K., Xu, C.Z.: Self-adaptive provisioning of virtualized resources in cloud computing. In: Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems. SIGMETRICS '11, New York, NY, USA, ACM (2011) 129–130.
10. Rao, J., Bu, X., Xu, C.Z., Wang, L., Yin, G.: Vconf: A reinforcement learning approach to virtual machines auto-configuration. In: Proceedings of the 6th International Conference on Autonomic Computing. ICAC '09, New York, NY, USA, ACM (2009) 137–146.
11. Zhang, J., Figueiredo, R.: Application classification through monitoring and learning of resource consumption patterns. In: Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International (2006).
12. Mishra, A.K., Hellerstein, J.L., Cirne, W., Das, C.R.: Towards characterizing cloud back-end workloads: insights from google compute clusters. *SIGMETRICS Perform. Eval. Rev.* **37** (2010) 34–41.
13. Kundu, S., Rangaswami, R., Gulati, A., Zhao, M., Dutta, K.: Modeling virtualized applications using machine learning techniques. *SIGPLAN Not.* **47** (2012) 3–14.
14. Liu, H., Xu, C.Z., Jin, H., Gong, J., Liao, X.: Performance and energy modeling for live migration of virtual machines. In: Proceedings of the 20th International Symposium on High Performance Distributed Computing. HPDC '11, New York, NY, USA, ACM (2011) 171–182.
15. Hastie, T., Tibshirani, R., Friedman, J.: *Unsupervised learning*. Springer (2009).
16. Benchmark: OLTPBench. <http://oltpbenchmark.com/wiki> (2015) [Online; accessed July-2015].
17. Benchmark: Cloudsuite. <http://parsa.epfl.ch/cloudsuite/cloudsuite.html> (2015) [Online; accessed July-2015].
18. Benchmark: High Performance Conjugate Gradients. <https://software.sandia.gov/hpcg/> (2015) [Online; accessed July-2015].
19. Benchmark: HPC Challenge Benchmark. <http://icl.cs.utk.edu/hpcc/> (2015) [Online; accessed July-2015].
20. Benchmark: WPMark Twenty Twelve. <http://mtekk.us/archives/wordpress/wpmark-twenty-twelve/> (2015) [Online; accessed July-2015].

21. Benchmark: Memtier benchmark. http://redislabs.com/blog/memtier_benchmark-a-high-throughput-benchmarking-tool-for-redis-memcached (2015) [Online; accessed July-2015].
22. Benchmark: PUMA. <https://sites.google.com/site/farazahmad/pumabenchmarks> (2015) [Online; accessed July-2015].
23. Czarnul, P.: Model of a computational node in a cloud evaluation of hardware metrics. In: TASK internal report, Gdansk University of Technology (2015) 1–26.
24. Kruskal, J.B., Wish, M.: Multidimensional scaling. Volume 11. Sage (1978).
25. Hartigan, J.A., Wong, M.A.: Algorithm as 136: A k-means clustering algorithm. *Applied statistics* (1979) 100–108.
26. Kotsiantis, S.B.: Supervised machine learning: A review of classification techniques. In: *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, IOS Press (2007) 3–24.
27. Rzeniewicz, J., Szymanski, J.: Selecting features with SVM. In: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications - 18th Iberoamerican Congress, CIARP 2013*. (2013) 319–325.
28. Czarnul, P.: Modeling, run-time optimization and execution of distributed workflow applications in the jee-based beesycluster environment. *The Journal of Supercomputing* **63** (2013) 46–71.

Development of a General Search Based Path Follower in Real Time Environment

B.B.V.L. Deepak, G. Raviteja, Upasana Behera and Ravi Prakash

Abstract The path planning problem of an Unmanned Ground Vehicle in a predefined structured environment is dealt in this paper. Here the environment chosen as the roadmap of NIT Rourkela obtained from Google maps as reference. An Unmanned Ground Vehicle (UGV) is developed and programmed so as to move autonomously from an indicated source location to the defined destination in the given map following the most optimal path. An algorithm based on linear search is implemented to the autonomous robot to generate shortest paths in the environment. The developed algorithm is verified with the simulations as well as in experimental environments.

Keywords Unmanned ground vehicle · NITR map · Path planning · MATLAB simulation

1 Introduction

In automated environments, Unmanned Ground vehicles (UGV) or mobile robots are used for several applications like transportation of materials, tour for elderly or disabled etc. from one specified location to destination following the optimal path [1]. UGV navigation environment includes perception, localization and map building, cognition and path planning and motion control. Accurate path planning

B.B.V.L.Deepak (✉) · G. Raviteja · U. Behera · R. Prakash
Department of Industrial Design, NIT Rourkela, Rourkela, Orissa, India
e-mail: bbv@nitrrkl.ac.in

G. Raviteja
e-mail: raviteja6911@gmail.com

U. Behera
e-mail: upasana.20behera@gmail.com

R. Prakash
e-mail: raviprakash.sunlight@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_32

315

enables autonomous mobile robots to follow or track an optimal collision free path from start position to the goal position without colliding obstacles in its workspace [2, 3]. An ideal path planner must be able to handle uncertainties in the sensed world model, to minimize the impact of objects on the robot and to find the optimum path in minimum time especially if the path is to be followed regularly.

The objective of this paper is to provide an approach for designing a path planning methodology for UGV in a complex environment with multiple branching paths to be chosen from, in order to reach the destination.

2 Related Work

A fundamental approach for formulating and solving the path planning problem is the configuration space (c-space) approach [4]. The central idea of this approach is the representation of the robot as a single point. There are numerous types of algorithms to search and manipulate the data structure used to store maps space environment zone/work. Examples of these algorithms is a graph search algorithm consisting of best first search algorithm traversing a graph using a priority queue to find shortest collision free path shorter collisions [4]. There are several artificial intelligence based approaches such as artificial immune system [5], Particle swarm optimization [6, 7], Fuzzy logics [8] etc. to solve robot motion planning problem.

But the above mentioned techniques are not applicable line follower robots. In this study a new approach has been introduced to solve motion control of a path follower in known environments.

3 Algorithm

There are several algorithms have been developed in the past [9, 10]. The algorithm that is proposed in this paper is similar to the binary search algorithm used to find a number in data structures. This algorithm uses both global path planner as known environment is used and the coordinate values are available, and local path planner since the path follows only the black coloured road provided.

Initially the environment in which the path has to be planned is verified and studied thoroughly with the source and goal positions in the environment. Different branches and paths at every connecting node are to be identified. The path starts and continues in the direction of goal until it reaches a node and a decision has to be made. When there are multiple branching paths to be chosen at a particular node then the destination point coordinates are checked and the corresponding path which leads to destination is chosen.

At this point choosing the branching path at a node uses the technique similar to that of the potential field approach. The path that leads to the goal position has attractive effect and among such paths the shortest one is chosen. For example

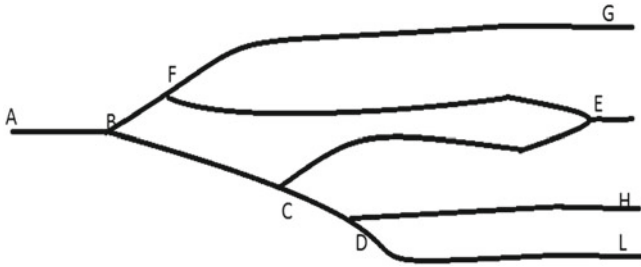


Fig. 1 General environment with multiple possible paths

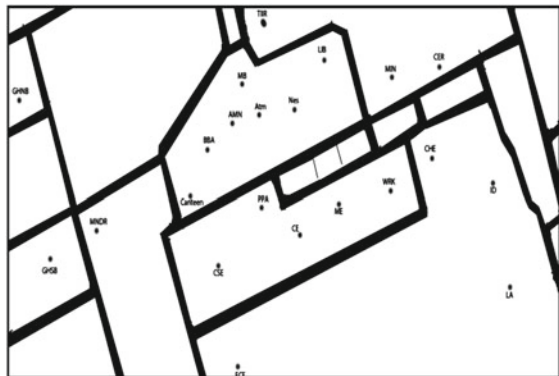
consider the environment shown in Fig. 1. With the source located at node A and goal is located at node E.

Now the choice has to be made at B whether to go towards F or C. Considering the branch from C to E there is a curve and has more length, so it will be better to choose the path along the F branch. After reaching F it will choose the branch towards E the goal position but does not continue towards G. When the environment is being analyzed the curve lengths are identified and hence helps in choosing the path in case of multiple choices as explained earlier.

4 Experimental Analysis

To validate the efficiency of the proposed algorithm, a known environment is considered as shown in Fig. 2. This environment is obtained from google maps, which represents the roadmap of National Institute of Technology, Rourkela. This roadmap contains the locations of various departments and other important places in the campus. The problem statement of the current research work is to decide the

Fig. 2 NIT Rourkela road map



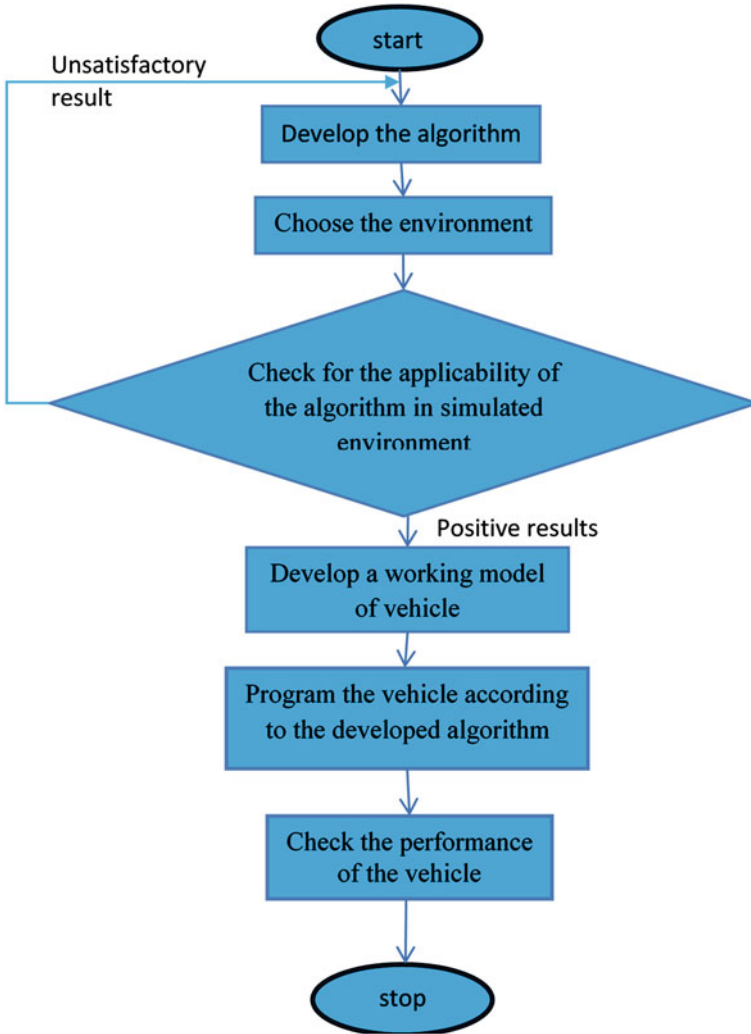


Fig. 3 Flow chart of the proposed algorithm

direction of the path by choosing the optimal of the available paths between the given departmental locations.

The roadmap is saved as an image file and then imported in MATLAB 2012b. For the various departments, corresponding coordinates are specified. First the map is studied and the location of the different departments are noted and verified. The path planning is done according to the process described in the flow chart as depicted in Fig. 3.

For simulation purpose the Unmanned Ground Vehicle (UGV) or the differential mobile robot considered here is represented as (robot x, robot y). Similarly all the

departments and other prominent locations on the map are represented by the x and y coordinates.

The source and goal positions can be defined by

Case 1: entering the locations with the keyboard during program execution.

Case 2: choosing the desired positions on the roadmap by using mouse.

Once the source and the goal positions are attained, their respective position with one another are checked and the direction of the motion is determined. Here the whole map is checked and verified globally. If the source coordinate value is less than the goal position coordinate values, then the UGV moves in the forward direction towards left to right. Otherwise the UGV moves from right to left. Once the source and destination positions are given on the map, the UGV will move towards its goal position. If there are branching paths along the way, then the robot decides its direction motion according to the shortest path traversal criteria. The most important factor to be considered here is that always the shortest path is chosen for deciding the direction of motion of the robot towards the goal.

When the vehicle is moving in the given path it will travel only along the road dedicated for its travel. It means that, the UGV always follows the black line indication on the road.

5 Results and Discussions

As described in the previous section, case 1 the source and goal position are entered with the mouse as shown in Fig. 4 and the simulation result containing the path for the corresponding positions is shown in Fig. 5. In this case it can be seen that goal coordinate value is higher than source position and hence the robot moves forward from left to right.

Here the source is at mechanical (ME) department with coordinates $x = 805$ and $y = 303$. And the destination is ceramic (CER) department with coordinates

Fig. 4 Source and goal positions

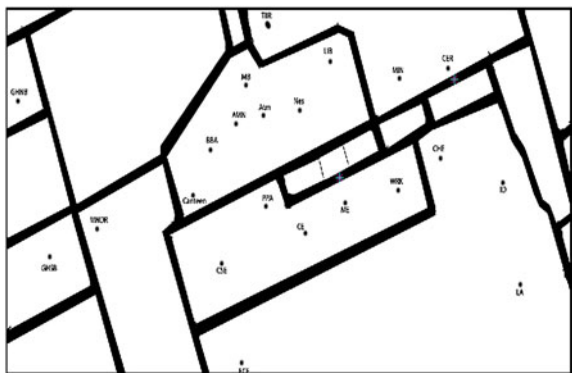


Fig. 5 Path for the positions indicated in case 1

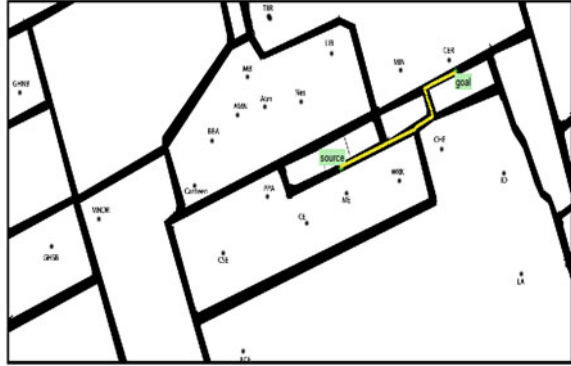


Fig. 6 Command window showing source and goal positions

```

Command Window
case 2
choice: 12
source position is at
ID
g choice: 21
goal position is at
bba
Warning: Image is too big to fit on screen; displaying at 67%
> In imutils\private\initSize at 72
  In imshow at 259
  In names at 158
>> t

t =

    20.2053
    
```

$x = 1074$ and $y = 127$. Time taken for the simulation to take place is 4.95 s. Similarly as described in case 2 the source and goal positions can be entered manually using keyboard during execution as shown in Fig. 6. The path obtained for the specified locations in Fig. 6. is shown in Fig. 7

For checking the output in case 2 the source position is at ID department with coordinates $x = 1264$ and $y = 292$ and the goal position is at BBA with coordinates $x = 459$ and $y = 171$. To check the validity of the proposed algorithm in the real time environment, a differential drive robot has been considered as shown in Fig. 8. The motion analysis has been performed as per the equations obtained from past research work [11–13]. The motion equation of the developed robot is represented in Eq. (1).

$$\dot{\xi}_T = \begin{Bmatrix} \dot{x} \\ \dot{y} \\ \dot{\psi} \end{Bmatrix} = \frac{1}{2s} \begin{bmatrix} s*\cos\psi & s*\cos\psi \\ s*\sin\psi & s*\sin\psi \\ -1 & 1 \end{bmatrix} * \begin{Bmatrix} v_{Rt} \\ v_{Lt} \end{Bmatrix} \tag{1}$$

4. Brooks, R. A. (1982). Solving the find-path problem by representing 3D space as generalized cones, Artificial Intelligence Laboratory, Massachusetts Institute of Technology. AI Memo 674, May.
5. Deepak, B. B. V. L., Parhi, D. R., & Kundu, S. (2012). Innate immune based path planner of an autonomous mobile robot. *Procedia Engineering*, 38, 2663–2671.
6. Deepak, B. B. V. L., & Parhi, D. R. (2013, December). Target seeking behaviour of an intelligent mobile robot using advanced particle swarm optimization. In *Control, Automation, Robotics and Embedded Systems, International Conference on* (pp. 1–6).
7. B. B. V. L., & Parhi, D. (2012). PSO based path planner of an autonomous mobile robot. *Open Computer Science*, 2(2), 152–168.
8. Kundu, S., Parhi, R., & Deepak, B. B. (2012). Fuzzy-neuro based navigational strategy for mobile robot. *International Journal of Scientific & Engineering Research*, 3(6).
9. Pakdaman, M., & Sanaatiyan, M. M. (2009, December). Design and implementation of line follower robot. In *Computer and Electrical Engineering, 2009. ICCEE'09. Second International Conference on* (Vol. 2, pp. 585–590). IEEE.
10. Bajestani, S. E. M., & Vosoughinia, A. (2010, August). Technical report of building a line follower robot. In *Electronics and Information Engineering (ICEIE), 2010 International Conference On* (Vol. 1, pp. V1–1). IEEE.
11. Deepak, B. B. V. L., Parhi, D. R., & Jha, A. K. (2011). Kinematic Model of Wheeled Mobile Robots. *Int. J. on Recent Trends in Engineering & Technology*, 5(04).
12. Parhi, D. R., & Deepak, B. B. V. L. (2011). Kinematic model of three wheeled mobile robot. *Journal of Mechanical Engineering Research*, 3(9), 307–318.
13. Deepak, B. B. V. L., Parhi, D. R., & Amrit, A. (2012). Inverse Kinematic Models for Mobile Manipulators. *Caspian Journal of Applied Sciences Research*, 1(13).

SDN Architecture on Fog Devices for Realtime Traffic Management: A Case Study

Kshira Sagar Sahoo and Bibhudatta Sahoo

Abstract Software Defined Network has become one of the most important technology to manage the large scale networks. The separation of the control plane from the data plane in networking devices is the main idea of SDN. Currently, Open Flow is the popular SDN standard, which has a set of functionalities. In the emerging cloud scenario smart devices plays an important role. But they are facing latency and intermittent connectivity. For this fog devices are placing in-between cloud and smart devices. Fog computing is currently applying on connected vehicles, sensor network etc. This article looks into the vehicular network area as a case study where SDN architecture can apply on fog devices for enhancement of the performance and betterment of traffic management and QoS on distribution of real time data.

Keywords Fog computing · SDN · Vehicular network · Openflow · DSRC

1 Introduction

Today's data is tremendously dispersed and delivered continuously, in large volumes and to a large number of users with different devices. Fog devices provide data, storage, computation and application services to the end-users at a distributed level. Thus, the idea of fog computing is to distribute all data and place it nearer to the user, which will remove network delays and jitter associated with data transfer [1–3]. When many users are simultaneously streaming the same content in a given cell in a cellular network, each user gets his own video stream and consumes his own portion of the cellular capacity. This unicast model and video's intensive bandwidth demands can cause frequent network congestion. Again, the initial establishment of cellular

K.S. Sahoo (✉) · B. Sahoo
Department of Computer Science and Engineering,
National Institute of Technology, Rourkela 769 008, India
e-mail: kshirasagar12@gmail.com

B. Sahoo
e-mail: bibhudatta.sahoo@gmail.com

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_33

323

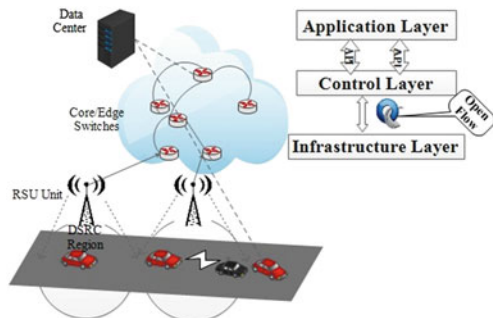
connection takes several seconds and the end-to-end latency is comparatively high [4]. DSRC (Dedicated Short Range Communication), also called 802.11P [5] is an alternate to cellular network, a short range communication services that supports in V2V environment [6]. Software Defined Network (SDN) is an emerging paradigm that makes the behavior of the network devices (such as routers/switches) programmable and allows them to be controlled by a central element, thus offering advanced customizability of network control and forwarding behaviors [7].

The rest of the paper is organized as follows. Section 2 describes the related work on content distribution. Section 3 discussed on proposed algorithm. Section 4 result analysis followed by future work and conclusion in Sect. 5.

2 Preliminaries

For large amount of content distribution, V2V model is the wiser option than V2I model. In the first SDN concept of vehicular networks and centralized control over V2V [8], RSU can serve as fog devices. By leveraging the control plane in SDN, the system can effectively collect and maintain individual vehicle states in a logically centralized way also control and optimize V2V/V2I multi-hop routing/switching. Basically, in this paper, we formulate the characteristics of the system is like this: Transmission of content can be done in two phases. In the first phase of the content server to RSUs and in the second phase Content will transfer from RSUs to vehicles. In Fig. 1, the core switches and aggregation switches will only route the content to the edge switches and edge switches are responsible to forward the content to RSUs. Then after control flooding is accomplished. RSUs will record all subscribers of a particular content service request. Each time the updates (a part of the content) will be broadcast to the subscribers. The more subscribers, the better efficiency can be achieved. Because of the intrinsic nature of the wireless communication, in a close environment the vehicles may suffer in transmission collision. Only one mode is active at a time i.e. V2V or I2V, for the above reason [9].

Fig. 1 DSRC and cellular link for content distribution along with SDN logic



2.1 Intelligent Capabilities in SDN

There are two types of entities in the SDN/OpenFlow network switches and controllers. The controller directs the switches to forward the flows. We are adding the third one, i.e. a analyzing server, to store the application information and perform traffic patterns analysis and prediction and classify the pattern which could improve the intelligence in SDN. To identify flows and detect applications, used Machine Learning (ML) techniques, which are based on statistical features. ML algorithms apply on true data set to train a classifier to classify flows, e.g. Support Vector Machine (SVM), Naive Bayes, neural networks etc. [10]. These techniques assume that the application typically sends data in some sort of pattern; these patterns can be used as a means of identification which would allow the connection to be classified by traffic class. Both supervised and unsupervised learning technique for the Internet traffic classification problem can be used. The unsupervised clustering technique use an Exception Maximization (EM) algorithm which classifies unlabeled training data into groups based on similarity [11].

3 Related Work

In [12] shows the CDS (Content Distribution Scheduling) is a NP-hard problem. CDS is a novel approach where each vehicle notify the current neighboring list to the RSU, then RSU select the sender and receiver vehicles and communicate either by using V2V or I2V channel. In this paper, we have used Type Based Content Distribution (TBCD) method [4] along with add more intelligent capabilities in SDN while forwarding the packets. The goal of our paper is to provide a content distribution model in a vehicular environment, which will provide high scalability, communication reliability under limited bandwidth using SDN paradigm.

4 System Model and Problem Analysis

For a clear understanding we discuss the notation used in the Table 1. The total number of data items D requested by a subscriber is denoted by $\{d_1, d_2, \dots, d_n\}$. The set of RSU in a city is denoted by $RSU = \{RSU_1, RSU_2, \dots, RSU_n\}$. The set of vehicles $V(t) = \{V_1, V_2, \dots, V_V(t)\}$, where $V_V(t)$ is the total number of vehicles at time t . The total number of vehicles can be grouped into either V2V or V2I mode; these two sets are denoted by $V_I(t)$ and $V_V(t)$ respectively. One vehicle has to be stay in one mode at a time; i.e. $V_I(t) \cap V_V(t) = \emptyset$ and $V_I(t) \cup V_V(t) = V(t)$. Each vehicle is having a set of request i.e. $RQ_{Vi}(t) = \{RQ_i^1, RQ_i^2, \dots, RQ_i^n\}$ where n is the total number of request send by the vehicle V_i at time t . Set of services again may be divided into satisfied

Table 1 Notations used in the algorithm

Notation	Description
$D = \{d_1, d_2, \dots, d_n\}$	Set of data items
RSU_n	Road Side Unit
SW_{Ci}	Core Switches
$V_I(t)$	Set of vehicles in I2V mode
$V_V(t)$	set of vehicles in V2V mode
$RQ_{Vi}(t)$	set of request submitted by V_i
Rq_i^j	jth request of i
SQR_{Vi}	Set of satisfied request of V_i
URQ_{Vi}	set of unsatisfied request of V_i
N_{Vi}	Set of neighbor of V_i
$DSRC_i$	i^{th} DSRC region1

request and unsatisfied request or pending request, they are denoted as SQR_{Vi} and URQ_{Vi} . So $SQR_{Vi} \cap URQ_{Vi} = \emptyset$ and $SQR \cup URQ = Q_{Vi}(t)$. Any vehicle V_i has a set of neighbor vehicles in V2V mode, they denoted by $N_{Vi}(t)$. V_{RSU} is the set of vehicles in the RSU region. To facilitate the above scenario the following set of conditions has to satisfy.

$$\{V_i | V_i \in V_{RSU} \wedge V_i \in V_I(t) \wedge d_I(t) \in URQ_{Vi}(t)\} \quad (1)$$

It tells that V_i must be in RSU region V_i is in I2V mode and $d_I(t)$ has not yet been serviced. In V2V mode the set of vehicles $SV(t) = \{SV_1, SV_2, \dots, SV_n\}$ are the designated sender vehicles. Set of data items to be transmitted by the set of vehicles is denoted by $D(SV(t)) = \{d(SV_1), d(SV_2), \dots, d(SV_n)\}$. Because of broadcast of packets, multiple data items may reach to the receiver, which may cause collision. Given a set of sender vehicles $SV(t)$, for any V_c in the V2V mode, SV_a and SV_b are the neighbors. So both must be in $SV(t)$ i.e. $SV_a, SV_b \in SV(t)$. The data collision might be occur in the following situation.

$$\{V_c | V_c \in V_{Vt} \wedge V_c \in N_{SVa}(t) \wedge V_c \in N_{SVb}(t)\} \quad (2)$$

Keeping in eye the dynamic nature of the traffic and heavy demand of the data services, it is required to enhance QoS via cooperative data exchange. So one of the objective is maximization of vehicle that either I2V or V2V mode during communication and simultaneously minimization of the packet transmission from server to the vehicle through RSUs.

5 Proposed Algorithm

When a vehicle sends a request to the server, the SDN agents keep track of vehicles' information in the cache. Then the server begins to push the content to the designated RSUs through switches. All these initial actions carried out between RSU and SDN agent is described in the Algorithm 1.

Algorithm 1 Initial action

```

 $V_{RSU}(t) \leftarrow \phi$ 
for each  $V_i \in V(t)$  do
  if RSU receives the updates periodically from  $V_i$  then
     $V_{RSU}(t) \leftarrow V_{RSU}(t) \cup V_i$ 
  end if
end for
for each  $V_j \in V_{RSU}(t)$  do
  for each  $q_{V_j}^m \in RQ_{V_j}(t)$  do
    Push the content request to the server
  end for
end for

```

Core switches send the data packet to the other switches whereas edge switches determine the updated copies and send the packets to the RSUs according to the current location of the vehicle. If the vehicle is in the DSRC region of an RSU, then it receive immediately, otherwise wait for update come from other vehicles in V2V mode. When a vehicle received a new update, it waits for a time interval W_s . It can be represented by

$$w_s = 1/d_s \quad (3)$$

where d_s is the distance from RSU. If the vehicle receives the same content during this period it will prohibits from rebroadcast. Otherwise the vehicle rebroadcast the packets to the others after W_s interval. The next action among RSU, vehicles and V2V are described in the Algorithm 2.

Algorithm 2 Content transmission from Switch to RSU and RSU to vehicle

```

for  $do$   $SQR_{V_i}(t) \subset RQ_{V_i}$ 
  if  $SW_i \in SW_c$  then
     $SW_k \leftarrow SQR_{V_i}(t)$  (where  $SW_k \notin SW_c$ )
  else
     $RSU_i \leftarrow SQR_{V_i}(t)$  (where  $SW_k \in SW_e$ )
  end if
  if  $V_{RSU_i} \in DSRC_i$  then
     $V_{RSU_i} \leftarrow SQR_{V_i}(t)$ 
  else
    wait for V2V mode
  end if
end for

```

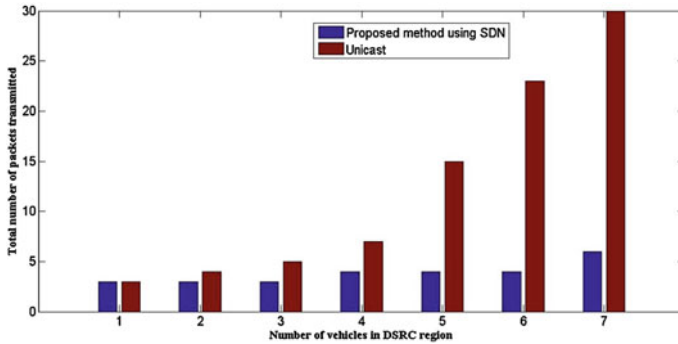


Fig. 2 Number of packets transmissions

6 Simulation Result

With a limited functionalities we have simulated this architecture using the NS3 simulator. The traffic characteristics are simulated based on Greenshield's model. This model is widely used in simulating macroscopic traffic scenario [13]. Using SDN technology in the said algorithms the total number of packet transmission from content server to the subscriber reduced drastically. Since the real time information is flowing from the content server, then would go to core switches and then to edge switches, to a RSU and finally reaches to the subscriber. So the packet is forwarded once in each hop. The graph obtained from the simulation shows that the packet is forwarded once for each hub, so the total number of packets transmitted is nearly equal in each time for content transmission which is depicted in Fig. 2. On the contrary, each subscriber, in the on-demand unicast method will retransmit once from the server. So the total number of packet transmission grows linearly with the number of subscribers increases.

7 Conclusion

For our work we have used Type-Based Content Distribution (TBCD) method [4], along with we have theoretically added ML classifier to classify the flow to support large scale real time content distribution in vehicular net-works. Theoretically, we have modeled ML classifier in our approach which our future research work. In future we will focus on the content delivery based on both ML classifier and DPI classifier. To extend this work, we will investigate the innovative approach to capture, aggregate, and analyses the fine grained real-time traffic to enhance the QoS in SDN architecture. Also current RSU model does not support multi hop V2V transmission, which is one of our future work.

References

1. fog computing is a new concept of data distribution <http://cloudtweaks.com/2013/12/fog-computing-is-a-new-concept-of-data-distribution/>, as accessed on 24th April 2015
2. Karthikeyan, B. Detecting and Isolating Distributed Denial of Service Attack in Smart Grid Systems. Diss. National Institute of Technology Rourkela, 2014.
3. Deswal, Reenu, Sambit Kumar Mishra, and Bibhudutta Sahoo. "Optimizing Power Consumption in Cloud Using Task Consolidation." *Networking and Communication Engineering* 7.4 (2015): 155-162.
4. Cao, Yi, Jinhua Guo, and Yue Wu. "SDN enabled content distribution in vehicular networks." *Innovative Computing Technology (INTECH)*, 2014 Fourth International Conference on. IEEE, 2014.
5. J. B. Kenney, Dedicated short-range communications (DSRC) standards in the United States. *Proceedings of the IEEE* 99.7 2011, pp.1162-1182.
6. D. Jiang and L. Delgrossi. IEEE 802.11 p: Towards an international standard for wireless access in vehicular environments. *Proceedings of IEEE Vehicular Technology Conference*, 2008.
7. K. Bakshi, Considerations for software defined networking (SDN): Approaches and use cases, in *Proc. IEEE Aerosp. Conf.*, Mar. 2013, pp. 19.
8. Liu, K., et al. "Cooperative data dissemination in hybrid vehicular networks: Vanet as a software defined network." Submitted for publication (2014).
9. Fujimura, Kaichi, and Takaaki Hasegawa. "A collaborative MAC protocol for inter-vehicle and road to vehicle communications." *Intelligent Transportation Systems*, 2004. *Proceedings. The 7th International IEEE Conference on. IEEE*, 2004.
10. Li, Yunchun, and Jingxuan Li. "MultiClassifier: A combination of DPI and ML for application-layer classification in SDN." *Systems and Informatics (ICSAI)*, 2014 2nd International Conference on. IEEE, 2014.
11. Erman, Jeffrey, Anirban Mahanti, and Martin Arlitt. "Qrp05-4: Internet traffic identification using machine learning." *Global Telecommunications Conference, 2006. GLOBECOM'06. IEEE. IEEE*, 2006.
12. Liu, Kai, et al. "Cooperative Data Scheduling in Hybrid Vehicular Ad Hoc Networks: VANET as a Software Defined Network."
13. P. Edara and D. Teodorovi Model of an advance-booking system for highway trips, *Transport. Res. Part C, Emerging Technol.*, vol. 16, no. 1, pp. 3653, 2008.
14. Cui, Hongyan, Xiang Liu, and Lixiang Li. "The architecture of dynamic reservoir in the echo state network." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 22, no. 3 (2012): 033127

Maximizing Network Lifetime of Wireless Sensor Networks: An Energy Harvesting Approach

Srikanth Jannu and Prasanta K. Jana

Abstract Energy preservation is very crucial in wireless sensor networks as they are operated in hostile and non-accessible areas. The use of renewable energy sources is an alternative technique for extending lifetime of a sensor network where the battery-driven sensor nodes run out of battery power faster. In this paper, we study and solve the problem of extending network lifetime by introduce energy-harvesting (EH) sensor nodes and propose a clustering algorithm to extend the network lifetime. In the proposed algorithm, we present an efficient scheme for cluster head selection by considering the locations of EH sensor nodes and all of these EH sensor nodes serve as relay nodes to the cluster heads. Simulation results and their theoretical analysis show that the proposed algorithm outperforms the existing algorithm.

Keywords Wireless sensor networks · Clustering · EH nodes · Network lifetime · Residual energy

1 Introduction

Recently energy harvesting (EH) of the sensor nodes [1–3] has drawn enormous attention for extending lifetime of the network. In this technique, EH sensor node harvests energy from the natural power sources such as solar, piezoelectric and thermal. Then, it converts that energy into electrical energy which is stored in devices such as super-capacitors to achieve almost infinite lifetime [1]. However, the deployment of large scale WSN composed of only EH sensors remains impractical due to high costs and low achievable duty cycles [4, 5]. Therefore,

S. Jannu (✉) · P.K. Jana
Department of Computer Science and Engineering, Indian School of Mines,
Dhanbad 826 004, India
e-mail: j.srikanth@live.com

P.K. Jana
e-mail: prasantajana@yahoo.com

deploying sensor nodes along with their certain percentage as EH nodes will be a cost effective system for prolonging life of WSNs [4–6]. However, using some EH nodes and clustering sensor nodes would be cost effective system for improving lifetime of the network.

In this paper, we study an effective combination of energy harvesting and clustering to create a cost effective system and propose an algorithm for maximizing network lifetime. Given the set of sensor nodes along with the set of EH nodes, the proposed algorithm selects CHs from the normal sensor nodes based on their residual energy and the distance from the EH nodes. Then, every sensor node is assigned to its nearest CH for the sake of cluster formation. Every CH collects the sensed data from its member sensor nodes and aggregates them. Then, the CH is assigned to its nearest EH node and that EH node forwards the data to the sink. Therefore, the CH consumes less energy by forwarding to the nearest EH node rather than sending it directly to the sink. As a result, the lifetime of the network increases significantly.

Several clustering algorithms have been developed in [7, 8]. In [9], the authors have proposed a cost-based energy balanced clustering and routing algorithm (CEBCRA). However, the algorithm does not consider connectivity among CHs. Voigt et al. proposed s-LEACH [10], where the authors extend LEACH protocol as solar aware. However, the authors have assumed that every sensor node of the network is solar powered which is very expensive. In [4], the authors have proposed an advanced-sLEACH protocol in which the authors select a EH node as the CH. However, the rate of charging is very sensitive to the environment [11]. It may not be practical to choose EH node as CH in most of the cases. Zhang et al. proposed clustering algorithms [12] to maximize the lifetime of the network. But, the algorithm selects fixed number of CHs in the network. The proposed method takes care of effective CHs selection based on sparsely deployed EH nodes and all EH nodes act as relay nodes for the CHs.

The rest of the paper is organized as follows: Sect. 2 shows system models. Section 3 describes the proposed clustering algorithm. In Sect. 4, simulation results are presented for verifying the performances of the proposed algorithm. Finally, Sect. 5 concludes the paper and presents several directions for future research.

2 System Models

2.1 Network Model

We assume a WSN which consists of sensor nodes and certain percentage of sensor nodes called EH nodes deployed randomly. Once deployment is completed they are assumed to be stationary. As a wireless network, the nodes do not have global information of the network. However, we consider that the nodes have the information of residual energy and the distance of its neighbor sensor nodes as in radio model [13]. Initially, each sensor node collects the local data and sends it to its

corresponding CH. Next, the CH aggregates the received local data and sends the aggregated data to the sink via an intermediate EH node. We use some terminologies which are useful to present the algorithm as follows:

- (1) $dist(S_i, S_j)$ is the Euclidean distance from the sensor S_i to S_j .
- (2) $Neighbour(S_i)$ is the set of all sensor nodes which are within the communication range of S_i .
- (3) $distbor(S_i)$ is the set of distances from all borders of the target area.
- (4) $distocen(S_i)$ is the distance from the sensor node S_i to the centre of the target area.

2.2 Harvesting Model

Solar radiation is an upper bound which depends on the environment and is determined by the maximum energy output of the solar collector E_{sol} . In this paper, we use an astronomical model to estimate the solar radiation which depends on the parameters such as the angle between the sunlight and the solar panel. If the angle of sunlight from the normal to the solar panel is θ then the effective sunlight can be proportional to $Cos\theta$ [14] which is used to estimate the solar radiation for a time period. The daily peak solar hours (PSH) can be given as follows:

$$PSH = \int_{t_{sunrise}}^{t_{sunset}} Cos\theta dt \tag{1}$$

Now, the solar energy is converted into the electric energy by the solar collector. Solar collector consists of a solar panel and a regulator. The solar panel converts photons into electric energy and the regulator regulates the output power of the solar panel P_{solar} and that energy is transferred to the storage. The solar energy of the solar collector E_{sol} is determined as follows

$$E_{sol} = PCH \times P_{solar} \tag{2}$$

In general P_{solar} is set to 1 kW per m^2 [14].

3 Proposed Algorithm

We define some terminologies which are useful to present the algorithm as follows.

Definition 1 (*Threshold Distance*) Threshold distance is a minimum permissible distance between two CHs which is used to distribute CHs evenly in the target area and it is denoted as *Threshold_distance*.

Definition 2 (*Threshold Energy*) Threshold energy is a minimum allowable residual energy of a sensor node which is required to become the CH for at least one transmission.

The foremost idea of our proposed algorithm is to choose k potential CHs from the total number of sensor nodes and it is as follows.

In this method, we use a weight function $W(S_i)$ to select CHs from normal sensor nodes by considering intra cluster distance and the residual energy of sensor node S_i ($\forall i, 1 \leq i \leq n$). The CH consumes more energy than its member sensor nodes. For that reason, we build the weight function by considering the following two factors.

- (1) A sensor node needs to hold adequate energy to become a CH. Therefore, the chance of choosing a sensor node as a CH is high when its residual energy is more. In other words,

$$W(S_i) \propto E_{resi}(S_i) \quad (3)$$

- (2) The selected CH has to run for a longer time than its member sensor nodes i.e., the average distance from any EH node EH_j of the sensor node has to be minimum than that of any of its neighbour sensor nodes. It means

$$W(S_i) \propto \frac{dist(S_i, EH_j)}{Neighbour(S_i)} \quad (4)$$

where, $Neighbour(S_i)$ and $E_{resi}(S_i)$ denotes the number of neighbour and residual energy of sensor node S_i as described in Sect. 2. To maximize the weight function we combined condition 1 and condition 2 as follows

$$W(S_i) \propto E_{resi}(S_i) \times \frac{Neighbour(S_i)}{dist(S_i, EH_j)}$$

i.e., $W(S_i) = K \times E_{resi}(S_i) \times \frac{Neighbour(S_i)}{dist(S_i, EH_j)}$

where, K is a proportionality constant. Without loss of generality, we assume that $K = 1$. Therefore,

$$W(S_i) = E_{resi}(S_i) \times \frac{Neighbour(S_i)}{dist(S_i, EH_j)} \quad (5)$$

Moreover, the next CH will be selected as follows: the new CH distance from the existing CH must be greater than threshold distance. If the distance between new CH and existing CH is less than the threshold distance then the CH which is near to the border of the target area or to the centre of the target area will be discarded. If the threshold distance is maintained between new CH and existing CH, then new CH exists. Similarly, we choose k number of CHs in the target area. This

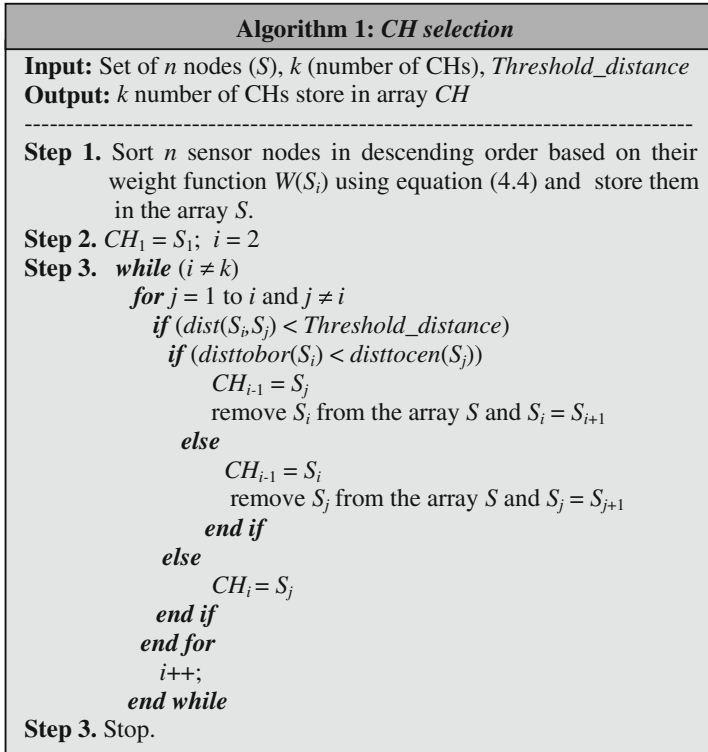


Fig. 1 CH selection algorithm

procedure assures the even distribution of the CHs in the target area. After selecting k number of CHs, the remaining normal sensor nodes are assigned to their nearest CH and form clusters. Finally, every CH is assigned to its nearest EH node.

Note that the number of EH nodes and the number of CHs are taken equal. Here, the CHs take responsibility of collecting the sensed data from their member sensor nodes, aggregate them and send those data to their assigned EH nodes. Then, these EH nodes forward the aggregated data to the sink. Therefore, the residual energy of the CHs is preserved as EH nodes are responsible of data forwarding to the sink. The detailed CH selection algorithm is presented in Fig. 1.

4 Simulation Results

The proposed algorithm was experimented through simulation run using MATLAB (version 7.5) and C programming language on Intel Core i7-2600 processor. For the experiments, we considered a 200×200 square meter area in which 100–400

sensor nodes are randomly deployed. Sink was assumed to be located at (100, 200) of the target area. Each sensor node was assumed to have an initial energy of 0.5 Joules. In the simulation run, the packet length be taken as 1000 bits, control bag length be 100 bits and the threshold distance is taken as 60 m. Here, we assumed the network lifetime as the number of rounds until 50 % of total sensor nodes die.

We present the simulation results by considering two scenarios of WSNs, WSN#1 and WSN#2 in which 8 % EH nodes are deployed and without any deployment of EH nodes in the target area respectively. At first, to show the significance of EH nodes, we ran our algorithm for WSN#1 and WSN#2. Then we compare the results of the proposed algorithm in terms of network lifetime and residual energy of the network. We also run an existing algorithm [12] for the sake of comparison.

4.1 Impact of EH Nodes on Lifetime and Residual Energy

We ran our proposed algorithm for scenario WSN#1 and scenario WSN#2. Table 1 shows the results of the proposed algorithm in terms of the network lifetime. Here, we observe that the results of WSN#1 are better than the WSN#2. This is due to the fact that in WSN#1, the EH nodes work as relay nodes for the CHs in the network where as in WSN#2, the CHs communicate the sink directly. As a result, the CHs of WSN#2 deplete energy faster than the CHs of WSN#1 because of long haul problem.

We also compare the results of WSN#1 and WSN#2 in terms of residual energy of the network. Table 2 presents the results in terms of residual energy of the network. Here, we observe that the results of WSN#1 are better than the WSN#2.

4.2 Comparison of the Network Lifetime and Residual Energy

We now show the comparison of the proposed method with the existing algorithm [12] in terms of the network lifetime for target area by varying sensor nodes from 100 to 400 as shown in Fig. 2.

Here, we compare the results of WSN#1 with the existing algorithm as WSN#1 is energy harvested network. It is obvious to note that the proposed method performs better. The rationale behind is that the existing algorithm selects single CH

Table 1 Network lifetime in terms of rounds

Sensor nodes	100	200	300	400
WSN#1	996	1185	1345	1411
WSN#2	664	733	773	866

Table 2 Comparison of residual energy

Sensor nodes	100	200	300	400
WSN#1	13.358	26.934	42.988	54.377
WSN#2	12.37	25.217	34.939	48.965

Fig. 2 Comparison of network lifetime

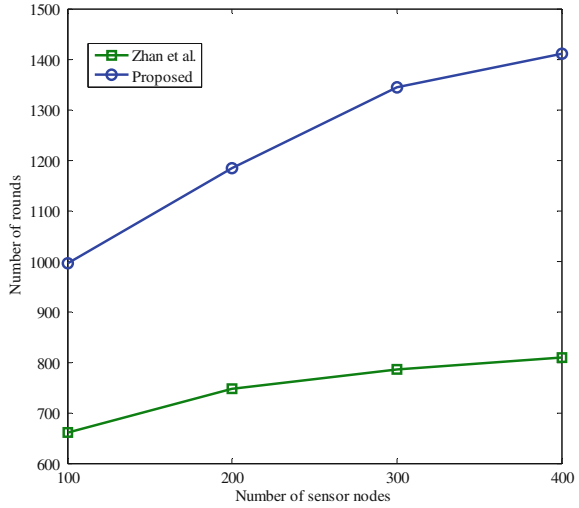
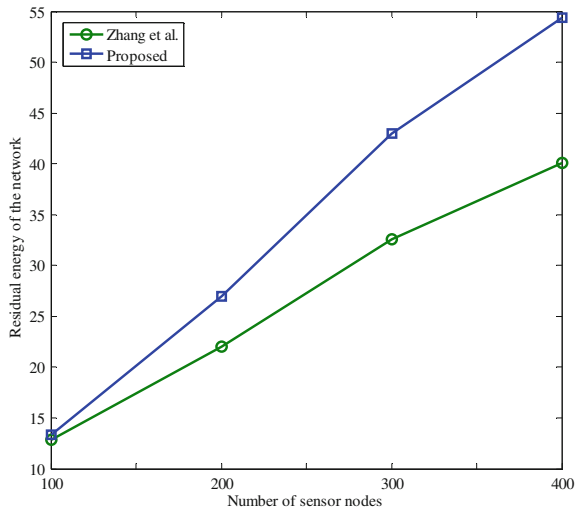


Fig. 3 Comparison of residual energy of the network



irrespective of the node density of the network. As a result, the CH dies faster due to over burden on that CH. Figure 3 shows the proposed algorithm performs better than the existing algorithm in terms of residual energy.

5 Conclusion and Future Work

In this paper, we have presented an energy efficient clustering algorithm. We have shown that the combination of energy efficient clustering and effective use of EH sensor nodes in the target area increase the network lifetime significantly. To show the result more realistic, we have used a basic astronomical prediction model for charging the EH sensor nodes. Our simulation results show that the superiority of the proposed algorithm in terms network lifetime and residual energy of the network over the existing method. For future work, we plan to integrate energy efficient routing protocol using EH sensor nodes to enhance the performance in terms of several performance metrics.

References

1. Sudevalayam, S., Kulkarni, P. Energy harvesting sensor nodes: Survey and implications. *Communications Surveys & Tutorials, IEEE*, (13) (3) 443–461 (2011).
2. V. Raghunathan, A. Kansal, J. Hsu, J. Friedman, M. Srivastava, Design considerations for solar energy harvesting wireless embedded systems, in: Fourth International Symposium on Information Processing in Sensor Networks, 457–462 (2005).
3. D. Hasenfratz, A. Meier, C. Moser, J.J. Chen, L. Thiele, Analysis, comparison, and optimization of routing protocols for energy harvesting wireless sensor networks, in: IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 19–26 (2010).
4. M. Islam, M. Islam, M. Islam, A-sleach: an advanced solar aware leach protocol for energy efficient routing in wireless sensor networks, in: Sixth International Conference on Networking, (4) (2007).
5. B. Medepally, N. Mehta, Voluntary energy harvesting relays and selection in cooperative wireless networks, *IEEE Tran. Wireless Communications*, (9) 3543–3553 (2010).
6. C. Bergonzini, D. Brunelli, L. Benini, Algorithms for harvested energy prediction in batteryless wireless sensor networks, in: 3rd International Workshop on Advances in Sensors and Interfaces, 144–149 (2009).
7. Rault, Tifenn, Abdelmadjid Bouabdallah, and Yacine Challal. Energy efficiency in wireless sensor networks: A top-down survey. *Computer Networks* (67) 104–122 (2014).
8. Afsar, M. Mehdi, and Mohammad-H. Tayarani-N. “Clustering in sensor networks: A literature survey. *Journal of Network and Computer Applications*, (46) 198–226 (2014).
9. P. Kuila, P.K. Jana, An energy balanced distributed clustering and routing algorithm for wireless sensor networks. 2nd IEEE Int. Conf. Parallel Distributed and Grid Computing (PDGC), 220–225 (2012).
10. Voigt, T., Dunkels, A., Alonso, J., Ritter, H., & Schiller, J. Solar-aware clustering in wireless sensor networks, in: IEEE International conference in Computers and Communications, (ISCC) (1) 238–243 (2004).
11. Bergonzini, D. Brunelli, and L. Benini, Algorithms for harvested energy prediction in battery less wireless sensor networks, International Workshop on Advances in sensors and Interfaces, (IWASI) 144–149 (2009).
12. Zhang, Pengfei, Gaoxi Xiao, Hwee-Pink Tan. Clustering algorithms for maximizing the lifetime of wireless sensor networks with energy-harvesting sensors, *Computer Networks*, (57) 2689–2704 (2013).

13. W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan.: An application specific protocol architecture for wireless microsensor networks, *IEEE Trans. Wirel. Commun.* 1 (4) 660–670 (2002).
14. Jeong, J., & Culler, D. Predicting the long-term behavior of a micro-solar power system. *ACM Transactions on Embedded Computing Systems (TECS)*, (2) 35 (2011).

Hybrid Network Intrusion Detection Systems: A Decade's Perspective

Asish Kumar Dalai and Sanjay Kumar Jena

Abstract With the increasing deployment of network systems, network attacks are increasing in intensity as well as complexity. Along with these increasing network attacks, many network intrusion detection techniques have been proposed which are broadly classified as being signature-based, classification-based, or anomaly-based. A deployable network intrusion detection system (NIDS) should be capable of detecting of known and unknown attacks in near real time with very low false positive rate. Supervised approaches for intrusion detection provides good detection accuracy for known attacks, but they can not detect unknown attacks. Some of the existing NIDS emphasize on unknown attack detection by using unsupervised anomaly detection techniques, but they can not distinguish network data as accurately as supervised approaches. Moreover they do not consider some other important issues like real time detection or minimization of false alarm. To overcome these problems, in the recent years many hybrid NIDS have been proposed which are basically aimed at detecting both known and unknown attacks with high accuracy of detection. In this literature review on hybrid network intrusion detection systems, we will discuss a few of the notable hybrid NIDS proposed in the recent years and will try to provide a comparative study on them.

Keywords Intrusion detection system · NIDS · Network security

1 Introduction

Network attacks are mainly of two types: the first category of network attacks, misuse the network resources such as network based DoS, probing, flooding, etc. In the second type attack, attackers exploits the vulnerabilities in the protocol. Network

A.K. Dalai (✉) · S.K. Jena
National Institute of Technology Rourkela, Rourkela, India
e-mail: dalai.asish@gmail.com

S.K. Jena
e-mail: skjena@nitrkl.ac.in

© Springer India 2017
D.K. Lobiyal et al. (eds.), *Proceedings of the International Conference on Signal, Networks, Computing, and Systems*, Lecture Notes in Electrical Engineering 395, DOI 10.1007/978-81-322-3592-7_35

341

defense systems, whether they are signature based or anomaly based, can be classified according to the attackers misuse type. The model to detect misuse of network resources uses the flow level information to make decisions and detect the attacks. To detect the exploitation of protocol vulnerabilities, the model uses information collected from the protocol headers. There is also one other type of network defense systems that detects applications vulnerabilities misuses by analyzing the information collected from the payloads. The well known open source system of this type is SNORT.

All these network intrusion detection systems use either signature based, classification based or anomaly based approaches for detecting intrusions. Signature based and classification based techniques are supervised, whereas anomaly detection technique is unsupervised intrusion detection techniques. Supervised intrusion detection techniques produce very high detection accuracy for known attacks, but they are not capable of detecting unknown attacks. On the other hand, unsupervised intrusion detection techniques are capable of detecting unknown attacks, but they can not categorize network data as accurately as supervised techniques for known attacks. So there is always a trade-off between detection accuracy and unknown attack detection when we employ either supervised or unsupervised approaches for network intrusion detection. In the recent years, a new type of intrusion detection approach known as hybrid intrusion detection systems has been proposed in the literature which combines more than one supervised and unsupervised interdependent intrusion detection approaches to increase the accuracy of detection as well as to detect unknown attacks. In this paper we will discuss some of the hybrid network intrusion detection approaches proposed in the recent years and will provide a comparative study on them.

1.1 Intrusions and Its Types

Intrusions are a set of actions particularly aimed to compromise the common security goals of the system like confidentiality, integrity, and availability. The attacks or intrusion to the system are mainly categorized into the following types:

- Denial of Service (DoS) A DoS attack is the class of attack that attempts to block access to the certain resource by legitimate users. Typically, the attack targets the particular network services, such as e-mail, web services or the temporary unavailability of all network connectivity and services. Common forms of denial of service attacks are TCP SYN flooding, buffer overflow, smurf, teardrop, ping of death, etc.
- User to Root Attacks (U2R) This is a type of attack where the attacker gains access to the normal user account on the victim system. The unauthorized access may be gained by social engineering, dictionary attack, sniffing passwords. The attacker gains the privileges of a super user of the system by exploiting the vulnerabilities.

- Remote to Local(R2L) R2L is a type of attack where the attacker sends packets to a remote machine over the network without having any account on the machine and then gains access (either as a root or as a user) to the machine and does malicious activities. The most common form of remote exploits is buffer overflow and other input modification attacks.
- Probe This type of attack scan the networks to identify valid network addresses and to collect information about the individual host (the e.g. operating system used and the services running on the host etc.). Some common probing attacks include Nmap (network mapper, a tool for scanning and enumeration), port sweep (scanning the ports to determine which services are running on the host), IP sweep (scanning the network hosts for services on ports of interest), etc.

2 Intrusion Detection Systems

An Intrusion Detection System (IDS) is a system that alerts the user about the intrusion by using either signature based or anomaly based method, or a combination of these techniques. IDSs are classified as Host-Based (HIDS) or Network-Based (NIDS) in the context of their place of deployment. Host-based systems guard a single host by monitoring its traffic while network-based systems obtain data by monitoring the traffic on the network.

Network intrusion detection system involves detecting intrusions in the whole network. Network intrusion detection systems are installed in the network where it can monitor the incoming and outgoing traffic of all devices on the network. Host Intrusion Detection Systems are placed on individual devices on the network. An HIDS analyzes the inbound and outbound traffic from the device and alert the user or administrator about the suspicious activity. Another approach for detection considers both the normal and anomalous patterns for training a system and then perform classification on the test data. Such a system combines the advantages of both the misuse-based and anomaly-based technique. In the design of an NIDS, the goal is to achieve the higher accuracy while analyzing the complete network traffic. Therefore, there is a need to design a hybrid intrusion detection model to meet this objective. The hybrid classifier usually combines several machine learning techniques, so that the overall system performance can be improved significantly. More specifically, a hybrid approach includes two different functional components. The first one inputs the raw traffic and produces the result and then the second one will then take the intermediate results as the input and takes the final decision.

2.1 Signature Based Detection Techniques

The Well-known signature-based detection techniques such as USTAT, SNORT and NetStat are limited in the way that they cannot detect unknown attacks. Also, when

a new attack is detected, it takes significant amount of time to update its signature database. This technique also produces number of false positive alerts.

2.2 Classification Based Detection Techniques

Classification-based detection techniques have both normal and abnormal data sets, and uses data mining techniques to train the system. This creates an accurate classification model than simple signature-based approaches. They are thus extremely useful in detecting known attacks where the attack signature is available with the IDS. However, they are still not capable to detect the zero day attacks.

2.3 Anomaly-Based Detection Techniques

Anomaly-based detection techniques develops a model of normal behavior, and then classify the abnormal which have statistically significant deviations from the normal. The advantage of these approach is their capability to detect zero day attacks. However, there is a potential chance of having a high false alarms when the model for generating the normal behaviors is not so accurate. Supervised learning approaches for anomaly detection, train the model with a known set of normal data and then use a testset to determine whether the data is normal or not. Unsupervised techniques are based on statistical approaches, outlier detection schemes, clustering, and state machines to detect the anomalous behavior without using any training data.

2.4 Hybrid Detection Techniques

In the recent years, researchers have tried to cascade the techniques mentioned above for network intrusion detection to enhance the performance of NIDS, which use hybrid classifiers to distinguish network data. To remove the unrepresentative training examples from each class, the hybrid classifiers uses clustering-based approach. Then, the results from the clustering are used as the training examples for the classifiers. Therefore, the first kind of hybrid classifiers can be based on either supervised or unsupervised learning methods. Finally, hybrid classifiers can also be based on the combination of the two different techniques in which the first one focus at optimizing the learning performance of the second model used for the attack prediction. Clustering is the process of a grouping the objects into clusters so that the objects in the same cluster similar to each other than to the objects in other clusters.

3 Existing Approaches for Hybrid Network Intrusion Detection Systems

Tsai and Lin [1] proposed a novel method based on the idea of the Triangle Area based Nearest Neighbors (TANN) by combining supervised and unsupervised learning methods for attack detection. They have used the k-Nearest Neighbor (k-NN) classifier to measure the similarity of attacks. Chen et al. [2] have applied Support Vector Machine (SVM) and Artificial Neural Network (ANN) for prediction of attacks based on frequency-based encoding methods. The goal of using SVM and ANN for attack detection is to develop a generalization capability from the limited set of training data. Two different encoding schemes: frequency-based and $tf \times idf$ scheme have been used to detect the intrusions. Herrero et al. [3] has proposed a novel hybrid artificial intelligent system for intrusion detection, called MOBILE-VISUALIZATION Hybrid IDS (MOVIH-IDS). The hybrid model is based on a multi-agent system that incorporates an unsupervised technique for effective detection of intrusions. It facilitates the detection of intrusions in dynamic networks, in a more flexible and adaptable manner. Ozyer et al. [4] have proposed an iterative rule learning method using a fuzzy rule-based genetic classifier. The approach consists of two steps: first, fuzzy association rule mining is used to generate a large number of rules, and then rule evaluation criteria are used to reduce the fuzzy rule search space. The candidate rules obtained after pre-processing are used in the genetic fuzzy classifier to generate class specific rules. Then for each class, the boosting genetic algorithm is employed to find its fuzzy rules meant to classify the data each time a fuzzy rule is extracted and incorporated in the model. Peddabachigari et al. [5] proposed two hybrid approaches for designing the IDS. SVM and Decision Trees (DT) are combined as a hierarchical hybrid intelligent system and an ensemble technique combining the standard classifiers. The motivation for using the hybrid technique is to improve the accuracy of the system in comparison to individual methods. The data set is first passed through the decision tree, and the information of the node is generated. The hybrid approach combines the results from the all individual systems resulting in better accuracy. Panda et al. [6] proposed a method to use multiple classifiers to make intelligent decisions. In this method, the data filtering is done after supervised classification or unsupervised clustering method to the training set. They have investigated the combination of DT, PCA, SPegasos (Stochastic variant of Piramol estimated sub-gradient solver in SVM), END, Random Forest and Grading for their method. Then the final classifier is used on the filtered data to get the final decision. Zhang et al. [7] proposed a hybrid intrusion detection system by combining the advantages of all different kinds of IDS. They have considered the accuracy and speed of data processing. The method combines the misuse detection technique and anomaly based technique. Gomez et al. [8] proposed a novel anomaly pre-processor that extends the functionality of the Snort IDS, resulting a hybrid detection system. This method, named H-Snort models the network traffic

at a high level, and then it stores the information to model the normal behavior of the system, it allows to configure it and adjust the sensitivity of the system to prevent the false alarms. Jawhar and Mehrotra [9] have presented an intrusion detection model based on hybrid fuzzy logic and artificial neural network. Fuzzy clustering is used to remove the overlap between normal and abnormal behavior, which in turn reduces the false alarms. Aydın et al. [10] have proposed the hybrid IDS by combining Packet Header Anomaly Detection (PHAD) and Network Traffic Anomaly Detection (NETAD). They have used Snort for misuse-based IDS and PHAD and NETAD as anomaly-based IDS. Hwang et al. [11] proposed a novel experimental hybrid IDS that combines the advantages of the low false-positive rate of the signature-based system and the unknown attack detection ability of anomaly based system. By mining anomalous traffic, they build the IDS that detects anomalies better than the SNORT or Bro systems. A weighted signature generation scheme is developed to integrate anomaly based technique with SNORT by generating signatures from anomaly based method. Chung et al. [12] have proposed an intelligent dynamic swarm based rough set (IDS-RS) based hybrid intrusion detection system. It uses simplified swarm optimization for intrusion data classification. Elbasiony et al. [13] proposed a model that depends on data mining classification and the clustering techniques. Random forest classification algorithm and weighted k-means clustering algorithm are used to for both misuse detection and anomaly detection. Kim et al. [14] proposed a new hybrid intrusion detection method that hierarchically integrates a misuse detection model and an anomaly detection. C4.5 decision tree algorithm is used for misuse detection, and then multiple one-class SVM models are designed for the decomposed subsets. Luo et al. [15] proposed a Four-Angle-Star based Visualized Feature Generation approach (FASVFG) with generalization accuracy of 94.3

4 General Comparison of the Existing Hybrid NIDSs

In the subsection, we will try to compare the existing hybrid network intrusion detection systems based on the issues related to NIDS as discussed above. While some of these approaches concentrate on only some particular issues, some other approaches takes some different issues into account. So finding out a general comparison among all these approaches is a difficult task. Most of these approaches have their own pros and cons, while they stand out in some particular issues, at the same time they are ignorant of some other important issues of an NIDS performance. A general comparison of the existing hybrid network intrusion detection approaches has been shown in Table 1. The comparison has been based upon the key issues related to an NIDS performance discussed in previous subsections.

Table 1 General comparison of existing hybrid NIDSs

Scheme	Hybrid detection approach	Detection accuracy	False alarm rate	Fastness of detection	Scalability	Capability to detect unknown attack	Dynamic update
[1]	k-means clustering + kNN classifier	99.01 %	2.99 %	Unknown	No	Yes	No
[2]	ANN + SVM + $tf \times idf$ scheme	100 %	8.53 %	Unknown	No	No	No
[3]	Unsupervised neural models	Unknown	Unknown	Near real time	Yes	Unknown	Yes
[4]	Fuzzy rule-based genetic classification + Boosting	92.09 %	4.20 %	Unknown	No	Unknown	Unknown
[5]	Decision trees and SVM	99.8 %	0.3 %	Unknown	No	No	No
[6]	Combination of classifiers like DT, PCA, RBFNN, SPegasos etc.	99.5 %	0.5 %	Unknown	No	No	No
[7]	SVM with RBF	90.28 to 99.76 %	0.47 %	0.25 to 1.47 ms	Yes	No	Yes
[8]	Statistical moving average method + SNORT	Unknown	Unknown	Unknown	No	No	No
[9]	FCM clustering and neural network	99.9 %	0.01 % (For known normal)	0.02ms	Yes	Yes (68.6 % detection accuracy)	No
[10]	Snort + PHAD + NETAD	72.63 %	Unknown	Unknown	Unkonown	No	No
[11]	Episode rule mining + SNORT	97 %	1 %	Approximately 0.0005 ms	Yes	Yes	Yes
[11]	Simplified swarm optimization	99.3 %	Unknown	Unknown	Yes	Yes	Yes
[13]	Random forest + wk-mean clustering	98 %	6 %	Unknown	Yes	No	Yes
[14]	Random forest + wk-mean clustering	99 %	0.3 %	Approximately 1.02 s	Yes	Yes	Yes
[15]	Visualized feature generalization	94.3 %	Unknown	Unknown	Yes	No	No

5 Conclusion

From the general comparison of the existing hybrid network intrusion detection systems, it has been observed that most of them performs very well in terms of detection accuracy and false alarm for already known attack and normal classes of traffic. But most of them are not capable of detecting unknown attacks, which is of prime importance with evolving network scenario. Also for those who are capable of detecting unknown attacks have low detection accuracy for unknown attacks. Time required for traffic categorization is a very important aspect of NIDSs. A NIDS can be scaled up to real networks only if it can detect attacks in real time.

References

1. Chih-Fong Tsai, C.Y.L.: A triangle area based nearest neighbors approach to intrusion detection. *Pattern Recognition* **43** 222–229 2010
2. Wun-Hwa Chen, Sheng-Hsun Hsu, H.P.S.: Application of svm and ann for intrusion detection. *Computers and Operations Research* **32** 2617–2634 2005
3. Alvaro Herrero, Emilio Corchado, M.A.P.A.A.: Movih-ids: A mobile-visualization hybrid intrusion detection system. *Neurocomputing* **72** 2775–2784 2009
4. Tansel Ozyer, Reda Alhajj, K.B.: Intrusion detection by integrating boosting genetic fuzzy classier and data mining criteria for rule pre-screening. *Journal of Network and Computer Applications* **30** 99–113 2007
5. Sandhya Peddabachigari, Ajith Abrahamb, C.G.J.T.: Modeling intrusion detection system using hybrid intelligent systems. *Journal of Network and Computer Applications* **30** 114–132 2007
6. M Panda, Ajith Abraham, M.R.P.: A hybrid intelligent approach for network intrusion detection. In: *Proc. International Conference on Communication Technology and System Design 2011*. ICCTSD 1–9 2011
7. Baojun Zhang, Xuezen Pan, J.W.: Hybrid intrusion detection system for complicated network. In: *Proc. of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery*. FSKD 2007
8. J. Gomez, C. Gil, N.P.R.B.C.J.: Design of a snort-based hybrid intrusion detection system. In: *Proc.of the IWANN 2009*. 515–522 2009
9. Jawhar, M., Mehrotra, M.: Design network intrusion detection system using hybrid fuzzy-neural network. *International Journal of Computer Science and Security* **4** 285 2010
10. Aydin, M., Zaim, A., Ceylan, K.: A hybrid intrusion detection system design for computer network security. *Computers & Electrical Engineering* **35** (2009) 517–526 2009
11. Hwang, K., Cai, M., Chen, Y., Qin, M.: Hybrid intrusion detection with weighted signature generation over anomalous internet episodes. *IEEE Transactions on Dependable and Secure Computing*, **4** (2007) 41–55
12. Yuk Ying Chung and Noorhaniza Wahid. A hybrid network intrusion detection system using simplified swarm optimization (sso). *Applied Soft Computing*, 12(9):3014–3022, 2012.

13. Reda M Elbasiony, Elsayed A Sallam, Tarek E Eltobely, and Mahmoud M Fahmy. A hybrid network intrusion detection framework based on random forests and weighted k-means. *Ain Shams Engineering Journal*, 4(4):753–762, 2013.
14. Gisung Kim, Seungmin Lee, and Sehun Kim. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, 41(4):1690–1700, 2014.
15. Bin Luo and Jingbo Xia. A novel intrusion detection system based on feature generation with visualization strategy. *Expert Systems with Applications*, 41(9):4139–4147, 2014.

Author Index

A

Abhishek, Kumar, 183
Agrawal, Smita, 25
Anirudh, U.K., 205
Arquam, Md, 171

B

Behera, Pratap Kumar, 285
Behera, Upasana, 315
Bharti, Manisha, 261
Bhattacharjee, Sanghita, 193

C

Chandrasekaran, K., 217
Chennakrishna, M., 13

D

Dalai, Asish Kumar, 341
Das, Himanish Shekhar, 193
Das, Manik Lal, 117
Deepak, B.B.V.L., 315
Deshmukh, Sachin, 297
Dhenakaran, S.S., 161
Dhingra, Atul, 85
Dora, Durga Prasada, 269
Doshi, Chintan Ketankumar, 217

G

Ganesh, L., 227
Gangodkar, Durgaprasad, 137
Gundlapalli, Meghana, 3
Gupta, C.P., 171

H

Hanmandlu, M., 85

I

Ingle, Vaishali, 297

J

Jana, Prasanta K., 331
Jannu, Srikanth, 331
Jayasree, M., 53
Jeevan, Medikonda, 85
Jena, S.K., 243
Jena, Sanjay Kumar, 341

K

Kaiwartya, Omprakash, 253
Kamath, Narasimha, 205
Kekre, Hemant B., 149
Khilar, Pabitra Mohan, 285
Kilaru, Seetaiah, 277
Krawczyk, Henryk, 303
Krishnan, Arun, 117
Kulkarni, Anoop C, 61
Kumar, Kirshna, 253
Kumar, Manoj, 25, 45, 261
Kumar, P. Ramesh, 161
Kumar, Sarowar, 183
Kumar, Sushil, 253, 269

L

Lakshman, Vidyashankar B., 217
Lawrence, Amith, 33

M

Mallik, Puspanjali, 269
Manikantan, K., 3, 13, 33, 61
Manoj Ashwin, N.V., 33
Mittal, Ankush, 137
Mukherjee, Anuta, 235
Mukherjee, Saswati, 235

N

Narayanan, N.K., 53, 93, 105
Natu, Shachi, 149

O

Orzechowski, Piotr, [303](#)

P

Padmaja, S., [277](#)

Panigrahi, B.K., [85](#)

Pant, Triloki, [127](#)

Parsola, Jyoti, [137](#)

Patnaik, K.S., [73](#)

Prakash, Ravi, [315](#)

Prasad, N.R., [61](#)

Proficz, Jerzy, [303](#)

Pukhrambam, Priyanka, [193](#)

Purohit, Nikhil, [13](#)

R

Raghu, Anunita, [3](#)

Ramesh, Rakshit, [61](#)

Raviteja, G., [315](#)

Reddy, K. Venugopal, [277](#)

Rout, Jitendra Kumar, [243](#)

S

Sahoo, Bibhudatta, [323](#)

Sahoo, Kshira Sagar, [323](#)

Sajwan, M., [73](#)

Sankaranarayanan, Sreecharan, [217](#)

Sarode, Tanuja, [149](#)

Sasibhushana Rao, G., [227](#)

Sharma, Ajay K., [261](#)

Singh, M.P., [183](#)

Sonam, [45](#)

Sreekanth, N.S., [93](#), [105](#)

Szymański, Julian, [303](#)

V

Vadda, Lavanya, [227](#)

Venkatesh, Rasula, [243](#)

Verma, Manish Kumar, [183](#)

Vyas, Laxita, [171](#)