Suresh Chandra Satapathy
Yogesh Chandra Bhatt
Amit Joshi
Durgesh Kumar Mishra   *Editors*

Proceedings of the International Congress on Information and Communication Technology

ICICT 2015, Volume 2

Advances in Intelligent Systems and Computing

Volume 439

About this Series

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

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

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

Suresh Chandra Satapathy
Yogesh Chandra Bhatt · Amit Joshi
Durgesh Kumar Mishra
Editors

Proceedings of the International Congress on Information and Communication Technology

ICICT 2015, Volume 2

 Springer

Editors

Suresh Chandra Satapathy
Department of Computer Science
  and Engineering
Anil Neerukonda Institute of Technology
  and Sciences
Visakhapatnam, Andhra Pradesh
India

Yogesh Chandra Bhatt
College of Technology and Engineering
Maharana Pratap University of Agriculture
  and Technology
Udaipur, Rajasthan
India

Amit Joshi
Sabar Institute of Technology
Sabarkantha, Gujarat
India

Durgesh Kumar Mishra
Sri Aurobindo Institute of Technology
Indore, Madhya Pradesh
India

Printed on acid-free paper

Preface

This AISC volume contains the papers presented at the ICICT 2015: International Congress on Information and Communication Technology. The conference was held during October 9 and 10, 2015 at Hotel Golden Tulip, Udaipur, India and organized by CSI Udaipur Chapter, Division IV, SIG-WNS, SIG-e-Agriculture in association with ACM Udaipur Professional Chapter, The Institution of Engineers (India), Udaipur Local Centre and Mining Engineers Association of India, Rajasthan Udaipur Chapter. It has targeted state-of-the-art as well as emerging topics pertaining to ICT and effective strategies for its implementation of engineering and managerial applications. The objective of this international conference is to provide opportunities for the researchers, academicians, industry persons, and students to interact and exchange ideas, experience and expertise in the current trends and the strategies for information and communication technologies. Besides this, participants will also be enlightened about vast avenues, current and emerging technological developments in the field of ICT in this era, and its applications being thoroughly explored and discussed. The conference has attracted a large number of high-quality submissions and stimulated the cutting-edge research discussions among many academic pioneering researchers, scientists, industrial engineers, and students all over the world and also has provided a forum to researcher. The goals of this forum, as follows, are to: Propose new technologies, share their experiences and discuss future solutions for design infrastructure for ICT; provide common platform for academic pioneering researchers, scientists, engineers, and students to share their views and achievements; enrich technocrats and academicians by presenting their innovative and constructive ideas; focus on innovative issues at international level by bringing together the experts from different countries. Research submissions in various advanced technology areas were received and after a rigorous peer-review process with the help of program committee members and external reviewer, 135 (Vol-I: 68, Vol-II: 67) papers were accepted with an acceptance ratio of 0.43. The conference featured many distinguished personalities like Dr. L.V. Murlikrishna Reddy, President, The Institution of Engineers (India); Dr. Aynur Unal, Stanford University, USA; Ms. Mercy Bere, Polytechnique of

Namibia, Namibia; Dr. Anirban Basu (Vice President and President Elect) Zera Gmbh, Germany; Dr. Mukesh Kumar, TITS, Bhiwani; Dr. Vipin Tyagi, Jaypee University, Guna; Dr. Durgesh Kumar Mishra, Chairman Division IV CSI; Dr. Basant Tiwari, and many more. Separate Invited talks were organized in industrial and academia tracks on both days. The conference also hosted few tutorials and workshops for the benefit of participants. We are indebted to ACM Udaipur Professional Chapter, The Institution of Engineers (India), Udaipur Local Centre and Mining Engineers Association of India, Rajasthan Udaipur Chapter for their immense support to make this Congress possible in such a grand scale. A total of 15 Sessions were organized as a part of *ICICT 2015* including 12 technical, one plenary and one Inaugural Session and one Valedictory Session. A total of 118 papers were presented in 12 technical sessions with high discussion insights. The total number of accepted submissions was 139 with a focal point on ICT. The Session Chairs for the technical sessions were Dr. Chirag Thaker, GEC, Bhavnagar, India; Dr. Vipin Tyagi, Jaypee University, MP, India; Dr. Priyanka Sharma, Raksha Shakti University, Ahmedabad, India; Dr. S.K. Sharma; Dr. Bharat Singh Deora; Dr. Nitika Vats Doohan, Indore; Dr. Mahipal Singh Deora; Dr. Tarun Shrimali; and Dr. L.C. Bishnoi.

Our sincere thanks to all sponsors, press, and print and electronic media for their excellent coverage of this congress.

October 2015   Suresh Chandra Satapathy
                                                            Yogesh Chandra Bhatt
                                                                        Amit Joshi
                                                           Durgesh Kumar Mishra

Organising Committee

Chief Patron

Prof. Bipin V. Mehta, President, CSI

International Advisory Committee

Chandana Unnithan, Victoria University, Melbourne, Australia
Dr. Aynur Unal, Standford University, USA
Dr. Malay Nayak, Director-IT, London
Chih-Heng Ke, MIEEE, NKIT, Taiwan
Dr. Pawan Lingras, Professor, Saint Mary University, Canada

National Advisory Committee

Dr. Anirban Basu, Vice President, CSI
Mr. Sanjay Mahapatra, Hon. Secretary, CSI
Mr. R.K. Vyas, Treasurer, CSI
Mr. H.R. Mohan, Immediate Past President, CSI
Prof. Vipin Tyagi, RVP III, CSI
Prof. S.S. Sarangdevot, VC, JRNVU, Udaipur
Dr. R.S. Shekhawat, RSC, Region 3, CSI
Prof. H.R. Vishwakarma, VIT, Vellore, India
Prof. Dr. P. Thrimurthy, Past President, CSI
Dr. G.P. Sharma, CSI Udaipur Chapter
Dr. Nilesh Modi, Chairman, CSI Ahmedabad Chapter
Prof. R.C. Purohit, CSI Udaipur Chapter
Prof. S.K. Sharma Director PIE, PAHER, Udaipur
Dr. T.V. Gopal, Anna University, Chennai

Dr. Deepak Sharma, CSI Udaipur Chapter
Prof. R.C. Purohit, CTAE, Udaipur
Prof. Pravesh Bhadviya, Director, Sabar Education, Gujarat
Dr. Bharat Singh Deora, Department of CS and IT, JRNRV University, Udaipur

Technical Program Committee

Chair
Dr. Suresh Chandra Satapathy, Chairman, Division V, CSI

Co-chairs
Dr. Nisarg Pathak, KSV, Kadi, Gujarat
Dr. Mukesh Sharma, SFSU, Jaipur

Program Secretary
Er. Sanjay Agal, Pacfic University, Udaipur

Members
Mr. Ajay Chaudhary, IIT Roorkee
Dr. Mahipal Singh Deora, BNPG College, Udaipur
Prof. D.A. Parikh, Head, CE, LDCE, Ahmedabad
Dr. Savita Gandhi, Head, CE, Rolwala, G.U., Ahmedabad
Prof. Dr. Jyoti Pareek, Department of Computer Science, Gujarat University
Prof. L.C. Bishnoi, Principal, GPC, Kota
Ms. Bijal Talati, Head, CE, SVIT, Vasad
Er. Kalpana Jain, CTAE, Udaipur
Dr. Harshal Arolkar, Immd. Past Chairman, CSI Ahmedabad Chapter
Mr. Bhavesh Joshi, Advent College, Udaipur
Prof. K.C. Roy, Director, Madhav University, Sirohi
Dr. Pushpendra Singh, Sunrise Group of Institutions
Dr. Sanjay M. Shah, GEC, Gandhinagar
Dr. Chirag S. Thaker, GEC, Bhavnagar, Gujarat
Mrs. Meenakshi Tripathi, MNIT, Jaipur
Mr. Jeril Kuriakose, Manipal University, Jaipur

Chair: Tracks Management
Prof. Vikrant Bhateja, SRMGPC, Lucknow, India

Organising Committee

General Chair
Dr. Dharm Singh, Convener, SIG-WNS, CSI
Dr. Durgesh Kumar Mishra, Chairman, Division IV, CSI

Organising Chair
Dr. Y.C. Bhatt, Chairman, CSI Udaipur Chapter

Organising Co-chairs
Dr. B.R. Ranwah, Immd. Past Chairman, CSI Udaipur Chapter

Local Arrangements Chair
Dr. S.S. Rathore, CTAE, Udaipur

Finance Chair
Prof. S.K. Sharma, Treasurer, CSI Udaipur Chapter

Organising Secretary
Mr. Amit Joshi, CSI Udaipur Chapter

Contents

About the Editors

Dr. Suresh Chandra Satapathy is currently working as Professor and Head, Department of CSE at Anil Neerukonda Institute of Technology and Sciences (ANITS), Andhra Pradesh, India. He obtained his Ph.D. in Computer Science and Engineering from JNTU Hyderabad and M.Tech. in CSE from NIT, Rourkela, Odisha, India. He has 26 years of teaching experience. His research interests are data mining, machine intelligence, and swarm intelligence. He has acted as program chair of many international conferences and edited six volumes of proceedings from Springer LNCS and AISC series. He is currently guiding eight Ph.D. scholars. Dr. Satapathy is also a Sr. Member, IEEE.

Dr. Yogesh Chandra Bhatt graduated from CTAE, Udaipur (1978) and M.Tech. (1980) and Ph.D. (1989) from IIT Kharagpur. Currently, he is serving as Dean (Student Welfare) and Chairman (University Sports Board), Professor and Project In-charge (FIM, FMTC) Department of FMP, CTAE, MPUAT, Udaipur with 36 years of service experience. He is also Professional Agricultural Engineer worked in all divisions of Teaching, Research and Extension wing of the University. He has served as Head of Department for 7 years. He has published more than 50 research papers in international and national journals, seminar and conferences, edited six books, five proceedings and has guided 15 PG students for M.Tech. and Ph.D. degrees. He has completed ten Adhoc research projects. Developed ten prototype technologies on farm mechanization. He has served as Honorary Secretary (2009–2011) and Chairman (2011–2013) of The Institution of Engineers India, Udaipur Local Centre. Organized three international conferences and 18 national conventions and All India Seminar of IEI and ISAE. Published Members Directory of IEI, ULC. He has started Er.M.P. Baya and Mrs. Sheela Baya National Award Rs. 50,000 and Rs. 25,000 and Scholarship for Engineering Students Rs. 60,000 per year from IEI Udaipur. He was Vice Chairman ISAE, Rajasthan Chapter for 8 years and Director of Farm Power and Machinery Group in ISAE for the year 2012–2015 and National Convener of SIG on e-Agriculture in the banner of CSI working on application of ICT techniques in farm machinery. He has received awards of appreciation certificate for outstanding services in MPUAT, Udaipur (2009) and Scroll of Honour from

with Recognition of Eminent Contribution in the field of Agri Engineering (2013) IEI Kolkata. Fellow and Chartered Engineers of IEI, and Life member of ISAE, SESI, CSI and Member ACM. Patron of Rajasthan Agricultural Machinery Association, Jaipur. Working as Chairman of Computer Society of India (CSI) Udaipur Chapter 2015–2016 and Vice Chairman of Association of Computing Machinery (ACM) Udaipur Chapter.

Mr. Amit Joshi has an experience of around 6 years in academic and industry in prestigious organizations of Rajasthan and Gujarat. Currently, he is working as an Assistant Professor in Department of Information Technology at Sabar Institute in Gujarat. He is an active member of ACM, CSI, AMIE, IEEE, IACSIT-Singapore, IDES, ACEEE, NPA and many other professional societies. He is Honorary Secretary of CSI Udaipur Chapter and Honorary Secretary for ACM Udaipur Chapter. He has presented and published more than 40 papers in national and international journals/conferences of IEEE, Springer, and ACM. He has also edited three books on diversified subjects, namely Advances in Open Source Mobile Technologies, ICT for Integrated Rural Development, and ICT for Competitive Strategies. He has also organized more than 25 national and international conferences and workshops including International Conference ETNCC 2011 at Udaipur through IEEE, International Conference ICTCS-2014 at Udaipur through ACM, International Conference ICT4SD 2015—by Springer recently. He has also served on Organising and Program Committee of more than 50 conferences/seminars/workshops throughout the world and presented 6 invited talks in various conferences. For his contribution towards the society, The Institution of Engineers (India), ULC, has given him Appreciation award on the Celebration of Engineers, 2014 and by SIG-WNs Computer Society of India on ACCE, 2012.

Dr. Durgesh Kumar Mishra has received M.Tech. degree in Computer Science from DAVV, Indore in 1994 and Ph.D. degree in Computer Engineering in 2008. Currently, he is working as a Professor (CSE) and Director, Microsoft Innovation Centre at Shri Aurobindo Institute of Technology, Indore, MP, India. He is also a visiting faculty at IIT-Indore, MP, India. He has 24 years of teaching and 10 years of research experience. He completed his Ph.D. under the guidance of late Dr. M. Chandwani on Secure Multi-Party Computation for Preserving Privacy. He has published more than 90 papers in refereed international/national journals and conferences including IEEE, ACM conferences. He has organized many conferences such as WOCN, CONSEG, and CSIBIG in the capacity of conference General Chair and editor of conference proceeding. His publications are listed in DBLP, Citeseer-x, Elsevier, and Scopus. He is a Senior Member of IEEE and has held many positions like Chairman, IEEE MP-Subsection (2011–2012), and Chairman IEEE Computer Society Bombay Chapter (2009–2010). Dr. Mishra has also served the largest technical and profession association of India, the Computer Society of India (CSI) as Chairman, CSI Indore Chapter, State Student Coordinator-Region III MP, Member-Student Research Board, Core Member-CSI IT Excellence Award Committee. Now he is Chairman CSI Division IV Communication at

National Level (2014–2016). Dr. Mishra has delivered his tutorials in IEEE
International conferences in India as well as abroad. He is also the programme
committee member, and reviewer of several international conferences. He visited
and delivered his invited talks in Taiwan, Bangladesh, Singapore, Nepal, USA, UK,
and France. He has authored a book on "Database Management Systems." He had
been Chief Editor of Journal of Technology and Engineering Sciences. He has been
also serving as a member of Editorial Board of many national and international
refereed journals. He has been a consultant to industries and government organi-
zations like Sales Tax and Labour Department of Government of Madhya Pradesh,
India. He has been awarded with "Paper Presenter award at International Level" by
Computer Society of India. Recently in month of June, he visited MIT Boston and
presented his presentation on security and privacy. He has also chaired a panel on
"Digital Monozukuri" at "Norbert Winner in twenty-first century" at Boston.
Recently, he became the Member of Bureau of Indian Standards (BIS), Goverment
of India for Information Security domain.

A Novel Methodology to Detect Bone Cancer Stage Using Mean Intensity of MRI Imagery and Region Growing Algorithm

K.E. Balachandrudu, C. Kishor Kumar Reddy, G.V.S. Raju and P.R. Anisha

Abstract Cancer has been a plague in our society since the dawn of recorded history. The radical surgical resection represents the only chance for cure but, unfortunately it is possible in only 15 % of patients. Even at experienced centers the 5 year survival rates for the most favorable patients who undergo resection and adjuvant therapy are less than 20 %. In this paper, a methodology is proposed for identifying the bone cancer affected part. The methodology involves scanned images captured at various locations of the human body which are collected from different diagnostic labs. Based on region growing algorithm, the segmentation is carried out to analyze the tumor part through which the intensity of cancer and the stage at which the tumor relies is empirically calculated.

Keywords Bone cancer · Gray scale · Mean intensity · Segmentation · Region growing algorithm · Tumor

1 Introduction

Cancer is a multifarious genetic disease which is caused primarily by environmental factors. The cancer causing agents (carcinogens) can be present in food, water, air, sunlight, and also in chemicals that people are exposed to. Possible signs and symptoms for such diseases include: a new lump, abnormal bleeding, a prolonged cough, unexplained weight loss, and a change in bowel movements, etc., while these symptoms may indicate cancer and may also occur due to other issues also [1, 2]. There are over 100 different known cancers that affect humans. In 2012, about 14.1 million new cases every year of cancer occurred globally. In the same

---

K.E. Balachandrudu
Malla Reddy Institute of Engineering & Technology, Hyderabad, India

C. Kishor Kumar Reddy (✉) · G.V.S. Raju · P.R. Anisha
Stanley College of Engineering & Technology for Women, Hyderabad, India
e-mail: kishoar23@gmail.com

year, this malignant tumor caused about 8.2 million deaths, i.e., 14.6 % of all human deaths [3].

There are different types of bone cancer that can affect different patient populations and they are often treated differently. Some of the most common types of bone cancer are osteosarcoma, chondrosarcoma, ewing's sarcoma, pleomorphic sarcoma, fibrosarcoma. The American Cancer Society (www.cancer.org) estimation for cancer related to bones and joints for the year 2014 represents that about 3,020 new cases shall be diagnosed, and also presumed that 1,460 deaths due to bone cancers were expected. Bone cancer like any other cancer predominantly occurs in four stages [4, 5].

Stage 1—The cancer has not spread out of the bone. The cancer is not an aggressive one.
Stage 2—The same as Stage I, but it is an aggressive cancer.
Stage 3—Tumors exist in multiple places of the same bone (at least two).
Stage 4—The cancer has spread to other parts of the body.

Since the exact cause of bone cancer is poorly understood, there are no lifestyle changes or habits that can prevent these uncommon cancers. The best way of facing it is to detect it at the earliest stage and take appropriate measure. The American Cancer Society's estimates for cancer of the bones and joints for 2015 are: About 2,970 new cases will be diagnosed and about 1,490 deaths from these cancers are expected. The estimated number of cancer cases and deaths across USA in 2014 are listed in Table 1 and across the globe is listed in Table 2.

To some extent the belief that cancer cannot be cured persists even today. In regard to the same, this paper contributes a methodology to detect cancer and estimate the mean intensity and stage of the cancer. This process is carried out in four stages namely seed point selection, collection of additional image data, determination of threshold value, checking the similarity threshold value [6–14].

The rest of the paper is organized is as follows: Sect. 2 presents the relevant work that was carried out all in various fields of research, Sect. 3 presents the methodology adopted for the experimentation process, Sect. 4 gives the result analysis and Sect. 5 put forwards the conclusion.

2 Relevant Work

Roy and Roy [1], proposed an automated method for detection of brain abnormalities from MRI scan images. They introduce a method to find appropriate threshold intensity value which is near the intensity value of the tumor border using standard threshold value. Leela and Veena Kumari [2] proposed a morphological approach for the detection of brain tumor and cancer cells. In this paper, the segmentation is carried out using k-means and fuzzy c-means clustering algorithm for better performance. Rohit et al. [6] implemented a method for calculating the brain tumor shape, tumor area, and its stages. Detection resists the accurate

Table 1 Estimated cancer cases and deaths across the USA

Cancer	Estimated cases		Estimated deaths	
	Male	Female	Male	Female
Pancreas	23,530	22,890	20,170	19,420
Stomach	13,730	8,490	6,720	4,270
Liver	24,600	8,590	15,870	7,130
Lung	116,000	108,210	86,930	72,330
Breast	2,360	232,670	430	40,000
Thyroid	15,190	47,790	830	1,060
Eye	1,440	1,290	130	180
Intestine	4,880	4,280	640	570
Brain	12,820	10,560	8,090	6,230
Bone	1,680	1,340	830	630

Table 2 Estimated cancer cases and deaths across globe

Cancer type	Estimated cases	Estimated deaths
Prostate	233,000	29,480
Breast	232,670	40,000
Lung	224,210	159,260
Colon	136,830	50,310
Melanoma	76,100	9,710
Bladder	74,690	15,580
Lymphoma	70,800	18,990
Kidney	63,920	13,860
Thyroid	62,980	1,890
Endometrial	52,630	8,590
Leukemia	52,380	24,090

determination of stage and size of tumor. To overcome that, this paper uses computer aided method for segmentation of brain tumor based on the combination of two algorithms, that is, k-means and fuzzy c-means algorithms.

Kaur and Juneja [15], explained how the gradient differential plays an indivisible part in demarking the tumor in brain areas. Here one benchmark set is established, if that part do not match with the benchmark set are skipped by the algorithm, i.e., eminent entropy and intensity which are taken as major feature for identification of tumor in brain. Zade and Wanjari [16] proposed a method for early detection of breast cancer from mammograms. Mammography is the best available technique to detect cancer cell in its earlier stages and here seeded region growing algorithm is used for cancer region detection based on pixel intensity identification regions are separated. Kowar and Yadav [3] proposed a method for brain tumor detection and segmentation using histogram thresholding. Manual segmentation of brain tumors from magnetic resonance images is a tough and time-consuming task.

Jose and Sambath [7] proposed a new method for brain tumor segmentation. They used k-means clustering and fuzzy c-means algorithms for segmenting the brain tumor; for removing the noise preprocessing methods are applied on the images. After segmenting the tumor part from brain they calculated the brain tumor size in millimeters. Based on the tumor size of brain, they estimated the stage of brain tumor. Patel and Doshi [8] applied different segmentation methods for segmenting the brain tumor from MRI and CT scan images. They applied thresholding methods like region growing, mean shift, and clustering methods like k-means and fuzzy c-means and all these methods are used with the help of image processing techniques for segmenting the tumor from brain biomedical images. Kaushik and Sharma [9] proposed a method for volume calculation of brain tumor. Region growing method is used for segmenting region of interest (ROI), and using edge detection method the boundary of tumor part is identified by edge detection method and volume of the tumor is calculated. Talegaonkar et al. [17] proposed a method for automatic brain tumor detection in magnetic resonance images. They detected the location, position, and size of the tumor in brain automatically. They segmented the brain tumor using k-means clustering algorithm. After that the number of white pixels in the brain tumor part is calculated based on that size of the tumor calculated.

Peskin et al. [4] proposed a method for robust volume calculations for lung CT images and various sizes of tumors in lung. Ponraj et al. [18] gave review on existing preprocessing approach of mammographic images of breast. Sasikala and Vasanthakumar [19] used a k-means clustering algorithm to detect cancer in a multi resolution representation of the original MRI, ultrasound, and mammogram images. Rajan and Prakash [20] described a new method that uses data mining predicting lung cancer at an early stage. Rajeswari and Reena [5] proposed a new method for classifying tumor areas. In their method, liver cancer cell classification is performed using a support vector mechanism and fuzzy neural network classifiers. MAPSTD is applied to provide association ranking, this method is applied to liver cancer patient datasets. Feldman et al. [21] introduced a boosted Bayesian multi resolution (BBMR) classifier for computerized recognition of prostate cancer (CaP) from digitized histopathology, an obligatory precursor of automated Gleason grading.

Ganesan et al. [22] proposed computer-aided diagnosis from brain cancer image segmentation and implemented a method for segmenting the brain tumor area in which the k-means algorithm used for mammograms is applied. This method improves accuracy and reduces computational time. A conventional decision tree approach is used for dataset analysis. Bandyopadhyay and Paul [23] applied K-means clustering and DBSCAN for segmenting a tumor in MRI images of the human brain. After segmentation, they compared the results of both algorithms and concluded that the computational time increases exponentially when DBSCAN is applied.

3 Proposed Methodology

The processing of bone cancer imagery proposed in this paper is to identify the region of bone cancer and evaluate its mean gray scale intensity and analyze at which stage the tumor is erupted. This method requires scan imagery of bone with or without noise. When an image containing noise, such as illumination variations, occlusions, scale variations, or deformation of objects, is analyzed, the information retrieved may not represent the original value. Hence, the image must be denoised. In order to carry out this, the images are segmented such that the pixels are classified based on standardized characteristics. In the present research, region growth algorithm is adopted and the steps involved in it are as follows: [6–14].

1. Seed point selection: The selection of seed points depends on the image. For a gray-level lighted image, segment the lightning from the background. Then observe the histogram and select the seed points from its highest range.
2. Collection of additional image data: The connectivity or pixel neighboring information facilitates the determination of the threshold and seed points.
3. Determination of threshold value: In the region growing method, all regions must be within the threshold value which is fixed determined prior to the process.
4. Checking the similarity threshold value: If the difference in the pixel-value is less than the similarity threshold, that region is treated as the same region.

   Algorithm for means intensity calculation and stage of cancer

1. Read the MRI bone cancer image I.
2. Apply region growing algorithm to find the region of interest (ROI) of bone cancer tumor.
3. Compute the mean intensity value I_μ, using Eq. (1)

$$I\mu = \frac{s}{N} \tag{1}$$

   S Sum of pixel intensities for extracted tumor
   N Number of pixels for extracted tumor part

4. If (Mean Intensity < 245), it is identified as stage 1 cancer, else it is stage 2 cancer.

Initially the collected images are processed individually so as to segment them and capture the effected part. Figure 1 is the original gray scale image (MRI image of Thorax) that is used for the experimentation. On applying the algorithm as specified, Fig. 1a is generated that extracts the actual tumor part and based on which the mean intensity of the tumor is analyzed.
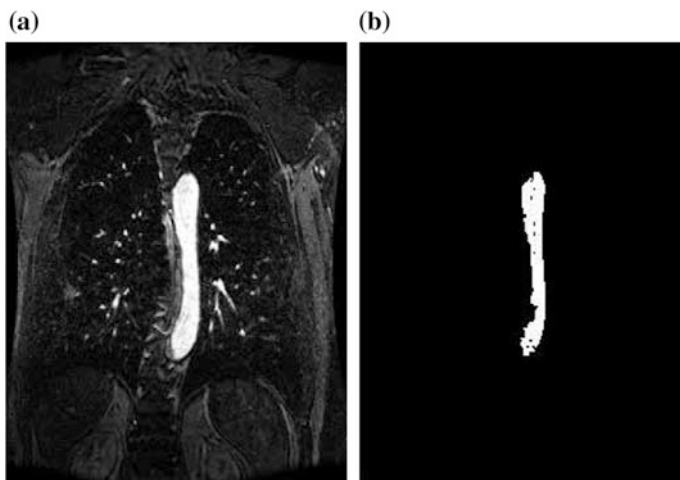
Fig. 1 **a** MRI image of thorax. **b** Segmented tumor part

Further, the images that contain tumor part are collected and analyzed in order to compute the number of pixels (N) and based on this value, the sum of pixels (S) is evaluated. Further, the mean intensity I_μ of the affected part is calculated, using Eq. (1) and the results are shown in Table 3. Upon the experimentation, if the mean intensity < 245, then it is identified as Stage 1 and if the mean intensity ≥ 245, it is determined to be at stage 2, shown in Table 3.

4   Results and Discussion

Research on bone cancer is now being done at many medical centers, university hospitals, and other institutions across the nation. There are several ongoing clinical trials focusing on bone cancer. It is to be identified at the earlier stages to prevent the death toll rate to a greater extent. Many researchers have put forward many theories which have turned to be inaccurate due to the lack of proper analysis and study of various parameters that involve in the formation of a bone cancer.

The bone cancer images obtained from radiology assistant website (www. radiologyassistant.nl) are used to analyze for the identification of cancer. The free accessibility of this information reflects the obligation of the radiology assistant to afford knowledge to wide viewers. Based on the experimentation, the values of mean intensity for the respective bone cancer images are tabulated in Table 3. Consider image 1, and as per the observed result it is affected with cancer and we evaluated mean intensity for the same after extracting the region of interest using region growing algorithm and the value is 243.22. In the similar manner, we

Table 3 Mean intensity evaluation and stage of cancer

Image No.	Observed	Mean intensity	Stage of cancer
1	Bone cancer	243.22	Stage 1
2	Bone cancer	246.84	Stage 2
3	Bone cancer	244.3	Stage 1
4	Bone cancer	248.13	Stage 2
5	Bone cancer	244.93	Stage 1
6	Bone cancer	245.82	Stage 2
7	Bone cancer	248.17	Stage 2
8	Bone cancer	246.03	Stage 2
9	Bone cancer	244.88	Stage 1
10	Bone cancer	244.24	Stage 1
11	Bone cancer	244.4	Stage 1
12	Bone cancer	244	Stage 1
13	Bone cancer	247.55	Stage 2`
14	Bone cancer	246.3	Stage 2
15	Bone cancer	244	Stage 1
16	Bone cancer	244.66	Stage 1
17	Bone cancer	244.43	Stage 1
18	Bone cancer	246.85	Stage 2
19	Bone cancer	245.81	Stage 2
20	Bone cancer	247.19	Stage 2
21	Bone cancer	244.38	Stage 1
22	Bone cancer	248.5	Stage 2
23	Bone cancer	246.47	Stage 2
24	Bone cancer	245.98	Stage 2
25	Bone cancer	245.6	Stage 2
26	Bone cancer	245.099	Stage 2
27	Bone cancer	249.2	Stage 2
28	Bone cancer	248.11	Stage 2
29	Bone cancer	247.86	Stage 2
30	Bone cancer	245.65	Stage 2

evaluated the mean intensities for different images and are tabulated in Table 3. Further we computed the stage of the cancer and the results are tabulated in Table 3.

5 Conclusion

Research on bone cancer is now being done at many medical centers, university hospitals, and other institutions across the nation. There are several ongoing clinical trials focusing on bone cancer. It is to be identified at the earlier stages to prevent

the death toll rate to a greater extent. Many researchers have put forward many theories which have turned to be inaccurate due to the lack of proper analysis and study of various parameters that involve in the formation of a bone cancer. Mining of data from MRI images is even more challenging and cumbersome task since it involves processing of not one but hundreds of MRI images in order to extract the required results. This paper presents a novel approach using region growing algorithm in the detection of mean intensity and stage of cancer.

References

1. Pabitra Roy and Sudipta Roy, "An Automated Method for Detection of Brain Abnormalities and Tumor from MRI Images", International Journal of Advanced Research in Computer Science and Software Engineering, 2013, pp. 1528–1589.
2. Leela G A and H.M Veena Kumari, "Morphological Approach for the Detection of Brain Tumour and Cancer Cells", Journal of Electronics and Communication Engineering Research, 2014, pp.07–12.
3. Manoj K Kowar and Sourabh Yadav, "Brain Tumor Detction and Segmentation Using Histogram Thresholding", International Journal of Engineering and Advanced Technology, 2012, pp. 16–20.
4. Adele P. Peskin and Karen Kafadar, "Robust Volume Calculations Of Tumors Of Various Sizes" Scientific Application and Visualization Group National Institute of Standards and Technology, Boulder, Colorado 80305 USA.
5. P. Rajeswari and G. Sophia Reena, "Human Liver Cancer Classification using Microarray Gene Expression Data," International Journal of Computer Applications, 2011, pp. 25–37.
6. Rohit S. Kabade and M. S. Gaikwad "Segmentation of Brain Tumour and Its Area Calculation in Brain MR Images using K-Mean Clustering and Fuzzy C-Mean Algorithm", International Journal of Computer Science & Engineering Technology, 2013, pp. 524–531.
7. Alan Jose, S. Ravi and M. Sambath. "Brain Tumor Segmentation Using K-Means Clustering And Fuzzy C-Means Algorithms And Its Area Calculation", International Journal of Innovative Research in Computer and Communication Engineering, vol. 2, issue 3, March 2014.
8. Jay Patel and Kaushal Doshi, "A Study of Segmentation Methods for Detection of Tumor in Brain MRI", Advance in Electronic and Electric Engineering, volume 4, number 3 (2014), pp. 279–284.
9. Aman Chandra Kaushik, and Vandana Sharma, "Brain Tumor Segmentation from MRI images and volume calculation of Tumor", International Journal of Pharmaceutical Science Invention, volume 2 issue 7 July 2013 pp. 23–26.
10. Milligan, G.W., Soon, S.C. and Sokol, M., "The Effect of Cluster Size, Dimensionality and the Number of Clusters on Recovery of True Cluster Structure," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1983, pp. 40–47.
11. Guralnik, V. and Karypis, G., "A Scalable Algorithm for Clustering Sequential Data," Proceedings of the IEEE International Conference on Data Mining Series, 2001, pp. 179–186.
12. Jarvis, R.A. and Patrick, E.A., "Clustering Using a Similarity Measure Based on Shared Nearest Neighbors," IEEE Transactions on Computers, 1973, pp. 1025–1034.
13. Modha, D. and Spangler, W., "Feature Weighting in k-means Clustering, "Machine Learning, Springer, 2002, pp. 217–237.
14. Adams Bischof, "Seeded Region Growing," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, pp. 641–647.

15. Navneet Kaur and Mamta Juneja, "A Novel Approach of Brain Tumor Detection Using Hybrid Filtering", Journal of Engineering Research and Applications, 2014, pp. 100–104.
16. Ashwini S. Zade and Mangesh Wanjari,"Detection of Cancer Cells in Mammogram Using Seeded Region Growing Method and Genetic Algorithm", Journal of Science & Technology, 2014, pp. 15–20.
17. Nikhil R Talegaonkar and Prashant N Shinde," an approach to automatic brain tumor detection inmagnetic resonance images", Proceedings of IRF International Conference, 13th April-2014, Pune, India, 17–19.
18. D. Narain Ponraj and M. Evangelin Jenifer, "A Survey on the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer", Journal of Emerging Trends in Computing and Information Sciences, vol. 2, December 2011, 656–664.
19. Sasikala and Vasanthakumar, "Breast Cancer-Classification and Analysis using Different Scanned Images," International Journal of Image Processing and Visual Communication, 2012, pp. 1–7.
20. Juliet R Rajan and Jefrin J Prakash, "Early Diagnosis of Lung Cancer using a Mining Tool," International Journal of Emerging Trends and Technology in Computer Science, 2013.
21. Michael Feldman, John Tomaszewski and Anant Madabhushi, "A Boosted Bayesian Multi resolution Classifier for Prostate Cancer Detection From Digitized Needle Biopsies," IEEE Transactions on Biomedical Engineering, 2012, pp. 1205–1218.
22. Karthikeyan Ganesan, U. Rajendra Acharya, Chua Kuang Chua, Lim Choo Min, K. Thomas Abraham and Kwan-Hoong Ng, "Computer-Aided Breast Cancer Detection Using Mammograms: A Review," IEEE Review in Biomedical Engineering, 2013, pp. 77–98.
23. Samir Kumar Bandyopadhyay and Tuhin Utsab Paul, "Segmentation of Brain Tumour from MRI image Analysis of k-means and DBSCAN Clustering," International Journal of Research in Engineering and Science, 2013, pp. 77–98.

BCI for Comparing Eyes Activities Measured from Temporal and Occipital Lobes

Sachin Kumar Agrawal, Annushree Bablani and Prakriti Trivedi

Abstract Brain–computer interface (BCI) is a system which communicates between user and machine. It provides a communication channel without using muscular activity. BCI uses brain rhythms as input of BCI system which are recorded by invasive or non-invasive BCI. Brain rhythms are always generated by brain when we are thinking, sleeping, deep sleeping, working and non-working states. In this paper, analysis of data recorded from electrodes placed at occipital and temporal lobes and comparison between both lobes using open and close eyes data.

Keywords EEG · Brain–computer interface · EEGLAB · Temporal lobe · Occipital lobe

1 Introduction

Brain–computer interface (BCI) is a system. As we know that a system is a middleware of two or more objects, as operating system communicates between user and computer hardware similarly BCI system communicates between user and machine through brain signals and these brain signals are called brain rhythms. BCI uses two types of communication: invasive and non-invasive BCI. Electroencephalography (EEG) is a non-invasive technique to record brain signals. Today non-invasive

S.K. Agrawal (✉) · Annushree Bablani · Prakriti Trivedi
Department of Computer Science and Engineering,
Government Engineering College Ajmer, Ajmer, Rajasthan, India
e-mail: sachinkumar.aggarwal@gmail.com

Annushree Bablani
e-mail: annushree.bablani@gmail.com

Prakriti Trivedi
e-mail: prakrititrivedi@rediffmail.com

technique is popular because it does not physically harm the subject, but signal quality is poor than invasive technique. BCI is useful for people, who are not able to control their body organs, but the mental capacity is normal. Every living being is doing some task like thinking, sleeping, working, etc. so brain generates many events when these tasks are performed. An event is triggered at specific time. For example when we open eyes and close eyes, our brain generates many events. In our work we have detected feature eyes open and eyes close effects on occipital and temporal lobes and compared the results retrieved from both lobes with open and close eyes.

2   Classifications of BCI

BCI can be classified in either of following three ways.

2.1   Input Types

BCI takes input in various ways, such as invasive and non-invasive techniques. In invasive, BCI signals are recorded with surgery of user or subject. This surgery can break the brain skin but it does not damage any neurons. Invasive BCI captures good quality signals, has very good spatial resolution and high frequency range, but can physically harm the subject. In non-invasive BCI, signals are recorded without surgery and do not physically harm the subject. Electroencephalography (EEG) is a way of recording electrical activity from the scalp using electrodes. This method is used for client and research purpose. Non-invasive BCI is very popular today but to only limitation is weak signal quality, more noise data and less frequency range.

2.2   Processing Types

BCI provides various types of methods for processing, and these methods use different purpose for different application area. BCI processing methods are synchronous, asynchronous, dependent and independent BCI.

In Synchronous BCI, Events are executed step by step: first step completes then the other step starts and as the other step completes the next step starts. It has to be done till the last step, if any of steps is not complete or fails then whole process fails. For example if I make my left hand up and then right hand up then machine starts. If I am not doing same as above then machine will not start. Synchronous BCI is a queue with specific order; queue contains set of subtask.
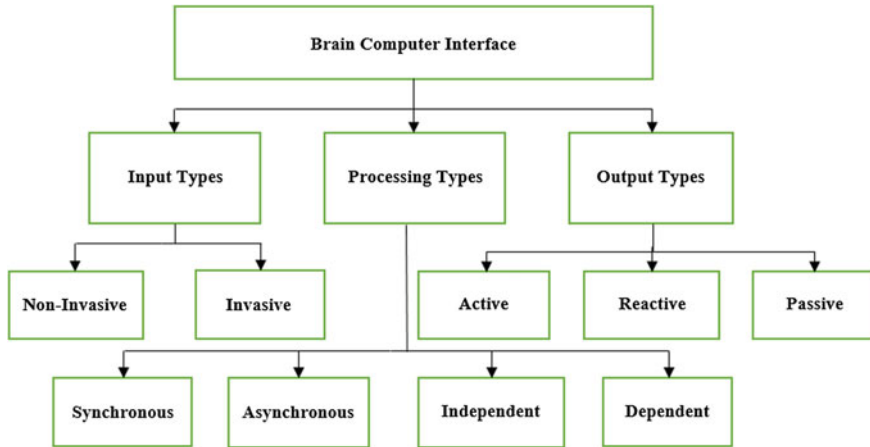
Fig. 1 Classification of BCI

When all subtasks have been completed (empty of queue) then only whole task is completed. Asynchronous BCI is not a queue-based processing. It does not any contain subtask; it is a set of tasks. Each task is run by single event. For example when I wake up then machine is started and when I sleep machine turns off (shut-down). Asynchronous BCI is more useful than synchronous BCI because it does not use queue whose event has to occur to complete task.

A dependent BCI does not use the brain's normal output pathway to carry the message, but activity in these pathways is needed to generate the brain activity that does carry it. EEG machine uses dependent BCI. For example when we open or close eyes then brain generates the muscular signal and via scalp these signals are recorded. Independent BCI does not depend on muscular activity. The user only thinks about the muscular action that user wants to perform and machine performs that action. It provides absolutely a new channel for communication. For example when we add two numbers our brain generates events, and these events are captured by device and perform task (Fig. 1).

2.3 Output Types

BCI can otherwise be classified as active, reactive and passive. Active BCI generates its output directly from brain activity, i.e. machine is controlled by thoughts of one's mind. It does not depend on external muscular events to trigger the generation of output. On the other hand, reactive BCI uses some command to be sent by focusing on some stimuli provided by the BCI system that will generate

event which we want to perceive. Passive BCI gives output of arbitrarily generated brain signals that are not under voluntary control.

3 Working

EEGLAB is an open-source toolkit which is run on MATLAB. EEGLAB is very flexible and easy to use because it provides graphical user interface (GUI) for new users. EEGLAB also provides tutorial, help windows and command history so that user can easily build a new script for own purposes. It provides various methods (independent component analysis, time/frequency analysis and artefact rejection) which helps user to analyse data.

3.1 Independent Component Analysis (ICA)

Independent component analysis algorithm can be applied to isolate artifactual and neurally generated EEG source. It maximises entropy of each component (independent of each other component) and it also decomposes speech and noise data. In EEGLAB, ICA works by default through runica() function. EEGLAB provides other functions that are similar to runica().

Figure 2 shows the continuous data and activity power spectrum. We can easily see that when frequency increases power decreases.
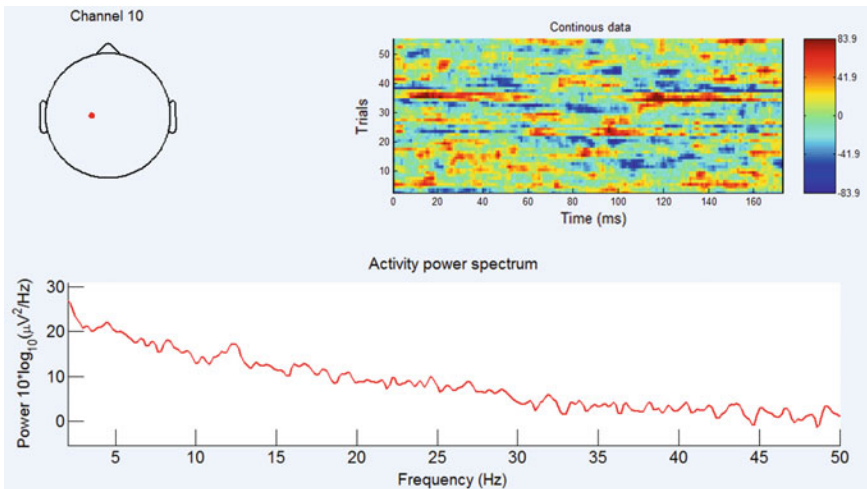
Fig. 2 Channel C1 (10) continuous data with activity power spectrum

3.2   Time and Frequency Decomposition

Its primary measures are event-related spectral perturbation (ERSP), inter-trial coherence (ITC) and event-related cross-coherence (ERCOH). ERSP measures mean event-related change in power spectrum at component or channel data. ITC measures magnitude and phase change in the power spectrum at a single channel or component data. ERCOH measures magnitude and phase change in the power spectrum at between two channels or component data. Normally, for n trials, if $f_k(f, t)$ is the spectral estimate of trial k at frequency f and time t then,

ERSP defined by

$$ERSP\ (f,t) = \frac{1}{n}\sum_{k=1}^{n} |f_k(f,t)|^2 \tag{1}$$

Inter-trial phase coherence is defined by

$$ITPC\ (f,t) = \frac{1}{n}\sum_{k=1}^{n} \frac{f_k(f,t)}{|f_k(f,t)|} \tag{2}$$

And inter-trial linear coherence defined by

$$ITLC\ (f,t) = \frac{\sum_{i=1}^{n} f_k(f,t)}{\sqrt{n \sum_{k=1}^{n} |f_k(f,t)|^2}} \tag{3}$$

To compute $F_k(f, t)$ EEGLAB uses the short-time Fourier transform, a sinusoidal wavelet transform decomposition that provides a specified time and frequency resolution.

3.3   Rejecting Artefacts

EEGLAB provides user to remove non-neural artefacts of channel, epoch and component data by GUI interface. EEGLAB provides reading data, events information, rejecting data and channel location file in different formats (Binary, Matlab, ASCII, Neuroscan, EGI, European Standard BDF, EDF, EDF+, etc.).

4   Result and Discussion

We have used EEG data for our research (experiment) which have been downloaded from PhysioNet. This dataset consists of 1 or 2 min EEG recordings and this data is in EDF+ format (European Data Format) containing 64 EEG electrodes as

per the international 10-10 system, each sampled at 160 samples per second. We are choosing eyes open and close effect on temporal lobe and occipital lobe. We have selected six electrodes T7, T8, T9, T10, TP7 and TP8 for temporal lobe and three electrodes O1, Oz and O2 for occipital lobe.

Table 1 Subject max and min values of opened and closed eyes with respective electrodes (all value in rms microvolt)

Electrodes name	Open eyes		Close eyes	
	Max	Min	Max	Min
T7	104.000	−80.000	169.000	−201.000
T8	79.000	−59.000	166.000	−124.000
T9	62.000	−48.000	154.000	−88.000
T10	118.000	−94.000	111.000	−131.000
TP7	78.000	−84.000	147.000	−121.000
TP8	81.000	−109.000	142.000	−97.000
O1	65.000	−58.000	105.000	−74.000
Oz	71.000	−58.000	97.000	−69.000
O2	66.000	−78.000	100.000	−86.000
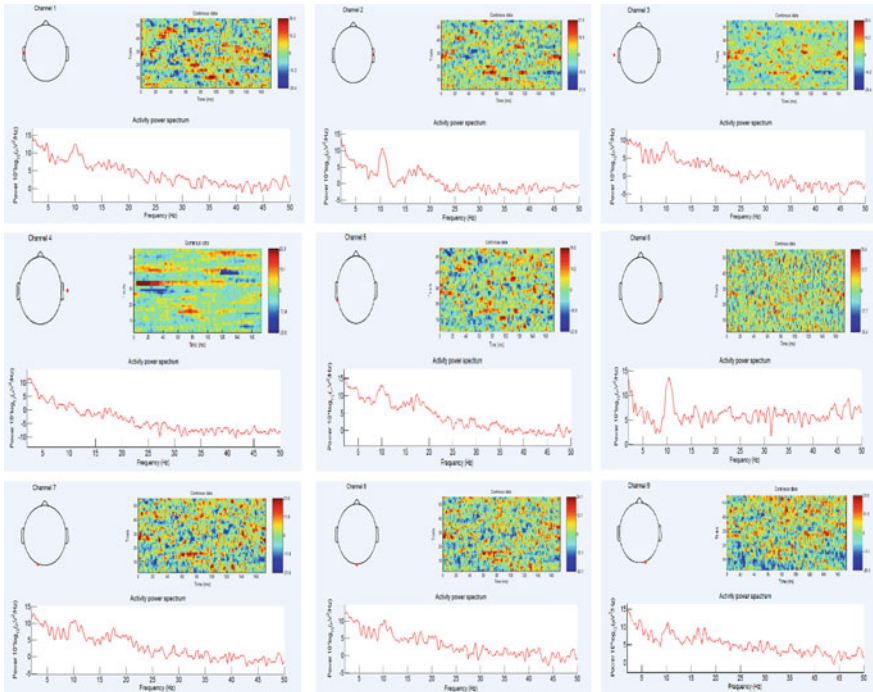
Fig. 3 Channel 1–9 continuous data and activity power spectrum of opened eyes
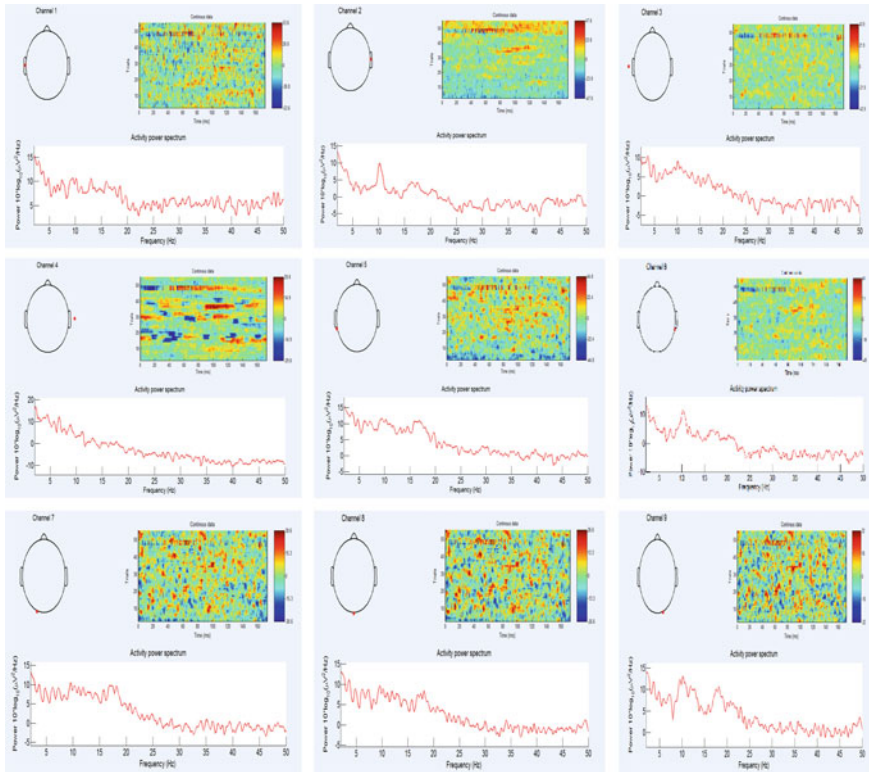
Fig. 4 Channel 1–9 continuous data and activity power spectrum of closed eyes

5 Conclusion

EEG is the best non-invasive BCI technique to record what is going on one's mind and to analyse these data we have EEGLAB, an interactive tool. For our work we have downloaded data recorded by EEG machine and investigated that data using EEGLAB toolbox. From our work it can be easily shown that occipital lobe is always more active when (eyes are closed) from the Table 1 data. Occipital lobe and temporal lobe both are more active at alpha wave (8–13 Hz) from Figs. 3 and 4. If max and min values of closed eyes are greater than min and max values of opened eyes, then we can say that occipital and temporal lobes become more active when eyes are closed and subject is in relax state and does not hear any sound.

Further Reading

1. http://sccn.ucsd.edu/eeglab.
2. http://www.physionet.org/pn4/eegmmidb/.
3. Vallabhaneni, Anirudh and Wang, Tao and He, Bin, Brain—Computer Interface (2005). DOI 10.1007/0-306-48610-5_3 © Springer US pages 85–121.
4. B. Graimann et al. (eds.), Brain–Computer Interfaces, The Frontiers Collection, DOI 10.1007/978-3-642-02091-9_1, © Springer-Verlag Berlin Heidelberg 2010.
5. Arnaud Delorme and Scott Makeig, "an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis", Journal of Neuroscience Methods 134 (2004), page: 9–21.
6. Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, Theresa M. Vaughan, Brain–computer interfaces for communication and control 2 March 2002.
7. Jonathan R. Wolpaw and Niels Birbaumer Laboratory of Nervous System Disorders, Wadsworth Center, NYS Department of Health, Albany, NY, USA and Institute Behavioural Neuroscience, Eberhard-Karls-University, Tubingen, Germany 21/10/2005.

An Adaptive Edge-Preserving Image Denoising Using Block-Based Singular Value Decomposition in Wavelet Domain

Paras Jain and Vipin Tyagi

Abstract Image denoising is a quite active research area in the domain of image processing. The essential requirement for a good denoising method is to preserve significant image structures (e.g., edges) after denoising. Wavelet transforms and singular value decomposition (SVD) have been independently used to achieve edge-preserving denoising results for natural images. Numerous denoising algorithms have utilized these two techniques independently. In this paper, a novel technique for edge-preserving image denoising, which combines wavelet transforms and SVD, is proposed. It is adaptive to the inhomogeneous nature of natural images. The multiresolution representation of the corrupted image in wavelet domain is obtained through the application of a discrete wavelet transform to it. A block-SVD based edge-adaptive thresholding scheme which relies on estimation of noise level is employed to reduce the noise contents while preserving significant details of the original version. Comparison of the experimental results with other state-of-the-art methods reveals the fact that the proposed approach achieves very impressive gain in denoising performance.

Keywords Wavelet · Image denoising · Edge-preservation · Singular values

1 Introduction

During the process of acquisition and transmission, a corruption due to noise can be occurred in digital images. The major challenge for a filtering technique is to suppress noise without eliminating the relevant details of an image. Numerous denoising algorithms [1–3] have been suggested during past few decades, linear

P. Jain (✉)
JK Lakshmipat University, Jaipur, Rajasthan 302026, India
e-mail: dr.parasjain.cse@gmail.com

V. Tyagi
Jaypee University of Engineering and Technology, Raghogarh, Guna, MP 473226, India
e-mail: dr.vipin.tyagi@gmail.com

spatial domain filters were the bases for many of them. Linear spatial filters [1, 2] are easy to implement but they tend to blur some fine image structures, such as edges [1].

To resolve the issues raised with linear filters, filtering based on nonlinear edge-preserving methods has gained the prime attention in the field and research is still continuing in this direction [4]. The wavelet transforms have received much attention in edge-preserving denoising problems in past few decades.

The basic approach of noise suppression using wavelet transform involves the following main stages. First, the multiscale decomposition of the corrupted image into the wavelet coefficients is obtained. Second, these wavelet coefficients are manipulated by applying a predetermined threshold on them according to a shrinkage rule. Finally, reconstruction from these manipulated coefficients then produces the desired denoised image. Numerous approaches use the concept of wavelet thresholding for image denoising [5–16].

Numerous image denoising approaches based on either wavelet transform or SVD have been independently developed; however, their connection has rarely been addressed. Hou [17] investigated a relation between wavelet transform and SVD and proposed an approach for image denoising by performing edge-adaptive SVD filtering in detail subbands of wavelet domain. Author established a fact that the combination of wavelet transform and SVD filtering could yield effective methods for image denoising. The method proposed in this paper is also inspired by the denoising approach in [17].

2 Proposed Technique

Generally, images are the signals that are nonstationary in nature and are composed of small homogeneous regions. There may be significant differences among these regions. This fact has encouraged the researchers to use the divide-and-conquer techniques for image restoration, for example, the piecewise and local image model suggested in [18]. The proposed method also follows divide-and-conquer strategy by dividing each detail subband into blocks and conquering each block individually by SVD-based edge-adaptive thresholding. This approach is time efficient in implementation and achieves better denoising performance. The main stages of the proposed denoising method are illustrated in Fig. 1.

Similar to the adaptive SVD filtering in wavelet domain (WASVD) proposed by Hou [17], the proposed approach: divides the detail subbands of wavelet domain into blocks; applies SVD-based edge detection to each block for checking the presence of edge structure in that block; associates different thresholds with edge-present and edge-absent blocks. However, unlike the WASVD, the proposed approach: instead of using fixed thresholds T_{nonedge} for all edge-absent blocks and T_{edge} for all edge-present blocks, determines these thresholds locally for each block by considering the coefficients in the block and noise variance; instead of using the fixed value of the term noise variance (used in the estimation of thresholds) for all
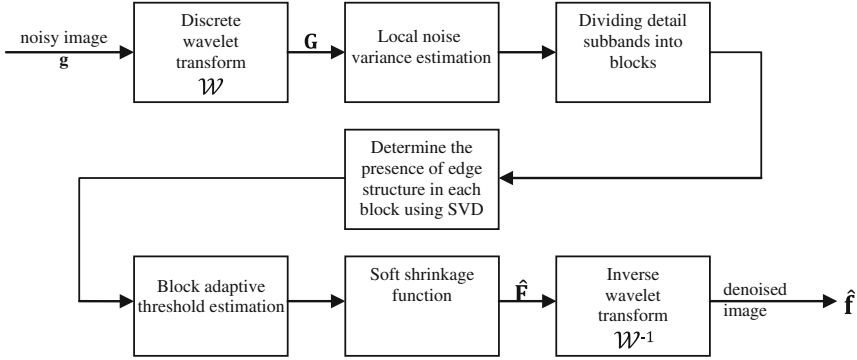
Fig. 1 Proposed denoising method

resolution scales, estimates its value locally for each resolution scale through a robust median estimator [5]; instead of using SVD filtering, performs the noise suppression through wavelet thresholding.

Let an additive noise ξ corrupt an image f according to the model as follows:

$$g = f + \xi, \tag{1}$$

where g is the observed (noisy) image. Here, the noise ξ is modeled as i.i.d (independently and identically distributed) Gaussian variable with zero mean and standard deviation σ, denoted as $\mathcal{N}(0, \sigma^2)$.

Initially, the input noisy image g is decomposed into different frequency subbands (one smooth subband at coarsest scale and three details subbands at each resolution scale) through a discrete wavelet transform $\mathcal{W}$ as

$$G = \mathcal{W}(g) = \left\{ \mathrm{LL}_J; \mathrm{LH}_j; \mathrm{HL}_j; \mathrm{HH}_j, j = J, J-1, \ldots, 1 \right\}, \tag{2}$$

where j is the scale and J is the largest scale in the decomposition. The lesser j is, the finer the scale is.

Let us use the notations g^J, S_1^j, S_2^j and $S_3^j, j = 1, \ldots, J$ for the subbands $\mathrm{LL}_J, \mathrm{LH}_j, \mathrm{HL}_j$ and HH_j, respectively. Consider a detail subband S at scale j, i.e., $S = S_l^j, l = 1, 2, 3, j = 1, \ldots, J$. Note that the additive noise model for image g in spatial domain in Eq. (1) is also applicable for the subband S_l^j in wavelet domain. Thus, we have

$$S_l^j = w_l^j + n_l^j, \tag{3}$$

where the noisy subband S_l^j is obtained after introducing the noise n_l^j in its noiseless counterpart w_l^j. The main concern of the proposed approach is to obtain $\hat{w}_l^j$, the estimate of w_l^j, from its noisy observation S_l^j. The proposed procedure works as follows.

The term local noise variance $\hat{\sigma}^2_{\text{noise},j}$ can be estimated as

$$\hat{\sigma}_{\text{noise},j} = \frac{\text{median}(|G_{xy}|)}{0.6745}, G_{xy} \in S_3^j \tag{4}$$

Let the subband S_l^j of size $m \times n$ be divided into blocks B_{kl} of size $p \times q (p \geq q), k = 1, \ldots, \lceil m/p \rceil, l = 1, \ldots, \lceil n/q \rceil$, where $\lceil \cdot \rceil$ is the ceiling function.

The threshold λ_{kl} on the block B_{kl} is computed by using the BayesShrink [7, 8] as threshold estimation criterion, expressed as

$$\lambda_{kl} = \frac{\hat{\sigma}^2_{\text{noise},j}}{\sigma_{\text{signal},kl}} \tag{5}$$

The term $\sigma_{\text{signal},kl}$ in Eq. (5) is the local signal deviation estimated on block B_{kl}, obtained as

$$\sigma_{\text{signal},kl} = \sqrt{\max\left(\hat{\sigma}^2_{B_{kl}} - \hat{\sigma}^2_{\text{noise},j}, 0\right)}, \tag{6}$$

where $\hat{\sigma}^2_{B_{kl}} = \frac{1}{N_{pq}} \sum_{x,y=1}^{N_{pq}} G_{xy}^2$ and N_{pq} is the number of wavelet coefficients G_{xy} in the block B_{kl}.

Now SVD will be used to check the presence of an edge structure in each block B_{kl}. In the theory of SVD [19], a real-valued $p \times q$ matrix B_{kl} can be decomposed uniquely as

$$B_{kl} = U\Sigma V^t = \sum_{i=1}^{s} \alpha_i \vec{u}_i \vec{v}_i^t, \tag{7}$$

where $\vec{u}_i$ and $\vec{v}_i$ are, respectively, the column vectors of the orthogonal matrices U and V. $\Sigma = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a diagonal matrix with diagonal elements $\alpha_i, i = 1, \ldots, s$. The diagonal elements of Σ are the singular values of B_{kl} which are arranged in a nonincreasing order. Let $\text{median}(\{\alpha_i\})$ denote the median of the singular values of block B_{kl}. If the condition given as

$$\text{median}(\{\alpha_i\}) > \beta \hat{\sigma}_{\text{noise},j}, \tag{8}$$

holds true then the underlying block B_{kl} contains an edge structure otherwise not. The parameter β is selected as suitable positive number. The block B_{kl} with edge structure must be provided a special treatment at the time of thresholding. A new threshold λ'_{kl} for these blocks (suggested in the proposed method) is determined as

$$\lambda'_{kl} = \gamma \lambda_{kl}, 0 \leq \gamma \leq 1 \tag{9}$$

where γ is a weight factor used to associate thresholds to the edge-present blocks smaller than the thresholds associated with edge-absent blocks. Now we will obtain the thresholded version $\hat{B}_{kl}$ of each block B_{kl} by applying its corresponding threshold according to a soft shrinkage function [9] as follows:

$$\hat{B}_{kl} = T_{\text{soft}}\big(B_{kl}, \lambda'_{kl}\big) = \text{sgn}\big(G_{xy}\big)\max\big(\big|G_{xy}\big| - \lambda'_{kl}, 0\big), \forall G_{xy} \in B_{kl} \qquad (10)$$

where the function $\text{sgn}(x)$ gives the information about the sign of argument x.

Therefore, thresholded results on each of the candidate block B_{kl} in subband S_l^j are obtained through Eq. (10). In all, we can obtain the thresholded version of subband S_l^j as

$$\hat{w}_l^j = \big\{\hat{B}_{kl}, k = 1, \ldots, \lceil m/p \rceil, l = 1, \ldots, \lceil n/q \rceil\big\} \qquad (11)$$

Thus, the final thresholded result of complete image in wavelet domain can be obtained as

$$\hat{F} = \big\{g^J; \hat{w}_1^j; \hat{w}_2^j; \hat{w}_3^j, j = J, J - 1, \ldots, 1\big\} \qquad (12)$$

Finally, the desired denoised image $\hat{f}$ can be determined through an inverse discrete wavelet transform $\mathcal{W}^{-1}$ to the thresholded wavelet domain image $\hat{F}$ as

$$\hat{f} = \mathcal{W}^{-1}\big(\hat{F}\big) \qquad (13)$$

3 Experimental Results

The block-SVD based noise suppression technique presented in this paper is tested on several images, corrupted due to the simulated noise $\mathcal{N}(0, \sigma^2)$, chosen from a test set which comprises the images from image set [20] along with popular images such as *Lena*, *cameraman*, and *peppers*. We started with estimation of the parameters that govern the method elaborated in Sect. 2: wavelet type, number of resolution levels J, block size $p \times q$, parameter β and weight factor γ. Various wavelets from different wavelet families [21] like Daubechies, Symlet, etc., are checked for making a choice. Symlet-8 (sym8) with four resolution levels (i.e., $J = 4$) is found suitable for our experiments. So this combination was fixed throughout the comparison process to make it fair.

The block size $p \times q$ is another parameter of concern. The detail discussion to make the optimal choice for this parameter under various conditions is given in [17]. Our experiments indicate that $p \times q = 8 \times 8$ is a better choice.

Based on the experimental observations, the best suitable values for β (Eq. 8) and γ (Eq. 9) came out to be 4 and 0.5, respectively.

To validate the effectiveness of our method, comparisons are made with some well-known denoising methods, namely SURELET [11], Bayes [8] and WASVD [17]. PSNR (in dB) and SSIM [22] results for the denoised images are depicted in Table 1. The boldfaced entries are demonstrating the best values among the values obtained for all the methods being compared. The quantitative analysis through the results given in Table 1 and the qualitative analysis through the visual results illustrated in Fig. 2 reveal the fact that the proposed approach outperforms other methods, especially considering Bayes and WASVD methods.

Table 1 Performance of various methods as measured by PSNR and SSIM

	PSNR results				SSIM results			
	[11]	[8]	[17]	Proposed	[11]	[8]	[17]	Proposed
Lena								
$\sigma = 10$	**34.76**	33.54	33.45	34.52	0.90	0.87	0.85	**0.91**
$\sigma = 15$	32.56	31.32	31.02	**32.86**	0.86	0.83	0.78	**0.88**
$\sigma = 20$	31.30	30.10	29.09	**31.79**	**0.84**	0.80	0.70	**0.84**
$\sigma = 25$	30.33	29.43	27.98	**30.85**	**0.82**	0.78	0.63	**0.82**
$\sigma = 30$	29.15	28.56	26.96	**29.71**	0.80	0.76	0.57	**0.81**
$\sigma = 35$	28.36	28.19	26.03	**28.82**	0.78	0.74	0.52	**0.79**
Cameraman								
$\sigma = 10$	33.77	35.16	34.74	**36.21**	0.93	0.89	0.87	**0.94**
$\sigma = 15$	31.44	32.93	32.11	**34.03**	**0.90**	0.84	0.78	**0.90**
$\sigma = 20$	29.79	31.58	30.20	**32.57**	0.87	0.80	0.69	**0.88**
$\sigma = 25$	28.55	30.34	28.79	**31.32**	0.85	0.76	0.62	**0.86**
$\sigma = 30$	27.28	29.54	27.68	**30.64**	**0.83**	0.74	0.55	**0.83**
$\sigma = 35$	26.78	29.03	26.84	**30.17**	**0.81**	0.59	0.49	**0.81**
Barbara								
$\sigma = 10$	32.28	31.10	32.00	**32.47**	0.90	0.85	0.88	<u>**0.91**</u>
$\sigma = 15$	30.02	29.54	29.57	**30.81**	0.85	0.79	0.81	<u>**0.88**</u>
$\sigma = 20$	28.47	27.50	27.83	**29.23**	0.80	0.74	0.74	<u>**0.85**</u>
$\sigma = 25$	27.30	26.67	26.68	**27.88**	0.75	0.70	0.68	<u>**0.80**</u>
$\sigma = 30$	25.94	25.47	25.74	**27.23**	0.71	0.67	0.62	<u>**0.78**</u>
$\sigma = 35$	24.85	24.97	25.03	**26.72**	0.67	0.64	0.57	<u>**0.74**</u>
Man								
$\sigma = 10$	**32.41**	31.44	31.35	32.16	**0.88**	0.86	0.83	**0.88**
$\sigma = 15$	**30.47**	29.76	29.09	30.05	0.83	0.80	0.75	**0.85**
$\sigma = 20$	**29.23**	28.50	27.49	28.82	**0.79**	0.75	0.68	**0.79**
$\sigma = 25$	**28.52**	27.36	26.39	27.96	0.75	0.71	0.61	**0.77**
$\sigma = 30$	27.32	27.04	25.52	**27.44**	**0.73**	0.69	0.56	**0.73**
$\sigma = 35$	**26.64**	25.94	24.50	26.59	0.70	0.67	0.51	**0.71**

(continued)

Table 1 (continued)

	PSNR results				SSIM results			
	[11]	[8]	[17]	Proposed	[11]	[8]	[17]	Proposed
Boat								
$\sigma = 10$	**33.21**	32.07	32.05	32.84	0.90	0.86	0.85	**0.92**
$\sigma = 15$	31.17	30.14	29.76	**31.51**	0.86	0.80	0.77	**0.89**
$\sigma = 20$	29.85	28.80	28.26	**30.67**	0.82	0.76	0.70	**0.84**
$\sigma = 25$	28.93	27.48	27.18	**29.38**	0.79	0.73	0.63	**0.81**
$\sigma = 30$	28.06	26.77	26.14	**28.62**	0.76	0.71	0.57	**0.78**
$\sigma = 35$	27.40	26.25	25.34	**27.88**	0.74	0.68	0.52	**0.75**
Peppers								
$\sigma = 10$	**33.38**	32.38	31.89	32.82	**0.88**	0.84	0.81	**0.88**
$\sigma = 15$	31.57	30.51	30.04	**31.67**	0.84	0.79	0.74	**0.85**
$\sigma = 20$	**31.05**	29.59	28.68	30.80	0.81	0.76	0.67	**0.82**
$\sigma = 25$	**29.65**	28.88	27.61	29.51	0.79	0.73	0.60	**0.80**
$\sigma = 30$	**29.77**	27.95	26.76	29.15	**0.77**	0.71	0.55	**0.77**
$\sigma = 35$	28.43	26.15	25.98	**28.77**	0.75	0.56	0.50	**0.76**

Lena Cameraman Barbara Peppers

Fig. 2 Denoising results for test images using proposed method

4   Conclusions

This paper presented an edge-preserving noise reduction technique using an edge-adaptive block-SVD in wavelet domain. With this technique, we have reached several remarkable conclusions. First, the approach of divide-and-conquer used in our method adapts to the inhomogeneous nature of natural images. Second, block-dependent noise suppression which involves the estimation of threshold locally for each block enhances the denoising results as it can depict the local structures well in comparison to a subband-dependent thresholding. Third, estimating the noise variance locally at each resolution level strengthens it, as it considers the noise strength of that resolution level. Fourth, edge-adaptive thresholding is applied to each block which considers the edge strength of that block for better edge-preservation.

References

1. Gonzalez, R.C., Woods, R.E.: Digital image processing. Prentice-Hall, Upper Saddle River (2008).
2. Shapiro, L., Stockman, G.: Computer Vision. Prentice Hall (2001).
3. Jain, P., Tyagi, V.: Spatial and frequency domain filters for restoration of noisy images. IETE Journal of Education 54(2), 108–116 (2013).
4. Jain, P., Tyagi, V.: A survey of edge-preserving image denoising methods. Information Systems Frontiers, 1–12 (2014), doi:10.1007/s10796-014-9527-0.
5. Donoho, D.L., Johnstone, I. M.: Ideal spatial adaptation via wavelet shrinkage. Biometrika 81, 425–455 (1994).
6. Donoho, D.L., Johnstone, I. M.: Adapting to unknown smoothness via wavelet shrinkage. Journal of the American Statistical Association 90(432), 1200–1224 (1995).
7. Chipman, H., Kolaczyk, E., McCulloc, R.: Adaptive Bayesian wavelet shrinkage. Journal of the American Statistical Association 440(92), 1413–1421 (1997).
8. Chang, S., Yu, B., Vetterli, M.: Adaptive wavelet thresholding for image denoising and compression. IEEE Transactions on Image Processing 9(9), 1532–1546 (2000).
9. Donoho, D.L.: De-noising by soft-thresholding. IEEE Transactions on Information Theory 41(3), 613–627 (1995).
10. Sendur, L., Seleselnick, I. W.: Bivariate shrinkage with local variance estimation. IEEE Signal Processing Letters 9(12), 438–441 (2002).
11. Blu, T., Luisier, F.: The SURE-LET approach to image denoising. IEEE Transactions on Image Processing 16(11), 2778–2786 (2007).
12. Chang, S., Yu, B., Vetterli, M.: Spatially adaptive wavelet thresholding based on context modeling for image denoising. IEEE Transactions on Image Processing 9(9), 1522–1531 (2000).
13. Jain, P., Tyagi, V.: An adaptive edge-preserving image denoising technique using tetrolet transforms. The Visual Computer 31(5), 657–674 (2015).
14. Jain, P., Tyagi, V.: LAPB: Locally adaptive patch-based wavelet domain edge-preserving image denoising. Information Sciences, 294, 164–181 (2015).
15. Jain, P., Tyagi, V.: An adaptive edge-preserving image denoising using epsilon-median filtering in tetrolet domain. Advances in Intelligent Systems and Computing 337, 393–400 (2015).

16. Jain, P., Tyagi, V.: An adaptive edge-preserving image denoising using arbitrarily shaped local windows in wavelet domain. International Journal of Computer Applications 114(15), 33–45 (2015).
17. Hou, Z.: Adaptive singular value decomposition in wavelet domain for image denoising. Pattern Recognition 36, 1747–1763 (2003).
18. Acton, S.T., Bovik, A. C.: Nonlinear image estimation using piecewise and local image models. IEEE Transaction on Image Processing 7, 979–991 (1998).
19. Golub, G. H., Van Loan C. F.: Matrix Computations. Baltimore MD John Hopkins University Press (1983).
20. http://decsai.ugr.es/cvg/CG/base.htm.
21. Nason, G. P., Silverman, B. W.: The stationary wavelet transform and some statistical applications. In lecture notes in statistics: wavelets and statistics, Springer-Verlag, Berlin, 281–300 (1995).
22. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004).

Performance Analysis of Voltage Controlled Ring Oscillators

Shruti Suman, K.G. Sharma and P.K. Ghosh

Abstract The voltage controlled ring oscillator (VCO) is a critical and necessary component in data communication systems and clock recovery circuits. It is basically an oscillator, whose output frequency is controlled by the input control voltage. Finally, it has been concluded that the selected VCOs have large varying characteristics. This paper comprehensively analyzes different ring VCOs and their performances have been summarized in terms of frequency range, power consumption and bandwidth, which enable the designers to select the most appropriate ring VCO for specific applications.

Keywords VCO · Ring VCO · Inverter · Delay stages

1 Introduction

Ring oscillator plays a very important role in analog and digital communication systems due to low power consumption, wide frequency range of operation and its high integration capability. Designing a ring VCO involves area, power and speed, and its applicability follows some trade-offs [1]. Some important applications are medical implant transceivers [2], radio frequency identification systems (RFID)

Shruti Suman (✉) · P.K. Ghosh
ECE Department, College of Engineering and Technology,
Mody University of Science and Technology, Lakshmangarh,
Sikar 332311, Rajasthan, India
e-mail: shrutisuman23@gmail.com

P.K. Ghosh
e-mail: pkghosh.ece@gmail.com

K.G. Sharma
ECE Department, Chandigarh College of Engineering & Technology,
Sector-26, Chandigarh 160019, India
e-mail: sharma.kg@gmail.com

29

transponders [3], etc. The working of each delay stage of VCO dictates the performance of voltage controlled ring oscillator. The aim of this paper is to compare different ring VCOs and analysis is done using single delay cell of each ring VCO. This compares the performance of different VCOs, so that the design solution can be achieved in terms of frequency range, power consumption and bandwidth.

The organization of the paper is summarized in other three sections: The overview and linear analysis of ring oscillator is done in Sect. 2. The different delay stages of ring VCOs are analyzed and results are summarized in Sect. 3 and finally Sect. 4 concludes.

2 CMOS Ring Oscillator

The ring oscillator consists of odd number of gain stages connected in a form of a chain [4]. Due to odd number of inversion, this circuit does not have stable operating point and so oscillates. The block diagram of N-stage (N = odd) ring VCO is shown in Fig. 1.

The output of Nth stage is fed back to the input of first stage. To sustain oscillation, Barkhausen criteria must be satisfied [5]. The propagation time (t_p) of transition of signal through the complete chain determines the period of ring oscillator and is given by the Eq. (1).

$$T = 2 \times N \times t_p \tag{1}$$

Here, N is the number of inverters (delay stages) in the chain. The factor 2 results due to the fact that a complete cycle requires a low to high and high to low transitions. The Eq. (1) is valid only for $2Nt_p \gg t_f + t_r$ (t_r and t_f are the rise and fall time periods, respectively).

2.1 Linear Model of Ring Oscillator

In order to derive the formula for oscillation frequency of ring oscillator, we use its linear model as shown in Fig. 2.

Fig. 1 Single ended N-stage ring VCO
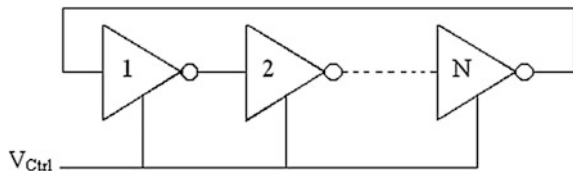
Fig. 2 Linear model of
N stage ring oscillator
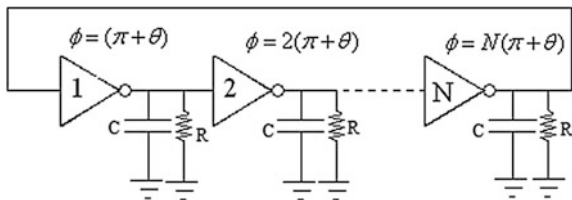
We assume that all the inverting stages are identical and we model them as a transconductance (g_m), loaded by a parallel combination of resistor R and capacitor C. The gain $L(s)$ of the inverting stage is defined by Eq. (2) as

$$L(s) = A_1(j\omega) = A_2(j\omega) = \cdots = A_N(j\omega) \tag{2}$$

The transfer function $A(j\omega)$ of each stage is given by

$$A(j\omega) = \left[\frac{-g_m R}{1 + j\omega RC}\right] \tag{3}$$

The conditions for sustained oscillation can be written as [6]

$$|A_N(j\omega)| = 1 \tag{4}$$

$$\angle A(j\omega) = \theta = \tan^{-1}(\omega RC) = \frac{2N\pi}{RC} \tag{5}$$

Thus for N identical stages of delay, the oscillation frequency can be calculated from the above equations and is given by

$$f_0 = \frac{1}{2Nt_d} \tag{6}$$

Here, t_d is the delay time per stage. Thus, the oscillation frequency is inversely related to and depends only on the delay time (t_d) and the number N of delay stages. Also for the given interval of time, the signal passing through each stage is twice.

3 Different Delay Stages of Ring VCOs

There are numerous ways by which a ring oscillator can be realized. This primarily depends on the implementation of different structural forms. The designs are analyzed and summarized in terms of relevant parameters.

3.1 Current Starved Ring VCO

The current starved ring VCO illustrates a wide oscillation frequency range of 66–875 MHz at 1.8 V supply voltage using 0.18 μM CMOS technology [7]. The circuit formed by using ring oscillator consists of odd number of gain stages connected in series and bias stage consists of current sink and current source. The oscillation is performed by ring oscillator and the frequency tuning is achieved by controlling the supply current. The variable bias currents are used to control the oscillation frequency of this ring VCO as illustrated in Fig. 3. The transistors M1 and M2 operate as inverters while M3 and M4 operate as current sink and current source, respectively. The current sources limit the current available to inverters. The drain currents of transistors M5 and M6 are same and are set by the input control voltage (V_{ctrl}). The current in transistors M4 and M5 is mirrored from bias stage to each cascaded inverting stage. The bias circuit is used to provide correct polarization for transistors M3 and M4 [8, 9]. The tuning of frequency of oscillation for a wide range can be done by changing the value of control voltage and this is the benefit of this configuration.

Fig. 3 Delay stage for current starved ring VCO

The total capacitance on the drains of M3 and M4 is given by Eq. (7) as

$$
\begin{aligned}
C_{\text{total}} &= C_{\text{out}} + C_{\text{in}} \\
&= C_{\text{ox}}(W_pL_p + W_nL_n) + 3/2C_{\text{ox}}(W_pL_p + W_nL_n)
\end{aligned}
\tag{7}
$$

where $C_{\text{in}}, C_{\text{out}}, C_{\text{ox}}$ and C_{total} are the input capacitance, output capacitance, oxide related capacitance, and total capacitance, respectively. The total capacitance is sum of the output and input capacitance of the inverter, and is given by

$$
C_{\text{total}} = \frac{5}{2} C_{\text{ox}}(W_pL_p + W_nL_n)
\tag{8}
$$

The oscillation frequency is determined by the bias current (I_d), number of stages (N), total capacitance (C_{total}) and control voltage (V_{ctrl}) as

$$
f_{\text{osc}} = \frac{I_d}{2NC_{\text{total}}V_{\text{ctrl}}}
\tag{9}
$$

The relation between the control voltage V_{ctrl} and supply current (I_c) is given by Eq. (10) [10].

$$
V_c - V_{\text{th}} = I_cR + \sqrt{\frac{2I_c}{k_n(W/L)}}
\tag{10}
$$

Here k_n is the transconductance and V_{th} is the threshold voltage of transistor.

The linearity and bandwidth of VCO are determined by variation of control voltage (V_{ctrl}). The main drawback of this circuit is that under low frequency, the current starved inverter suffers from slow rise and falls at its output.

3.2 Combined Ring VCO

This type of ring VCO is a combined structure [11] composed of odd number of stages implemented with both basic inverters (M1; M2 and M3; M4) and the current starved VCO in which M7, M5 and M6, M8 act as current mirror and controlling of VCO can be done by using transistor M9 as depicted in Fig. 4. A combined VCO provides better frequency stability as compared to the simple oscillators that use only one type of inverter stage. The delay can be considered as the sum of the time delays of all the inverters that are used in ring VCO and additional circuitry. The frequency range of the combined ring VCO was observed from 22.4 to 954.50 MHz with supply voltage of 3 V. The frequency stability can be achieved in terms of the

Fig. 4 Delay stage for combined ring VCO
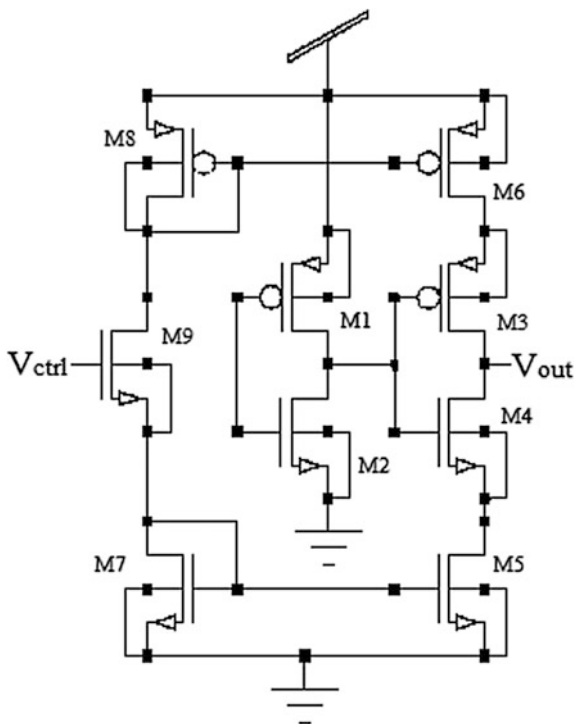
frequency deviation $(\Delta f_{\mathrm{osc}})$ with respect to the oscillation frequency (f_{osc}), and is defined as

$$\frac{\Delta f_{\mathrm{osc}}}{f_{\mathrm{osc}}} = \frac{f(V_{\mathrm{DD}} + \Delta V_{\mathrm{DD}}) - f(V_{\mathrm{DD}})}{f(V_{\mathrm{DD}})} \tag{11}$$

Here $f(x)$ is the oscillation frequency when supply voltage is x.

3.3 Ring VCO Based on Transmission Gate

The transmission gate is used with each inverting stage as delay element for the VCO. This achieves an operating frequency range of 4.8 Hz–256 MHz at 3.3 V supply voltage and 0.35 µm CMOS technology. In transmission gate-based ring VCO, each delay stage is composed of two transistors of inverter (M1; M2) and one transmission gate formed by (M3; M4) as shown in Fig. 5. Wide frequency range of ring VCO can be obtained by varying the resistance of transmission gate.

The resistance of transmission gate can be adjusted by the two voltages applied on the transistors gate whose sum should be fixed V_{DD}. The charging and

Fig. 5 Delay cell for
TG-based RVCO
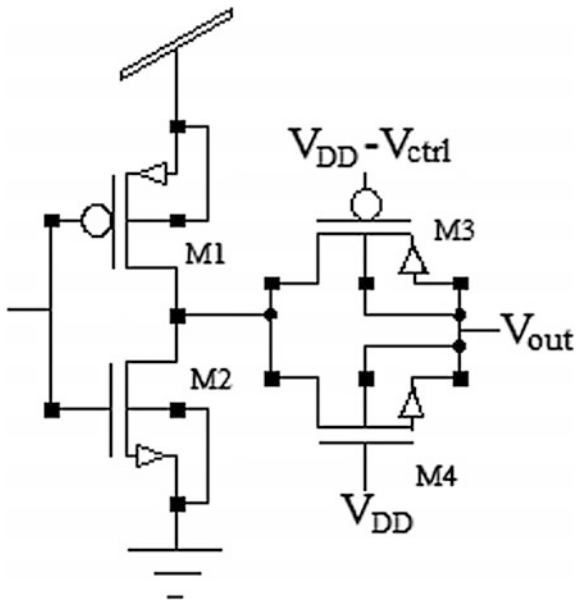
discharging path is symmetrical so that the difference of charging and discharging
time is independent of transmission gate [12, 13]. The ring VCO based on this delay
cell is power efficient due to reduction of short circuit current. The delay due to each
inverting stage is sum of the delay due to transmission gate and delay of inverter.
The delay of transmission gate is effectively determined by time to charge or
discharge the load capacitance (C_L) at its output through the parallel connection of
equivalent resistance (R_{eq}) of the two transistors. Thus, the propagation delay
($V_{out}(t_p) = V_{DD}/2$) is given by

$$t_p = \ln(2) \frac{2V_{DD}}{k_n(V_{DD} - V_{tn})^2 + k_p(V_{DD} - |V_{tp}|)^2} C_L \qquad (12)$$

$$\tau_{Total}(\text{delay}) = \tau_{TG}(\text{delay}) + \tau_{inv}(\text{delay}) \qquad (13)$$

Where V_{tn}, V_{tp} are threshold voltages for nMOS and pMOS transistors, respec-
tively, and k_n, k_p are the gain factors (proportional to the ratio of width and length of
the channel between source and drain (W/L) of the two transistors).

3.4 Ring VCO Based on NAND Logic Gate

The NAND logic gate-based VCO [14] topology exhibits frequency variation in the
range of 3290.9–4228 MHz with supply voltage variation from 1.8 to 2.4 V. In this
type of ring VCO, three-transistor (M1, M2 and M3)-based NAND gate is used as

Fig. 6 Delay cell for NAND
gate-based RVCO
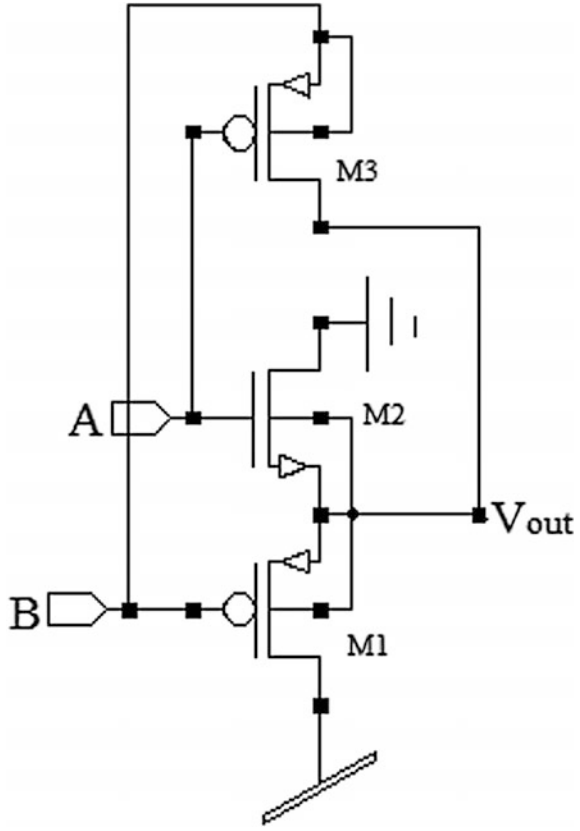
delay element and inverter operation has been performed by three-transistor NAND
gate as shown in Fig. 6. The circuits have been designed using 0.18 μm CMOS
technology. Out of the two input terminals of NAND gate, one terminal is con-
nected to logic high (i.e., 1.0 V) and the feedback signal is applied to the other
terminal. This circuit works as inverter without having direct path between V_{DD} and
ground, so the circuit is power efficient. The NAND gate delay stage in ring VCO is
based on pass transistor logic which has the ability to implement logic function with
a smaller capacitance value and less transistor count, and hence has better perfor-
mance but the drawback of this circuit is that the voltage difference of high and low
logic levels decreases at each stage. By the pass gate logic, the delay introduced by
the NMOS pass gate M2 is given by [15, 16]:

$$t_n = \frac{2C_{ox}}{k_n} \left[\left(\frac{1}{V_{DD} - V_x - V_{t,n}} \right) - \left(\frac{1}{V_{DD} - V_{t,n}} \right) \right] \tag{14}$$

where C_{ox} is the parasitic capacitance, $V_{t,n}$ is the threshold voltage of nMOS
transistor.

Table 1 Summary of parameters in different ring VCOs

Reference no.	No. of transistors used in three stage ring VCO	Power consumption (μW)	Supply voltage (V)	Maximum frequency (MHz)	Bandwidth (MHz)	Technology (μm)
[7]	14	260.000	1.8	875.00	809.00	0.18
[11]	21	65.9800	3.0	954.50	212.14	0.18
[12]	12	1760.00	3.3	256.00	255.99	0.35
[14]	09	486.180	2.4	4228.00	937.10	0.18

The performance parameters of different ring VCOs used in the articles are summarized in Table 1.

4 Conclusion

The performances of different ring VCOs have been presented in this paper, and the logic gate-based ring VCO shows the maximum frequency of oscillation and bandwidth than other ring VCO types. The transmission gate-based VCO offers better results when power requirement is minimum and frequency range is not an issue. Also, we found that the current starved ring VCO has highest power dissipation but provides very high frequency range of operation. Combined delay stage-based VCO shows the better frequency stability than current starved ring VCO. Therefore, ultimate applications to specific areas will be decided by the trade-off among the range of frequency of oscillation, power dissipation, bandwidth requirements, etc.

References

1. Stephen Docking and Manoj Sachdev, "A Method to Derive an Equation for the Oscillation Frequency of a Ring Oscillator" IEEE transactions on circuits and system-fundamental theory and applications, Vol. 50, No. 2, pp. 259–264, Feb. (2003).
2. Ahmet Tekin, Mehmet R. Yuce, and Wentai Liu, "Integrated VCOs for Medical Implant Transceivers" *Hindawi Publishing Corporation*, Vol. 2008, Article ID 912536, May (2008).
3. Jubayer Jalil, Mamun Bin Ibne Reaz, Mohammad Arif Sobhan Bhuiyan Labonnah Farzana Rahman, and Tae Gyu Chang, "Designing a Ring-VCO for RFID Transponders in 0.18 μm CMOS Process" *Hindawi Publishing Corporation,* Vol. 2014, Article ID 580385, Jan. (2014).
4. B. Razvi, Design of Analog CMOS Integrated Circuits, *Tata McGraw- Hill*, Third edition, (2001).
5. M. Tiebout, "Low Power VCO Design in CMOS", Springer Series in Advance Microelectronics, Volume 1–17, ISBN-13 978-3-540-24324-3, (2011).

6. Bhawika Kinger, Shruti Suman, K. G. Sharma and P. K. Ghosh "Design of Improved Performance Voltage Controlled Ring Oscillator" 5th IEEE International Conference on Advanced Computing &Communication Technologies, Rohtak, (India), *IEEE Digital library*, ISBN No. 978-1-4799-8487-9, pp. 441-445, Feb. (2015).

7. Gudlavalleti Rajahari, Yashu Anand Varshney, and Subash Chandra Bose, "A Novel Design Methodology for High Tuning Linearity and Wide Tuning Range Ring Voltage Controlled Oscillator", *Springer-Verlag Berlin Heidelberg* 2013, CCIS 382, pp. 10–18, (2013).

8. G. S. Jovanovic and M. K. Stojcev, "Current starved delay element with symmetric load" *International Journal of Electronics,* Taylor and francis, Vol. 93, No. 3, pp. 167–175, March (2006).

9. Cheng Zhang, Ming-Chen, Lin Marek Syrzycki "Process Variation Compensated Voltage Controlled Ring Oscillator with Subtraction-Based Voltage Controlled Current Source", *IEEE CCECE*, (2011).

10. Shruti Suman, Monika Bhardawaj Prof. B. P. Singh, "An Improved performance Ring Oscillator Design", in International Conference on Advance Computing and Communication Technology, Rohtak, India, *IEEE digital library,* pp. 236–239, Jan. (2012).

11. Abbas Ramazani, Sadegh Biabani, Gholamreza Hadidi, "CMOS Ring oscillator with combined delay stages", *Elsevier GmbH*, Dec. (2013).

12. Meng-Lieh Sheu, Ta-Wei Lin, Wei-Hung Hsu, "Wide Frequency Range Voltage Controlled Ring Oscillators based on Transmission Gates", on *IEEE symposium Circuit and System*, Vol. 323–326, pp. 2731–2734, May (2005).

13. Minh-Hai Nguyen, Cong-Kha Pham, *"*A Wide Frequency Range and Adjustable Duty Cycle CMOS Ring Voltage Controlled Oscillator", *Third International Conference on Communication and electronics*, *pp*. 107–109, Aug. (2010).

14. Manoj Kumar, "A Low Power Voltage Controlled Oscillator Design, "*Hindawi Publication Corporation, ISRN electronics*, Vol. 2013, April (2013).

15. J. M. Rabaey, "Digital integrated circuits: A design perspective," Prentice-Hall Book Company, 1st edition, ISBN No. - 0-13-178609-1, (1996).

16. Sung-Mo Kang, Yusuf Leblebici, CMOS Digital Integrated Circuit, *TATA McGraw-Hill*, third edition, (2003).

Modified Design of Integrated Ultra Low Power 8-Bit SAR ADC Architecture Proposed for Biomedical Engineering (Pacemaker)

Jubin Jain, Vijendra K. Maurya, Rabul Hussain Laskar and Rajeev Mathur

Abstract An energy efficient modified architecture of 8-bit 100 kS/s SAR ADC for the biomedical implant pacemaker is presented in this paper. With the stringent need to prolong the battery life of portable battery operated biomedical implants such as pacemaker, an improved architecture of SAR ADC is proposed which ensures improved performance than other reported SAR ADC architectures. The ADC employed in the pacemaker drains huge amount of power from battery during the time of analog to digital signal conversion. The work presents ADC design which ensures the microwatt operation which in turn makes the pacemaker to run on small battery. The ADC is realized in 180 nm CMOS technology operated at 1.8 V. The power consumption and energy efficiency reported during simulation are 2.5 µW and 0.77 pJ/state having precision of 6.68 bits.

1 Introduction

With the rising inclination of development in the field of biomedical electronics especially battery operated biomedical implants, which propose stringent constraints associated with the design amongst which the most important parameters are power consumption and area. In accordance to meet the low power consumption, an energy efficient breed of ADC is needed to be introduced. Therefore, the energy efficient data converters required for the acquisition of cardiac pulses play important role at the front and back end of the systems. The paper aims to design energy efficient ADC which comes from the low power design of building blocks of ADC which are

Jubin Jain (✉) · V.K. Maurya · Rajeev Mathur
Geetanjali Institute of Technical Studies, Udaipur, India
e-mail: jubincb2@gmail.com

R.H. Laskar
National Institute of Technology Silchar, Silchar, Assam, India

comparator, sample and hold circuit, digital to analog converter, and successive approximation register. The biomedical inspired ADCs follow the specification of (<8 bits) resolution and sampling rate (1–1000 kS/s). The proposed SAR ADC fits for the low power application in biomedical implants with respect to comparison between other ADC architectures such as flash ADC, pipelined ADC, and algorithmic ADC. The proposed ADC architecture features noise elimination in comparator, transconductance compensation amplifier employed in S/H circuit design with clock feed through cancellation, binary weighted charge scaling DAC design with MUX as switch for charge injection removal, and non-redundant SAR register design with such architectural changes the ADC reports improved performance.

2 Successive Approximation Register ADC Design

2.1 Principle of SAR ADC

The principle of SAR ADC is based on the concept of binary search algorithm; it passes through all the levels of quantization before converging to get the final digitized response. The SAR ADC comprises of comparator, sample and hold circuit, SAR logic, and DAC converter. The SAR ADC operates successively on an account to divide the voltage range by half value. The three basic operations performed by ADC; first is settling of DAC response, second is decision made by comparator and finally, the control logic determines the next level of DAC. Initially the MSB is set to logic '1' and digital equivalent obtained is compared with the unknown input voltage analog in nature. If the output obtained from the DAC is greater than analog input, the MSB tends to remain in ON state and second MSB tends to remain in previous state. The process is conducted repeatedly till it goes down to LSB.

The output obtained from the SAR is fed to the DAC whose response is considered as a variable reference for the comparator, while the other input connected to the comparator first passes through the S/H circuit then after being sampled it is applied to the comparator which is considered as second input of comparator. The output response is taken off from the comparator and is approximated with the unknown analog input voltage with the n-bit digital value of SAR. An 8-bit register is responsible for the control of timing during conversion. An analog input V_{IN} is compared with the output of the DAC. The direction of binary search is controlled by output of comparator and the SAR is responsible for the digital control and holds the output code obtained after the conversion being finished.

2.2 Comparator Using Hysteresis for Noise Compensation

Often it happens, comparator is operated in noisy environment and is needed to detect transitions at the threshold points. Depending upon the frequency of detected signal, if comparator is fast enough and amplitude of noise being large will result in

the presence of noise in output of comparator. Efforts in modification of transfer characteristics are desired which can be achieved by hysteresis.

In conventional design of comparator, characteristics are affected with the presence of noise. In order to eliminate the noise, concept of hysteresis is introduced. Circuit sensitivity towards the noise and multiple transitions is reduced at the output by using this technique. Hysteresis is considered as degree of quality of comparator in which input threshold changes with the change in input signal. When the input crosses the threshold, results change in output and subsequently threshold is reduced in turn the input must return to previous threshold and again the output changes. The intention behind this technique is that output must follow the low frequency biomedical signal. As a result, proposed comparator proves to be energy efficient and suitable for the design of ADC.

2.3 Digital to Analog Converter

The section deals with the design of R-2R ladder DAC realized in 250 nm CMOS technology operated at 3.3 V and the charge scaling DAC realized in 180 nm CMOS technology operated in 1.8 V. The comparative analysis is performed between both designs. The aim is to operate the DAC at low voltage with low power consumption which also comes from reduction in charge injection error using MUX as switch.

R-2R Ladder, when repeatedly the structures of the resistor are cascaded, the design obtained is known as binary weighted R-2R DAC which is considered the most common type of DAC with low cost, the only advantage is that it makes use of two different values of resistors. In ladder structure, input is fed through the N-bit inputs where the range of input is near 2–3 V. The configuration is designed in such a way that the R and 2R alternately appears. Voltage at each node is related to the V_{ref} and voltage division of the ladder network binary weighted relationship exists. The biggest disadvantage associated with the design is mismatching of resistor values which degrades the resolution and resistor fabrication difficulty.

Charge scaling DAC design consists of binary weighted linear capacitors which are arranged in parallel in order to achieve high resolution. The concept has the biggest advantage that the occupied total area is reduced by the use of capacitor which is essential in high resolution design. The proposed DAC topology reduces size and the power consumption. Initially the capacitor is discharged and the array of capacitor is switched to either V_{ref} or ground causing the output voltage (V_{out}) obtained due voltage division existing between the capacitors. The unit capacitor value is assumed to be 20 fF and the values of capacitor are multiple of 20 fF. The charge scaling DAC converter makes the conversion of 8-bit digital word to respective analog signal; scaled voltage reference is obtained from the capacitive network. The DAC consists of capacitor and MUX switches to which the digital word is applied. Initially digital word is applied to MUX circuit; the particular voltage to which the capacitor is to be charged is decided by the MUX which in

turn depends upon the logic value present in digital word. If the digital word contains logic '0' then the input is connected to the ground 'GND' and the capacitor voltage is '0'. If the digital word contains the input bit '1' then the capacitor is charged to V_{ref}.

2.4 Sample and Hold Circuit

The section deals with the design of sample and hold circuit realized in 130 nm CMOS process operated at supply of 3.3 V and proposed design of S/H circuit with clock feed through cancellation realized in 180 nm CMOS process operated at supply of 1.8 V.

Sample and hold circuit consists of capacitor and analog switch which is responsible to propose input isolation with the capacitor. An operational amplifier inserted in the circuit which is designed in voltage follower topology avoids the loading effect on the output capacitor. When the circuit is driven in standby mode, the amplifier circuit is powered down which causes the reduction in the power consumption.

The operation of S/H circuit is simple, as the clock signal clk is '1', the capacitor starts charging such that it samples the input signal and tracks the changes in input signal. When clkbar signal is activated (clk=0) the output is held at the voltage to which it was previously charged. The time interval during which the sampler tends to remain closed is known as sampling period. The rise in power consumption comes from the design of OPAMP, charge injection error, and clock feed through error in order to remove these errors, a new design of S/H circuit is proposed featuring low power consumption.

The proposed S/H circuit design is responsible to reduce clock feed through errors which means clock transitions are coupled to capacitor through MOS switch as a result error is produced at the output voltage. The design deals with the dummy switch NMOS_2 and a low power transconductance compensation unity gain OPAMP is employed in the circuit which buffers voltage across capacitor, prevents voltage leaking off from the capacitor, and provides low resistive path for the held voltage at output.

2.5 Successive Approximation Register (SAR)

The principle of SAR based on a binary search algorithm employed in the design of SAR arranged in a feedback loop includes 1-bit ADC. The architecture comprises of the front end sample and hold circuit, comparator, DAC, and SAR logic. Shift register can simply be considered as SAR logic with combination of decision logic and decision register. The change in the last bit of register is pointed by the pointer and all the results of comparison performed during the time of conversion are stored

in the register. The difference between the sampled signal (V_{IN}) and DAC output voltage (V_{DAC}) is made half during binary search. The MSB is set to '1' initially during conversion such that DAC produces the midscale at analog output. On an account to determine the polarity of $V_{IN} - V_{DAC}$, the comparator is strobed. The pointer and the decision logic direct the logical output of the comparator to the MSB. In case when $V_{IN} > V_{DAC}$ the register maintains the MSB at logic '1' or else sets it to '0'. Subsequently, the pointer is determined to choose MSB as '1'. After the DAC output being settled to its new value again the comparator is strobed and the above sequence is repeated.

The section deals with the design of proposed non-redundant SAR architecture, basically there exist two different structures namely sequence code register and non-redundant structure. Both the structures are composed of particular number of registers which is responsible to save and determine the value of the bits existing in the digital word which is produced by the SAR ADC. Since power consumption is prime concern, so the proposed architecture makes use of less number of registers than the sequence code register which is advantageous in case of reduction in size. In this architecture, initially at the beginning of conversion it is assumed that the MSB is set to logic level '1' and rest of all other bits are reset and then the digital word obtained is proposed to the D/A converter on an account to obtain the analog equivalent. Such that the value obtained is compared with the sampled input signal. In accordance with the output of the comparator the MSB bit is decided and then again the whole procedure is repeated. In order to reduce the number of clock cycles instead of loading value in shift register the logic '1' is shifted in the registers.

2.6　Figures and Tables

2.6.1　Design of Comparator

See Figs. 1 and 2.
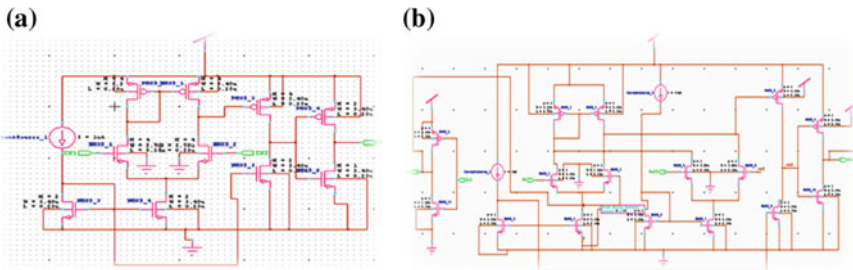
Fig. 1 **a** Schematic of comparator without hysteresis (250 nm). **b** Schematic of comparator using hysteresis (180 nm)
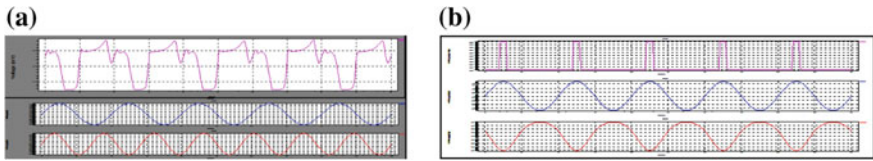
Fig. 2 a Output waveform of comparator without hysteresis. **b** Output waveform of comparator using hysteresis

2.6.2   Design of Digital to Analog Converter

See Figs. 3 and 4.

2.6.3   Design of Sample and Hold Circuit

See Figs. 5, 6 and 7.

2.6.4   Design of Non-redundant SAR Logic

See Figs. 8 and 9.

2.6.5   Schematic of SAR ADC

See Fig. 10.
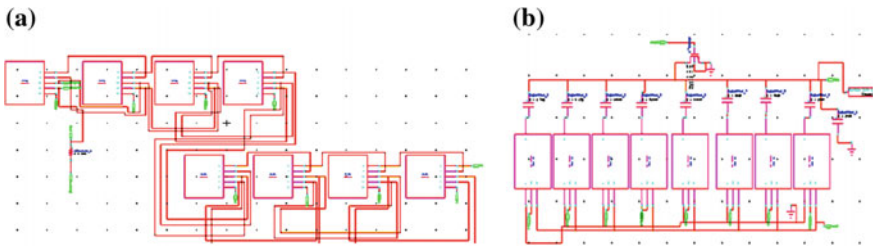
Fig. 3 a Schematic of R-2R ladder (250 nm). **b** Schematic of charge scaling (180 nm)
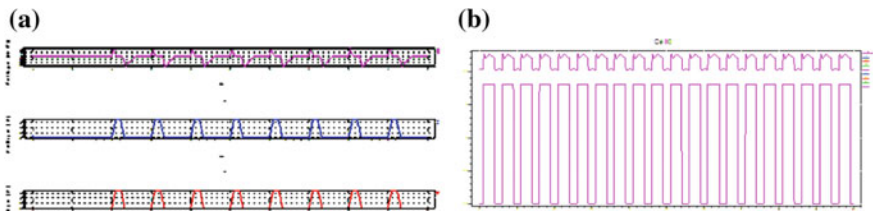
Fig. 4 a Output waveform of charge scaling. **b** Output waveform of R-2R ladder
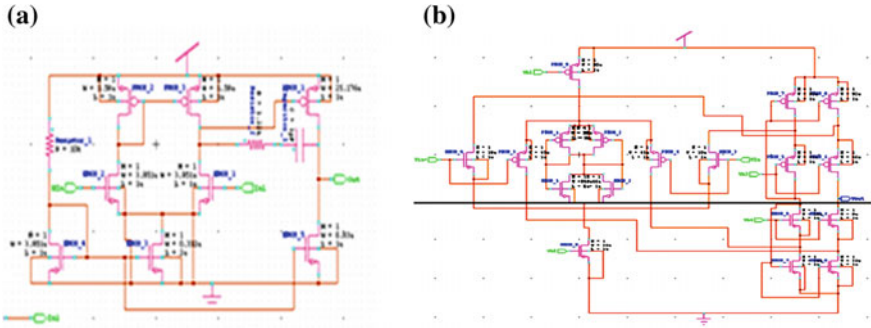
Fig. 5 a Schematic of OPAMP using miller compensation (130 nm). **b** Schematic of proposed low power OPAMP with input compensation for S/H circuit (180 nm)
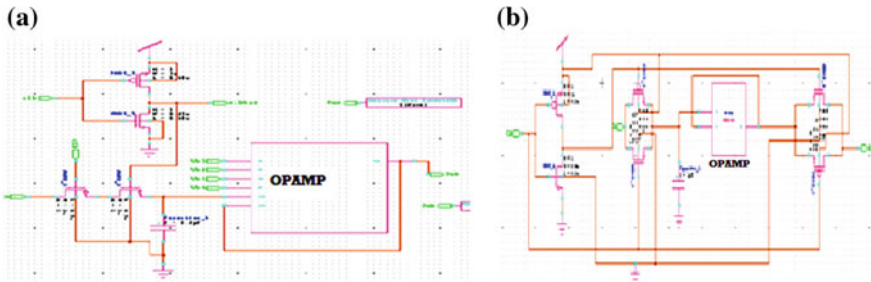
Fig. 6 a Schematic of proposed S/H Circuit with clock feed through cancellation. **b** Schematic of existing S/H circuit
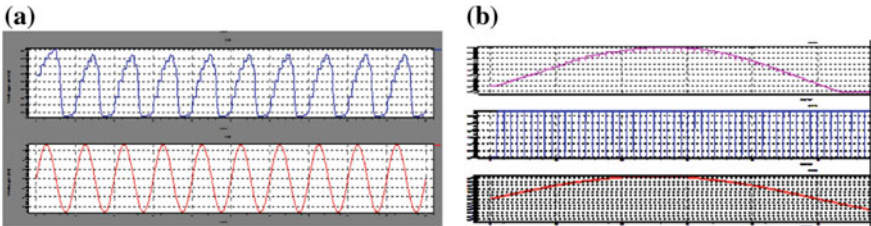
Fig. 7 a Output waveform of proposed S/H circuit. **b** Output waveform of existing S/H circuit
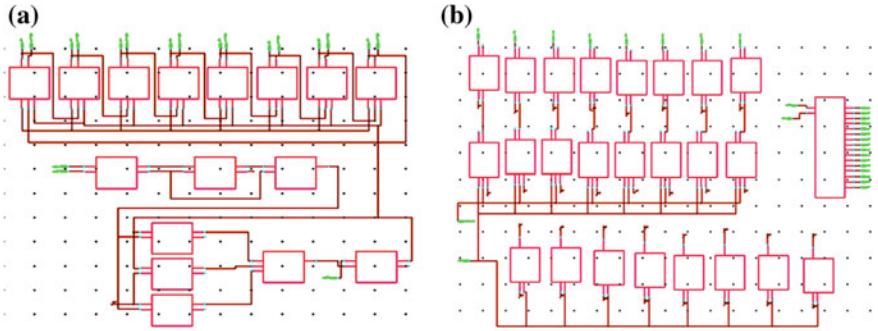
Fig. 8 **a** Schematic of shift register. **b** Schematic of SAR control logic

Fig. 9 Output waveform of SAR logic
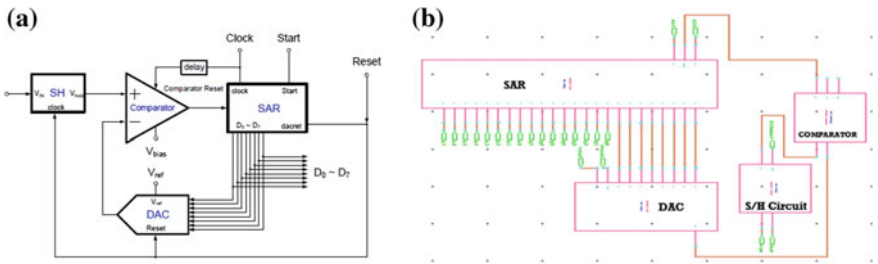
Fig. 10 **a** Block diagram of SAR ADC block. **b** Schematic of SAR ADC (180 nm)

2.6.6 Simulation Waveform of SAR ADC

See Fig. 11.

2.6.7 Performance Analysis Table for Proposed SAR ADC

S. No	Parameters	Units	Simulated results (proposed work)	Work [8]	Work [11]	Work [9]	Work [10]
1.	CMOS technology	m	0.18 μ	0.8 μ	0.18 μ	0.18 μ	65 n
1.	Supply voltage	V	1.8	2.8	2.5	0.6	1.2
2.	Resolution	bits	8	10	9	NA	6
3.	ENOB	bits	6.68	8.4	NA	NA	NA
4.	SNR	dB	36	NA	NA	NA	NA
5.	Sampling rate	S/s	100 k	2.9 k	NA	100 k	NA
6.	Energy per quantization	pJ/state	0.77	NA	NA	NA	NA
7.	Conversion time	μs	2.5	NA	55	NA	NA
8.	Power consumption	W	2.5 μ	8.81 μ	130 m	6.45 μ	2.1 m

2.7 Program Code

MNMOS_1 Out In1 Gnd Gnd NMOS W=1u L=180n AS=900f PS=3.8u AD=900f PD=3.8u

MNMOS_2 Out In2 Gnd Gnd NMOS W=1u L=180n AS=900f PS=3.8u AD=900f PD=3.8u

MPMOS_1 N_1 In1 Vdd Vdd PMOS W=4u L=180n AS=3.6p PS=9.8u AD=3.6p PD=9.8u
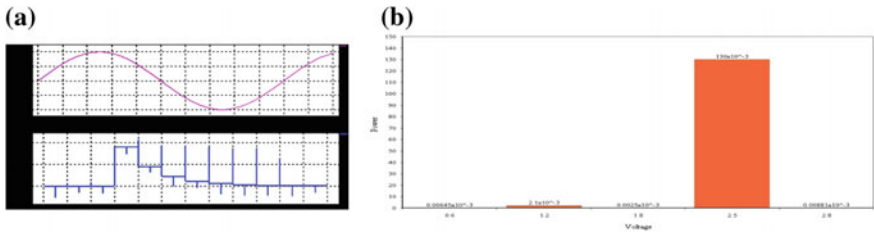
Fig. 11 **a** Output waveform of SAR ADC. **b** Graph of power consumption analysis for proposed ADC

MPMOS_2 Out In2 N_1 Vdd PMOS W=4u L=180n AS=3.6p PS=9.8u AD=3.6p PD=9.8u

MNMOS_1 N_5 B N_6 Gnd NMOS W=1u L=180n AS=900f PS=3.8u AD=900f PD=3.8u

MNMOS_2 N_7 sel Gnd Gnd NMOS W=1u L=180n AS=900f PS=3.8u AD=900f PD=3.8u

MNMOS_3 N_5 A N_7 Gnd NMOS W=1u L=180n AS=900f PS=3.8u AD=900f PD=3.8u

MNMOS_4 N_6 selbar Gnd Gnd NMOS W=1u L=180n AS=900f PS=3.8u AD=900f PD=3.8u

CCapacitor_1 N_18 Vout 2p

CCapacitor_2 N_22 Vout 1p

CCapacitor_3 Vout Gnd 2p

MNMOS_1 N_4 N_4 N_10 Gnd NMOS W=46u L=1u AS=41.4p PS=93.8u AD=41.4p PD=93.8u

MNMOS_2 N_4 N_4 N_10 Gnd NMOS W=46u L=1u AS=41.4p PS=93.8u AD=41.4p PD=93.8u

MNMOS_17 N_21 Vb5 Gnd Gnd NMOS W=9u L=1u AS=8.1p PS=19.8u AD=8.1p PD=19.8u

.subckt comparator IP In1 Out Out3 Out4 Gnd Vdd

MNMOS_3 N_12 N_5 3 3 NMOS W=5.4u L=180n M=2 AS=3.375p PS=6.65u AD=4.86p PD=12.6u

MNMOS_5 N_2 Out3 N_25 Gnd NMOS W=3.6u L=180n M=2 AS=2.25p PS=4.85u AD=3.24p PD=9u

MNMOS_7 N_25 N_7 3 3 NMOS W=5.4u L=180n M=2 AS=3.375p PS=6.65u AD=4.86p PD=12.6u

MNMOS_8 N_7 N_7 3 3 NMOS W=5.4u L=180n M=2 AS=3.375p PS=6.65u AD=4.86p PD=12.6u

MNMOS_9 vo2 N_12 3 3 NMOS W=5.4u L=180n M=2 AS=3.375p PS=6.65u AD=4.86p PD=12.6u

MPMOS_2 N_2 N_2 Vdd Vdd PMOS W=5.4u L=180n M=4 AS=3.375p PS=6.65u AD=4.1175p PD=9.625u

CCapacitor_1 N_2 N_1 1.28p

CCapacitor_2 Out N_3 2.56p

CCapacitor_3 N_2 N_4 640f

CCapacitor_4 Out N_5 320f

CCapacitor_5 Out N_6 160f

CCapacitor_6 Out Gnd 20f

CCapacitor_7 Out N_7 80f

CCapacitor_8 Out N_8 40f

CCapacitor_9 Out N_9 20f

XCell0_1 N_4 Vb1 Vb2 Vb3 Vb4 N_5 N_1 N_3 Out Gnd Vdd thesisnew opamp_Cell0_view0

.tran 1n 100n

.include "C:\Users\Jubin\Documents\Tanner EDA\Tanner Tools v14.1\L-Edit and LVS\LVS\SPR_Core\hp05.md"
Vdd Vdd Gnd 1.8
Vstr str Gnd 1.8
Vref Vref Gnd 1.1
VVIN VIN Gnd dc 1 SIN (0 1 0.0001G 0)
Vclk clk Gnd PULSE (0 1.8 1n 20p 20p 0.2n 1n)
.PRINT TRAN v(VIN)v(q0) v(q1)v(q2) v(q3) V(q4) v(q5) v(q6) v(q7) v(dacout)
.end

3 Conclusion

The work presents the biologically inspired ADC, the proposed and existing schematics of the circuit have been designed using TANNER EDA tool implemented in 0.18 μ CMOS process with a supply voltage of 1.8 V. With respect to the current scenario, pacemaker operates at the battery voltage of 3–5 V as a result lifetime of pacemaker is needed to be improved because surgical operations cannot be performed within short span of time. The experimental results obtained after simulation have been mentioned, the proposed ADC is characterized by ultra low power consumption of 2.5 μW with the precision of 6.68 bits. In accordance with comparative analysis made with other recent research works, the proposed work proves to be well suited for pacemaker.

References

1. Shouli Yan and Edgar Sanchez-Sinencio: Low Voltage Analog Circuit Design Techniques: A Tutorial, IEICE Transactions on Analog Integrated Circuits and Systems, vol. E00-A, no. 2, February (2000).
2. Philip E Allen & Douglas R Holberg: CMOS Analog Circuit Design, ISBN 0-19-511644-5, Oxford University Press, 2nd Edition, 2002.
3. Jens Sauerbrey, Doris Schmitt-Landsiedel, Roland Thewes: A 0.5 V, 1 μW Successive Approximation ADC, IEEE Journal of Solid State Circuit (2002).
4. Sandro A.P Haddad, Richard Houben and Wouter A. Serjdin: The evolution of pacemaker: an electronics perspective, from hand crank to advanced wavelet analysis, DISens symposium-book, 2005.
5. Kuttner, A. S. T. H. F., Sandner, C., & Clara, M.: A 6 bit, 1.2 gsps low-power flash-adc in 0.13 μm digital cmos, IEEE Journal of Solid-State Circuits Vol. 40, No. 7., 2005.
6. Deguchi, K., Suwa, N., Ito, M., Kumamoto, T., & Miki, T.: A 6-bit 3.5-GS/s 0.9-V 98-mW flash ADC in 90-nm CMOS. IEEE Journal of Solid-State Circuits, Vol. 43, No. 10, pp 2303–2310., 2008.
7. Chen, Fred, Anantha P. Chandrakasan, and Vladimir Stojanovic: A Low-power Area-efficient Switching Scheme for Charge sharing DACs in SAR ADCs, IEEE (2010).

8. G.Bonfini, C.Garbossa, R.Saletti, "A Switched OPAMP-based 10-bit Integrated ADC for Ultra Low-power Applications", Via Giuntini 13, I-56023 Navacchio (Pi), Italy (2011).
9. RVNR Suneel Krishna, Aleti Shankar: Design of Low Power SAR-ADC in 0.18 μm, Mixed-Mode CMOS Process (2013).
10. Gulrej Ahmed and Rajendra Kumar Baghel: Design of 6-bit flash ADC converter using variable switching voltage CMOS comparator, VLSICS, Vol. 5, No. 3 (2014).
11. Akhil A, Sunil Jacob: Design of 9 bit SAR ADC using high speed and high resolution open loop CMOS comparator in 180 nm technology with R-2R DAC topology, International Journal of VLSI and Embedded Systems-IJVES, Vol. 05, Article 11492 (2014).

A Novel Symmetric Key Cryptography Using Dynamic Matrix Approach

Neetu Yadav, R.K. Kapoor and M.A. Rizvi

Abstract One of the most challenging aspects in today's information and communication technology is data security. Encryption is one of the processes to secure the data before it is communicated. It is a process of altering an intelligible form of data into an unintelligible form based on encryption algorithm using a key. To get the original data back, a decryption process is used. Encryption algorithms are of two types: symmetric key encryption algorithm which uses same key and asymmetric key encryption algorithm which uses different keys for encrypting and decrypting the data. In this paper, an algorithm has been proposed based on several factors which are dynamic square matrix creations depending on the length of information, performing ASCII conversions, applying XNOR operation and then transposing the matrix to enhance the security and efficiency of data. The performance of proposed algorithm is compared with existing algorithm and found to be superior on various parameters.

Keywords Dynamic square matrix · ASCII value · Encryption · Decryption · XNOR operation · Transpose

1 Introduction

In today's world where most of the personal information and financial transactions are carried over the Internet, encryption of data is a requisite element to any effective computer security system [1]. Sending data through Internet has higher chances of getting hacked as the hackers are becoming efficient in their job that they

Neetu Yadav (✉) · R.K. Kapoor · M.A. Rizvi
National Institute of Technical Teachers' Training and Research, Bhopal, MP, India
e-mail: neetuy596@gmail.com

R.K. Kapoor
e-mail: rkkapoor@nitttrbpl.ac.in

M.A. Rizvi
e-mail: marizvi@nitttrbpl.ac.in

can easily hack the unencrypted data over the internet. If the hacked data contain sensitive information it may be misused by the hackers. This requires the security of the data before communicating over the network. So with the increased use of Internet the major concern is the security of data. Many organizations, whether they are large, small, or government organizations are affected by the network security. An intruder can do all sorts of harm if the security is broken [2]. Incorporating a strong security mechanism can protect the data from any alterations. Encryption is one of the effective ways to secure the data. It transforms an intelligible form of data into unintelligible form that is unreadable to the others using an encryption algorithm and key. Encryption occurs when some replacement technique, shifting technique and mathematical operations are applied on the data [3]. The main purpose of every encryption algorithm is to make the decryption of generated cipher text as difficult as possible without using the key. Depending on the use of keys encryption algorithms are categorized into two categories: symmetric key encryption algorithm and asymmetric key encryption algorithms. If a single key is used for encryption as well as for decryption it is called symmetric key encryption algorithm. If different keys are used at sending and receiving end it is called asymmetric key encryption algorithm.

2 Literature Survey

Many attempts have been done by researchers to develop a more secure encryption algorithm in spite of existing algorithms. Some of these research studies are presented in this section.

Authors in a study [4] have proposed an algorithm in which variable length key is used whose length depends on the message length and data are converted into ASCII values. Then encrypted data are operated with XOR operation with variable length key generated randomly.

Srikantaswamy et al. [5] proposed an algorithm that generates key of any length using seed value to encrypt the data. It improves one time pad by generating different key values using a seed value. But if a seed value is known the entire key values can be predicted. In this study, the system can be made more secure if the random keys are generated depending on the order of the matrix created dynamically so that even if the key length is known then it will not be possible to guess the value of the key.

Mahmood et al. [6] proposed an algorithm which uses linear congruential generator (LCG) for generating key. As per this algorithm, it is impossible to detect the pattern which enhances the security. But to generate the dynamic keys the random function requires an initial value. If this value is known then all the keys can be predicted. The random keys generated based on the order of the matrix created dynamically can further increase the security.

In a study, Gitanjali et al. [7] proposed an encryption technique based on ASCII values and these values are converted into cipher text using the palindrome number.

The matrix multiplication is used to secure the data. The system can be further improved if the plain text matrix is created dynamically and keys are generated randomly.

A new 32 bit encryption and decryption algorithm which applies 1's and 2's complements and XOR operation to convert the plain text into cipher text was proposed by Kumar et al. [8]. In this algorithm fixed length keys are used for encryption and decryption. In this study, the security can further be improved if the keys are generated randomly.

Authors [9] proposed an algorithm based on ASCII conversions and a simple cyclic mathematical function for encrypting the data. The output of the cyclic function rotates between 0 and 31. In this algorithm, it may be possible to detect the pattern of the cipher text. The security can be improved further if random keys are used to encrypt the text and logical XNOR operation is used in order to make the pattern of the cipher text unpredictable.

Authors in study [10] proposed an algorithm in which logical OR operation is used to generate the key and encrypt the text using XOR operation. Different keys used for each round make the data more secure. This algorithm has limitation that the size of matrices used for generating key should be same whereas the proposed algorithm generates key of the order of the matrix created dynamically.

3 Methodology

The proposed algorithm is based on dynamically creating matrix for plain text based on its length and generating a variable length key depending on the order of the matrix generated. Then incorporating XNOR operation on each row element with key and converting the values into decimal. Incorporate an addition operation with generated encoding matrix and transpose the resultant matrix. Sending the resultant matrix to the receiving end to decrypt the data, both the key and the encoding matrix are required. Hence, the proposed method will improve the security of data and gives better results as compared to existing methods.

3.1 Encryption

Dynamically plain text is converted into square matrix based on its length and ASCII values are taken. A binary matrix of the same order is created by converting ASCII values into 8-bit binary. Then apply 1's complement on elements and perform XNOR operation with 8-bit binary form of randomly generated key whose length is equal to the order of the matrix. Convert binary matrix into decimal matrix and perform addition operation with encoding matrix of the same order. Transpose the matrix then final cipher text matrix is obtained.

3.1.1 Encryption Algorithm

1. Input the plaintext message.
2. Dynamically convert the plain text message into square matrix [A]of maximum possible size of order m×n where m=n
3. Convert all elements of matrix [A] into their ASCII values.
4. Create matrix [B] of same order by converting the ASCII values of elements of matrix [A].
      (i)Convert ASCIIs of matrix [A] elements into binary form.
      (ii)Apply 1's complement on binary values of each row element of the matrix.
5. Repeat step4 for all the remaining rows of the matrix [B].
6. Create matrix [I] of the obtained values of same order.
7. Generate any random number key K1 whose length is equal to the order of the matrix [I]. Add its elements and convert the sum obtained into binary form.
8. Perform XNOR operation of each row element of matrix [I] with the binary sum of elements of key K1.
9. Create a decimal matrix [M] of same order m×n of values obtained in step8 by converting binary values into decimals.
10. Generate an encoding matrix [E] of same order m×n from the matrix [M] as-
          Initialize i=0,j=0;
          For(i=0 to m-1)
            Increment i by 1
            NewRow[i]=0;
                For(j=0 to n-1)
                 NewRow[i]= NewRow[i]+i*j
               End inner loop
          End outer loop
11. Add the elements of matrix [M] with the elements of encoding matrix[E] which will generate the resultant matrix[P] after addition as [M]+[E]=[P]
12. Take transpose of matrix [P] obtained in above step11.
13. Send the cipher text matrix[C] obtained above to the receiving end.

3.2 Decryption

In decryption phase, the transpose of the received cipher text matrix [C] of order $m \times n$ is taken to obtain matrix [P]. On subtracting the encoding matrix [E] from matrix [P], the matrix [M] is obtained. The decimal values of elements of matrix [M] are converted into binary values to obtain matrix [R]. On performing XNOR operation of each row element of matrix [R] with key K1 and applying 1's complement operation on each element, the resultant matrix ASCII values are converted into character to obtain the plain text matrix [A].

3.2.1 Decryption Algorithm

1. Input the cipher text matrix [C] of order m × n.
2. Transpose the Cipher text matrix [C] that will generate matrix [P].
3. Subtract the row elements of encoding matrix [E] from the row elements of matrix [P], [P]-[E] = [M].
4. Create binary matrix [R] of order m × n of the converted decimal values of each element of matrix [M] into their 8-bit binary equivalent.

    (i) Perform XNOR operation of each row element of matrix [R] with the binary sum of elements of random number key K1.
    (ii) Apply 1's complement on each row element of the matrix [R].
    (iii) Convert the obtained values into ASCIIs.
    (iv) Convert the ASCIIs into their equivalent character.

5. Repeat Step 4 for all the remaining rows of the matrix [R] to obtain the final plain text matrix [A].

4 Case Study

4.1 Encryption Case

1. Input plain text message "Cryp2Hy ischniQ$ecurD@t@".
2. Dynamically create square matrix [A] of order 5 × 5 so that it can accommodate the plain text having length 24 as shown in Table 1.
3. Convert all the elements of matrix [A] into their equivalent ASCII as shown in Table 2 where in Table 1 " " denotes space whose ASCII is 32 and take ASCII for null values as 0.

Table 1 Plain text matrix [A] of order 5 × 5

C	r	y	p	2
H	y		i	s
c	h	n	i	Q
$	e	c	u	r
D	@	t	@	

Table 2 Matrix [A] ASCII values representation

67	114	121	112	50
72	121	32	105	115
99	104	110	105	81
36	101	99	117	114
68	64	116	64	0

4. Create binary matrix [B] by converting the values obtained above in Table 2 into binary form as shown in Table 3.
5. Take 1's complement of the each row element of the above binary matrix [B] (Table 4).
6. Generate any random number key K having length equal to the matrix order such that 1 ≤ elements of key ≤ 9. Let the generated key K be [9 8 4 9 8] whose elements sum is 38. Perform XNOR of each element of matrix [I] with resultant key K1 = 00100110 (binary of sum = 38), the matrix obtained is shown in Table 5.
7. Obtain a decimal matrix [M] of above calculated binary values (Table 6).
8. Generate an encoding matrix [E] of same order 5 × 5 from matrix [M] (Table 7).
9. Add each row element of matrix [M] with encoding matrix [E], matrix [P] is obtained (Table 8).

Table 3 Binary matrix [B] of order 5 × 5

01000011	01110010	01111001	01110000	00110010
01001000	01111001	00100000	01101001	01110011
01100011	01101000	01101110	01101001	01010001
00100100	01100101	01100011	01110101	01110010
01000100	01000000	01110100	01000000	00000000

Table 4 Binary matrix [I] of order 5 × 5 obtained after 1's complement

10111100	10001101	10000110	10001111	11001101
10110111	10000110	11011111	10010110	10001100
10011100	10010111	10010001	10010110	10101110
11011011	10011010	10011100	10001010	10001101
10111011	10111111	10001011	10111111	11111111

Table 5 Matrix [I] of order 5 × 5 after applying XNOR operation on each element with key K1

01100101	01010100	01011111	01010110	00010100
01101110	01011111	00000110	01001111	01010101
01000101	01001110	01001000	01001111	01110111
00000010	01000011	01000101	01010011	01010100
01100010	01100110	01010010	01100110	00100110

Table 6 Decimal matrix [M] of order 5 × 5

101	84	95	86	20
110	95	6	79	85
69	78	72	79	119
2	67	69	83	84
98	102	82	102	38

Table 7 Encoding matrix [E] of order 5 × 5

102	86	98	90	25
112	99	12	87	95
72	84	81	91	134
6	75	81	99	104
103	112	97	122	63

Table 8 Resultant matrix [P] of order 5 × 5

203	170	193	176	45
222	194	18	166	180
141	162	153	170	253
8	142	150	182	188
201	214	179	224	101

Table 9 Cipher text matrix [C] after transpose

203	222	141	8	201
170	194	162	142	214
193	18	153	150	179
176	166	170	182	224
45	180	253	188	101

10. Take the transpose of the above obtained matrix [P], the final cipher text matrix [C] of order 5 × 5 is obtained as shown in Table 9.
11. Send the above cipher text matrix [C] at the receiving end.

4.2 Decryption Case

1. Take the transpose of the cipher text matrix [C], the resultant matrix [P] will be obtained as shown in Table 10.
2. Subtract encoding matrix [E] from the above matrix [P] that is [P]-[E] = [M], the decimal matrix [M] is obtained as shown in Table 11.
3. Convert each row of matrix [M] into 8-bit binary equivalent to obtain the binary matrix [I] as shown in Table 12.
4. Perform XNOR of each row element of above matrix with the key K1 = 00100110, the binary matrix [R] is obtained as shown in Table 13.

Table 10 Resultant matrix [P] of order 5 × 5

203	170	193	176	45
222	194	18	166	180
141	162	153	170	253
8	142	150	182	188
201	214	179	224	101

Table 11 Decimal matrix [M] of order 5 × 5

101	84	95	86	20
110	95	6	79	85
69	78	72	79	119
2	67	69	83	84
98	102	82	102	38

Table 12 Binary matrix [I] of order 5 × 5

01100101	01010100	01011111	01010110	00010100
01101110	01011111	00000110	01001111	01010101
01000101	01001110	01001000	01001111	01110111
00000010	01000011	01000101	01010011	01010100
01100010	01100110	01010010	01100110	00100110

Table 13 Binary matrix [R] of order 5 × 5 after applying XNOR operation

10111100	10001101	10000110	10001111	11001101
10110111	10000110	11011111	10010110	10001100
10011100	10010111	10010001	10010110	10101110
11011011	10011010	10011100	10001010	10001101
10111011	10111111	10001011	10111111	11111111

5. Now take the 1's complement of each row element of above matrix [R] as shown in Table 14.
6. Convert binary values of each row element of binary matrix into their ASCII equivalent as shown in Table 15.
7. Convert the ASCII into the equivalent characters, the plain text matrix [A] is obtained as shown in Table 16.
8. Hence from the above matrix, we obtain the plain text message "Cryp2Hy ischniQ$ecurD@t@".

Table 14 Binary matrix [R] of order 5 × 5 after taking 1's complement

01000011	01110010	01111001	01110000	00110010
01001000	01111001	00100000	01101001	01110011
01100011	01101000	01101110	01101001	01010001
00100100	01100101	01100011	01110101	01110010
01000100	01000000	01110100	01000000	00000000

Table 15 ASCII values of binary matrix [R]

67	114	121	112	50
72	121	32	105	115
99	104	110	105	81
36	101	99	117	114
68	64	116	64	0

Table 16 Plain text matrix [A] of order 5 × 5

C	r	y	p	2
H	y		i	s
c	h	n	i	Q
$	e	c	u	r
D	@	t	@	

5 Results

The proposed method encrypts and decrypts the data and works satisfactorily. It enhances the security of data against cryptanalysis attack because of the generation of variable length key based on matrix order. The unpredictability of key length until the order of the generated matrix is known. Each time different cipher text is produced for the same text data so that it is not possible to guess the relationship between cipher text and plain text for cryptanalysis hence the proposed method provides better security for protecting the data.

6 Conclusion and Future Scope

The proposed method solves the various aspects of data security. The plaintext matrix created dynamically makes the job of attacker difficult. The use of random key and the encoding matrix makes the algorithm strong and the operation applied creates randomness in the text. The algorithm can work on message of any length. In future we can use it for other data types.

References

1. Why Data Encryption is Important—Streetdirectory.com, http://www.streetdirectory.com/travel_guide/125729/security/why_data_encryption_is_important.html.
2. Computer Network Attacks—A Study—Academia.edu, http://www.academia.edu/9128513/Computer_Network_Attacks_-_A_Study.
3. Guide to Industrial Control Systems (ICS) Security—CiteSeer, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.224.673&rep=rep1&type=pdf.
4. Charru, Paramjeet Singh, Shaveta Rani "Efficient Text Data Encryption System to Optimize Execution Time and Data Security" International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), Volume 4, Issue 7, July 2014, ISSN: 2277 128X.
5. S.G. Srikantaswamy, Dr. H.D. Phaneendra "Enhanced One Time Pad Cipher with More Arithmetic and Logical Operations with Flexible Key Generation Algorithm", International Journal of Network Security & Its Applications (IJNSA), Vol. 3, No. 6, November 2011.

6. Zeenat Mahmood, J. L Rana, Prof. Ashish Khare," Symmetric Key Cryptography using Dynamic Key and Linear Congruential Generator (LCG)", International Journal of Computer Applications (0975-8887) Volume 50, No. 19, July 2012.
7. J Gitanjali, N Jeyanthi, C Ranichandra "ASCII based cryptography using unique id, matrix multiplication and palindrome number." Networks, Computers and Communications, The 2014 International Symposium on. IEEE, 2014.
8. Niraj Kumar, Pankaj Gupta, Monika Sahu, Dr. M A Rizvi, "Boolean Algebra based Effective and Efficient Asymmetric Key Cryptography Algorithm: BAC Algorithm", Automation, Computing, Communication, Control and Compressed Sensing (iMac4 s), 2013 International Multi-Conference, 2013.
9. Md. Palash Uddin, Md. Abu Marjan, Nahid Binte Sadia, Md. Rashedul Islam, "Developing a Cryptographic Algorithm Based on ASCII Conversions and a Cyclic Mathematical Function", Third International conference on informatics, electronics & vision, IEEE, 2014.
10. Dr. S. Kiran, R. Pradeep Kumar Reddy, I. Raja Sekhar Reddy," Multiple Text Cryptanalysis By Using Matrix And Logical Operations", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 5, September-October 2014, ISSN 2278-6856.

Experimenting Large Prime Numbers Generation in MPI Cluster

Nilesh Maltare and Chetan Chudasama

Abstract Generating large prime number is a time consuming problem. It is useful in key generation in network security. The problem of generating prime number is easy to parallelize. Use of high performance computing (HPC) facilities can be applied for getting faster results. Embarrassing parallel pattern is applied to categories of algorithms where concurrency is explicit. We have conducted experiments on sequential and parallel approaches of prime numbers generation. In this paper, our objective is to analyze time taken in generating large prime numbers using multicore cluster. The result could be reusable in estimating time required for key generation in cryptography.

Keywords Prime number generation · Multicore cluster · MPI · Parallel Design Patterns · HPC

1 Introduction

Prime number [1] is a natural number that can be divided by exactly two distinct natural number divisors: 1 and itself. The property of number is a prime number or composite number called primality. Primality of a number is tested with trial division. This simple method of verifying the primality of a given number n is easy to implement with computers but performs very slow. It consists of testing whether n is a multiple of any integer between 2 and $\sqrt{n}$. Apart from that, there are various methods [2] for generating prime numbers: sieve of Atkin, sieve of Eratosthenes, sieve of Sundaram, and wheel factorization. Algorithms much more efficient than trial division have been devised to test the primality of large numbers. However,

Nilesh Maltare (✉) · Chetan Chudasama
MBICT, New V. V. Nagar, Anand 388120, Gujarat, India
e-mail: nmaltare@mbict.ac.in
URL: http://www.nileshmaltare.co.in

Chetan Chudasama
e-mail: cpchudasama@mbict.ac.in

fast methods are available only for some cases of prime numbers of special forms, such as Mersenne numbers and they are difficult to parallelize.

The prime numbers are very important in cryptography [3]. Many popular algorithms used in public-key cryptography are based on the fact that integer factorization [4] is a "hard" problem. This means that the time required to factorize integers into their prime factors increases exponentially with the size of the integer (i.e., the numbers of bits needed to encode it). So if the encryption uses very large integers, it would take a large amount of time to "crack" it. There are other practical uses of prime numbers. Most of them are related to the fact that prime factorization is unique.

In this paper, we are analyzing time required to generate large prime numbers on multicore clusters. We are considering sieve of Eratosthenes for generating prime. All the experiments are performed on *PARAM Yuva II Cluster*.

2 Parallel Design Patterns Used

Recent trends in hardware design [5, 6] are toward multicore CPUs with hundreds of cores. It demands for better programs which can exploit multicore. In other words, sequential programs need to be refactored for parallelism [7]. Parallel software that fully exploits the hardware is difficult to write. A huge increase occurs in complexity and work for programmer. Humans are sequential beings, deconstructing problems into parallel tasks is hard for many of us. Parallelism is not easy to implement. Most parallel programming environments focus on the implementation of concurrency rather than high-level design issues. A better design leads to reliable, efficient, and more maintainable program. A design pattern describes a high-quality solution to a frequently occurring problem in some domain. Pattern for parallelism has been explored in [8–11]. Parallel design patterns provide highly parallel implementation with flexibility to adopt different platforms.

2.1 *Embarrassingly Parallel*

An *embarrassingly parallel* algorithm contains obvious concurrency which can be exploited easily. Once these independent tasks have been defined, it is easy to distribute them to different processor. Nevertheless, while the source of the concurrency is often obvious, taking advantage of it in a way that makes for efficient execution can be difficult. The embarrassingly parallel pattern shows how to organize such a collection of tasks so that they can execute efficiently. The challenge is to organize the computation so that all units of execution finish their work at about the same time. If they finish at the same time, then the computational load is balanced among processors. The problem is decomposed into a set of independent tasks. Most algorithms based on task queues and random sampling are

instances of this pattern. Our problem in prime number generation is embarrassingly parallel and can be implemented with same. We have devised a solution which adopts hardware environment and available cores to divide task dynamically.

3 Hardware Used for Implementation

All the programs are implemented in PARAM Yuva II. PARAM Yuva II [12] is a high performance computing cluster that is among the latest addition to the series of prestigious PARAM series of supercomputers built in India. PARAM Yuva II constitutes Intel Xeon processors, RCS-based accelerator cards, Intel Xeon Phi coprocessors, NVIDIA GPUs, and AMD SMP server. This provides a heterogeneous platform for users to explore a wide range of scientific applications.

PARAM Yuva II has four subclusters:

1. The first subcluster is a 218 node cluster of Intel server system R2000GZ with two Intel Xeon E5-2670 (Sandy Bridge) processor (eight cores each), with FDR Infiniband interconnect. Each of these nodes has two Intel Xeon Phi 5110P coprocessors, each with 60 cores and 8 GB RAM to boost the computing power.
2. The second one is a 100+ node cluster with HP Proliant DL580 G5 nodes, with each of the nodes having four Intel Xeon X7350 (Tigerton) processor (four cores each), with PARAMNet3 as well as DDR Infiniband interconnects.
3. The third subcluster is a 4 node cluster of Supermicro SuperServer 1027GR-TRF with two Intel Xeon E5-2650 (Sandy Bridge) processor (eight cores each), with FDR infiniband interconnect. Each of these nodes has two NVIDIA GPU Tesla M2090 cards to accelerate the performance.
4. The fourth one is a Supermicro 4U AMD SR5690 server. This system consists of four AMD Opteron 6276 processors (sixteen cores each). This is connected with the rest of the cluster by Gigabit Ethernet connection as well as FDR Infiniband interconnect.

4 Results

All experiments are done on PARAM Yuva II HPC. First prime number generation C Program generates first 25,000,000 prime numbers executed on sequential manner and it takes 36.73 s execution time (refer Table 1). For two cores it takes 18.62 s execution time. And so on, we observe that increase in number cores will result in better execution time (faster) of prime number generation program. At some point of time performance, gain tends to become constant. This can be justifying if we consider latency in distribution and collection of results from multiple cores. The results encourage to work in establishing highest speedup achieved by

Table 1 Time required to generate first 25,000,000 prime numbers

Number of cores	Time required to compute in seconds
1	36.73
2	18.62
4	4.5
8	2.3
16	1.2
32	0.4
64	0.3
128	0.3

parallelizing application in any number of cores. In the second experiment, we have calculated time required to generate largest n digit prime number and total number of prime found (refer Table 2). We have performed experiment with 5, 6, 7, and 8 digits. The time required increases rapidly with increase in number of digits. These results will also indicate time required for factorization of *n* digit number (Figs. 1 and 2).

Table 2 Time required for n digit prime number

N digit prime	5	6	7	8
Core	Time required to compute in seconds			
Sequential (gcc compiler is used)	0.046	1.94	50.73	1362.82
2	0.024738	0.596771	15.915941	427.063786
4	0.005743	0.28931	7.773494	212.445226
8	0.003014	0.14869	1.91788	105.704038
16	0.001574	0.064856	0.919607	34.062
32	0.001204	0.04228	0.925395	29.485275
64	0.003484	0.023725	0.9034722	24.749449

Fig. 1 This figure is showing time required to generate first 25,000,000 prime numbers in seconds on number of cores shown in horizontal axis
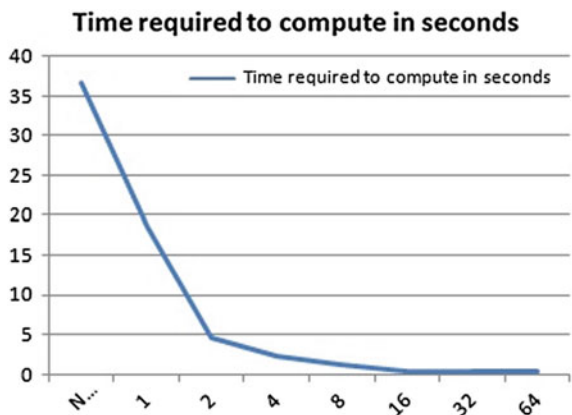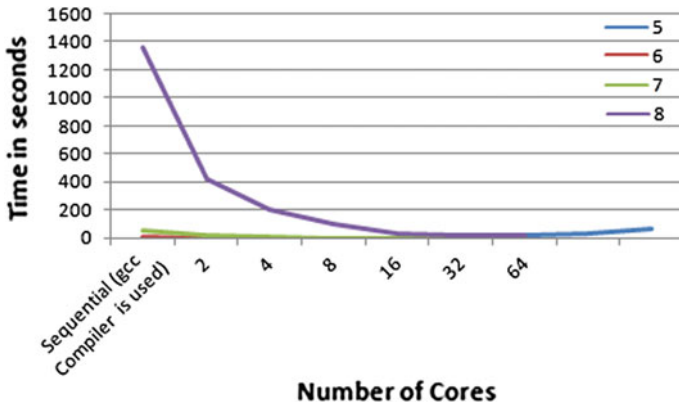
Fig. 2 This figure is showing time required to generate n digit prime numbers. We have calculated time required for 5, 6, 7, and 8 digits

5 Conclusion and Future Work

The prime numbers are important in cryptography for encrypting information. Most algorithms [13] use prime numbers for encryption and their secrecy depends upon time required to generate prime or factorization. A supercomputer can be used in cracking of information as the problem is embarrassingly parallelizable. Although, we have observed methods which are not efficient are easy to parallelize, while more efficient methods are difficult to parallelize. We have used HPC facility to generate large prime numbers.

Main objective is to demonstrate time taken by prime generation with increasing number of cores in multicore platform. Results can be useful at the time of generating keys in cryptographic algorithms (e.g., Diffie Hellman). We have experimented with finding prime numbers up to 8 bit. To calculate very large prime, we need to store number in file because it may not be stored in built in data types. We can also focus on devising a pattern or change in programs structure, so that all processors work together to find out if numbers are prime one at a time, instead of checking primality of different numbers. We have used up to 128 cores which can be extended to get better results.

References

1. Gowers, T.: Mathematics: A Very Short Introduction, Oxford University Press, ISBN 978-0-19-285361-5 (2002).
2. Narkiewicz, W.: The development of prime number theory: from Euclid to Hardy and Littlewood, Springer Monographs in Mathematics, Berlin, New York, Springer-Verlag, ISBN 978-3-540-66289-1 (2000).
3. Stallings, W.: Cryptography and Network Security: Principles and Practice, Prentice Hall, ISBN 978-0133354690 (2013).
4. Crandall, Richard; Pomerance, Carl (2005), Prime Numbers: A Computational Perspective (2nd ed.), Berlin, New York: Springer-Verlag, ISBN 978-0-387-25282-7.
5. Kozyrakis, C.E. and Patterson, D.A.: A new direction for computer architecture research. Computer Volume: 31. Digital Object Identifier: 10.1109/2.730733 (1998).
6. Grama A, George Kar.: Introduction to Parallel Computing, Addison-Wesley, ISBN 13: 9780201648652 (2003).
7. Amrasinghe Saman.: Multicore Programming Primer, MIT, OCW Video Course, January IAP (2007).
8. Mattson T. G., Sanders A., Massingill B.: Patterns for Parallel Programming, Addison-Wesley Professional, ISBN-13: 9780321228116 (2005).
9. Gamma Eric: Design Patterns, Prentice Hall (1994).
10. Macdonald S., Szafron D., Schaeffer J., and Bromling S.: From patterns to frameworks to parallel programs, Journal of Parallel and Distributed Computing (2001).
11. Goswami D., Singh A.: Composing Parallel Application using Design-Patterns, IEEE Conference (1997).
12. CDAC, Pune: PARAM Yuva II User Manual Version 1.0 (2015).
13. Hwu W., Keutzer K., Mattson T., The Concurrency Challenge, IEEE Design and Test, 25, 4, pp. 312–320 (2008).

Design and Analysis of High Performance CMOS Temperature Sensor Using VCO

Kumkum Verma, Sanjay Kumar Jaiswal, K.K. Verma and Ronak Shirmal

Abstract This paper presents a CMOS temperature sensor which is designed using self-bias differential voltage controlled ring oscillator at 180 nm TSMC CMOS technology to achieve low power. This paper focuses on design, simulation, and performance analysis of temperature sensor and its various components. In this used VCRO has full range voltage controllability along with a wide tuning range from 185 to 810 MHz, with free running frequency of 93 MHz. Power dissipation of voltage controlled ring oscillator at 1.8 V power supply is 438.91 µW. Different parameters like delay and power dissipation of individual blocks like CMOS temperature sensor component, voltage level shifter, counter and edge triggered D flip-flop are also calculated with respect to different power supply and threshold voltages. Power dissipation and delay of VCRO-based temperature sensor at 5 V power supply is 80.88 mW and 7.656 nS, respectively, and temperature range is from −175 to +165.

Keywords Voltage-to-digital converter · VCO · WSN · ADC · VLSI

1 Introduction

Temperature does not depend on any material. It is physical quantity which we use in our daily routine. Because of its independent behavior it has intensive property. Temperature sensor nowadays are used in VLSI implementation in the RFID and

Kumkum Verma (✉) · S.K. Jaiswal
ECE Department, Sangam University, Bhilwara, Rajasthan, India
e-mail: kumkum.verma1983@gmail.com

K.K. Verma
ECE Department, Dr.R.M.L.Avadh University, Faizabad, UP, India

Ronak Shirmal
ECE Department, Geetanjali Institute of Technical Studies, Udaipur, Rajasthan, India

wireless sensor network (WSN) application. On a single thin silicon wafer thousands of components are introduced [1–14].

So for the accuracy purpose importance of design of low voltage, a low-power circuit exists. The power consumed by batteries is high and supply voltage is comparatively low, demand of life time of the battery also takes place, all these factors show the requirement of low power system. The decreasing issue of power supply voltage is it stops the flow of signal in circuit. It creates problem for analog circuit design. Transistor characteristic also decreased because of low-voltage supply.

The value of device characteristic decreases due to scaling down of CMOS technology. All the high-resolution CMOS temperature sensors, like band gap temperature sensor based on analog-to-digital (ADC) and temperature sensor using thermal diffusivity sensing have been reported. High resolution temperature sensor consumes large area and significant power, due to their complex structure conversion rate is low because of all these it is difficult to add high resolution temperature sensor into other system. BJT-based temperature sensor that accepts the voltage-to-digital converter (VDC) is largely improved and rarely used. The VDC consists of a complex analog technique that requires more power and area temperature sensor based on the delay line. The delay line requires a stable reference signal to proceed. The ring oscillator can use main clock of microprocessor, or DRAMs as a reference signal. Temperature sensor which is based on a ring oscillator needs only the frequency-to-digital converter (FDC), i.e., mainly consists of a counter. This ring oscillator-based temperature sensor is useful and a better way to achieve high portability. A block diagram of temperature sensor is shown in Fig. 1.

Temperature is a physical quantity that measures hot and cold property of substance on a numeric scale. Output of the ring oscillator is proportional to the change in temperature. The positive rise in voltage shifter is determined by counter and after that it is saved in a register. The difference between outputs of register is defined as temperature.
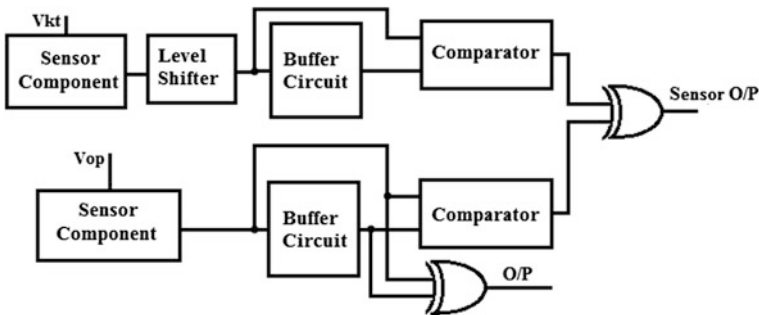
Fig. 1 Block diagram of temperature sensor

2 Design of Circuit

Different components used in designing of temperature sensors are given.

2.1 Ring Oscillator

Voltage control oscillator (VCO) takes one of the most important places in digital and analog circuit. The ring oscillator-based VCO, i.e., an implementation of VCO is usually used in the clock generation subsystem. The ring oscillator has integrating nature. Due to this nature the ring oscillator takes main place in many digital communication systems. They are used as VCO in many applications; one of them is clock recovery circuits for serial data communication.

In this Fig. 2, the MOSFETs M7 and M13 work as inverter while MOSFETs M2 and M18 are current source. To limit the current available to the inverter the current sources are used. The drain currents of MOSFETs are M1 and M12. The other transistors are added just because to form the 5-stages of ring oscillator. The 5-stage ring oscillator gives better VCO characteristic and frequency range.

2.2 Comparator

In the process of converting analog signal to digital signals CMOS comparator is commonly used. In an analog-to-digital conversion process, it is necessary to first sample the input.
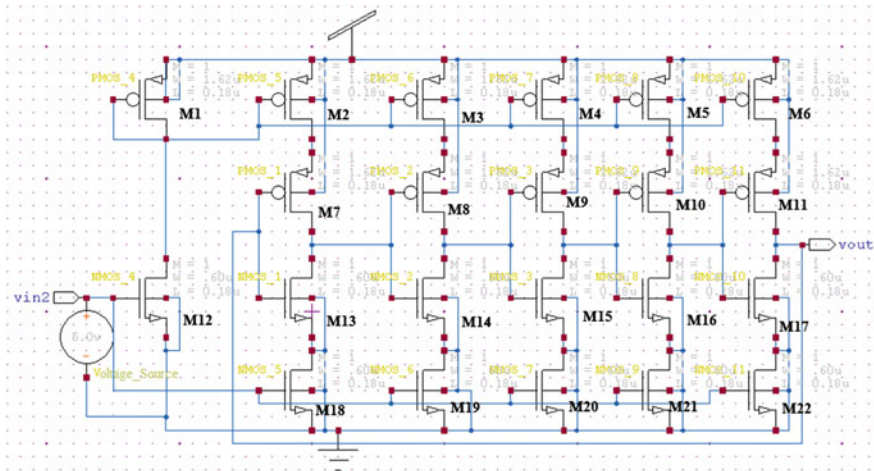
Fig. 2 Voltage controlled ring oscillator
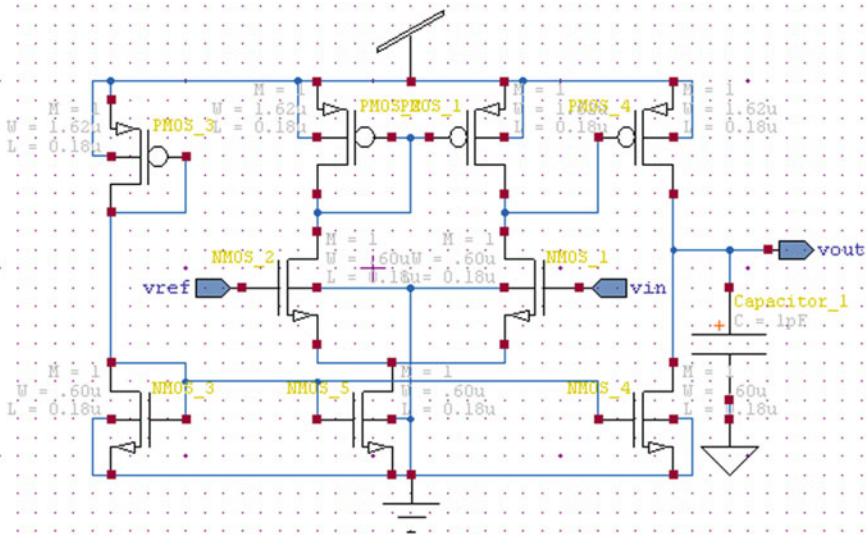
Fig. 3 CMOS comparator

This sampled signal is then used in a combination of comparators to calculate the digital equivalent of the analog signal. Comparator compares an analog signal with reference signal to generate an output based on the comparison. Schematic diagram of 1-bit comparator circuit is shown in Fig. 3.

The use of compressor in temperature sensor is comparing the current temperature value with the previous temperature value stored in buffer.

2.3 Temperature Sensor

The diagram of temperature sensor consists of two temperature sensor component, level shifter, two buffers, two comparator circuits, and two XOR gate. To sense the change in temperature (V_{ST}), temperature sensor component 1 is used and temperature sensor component 2 is kept at constant operating voltage (V_{OP}) of system.

Voltage level shifters convert the voltage which is taken from output of temperature sensor component 1 to provide perfect logic zero or perfect logic one output. Comparator compares every single temperature reading with previous temperature reading to produce output. These previous temperature readings are stored in buffer circuit.

The two comparator outputs are then passed into an XOR gate circuit which calculates on the basis of satisfactory condition. If both V_{ST} and V_{OP} are same then XOR gate gives logic zero output and if they are different then gives logic '1' output. The circuit diagram of temperature sensor is shown in Fig. 4.
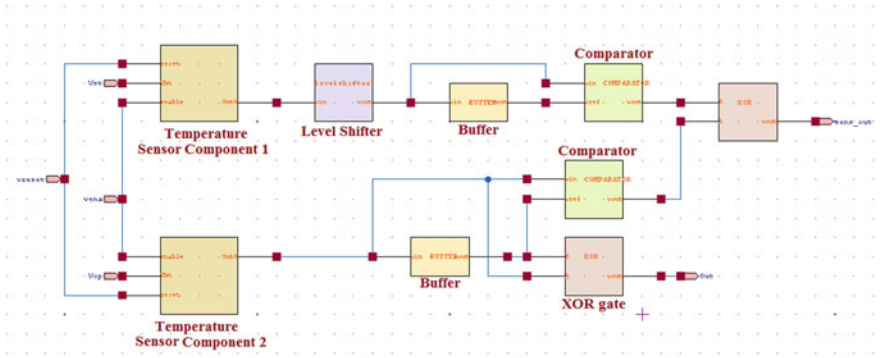
Fig. 4 Temperature sensor

3 Result and Discussion

3.1 Ring Oscillator

The voltage controlled oscillator is used to generate a specific frequency signal. The VCO is designed by 5-stage ring oscillator and that ring oscillator gives better VCO characteristic and frequency range. The simulated waveform of VCO and its characteristic is shown in Figs. 5 and 6.
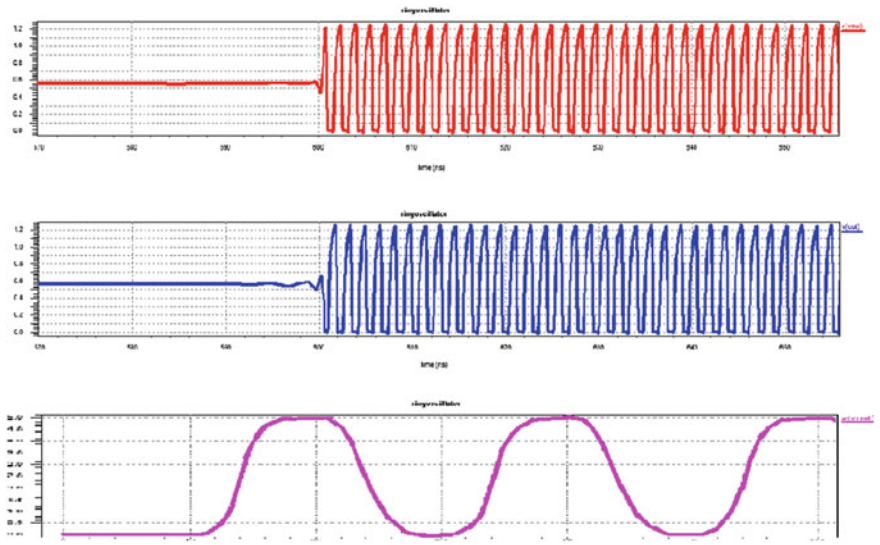
Fig. 5 Voltage controlled ring oscillator

Fig. 6 Characteristic of VCO
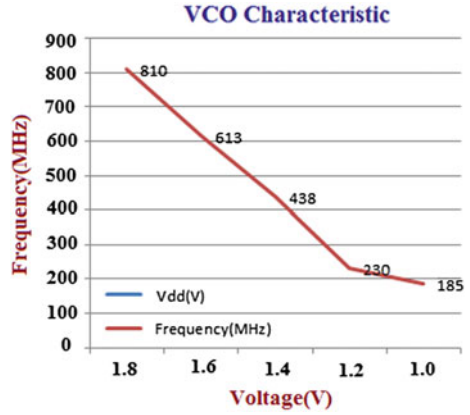
Table 1 Parameter of VCO at different voltage

Vdd (V)	Delay (nS)	Frequency (MHz)	Power dissipation (nW)
1.8	0.117	810	438.91
1.6	0.154	613	274.41
1.4	0.195	438	168.53
1.2	0.342	230	113.27
1.0	0.523	185	63.48

The value power dissipation, delay, and operating frequency range of voltage controlled ring oscillator at different power supply voltage is shown below in Table 1.

By increasing power supply voltage delay decreases, i.e., delay is inversely proportional to the power supply while power dissipation is directly proportional to square of power supply voltage. These parameters are also calculated at different threshold voltages as shown below in Table 2.

Table 2 Parameter of VCO at different threshold voltage

Vdd (V)	Delay (nS)	Frequency (MHz)	Power dissipation (nW)
0.17	0.092	746	492.62
0.27	0.131	599	315.02
0.37	0.195	438	168.53
0.47	0.232	251	67.85
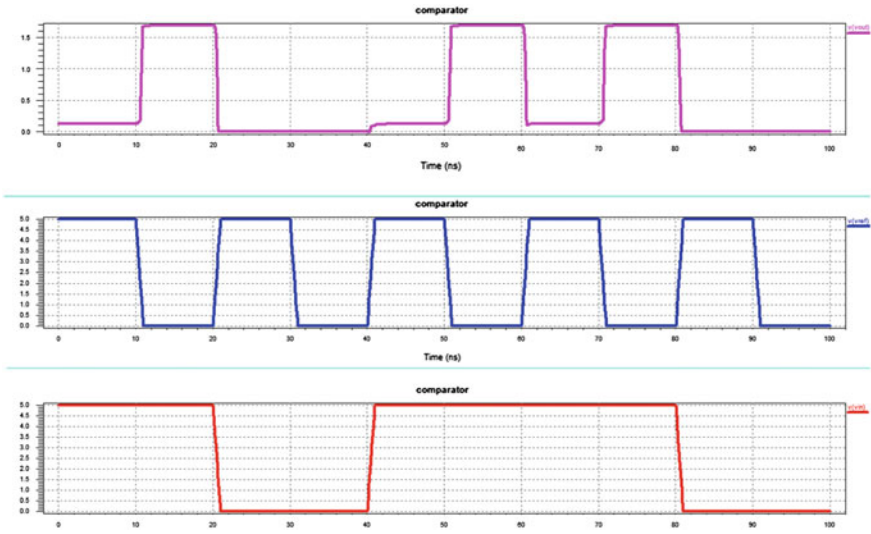
Fig. 7 Waveform of comparator
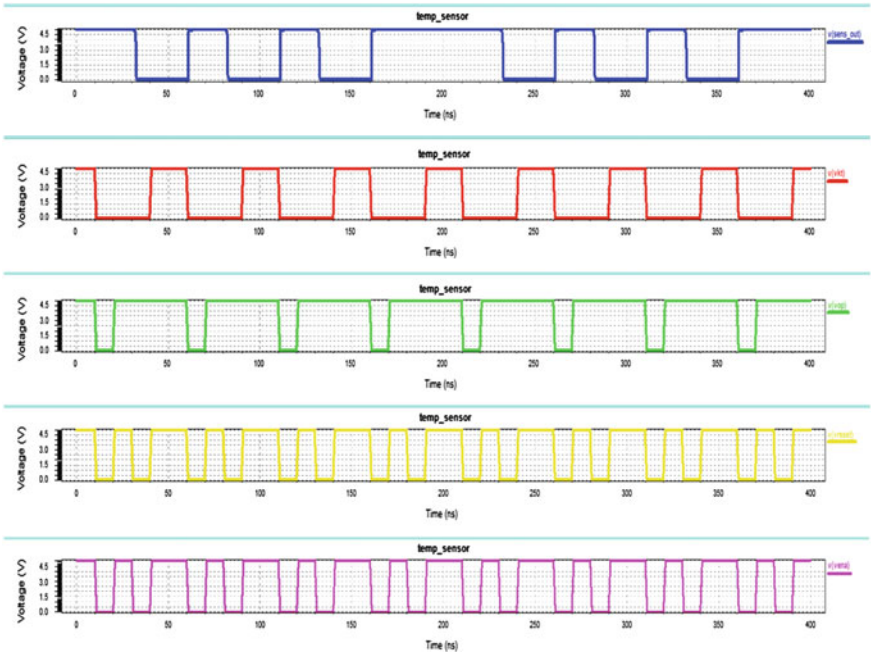
Fig. 8 Waveform of temperature sensor

3.2 Comparator

Comparator circuit is used to compare two signals to produce an output. It has two inputs v_{in} and v_{ref} and one output v_{out}. When input v_{in} is greater than v_{ref}, then output is high otherwise output is low as shown below in Fig. 7.

3.3 Temperature Sensor

Waveform of temperature sensor component is shown in Fig. 8.

Delay and power dissipation of temperature sensor component are determined by various parameters like power supply voltage, temperature, and threshold voltage as shown in Tables 3, 4 and 5, respectively.

Table 3 Parameters of temperature sensor at different temperature

Temperature (°C)	Delay (nS)	Power dissipation (mW)
27	7.656	80.88
37	6.923	83.235
47	5.357	89.56
57	3.514	102.35

Table 4 Parameters of temperature sensor at different voltage

V_{TH} (V)	Delay (nS)	Power dissipation (mW)
0.37	7.656	80.88
0.47	9.534	76.475
0.57	10.916	65.32
0.67	12.546	60.44

Table 5 Parameters of temperature sensor at different V_{TH}

V_{DD} (V)	Delay (nS)	Power dissipation (mW)
5	7.656	80.88
4	8.454	45.1
3	9.417	14.47
2	11.967	0.78

4   Conclusion

A voltage controlled ring oscillator-based CMOS temperature sensor has been designed at 180 nm CMOS TSMC technology in Tanner Tool 13.1. The proposed temperature sensor takes smaller silicon area with higher resolution than the conventional temperature sensor based on band gap reference. The characteristic of VCRO is designed between its control voltage and frequency. The frequency range of VCRO is determined as 185–810 MHz by its characteristic, with free running frequency of 93 MHz. Power dissipation of voltage controlled ring oscillator at 1.8 V power supply is 438.91 µW.

Various parameters like delay and power dissipation of other circuits are also calculated with respect to different power supply and threshold voltages. Power dissipation and delay of VCRO-based temperature sensor at 5 V power supply are 80.88 mW and 7.656 nS, respectively, and temperature range is from −175 to +165. From results we see that by increasing temperature power dissipation of circuit increases while delay decreases.

Reference

1. Monica Rose Joy and Thangamani M., "Design and Analysis of Low Power Comparator Using Switching Transistors", IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Vol. 4, Issue 2, PP 25–30, ISSN: 2319–4200, 2014.
2. Madhumathi S. and Ramesh Kumar, "Design And Analysis Of Low Power And High Speed Double Tail Comparator", International Journal Of Technology Enhancements And Emerging Engineering Research, Vol. 2, Issue 5, ISSN: 2347-4289, 2014.
3. Abhishek Rai, and B Ananda Venkatesan, "Analysis and Design of High Speed Low Power Comparator in ADC", IJEDR, Vol. 2, Issue 1, ISSN: 2321-9939, 2014.
4. Neil Weste and David Harris, "CMOS VLSI Design: A Circuits and Systems Perspective", 4 edition, Addison-Wesley publication.
5. M. Rabaey, Anantha Chandrakasan and Borivoje Nikolic, "Digital Integrated Circuits", 2nd Edition, Prentice Hall publication.
6. Poorvi Jain and Pramod Kumar Jain, "Design and Implementation of CMOS Temperature Sensor", International Journal of Current Engineering and Technology, Vol. 4, No. 2 (April 2014).
7. Young-Jae An, Kyungho Ryu, Dong-Hoon Jung, Seung-Han Woo and Seong-Ook Jung, "An Energy Efficient Time-Domain Temperature Sensor For Low-Power On-Chip Thermal Management", IEEE Sensors Journal, Vol. 14, No. 1, January 2014.
8. Xuehui Zhang and Mohammad Tehrani poor, "Design Of On-Chip Lightweight Sensors For Effective Detection Of Recycled ICS", IEEE Transactions on Very Large Scale Integration (VLSI) Systems.
9. Ruxi Wang, Dushan Boroyevich, Puqi Ning, And Kaushik Rajashekara, "A High-Temperature SiC Three-Phase AC–DC Converter Design For >100 °C Ambient Temperature", IEEE Transactions On Power Electronics, Vol. 28, No. 1, January 2013.
10. Kisoo Kim, Hokyu Lee and Chulwoo Kim "366-Ks/S 1.09-Nj 0.0013-Mm2 Frequency-To-Digital Converter Based Cmos Temperature Sensor Utilizing Multiphase Clock", IEEE Transactions On Very Large Scale Integration (VLSI) Systems, Vol. 21, No. 10, October 2013.

11. Hua Wang, Ching-Chih Weng, And Ali Hajimiri, "Phase Noise And Fundamental Sensitivity Of Oscillator-Based Reactance Sensors", IEEE Transactions on Microwave Theory And Techniques, Vol. 61, No. 5, May 2013.
12. Nima Sadeghi, Alireza Sharif-Bakhtiar, And Shahriar Mirabbasi "A 0.007-108-1-MHz Relaxation Oscillator For High-Temperature Applications Up To 180 In 0.13- CMOS", IEEE Transactions On Circuits And Systems, Vol. 60, No. 7, July 2013.
13. Clifton L. Roozeboom, Matthew A. Hopcroft, Wesley S. Smith, Joo Yong Sim, David A. Wickeraad, Peter G. Hartwell, And Beth L. Pruitt, "Integrated Multifunctional Environmental Sensors", Journal Of Micro electromechanical Systems, Vol. 22, No. 3, June 2013.
14. Matthew Spencer and Steven Callender, "Digital Temperature Sensing in a Variable Supply Environment", IEEE Transactions On Electron Devices, Vol. 52, No. 11, November 2010.

High Performance Add Drop Filter Based on PCRR for ITU-T G.694.2 CWDM System

Ekta Kumari and Pawan Kumar Inaniya

Abstract In this work, we demonstrated add drop filter using photonic circular ring resonator. The proposed design incorporated square lattice structure and silicon rods that are enclosed by air. Dropping efficiency and coupling efficiency are explored by utilizing 2D finite-difference time-domain (FDTD). The designed filter gives nearly 100 % of dropping efficiency at 1500 nm and coupling efficiency at 1471 nm with refractive index of 3.47. Photonic band gap of the proposed design has been assessed through plane wave expansion method (PWE).

1 Introduction

Photonic Crystal has received an impetus and attracted academia in the optical world and numerous investigators have gainful attention for the investigation of filters. Photonic crystals are nano structures or dielectric materials which have a particular dielectric constant in some specific dimensions [1]. Because of the periodicity, it becomes in the form of photonic band gap (PBG). Through inserting the defects such as line defect and point defect in the structure, the light is propagated in the PBG region. The two-dimensional photonic crystal played an important role in making the photonic devices.

In the time of the literature, photonic crystal (PC)-based ADF is designed and investigated using square lattice [2, 3] and triangular lattice [4, 5]. Amidst this, two-dimensional square lattice PC-based add drop filter is easy to fabricate and has simple design as compared to others. The square lattice photonic crystal-based ADF is designated using the PCRR [6–8].

Add drop filter is the most prominent device that is used to either drop or select the desired wavelength. In CWDM systems, respected wavelengths are added or

Ekta Kumari (✉) · P.K. Inaniya
Govt. Women Engineering College Ajmer, Ajmer, India
e-mail: ecektachoudhary@gmail.com

P.K. Inaniya
e-mail: pawan.inaniya@gmail.com
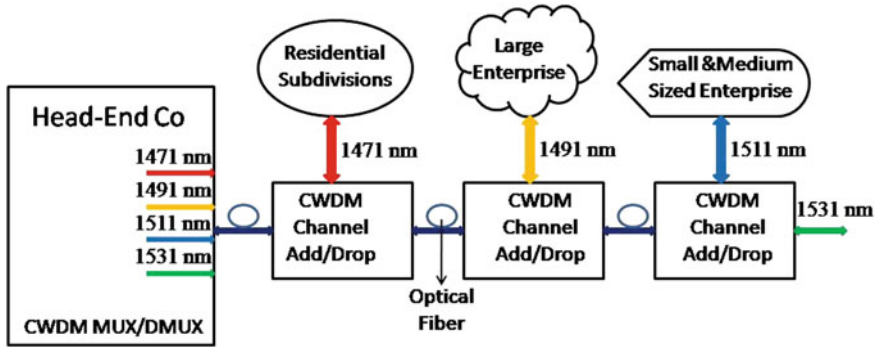
Fig. 1 Basic structure of CWDM network

rerouted by add drop filter. More prevailing add drop filters are such as ring resonator-based filters, acousto-optic filters, thin film filters, arrayed waveguide grating-based filters, liquid crystal filters, etc. For photonic integrated circuits, conventional filters are not acceptable. But for filtering basis, ring resonators depend on add drop filter which is an irresistible applicant and provides better retort [9–11].

In the Fig. 1 shows the overview of the CWDM network. CWDM network is incorporated with CWDM Multiplexer or Demultiplexer and CWDM add drop Filter (ADF). Hence, for much kind of applications, the four different wavelengths such as 1471 nm, 1491 nm, 1511 nm, and 1531 nm are sent from central office to many numbers of places such as large, small/medium enterprises, and residential subdivisions. ADF is used to add and drop the respective wavelength bands. These entire three wavelengths are generated by the CWDM Multiplexer/Demultiplexer device [16, 17]. The ring resonator-based ADFs having square lattice are designated in the literature using square cavity [6], hexagonal cavity [12], quasi square cavity [13, 14], or dual curved cavity [15].

Single point-defect cavity is used to make the smallest photonic crystal ring resonator that had ultra small cavity size. A small-scale photonic crystal has extremely low loss and highest quality factor that is much preferable.

2 Structure of Circular Photonic Crystal Ring Resonator

In this proposed work, we introduced a modern type of circular photonic ring resonator, bus waveguide, and drop waveguide. Finite element method is used to determine band gap of the structure. In this design, we have 17 and 21 numbers of silicon rods in X-axis and Z-axis, respectively. The distance betwixt the two adjoining rods of silicon is 540 nm, that is referred as lattice parameter, denoted by "a". Radius of silicon rods is 0.1 μm having index of refraction 3.47 which is enclosed in air.
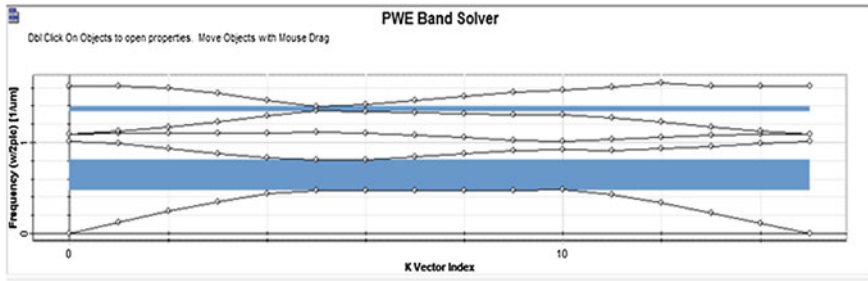
Fig. 2 Band diagram of 17 × 21 photonic crystal ring resonator

Figure 2 represents the band gap diagram of proposed structure by inserting any defects. Traverse electric (TE) mode is only used for measuring the band gap in PBG region. Here, only two TE modes are obtained in the structure. The first one TE PBG is obtained from 0.487 a/λ to 0.805 a/λ whose respective wavelengths are 670 to 1108 nm. Second one PBG is obtained from 1.352 a/λ to 1.395 a/λ whose respective wavelengths are 387 to 399 nm.

Proposed design is incorporated by two waveguides and a circular shaped photonic crystal ring resonator which is placed betwixt these two waveguide in Γ–X direction. Input signal is applied as a Gaussian signal that is noted by port "A" in bus waveguide. Output is provided through power monitor that is obtained at received port, noted by port "B". The top most waveguide is known as bus waveguide and the bottom waveguide that is placed at the right field of ring resonator in upright direction is known as dropping waveguide.

Both bus waveguide and drop waveguide are created by using line defects and a circular ring resonator is made up of point defect. In circular ring design, inner rods and outer rods are formed by shifting 25 % of lattice constant 'a' toward the center. Scatter rods are used to minimize the backward propagation, which are deposited at four sides having half lattice constant. Dropping efficiency and coupling efficiency are examined at port B and port C.

In the designed structure we are using 9 × 9 number of silicon rods in a circular ring resonator. Silicon rods are arranged in ring by using point defect. By varying the pitch of air holes, silicon rods will be shifted and photonic crystal ring resonator is generated and in dropping waveguide one silicon rod will be shifted by altering the lattice parameter of 50 % to minimize the backscatter (Fig. 3).

3   Simulation and Results

A Gaussian signal is applicable as input port and for obtaining the coupling/dropping efficiency at output is captured by power monitor at port "B" and "C". The normalized transmission spectrums at port "B" and port "C" are incurred using fast Fourier transform (FFT) that is reckoned by 2D finite-difference time-domain method.

Fig. 3 **a** Design of photonic crystal ring resonator. **b** Refractive index view
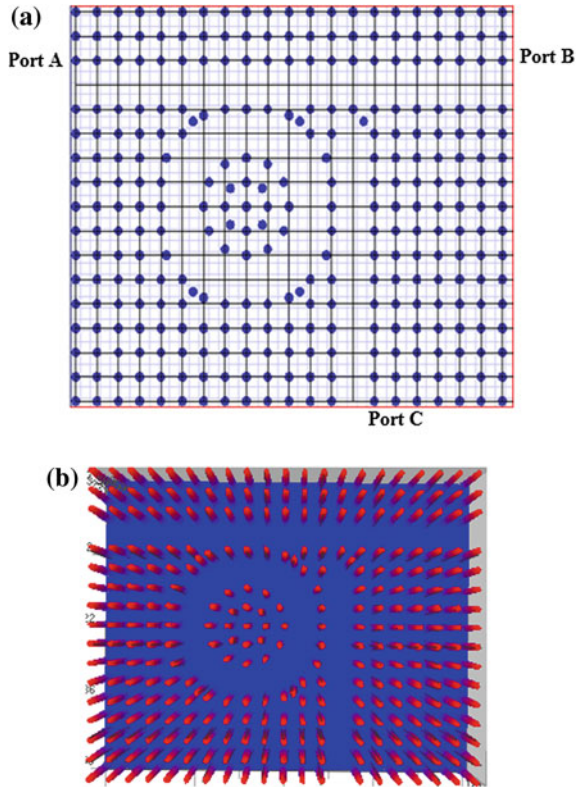
Fig. 4 Normalized transmission spectrum of circular ring resonator
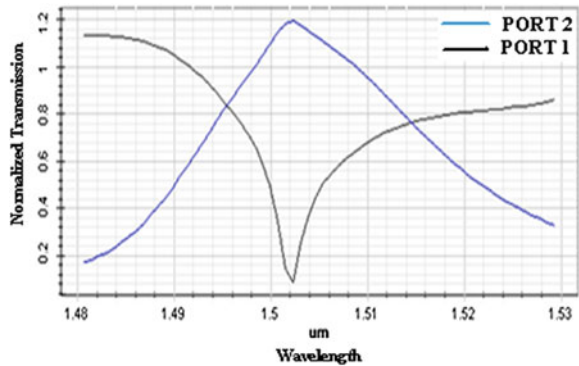
Figure 4 shows normalized transmission spectrums of circular ring resonator (PCRR). Simulation represents that coupling and dropping efficiency are close to 100 %.

When there is a change in the index of refraction of the designed structure, the wavelength of resonant filter has been changed because refractive index is restricted to the light propagation with certain wavelength in designed structure. At certain

Fig. 5 a Electric field pattern at resonant wavelength 1500 nm. b Electric field pattern at resonant wavelength 1471 nm
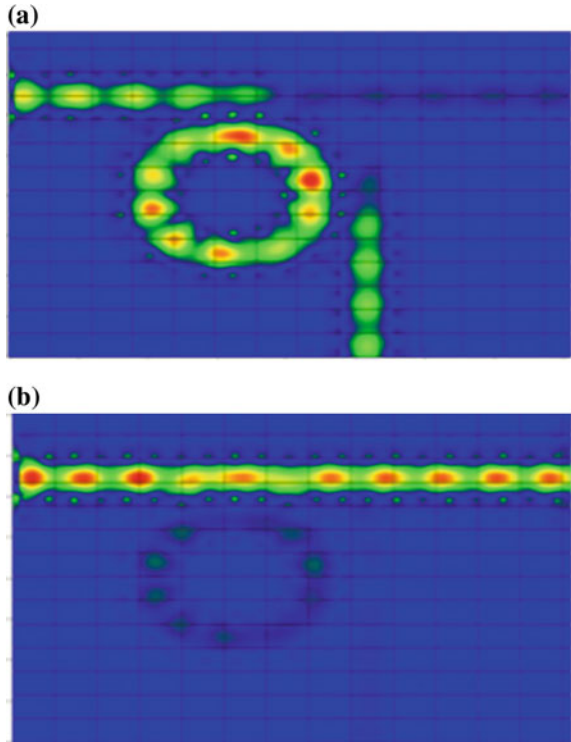
wavelength, light is transferred through bus waveguide without disturbing the circular ring resonator that arrived at port "B" this condition is known as off-resonant condition. When the light is directly coupled to the ring resonator and propagates through the dropped waveguide and appeared at port "C", this condition is known as on-resonant condition.

Figure 5a, b show electric field pattern of circular PCRR at resonant wavelength 1500 nm and 1471 nm, respectively.

At resonant wavelength 1500 nm, signal is fully coupled with circular PCRR and obtained at port "C" in dropping waveguide. The coupling efficiency that is given by dropping waveguide is 100 %. On resonant condition signal provides forward dropping efficiency.

At resonant wavelength 1471 nm, signal does not couple to any ring resonator and directly passed through the bus waveguide, obtained at port "B". Output port "B" also provides dropping efficiency that is close to 99 %. It is also termed as off-resonant condition.

Hence, the strength of an electric field is ascertained distinctly and designated add drop filter (ADF) demonstrates favorable property afterward optimization.

4   Conclusion

In this work, the proposed design based on circular ring resonator (PCRR) depends on add drop filter (ADF). The study and simulation show the 100 % coupling efficiency at 1500 nm and 1470 nm wavelengths that have been obtained by utilizing a PCRR in proposed design. The spacing between two wavelengths is 30 nm that is more preferable for CWDM system. These simulation results are analyzed by using 2D finite-difference time-domain (FDTD) method. The gross domain of wafer dimension is 11.2 μm × 8.9 μm. Hence this type of device would be further practical for the recognition of CWDM systems in integrated optic circuits.

References

1. D. Joannopoulos, R. D. Meade and J. N. Winn, 'Molding the Flow of Light Photonic Crystals', 2nd edition, 2008.
2. Chun-Chih Wang, Lien-WenChen, "Channel drop filters with folded directional couplers in two-dimensional photonic crystals", Physica B, vol. 405, pp. 1210–1215, 2010.
3. Min Qiu and Bozepa Jaskorzynska, "Design of a channel drop filter in a two-dimensional triangular photonic crystal", IEEE Journal of Applied Phys. Lett, vol. 83, no. 6, pp. 1074–1076, 2003.
4. Hongliang Ren, Chun Jiang, Weisheng Hu, Mingyi Gao, Yang Qu, and Fanghua Wang, "Channel drop filter in two-dimensional triangular lattice photonic crystals," Opti Express, vol. 24, no. 10, pp. A7–A11, 2007.
5. M. S. Vasconcelos, P. W. Mauriz, E. L. Albuquerque, "Optical filters based in quasi periodic photonic crystal", Journal of Microelectronics, vol. 40, pp. 8 51–8 53,2009.
6. V. Dinesh Kumar, T. Srinivas, A. Selvarajan, "Investigation of ring resonators in photonic crystal circuits," Photonics and Nanostructures 2, pp. 199–206, 2004.
7. M. David, A. Ghalfari, F. Monif M. S. Abriarnian, "T-Shaped channel drop filters using photonic crystal ring resonators", Physics E, vol. 40, pp. 3151–3154, 2008.
8. Juan Joe's Vegas Ohnos, Masatoshi Tokushima, and Kenichi Kit yarn, "Photonic add-drop filter based on integrated photonic Crysta1 structures", IEEE Journal of Selected Topics in Quantum Electronics, vol 16, no. 1, pp. 332–337, 2010.
9. Nakkeeran R and Robinson S "Add Drop Filter for ITU-T G.694.2 CWDM Systems" ©2011 IEEE.
10. D. Sadot, E. Boimovich, "Tunable optical filters for dense WDM networks", IEEE Communication Magazine, vol. 36, no. 12, pp. 50–55, 1998.
11. S. Fan, P. R Villeneuve, J. D. Joannopoulos s, H. A. Haus "Channel drop filters in photonic crystals", Opti. Express, vol. 3, no. 1, pp. 4–11, 1998.
12. Trong Thi mai, Fu-Li Hsiao, Chengjuo Lee, Wenfeng Xiang, ChiiChang Chen, W. K. Cho "Optimization and comparison of photonic crystal resonators for silicon micro cantilever sensors", Semvrs and Actual Ova, pp. l–l0, 2010.
13. B.S. Darki, N. Granpaye, "Improving the performance of a photonic crystal ring— resonator-based channel drop filter using particle swan optimization method", Optics Communications, in press, vol. 15232, pp. I–5,2010.
14. M. David, A Ghalfar F. Monif M.S. Abriarnian, "A new broadband photonic crystal add drop filter", Journal of Applied Sciences, vol. 8, no. 11, pp. 2178–2182, 2008.

15. P. Andalib and N. Granpayeh, "Optical add/drop filter based on dual curved photonic crystal resonator," 5" IEEE international Conference on Photonics, pp. 249–250, 2008.
16. Robinson S and Nakkeeran R, "Channel Drop Filter based on 2D Square-Lattice Photonic Crystal Ring Resonator" ©2010 IEEE.
17. S. Robinson and "Hetero structure Based Add Drop Filter for ITU-T G.694.2 CWDM Systems Using PCRR" R. Nakkeeran2 IEEE-2018.

Efficient Data Dissemination in Wireless Sensor Network Using Adaptive and Dynamic Mobile Sink Based on Particle Swarm Optimization

Nivedita Kumari and Neetu Sharma

Abstract Wireless sensor networks are an emergent technology for monitoring physical world. The energy constriction of Wireless sensor networks makes energy (resource) saving and elongating the network lifetime becomes the utmost imperative goalmouths of numerous routing conventions. Network segment division is a strategic technique used to prolong the lifetime of a sensor network by plummeting energy consumption. Subsequently sensor nodes adjacent to sink have to tolerate more traffic encumbrance to forward data, they will hastily diminish their energy which directs network partition and energy hole problem. Introducing mobility into sensor networks brings in new prospects to mend network performance in rapports of energy consumption, network lifetime, latency, throughput, etc. In this paper, a protocol is proposed for efficient data dissemination using mobile sink in wireless sensor networks. We used multiple mobile sink in order to gather the sensed data from predetermined paths and it moves in forward direction and returns back to the same position. The predetermined paths are calculated intelligently through particle swarm optimization (PSO) based on heterogeneity of energy (current resources with respect to average resources) and network density as an objective functions. In order to gather data from reference area, mobile sinks stay temporarily at some fixed point. The potency of our proposed algorithm will be in terms of network lifetime, number of dead sensors, energy consumption, end-to-end delay, network throughput, etc.

Keywords Wireless sensor network · Energy efficient routing · Mobile sink · Heterogeneous network · Energy efficiency · Particle swarm optimization · Artificial intelligence · Energy ratio distribution

Nivedita Kumari (✉) · Neetu Sharma
Department of Computer Science and Engineering,
Govt. Engineering College Ajmer, Ajmer, India
e-mail: nivedita2042@gmail.com

Neetu Sharma
e-mail: neetucom10@gmail.com

1 Introduction

Wireless sensor networks (WSNs) is one of the developing research areas, it contains a lot of autonomous and distributed sensor nodes. These nodes are power-limited, small size, low-cost, and multifunctional devices with processing, sensing, and wireless communication potentialities. Each sensor node has some work to measure and monitor the ambient circumstances and forward the collected data to a sink node in a direct infection manner or multiple hop infection manner. Nowadays, wireless sensor network has wide range of applications, such as surveillance and military target tracking, biomedical health monitoring, smart home management, disaster warning system, greenhouse monitoring, wildlife animal protection, etc. [1, 2].

Sensor nodes have limited battery power. This is one of the most substantial confining roles for WSNs. Since sensor nodes are normally spread over special areas where human cannot get close to, it is very hard, it is not possible to recharge and replace the battery [3, 4]. So, in order to increase the network lifespan, it must be necessary to design and manage the energy resource of sensor networks carefully.

Sensor nodes which are close to sink, take a participation in forwarding the data to sink node for other sensor nodes. There is a result that they deplete their energy rapidly and lead to the energy hole problem because they die soon [5, 6].

Introducing mobile sinks in spite of static sinks to expand network lifespan for wireless sensor networks has attracted increased attention in recent years [7]. Here the mobile strategies for wireless sensor networks do not deploy the mobile sensor nodes. It is the strategies for deployment of mobile sink node(s). Implementation of the deployment of mobile sink node is easier than the deployment of mobile sensor nodes. Since sink nodes are mobile so sink's neighbor nodes change time by time, so that problem of energy hole due to static sink can be extenuated and energy expenditure among sensor nodes can be equilibrated. Therefore, the network lifespan can be increased.

In this paper, to collect the sensed data from sensor nodes, three mobile sinks are used instead of one static sink. The mobile sinks will move in forward direction and return back to the same position along preset paths and in order to collect the data from reference area, mobile sink stays temporarily at some fixed point. Here, two mobile sink are placed on arc lines and one mobile sink placed on diameter of circular area and they move in preset path, respectively. Whole network area is divided into two parts: one part is the concentric circle of the sensor field area with radius r and second one is the remaining part of the sensor field area. When the mobile sink stays temporarily at the midpoint (center) of the circular area, it receives the data sensed by the sensor nodes in the concentric circle of the area and when the mobile sink reaches to the other fixed points, it collects the sensed data from sensor nodes in the second part.

Fig. 1 Network model
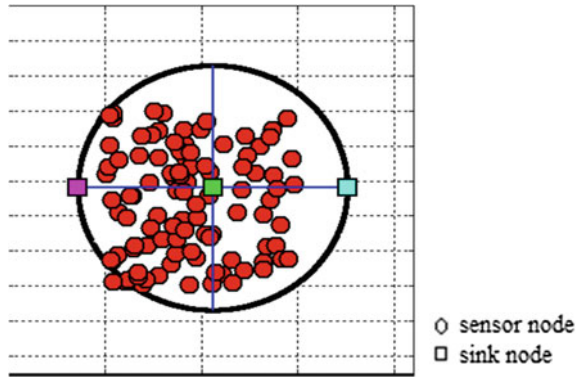
2 System Model

Basic Assumptions:

- Sensor nodes have unique ID and they are heterogeneous.
- Sinks are mobile.
- Predetermined paths are calculated intelligently through particle swarm optimization (PSO) based on heterogeneity of energy (current resources with respect to average resources) and network density as an objective functions.
- Sinks are placed at certain points and sinks are mobile so it moves from one site to another freely.

3 Network Model

Assuming that a number of sensor nodes are randomly spread over the network field to monitor the surrounding environment. Figure 1 shows the network model where sink node moves along periphery or diameter of the circle.

4 Procedure of Proposed Model

We are going to apply the method in a sensor field. Area of sensor field is about 100×100 m. However the area of field can be changed as per the variations of result. Initially, the centre of sensor field is occupied by the base station, the next position of sink can be decided through the optimum path and this is selected by PSO. Quantity of initial energy in a node is the residual energy and initially the degraded energy is zero, Hence Total energy E_t is the sum of degraded and residual

energy. So we can say that total energy is the amount of residual energy. Also after the particular cycle, average energy Ea of a node can be calculated with a particular number of cycle and the knowledge of total energy.

$$\text{Ea} = E_t \times \left(\frac{1 - (r/R_{\max})}{n} \right)$$

Before assigning a cluster head, and at every new cycle its value renewed, we calculate the dead statistics. For optimum probability, the new expression can be calculated from different levels of energy and optimum probability defined earlier. In traditional protocol, it is considered to be random.

$$p(i) = \frac{n \times \text{current Energy}}{\text{distance from BS} \times \text{average energy}} \times \text{Node Density}$$

Based on agent selection (with respect to network density, distance from sink, and energy criteria) objectives above, particle swarm optimization calculate the optimum path for mobile sink within the wireless sensor network. The distance from sink is only considered so that sink should not have to move a longer path in every instance of communication. Here, if the probability structure is greater than temporary number (between 0 and 1) assigned to node below, that node will become an agent.

$$T(s_i) = \begin{cases} \dfrac{P_i}{1 - P_i \left(r \bmod \frac{1}{P_i} \right)} & \text{if} \in G \\ 0 & \text{otherwise} \end{cases}$$

Here P_i appears from new expression of optimum probability $P(i)$.

Hence the above criteria can be fulfilled by the node with higher energy, which is in denser area and lesser distance from sink amongst the other nodes and therefore that node acts as agent to transmit data for a more retentive period which gives results as throughput increment and maximize network lifespan.

After a more prominent criteria sensors are selected as agents, energy expended by it on that particular cycle and finishes the cycle of steady state phase and that spent amount of energy is calculated by energy model.

$$E_{TX}(l, d) = \begin{cases} lE_{\text{dec}} + l\varepsilon_{fx}d^2, & d < d_0 \\ lE_{\text{dec}} + l\varepsilon_{fx}d^4, & d \geq d_0 \end{cases}$$

The node which goes to the set of member node, and acts as the normal node and finishes the cycle of steady state phase, If a node is not a more prominent criteria node it is discarded from the criteria above. By replacing several sinks with a dynamic one which can move in field area based on actual network statistics can increase the network lifetime and efficiency drastically. Figure 2 shows the basic flow diagram of communication procedure.
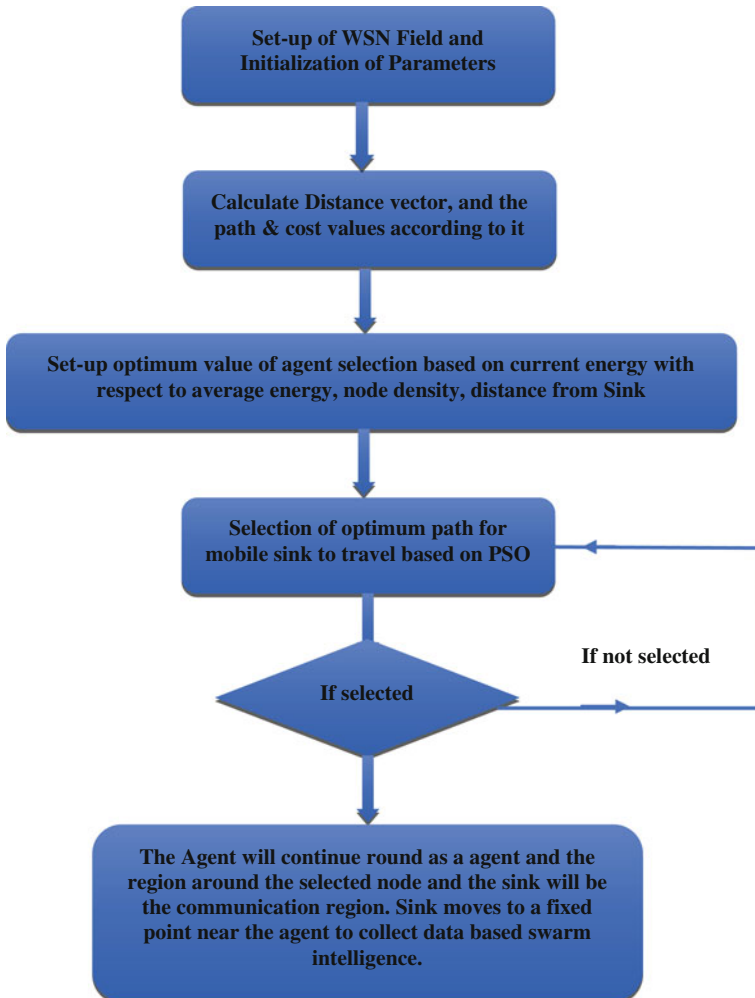
Fig. 2 Basic flow diagram of proposed algorithm

5 Performance Evaluations

In this section, to evaluate the performance of our proposed algorithm we use MATLAB. Suppose that 100 sensor nodes are randomly spread over the circular field and the radius of circular field is 100 m, initial energy of sensor nodes 0.5 J.

The performance of the network can be analyzed, to Analyze we have to plan to take a comparability between our proposed algorithm and other two different algorithms: the first one is by using mobile sink routing (Jin Wang) which divides the circle in three parts and travels along the fringe of the network and second one is, using only one mobile sink (Random Sink Position Based Routing), In which the

Fig. 3 Lifespan of sensors
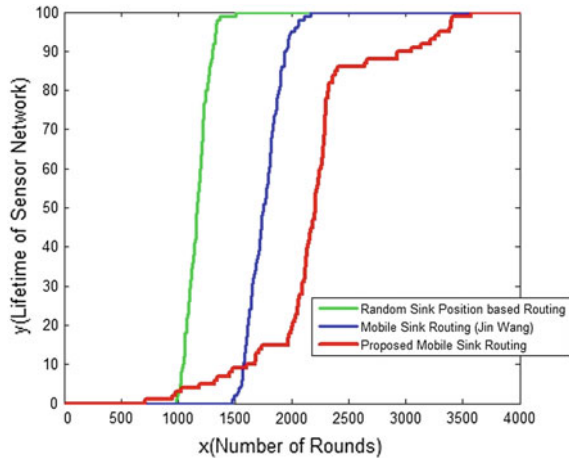
mobile sink moves along the periphery of the network. Both the methods have almost same result, except the performance time. One mobile sink takes more time to collect the data from sensor nodes from reference area than the time taken by three mobile sinks (as shown in the Fig. 3). Therefore three mobile sink can save more time.

Comparison of the above two algorithms in terms of network lifespan is shown in Fig. 3. In the given figure network lifespan is defined as the round. From the figure it is clear that the network lifespan using three mobile sink is better than the lifetime using one mobile sink. This figure also indicates that the sensor nodes die much faster by using one mobile sink, as the round number increases.

Comparison of end-to-end delay in different conditions is shown in Fig. 4. Latency is an important issue for wireless sensor networks. From Fig. 4 we can see that end-to-end delay decreases when the number of round increases. Delay of using only one mobile sink is about 1.5 times greater than by using our proposed

Fig. 4 End-to-end delay comparison
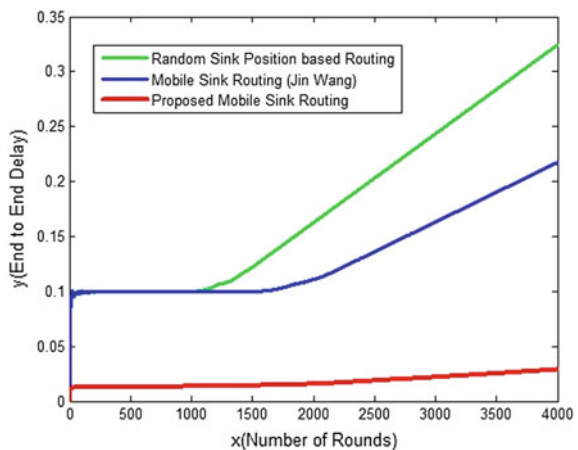
Fig. 5 Network throughput
comparison
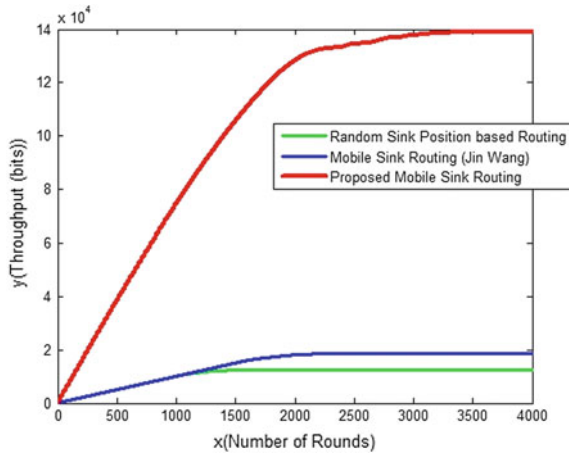
algorithm (As shown in the Fig. 4). Means that our proposed algorithm gives much
better performance in terms of end-to-end delay.

Comparability of our proposed algorithm with the other two methods is in terms
of network throughput. The rate of the sensed data collection by using the one
mobile sink is minimum as compared to the rate of data collection done by three
mobile sinks. As the number of round increases, more data are collected by three
mobile sink (shown in Fig. 5).

6 Conclusion

It is a better option to deploy more than one mobile sink to the WSNs to increase
the network lifespan. We can see that end-to-end delay decreases when the number
of round increases. Delay by using only one mobile sink is about 1.5 times greater
than by using our proposed algorithm. Figure 3 also indicates that the sensor nodes
die much faster by using one mobile sink, as the round number increases. In this
paper, we presented our efficient data dissemination in wireless sensor network
using adaptive and dynamic mobile sink based on particle swarm optimization
algorithm, where the mobile sinks proceed along preset routes to collect the data
sensed by sensor nodes. The functioning of our proposed algorithm is demonstrated
through simulation results.

References

1. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, Wireless sensor networks: a
   survey, *Computer Networks*, Vol. 38, No. 4, 2002, pp. 393–422.
2. K. Akkaya and M. Younis, A survey on routing protocols in wireless sensor networks, *Elsevier
   Ad Hoc Network*, Vol. 3, No. 3, 2005, pp. 325–349.

3. P. T. A. Quang and D. S. Kim, Enhancing Real-Time Delivery of Gradient Routing for Industrial Wireless Sensor Networks, *IEEE Transactions on Industrial Informatics*, Vol. 8, No. 1, 2012, pp. 61–68, Feb.

4. S. Lindsey, C. Raghavendra, K. M. Sivalingam, Data gathering algorithms in sensor networks using energy metrics, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 13, No. 9, 2002 pp. 924–935.

5. D. Wei, Y. Jin, K. Moessner, R. Tafazolli, An Energy-Efficient Clustering Solution for Wireless Sensor Networks, *IEEE Transactions on Wireless Communications*, Vol. 10, No. 11, 2011, pp. 3973–3983.

6. W. R. Heinzelman, A. Chandrakasan and H. Balakishnan, Energy-efficient communication protocol for wireless microsensor networks, *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 2002.

7. Chatzigiannakis, A. Kinalis, S. Nikoletseas, Efficient data propagation strategies in wireless sensor networks using a single mobile sink, *Computer Communications*, Vol. 31, No. 5, 2008, pp. 896–914.

A New Approach to Intuitionistic Fuzzy Soft Sets and Its Application in Decision-Making

B.K. Tripathy, R.K. Mohanty, T.R. Sooraj and K.R. Arun

Abstract Soft set theory (Comput Math Appl 44:1007–1083, 2002) is introduced recently as a model to handle uncertainty. Recently, characteristic functions for soft sets and hence operations on them using this approach were introduced in (Comput Math Appl 45:555–562, 2003). Following this approach, in this paper we redefine intuitionistic fuzzy soft sets (IFSS) and define operations on them. We also present an application of IFSS in decision-making which substantially improve and is more realistic than the algorithms proposed earlier by several authors.

Keywords Soft sets · Fuzzy sets · FSS · IFSS · Decision-making

1 Introduction

Notion of fuzzy sets introduced in [1] is one of the most fruitful models of uncertainty and has been extensively used in real-life applications. In order to bring topological flavour into the models of uncertainty and associate family of subsets of a universe to parameters, soft sets were introduced in [2]. Subsequently, operations on these sets were introduced [3, 4]. Hybrid models obtained by suitably combining individual models of uncertainty have been found to be more efficient than their

B.K. Tripathy (✉) · R.K. Mohanty · T.R. Sooraj
School of Computing Science and Engineering, VIT University, Vellore,
Tamil Nadu, India
e-mail: tripathybk@vit.ac.in

R.K. Mohanty
e-mail: rknmohanty@gmail.com

T.R. Sooraj
e-mail: soorajtr19@gmail.com

K.R. Arun
School of Information Technology and Engineering, VIT University, Vellore,
Tamil Nadu, India
e-mail: arun.kr@vit.ac.in

components. Several such hybrid models exist in literature. Following this trend the notion of FSS was put forward in [5]. In [6] soft sets were defined through their characteristic functions. This approach has been highly authentic and helpful in defining the basic operations on soft sets. Similarly, it is expected that defining membership function for fuzzy soft sets will systematize many operations defined upon them as done in [7]. Extending this approach further, we introduce the membership functions for IFSS in this paper. Decision-making using soft sets was discussed in [3]. This study was further extended to the context of FSSs by Tripathy et al. in [7], where they identified some drawbacks in [3] and took care of these drawbacks while introducing an algorithm for decision-making. In this paper, we have carried this study further by using IFSS instead of FSS in handling the problem of multi-criteria decision-making.

Intuitionistic fuzzy set (IFS) was introduced in [8]. It is a fruitful and more realistic model of uncertainty than the fuzzy set. The notion of non-membership function introduced, which does not happen to be one's complement of the membership function introduces more generality and reality to IFS. The hesitation function generated as a consequence is what real-life situations demand. In case of fuzzy sets the hesitation component is zero. Among several approaches, Maji et al. [5] have defined FSS and operations on it. Here, we follow the definition of soft set from [6] in defining IFSS and redefine their union, intersection, complement, and some other operations on them. It may be noted that following the same approach we have extended the definition in [6] to the context of FSS. The major contribution in this paper is introducing a decision-making algorithm which uses IFSS and we illustrate the suitability of this algorithm in real-life situations. Also, it generalises the algorithm introduced in [7] while keeping the authenticity intact.

2   Definitions and Notions

Let U, E denote the universal set and parameter set respectively. Elements of U are denoted by x and those of E are denoted by e. Let $P(U)$ and $I(U)$ be the power set and fuzzy power set of U respectively.

Definition 1 (Soft Set) We denote a soft set over (U, E) by (F, E), where

$$F : E \rightarrow P(U). \tag{1}$$

Definition 2 (FSS) We denote a FSS over (U, E) by (F, E) where

$$F : E \rightarrow I(U). \tag{2}$$

3 Intuitionistic FSS

An IFS, A is characterized by two functions μ_A and v_A called the membership and non-membership function of A respectively such that $\mu_A : U \rightarrow [0, 1]$ and $v_A : U \rightarrow [0, 1]$. For any $x \in U$, We have $0 \le \mu_A(x) + v_A(x) \le 1$. The hesitation function of A is denoted by π_A and for any $x \in U$, is given by $\pi_A(x)$. The indeterministic part of x is $\pi_A(x) = 1 - \mu_A(x) - v_A(x)$.

In this section, we describe the main notions used and the operations of IFSSs. Let IF (U) be the intuitionistic fuzzy power set of U.

Definition 3 A pair (F, E) is an IFSS over (U, E), where F is given by

$$F : E \rightarrow \text{IF}(U). \tag{3}$$

$\mu^a_{(F,E)}(x)$ and $v^a_{(F,E)}(x)$ denote the membership value and non-membership value of x in (F, E) with respect to parameter a respectively.

For two IFSSs (F, E) and (G, E), we have,

Definition 4 We say that (F, E) is an IF soft subset of (G, E) if

$$\mu^a_{(F,E)}(x) \le \mu^a_{(G,E)}(x) \text{ and } v^a_{(F,E)}(x) \ge v^a_{(G,E)}(x) \tag{4}$$

Definition 5 (F, E) is equal to (G, E), that is $(F, E) = (G, E)$ if $\forall x \in U$,

$$\mu^a_{(F,E)}(x) = \mu^a_{(G,E)}(x) \text{ and } v^a_{(F,E)}(x) = v^a_{(G,E)}(x) \tag{5}$$

Definition 6 $\forall x \in U$ and $\forall e \in E$, $(F, E)^c$ denotes the complement of (F, E) and is given by

$$\mu^a_{(F,E)^c}(x) = v^a_{(F,E)}(x) \text{ and } v^a_{(F,E)^c}(x) = \mu^a_{(F,E)}(x) \tag{6}$$

Definition 7 (F, E) is said to be the absolute IFSS if $F(e) = U, \forall e \in E$. So, we have

$$\mu^e_{(F,E)}(x) = 1 \text{ and } v^e_{(F,E)}(x) = 0 \tag{7}$$

Definition 8 (F, E) is said to be the null IFSS if $F(e) = \phi, \forall e \in E$; that is

$$\mu^e_{(F,E)}(x) = 0 \text{ and } v^e_{(F,E)}(x) = 1 \tag{8}$$

Definition 9 For any two IFSSs (F, E) and (G, E), their intersection is IFSS (H, E) and $\forall a \in E, \forall x \in U$, we have

$$\begin{aligned} \mu^a_{(H,E)}(x) &= \min\left\{\mu^a_{(F,E)}(x), \mu^a_{(G,E)}(x)\right\} \text{ and} \\ v^a_{(H,E)}(x) &= \max\left\{v^a_{(F,E)}(x), v^a_{(G,E)}(x)\right\} \end{aligned} \tag{9}$$

Definition 10 For any two IFSSs (F, E) and (G, E), their union is IFSS (H, E) and $\forall a \in E, \forall x \in U$, we have

$$
\begin{aligned}
\mu^a_{(H,E)}(x) &= \max\left\{\mu^a_{(F,E)}(x), \mu^a_{(G,E)}(x)\right\} \text{ and} \\
v^a_{(H,E)}(x) &= \min\left\{v^a_{(F,E)}(x), v^a_{(G,E)}(x)\right\}
\end{aligned}
\tag{10}
$$

4 Application of IFSS in Decision-Making

Several applications of soft sets are discussed in [2]. A new approach for decision-making problem is proposed in this paper.

Suppose a person wants to identify the best house to buy, then the person has to compare different parameter values for each house. The parameters can be beautiful, furnished, expensive, distance, green surroundings, etc. If there is one or more parameters like "Expensive" or "Distance", the values of those parameters affect the decision inversely. We call these parameters as negative parameters. So, the parameters can be categorized as two types. (i) If the value of the parameter is directly proportional to the interest of the customer then we say that is a positive parameter. (ii) If the value of the parameter is inversely proportional to the interest of the customer then we say that is a negative parameter. For example "Beautiful" is a positive parameter. If the value of parameter "Beautiful" increases then the customer's interest will also increase. Whereas 'Expensive' is a negative parameter. Here, if the value of the parameter 'Expensive' increases then the interest of customer will decrease.

We prioritize the parameters by multiplying with priority values given by the customer. The customer has to give the priorities for each parameter. The priority is a real number in $[-1, 1]$. When a parameter value does not affect the customer's decision then the priority will be 0 (zero). If a parameter value affects positively to customer's decision then the priority will be $(0, 1]$ and if a parameter value affects negatively to customer's decision then the priority will be $[-1, 0)$. If priority value is not given for one or more parameters then the value of the priority is assumed to be 0 by default and that parameter can be eliminated from further computation. To get even more reduction in computation we can keep only one object if there are some objects with same values for all parameters.

The customer's priority value will be multiplied by the parameter values to obtain the priority table. In this paper, three different Tables (2,3 and 4) are computed separately for μ, v and π values. The row sums are calculated for each row of every table.

The comparison tables for membership, non-membership and hesitation values can be obtained by taking the difference of row sum of an object with others in respective priority table.

In decision table the score of each object can be obtained by calculating row sum in all comparison tables. The object having highest score will be more suitable to customer's requirement and subsequent values are the next choices. If more than one object is having the same highest value in the score column then the object having higher value under the highest priority column is selected and will continue like this.

$$\text{Score} = (\text{Membership Score} - \text{Non-Membership Score} \\ + \text{Hesitation Score} + 1)/2 \tag{11}$$

Equation (11) reduces to only membership score in case of a FSS.

4.1 Algorithm

1. Input the IFSS (*F, E*) in tabular form.
2. Input the priority given by the customer for every parameter. If priority for any parameter has not given then take it as 0 and opt out from further computations. For negative parameters the priority value must have to lie in the interval $(-1, 0)$
3. Multiply the priority values with the corresponding parameter values to get the priority tables for membership, non-membership and hesitation values.
4. Compute the row sum of each row in all of the priority tables.
5. Construct the respective comparison tables by finding the entries as differences of each row sum in priority tables with those of all other rows.
6. Compute the row sum for each row in all of the comparison tables to get the membership, non-membership and hesitation values for decision table.
7. The decision table can be constructed by taking the row sum values in comparison tables and compute the score using (11).
8. The object having highest value in the score column is to be selected. If more than one object is having the same highest value in the score column then the object having higher value under the highest rank priority column is selected and will continue like this.

Let the set of houses be given by $U = \{h_1, h_2, …, h_6\}$ and $E = \{$Beautiful, Wooden, Green Surrounded, Expensive, Distance$\}$. Consider an IFSS (*F, E*) as given in Table 1.

Suppose a customer wants to buy a house out of given houses in U which suits his needs on the basis of values parameters given in Table 1. That means out of all available houses, he needs to select a house according to his priorities.

The priority given by the customer for all parameters, respectively, are 0.7, 0.0, 0.2, -0.5, -0.2. The priority table can be obtained by multiplying the values in Table 1 with respective parameter priority values given by the user. This way Tables 2, 3 and 4 are obtained for membership, non-membership and hesitation

Table 1 Tabular representation of the lFSS (F, E)

U	Beautiful		Wooden		Green surrounded		Expensive		Distance	
h_1	0.10	0.70	0.00	0.90	0.20	0.70	0.10	0.80	0.80	0.20
h_2	0.90	0.00	0.60	0.40	0.80	0.20	0.80	0.10	0.30	0.60
h_3	0.30	0.50	0.10	0.80	0.20	0.70	0.30	0.40	0.40	0.60
h_4	0.70	0.10	0.70	0.20	0.60	0.40	0.60	0.20	0.60	0.30
h_5	0.30	0.40	0.40	0.50	0.40	0.50	0.50	0.20	0.10	0.90
h_6	0.90	0.10	0.50	0.40	0.60	0.30	0.60	0.20	0.50	0.40

Table 2 Priority table for membership values

U	Beautiful	Wooden	Green surrounded	Expensive	Distance	Row sum
h_1	0.07	0.0	0.04	−0.05	−0.16	−0.10
h_2	0.63	0.0	0.16	−0.40	−0.06	0.33
h_3	0.21	0.0	0.04	−0.15	−0.08	0.02
h_4	0.49	0.0	0.12	−0.30	−0.12	0.19
h_5	0.21	0.0	0.08	−0.25	−0.02	0.02
h_6	0.63	0.0	0.12	−0.30	−0.10	0.35

Table 3 Priority table for non-membership values

U	Beautiful	Wooden	Green surrounded	Expensive	Distance	Row sum
h_1	0.49	0.0	0.14	−0.40	−0.04	0.19
h_2	0.00	0.0	0.04	−0.05	−0.12	−0.13
h_3	0.35	0.0	0.14	−0.20	−0.12	0.17
h_4	0.07	0.0	0.08	−0.10	−0.06	−0.01
h_5	0.28	0.0	0.10	−0.10	−0.18	0.10
h_6	0.07	0.0	0.06	−0.10	−0.08	−0.05

Table 4 Priority table for hesitation values

U	Beautiful	Wooden	Green surrounded	Expensive	Distance	Row sum
h_1	0.14	0.0	0.02	−0.05	0	0.11
h_2	0.07	0.0	0	−0.05	−0.02	0.00
h_3	0.14	0.0	0.02	−0.15	0	0.01
h_4	0.14	0.0	0	−0.1	−0.02	0.02
h_5	0.21	0.0	0.02	−0.15	0	0.08
h_6	0	0.0	0.02	−0.1	−0.02	−0.1

Table 5 Comparison table for membership values

h_j h_i	h_1	h_2	h_3	h_4	h_5	h_6	Row sum
h_1	0.00	−0.43	−0.12	−0.29	−0.12	−0.45	−1.41
h_2	0.43	0.00	0.31	0.14	0.31	−0.02	1.17
h_3	0.12	−0.31	0.00	−0.17	0.00	−0.33	−0.69
h_4	0.29	−0.14	0.17	0.00	0.17	−0.16	0.33
h_5	0.12	−0.31	0.00	−0.17	0.00	−0.33	−0.69
h_6	0.45	0.02	0.33	0.16	0.33	0.00	1.29

Table 6 Comparison table for non-membership values

h_j h_i	h_1	h_2	h_3	h_4	h_5	h_6	Row sum
h_1	0.00	0.32	0.02	0.20	0.09	0.24	0.87
h_2	−0.32	0.00	−0.30	−0.12	−0.23	−0.08	−1.05
h_3	−0.02	0.30	0.00	0.18	0.07	0.22	0.75
h_4	−0.20	0.12	−0.18	0.00	−0.11	0.04	−0.33
h_5	−0.09	0.23	−0.07	0.11	0.00	0.15	0.33
h_6	−0.24	0.08	−0.22	−0.04	−0.15	0.00	−0.57

Table 7 Comparison table for hesitation values

h_j h_i	h_1	h_2	h_3	h_4	h_5	h_6	Row sum
h_1	0.00	0.11	0.10	0.09	0.03	0.21	0.54
h_2	−0.11	0.00	−0.01	−0.02	−0.08	0.10	−0.12
h_3	−0.10	0.01	0.00	−0.01	−0.07	0.11	−0.06
h_4	−0.09	0.02	0.01	0.00	−0.06	0.12	0.00
h_5	−0.03	0.08	0.07	0.06	0.00	0.18	0.36
h_6	−0.21	−0.10	−0.11	−0.12	−0.18	0.00	−0.72

values respectively. Note that, two parameters, 'Expensive' and 'Distance' having negative priorities, show negative parameters.

Priority tables can be obtained by multiplying the priority values with the original values. Row sums in each table are also computed.

The respective comparison tables are constructed by finding the entries as differences of each row sum in priority tables with those of all other rows and compute row sum in each of Tables (5, 6 and 7).

By using the formula (11), the decision table can be formulated (Table 8).

Table 8 Decision table

Houses	Membership score	Non-membership score	Hesitation score	Score
h_1	−1.41	0.87	0.54	−0.37
h_2	1.17	−1.05	−0.12	1.55
h_3	−0.69	0.75	−0.06	−0.25
h_4	0.33	−0.33	0.00	0.83
h_5	−0.69	0.33	0.36	0.17
h_6	1.29	−0.57	−0.72	1.07

Decision-Making The Customer should go for the house h_2 which has highest score. If for some reason, the customer does not want that house then subsequent scored houses can be considered.

5   Conclusion

In this paper we introduced the membership function for IFSS which extends the notion of characteristic function for FSS introduced by Tripathy et al. in [7]. With this new definition we redefined many concepts associated with IFSSs and established some of their properties. The notion of IFSS introduced in [3] and their application to decision-making had many flaws. In this paper we pointed out these flaws and provided solutions to rectify them, so that the decision-making becomes more efficient and realistic.

References

1. Zadeh L.A.: "Fuzzy sets", Information and Control, 8, (1965), pp. 338–353.
2. Molodtsov D.: "Soft Set Theory—First Results", Computers and Mathematics with Applications, 37, (1999), pp. 19–31.
3. Maji P.K, Biswas R and Roy A.R: An Application of Soft Sets in a Decision Making Problem", *Computers and Mathematics with Applications*, 44, (2002), pp. 1007–1083.
4. Maji P.K., Biswas R. and Roy A.R.: "Soft Set Theory", Computers and Mathematics with Applications, 45, (2003), pp. 555–562.
5. Maji P.K., Biswas R. and Roy A.R.: "Fuzzy Soft Sets", *Journal of Fuzzy Mathematics*, 9(3), (2001), pp. 589–602.
6. Tripathy B.K., Arun K.R, "A New Approach to Soft Sets, Soft Multisets and Their Properties", International Journal of Reasoning-based Intelligent Systems, Vol. 7, no. 3/4, 2015, pp. 244–253.
7. Tripathy B.K, Sooraj.T.R, Mohanty, "A New Approach to Fuzzy soft set theory and its applications in decision making" accepted for presentation at ICCIDM2015.
8. Atanassov K.: "Intuitionistic Fuzzy Sets", *Fuzzy Set Systems*, Vol. 20, (1986), pp. 87–96.

Categorical Data Clustering Based on Cluster Ensemble Process

D. Veeraiah and D. Vasumathi

Abstract In spite of the fact that endeavors have been made to take care of the issue of clustering straight out information by means of group gatherings, with the outcomes being focused on customary calculations, it is observed that these procedures sadly create a last information parcel taking into account deficient data. The fundamental gathering data network exhibits just group information point relations, with numerous passages left obscure. Downright Data clustering and Cluster ensemble approach have been related and partitioned with examination in application areas with respect to related data. The fundamental aim of this paper is to examine and share information between these two data points and use this shared information for making novel clustering calculations for absolute information in light of the cross-preparation between the two subsequent item sets with exploratory analysis. All the more decisively, we normally characterize the Categorical Data Clustering (CDC) issue with improvement issue from the perspective of CE, and calculate with a CE approach for grouping clear-cut information.

Keywords Cross-fertilization · Categorical information · Group outfits · Connection-based closeness · Information mining

1 Introduction

Clustering is the procedure of making a gathering of conceptual items into classes of comparable articles. A cluster of information items can be dealt with as one gathering. While doing group examination, we first segment the arrangement of information into gatherings in view of information closeness and later dole out the
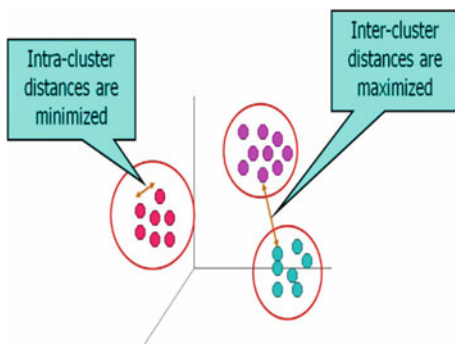
D. Veeraiah (✉)
Jawaharlal Nehru Technological University Kakinada, Kakinada, Andhra Pradesh, India
e-mail: veeraiahdvc@gmail.com

D. Vasumathi
JNTU College of Engineering Hyderabad, Hyderabad, Telangana, India
e-mail: rochan44@gmail.com

Fig. 1 Data clustering group connections which iterate inter- and intra-cluster formation

names to the gatherings. Data grouping is one of the essential arrangements in real-time applications we have in comprehending the structure of an information sheet. It plays a pivotal, basic part in hardware identification with machine learning, information mining, data recovery, and example acknowledgment (Fig. 1).

Clustering means to arrange information into gatherings or cluster such that the information in the same group is more like one another than to those in distinctive groups. Clustering is a helpful system for gathering information focuses such that focuses inside of a solitary gathering/cluster have comparable qualities (or are near to one another), while focuses on distinctive gatherings are unique. For example, consider a business sector created database containing one exchange per client, where every exchange containing the arrangement of things is acquired by the client. The groups can then be utilized to portray the distinctive client clusters, and these portrayals can be utilized as a part of the focus on showcasing and publicizing such that particular items are coordinated toward particular client clusters. The portrayals can likewise be utilized to anticipate purchasing examples of new clients in light of their profiles. For instance, it might be conceivable to presume that high-salary clients purchase imported sustenance, and later mail altered inventories for imported nourishments to just these high-paying clients (Fig. 2).

Fig. 2 Link clusters using data structure for processing similarity metrics
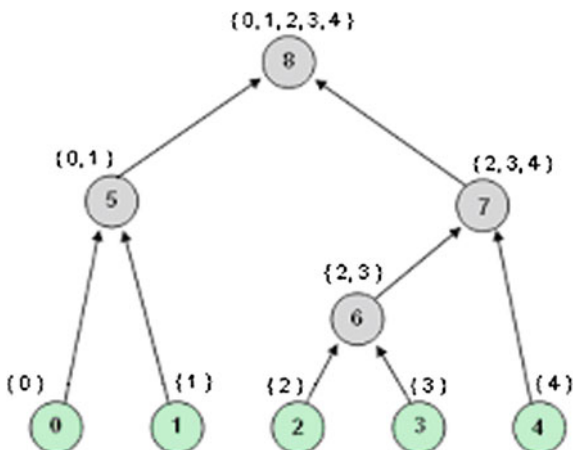
Most past clustering calculations focus on efficient numerical information whose scientific and geometric properties can be misused normally to characterize separation capacities between information focuses. Nonetheless, a great part of the information existed in the databases is clear cut, where property estimations cannot be normally requested as numerical qualities. A case of clear-cut property is shape whose qualities incorporate circle, rectangle, oval, and so forth. Because of the exceptional properties of all out characteristics, the clustering of unmitigated information appears to be more interesting than that of numerical information. A couple of calculations have been introduced lately for grouping clear-cut information.

Cluster ensemble (CE) is the system to consolidate a few keeps running in diverse clustering calculations to produce a typical piece of the first elected dataset, going for a combination of results from an arrangement of commit of individual results. In spite of the fact that the examination on cluster troupe has not been generally perceived as that brushing different classifier or relapse models, all the more as of late, a few exploration endeavors have been made freely.

Recently, CDC and CE have long been regarded as distinct evaluation and program areas. The starting level in this document is the understanding of some key invisible similarities between these two exclusive areas. This understanding makes possible the research of CDC issue from a CE perspective. The CDC computation for managing team details unreliable is with two example responsibilities. That is, our first dedication is the research on invisible qualities, similarities, and differences in the center of CDC and CE, which creates the assumption for the undertaking of CE centered clustering computations for overall details. All the more decisively, despite the fact that CE is a general structure with numerous applications and CDC is a unique case in clustering examination, from a limited perspective, these two issues are fundamentally proportionate.

Our second dedication is the immediate modification and usage of CE viewpoint for clustering overall details. We officially define the CDC problem as an optimizing problem from the viewpoint of CE, and apply a CE strategy for collection of unmitigated details. Our test results demonstrate the new absolute information clustering techniques to accomplish preferable grouping exactness over past calculations, which affirms our instinct that CE methodologies and CDC strategies can be utilized conversely.

The remainder of this chapter is organized as follows: Sect. 2 introduces a discriminating audit on related work. Section 3 is a fascinating view on the hidden properties, likenesses, and contrasts in the middle of CDC and CE. In Sect. 4, we characterize the CDC issue as an improvement issue and portray the CE-based calculations for clustering all our information. Trial results are given in Sects. 5 and 6 concludes the paper.

2 Background Approach

In this segment we display the gathering accumulation structure whereupon the present examination has been perceived [1]. The recommended connection-based methodology, for example, the genuine intuition of refining a troupe data network and points of interest of a connection-based similarity measure.

2.1 Issue Formation and Common Framework

Let $X = (x1, \ldots, x N)$ be an arrangement of N points of interest variables and $\pi = (\pi1; \pi2, \pi3, \pi4, \ldots \pi n;)$ Mg be a gathering accumulation with M stage clusterings, each of which is for the most part known as a gathering member. Every base grouping benefits an arrangement of gatherings $\pi i = \{C_1^n, C_2^n, C_3^n, \ldots C_{k_i}^n\}$, such that $\cup_{j=1}^{k_i} C_j^i = X$, where ki is the mixed bag of gatherings in the required clustering. For every $x 2 X$, $C(x)$ means the gathering brand to which the subtle elements variable x joined. In the ith grouping, $C(x) ='' j''(or'' C_j^{i''})$ if $x \in C_j^i$. The issue is to find another parcel $\pi*$ of a point of interest set X that comprises the subtle elements of the cluster gathering π [2, 3].

Fundamentally, arrangements obtained from the diverse base clustering are accumulated to frame any parcel. This met level system incorporates two noteworthy ventures of: (1) making gathering accumulation, and (2) delivering a definitive segment, regularly alluded to as an understanding capacity (Fig. 3).
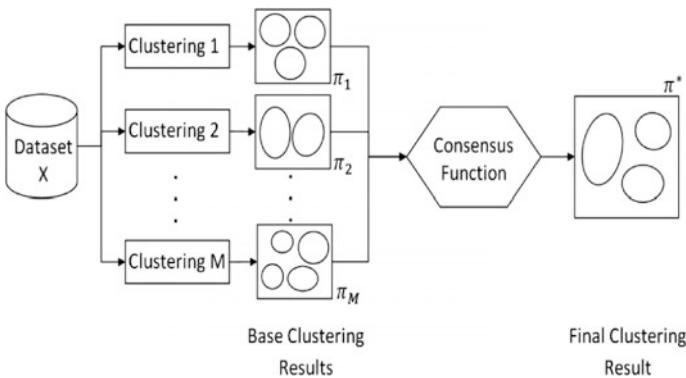
Fig. 3 The essential methodology of gathering outfits. It first is material various base groupings to a dataset X to secure inverse clustering choices. At that point, these choices are blended to situated up the last clustering result $\pi*$ utilizing an assigned word

2.2 Collection Creation Methods

It has been confirmed that clothing is most effective when designed from an agreement of signs whose mistakes are different [4, 1]. Especially for information collection, the outcomes acquired with any individual requirements over several versions are usually very much as well. In such a circumstance where all accumulation individuals concede to how a dataset ought to be divided, conglomerating the stage grouping results will show no upgrade over any of the part relates. Thus, a few heuristics have been proposed to present engineered dangers in grouping routines, giving an assortment inside of a group outfit. A portion of the continuous capacities was utilized for absolute information grouping detail.

2.3 Agreement Functions

Having obtained the group troupe, a variety of conforming capabilities have been created and made available for identifying a specified information package. Every accord capacity uses a particular type of data grid, which condenses the base grouping results. In regard to these capabilities, agreement strategies can be named takes over the agreement functions.

Highlight-based procedure. It changes over the issue of group outfits to grouping specific data. In particular, every stage grouping gives a cluster name as another capacity clarifying every data variable.

2.4 Group Outfits of Particular Data

While an enormous variety of collecting buildup workouts for mathematical simple elements have been placed ahead in past several years, there are just a couple examine that perform such a viewpoint to specific neat places to see clustering. The last clustering result has been delivered utilizing the diagram-based accord strategies.

Specific to this alleged "direct" accumulation, creation strategy, a given straight out information sheet can be indicated utilizing a parallel cluster affiliation framework. Such a point of interest framework is practically identical to the "business wicker container" numerical impression of specific subtle elements, which has been the concentrate on routine specific subtle elements examples.

3 Cluster Ensemble-Based Approach

In this area, we obtain that considered by the selection group to formalize the CDC problem as an improvement problem regarding allocated common information and signify those CE centered calculations for selection all our information.

3.1 Object Function for CE

Consider the dataset $X = \{x1, x2 \ldots xn\}$ is a contract of material shown by being straight out features, $A1, \ldots Ar$, with locations $D1, \ldots, Dr$ individually [5]. The value set V_i is a situated of reports of Ai that exist in X. As defined in Area 3.2, on the off chance that we determine each clustered (i) as a prospective that applies concepts in VI to particular common numbers, we can get the best apportioning (i) identified by every top quality Ai. Thus the last clustering generate can be considered as the team collecting outcome by combining the categories given by (i). Intuitively, an excellent combined collection ought to discuss, however much details as could reasonably be predicted with the given r labeling. Strehl and Ghosh implement the typical details in detailed speculation to evaluate the typical details, which can be particularly linked in this composing.

All the more compactly, as indicated in Strehle's documents, given 'r' categories with the qth gathering $r(q)$ having $k(q)$ categories, a conform potential is recognized as a potential $N\ n \times r\ N$ mapping an agreement of clusterings to an integrated clustering:

$$\tau : \{\beta^q | q \in \{1, 2, \ldots n\}\} \rightarrow \beta \tag{1}$$

The agreement of categories is signified as $\wedge = \{\lambda^q | q \in \{1, 2, \ldots n\}\}$, the perfect signed up with grouping ought to provide the most details to the first cluster.

3.2 Group Collection-Based Algorithm

In this way, there are a few calculations for selection outfits. The strategy structured in past techniques developed for mixing keeps working on selection calculations with the same wide range of packages. In this way, it is not appropriate in our composing, as the quality of bundles determined by different absolute feature can be unique.

In the case that two products are in the same collection they are believed to be absolutely relative, and if not they are different. This is the most uncomplicated heuristic and is used as an aspect of the Cluster-based Likeness Dividing Criteria (CSPA). With this viewpoint, one can usually determine a single collection into a dual similitude lattice. Assessment between the two products is 1 on the off opportunity that they are in the same team and 0 usually. For every collection, a combined similitude $n \times n$ line is created.

The area smart regular of r such lattices discussing to the r places of categories generates a bye and huge comparability lines. In that factor, the METIS computation is used to area the similitude graph (vertex = content, advantage bodyweight = closeness) to get the last categories. Every team is verbal to a hyperedge with the same loads, the detailed products are considered as vertices with the same

loads. At that factor, a hyper graph splitting computation, HMETIS is used to package the hyper graph such that the whole of loads hyper edge cut is reduced. As in HGPA, every collection is verbal to a hyperedge. The believed in MCLA is to collect furthermore, crumple relevant hyperedges and relegate every product to the caved in hyperedge in which it takes attention in it most strongly. The hyperedges that are regarded as relevant with the end objective of caving in are determined by a graph centered collection of hyperedges. Every variety of hyperedges is referred as meta-groups. For the proper procedure in relevant data grouping, consider the procedure in the above paragraphs where CE is the best effective data clustering for grouping individual item selection based on attributes of uploaded data.

4 Performance Evaluation

A thorough efficiency research has been done to assess our technique. In this section, we signify those assessments and outcomes. We ran our calculations on authentic datasets acquired from the UCI System Learning Data source to analyze its clustering efficiency against different calculations.

4.1 Genuine Datasets and Evaluation Technique

We tried different factors with four authentic datasets: the Congressional Ballots dataset, the Wi Bosom Cancer dataset, the Mushroom dataset, and the Zoo dataset, which were obtained from the UCI System Learning Data source. Here we provide a brief business presentation of these datasets.

Congressional Votes It is the United Declares Congressional Voting Details 26 years ago. Each record talks to one Congressman's vote on 16 issues. All features are Boolean with Yes (signified as y) and No (indicated as a) features. Despite indicators of Republications or Democratic properties given using their history, the dataset contains 435 records with 168 Conservatives and 267 Democrats.

WI Breast Cancer Data 2 It has 699 situations with 9 functions. Every record is noticeable as kindhearted (458 or 65.5 %) or risky (241 or 34.5 %). In our composing, all functions are regarded as directly out with functions 1, 2, …, 10. The Mushroom Data Set has 22 functions and 8124 information [6]. Every record speaks of the actual functions of an individual mushroom. A depiction name of dangerous or usable is furnished with every history. The amounts of usable and dangerous weeds in the dataset re 4208 and 3916, independently.

Zoo information includes 101 cases of creatures with 17 elements and 7 produce classes. The name of the creature includes the first property. There are 15 Boolean elements associated with the area of hair, quills, egg, milk, central source, rotor blades, tail; and whether popular, ocean, predator, toothed, inhales in, venomous,

personal, cat size. The character top high quality analyzes to the high top high quality of feet, soothing in the set {0, 2, 4, 5, 6, 8}. Accepting collection results is a non-minor challenge. In the area of authentic titles, as in the situation of the detail sets we used, the collection perfection for calculating the clustering results was realized as takes after. Given the last variety of categories, *k*, collection perfection '*r*' is recognized as: $r = \frac{\sum_{i=1}^{k} ai}{n}$ where n is the high top quality of records in the data set, *ai* is the high top quality of situations occurring in both team I and its evaluating category, which had the maximum top quality. As such, *a* is the number of records with the category indicates that instructions team *I*. Consequently, the clustering slip up is recognized as *e = 1*.

5   Experimental Results

Up to this factor, there is not much recognized conventional way of Categorical Data Clustering (CDC) assessments [7]. However, we viewed that most collection computations assist the quality of the categories as a detail parameter, so in our assessments, our collection of each dataset into a unique variety of bundles moves from 2 to 9. For each changed variety of bundles, the collection drops of different computations were believed about.

In all the tests, apart from the high quality of packages, all the aspects needed by the Cluster Ensemble selection calculations are situated to be default 3. The Squeezer calculations need just a likeness limit as detail parameter, so we set this parameter to an authentic value to get the craved wide range of categories (For the Squeezer calculations, if the produce of wide range of packages is the same, the gathering excellence verges on the indistinguishable. Thus, we can apply any similitude limit assurance that can make the calculations get the craved wide range of groups). For the GA Cluster calculations, we set the inhabitants sizing to be 50, and set different aspects of their traditional principles.

5.1   Clustering Results of Congressional Voting (Votes) Data

Figure 4 shows effective voting data allots on the ballots dataset of different clustering computations. From Fig. 4 we can abridge the comparative performance of these computations as required after Table 1 [7, 5]. Contrasting with the Squeezer computation and the GA Cluster computation, Cluster Ensemble calculation is conducted best in four situations and second best in four situations. It is never conducted most extremely terribly. Furthermore, the regular collection mistake of the Cluster Ensemble computation was reasonably lesser than that of different computations.

Fig. 4 Cluster errors with number of clusters
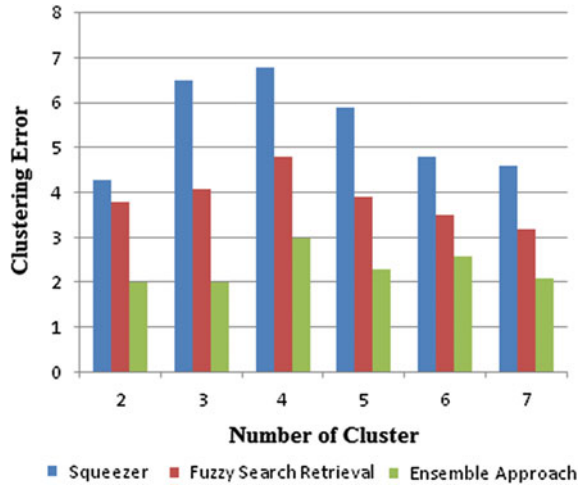
Table 1 Relative performance of different clustering algorithms

| Ranking	1	2	3	Clustering error
Squeezer	2	1	5	0.163
Fuzzy search(FS)	3	2	3	0.136
Cluster ensemble	4	4	0	0.115

Since in the incorporated CE strategy, by Collection, we first run Squeezer, Fuzzy Search, and MCLA independently, and selecting the one with the best ANMI as the last result, in this dataset, it is viewed that Fuzzy Search (FS) has the best ANMI for six periods, and MCLA has the best ANMI for two periods. Consequently, the exposed effects of Cluster Ensemble collection data puzzled with Fuzzy Search gives effective outcome results compared to Squeezer and Fuzzy Search methods. Let us see more empirical studies of CE discussed in the following sections with suitable datasets and item sets.

5.2 Clustering Outcomes of Cancer Data

The evaluate results on the growth dataset are represented in Fig. 5 and the explanation of the comparative performance of the three computations is given in Desk 2 [2]. From Fig. 5 and Table 1, despite the factor that the regular clustering perfection of our computation is just a bit excellent for anything that of the Squeezer and GA Cluster computation, while the circumstances of our computation that defeat the other two calculations are prevalent in this research.

In this dataset, CE has the best ANMI for all circumstances while the clustering outcomes of fuzzy and Squeezer selection absolutely. From this assessment and

Fig. 5 Comparative time
Efficiency with previous
techniques and our proposed
approach in time
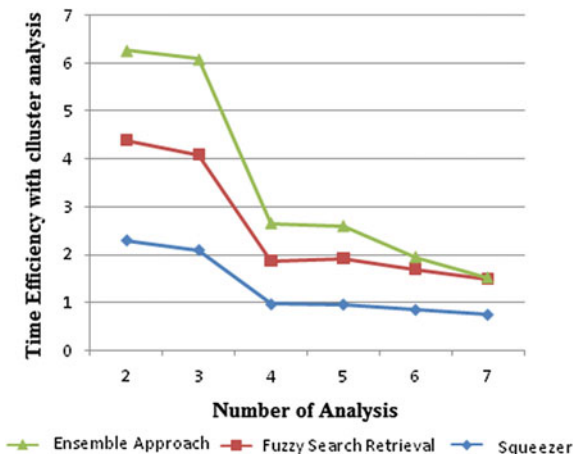
outcomes, it is apparent that the gathering produce of Cluster Ensemble is usually given the best results compared with previous techniques. That is, Cluster Ensemble can defeat the Squeezer and FS Clustering is generally because of the adequacy of CE. We can see that the performance of Cluster Ensemble collection, computation on the zoo details set is not tasty compared with two other computations. It reveals that Cluster Ensemble gives the best effective performance in out coming results in terms of time in overall datasets present in datasets. In any scenario, it ought to be noticed that the clustering efficiency of Cluster Ensemble is near to that of the other two calculations. That is, even in a dataset with unbalanced classification submission, our calculations can accomplish similar efficiency.

6   Conclusion

Categorical data clustering is the main focusing term data analysis in clustering. This paper presents an effective cluster ensemble approach categorical data clustering. CE is a common detail recycling framework with several programs, and CDC is an exclusive case in collection similar weight of categorical items. It changes the first absolute information network to a data saving numerical variety (RM), to which a successful diagram parceling strategy can be specifically connected. The experimental study, with distinctive troupe sorts, legitimacy measures, and information sets, recommends that the proposed Cluster outfit approach for the most part accomplishes better grouping results looked at than those of the conventional absolute information calculations and benchmark bunch gathering strategies. For upcoming perform, we seek to summarize Fuzzy c-means like clustering computations for all out details that directly enhance the common information discussing-based product perform.

References

1. M. Dutta, A.K. Mahanta, and A.K. Pujari," QROCK: A Fast Addition of the ROCK Criteria for Clustering of Express Information in Design Identification Characters, vol. 26, pp. 2364–2373, 2005.
2. Y. Zhang, A. W. Fu, C. H. CAI, P. A. Heng. "Clustering particular data" In: Proc of ICDE'00, pp. 305–305, 2000.
3. E. Minkov, W.W. Cohen, and A.Y. Ng "Contextual Search and Name Disambiguation in E-mail Using Charts", in Proc. Int'l Conf. Research and GrowthiIR,pp. 27342006.
4. P. Reuther and B. Wally, in Int'l J. "Survey on Analyze Selections and Methods for Individual Name Related," by metadata, Semantics and Ontologies, vol. 1, no. 2, pp. 89–99, 2006.
5. A. Strehl, J. Ghosh "Cluster outfits - an understanding recycling structure for mixing several partitions". Publication of Device Studying Research, 2002, 3:583–617.
6. F. Fuss, A. Pirotte, J.M. Provides, and M.Saerens"Random-Walk Calculations of Resemblances between Nodes of a Chart with Program to Collaborative Suggestions," Proceedings in IEEE Trans. Information and Information Eng., vol. 19, no. 3, pp. 355–369, Mar. 2007.
7. P-E. Jouve, N. Nicoloyannis " A new means for mixing categories, programs for group outfits in KDD". In: Similar and Allocated processing for Device Studying Workshop, combined with ECML'03 and PKDD'03, pp. 35–46, 2003.
8. V. Castelli, A. Thomasine, C-S Li. CSVD: Grouping and unique value breaking down for estimating likeness look for in high-dimensional spaces" buy in IEEE Dealings on Information and Information Technological innovation, 2003, 15(3): 671–685.
9. T. Boongoen, Q. Shen, and C. Cost "Disclosing Incorrect Identification through Multiple Web link Research," in Synthetic Intellect and Law, vol. 18, no. 1, pp. 77–102, 2010.
10. E. Abdu and D. Salane, "A Spectral-Based Clustering Criteria for Express Information Using Information Summaries," Proc. Workshops Information Exploration using Matrices and Tensors, pp. 1–8, 2009.

Region-Based Clustering Approach for Energy Efficient Wireless Sensor Networks

Kalyani Wankhede and Sumedha Sirsikar

Abstract Nowadays there is a huge increase in the use of sensors in various applications such as remote monitoring of environment, automobiles, disaster prone zones, home control, and military applications. The capabilities of Wireless Sensor Network (WSN) can be extended using self-organization to change their behavior dynamically and achieve network wide characteristics. Clustering techniques show the self-organized behavior of WSNs. Sensor nodes are grouped into disjoint, nonoverlapping subsets called clusters. Cluster Heads (CHs) collect data from the sensor nodes present in the cluster and forward it to the neighboring nodes using shortest path distance calculation and finally to the Base Station (BS). In the proposed clustering technique, network area is divided into regions. Cluster Head (CH) is elected by using the highest residual energy and the node degree. Cluster communication is at the most two-hop which results in less number of messages from member nodes to BS. Within a cluster, data is sent to CH by member nodes. After some time CH's energy level goes below threshold value, then new CH is elected and clusters are reformed within each region. Our proposed clustering technique can be used in military applications to detect and gain information about enemy movements. Results are obtained by varying the number of nodes at different transmission ranges. The simulation results show that the proposed clustering algorithm reduces energy consumption, prolongs network lifetime, and achieves scalability.

Keywords Wireless sensor networks · Clustering · Energy efficiency · Scalability · Self-organization

Kalyani Wankhede (✉) · Sumedha Sirsikar
Department of Information Technology, Maharashtra Institute
of Technology, Pune, India
e-mail: er.wankhedekalyani@gmail.com

Sumedha Sirsikar
e-mail: sumedha.sirsikar@mitpune.edu.in

1 Introduction

Wireless sensor network is a large-scale network of small embedded devices. Each device possesses sensing, computing, and communicating capabilities. Sensor nodes are constrained in terms of power, communication bandwidth, and energy, and storage space. Thus WSN requires very efficient resource utilization. Current advances in reduction and low-power design have led to the improvement of small-sized sensors that are capable of detecting ambient conditions such as temperature, pressure, humidity, moisture, and sound. Sensors are generally equipped with data processing, storage, and communication capabilities [1].

This technological development has encouraged researchers to visualize aggregating the limited capabilities of the individual sensors in a large-scale network that can operate unattended. WSNs have both civilian and military applications which include scene reconstruction, motion tracking and detection, battlefield monitoring, and remote sensing. In disaster management situation such as earthquakes, sensor networks can be used to selectively map the affected regions directing crisis response units to survivors. In military circumstances, sensor networks can be used in surveillance and to detect movements of enemy, chemical gases, unknown troop, vehicle activity, or the presence of micro-agents. Information about enemy movement detection and vigilance for unknown troop is sensed by sensor nodes. Sensor nodes send data to CH node and then CH sends to BS. Sensed data reached BS by using clustering technique. BS is connected to command nodes so that they can communicate with each other. Hence, command node get information about enemy position, movements then takes appropriate decisions on it. Figure 1 shows the proposed system architecture for military application.
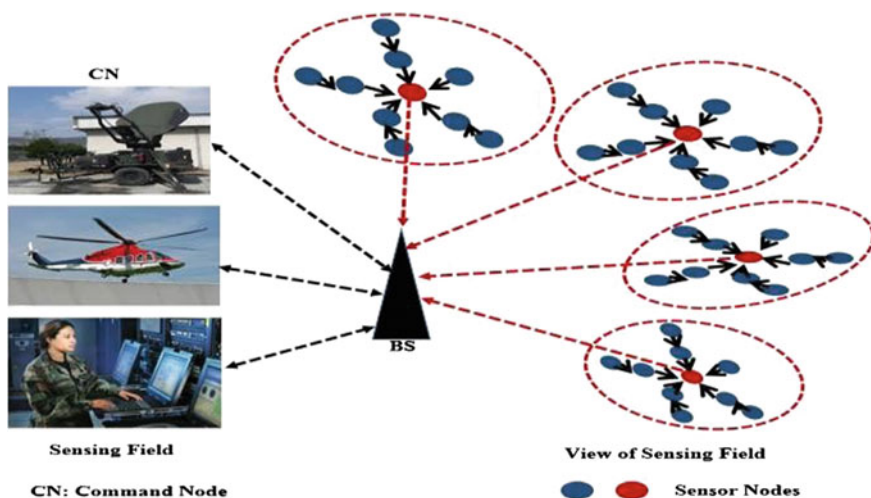

Fig. 1 Proposed system architecture

2   Related Work

LEACH [2, 3] is a randomized clustering algorithm which uses adaptive cluster formation. It uses two phases: setup phase and steady phase. In setup phase CH election and cluster formation take place and actual data is transferred in steady phase.

Divide-and Rule Scheme [4] uses static clustering for clustering technique and dynamic CH election. Two CHs are elected, namely primary level CH and secondary level CH. Primary CH sends data to secondary CH, then combined data is sent toBS.

In DDAR [5] protocol, the distance and energy of the node is used as a factor to select CH. Protocol operation has three phases, namely network setup, routing path construction, and schedule creation. Results show that DDAR protocol is more energy efficient than MTE, LEACH, and LEACH-C protocols.

Two-Hop Clustering (THC) [6] protocol employs two-hop clustering method, which reduces the amount of energy required by the member node to communicate with the respective CH nodes. All the two-hop away member nodes send their data to the CH nodes via one-hop nodes, which considerably reduces the amount of energy consumption.

Cluster of cluster heads approach [7] is explained by using three phases as initialization, setup, and steady. Setup phase contains Cluster formation and CH election phases. In this approach clustering technique is used which increases the lifetime of the network.

HEED is used to improve network lifetime by distributing energy consumption. It uses distributed methodology for clustering and probability approach for CH selection. Clusters are well structured at network area in HEED [8].

3   Proposed System

3.1   Assumptions

- The network area is fixed
- All sensor nodes are homogeneous in the network
- All sensor nodes are stationary
- BS is stationary and has information about location of all nodes
- The intra-cluster communication is two-hop
- Sensors compute the approximate distance to other sensors, based on received signal strength and transmitting power
- Symmetric Communication between any two sensors means the sensor node SNi can listen to the sensor node SN_j, and SN_j to SNi.

4 Approach

Energy is one of the major constraints of WSN. This problem is addressed in our work and a solution for limited energy source has been proposed. High energy efficiency in the network is maintained using cluster-based algorithms that adapt the size of the cluster by itself and reform the clusters. This global decision is taken by the network itself through local interactions among themselves. Hence, sensor network behaves as self-organized.

The proposed solution is to group the sensor nodes into various clusters that consist of five phases.

4.1 Division of Network into Fixed Regions

Whole network area is divided into 16 rectangular regions. In the network area node deployment is random. Nodes are deployed in such a way that all regions are covered. Deployment of nodes is shown in Fig. 2. Every region consists of one or more clusters. Each cluster has a cluster head. Among all the nodes present in that region CH has more residual energy and node degree.

4.2 Cluster Head Election and Advertisement of CH

The first step toward clustering is Cluster Head (CH) election. CH is elected with the collaboration of Base Station (BS), i.e., sink node. Network deployment is
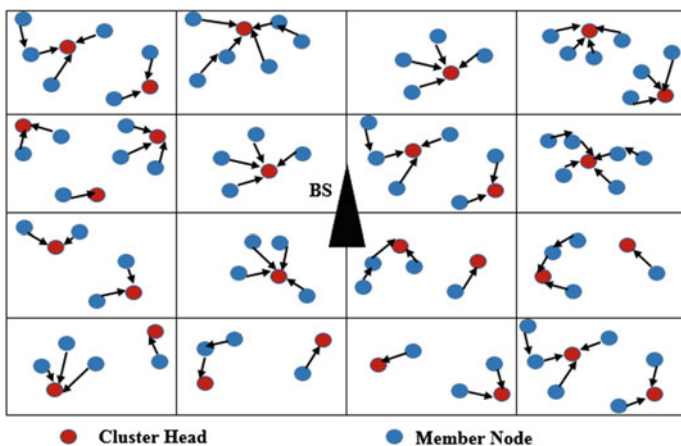
Fig. 2 Cluster formations in region

Fig. 3 Two-hop
communication Within cluster
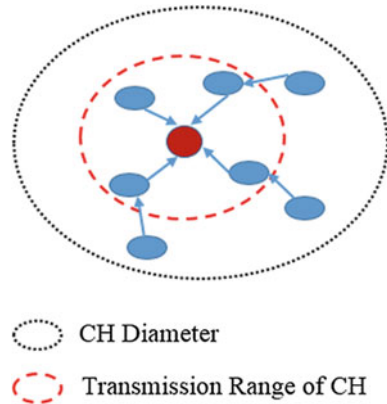
considered as manual, so the BS is well informed about the geographical location of the nodes. In the cluster head election phase all the deployed nodes send their energy level to the BS. CH is elected based on two parameters: highest residual energy and node degree. Then BS broadcasts this information to all nodes.

4.3 Cluster Formation

The clusters are formed on the basis of node distance. The distance between two nodes is calculated by Euclidian distance formula. Initially single-hop member is identified within the transmission range of the CH. Next hop nodes which are within cluster diameter are allowed to join the cluster. Thus next hop nodes are joining to single hop member nodes which are in the transmission range of CH. The nodes coming in the cluster diameter will join the single hop member nodes by two-hop communication as shown in Fig. 3.

4.4 Actual Data Transmission

In this phase, the CHs create a time slot, gather information from the member nodes, and finally transmit the aggregated data to the base station. CH generates time slot for each node when it has data to transmit. This produces data collision at the different nodes. Sensed data from all the member nodes during their time slot is collected by the CHs. The node which is two hop away sends its data to the one-hop neighbor. The one-hop neighbor updates its data to the received one and then transmits to the CH. The CH nodes aggregate the data received from the member nodes. This reduces the number of transmissions to the base station and also consumption of energy. Thus CH nodes send their data to the base station using multi-hop communication.

4.5 Reclustering

During the continuous operation of sensor network, the battery level of sensor
nodes as well as cluster heads decreases. Eventually sensor nodes and cluster heads
become inactive. If the node energy goes below the threshold value, the sensor
network is to be reorganized into new clusters. Meanwhile the node with highest
residual energy and degree will become the new cluster head.

5 Simulation and Results

The simulation of proposed clustering algorithm is carried out using NS2. The
parameters used in the simulation are shown in Table 1.

Figure 4 shows the cluster formation in the WSN based on residual energy and
node distance.

Our algorithm is applied for different number of nodes at different transmission
range. The average energy consumption for each simulation is calculated. Table 2

Table 1 Simulation parameters

Parameter	Value
Network area	3400 m × 1800 m
Channel type	Wireless
Number of nodes	40, 50, …, 120
Initial energy	1000 J
Threshold energy	100 J
Transmission range	250, 300, 350, 400 m
Transmitting and receiving power	1 W
Simulation period	100 S
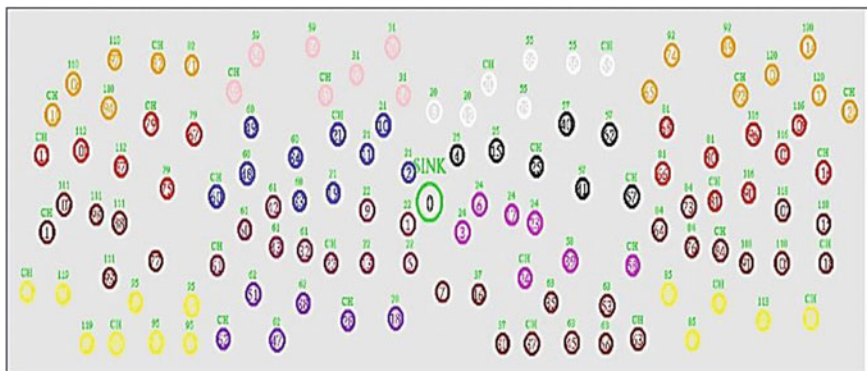
Fig. 4 Cluster formation

Table 2 Average energy consumed

Number of sensor nodes	Average energy consumed (J)			
	Range: 250 m	Range: 300 m	Range: 350 m	Range: 400 m
40	7.51	9.86	11.73	13.79
50	8.28	10.24	11.33	13.65
60	8.56	10.84	11.23	13.45
70	8.04	9.26	11.15	12.87
80	8.54	10.32	10.72	12.87
90	8.71	10.28	10.64	12.47
100	8.54	10.20	10.45	12.31
110	8.15	10.17	10.40	12.05

Table 3 Network lifetime

Number of sensor Nodes	Network lifetime (s)			
	Range: 250 m	Range: 300 m	Range: 350 m	Range: 400 m
40	3667.23	2613.73	2254.67	2014.18
50	4067.81	2757.35	2490.65	2267.85
60	4431.19	3501.38	2980.53	2755.35
70	5001.05	4342.75	3466.40	3271.28
80	5853.03	4846.88	4429.31	3725.03
90	6212.54	5519.86	5232.72	4650.26
100	6367.97	5559.83	5105.35	3935.22
110	7882.66	6064.38	5197.75	4040.72

shows the values obtained for Average Energy Consumed by different number of nodes 40, 50, 60, …, 120 with transmission ranges 250, 300, 350 and 400 m.

Table 3 shows the values calculated for Network Lifetime for number of nodes 40, 50, 60, …, 120 with transmission ranges 250, 300, 350 and 400 m.

Figures 5 and 6 show graphs corresponding to Tables 2 and 3 respectively.

Fig. 5 Graph plot for number of nodes versus Average energy consumed
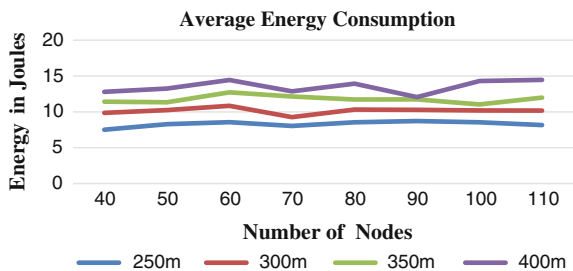
Fig. 6 Graph plot for
number of nodes versus
Network lifetime
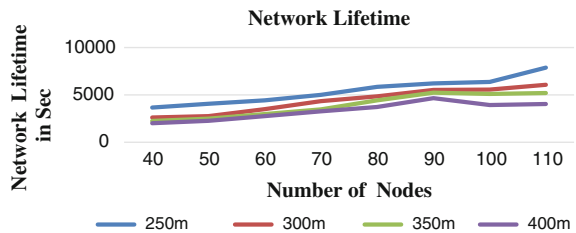
6 Conclusion

In this paper, clustering algorithm uses the highest residual energy and the highest
degree of node to elect as a CH. The next CH will be elected when previous energy
reaches up to the threshold value. In our algorithm, two-hop clustering reduces the
number of communication from node to base stations. It helps to save energy where
energy of node is an important resource. Our clustering algorithm tried to minimize
the average energy consumption in WSN. Also system performance is compared
for average energy consumption and network lifetime by using different number of
nodes at different transmission ranges. The simulation results show that the pro-
posed algorithm is able to reduce the average energy consumption to prolong
network lifetime and achieves scalability.

References

1. Krishna Doddapaneni, Fredrick A. Omondi, Enver Ever, Purav Shah, Orphan Gemikonakli,
   Roberto Gagliardi.: Deployment Challenges and Developments in Wireless SensorNetworks
   Clustering, 24th International Conference on AINA Workshops, IEEE (2014).
2. Naveen Sharma and A. Nayyar.: A Comprehensive Review of Cluster Based Energy Efficient
   Routing Protocols for Wireless Sensor Networks, IJAIEM, Volume 3, Issue 1, January (2014).
3. Wendi B. Heinzelman, Anantha P. Chandrakasan, Hari Balakrishnan.: An Application-Specific
   Protocol Architecture for Wireless Micro-sensor Networks, IEEE Transactions on volume 1,
   issue 4, pp. 660–670 (2002).
4. K. Latif, A. Ahmad, N. Javaid, Z.A. Khan, N. Alrajeh.: Divide-and-Rule Scheme for Energy
   Efficient Routing in Wireless Sensor Networks, 4th International Conference on ANT 2013,
   Published by Elsevier (2013).
5. Navin Gautam, Won-Il Lee and Jae-Young Pyun.: Dynamic Clustering and Distance Aware
   Routing Protocol for Wireless Sensor Networks, ACM Spain, pages 9–14, (2009–2010).
6. Babu N V, Puttamadappa. C, Bore Gowda S B.: Energy Efficient Two Hop Clustering for
   Wireless Sensor Networks, IJCSNS International journal of Computer Science and Network
   Security, vol. 13 no. 9, September (2013).
7. Vaibhav Deshpande, Arvind R. Bhagat Patil.: Energy Efficient in Wireless Sensor Network
   Using Cluster of Cluster Heads, IEEE (2013).
8. Ossama Younis and Sonia Fahmy.: HEED: A Hybrid, Energy-Efficient, and Distributed
   Clustering Approach for Ad Hoc Sensor Networks, IEEE transaction on Mobile Computing,
   vol. 3, no. 4, October–December (2004).

Analyzing Complexity Using a Proposed Approximation Algorithm MDA in Permutation Flow Shop Scheduling Environment

Megha Sharma, Rituraj Soni, Ajay Chaudhary and Vishal Goar

Abstract Sequencing and scheduling play a very important role in service industries, planning, and manufacturing system. Better sequencing and scheduling system have a significant impact in the marketplace, customer satisfaction, cost reduction, and productivity. Therefore, proficient sequencing directs to enhance in utilization effectiveness and therefore brings down the time required to complete the jobs. In this paper, we are making an attempt to bring down the time complexity in m-machine flow shop environment by reducing the sequences and to find an optimal or near optimal make-span. For effective analysis, a well-known heuristic algorithm that is, CDS is considered.

Keywords Sequencing · Permutation flow shop scheduling · Time complexity · CDS heuristic · Decomposition method

1 Introduction

The classical method of flow shop scheduling problem (FSP) is one of the most fascinating areas of exploration for more than 50 years [1]. Flow shop scheduling problem is a process in which group of n jobs is processed by a group of m

Megha Sharma (✉) · Rituraj Soni
Computer Science Department, Engineering College Bikaner, Bikaner, India
e-mail: meghasharma.622@gmail.com

Rituraj Soni
e-mail: rituraj.soni@gmail.com

Ajay Chaudhary
Core Techies, Bikaner, India
e-mail: ajaychaudharybkn@gmail.com

Vishal Goar
MCA Department, Engineering College Bikaner, Bikaner, India
e-mail: goar.vishal@gmail.com

machine. There are some assumptions which are considered in any flow shop scheduling environment [2].

(1) Machines are always available throughout the scheduling period.
(2) Order of sequence in all machines is same throughout the scheduling operation.
(3) Preemption is not permitted, i.e., once an operation begins on the machines it should be completed before other operation can start on that particular machine.
(4) All processing time of jobs on machine should be known, in advance.
(5) A set of n jobs are accessible at time zero for processing.
(6) Each job will be processed by each machine once and only once.

Managers usually opt for job sequences and schedules for loading machine that grant total provision for mean flow time, processing time, average lateness, and average tardiness to be minimized. So, in flow shop environment sequencing and scheduling plays a significant role. There are $(n!)$ possible schedule for a sequence of jobs in permutation flow shop scheduling [3]. Therefore, efforts have been made by researchers in the past to reduce these schedules as much as possible to get optimal result.

For small number of problem many exact methodologies have been put forward and they are fruitful. Johnson's algorithm [4] is the primary algorithm and yields the optimal result for two machines and n jobs flow shop scheduling problem. But when size of problem increases exact methods for solving them becomes complicated and computational time also increases. For scheduling n-jobs on m-machine many heuristics have been put forward over the years. In such problems there are many objectives which can be minimized, some of them are: total flow time, total tardiness, total completion time, etc. but minimizing make-span, i.e., total completion time is one of the most widely considered performance measure [5].

Minimizing make-span for m-machine permutation flow shop scheduling has been addressed by many researchers [6–8]. In this paper an approach has been made to reduce time complexity by introducing a concept of decomposition method [9] in heuristic environment, to get optimal or near optimal make-span in less time interval. Using this approach sequences will get reduced as the number of machines increases. That is, less computational effort. Since, the problem for $m \geq 3$ is NP hard, heuristic algorithm requires further improvement so that more optimal results can be obtained.

With this paper, future work of [9] has been put forward and implemented.

2 Literature Review

Many heuristics have been developed in this field so that problem can be solved with minimum usage of processing equipment. Johnson [4] was the first researcher to consider an algorithm by which optimal sequencing can be calculated for n-jobs

and 2-machine problem. Then researcher developed different heuristics for minimizing make-span for m-machine.

Palmer [8] developed a slope index method to calculate sequences of jobs on machines which are also termed as palmer's heuristic. In this method priority was given to those jobs whose operation time tends to grow from one machine to other.

Campbell et al. also known as CDS [7] were an extension of Johnson's algorithm but in these heuristic m-machines were considered. In this method at most (*m*-1) different sequences were developed and from those sequences best sequence is chosen.

Nawaz et al. (NEH) [6] developed this heuristic in which total processing time is considered. In this method higher priority will be given to the job with prominent total processing time on all machines rather than the job with less total processing time on all machines. An analysis based on heuristic algorithms and there comparison [10] with each other using RPD is also studied and published.

In this paper, we gave emphasis on CDS heuristic algorithm.

3 Johnson's Algorithm

In flow shop there are n jobs concurrently available at time zero and to be scheduled by 2-machine arranged in series. Its foremost goal is to obtain optimal sequence of jobs so that make-span can be reduced [4]. This is the most significant result which has become standard theory of scheduling. Johnson's rule is as follow:

Minimum process time for machine 1 should be greater than or similar as that of maximum process time for machine 2.
Minimum process time for machine 3 should be greater than or similar as that of maximum process time for machine 2.

3.1 Step by Step Explanation

(1) Let set $a = \{j, T_{i1} < T_{i2}\}$.
(2) Let set $b = \{j, T_{i1} > T_{i2}\}$.
(3) Sort the jobs in 'a' by ascending order of T_{i1}.
(4) Sort the jobs in 'b' by descending order of T_{i2}.
(5) Then merge both the sorted order that is, $\{a\}$ followed by $\{b\}$.

Let us take an example (Table 1).
The solution will be:

(A) Job set $a = \{1, 4\}$ and $b = \{2, 3, 5\}$.
(B) Optimal sequence form will be $\{1, 4, 5, 3, 2\}$.

Table 1 5-Jobs and
2-machine problem

Jobs	Machine 1 (T_{i1})	Machine 2 (T_{i2})
1	3	5
2	7	4
3	2	1
4	5	6
5	5	5

4 Campbell, Dudek, and Smith Heuristic Algorithm (CDS)

It is an extension of Johnson's algorithm which is applied to m-machine and n-job problem [1, 7]. In this method at most (m-1) separate sequences are developed and from those sequences finest sequence is identified [7]. The main emphasis of this heuristic is to reduce make-span in flow shop problem.

4.1 Step by Step Explanation

(1) Create additional number of n-job and m-machine problems, P, where $P \leq m$-1.
(2) For first problem set $K = 1$.
(3) Now calculate total processing time for each job (i) on machine-1 and machine-2

$$M1 = \sum_{j=1}^{K} Tij \tag{1}$$

$$M2 = \sum_{j=m-k+1}^{m} Tij \tag{2}$$

(4) Now apply Johnson rule to each (m-1) sequences. Select the minimum processing time from two column matrix.
(5) Increment $K = K + 1$ and repeat until $K = P$.
(6) Choose the minimum total processing time out of them and that will be the best sequence.

5 Proposed Approximation Approach: MDA

In this paper, we are using an approach which is termed as decomposition method [3] which is extended to provide an exposition in flow shop environment. In our new approximation algorithm decomposition technique is applied on classical CDS

heuristic algorithm. The new algorithm is termed as MDA that is, modified decomposition algorithm. The foremost purpose of this method is to reduce time complexity so that optimal solution can be obtained in less computational efforts. It will be done by reducing sequences, so in a meantime two objectives will be reduced. In our paper time complexity and sequences of MDA are compared with CDS heuristic algorithm because CDS is a sterling heuristic algorithm which always contributes optimal or near optimal results as compared to other heuristic algorithm.

5.1 Step by Step Explanation

(1) Convert the given m-machine and n-job problem where $m > 2$, into 2-machine problem where $P = m$-1, m-2, m-3 … m-K and $K = 1, 2, 3 … N$. Value of K depends on number of machines.
(2) Split the set of m-machine into $m/2$ pair so that it can be converted into 2-machine problem.
(3) If m is odd then apply SPT rule to the last machine which is left unpaired.
(4) Then apply Johnson's rule to each set of pair to get optimal sequence (s_K).
(5) By using those obtained sequences that is, s_K we will calculate make-span from original problem as in CDS. Minimum make-span will be the best solution.

6 Results and Analysis

For understanding the above algorithms we have taken an example with 5 jobs and 5 machines problems. On this test case we will apply both the algorithms and then we will calculate their make-span and compare.

By applying rules of MDA approach we will split the above problem into m/2 set of pairs and we will get three set of pairs as follows:

(1) For M1 and M2 sequence is {J4-J5-J1-J2-J3} and calculated make-span is 38.
(2) For M3 and M4 sequence is {J2-J1-J3-J5-J4} and calculated make-span is 39.
(3) For M5 sequence is {J3-J1-J2-J4-J5} and make-span is 40.

We apply CDS heuristic to the same problem by adding each machine from left and right side till (m-1) sequences. By doing this we will get four sequences.

(1) For M1 and M5 sequence is {J5-J4-J2-J1-J3} and make-span is 41.
(2) For M1+M2 and M5+M4 sequence is {J5-J2-J1-J4-J3} and calculated make-span is 40.
(3) For M1+M2+M3 and M5+M4+M3 sequence is {J5-J2-J1-J4-J3} and calculated make-span is 40.

Table 2 Problem with 5-Jobs and 5-machines

Jobs/i	M1	M2	M3	M4	M5
J1	5	5	3	6	3
J2	4	4	2	4	4
J3	4	4	3	4	1
J4	3	6	3	2	5
J5	3	5	6	3	7

(4) For M1+M2+M3+M4 and M5+M4+M3+M2 sequence is {J4-J5-J2-J1-J3} and calculated make-span is 38.

So, by using MDA approach make-span is 38 and when we apply CDS approach on same problem, make-span obtained is also 38. But with MDA we need to calculate only three sequences that is, less computational effort is done with MDA than benchmark algorithm because in that $(m-1)$ sequences are calculated.

For the algorithms considered, codes were generated in java because java is open source, platform independent, and object-oriented, it has built in ability to support national character set, portable, it has built in collection classes, java virtual machine JVM is available on almost all platforms so java codes can run on any machine and JVM prevents machine or system from an incorrectly written application which can cause harm or problem to computing environment. MDA and CDS are a new approach which is generated in java. Codes were run in an i5 PC with 4 GB RAM. The following example in Table 2, is implemented in java, its make-span and total time is calculated and an interface is developed. The interface developed is shown below.

In Fig. 1 make-span and total time for CDS and MDA are taken and we can easily see the difference. CDS take more time than MDA. CDS take 0.982899 ms for calculating make-span and MDA take 0.622462 ms for calculating make-span, difference between total time is 0.360437 ms. Improvement of MDA on test cases is shown and calculated using RPD [10] that is, relative percentage deviation as performance measure so that proposed algorithm can be differentiated with benchmark algorithm solutions and it is defined as:

$$RPD = \frac{SR - SH}{SR} * 100\%$$ (3)

SH Solution of proposed heuristic algorithm.
SR Solution of reference heuristic algorithm.

Short improvement of MDA in Fig. 1 is 40 % as compared with CDS algorithm. We have applied these algorithms on several test cases and those test cases are taken from references which are published. Some of them are listed below with total time in nanoseconds and improvement.
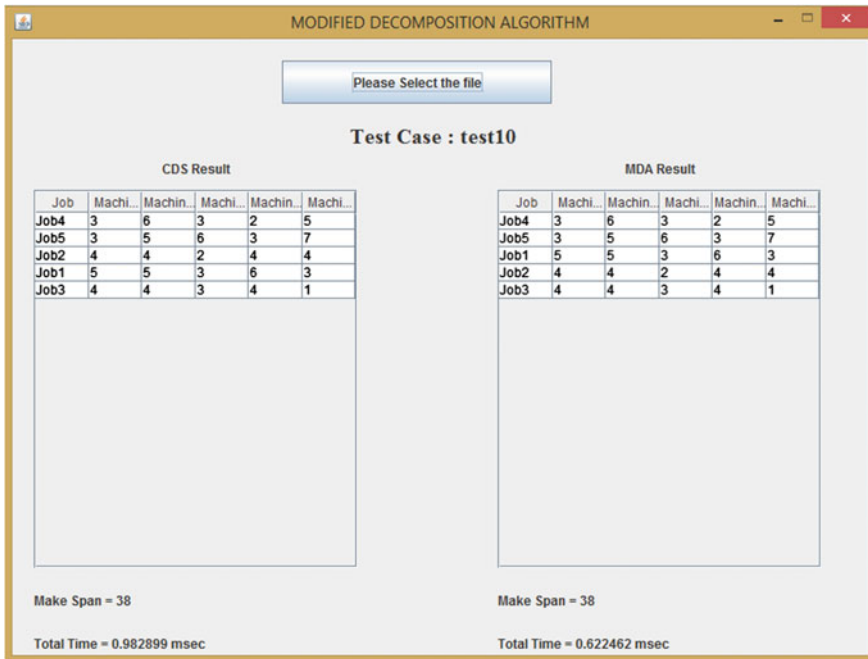
Fig. 1 Interface for CDS and MDA algorithm

Table 3 Improvement of proposed algorithm

Test cases	Make-span of CDS	Time taken by CDS	Make-span of MDA	Time taken by M	Improvement of MDA (%)
1	170	863360	170	228821	73
2	32	367079	32	172068	53
3	53	281347	53	103240	63
4	79	233047	79	117127	50
5	43	251159	43	123164	51
6	50	259612	50	95996	63
7	54	1058974	54	382172	64
8	38	633935	38	254781	60
9	40	271082	40	148522	45
10	70	227009	70	112297	51

In Table 3 it is concluded that MDA is improving with each test case which is involved, so MDA yields feasible make-span. Improvement of MDA is shown (Fig. 2).

Average of improvement is also considered for test cases and for that average of best 90 %, best 60 % and, best 50 % test cases is taken so that comparison of

Fig. 2 Improvement In MDA
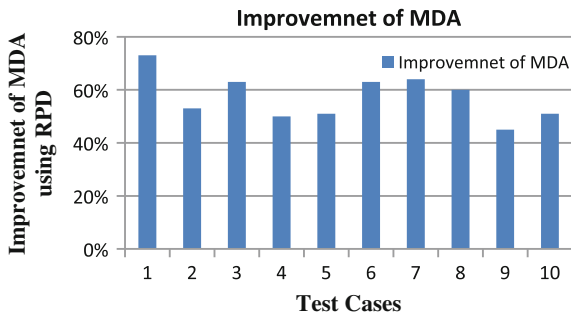
Table 4 Comparison of improvement of mda

Total number of test cases	Average of improvement in 100 % test cases (%)	Average of improvement in best 90 % test cases (%)	Average of improvement in best 60 % test cases (%)	Average of improvement in best 50 % test cases (%)
80	57	59	63	65

improvement between CDS and MDA can be easily understood. They are described below in Table 4.

In this paper, we are making an attempt to reduce time complexity by reducing sequences. When a comparison is made between MDA and CDS then more computational effort is made by CDS as compared with MDA that is, we are obtaining same make-span with both algorithms but in less time and it can be concluded from Fig. 1 and Table 3. In each and every case MDA is improving. We summarize these results with graphs. A comparison is made between both the algorithms in terms of total time taken by algorithms and number of sequences generated to compute make-span. Time will be in nanoseconds. Generation of sequences depends upon number of machines taken in test cases.

Fig. 3 Difference of total time taken by CDS and MDA
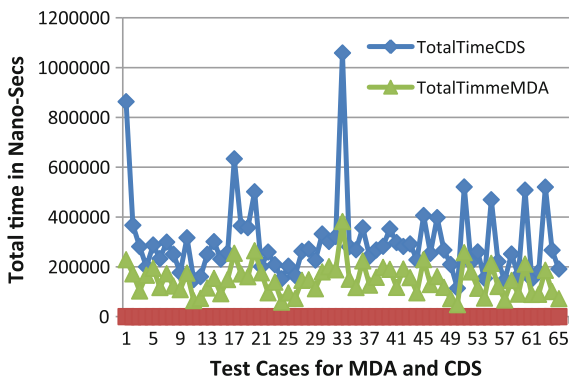
Fig. 4 Difference of sequences generated by CDS and MDA
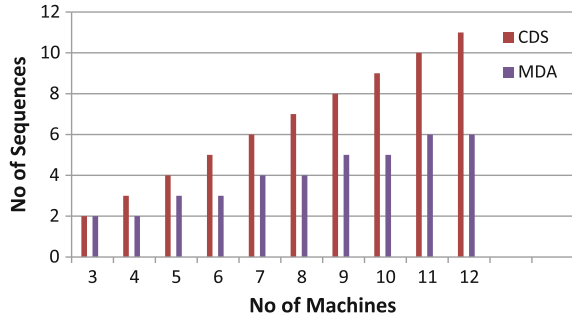
In Fig. 3 we are explaining difference in total time for calculating make-span. A difference is seen in total time taken by both the algorithms in graph that is, CDS take more time than MDA and in Fig. 4 we are showing results in terms of sequences and it can be seen that in CDS more sequences are generated than MDA and generation of sequences depend upon number of machines.

7 Conclusion and Future Work

In this paper, an effort has been made to study about sequence-dependent operations and main idea of this paper is to reduce time complexity by reducing sequences for a problem. Here, we conclude that MDA is a robust and multi-objective algorithm, with which we can find quickly, and can obtain feasible results to sequencing problem that includes multiple machines and multiple jobs in less time. When clarity of algorithm and good results of problem are concerned the proposed solution is better than CDS algorithm because CDS at some point becomes complicated when machine increases. Complexity calculated for MDA is $O\left(\frac{n^2}{2}\right)$ and complexity calculated for CDS is $O(2n^2)$. In CDS heuristic algorithm we need to calculate (m-1) sequences [7] but in proposed method we need to calculate less sequences than CDS to obtain feasible solution. Calculation of sequences in MDA depends upon number of machines. So, we can see that proposed method yields better and faster results than benchmark algorithm that is, with less computational efforts.

With this we can examine that proposed technique is a time saving method and it reduces time complexity of a problem.

Proposed method proves to be good method in terms of complexity. In future work, we will work more on MDA and manipulate sequences by using splitting techniques on processing time of problems so that we can reduce make-span and idle time of algorithm when compared with heuristic algorithm.

References

1. Pinedo, M. Scheduling: Theory, Algorithms and Systems. Prentice Hall, New Jersey, second edition 2002.
2. Addison. Wesley, R. L. Conway, W.-L. Maxwell, L,. W,: Miller,.'Theory-of. Scheduling',. reading., May 1967.
3. Michael Pinedo–Flow Shop- community. Stern. Nyu. Edu fculty/scheduling/shakhlevich/handout.
4. Johnson S.M., (1954). "Optimal two and three stage production schedules", Noval Research logistics quarterly 1, 61.
5. "A New Heuristic Algorithm using Pascal's Triangle to determine more than one Sequence having Optimal/ near Optimal Make-span in Flow Shop Scheduling Problems" Baskar A and Anthony Xavior M vol 39, Feb 2012.
6. Nawaz. M.; Enscore, E.; Ham, I. Aheuristic algorithm for the m machine, n job flow shop sequence problem, OMEGA, 11, 1(1983), 91–95.
7. Campbell, H. G.; Dudek, R. A.; Smith, M. L. A Heuristic Algorithm for the n Job, m Machine Sequencing Problem, Management Science, 16 10(1970), 630–637.
8. Palmer, D. S. Sequencing jobs through a multi-stage process in the minimum total time – a quick method of obtaining a near optimum, Operations Research. Q. 16(1965), 101–107.
9. Megha Sharma, Rituraj Soni and Ajay Chaudhary-An Experimental Analysis of Heuristic on Time Complexity For Permutation Flow Shop Scheduling No. 18; July-September, 2015.
10. Pavol Semanco and Vladimir Modrak-A Comparision of Constructive Heuristic with the Objective of Minimizing Make-span in the Flow Shop Scheduling Problem.

Empirical Evaluation of Threshold and Time Constraint Algorithm for Non-replicated Dynamic Data Allocation in Distributed Database Systems

Arjan Singh

Abstract Data allocation plays a significant role in the design of distributed database systems. Data transfer cost is a major cost of executing a query in a distributed database system. So the performance of distributed database systems is greatly dependent on allocation of data between the different sites of the network. The performance of static data allocation algorithms decreases as the retrieval and update access frequencies of queries from different sites to fragments changes. So, selecting a suitable method for allocation in the distributed database system is a key design issue. In this paper, the data allocation framework for non-replicated dynamic distributed database system using threshold and time constraint algorithm (TTCA) is developed and the performance of TTCA is evaluated against the threshold algorithm on the basis of total cost of reallocation and the number of migrations of fragments from one site to another site.

Keywords Distributed database system · Static allocation and dynamic allocation

1 Introduction

Data allocation in a distributed database design can be categorized in two different types: static and dynamic [1, 2]. In the static environment, optimum solution can be obtained through static data allocation in which the retrieval and update access frequencies of queries from different sites to fragments never change. But in a dynamic environment where these access patterns change over time, the static allocation solutions would degrade the performance of distributed database system. The dynamic environment requires the change in data allocation schema in order to reduce the data transmission cost or communication cost as one of the main aims of

Arjan Singh (✉)
Punjabi University, Patiala 147002, Punjab, India
e-mail: arjanpu@gmail.com

distributed database is to provide the locality of the data. Therefore, reorganization or reconfiguration of distributed database system is required to maintain the performance of system during its lifetime. The reconfiguration process entails the following [3]:

a. Physical change
b. Logical change
c. Combination of both (a) and (b).

Wilson and Navathe [4] have classified the reorganization and reconfiguration of distributed database into two different categories: Full/Total redistribution and limited redistribution.

- *Full/Total Redistribution* It involves the new fragmentation and allocation of data.
- *Limited Redistribution* It involves only the reallocation of data/fragments.

The present study concentrates only on the limited redistribution or reallocation of data. The reallocation process requires the monitoring of dynamic distributed database environment to detect the change in access patterns of different queries in the system. Statistics of these changes are evaluated then reallocation process is initiated if needed to improve the performance of dynamic distributed database.

During the past three decades, lot of work has been done for dynamic allocation of data in distributed database systems. Wolfson et al. [5] have introduced a model for dynamic data allocation for data redistribution. Brunstroml et al. [6] have proposed an algorithm named optimization algorithm for dynamic data allocation. Chaturvedi et al. [7] have presented a machine learning-based approached. Chin [8] has proposed an incremental allocation and reallocation of data based on the changes in workload. Lin and Veeravalli [9] have given centralized control-based algorithm for dynamic object allocation. Mei et al. [10] have included security criteria into the allocation of dynamic and replicated distributed file systems. Ulus and Uysal [11] have proposed a threshold algorithm for non-replicated distributed databases. Threshold algorithm reallocates the data fragments according to the changing data access patterns. Singh and Kahlon [12] have proposed an algorithm named threshold and time constraint algorithm (TTCA) for non-replicated dynamic data allocation. The TTCA algorithm of Singh and Kahlon [12] is an extension of two data reallocation algorithms: Optimal algorithm of Brunstrom et al. [6] and threshold algorithm of Ulus and Uysal [11].

In this paper, the data allocation framework for non-replicated dynamic distributed database system using threshold and time constraint algorithm (TTCA) [12] is developed and the performance of TTCA is evaluated against the threshold algorithm of Ulus and Uysal [11] on the basis of total cost of reallocation and the number of migrations of fragments from one site to another site.

2 Threshold and Time Constraint Algorithm (TTCA) [12]

TTCA algorithm includes the time constraint to the threshold algorithm and uses an access counter to keep track of access history of fragments. TTCA algorithm takes into consideration both the threshold value (T) as well as the time (t) at which accesses are made to a particular fragment before reallocating the fragment from one site to another site. Initially all the fragments are distributed over different sites using any static allocation method in a non-redundant manner. TTCA maintains an access counter matrix M of size $m \times n$, where m denotes the total number of fragments and n denotes the total number of sites. M_{ij} is the number of accesses to fragment i by site j. Figure 1 shows the data allocation framework for dynamic distributed database system using TTCA algorithm, where CRS is the counter of remote site and CCO is the counter of the current owner site of the fragment.

3 Experimental Setup and Results

A fully connected network consisting of four sites (S_1, S_2, S_3, and S_4) and a database consisting of five fragments (F_1, F_2, F_3, F_4, and F_5) is taken into consideration. The unit transmission cost between different sites is given in the Fig. 2. The size of each fragment is given in Table 1. Initial fragments allocation schema is shown in Fig. 2. According to initial allocation, fragment F_1 is allocated to site S_1; fragment F_3 is allocated to site S_2; fragments F_4 and F_5 are allocated to site S_3, and fragment F_2 is allocated to site S_4. A set of 100 independent randomly generated transactions for each fragment is initiated from distinct sites to check the performance of TTCA algorithm and threshold algorithm.

Performance comparison of TTCA algorithm and threshold algorithm is done on the basis of two different measures: total migration cost to reallocate a fragment and the number of migrations. Results are obtained after the completion of 100 transactions for each five fragments. Shortest path from current owner to remote owner is used for the migration of a fragment. Behavior of both the algorithms is observed based on two different threshold values 5 and 7. The specified time (t) for TTCA algorithm is taken as 2 hours.

Figure 3a shows that the total cost of reallocation of fragment F_1 taken by threshold algorithm is more than that of TTCA algorithm for both the threshold values. Reallocation cost in case of threshold algorithm increases, as the threshold value increases from 5 to 7. But reallocation cost in case of TTCA algorithm remains same as the threshold value increases from 5 to 7. Figure 3b shows that TTCA algorithm decreases the number of migrations as compare to threshold algorithm for both the threshold values.

Figures 4a, 5a and 6a show that the total cost of reallocation of fragments F_2, F_3, and F_4 taken by threshold algorithm is more than that of TTCA algorithm for both the threshold values. Reallocation cost of threshold algorithm and TTCA algorithm
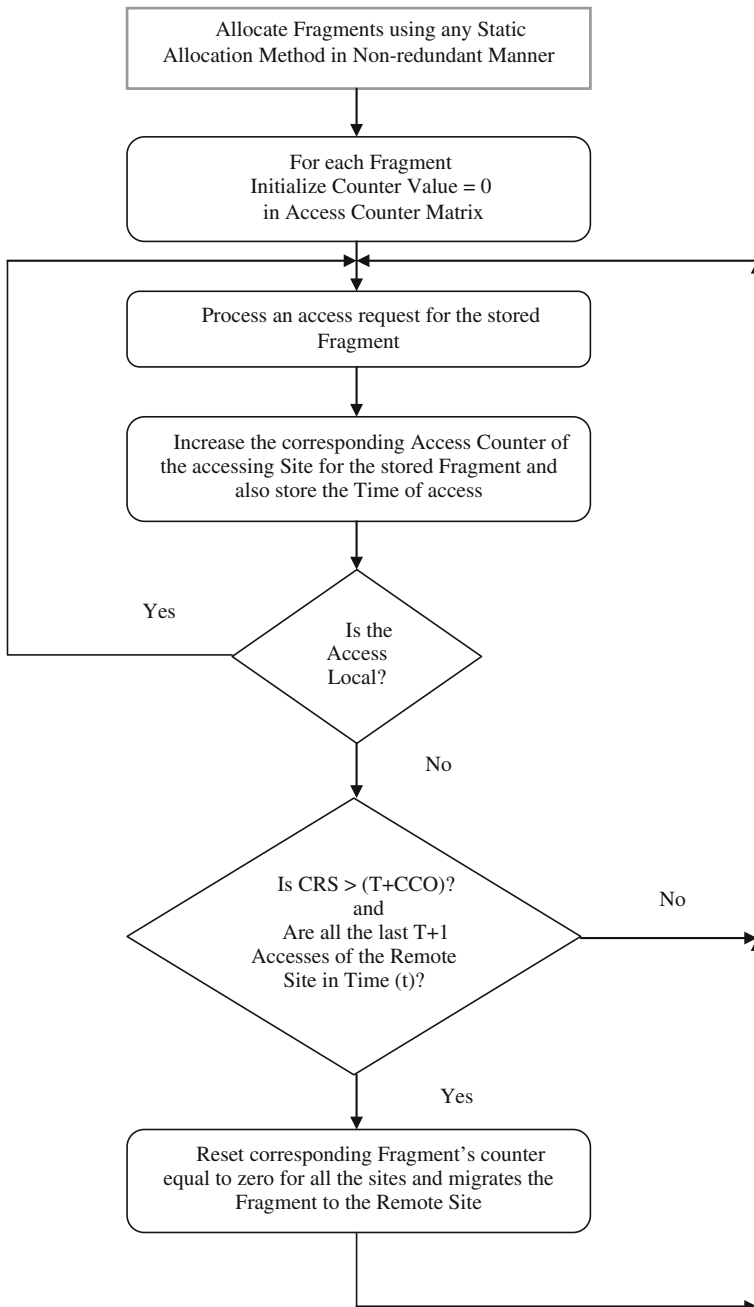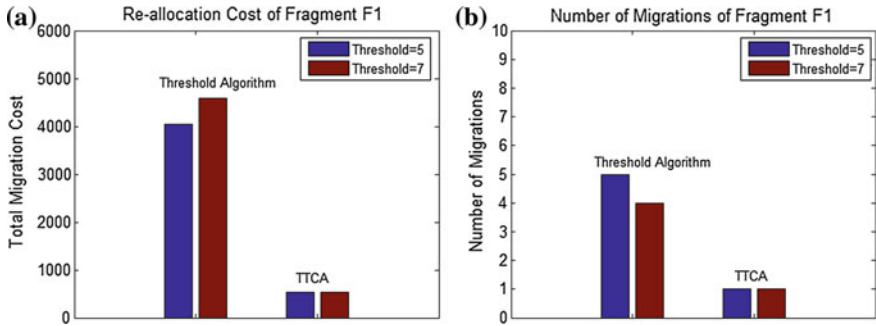
Fig. 1 Data allocation framework using TTCA

Fig. 2 Fully connected networks
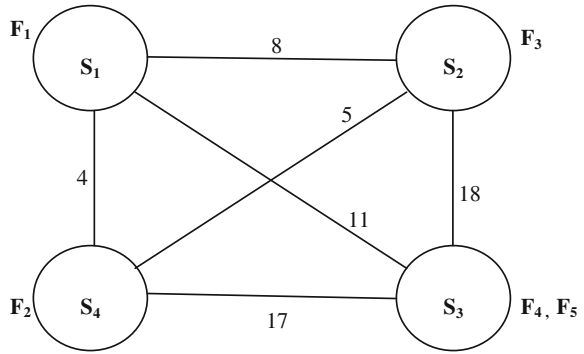
Table 1 Size of each fragment

Fragments	F_1	F_2	F_3	F_4	F_5
Size	135	245	320	160	57

Fig. 3 Reallocation cost and number of migrations of fragment F_1

decreases as the threshold value increases from 5 to 7. Figures 4b, 5b, and 6b also show that the number of migrations of fragments F_2, F_3, and F_4 decreases as the threshold value increases in both the algorithms. But TTCA algorithm further decreases the number of migrations as compare to threshold algorithm.

Figure 7a shows that the total cost of reallocation of fragment F_5 taken by threshold algorithm more than that of TTCA algorithm for both the threshold values. Reallocation cost of threshold algorithm decreases as the threshold value increases from 5 to 7. On the other hand, reallocation cost of TTCA algorithm remains same as the threshold value increases from 5 to 7. Figure 7b shows that TTCA algorithm decreases the number of migrations as compare to threshold algorithm. The number of migrations of fragment F_5 increases as the threshold value increases in case of threshold algorithm. But the number of migrations of fragment F_5 remains same as the threshold value increases in case TTCA algorithm.
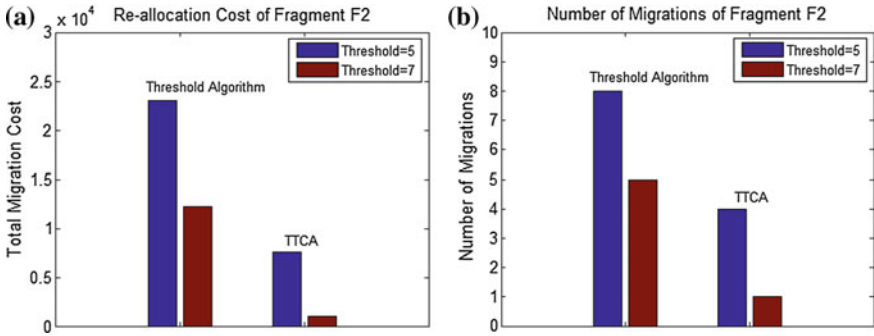
Fig. 4 Reallocation cost and number of migrations of fragment F_2
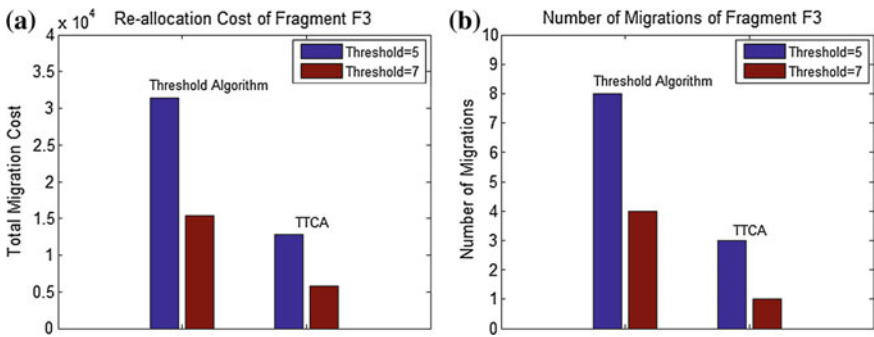
Fig. 5 Reallocation cost and number of migrations of fragment F_3
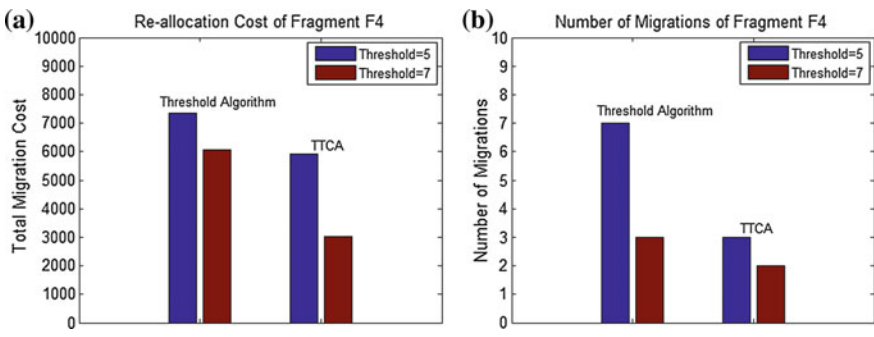
Fig. 6 Reallocation Cost and Number of Migrations of fragment F_4

Form the above discussion; it is clearly evident that TTCA algorithm is an effective method for reallocation of data as compare to threshold algorithms for distributed database system where the frequency of access pattern changes rapidly. TTCA algorithm decreases the total cost of reallocation and the number of
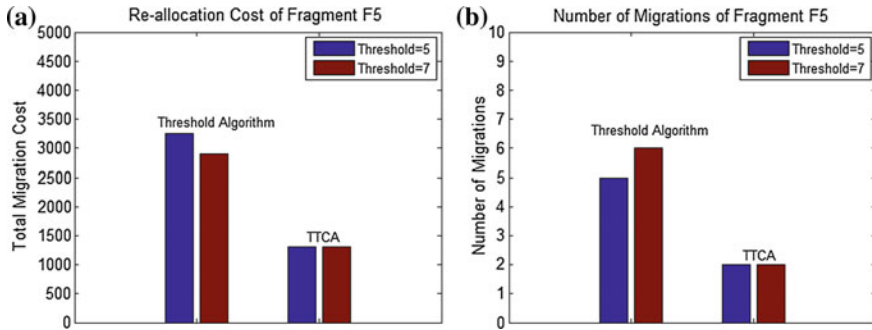
Fig. 7 Reallocation cost and number of migrations of fragment F_5

migrations of fragments from one site to another site as compare to threshold algorithm. TTCA algorithm improves the performance of the distributed database system by decreasing the network traffic and reduce the data transfer cost compared to threshold algorithm.

TTCA gives the ownership of fragments to the site having maximum access rate but on the other hand threshold algorithm gives the ownership of the fragment to last accessing site which may or may not be the frequent user of the fragment. TTCA ensures that the migration rate of the fragments decreases or remains same as the threshold value increases. But in case of threshold algorithm migration rate can be increased as the threshold value increase. By minimizing the number of migrations, TTCA algorithm further improves the overall performance of the distributed database system.

4 Conclusion

A framework for non-replicated dynamic allocation of data in distributed database has been developed using TTCA algorithm. TTCA algorithm reallocates data with respect to the changing data access patterns with time constraint. Results show that the TTCA algorithm is an effective method for reallocation of data as compare to threshold algorithms for distributed database system where the frequency of access pattern changes rapidly. TTCA algorithm decreases the total cost of reallocation and the number of migrations of fragments from one site to another site as compare to threshold algorithm. TTCA algorithm improves the performance of the distributed database system significantly by decreasing the network traffic and reduce the data transfer cost as compared to threshold algorithm. TTCA gives the ownership of fragments to the site having maximum access rate, but on the other hand threshold algorithm gives the ownership of the fragment to last accessing site which may or may not be the frequent user of the fragment. TTCA ensures that the migration rate of the fragments decreases or remains same as the threshold value

increases. But in case of threshold algorithm migration rate can be increased as the threshold value increase. By minimizing the number of migrations, TTCA algorithm further improves the overall performance of the distributed database system. In future, TTCA can further be extended for replicated dynamic allocation of data.

References

1. Ceri, S., Pelagatti, G.: Distributed Databases: Principles & Systems. McGraw-Hill International Editions.
2. Ozsu, M., Valduriez, P.: Principles of Distributed Database Systems. Prentice Hall, second ed. 1999.
3. Rivera-Vega, P.I., Varadarajan, R., Navathe, S.B.: Scheduling Data Redistribution in Distributed Databases. In Proceedings of 6th International Conference on Data Engineering, pp.166–173, 5-9 February 1990.
4. Wilson, B., Navathe, S.B.: An Analytical Framework for the Redesign of Distributed Databases. In Proceedings of the 6th Advanced Database Symposium, pp. 77–83, 1986.
5. Wolfson, O., Jajodia, S., Huang, Y.: An Adaptive Data Replication Algorithm. ACM Trans. Database Systems, vol. 22, no. 2, pp. 255–314, 1997.
6. Brunstroml, A., Leutenegger, S.T., Simhal, R.: Experimental Evaluation of Dynamic Data Allocation Strategies in a Distributed Database with changing Workload. ACM Trans. Database Systems, 1995.
7. Chaturvedi, A., Choubey, A., Roan, J.: Scheduling the Allocation of Data Fragments in a Distributed Database Environment: A Machine Learning Approach. IEEE Trans. Eng. Management, vol. 41, no. 2, pp. 194–207, 1994.
8. Chin, A.: Incremental Data Allocation and Reallocation in Distributed Database Systems. Journal of Database Management, vol. 12, no. 1, pp. 35–45, 2001.
9. Lin, W.J, Veeravalli, B.: A Dynamic Object Allocation and Replication Algorithm for Distributed System with Centralized Control. International Journal of Computer and Application, vol. 28, no. 1, pp. 26–34, 2006.
10. Mei, A., Mancini, L., Jajodia, S.: Secure Dynamic Fragment and Replica Allocation in Large-Scale Distributed File Systems. IEEE Trans. Parallel and Distributed Systems, vol. 14, no. 9, pp. 885–896, Sept. 2003.
11. Ulus, T., Uysal, M.: Heuristic Approach to Dynamic Data Allocation in Distributed Database Systems. Pakistan Journal of Information and Technology, 2(3): pp. 231–239, 2003.
12. Singh, A., Kahlon, K.S.: Non-replicated Dynamic Data Allocation in Distributed Database Systems. International Journal of Computer Science and Network Security, VOL.9 No.9, pp. 176–180, September 2009.

Automated Usability Evaluation
of Web Applications

Sanchita Dixit and Vijaya Padmadas

Abstract Evaluating the usability of web applications in terms of their learn-ability and navigability with the current design principles is a prospective area where improvement in the web designing principles should be revolutionised. The availability of any help system while navigating the web applications could be a progressive change towards providing better experience to the end users. Providing such system would be cost effective rather than suggesting the improvements in the design process which might not comply evenly with the web application developers across the globe. The research work presented here aims at providing such a help system to users which improves the user's experience of the web application in terms of two factors: learn-ability and navigability. The proposed help system provides the suggestions in the form of links to be clicked next according to the user goal. The test-bed for the experiments conducted is Virtual Labs web application which is designed as a set of virtual laboratories aimed at users who are deprived of physical infrastructure for carrying out laboratory experiments in their institutes.

Keywords Usability · Learn-ability · Navigability

1   Introduction

The web has been experiencing a continuous and impressive growth in the last decade because of the easily available tools for the creation and publishing of information on the web. The end user finds it difficult to understand and learn the functioning of different services provided by the web applications. The goal of getting a set of design specifications which can help the developers create websites favourably accepted by the users was the motivation behind taking up this research work.

Sanchita Dixit (✉) · Vijaya Padmadas
Thadomal Shahani Engineering College, Mumbai, India
e-mail: dixit.sanchita90@gmail.com

Vijaya Padmadas
e-mail: vijayapadmadas@gmail.com

To analyse the above problem, the case study considered here is the Virtual Labs (http://virtual-labs.ac.in) web application which is a learning management system. Here all the relevant information like different course experiments, video lectures and animated demonstrations are found at a common place. The labs provide students a visual demonstration of techniques and concepts of their concerned subjects bringing in a self-paced learning environment.

With the presence of huge number of learning web applications, it can be debated as to why a particular application is preferred over another for achieving an objective and it is said to be successful if it serves its purpose completely and have a good number of visitors with enough usage time. There could be various influencing factors like purpose, clarity, user focus, navigation, content, usability, accessibility, appearance, credibility, regular updates, etc. [1–3]. We have focused on usability. According to Neilsen, usability is a quality attribute that assesses how easy user interfaces are to use [4].

Some of the parameters being considered for study are:

- *Usability:* During the navigation of a web application, if the user gets a feeling of being lost, or gets lots of alert/error messages, they will most likely stop accessing the web application. Hence it is important to assess the extent to which a particular web application is usable. Several usability assessment mechanisms are available which involve automated evaluations as well as manual evaluations using human subjects that form the representative set of the intended audience of the website. Based on their feedback the structure, design and process of the web application development can be iterated.
- *Learn-ability:* Learn-ability = Understanding + Remembering learn-ability refers to how recurring users can get back to the web application and already know how to use it. This could be measured by comparing performances of the user's first session against a later one.
- *Navigability:* Navigability refers to the ease of navigation while browsing through different web pages while the user is trying to accomplish his/her goal. Moving in and out of pages should be intuitive enough so that user feels the ease of entry and exit from a particular page at his/her own wish.

The Virtual Labs have been in the development phase only and not completely deployed for the end users, hence the question arises as to how much the user interface of Virtual Labs is learn-able. The initial user testing performed with the live participants highlighted the issues with the Virtual Labs web pages.

The subsequent experiment is carried out with the same set of participants but providing them with a help system in the form of a tool which aids in the completion of the tasks given to them. The system developed by us takes the input in the form of user goal and provides the link to be clicked next. The next link provided is in the form of a suggestion where the users are free to either go with the suggested link or try to navigate according to their own understanding without considering the help system's suggestion.

An overview of different methods already used and developed is explained in next section. Section 3 describes the hypothesis formulated. The design and methodology of the experiments is covered in Sect. 4. Section 5 describes the results followed by conclusion along with future work in Sect. 6.

2   Related Work

According to Nielsen, Usability refers to the extent to which a product can be used by specified users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specified context of use [4].

Usability is defined by five quality components as per the Nielsen description:

1. *Learn-ability:* Ease of accomplishing basic tasks by the end users when they use the web application for the first time.
2. *Efficiency:* Once users have learned about the web application, how quickly can they perform tasks?
3. *Memorability:* When users return to the web application after a period of not using it, how easily can they re-establish proficiency?
4. *Errors:* How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
5. *Satisfaction:* How pleasant is it to use the design of the web application?

Out of the 10 general heuristics of usability evaluation as given by Nielsen, three evaluated here are: (A) user control and freedom, (B) consistency and design and (C) help and documentation.

The two parameters that are considered for evaluation in this research work are:

(1) Learn-ability

Nielsen defines learn-ability as easy first time use and list learn-ability as a subcomponent of the construct of usability. Another definition of learn-ability is usability over time. A more learn-able system is one that reduces the time it takes to complete tasks.

(2) Navigability

Web navigation refers to the process of navigating a network of information resources in the World Wide Web, which is organised as hypertext or hypermedia. A website's overall navigational scheme includes several navigational pieces such as global, local, supplemental, and contextual navigation [5].

The different approaches which provide navigation support to the user are briefly mentioned below:

1. *Using information retrieval techniques and search mechanisms* [6]: In order to provide better web navigation to blind people, Zajicek et al. (2007) developed a tool which provides a list of links, headings and a summary of the web page [6].

2. *Recording: browsing steps and analysing usage patterns* [7]: Another method is by recording the browsing steps and later allowing them to replay [7].
3. *Constructivist or Graphical approach:* It is proposed by Zeiliger et al. [8]. Basically they gather, represent, structure and create navigational objects with a graphical user interface.
4. *Using training and heuristic approach:* A tool called WebWatcher [9] provides support to the user by employing knowledge about which hyperlinks are likely to lead to the target information [9]. Letizia [10] tool uses heuristic approach to learn user's interest through their behaviour [10].
5. *Using Information Foraging Theory (IFT)*: ScentTrails [11] provides support by merging browsing and searching techniques, and using information foraging theory [11].

2.1   USABILICS

USABILICS provides an interface model that supports the definition of tasks in a simple and intuitive way. Based on this model, evaluators are able to define tasks by simply interacting with the application's graphical interface, in the same way end users are supposed to do. Another important aspect is that it considers the similarity among the possible paths of a given task, allowing a generic approach for the definition of similar paths [12].

2.2   WebRemUSINE

WebRemUSINE is a method and an associated tool to detect usability problems in Web interfaces through a remote evaluation where users and evaluators can be separated in time and space. The approach combines two techniques that are usually applied separately: (a) empirical testing and (b) model-based evaluation.

3   Hypothesis

1. The search mechanism in any web application gets affected by the position of the search box on the landing page in successful completion of the goals by the users. We predict that a help system would improve the search process of different elements in a web page by providing the users with suggestions in the form of links to be clicked next.

2. Presence of redundant information on the different pages in the left and right hand sided widgets results in greater effort to reach the goal by the users. We hypothesise that this would decrease the learn-ability of the web application measured in terms of the effort required by the end users to achieve their goals.
3. Meaning of the icons and the text description must be intuitive for every category of users, absence of which reduces the learn-ability as well as results in failure of goals by the end users. We predict that the users would eventually leave the site if they could not comprehend the meaning of the icons/links.

4 Method

Figure 1 represents the overall formulation of the research methodology. Initially around 500 web pages of Virtual Labs were selected for initial assessment of stage 1. It included labs from all the different domains, e.g. Computer Engineering, Electronics Engineering, etc. Around 20 different labs with each lab of average five experiments and each experiment of average five web pages were studied initially. It helped us in general understanding of the prospective problems related to usability.

For stage 2, once identified with the potential usability issues, a subset of 200 web pages was considered. Based on the above assessment, solutions are to be proposed for each problem detected and thereby automated for improved usability of web applications.

Fig. 1 Diagram for research methodology
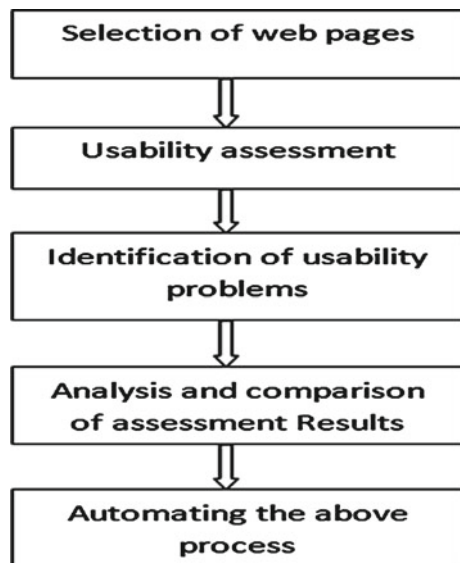
The overall methodology followed is as follows:

1. For learn-ability: First two rounds of experiments are conducted with the same set of participants used for the evaluation of external learn-ability metrics (defined by ISO/IEC) by user observation and Selenium IDE evaluation tool. Same tasks were given in both the rounds of experiments to assess how much the test-bed of virtual labs web application is learn-able.
2. For navigability: Navigation support system/help system is used in the second experiment. The help system is based on the semantic similarity between user goal and hyperlinks present in the web pages where the selection process of hyperlinks is focussed. The system sends the query to goal-specific competing headings/link-analysis tool by Colorado University Database. Based on the latent semantic analysis (LSA) scores generated, it gives suggestions in the form of link to click next and thereby helping the user to locate the web page to reach their goal.

4.1   Design

The experiment followed a A/B testing methodology where two versions (A and B) are compared. Version A is the currently used version (control), while Version B is modified in one respect (treatment). The first experiment carried out is the initial assessment phase where the Virtual Labs web application is tested as it is without any modifications. The subsequent experiment is carried out with the same set of participants and similar constraints with only the difference of help system. The help system developed by us is a navigation tool wherein the users can provide their input in the form of website URL as well as their goals. Total of 26 participants participated. Out of which 15 were male and 11 were female candidates. The age group considered is between 17 to 28 years.

4.2   Technical Requirements

The most basic requirement is the internet access. For recording the entire navigational history along with the clicks, Selenium IDE tool is installed on the machine on which the experiment is performed. The entire recordings can be replayed back in the form of test suites saved for each participant. The complete action generated is logged in the form of HTML tables and stored in the local database provided by the Selenium IDE tool itself. For the second experiment in addition to the above, the help system has to be run from the evaluators' machine which requires python interpreter, HTTPlib-2.0 and Beautiful Soup as the necessary packages.

4.3   Role of Evaluator

The role of evaluator is being performed by us by observing each participant's activity and noting down mistakes while they perform their tasks. The observations include long inaction time and the total time taken by each participant to complete the task.

4.3.1   Procedure

Prior to the beginning of the experiment, the participants were asked about their general information like name, age, email id and present occupation. Each participant was given five similar tasks to perform mostly related to accessing a particular experiment page in the website. While the tasks were performed, clicks to navigate to a different page or within the same page are recorded. The participants were given a choice of leaving the website whenever they wish to if not able to complete the tasks successfully. The total time taken to complete each task either successfully or not is also recorded by the evaluator. The raw data given by the Selenium IDE were exported in the form of HTML files.

5   Results

5.1   Ease of Function Learning (ISO/IEC 12207 SLCP Metric)

How long does the user take to learn to use a function? (ISO/IEC Definition)

Function here is the ability to access an experiment and the time referred is the mean time in successive experimental rounds conducted with the same set of participants (Fig. 2).

In the above figure, the first column indicates the five different tasks given to the participants. The second and third columns represent the time taken by successful participants of each task where first value represents the number of successful participants and second value represents the mean time taken to complete the task. The fourth column gives the mean time taken to learn to use a function correctly where,

T = mean time taken to complete a task in experiment 1/mean time taken to complete a task in experiment 2 (represented by ratio R)

Last column represents the relative difficulty of tasks where the results indicate that Task 5 proved to be the most difficult leading to maximum failure rate initially but also indicating the highest improvement in the success rate of the task after using help system.

Tasks	Experiment 1	Experiment 2	Ratio	Relative Difficulty of Tasks
Task 1	26, 2.52 mins	26, 2.00 mins	1.26	2.26 mins
Task 2	10, 2.20 mins	26, 1.65 mins	1.33	1.93 mins
Task 3	23, 1.16 mins	26, 0.54 mins	2.14	0.85 mins
Task 4	12, 2.79 mins	20, 2.50 mins	1.16	2.65 mins
Task 5	05, 3.17 mins	15, 2.95 mins	1.07	3.04 mins

Fig. 2 Table showing the time taken by successful participants in experiments 1 and 2

In experiment 2, total number of participants who could complete all the tasks successfully is 12 compared to the experiment 1 where only 2 participants could complete all the tasks. Hence the help system has improved the success rate of completion of all tasks from 7.7 to 46.2 %.

5.2 Ease of Learning to Perform a Task in Use (ISO/IEC 12207 SLCP Metric)

How long and how many total numbers of useful clicks are required to learn how to perform the specified task efficiently? (ISO/IEC Definition)

From the Fig. 3, Effort = total time and number of useful clicks

Tasks	Initial Effort	Threshold	Final Effort
Task 1	2.52 mins, 9 clicks	1.76 mins	2.00 mins, 5 clicks
Task 2	2.20 mins, 9 clicks	1.54 mins	1.65 mins, 5 clicks
Task 3	1.16 mins, 8 clicks	0.81 mins	0.54 mins, 5 clicks
Task 4	2.79 mins, 11 clicks	1.95 mins	2.50 mins, 6 clicks
Task 5	3.17 mins, 13 clicks	2.22 mins	2.95 mins, 7 clicks

Fig. 3 The initial and final effort by the participants in the experiments 1 and 2

where, Total Time = mean time to perform a task successfully

No. of useful clicks = total no. of useful clicks required to reach to the function given as a task to the participant.

Useful clicks are defined as: Any click on a dynamic link on the web page which leads to an information page and not an error page. The number of useful clicks is given by Selenium IDE.

The second column of the table shows the initial effort after the first round of experiment where the first value gives the mean time to complete a task successfully and second value represents the mean clicks required to complete the task successfully. No help system was provided initially in order to aid in the tasks. The participants had no prior knowledge about the web application and were using it for the first time.

The next column represents the threshold which is defined as:

Threshold = 70 % of the users effort of the first time use of the web application (ISO/IEC Reference)

The last column shows the final effort required by the same participants after the second experiment to complete the tasks successfully. The effort value has decreased for each task after the help system was provided to the participants.

The final effort for each task is well within the threshold both in terms of mean time as well as mean number of clicks required after the help-system has been used. In Experiment 1, 90 % of the participants did not use the search box provided on the home page. The results validate our first hypothesis of improvement in the location of the web pages which is evident by the improvement in the number of successful tasks by same participants in the successive two experimental rounds.

From Fig. 4, the results of the user feedback depict the users willingness and need for a help system/help documentation, lack of which could result in leaving the website by the user.

The lack of any help could result into high likely possibility of leaving the site as evident by the 42 % participants. 34 and 30 % participants agreed to high likely to the fact that any such help system would reduce their efforts in terms of completion of the tasks.

From the user feedback and the difficulties listed by the participants, our hypothesis is that the presence of redundant information in different pages resulted in increased effort to complete a task successfully and hence decreased learn-ability of the web application, is validated. 99 % of the participants did not click on the links present on the left/right hand side widgets of home page. (Institute wise distribution of labs, FAQs etc.)

Our hypothesis that the absence of intuitive icons/text description/links would lead to leaving the website by the end users is validated from the fact that 95 % of the participants asked for help by the evaluator when reached till the respective lab page as to which link to be clicked for the particular experiment page. (e.g. SIMULATION icon) which highlights the limitation of the help system. For the successful participants, when asked by the evaluator the meaning of icon/text description "SIMULATION", 90 % guessed it based on the other icons present and agreed that they could not decipher the meaning by the text itself.

Questions	Least Likely (In %)	Somewhat Likely (In %)	Very Likely (In %)	Highly Likely (In %)
How likely useful the help documentation/help system would be, if provided with one during the task?	7.7	26.9	53.8	11.5
If provided with one such help during the task, would you still move ahead with your own understanding, neglecting the help?	34.6	38.5	19.2	7.7
Would the lack of help affect your willingness to visit the site again?	11.5	15.4	30.8	42.3
Do you think the help would reduce the time required to perform the same task again?	3.8	30.8	34.6	30.8

Fig. 4 User feedback after first round of experiment

6 Conclusion and Future Work

A help system in the form of a navigability tool has been developed to facilitate the user interaction with the web application. Metrics have been proposed and their calculations have been done using the existing as well as proposed tool to study the learn-ability aspect. We observed that the incorporation of a help system improved the efficiency of the end users while using the website in terms of reduced efforts. It not only helped in making the website better navigable but also helped in better learn-ability by the users.

As a future work, we intend to propose a framework or model which could automate the complete process of providing guidelines to the designers for better designed interface including intuitive as well as comprehensible links/icons for better understanding by the end users.

References

1. 6 Website Success Factors. From: http://www.fireflyweb.ie/Blog-Posts/6-website-success-factors.html. Last Accessed: March, 2014.
2. Nielsen. J. Usability 101: Introduction to usability. Jakob Nielse's Alertbox (2003).
3. Virtual Labs web site. [Online]. Available: https://virtual-labs.ac.in.

4. A. Fernandez, et al. Usability evaluation methods for the web: A systematic mapping study. Information and software Technology, 2001.
5. Web Navigation. From: https://en.wikipedia.org/wiki/Webnavigation.html.
6. Zajicek, M., Powell, C., and Reeves, C.A Web navigation tool for the blind. In Proc. of third international ACM conference on Assistive technologies, (1998), 204–206.
7. Anupam, V., Freire, J., Kumar, B., and Lieuwen, D. Automating Web navigation with the WebVCR. Computer Networks 33, no. 1 (2000), 503–517.
8. Zeiliger, R., Belisle, C., and Cerratto, T. Implementing a Constructivist Approach to Web Navigation support. World Conference on Educational Multimedia, Hypermedia and Telecommunications. (1999).
9. Armstrong, R., Freitag, D., Joachims, T., and Mitchell, T. Webwatcher: A learning apprentice for the world wide web. In AAAI Spring symposium on Information gathering from Heterogeneous, distributed environments, (1995) 6–12.
10. Lieberman, H. Letizia: An agent that assists web browsing. IJCAI (1) (1995), 924–929.
11. Leandro Guarino de Vasconcelos, Larcio Augusto Baldochi Jr., Universidade Federal de Itajub, Towards an automatic evaluation of Web applications in 2012 ACM Proceedings of the 27th Annual ACM Symposium on Applied Computing.
12. Hamilton Fernandes de Moraes Junior, Fabia Lika Nishida and Ana Cristina Vieira de Melo, Department of Computer Science University of So Paulo USP, So Paulo, Brazil, Modelling Websites Navigation Elements According to Usability Aspects in 2012 IEEE Eighth International Conference on the Quality of Information and Communications Technology.

An Efficient Approach for Frequent Pattern Mining Method Using Fuzzy Set Theory

Manmay Badheka and Sagar Gajera

Abstract In methods of data mining, association rule mining is well suited for applications such as decision making, catalog design, forecasting, etc. Existing association rule mining (ARM) methods are only applicable to a dataset consisting of binary attributes. But for the real-world application, traditional methods of ARM should be modified to handle numerical attributes as well. Fuzzy set concepts provide solution to some disadvantages of traditional ARM methods. Fuzzy membership function is providing convenient way to quantize numerical attribute of a dataset compared to other traditional discretization technique. Association rules generated by using fuzzy membership function with linguistic label have a better interpretability.

Keywords Quantitative association rule · Fuzzy set concept · Fuzzy support · Fuzzy confidence · Certainty factor

1 Introduction

Data mining is a subdiscipline of computer science. It is a process of uncovering patterns from the data in abundance. It involves methods of combining various fields such as database system, statistics, and machine learning.

Frequent pattern mining is a technique of finding hidden pattern from a data which occurs frequently. It is presented in the form of association rule ($C \Rightarrow D$). So, it is also known as association rule mining. It is useful in many sectors such as banking, bioinformatics, financial corporation, etc. Basic approach in frequent

Manmay Badheka (✉) · Sagar Gajera
L. J. Institute of Engineering and Technology, Ahmedabad, India
e-mail: manmaybadheka@gmail.com

Sagar Gajera
e-mail: sdgajera14@gmail.com

pattern mining is to determine a relation among data in a database but it should be interesting. There are various interesting measures such as

$$\text{Support}(C \Rightarrow D) = P(C \cup D) \tag{1}$$

$$\text{Confidence}(C \Rightarrow D) = P(C|D) \tag{2}$$

An item which has a value of interesting measures above predetermined threshold is said to be the frequent item. Apriori algorithm is one of the most popular methods for finding a frequent pattern [8].

Traditional methods of ARM have many limitations such as, (1) problem of sharp boundary (2) cannot provide association between numeral attributes (3) significance of the item is not being considered, only frequency of the item is taken into account (4) lack of strong interesting measures. Most of the real-world data consist of quantitative attributes. So, it is very essential to device appropriate approach that can be dealt with quantitative values.

The structure of the paper is as follows. Section 2 represents existing work and theories. Section 3 presents designed algorithm. Section 4 gives a complete analysis proposed algorithm. Section 5 concludes a paper and presents direction for the future.

2 Related Work

The various techniques for frequent pattern mining are described below.

2.1 Quantitative Association Rule Mining

Quantitative association rule contains quantitative attribute such as age, income, etc. In this method, numerical attributes are converted into nominal values by using predefined concept hierarchies, for e.g,

$$\text{age}(X, ``30-40") \bigwedge \text{income}(X, ``50\,K-60\,K") \Rightarrow \text{invest}(X, ``GOLD")$$

After converting numeral attributes to the categorical attributes, it can be handled in a same manner like traditional methods of ARM [1]. It only considers presence and absence of the item rather than considering significance of the item [2].

2.2 Fuzzy Association Rule Mining

In the recent years, computational intelligence has been arising as a solution to overcome the problem of searching for optimal intervals, inducing rules and also

modeling quantities with fuzzy sets. It also provides facility of linguistic label which makes the rules more interpretable. General form of fuzzy association rule is given as below

$$\text{If} <P, X> \text{ then } <Q, Y>$$

Where P and Q are disjoint itemsets and X and Y are fuzzy sets. It follows the same procedure like Apriori algorithm. In this method, both significance and frequency of the item are being considered for generation of rule.

2.3 Fuzzy Membership Function

Given an item $i_j \in I$ with value v and fuzzy sets $\{x_1, \ldots, x_n\}$, the degree of membership α is given by a function $\alpha = f(v)$, where the function is either a triangular, trapezoidal, etc. It transforms larger values into smaller values. In this item can be a member of more than one membership function. It provides the significance of an individual item. It also gives the graphical representation of each item. Membership degree is in the range of interval [0, 1] as shown in Fig. 1. When membership value is 0 it means there is no membership and value 1 represents the full membership [9, 10].

2.3.1 Calculation of Membership Degree

A triangular membership function is calculated as shown in Fig. 2.

2.3.2 Quantitative Dataset

Transactional dataset M is initially transformed into a quantitative dataset M_Q with quantitative transaction $T_q = \{t1_q, t2_q \ldots, tn_q\}$ and quantitative attributes.

Fig. 1 Triangular membership function [5, 6, 12]
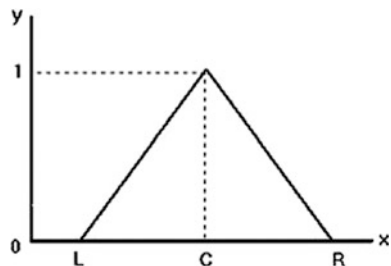
Fig. 2 Equations of triangular membership function [11]

$$
\text{triangle}(x; a, b, c) = \begin{cases} 0, & x \le a. \\ \frac{x-a}{b-a}, & a \le x \le b. \\ \frac{c-x}{c-b}, & b \le x \le c. \\ 0, & c \le x. \end{cases}
$$

Table 1 Transactional dataset

TID	Record
1	{<A,{1,3,2}>, <B,{4,3,1}>, <C,{6,2,5}>}
2	{<B,{4,3,1}>, <D,{3,6,1}>}
3	{<B,{4,3,1}>, <C,{6,2,5}>}
4	{<A,{1,3,2}>, <D,{3,6,1}>}

Table 2 Quantitative dataset

TID	P	Q	R
1	3.67	2.67	2.67
2	3.5	4.5	1.0
3	5.0	2.5	3.0
4	2.0	4.5	1.5

$$
\text{Quantitative Attribute Value} = \frac{\text{Sum of each attribute of all Item in Transaction}}{\text{Transaction Length}} \quad (3)
$$

In transactional dataset items are given with value of their quantitative attributes as shown in Table 1.

A transformed dataset with the quantitative attribute value is shown in Table 2.

A quantitative dataset is transformed into fuzzy dataset by using overlapping triangular membership functions as shown in Table 3.

2.3.3 Fuzzy Support

Fuzzy support $FS(P)$ of an itemset P is a number that measures the degree of membership $m(P, v)$ of the itemset for all values v with respect to the set of transaction in T [3]. This is calculated as

Table 3 Fuzzy transaction dataset with membership

FTID	P			Q			R		
	Low	Med	High	Low	Med	High	Low	Med	High
1	0	0.88	0	0.22	0.44	0	0.22	0.44	0
2	0	1	0	0	0.33	0.33	0.66	0	0
3	0	0	0.66	0.33	0.33	0	0	0.66	0
4	0.33	0	0	0	0.33	0.33	1	0	0

$$FS(P) = \frac{\sum_{j=1}^{n} m_j(x_j, v)}{|T|} \tag{4}$$

Fuzzy support for a rule is denoted as $FS(P \rightarrow Q) = FS\ (P \cup Q)$

$$FS(P \rightarrow Q) = \frac{\sum_{j=1}^{n} \prod m_j(x_j, v)}{|T|} \tag{5}$$

2.3.4 Fuzzy Confidence

Fuzzy confidence of rule is calculated as below

$$FC(P \rightarrow Q) = \frac{FS(P \cup Q)}{FS(P)} \tag{6}$$

2.4 Certainty Factor

To complement support-confidence framework, new measures for evaluation of association rules were proposed. Among these, certainty factor is regarded as one of the most effective measures. It ranges between $[-1, 1]$. Certainty factor provides the statistical correlation between the items in the frequent item [4, 7].

$$CF(P \rightarrow Q) = \frac{\text{Conf}(P \rightarrow Q) - \text{Sup}(Q)}{1 - \text{Sup}(Q)}$$

If Conf $(P \rightarrow Q) > \text{Sup}(Q)$

$$CF(P \rightarrow Q) = \frac{\text{Conf}(P \rightarrow Q) - \text{Sup}(Q)}{\text{Sup}(Q)}$$

If Conf $(P \rightarrow Q) \leq \text{Sup}(Q)$

$$\begin{aligned} CF(P \rightarrow Q) &= 1; \quad \text{If} \quad \text{Sup}(Q) = 1 \\ CF(P \rightarrow Q) &= -1; \quad \text{If} \quad \text{Sup}(Q) = 0 \end{aligned} \tag{7}$$

3   Proposed Algorithm

The generation of interesting rule is given by the following steps:

Input: Quantitative dataset, Fuzzy support, Fuzzy confidence, Certainty Factor
Output: Fuzzy Association Rule
Procedure:
Conversion of raw dataset to quantitative dataset
For each transaction in dataset
    For each item in the transaction
       For each quantitative attribute of each item
       Take summation quantitative attribute
       Divide it by a transaction length
      End
    End
End
Conversion of quantitative dataset to fuzzy dataset
For each transaction in quantitative dataset
    For each quantitative value in transaction
      Determine ranges for each quantitative attribute
      Apply fuzzy triangular membership function
    End
End
Determination of fuzzy frequent itemset
For each transaction in fuzzy dataset
Calculate fuzzy support for each item
Selection of candidate-m itemset (m=1)
For each item in candidate-m itemset
When m>1 make combinations of items in candidate-m itemset
If support of item > min fuzzy support
    Add it to frequent-m itemset
else
    Discard itemset
    Increment m
    End
End
Generation of fuzzy association rule:
For frequent-m itemset (where m>1)
Generate all possible association rules
Calculate support and confidence of association rules
Calculate certainty factor of association rules
If certainty factor of rule > min certainty factor
    Strong association rule
else
    Weak association rule
    Discard weak association rule
End

4 Experimentation

To exhibit the efficacy of the research work, we have carried out many experiments on a transactional dataset. The dataset is a synthetic dataset which is generated by using random function of Java. The dataset is a transactional dataset containing 5000 records, 150 items, and maximum transaction length is 25. For the purpose of the experiment we take 30 quantitative attributes. The experiment is performed on core 2 duo processor with 4 GB RAM.

Experiment 1: (Measures of Quality)

Experiment in the first phase compares the proposed approach with the traditional fuzzy ARM approach with respect to number of the rules generated by individual approaches. In proposed approach, certainty factor is taken as a quality measure to generate interesting rules. Here minimum support is 0.15. In traditional fuzzy ARM approach fuzzy confidence is taken as a quality measure to generate interesting rules. From the Fig. 3 it is clearly shown that certainty factor provides better performance than the fuzzy confidence.

Another comparison is done on the basis of the operator selected for the calculation of the fuzzy support as shown in Fig. 4. In proposed approach, mul operator is selected for the calculation of the fuzzy support. Min and max operators are other options for the calculation.

Experiment 2: (Measures of Performance)

In the second phase of experiment, the influence of the number of transaction and number of items on the execution time is investigated with the support threshold of 0.2 and certainty factor value to 0.20. Figure 5 clearly shows that execution time varies linearly with the number of transactions.
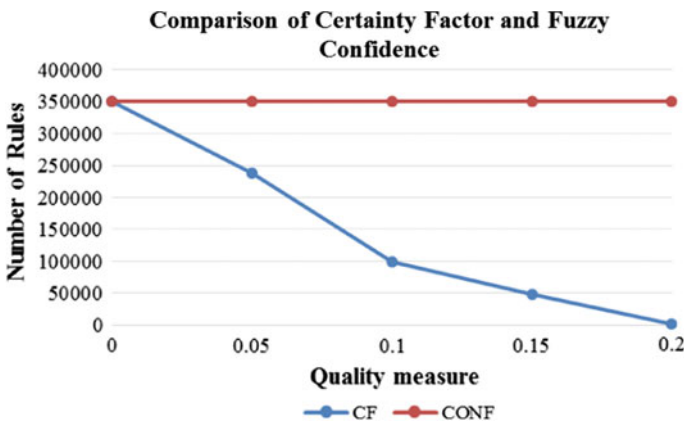
Fig. 3 Comparison of certainty factor and fuzzy confidence
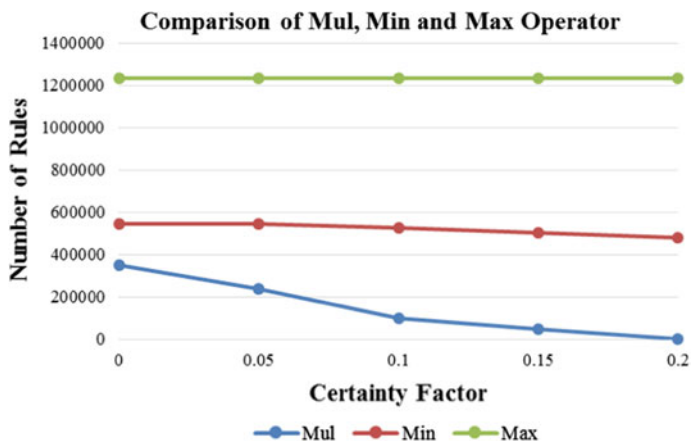
Fig. 4 Comparison of mul, min, and max operators
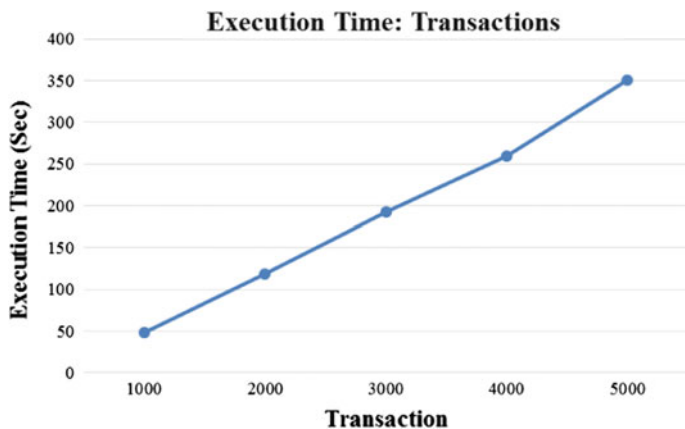
Fig. 5 Execution Time: Transactions

Figure 6 clearly shows that execution time varies linearly with increase in number of items.

5 Conclusion

As per research and from generated results we conclude that it solves some of the problem of the quantitative association rule mining. It generates fuzzy association rule with the linguistic labels which provide better interpretability. It also considers the significance of the item to generate association rules. It also resolves sharp
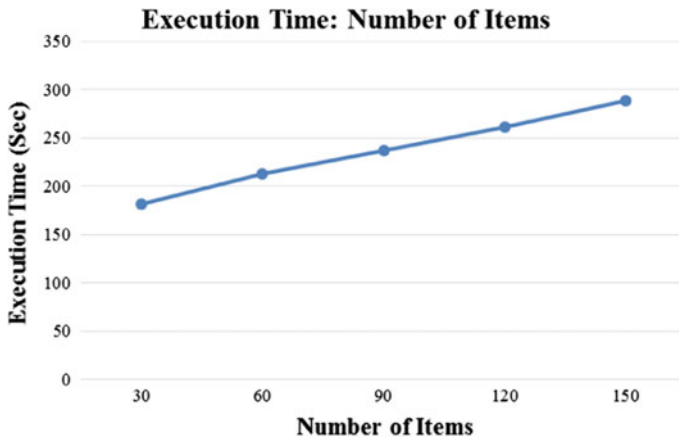
Fig. 6 Execution Time: Number of Items

boundary problem by using overlapping membership function. From experiment it is observed that it generates less number of interesting association rules which are desirable for analysis purpose. In future work, design a more sophisticated system that provides more effective partitioning techniques and generates the strongest interesting association rules.

References

1. Li, D., Zhang, M., Zhou, S., Zheng, C.: A new approach of self-adaptive discretization to enhance the apriori quantitative association rule mining. In: 2nd International Conference on Intelligent System Design and Engineering Application, pp. 44–47. IEEE (2012).
2. Rathod, A., Dhabariya, A., Thacker, C.: An Approach to Mine Significant Frequent Patterns by Quantity Attribute. In: 4th International Conference on Communication Systems and Network Technologies, pp. 414–418. IEEE (2014).
3. Muyeba, M. K., Lewis, S. J. G., Han, L., Keane, J.: Understanding Low Back Pain Using Fuzzy Association Rule Mining. In: IEEE International Conference on Systems, Man, and Cybernetics, pp. 3265–3270. IEEE (2013).
4. Watanabe, T., Fujioka, R.: Fuzzy association rules mining algorithm based on equivalence redundancy of items. In: IEEE International Conference on Systems, Man, and Cybernetics, pp. 1960–1965. IEEE (2012).
5. Dash, S. K., Mohanty, G., Mohanty, A.: Intelligent air conditioning system using fuzzy logic, International Journal of Scientific & Engineering Research, vol. 3, (12) (2012) 1–6.
6. Das, T. K., Das, Y.: Design of a Room Temperature and Humidity Controller using Fuzzy Logic, American Journal of Engineering Research, vol. 2, (11) (2013) 86–97.
7. Watanabe, T.: An improvement of fuzzy association rules mining algorithm based on redundancy of rules. In 2nd International Symposium on Aware Computing, pp. 68–73, IEEE. (2010).
8. Han, J., Kamber, M., Pei, J.: Data mining: concepts and techniques. Morgan Kaufmann, San Francisco (2006).

9. Ross, T. J.: Fuzzy logic with engineering applications. John Wiley & Sons. United Kingdom (2004).
10. Rajasekaran, S., Pai, G. V.: Neural networks, fuzzy logic and genetic algorithm: synthesis and applications. PHI Learning Pvt. Ltd., New Delhi (2012).
11. Fuzzy Logic Membership function, http://www.bindichen.co.uk/post/AI/fuzzy-inference-membership-function.html.
12. Application of Fuzzy controller, http://dimes.lins.fju.edu.tw/pub/Fzysas-96j/Fzysas96.htm.

Virtual Machine Migration: A Green Computing Approach in Cloud Data Centers

Minu Bala and Devanand

Abstract A recent fast growing development and demand in high performance computing has brought IT technocrats on forefeet to devise energy aware mechanisms so that CO_2 emission can be reduced to a great extent. The resources in cloud data centers are always over provisioned in order to meet the peak workload. These resources consume a huge amount of energy, if used in their full capacity. By dynamically adopting the green computing policies as per current workload, the energy consumption of cloud data center can be reduced. In the present study, VM migration process has been discussed and the simulation driven results for evaluation of the proposed heuristic on the basis of static upper and lower limits allowed for CPU utilization have been presented. The comparative analysis of resource utilization and power consumption of a data center, with and without migration policy, reveals that a significant amount of power consumption can be reduced by VM migration and utilization of resources can be optimized.

Keywords Cloud computing · Energy consumption · Energy efficiency · Green computing · Virtualization · Live migration of virtual machines

1 Introduction

In the past few decades, there is an enormous growth in cloud data centers due to demand for high performance computing created by various scientific and compute-intensive applications. These computing infrastructures consume large amounts of energy leading to increase in overall operational costs of the infrastructure [1, 2]. The functionality of equipments and services involved in ICT

Minu Bala (✉)
Department of C.Sc, G.G.M. Science College, Jammu, J&K, India
e-mail: ind_minu@yahoo.com

Devanand
Department of C.Sc. & I.T, Central University of Jammu, Jammu, J&K, India
e-mail: devanandpadha@gmail.com

depends on electricity and a significant portion of electricity production is dependent on nonrenewable energy sources (like coal, oil, etc.). Therefore, ICT contributes in global warming and green house gas emissions. According to Gartner's study, devices and services involved in ICT contributes 2 % in global CO_2 emissions [3]. These data centers are further composed of large number of computing and communicating resources and produce enormous heat when operational. Large cooling systems are, therefore, required for these vast infrastructures, which add a lot to the energy consumption of whole cloud computing environment. There is need of devising new innovative ideas of designing modern data centers with new architectural designs and intelligent cooling systems, low power consuming hardware resources, dynamic workload distribution policies to balance load, power management policies, etc., to reduce the energy consumption while keeping the SLAs and customers interest intact. The focus of this work is on energy and performance efficient resource management policies that can be applied in a virtualized data center by a cloud provider. An adaptive threshold-based approach has been used in the proposed heuristic for dynamic VM consolidation. The study is organized into four main sections. Section 2 discusses the related work. Section 3 throws some light on virtualization and virtual machine migration approaches. Section 4 discusses about the proposed method of VM migration and methodology used in the experimentation work. In Sect. 5, experimental setup and results obtained during experimentation have been discussed. Final section presents the conclusion on the basis of the results obtained in Sect. 6.

2 Related Work

Jung et al. [4] have studied the problem of dynamic consolidation of VMs executing multi-tier web application using live migration while meeting SLA requirements. They have devised a new method of placing VM from source to destination machine using bin-packing algorithm. They made use of gradient search techniques for locating the new place for VM. The limitation of this new technique is its limitation to serve multitenant infrastructure as a service environment.

Zhu et al. [5] have proposed a new automated integrated system for capacity and workload management using resource controllers working on different time scales and scopes. These controllers place compatible workloads onto server and do redistribution and consolidation of VMs so as to fulfill the imposed SLAs. However, static thresholds are not good for environments with varying workloads that exhibit dynamic resource utilization patterns [6].

Elmroth et al. [7] have suggested new technology independent interfaces and architectural designs for virtual machine migration in case of federation of clouds. Different interoperability issues in federation of clouds have been discussed. They have suggested that different cloud service providers involved in federation of clouds have equal opportunity of providing virtualized services remotely or locally.

Beloglazov et al. [8] have suggested technique for dynamic consolidation of VMs based on adaptive utilization thresholds. They analyzed their proposed dynamic virtual machine consolidation mechanism by carrying out their experimental work based on the workload samples from about 1000 servers from Planet Lab. The suggested method ensures to satisfy the customers' requirements to a great extent.

Beloglazov et al. [9] have observed in their work that energy efficiency and QoS like issues in cloud data centers are closely related to each other. They have developed an energy efficient virtual machine consolidation policy for live migration of VMs so that underutilized hosts can be made passive or in inactive mode to reduce the power usage while ensuring the QoS unaffected.

Voorsluys et al. [10] have studied the effect of live virtual machine migration technique on applications running on Xen-based virtual machines. They have come up with an idea that computing overhead due to migration of VMs is very critical in scenarios where SLAs are of major concern. The systems where SLAs are not as important, the live migration overhead is within acceptable range.

3   Virtualization

Virtualization is a software-based technique that creates an abstraction layer known as virtual machine monitor (VMM) or hypervisor which is responsible for gathering the physical details of hardware, thereby creating multiple virtual machine instances. It is responsible for defining and supporting execution of virtual machines. This technology optimizes the resource utilization of cloud data centers by managing the VMs as per the workload.

3.1   Virtual Machine Migration Approaches

Virtual machine migration technique is used in cloud data center for performing various activities like power management, system maintenance, resource sharing, load balancing, etc. In this technique, the virtual machine monitor is responsible for migrating the VM's current working state from source machine to destination machine. The migration process can be a live one or nonlive. The application service is not suspended in case of live migration where as in case of nonlive migration, the application service is paused at source machine, VM state is copied at target machine, and then the service is resumes at target machine. The live VM migration process can be categorized as precopy virtual machine migration, post-copy virtual machine migration, and hybrid virtual machine migration.

In pre-copy method, the whole memory of the VM (which is to be migrated) is captured as memory pages and marked as dirty pages [11]. These dirty pages

logging is transferred to target machine in several rounds. The dirty pages which are recently updated and marked dirtied in the last round at the source machine are also shifted simultaneously with the dirty pages logging to the target machine. This process continues until some termination criterion is reached. After the termination of copying process, the current state of the VM, along with any leftover dirty pages, is sent to the target machine and then VM is resumed at the new location.

In post-copy method, the execution of VM (which is to be migrated) is suspended, its minimum state is captured, transferred to the target machine, and then its execution is resumed at target machine [12]. The memory pages which are requested by application running at new location are pushed from source machine to target machine's local storage and are then served to the application. The link with the source machine is detached only when whole memory of VM is updated at local storage of target machine.

4 Proposed Heuristic for VM Migration

The virtual machine migration process involves three major decisions for migrating VMs in cloud data center. When to migrate? Which VM to migrate? Where to migrate? In the present study, CPU utilization has been taken as the decision parameter for virtual machine migration process. CPU utilization of host machine varies with time because the applications running on VMs face dynamic workload over time. Overloaded machines get heated and also affect QoS. Redistribution of workload by VM migration can optimize the resource utilization. VMs on underutilized hosts can be consolidated by migrating them to other hosts in the data center, thereby, switching off few hosts and reducing overall power consumption. The proposed heuristic is based on static upper and lower CPU utilization thresholds for redistribution and consolidation of VMs in cloud data center. A number of simulators are available as open source to support energy aware experimentation with various cloud data center frameworks. CloudReports, a CloudSim-based simulator has been used in the present study to conduct performance evaluation of different cloud scenarios through simulation. It is very efficient tool to simulate energy aware cloud environment. It provides an environment in which researchers can test its policies like scheduling resources, balancing workload at different levels, etc.

5 Performance Evaluation

The simulation setup for present study consists of a data center with four hosts each having four processing units, number of VMs = 32, average image size = 1000, average RAM = 512 MB, average bandwidth = 100,000, cloudlets sent per

minute = 200, average length of cloudlets = 50,000, average cloudlet's file size = 500, average cloudlet's output size = 500. The upper and lower thresholds are set as 0.8 and 0.2, respectively. The simulation runs have been made for two scenarios, S1 for data center without any migration policy and S2 for data center with proposed VM migration policy.

5.1   Results from Simulation

Table 1 below depicts the results obtained after implementing the proposed VM migration policy based on static CPU utilization thresholds. Overall, 10 VM migrations have been taken place in the scenario under study, four migrations are of distribution type and six migrations are of consolidation type. Here, distribution type migration means that VMs from hosts, running above upper CPU utilization threshold limit, are migrated to the other Hosts in the data center and consolidation type migration means that VMs from hosts, running below lower CPU utilization threshold limit, are migrated to the other hosts so that resources and energy can be saved. In Table 1, Dist. is for distribution, Consol. is for consolidation.

Figures 1 and 2 depict the overall CPU utilization in a data center without VM migration policy and with proposed migration policy, respectively. Figures 3 and 4 depict the overall power consumption in a data center without VM migration policy and with proposed migration policy, respectively. It is clear from the graphs that results obtained after implementing migration policy are better than the results obtained without implementing VM migration policy.

Table 1 VM migration description and details

Mig. ID	Mig. type	Mig. details	Source host		Target host	
			CPU Utilz. in %	P.C. in %	CPU Utilz in %	P.C in %
0	Dist.	VM0 from H0-H1	90	98.125	40	82.50
1	Dist.	VM1 from H0-H1	90	98.125	40	82.50
2	Dist.	VM2 from H0-H1	90	98.125	40	82.50
3	Dist.	VM3 from H0-H2	90	98.125	10	73.125
4	Consol.	VM6 from H0-H1	20	76.25	10	73.125
5	Consol.	VM7 from H0-H1	20	76.25	10	73.125
6	Consol.	VM8 from H0-H1	20	76.25	10	73.125
7	Consol.	VM13 from H0-H1	20	76.25	10	73.125
8	Consol.	VM14 from H0-H1	20	76.25	10	73.125
9	Consol.	VM15 from H0-H1	20	76.25	10	73.125
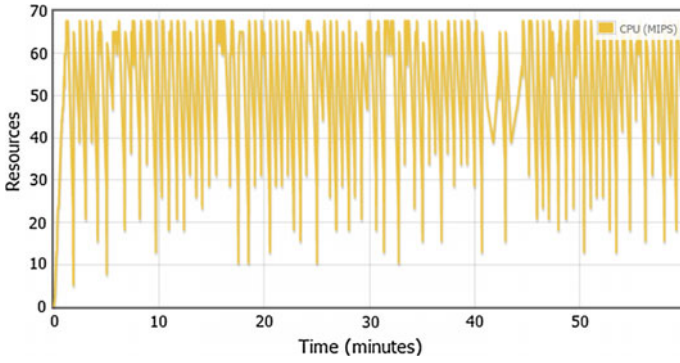
Fig. 1 Overall CPU utilization of hosts in a data center without VM migration policy
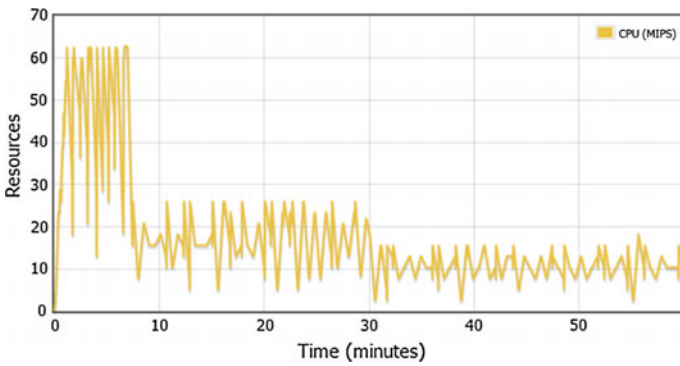
Fig. 2 Overall CPU utilization of hosts in a data center with VM migration policy
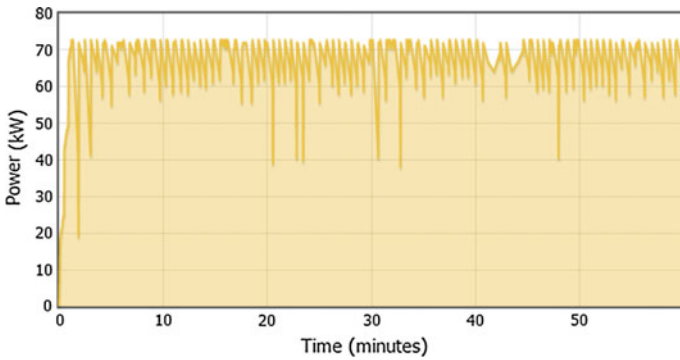
Fig. 3 Overall power consumption in data center without VM migration policy
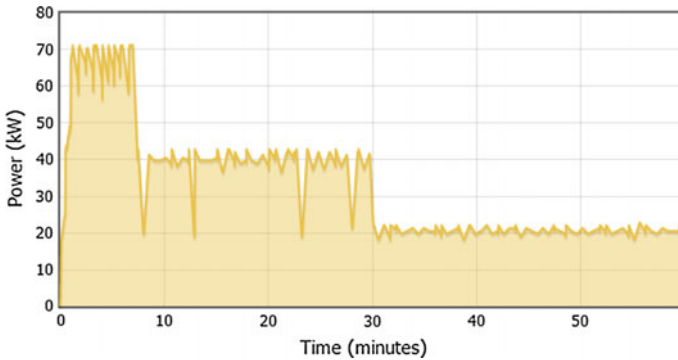
Fig. 4 Overall power consumption in data center with VM migration policy

5.2 Comparative Analysis

The CPU utilization of Host0, Host1, Host2, and Host3, in scenario S1, is approximately 95 %, 95 %, 85 %, and 0 %, respectively, as per the workload considered for experimentation. Power consumption of Host0, Host1, and Host2, in scenario S1, is almost same and is approximately 98 % and is 0 % for host3 as per the workload considered in the present study. With the proposed policy, the CPU utilization and power consumption of the data center decreases significantly. In scenario S2, initially, the CPU utilization of Host0, Host1, Host2, and Host3 is 93.75 %, 93.75 %, 63.54 %, and 0 %, respectively, but migration of few VMs decreases CPU utilization to 93.75 %, 44.79 %, 0 %, and 0 %, respectively. At second time frame, few more migrations decreases the CPU utilization of Host0, Host1, Host2, and Host3 to 62.50 %, 52.08 %, 0 %, and 0 %, respectively, and in the similar way, to 41.67 %, 35.42 %, 0 %, and 0 % and finally to 0 %, 72.92 %, 0 %, and 0 %. The power consumption of hosts in the data center also decreases in scenario S2. The power consumption of Host0, Host1, Host2, and Host3 is initially as 6123.00 W, 5664.75 W, 2778.75 W, and 0 W, respectively. It decreases to 3498 W, 3484.75 W, 0 W and 0 W for Host0, Host1, Host2, and Host3, respectively.

6 Conclusion

In the present study, the important role of live virtual machine migration in dynamic resource management of virtualized cloud systems has been discussed and a upper lower utilization Thresholds heuristic for dynamic reallocation of VMs to minimize energy consumption of cloud data center, has been proposed and evaluated. The results demonstrate that energy consumption of the data centers can be minimized by dynamic migration of virtual machines from over and underutilized hosts to the

other hosts, thereby shutting down few hosts in the data centers and reduce overall energy consumption of the data center. The utilization of resources in the data center is also optimized by adopting proposed VM migration policy. In future, it is proposed to investigate further on the dynamic CPU utilization threshold migration heuristics to optimize the results and minimize the percentage of SLA violations.

References

1. Vaquero, L.M., Merino, L.R., Caceres, J., Lindner, M.: A Break in the Clouds: Towards a Cloud Definition. In: ACM SIGCOMM Computer Communication Review, vol. 39(1), pp. 50–55 (2009).
2. Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., Brandic, I.: Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility, Future Generation Computer Systems. In: Elsevier Science, vol. 25(6), pp. 599–616. Amsterdam, The Netherlands (2009).
3. Pettey, C.: Gartner estimates ICT industry accounts for 2 percent of global CO_2 emissions, http://www.gartner.com/it/page.jsp?id=503867.
4. Jung, G., Joshi, K. R., Hiltunen, M. A., Schlichting, R. D., Pu, C.: A cost-sensitive adaptation engine for server consolidation of multitier applications. In: 10th ACM/IFIP/USENIX International Conference on Middleware, pp. 1–20, Urbana Champaign, USA (2009).
5. Zhu, X., Young, D., Watson, B.J., Wang, Z., Rolia, J., Singhal, S., Mckee, B.: Integrated capacity and workload management for the next generation data center. In: 5th International Conference on Autonomic Computing, pp. 172–181, Chicago, USA (2008).
6. Jung, G., Joshi, K. R., Hiltunen, M. A., Schlichting, R. D., Pu, C.: Generating adaptation policies for multi-tier applications in consolidated server environments. In: 5th IEEE International Conference on Autonomic Computing, pp. 23–32, Chicago, IL, USA (2008).
7. Elmroth, E., Larsson, L.: Interfaces for Placement, Migration, and Monitoring of Virtual Machines in Federated Clouds. In: 8th International Conference on Grid and Cooperative Computing, pp. 253–260 (2009).
8. Beloglazov, Buyya, R.: Adaptive Threshold-Based Approach for Energy- Efficient Consolidation of Virtual Machines in Cloud Data Centres. In: 8th International Workshop on Middleware for Grids, Clouds and e-Science, pp. 1–6 (2010).
9. Beloglazov, Buyya, R.: Energy Efficient Allocation of Virtual Machines in Cloud Data Centres. In: 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, pp. 577–578 (2010).
10. Voorsluys, W., Broberg, J., Venugopal, S., Buyya, R.: Cost of Virtual Machine Live Migration in Clouds: A Performance Evaluation. In: 1st International Conference on Cloud Computing, pp. 254–265, Beijing (2009).
11. Perez-Botero, D.: A Brief Tutorial on Live Virtual Machine Migration From a Security Perspective. Princeton University, USA (2011).
12. Mishra, M., Sahoo, A.: On Theory of VM Placement: Anomalies in Existing Methodologies and Their Mitigation Using a Novel Vector Based Approach. In: 4th International Conference on Cloud Computing. pp. 275–82 (2011).

Detecting Myocardial Infarction by Multivariate Multiscale Covariance Analysis of Multilead Electrocardiograms

L.N. Sharma and S. Dandapat

Abstract In this work, multiscale covariance analysis is proposed for multilead electrocardiogram signals to detect myocardial infarction (MI). Due to multiresolution decomposition, diagnostically important clinical components are grossly segmented at different scales. If multiscale multivariate matrices are formed using all ECG leads and subjected to covariance analysis at wavelet scales, covariances change from normal as MI evolves. This is due to the underlying pathology which is seen in few ECG leads. To capture the changes that occur during infarction, multiscale multivariate distortion metric is applied on covariance structures. To evaluate the proposed method, data sets are adopted from PTB diagnostic ECG database. This includes healthy control (HC), myocardial infarction in early stage (MIES), and acute myocardial infarction (AMI). The results show that the proposed method can detect the pathological MI subjects. For MI detection, the accuracy, the sensitivity, and the specificity is found to be 80, 76, and 84 %, respectively. The proposed method is simple and can be easily implemented for offline analysis for diagnosis of infarction using multiple leads.

Keywords ECG · Multilead ECG · Myocardial infarction · Covariance · Multiscale distortion

1 Introduction

Standard 12-lead electrocardiogram contains the relevant diagnostic information of the heart. Looking at these leads, a physician tries pathological conclusion. Each lead views the heart at a unique angle and it is sensitive to that particular region of the heart. Inferior view is represented by leads II, aVF, and III. Lead-I and lead-aVL

L.N. Sharma (✉) · S. Dandapat
Indian Institute of Technology, Guwahati 781039, India
e-mail: lns@iitg.ernet.in

S. Dandapat
e-mail: samaren@iitg.ernet.in

give left lateral view. The chest leads (V1, V2, V3, V4, V5, and V6) record the electrical potential of the cardiac musculature beneath the electrode. This gives information about small cardiac disorder in the ventricles mainly in the anterior ventricular wall [1]. The electrical information captured by these leads may carry pathological information of different regions of the heart. The acute myocardial infarction may be diagnosed looking at the signals in different leads. The reason behind a myocardial infarction (commonly known as heart attack) is the occlusion of one of the coronary arteries [1]. Due to progressive atherosclerosis in coronary arteries, myocardium loses perfusion and is deprived of oxygen and other nutrients. The multilead recordings of heart potentials due to infarctions may deviate from its standard characteristics. To detect MI several attempts were made, such as time-domain method [2, 3] frequency-domain ST-segment analysis [4], wavelet transform-based method [5, 6], and neural network approach [7–9]. Most of these methods employ modeling techniques, by training and testing the system. Also, most of them are based on few electrocardiogram (ECG) leads. These methods consider only important ECG components rather than entire ECG segments, such as ST-segment, ST-T complex, etc. There are only few studies carried-out for detection of myocardial infarction by monitoring all the 12-leads over time. Thus, there is a need to develop an algorithm to detect the MI using all standard clinical leads. The success of multiscale multivariate analysis [10, 11] motivates to develop a computerized method for detection of MI. In this work, an algorithm is developed for detecting MI using 12-lead ECG which analyzes the covariance structures at wavelet scales.

2 Method

During an acute myocardial infarction, the ECG evolves typically in stages; (1) T-wave peaking (hyperacute T-wave) followed by T-wave inversion (Fig. 1), (2) ST-segment elevation (Fig. 1) and (3) appearance of pathological Q-waves (Fig. 1) [6, 12, 13]. During an infarction, any one of these changes may be present. In Fig. 1, signals recorded from a patient with inferior myocardial infarction are plotted and marked in the pathological segments. The leads II, III, and aVF in (Fig. 1) reflect electrocardiogram changes (Fig. 1b, c, f, marked with circles) associated with inferior infarction. ST-elevation in lead-III is more than lead-II with negative T-wave. Also, ST-segment is depressed in lead-I and lead-aVL (Fig. 1a, e, marked with arrows).

2.1 Preprocessing

To remove the noise due to base line drift and wandering, a moving average filtering approach was employed [14]. In order to remove the higher frequency
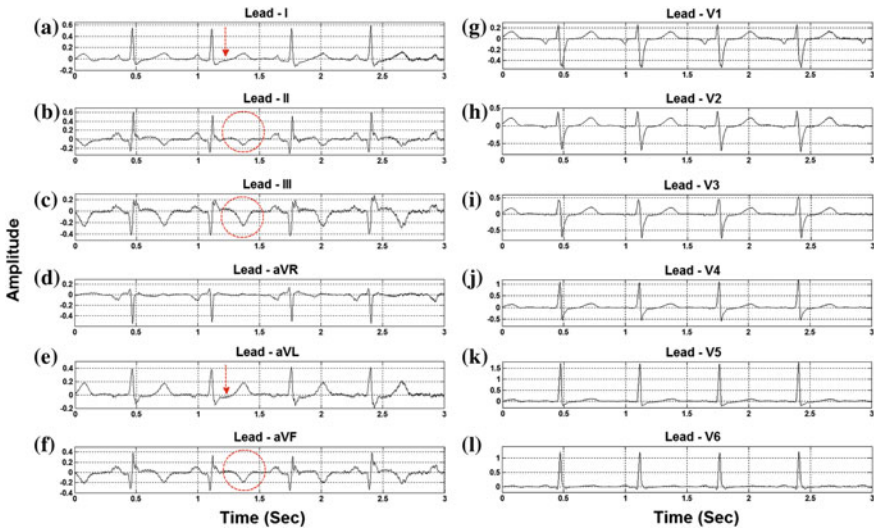
Fig. 1 Multilead ECG signal with Inferior myocardial infarction, database PTB Diagnostic ECG Database (PTBDB), and dataset-patient021/s0073lrem

noise a denoising algorithm based on relative energies of wavelet subbands and estimated noise variance is used to denoise the ECG signals [15]. After that, R peaks of the ECGs were detected using Tompkins's algorithm [16]. From detected R peaks, signal is segmented into frames each having length of four ECG beat, such that the frame covers most of the characteristic of an ECG cycle.

2.2 Multiscale Covariance Analysis

Multiresolution analysis of an ECG signal with L level decomposition [11] using wavelet transform gives kth wavelet coefficient at jth level, $W_{j,k}$. It results in $L + 1$ subbands, an approximation subband at level L and details subband at level j, where $j = 1, 2 \ldots L$. For all n number of frames, numbers of wavelet coefficients are equal at each wavelet scale [11]. So, wavelet coefficients obtained with L level wavelet decomposition are arranged in $L + 1$ subband matrices for each frame. Each column in a subband matrix corresponds to subband of one channel and each row represents coefficients. The approximation subband matrix of ith frame is denoted as A_L^i and details subband matrices are D_j^i. Wavelet decomposition of an ECG signal grossly segments its clinical components into different subbands. Thus, multiscale matrices contain different parts of information segmented from original signals. Higher-order wavelet subband matrices contain more signal energy [11] whereas lower order subband matrices show very low energy contribution. For a six level wavelet decomposition, there is an approximation subband, cA_6 and six detailed subbands,

cD_6, cD_5, cD_4, cD_3, cD_2, and cD_1. It is noticed that cA_6 subband reflects the parts of low-frequency components such as P- and T-wave. cD_6 and cD_5 subbands show the lower frequency part of QRS-complexes with higher frequency part of T-wave. cD_4 and cD_3 subbands show the significant higher frequency part of QRS-complexes. cD_2 and cD_1 subbands contain some higher frequency part of QRS-complexes and noise. Taking advantage of this feature, multiscale matrices are subjected for covariance analysis [17] which may reflect the changes of 'PQRST' morphologies. Myocardial infarction is characterized by specific waves or segments, Q-wave, T-wave, and ST-elevation or depression [18] in the ECG beats. Thus, changes in 'PQRST' morphologies are expected to affect the covariance structures at wavelet scales. The infarction detection is considered in two classes as infracted and healthy (non-infarcted). After forming frames for each cases subband matrices, (A_L and D_j), are subjected for covariance analysis. The covariance matrices for each frame from mean removed data are evaluated as

$$C_{A_L}^i = \frac{1}{(N_L - 1)} \left([A_L^i][A_L^i]^T \right) \tag{1}$$

$$C_{D_j}^i = \frac{1}{(N_j - 1)} \left([D_j^i][D_j^i]^T \right) \tag{2}$$

where $C_{A_L}^i$ is the covariance matrix at Lth approximation scale and $C_{D_j}^i$ is the covariance matrices at jth details scale. N_L and N_j are the number of coefficients at approximation and details, respectively, for ith frame. It can be applied for proposed algorithm to observe the changes that occurs when ECG evolves through different stages (Fig. 1). After covariance analysis, it is essential to capture the changes that occur due to infarction, in different leads. The multiscale multivariate distortion (MMD) [19] can capture the changes in matrices. It is defined as

$$\varepsilon_{A_L^i} = \frac{\left\| C_{A_L}^i - C_{A_L}^{i-1} \right\|}{\left\| C_{A_L}^i - \bar{C}_{A_L}^i \right\|} \tag{3}$$

$$\varepsilon_j^i = \frac{\left\| C_{D_j}^i - C_{D_j}^{i-1} \right\|}{\left\| C_{D_j}^i - \bar{C}_{D_j}^i \right\|} \tag{4}$$

where, $\| . \|_2$ is the matrix 2-Norm (Spectral Norm). $\bar{C}_{A_L}^i$ and $\bar{C}_{D_j}^i$ are mean values. It is expected that this measure can capture the changes occurred in ECG leads over time when infarction evolves. A suitable threshold on MMD can detect the signal with infarction and can classify as normal or healthy.

2.3 Performance Measure

The performance of classification is measured in terms of sensitivity, specificity and accuracy. Sensitivity relates to the ability of the test to identify positive results [20]. That is the percentage of correctly identified MI. Sensitivity (SE) is calculated as

$$SE = \frac{TP}{TP + FN} \tag{5}$$

where *TP* and *FN* represent the number of true positives and false negatives. Specificity relates to the ability of the test to identify negative results [20]. That is, the percentage incorrectly identified as MI. Specificity (*SP*) is calculated as

$$SP = \frac{TN}{TN + FP} \tag{6}$$

where *TN* and *FP* represent the number of true negatives and false positives. The accuracy of a measurement system is the degree of closeness of measurements of a quantity to that of its actual (true) value. Classification accuracy (Acc) is defined as

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

Performance of the proposed method is evaluated using results in a confusion matrix. A confusion matrix may be helpful in summarizing the actual and predicted classifications. It contains information about actual and predicted classifications by a system.

3 Results and Discussion

The standard 12-lead ECG signals are taken from PTB diagnostic ECG database [21]. The number of the total subjects is 294 in the database. The ECG data sets used in the analysis are healthy control (HC) and acute myocardial infarction (AMI). The multilead signals in the database are digitized at 1000 samples per second. The datasets patient105/s0303lre, patient021/s0073lrem, and patient026/s0088lrem are Healthy, Inferior MI with T-wave inversion, and Inferior MI with pathological Q-wave, respectively. They are subjected to amplitude normalization and mean removal. The frames are formed through preprocessing. Each frame contains four ECG beat. Wavelet decomposition using Daubechies 9/7 biorthogonal wavelet

filters up to six levels are used for each channel data for each frame. To capture the changes, covariance matrices for healthy and inferior MI cases are evaluated. In Tables 1 and 2, covariances for Healthy and MI cases are shown for A_6 matrix. In matrix A_6, covariances normally capture the T-wave morphology [22]. The values of covariances of leads II, III, aVF with lead-I are 4.126, 0.935, and 2.946, respectively, for healthy subject (Table 1). The covariances of leads III and aVF with lead-II are 2.640 and 3.952, respectively. The covariance of lead-III and aVF is 2.894. Due to infarction (with T-wave inversion), values of covariances of leads II, III, aVF with lead-I have changed to −3.758, −1.842, and −1.498, respectively (Table 2). The values of covariances of leads III and aVF with lead-II have become 0.168 and 1.351. Similarly, covariance of lead-III and lead-aVF is 0.552. It is seen that the covariance structures are changed exactly opposite in leads which were infarcted (Table 2). Other subband matrices are also subjected to covariance analysis. But no significant changes are found as in A_6 matrix. To quantify the changes in covariance values, an error measure is essential. The changes in covariance structures due to change in ECG morphologies are evaluated using multiscale multivariate distortion.

Figure 2 shows the plot of 'MMD versus Frame' for all wavelet subband matrices for dataset patient021/s0073lrem (Inferior MI with T-wave inversion). The peak MMD was observed in the subband matrix A_6 when the signal evolves from normal to infarction (Fig. 2a). The subband matrices D_6 and D_5 show a small change in MMD values. This may be due to the part of T-waves and QRS-complexes which are grossly segmented during wavelet transform and are appearing in subbands cD_6 and cD_5. Remaining other subband matrices show negligible MMD values. A threshold based on maximum peak is applied to classify the signal as MI.

3.1 Testing with Healthy Control and Synthetic ECG

In Fig. 3, changes in MMD values with each frame are evaluated for healthy subject (patient105/s0303lre) and simulated synthetic multilead ECG signals. These data does not have ST-elevation, peaked T-wave, pathological Q-wave, or inverted T-wave. Since the changes occur due to pathology in A_6, D_6, and D_5 subband matrices, 'MMD versus Frame' plots are shown for these matrices only. For healthy subject, the changes in MMD values are negligible (Fig. 3a). For synthetic ECG, the changes are almost zero (Fig. 3b). This proves that the infarction for the MI subject can be captured by looking at the changes in MMD values. So, a simple threshold is applied to classify the MI and non-MI subjects (healthy). The accuracy, sensitivity, and specificity for classification of subjects are found to be 80, 76, and 84 %, respectively. The classification results in the form of confusion matrix are shown in Table 3.

Table 1 Covariance Matrix of A6 of Healthy Person with Normal T-wave

Leads	I	II	III	aVR	aVL	aVF	V1	V2	V3	V4	V5	V6
I	5.13	4.12	0.93	−5.27	2.86	2.94	−4.30	−2.08	0.19	2.87	3.89	5.02
II	4.12	4.58	2.64	−5.13	0.93	3.95	−4.65	−3.23	−1.61	1.39	2.59	4.24
III	0.93	2.64	2.98	−2.28	−1.52	2.89	−2.56	−2.68	−2.58	0.84	−0.09	1.21
aVR	−5.27	−5.13	−2.28	6.04	−1.98	−4.13	5.26	3.23	1.04	2.30	−3.62	−5.30
aVL	2.86	0.93	−1.52	−1.98	3.09	−0.06	−1.10	0.50	2.00	2.60	2.76	2.59
aVF	2.94	3.95	2.89	−4.13	−0.06	3.66	−3.96	−3.14	−2.09	0.50	1.56	3.14
V1	−4.30	−4.65	−2.56	5.26	−1.10	−3.96	4.76	3.22	1.50	1.55	−2.76	−4.40
V2	−2.08	−3.23	−2.68	3.23	0.50	−3.14	3.22	2.84	2.18	0.04	−0.96	−2.33
V3	0.19	−1.61	−2.58	1.04	2.00	−2.09	1.50	2.18	2.61	1.43	0.89	−0.12
V4	2.87	1.39	−0.84	−2.30	2.60	0.50	−1.55	−0.04	1.43	2.35	2.66	2.71
V5	3.89	2.59	−0.09	−3.62	2.76	1.56	−2.76	−0.96	0.89	2.66	3.31	3.80
V6	5.02	4.24	1.21	−5.30	2.59	3.14	−4.40	−2.33	−0.12	2.71	3.80	5.00

Table 2 Covariance Matrix of A6 of Inferior MI Person with T-wave Inversion in Leads II, III, and aVF

Leads	I	II	III	aVR	aVL	aVF	V1	V2	V3	V4	V5	V6
I	7.65	-3.75	-1.84	-6.16	6.55	-1.49	-4.29	-0.35	4.20	7.00	7.68	7.46
II	-3.75	2.34	0.16	3.36	-2.69	1.35	-2.76	-0.71	-1.78	3.52	-3.99	-3.96
III	-1.84	0.16	1.53	0.98	-2.35	0.55	0.05	-0.72	-1.42	1.55	-1.51	-1.35
aVR	-6.16	3.36	0.98	5.19	-4.92	1.62	3.90	0.65	-3.19	5.69	-6.33	-6.21
aVL	6.55	-2.69	-2.35	-4.92	6.17	-0.63	-2.96	0.28	3.90	5.90	6.33	6.07
aVF	-1.49	1.35	0.55	1.62	-0.63	1.05	-1.65	-0.74	-0.48	1.48	-1.78	-1.83
V1	-4.29	-2.76	0.05	3.90	-2.96	-1.65	3.52	1.11	-1.80	4.00	-4.65	-4.66
V2	-0.35	-0.71	-0.72	0.65	0.28	-0.74	1.11	0.77	0.26	0.38	-0.63	-0.73
V3	4.20	-1.78	-1.42	-3.19	3.90	-0.48	-1.80	0.26	2.64	3.87	4.09	3.90
V4	7.00	3.52	1.55	5.69	5.90	1.48	-4.00	-0.38	3.87	6.49	7.11	6.91
V5	7.68	3.99	-1.51	-6.33	6.33	1.78	-4.65	-0.63	4.09	7.11	7.86	7.68
V6	7.46	3.96	-1.35	-6.21	6.07	1.83	-4.66	-0.73	3.90	6.91	7.68	7.52
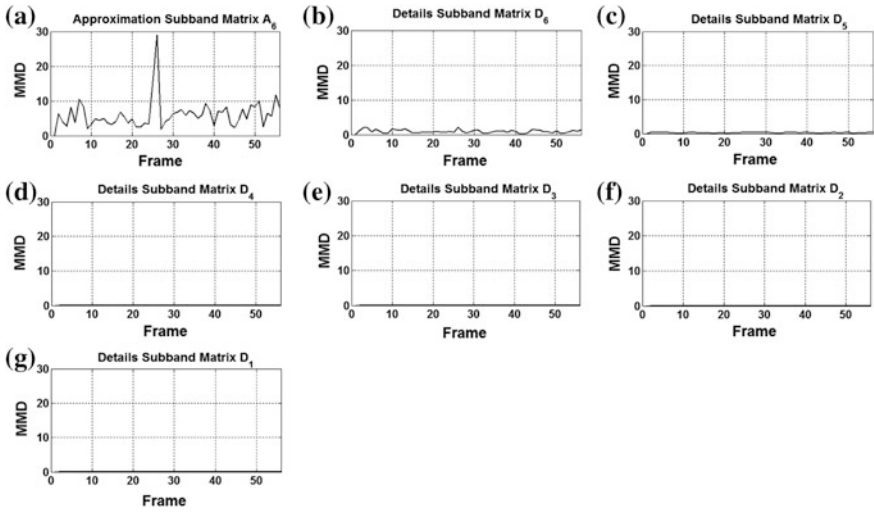
Fig. 2 MMD for wavelet subbands (Decompositions up to six levels) PTB multilead ECG database, Dataset-patient021/s0073lrem-MI
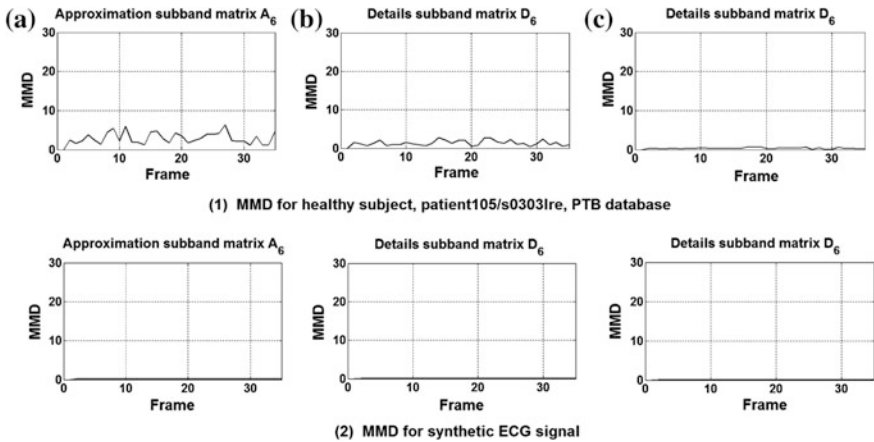
Fig. 3 MMD for wavelet subbands (MMD for wavelet subbands (Decompositions up to six levels) PTB multilead ECG database, Dataset-patient021/s0073lrem-Healthy

Table 3 Confusion matrix for detection of MI

Original class	Predicted class	
	Inferior MI	Healthy
Inferior MI	19	6
Healthy	4	21

Table 4 Comparison of proposed method with existing method for detection of MI

Metric	Reddy et al.	Lu et al. (%)	Heden et al. (%)	Proposed method (%)
Sensitivity	79	84.6	95	76
Specificity	97 %	90.0	86.3	84
Accuracy	NA	NA	NA	80

3.2   Testing with Healthy Control and Synthetic ECG

The proposed method to detect myocardial infarction is simple and can be easily implemented for offline analysis. Most of the methods reported in literature are based on existing classifier and neural network. This requires training and testing of the system. In Table 4, the results found using reported algorithm for MI detection is compared with other methods. Results reported in this paper are comparable to the results of MI detection in the literature.

4   Conclusion

The work presented in the paper discusses a new MI detection method. The detection method is proposed by noninvasive, readily available 12-lead clinical ECG and the detection is treated in general. There are various type of anatomic categories of infarctions (localization). Based on the affected areas and occlusion, there are inferior infarction, lateral infarction, anterior infarction, and posterior infarction or combinations of two like anterolateral infarction. With the help of all 12-leads, physicians try to identify the area of infarction (localization). We have not addressed this problem. But encouraging results of this simple and easy technique to detect myocardial infarction can further be extended.

References

1. Guyton and Hall, Text Book of Medical Physiology, Eleventh ed. Elsevier, 2006.
2. P. De Chazal, M. O'Dwyer, and R. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," Biomedical Engineering, IEEE Transactions on, vol. 51, no. 7, pp. 1196–1206, 2004.
3. S. Mitra, M. Mitra, and B. Chaudhuri, "A rough-set-based inference engine for ECG classification," Instrumentation and Measurement, IEEE Transactions on, vol. 55, no. 6, pp. 2198–2206, 2006.
4. F. Badilini, M. Merri, J. Benhorin, and A. Moss, "Beat-to-beat quantification and analysis of st displacement from holter ECGs: a new approach to ischemia detection," in Computers in Cardiology 1992, Proceedings of, 1992, pp. 179–182.
5. C. Papaloukas, D. Fotiadis, and A. Likas, "Automated methods for ischemia detection in long duration ECGs," Cardiovascular Review and Reports, vol. 24, no. 6, pp. 313–319, 2003.

6. Sharma, L. N. and Tripathy, R. K. and Dandapat, S., "Multiscale Energy and Eigenspace Approach to Detection and Localization of Myocardial Infarction", Biomedical Engineering, IEEE Transactions on, vol. 62, no. 7, pp. 1827–1837, 2015.
7. M. Reddy, L. Edenbrandt, J. Svensson, W. Haisty, and O. Pahlm, "Neural network versus electrocardiographer and conventional computer criteria in diagnosing anterior infarct from the ECG," in Computers in Cardiology 1992, Proceedings of, Oct 1992, pp. 667–670.
8. H. Lu, K. Ong, and P. Chia, "An automated ECG classification system based on a neuro-fuzzy system," in Computers in Cardiology 2000, 2000, pp. 387–390.
9. R. R. Bo Heden, Hans Hlin and L. Edenbrandt, "Acute myocardial infarction detected in the 12-lead ECG by artificial neural networks," in Circulation, 1997, pp. 1798–1802.
10. B. R. Bakshi, "Multiscale PCA with application to multivariate statistical process monitoring," AIChE Journal, vol. 44, pp. 1596–1610, 1998.
11. L. N. Sharma, S. Dandapat, and A. Mahanta, "Multichannel ECG data compression based on multiscale principal component analysis," Information Technology in Biomedicine, IEEE Transactions on, vol. 16, no. 4, pp. 730–736, July 2012.
12. F. Morris, J. Eldhouse, W. Brady, and J. Camm, ABC of Clinical Electrocardiography, ser. ABC Series. John Wiley & Sons, 2003.
13. M. Thaler, The Only EKG Book You'll Ever Need, ser. The Only EKG Book You'll Ever Need. Wolters Kluwer, Lippincott Williams & Wilkins, 2010, Available: http://books.google.co.in/books?id=wQfzn-2UIaoC.
14. R. Rangayyan, Biomedical signal analysis: a case-study approach, ser. IEEE Press series in biomedical engineering. IEEE Press, 2002.
15. L. N. Sharma, S. Dandapat, and A. Mahanta, "Multiscale wavelet energies and relative energy based denoising of ECG signal," in International Conference on Communication, Control and Computing Technologies. (ICCCCT), October 2010.
16. W. Tompkins, Ed., Biomedical digital signal processing. New Jersey: Prentice Hall, 1993.
17. I. Jolliffe, Principal Component Analysis, ser. Springer Series in Statistics. Springer, 2002.
18. JP Martinez, R. Almeida, S. Olmos, AP Rocha, and P. Laguna, "A wavelet-based ECG delineator: evaluation on standard databases," IEEE Transactions on Biomedical Engineering, vol. 51, no. 4, pp. 570–581, 2004.
19. L. N. Sharma and S. Dandapat, Multiscale distortion measure for multichannel electrocardiogram signals. Biomedical Engineering, Narosa Publishing House Pvt. Ltd., 2011.
20. D. G. Altman and J. M. Bland, "Diagnostic tests 1: Sensitivity and specificity." Medical Statistics Laboratory, Imperial Cancer Research Fund, London., vol. 308, p. 1552, June 1994.
21. A. L. Goldberger, et al, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," Circulation, vol. 101, no. 23, pp. e215–e220, doi: 10.1161/01.CIR.101.23.e215, 2000.
22. L. N. Sharma, S. Dandapat, and A. Mahanta, "ECG signal denoising using higher order statistics in wavelet subbands," Biomedical Signal Processing and Control, Elsevier, vol. 5, no. 3, pp. 214–222, 2010.

QoS Improvement in MANET Using Particle Swarm Optimization Algorithm

Munesh Chandra Trivedi and Anupam Kumar Sharma

Abstract MANET is the type of communication network in which communicating device can move anywhere. Movability of devices plays a key role when we discussed about the performance of any routing protocol used in MANET. Performance of any MANET routing protocol depends on the values of its parameters. What will be the value of these parameters for which the performance of protocol is optimal? In this work the particle swarm optimization technique is used to select the optimal value of parameters of AODV routing protocol to improve the QoS in MANET. Java programming language is used to implement optimization technique and then output value is used as input into the network simulator tool NS2.35 for majoring the performance of AODV. The experimental results show 70.61 % drop in Average End-to-End delay (AE2ED), 34.06 % drop in Network Routing Load (NRL), and slight improvement (1.81 %) in Packet Delivery Ratio (PDR) using optimal combination of value of parameters.

Keywords MANET · AODV routing protocol · Particle swarm optimization (PSO) · Performance metrics

1 Introduction

Mobile ad hoc network [1, 2] is the class of wireless network in which communicating devices are mobile in nature and there is not a concept of fixed network topology. Nowadays, MANET gets huge popularity due to its excellent characteristics [3] like

M.C. Trivedi (✉)
ABES Engineeing College, Ghaziabad, India
e-mail: munesh.trivedi@gmail.com

A.K. Sharma
NIEC, New Delhi, India
e-mail: mailanupam2000@gmail.com

fast and cost-effective deployment, no need for base station, infrastructureless, self-manageable, etc. On the contrary of these beautiful characteristics MANET also have some challenges like open-shared wireless medium, limited energy (battery power), limited memory, processing power of communicating devices, movability, etc. Movability of devices plays a key role when we discuss the performance of MANET routing protocols. As far as applications of MANET are concerned, MANET is used in large numbers of applications [3] such as defense applications, rescue operations like in the case of tsunami, earthquake, disaster relief operations, etc. Numbers of routing protocols have been proposed for such types of communications and the protocols are classified as table-driven (proactive), on demand-driven (reactive), and hybrid routing protocols. In this paper, AODV (Ad hoc On-demand Distance Vector) routing protocol which is a type of reactive routing protocol [4] has been used. The detail about AODV is given in next section.

The performance of AODV routing protocol is based on the value of its parameters. What value will we use for these parameters so that protocol gives best performance? Various optimization techniques have been proposed such as particle swarm optimization, ant colony optimization, fish optimization, etc. that can be used for this purpose.

Being very large combination of these values, it is difficult to find an optimal combination for better QoS in MANET. Therefore, this paper presents method that used particle swarm optimization technique (PSO) [5] for selecting the optimal values of these parameters to improve the QoS.

The rest of this paper is organized as follows: In Sect. 2 brief overview of previous work is given. Section 3 gives a brief introduction of AODV routing protocols. Section 4 describes optimization techniques. The proposed model is described in Sect. 5. Section 6 gives the list of performance metrics used in this work to evaluate the performance of AODV routing protocol. The details of simulation and performance analysis are given in Sect. 7. Section 8 gives the conclusions of this research work.

2 Literature Review

Alba et al. [6] used genetic algorithm to find the optimal broadcasting strategy to improve QoS in MANET. In this using GA select the preferred configuration of the network so that QoS is enhanced. To improve the QoS of Ad hoc injection network by optimizing its properties, Dorronsoro et al. [7] used six versions of genetic algorithm (GA). Toutouh et al. [8] evaluate GA, simulated annealing (SA), differential evolution (DE), particle swarm intelligence (PSO), and random metaheuristic algorithms to improve the QoS of OLSR protocol by finding the optimal set of parameter values.

To improve the performance and QoS of AOMDV routing protocol, Lobiyal et al. [9] used particle swarm optimization (PSO). In this paper PSO is used to choose the optimal value to the parameters of AOMDV routing protocol so that performance becomes maximum. Garc et al. [10] have successfully implemented GA, SA, PSO, and rand metaheuristic to enhance the performance of AODV routing rotocol in MANET. The most common problem of MANET, i.e., multicast routing, which was efficiently solved by the Cheng et al. [11] using GA. Shokrani and Jabbehdari [12] have developed a routing protocol for MANET that gives the better performance in terms of QoS. This used the ant colony optimization technique to improve that QoS of MANET.

3 AODV Routing Protocol

Ad hoc On-demand Distance Vector (AODV) [13] is the reactive routing protocol. It used in the concept of reactive routing to establish the route, maintain, and recover the route between source and destination. For this purpose AODV used several packets given as: route request packet (RREQ), route reply packet (RREP), route error packet (RERR), route reply acknowledge (RREP-ACK). Whenever any node in the network wants to share some information with other node in the network, first it checks its own catch for fresh enough route to destination. If they have information, used that route for communication. Otherwise source node initiates route formation process by broadcasting the RREQ packet to all of its neighbor's nodes. Job performed by neighbor nodes: if the neighbor node itself is the final destination of the packet then it sends back the RREP packet. Otherwise, first it checks its own catch for fresh enough route to destination. If they have a, it used that route to forward the RREQ packet else it broadcasts RREQ packet to all of its neighbor's nodes. Once the route is successfully formed source node used this route for communication. When intermediate node forwards the RREQ it adds its broadcast ID and address of the previous node in the RREQ packet. The freshness of the route at the intermediate node is determined by comparing the destination sequence number in RREQ with sequence number at intermediate node. Same RREQ packet at the destination may be reached from the different path. The destination node compares the hop count values in these multiple upcoming RREQ packets and it selects the route for sending the RREP packet for which the hop count value in RREQ packet is minimum. If the minimum values for two or more routes are same then decision of selecting route is based on the sequence number. When the RREP is received at the source node then route is established and sender used this route to send the data packets to the destination node. Due to the mobility of the nodes or any other regions, like battery, route may break during the communication. For this purpose AODV used the RERR packet to indicate the error in the path.

4 Particle Swarm Optimization

In computer science, basically PSO is a computational method that optimizes a problem by repeatedly trying to improve a user solution related to given measurement values. It was developed in 1995 by James Kennedy (Social Psychologist) and Russel Eberhart (Electrical Engineer). PSO is an effective optimization technique based on the movement and intelligence of swarm (SWM). Generally, SWM in common words is a cluster, here clusters of user solution. And PSO optimizes a problem by having these clusters of user solutions; here select particles (Prtc) and move these Prtc around the search space by means of Prtc's position and Prtc velocity given in simple mathematical formulae.

PSO uses a number of Prtc that formed an SWM moving around in the SS to find out the optimal solution. Each Prtc is treated as a point in a D-dimensional SS which adjusts its "flying" according to its own flying experience as well as the flying experience of other Prtcs. Whenever improved position are found these will then come to guide the movements of the SWM. This process is repeated so many times and by doing so, it is hoped but not confirmed, that a satisfactory solution will be found. The basic concept of PSO lies in accelerating each Prtc toward its pbest (its own best past location) and gbest (the best past location of the whole swarm or a close neighbor).

ALGORITHM 1: Particle Swarm Optimization

Step 1: Start
Step 2: Initialize PAP[] with RPV on D dimensions in the SS
Step 3: Loop:
Step 4: Calculate OFF in D variables for each particle
Step 5: Compare particles fitness evaluation with its $pbest_i$.
　　　　IF (Current Value of OFF > $pbest_i$) Then Set $pbest_i$ = Current Value and P_i = current value of xi in D dimension space
Step 6: Identify the particle in the neighbourhood with the best success so far, and assign its index to the variable g.
Step 7: calculate new velocity as
$$vi = vi + U(0,\phi1) \times (Pi - xi) + U(0,\phi2) \times (Pg - xi)$$
Step 8: calculate new position as
$$xi = xi + vi$$
Step 9: If a criterion is met, exit loop
Step 10: End Loop
Step 11: End
Abbreviations used in PSO algorithm:
PAP: Population array of particles , RP: Random Positions and velocity, SS: Search Space, OFF: Optimization Fitness Function
Input/ Output of PSO Algorithm
Input: 10 sets of 7 parameters value, Output: optimized 10 sets of 7 parameters value.

Abbreviations used in PSO algorithm

PAP: Population Array of Particles, RP: Random Positions and velocity, SS: Search Space, OFF: Optimization Fitness Function.

Input/Output of PSO Algorithm

Input: 10 sets of 7 parameter values, Output: optimized 10 sets of 7 parameter values.

5 The Proposed Method

In this work, we have proposed the method to improve the QoS in MANET. The proposed model consists of two parts. First part is used for optimization purpose and second for solution evaluation. For optimization purpose, proposed method imposes particle swarm optimization (PSO) technique and network simulator is used for solution evaluation.

PSO is used to find the optimal value of parameters in search space. For this it used fitness function given in Eq. 1. Then result obtained from PSO is passed one by one to the parameter value of AODV routing protocol and then test the performance of AODV protocol with optimized value using the network simulator NS2. For evaluating the performance NS2 setup mobile ad hoc network and configure this with AODV routing protocol. Simulation setup for this configuration is given in next section. Then by applying the performance majoring parameter like AE2ED, NRL, and PDR evaluate the performance of AODV in MANET. Figure 1 shows the working model of proposed model.

Fitness function for this particular problem is defined as follows:

$$\text{fitness} = \text{w1.PDR} - (\text{w2.AE2ED} + \text{w3.NRL}) \qquad (1)$$

This fitness function is defined with the objective to maximize PDR and minimize AE2ED and NRL. Equation 1 is the maximizing function because PDR is
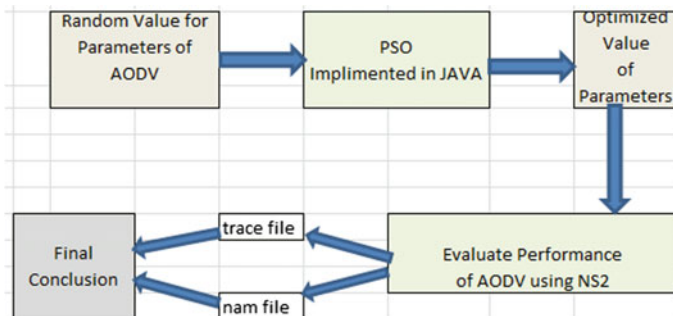
Fig. 1 Working model of proposed method

used with positive sign and AE2ED and NRL are used with negative sign. The variable w1, w2, and w3 are the weight variables that define the impact of performance metrics on the resultant fitness value. For simulation we have taken the values of w1, w2, and w3 as 0.5, 0.25, and 0.2, respectively.

ALGORITHM 2: Mapping of Proposed Method with PSO Algorithm

Step 1: Initialize $x_i = 10$ sets (random sets) of 7 parameter values as input. Here the value of D is 7 and i=1 to 10
Step 2: start loop
Step 3: Calculate OFF of 10 sets (x_i) in the form of communication Overhead (C_i)
Step 4: *In first generation*
    IF ($pbest_i == C_i \&\& P_i == x_i$)
    Then value of xi denote current position of ith set in 7 dimension search space
    For successive generations
    Compare fitness value C_i of x_i with its $pbest_i$.
        IF ($C_i >$ $pbest_i$) Then Set $pbest_i = C_i$ and $P_i =$ current value of x_i in D (7) dimension space
Step 5: The value of i corresponding to minimum pbesti is selected and assigned to g. Where i belongs to 1 to 10
Step 6: *In first generation vi = 0*
    For successive generations calculate new velocity as

$$vi = vi + U(0,\phi 1) \odot (Pi - xi) + U(0,\phi 2) \odot (Pg - xi)$$

    Symbol $\odot$ denotes component wise multiplication. $U(0,\phi 1)$ denotes the random number between 0 and $\phi 1$
Step 7: calculate new position as

$$xi = xi + vi$$

Step 8: If a criterion is met after 50 successive generations, exit loop
Step 9: End Loop

6   Performance Metrics

The following metrics are used to evaluate the performance of the AODV routing protocol in MANET.

A. **Average End-to-End Delay (AE2ED)** This is the average time taken (for transmission from source to destination) by packets that are received successfully at the destination node.
B. **Network Routing Load (NRL)** This is defined as the ratio of administrative routing transmission and data packets delivered. Here, transmission counting is done by counting each hop separately.
C. **Packets Delivery Ratio (PDR)** It is the ratio of number of packets received successfully at the destination and number of packets sent by the source node.

7 Experimental Setup and Result Analysis

The implementation of proposed model is done in two parts, first part that used PSO for optimization is implemented using Java programming language and second part that is used for evaluating the performance of network is implemented using network simulator NS2 version 2.35 [14, 15] on Ubuntu 11.04. These two parts are interlinked using shell programming.

For evaluating the performance of network we have taken three cases with different network sizes. In case 1, case 2, and case 3 network sizes are 50 nodes, 75 nodes, and 100 nodes, respectively. List of seven major parameters of AODV routing protocol which affects the performance of network is given in Table 1. Simulation parameters used to build network for our problem are given in Table 2.

The above experiment is executed 10 times (for 10 different seed values) for the population size of 10. After simulation we get the optimized value of parameters for which the performance of protocol is maximum.

Table 3 shows the result which we get after simulation. In table, the values of PDR, AERED, and NRL are given for default and optimized value of parameters of AODV. The obtained value of parameters shows drastic improvement in QoS compared to the default value of parameters. There is 70.61 % drop in Average End-to-End delay (AE2ED), 34.06 % drop in Network Routing Load (NRL), and

Table 1 AODV parameters

Parameter	Default value	Range
ACTIVE_ROUTE_TIMEOUT	3 s	1–10
ALLOWED_HELLO_LOSS	2	1–10
NODE_TRAVERSAL_TIME	0.4	0.01–1.00
RREQ_RETRIES	2 tries	1–10
TTL_START	1 s	1–10
TTL_INCREMENT	2 s	1–10
TTL_THERSHOLD	7 s	1–20

Table 2 Simulation parameters

Simulator	NS2(v-2.35)
Simulation time	600
Number of nodes	50, 75, 100
Simulation area	800 m × 800 m
Transmission range	250 m
Channel bandwidth	2 Mbps
Propagation model	Two ray ground
Application traffic	CBR
Packet size	512 bytes
Traffic rate	4 packet/s
Pause time	300 s

Table 3 Simulation results

Performance metrics	Default values			Optimized values		
	No. of nodes 50	No. of nodes 75	No. of nodes 100	No. of nodes 50	No. of nodes 75	No. of nodes 100
PDR	0.9	0.79	0.7	0.99	0.81	0.72
AE2ED	0.045	0.72	1	0.03	0.23	0.3
NRL	0.32	0.45	0.6	0.21	0.32	0.39

Fig. 2 Communication overhead
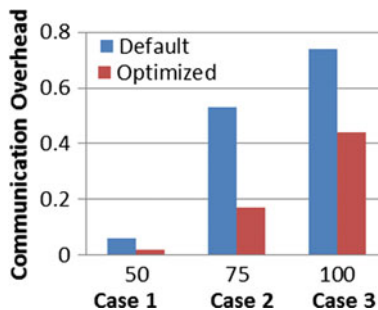
slight improvements (1.81 %) in Packet Delivery Ratio (PDR) using optimal combination of value of parameters of AODV routing protocol.

After obtaining result from simulation we calculate the communication overhead using the fitness function given in Eq. 1 for both default value and optimized value. Then the comparison of communication overhead with default value and optimized value is shown in Fig. 2. Result shows that AODV with optimized value of parameters gives less overhead than with default value of parameters.

8 Conclusion

The performance of routing protocol in MANET depends on the value of its parameters. What will be the value of parameter for with the result will be maximum? For this purpose, proposed model imposed PSO algorithm to select the optimized value of parameters for which the result is better than the performance of protocol with default value. Simulation result shows that proposed method improved QoS in MANET. The obtained value of performance metrics shows drastic improvement in QoS compared to the default value of parameters. There is 70.61 % drop in Average End-to-End delay (AE2ED), 34.06 % drop in Network Routing Load (NRL), and slight improvements (1.81 %) in Packet Delivery Ratio (PDR) using optimal combination of value of parameters.

References

1. S. K. Sarkar, T. Basavaraju, and C. Puttamadappa, "Ad Hoc Mobile Wireless Networks: Principles, Protocols and Applications," Auerbach Publications, MA, USA, 2007.
2. K. V. Arya and S. S. Rajput, "Securing AODV Routing Protocol in MANET using NMAC with HBKS technique," in Proc. of the IEEE International Conference on SPIN, pp. 281–285, 2014.
3. S. S. Rajput et al., "Performance comparison of AODV and AODVDOR routing protocol in MANET," International Journal of Computer Applications (IJCA), vol 63, no 22, pp. 19–24, 2013.
4. J. -H. Song, V. W. Wong, and V. C. Leung, "Efficient on-demand routing for mobile ad hoc wireless access networks," IEEE Transactions on Selected Areas in Communications, vol. 22, no. 7, pp. 1374–1383, 2004.
5. Poli, Riccardo, James Kennedy and Tim Blackwell. Particle swarm optimization, Swarm intelligence 1.1; 2007; pp. 33–57.
6. Alba E, Dorronsoro B, Luna F, Nebro A, Bouvry P, and Hogie L. A Cellular MOGA for Optimal Broadcasting Strategy in Metropolitan MANETs, Computer Communications 2007; 30: pp. 685–697.
7. Dorronsoro B, Danoy G, Bouvry P, and Alba E. Evaluation of different optimization techniques in the design of ad hoc injection networks, in Workshop on Optimization Issues in Grid and Parallel Computing Environments, part of the HPCS, Nicossia, Cyprus; 2008; pp. 290–296.
8. Toutouh J, Garcia-Nieto J and Alba E. Intelligent OLSR Routing Protocol Optimization for VANETs, in IEEE Transaction on Vehicular Technology 2012; 61: pp. 1884–1894.
9. D K Lobiyal, C P Katti, and A K Giri, "Parameter Value Optimization of Ad-hoc On Demand Multipath Distance Vector Routing using Particle Swarm Optimization" in Proc. of ELSEVIER International Conference on Information and Communication Technologies (ICICT 2014), Procedia Computer Science 46 (2015) pp. 151–158.
10. Garc´a-Nieto, Toutouh J, and Alba E. Automatic Parameter Tunning with Metaheuristics of the AODV Routing Protocol for Vehicular Ad hoc Networks, Evo Applications, part II. LNCS 6025; 2010; pp. 21–30.
11. Cheng H and Yang S. Genetic algorithms with immigrant schemes for dynamic multicast problems in mobile ad hoc networks, Eng. Appl. Artif. Intell 2010; 23: pp. 806–819.
12. Shokrani H and Jabbehdari S. A novel ant-based QoS routing for mobile ad hoc networks, Proceedings of the first international conference on Ubiquitous and future network. Piscataway, NJ, USA: IEEE Press; 2009; pp. 79–82.
13. S. Kalwar, "Introduction to reactive protocol," IEEE Potentials, vol. 29, no. 2, pp. 34–35, 2010.
14. The Network Simulator Project—Ns-2, online http://www.isi.edu/nsnam/ns/.
15. S. S. Rajput et al., "Comparative analysis of Random Early Detection (RED) and Virtual Output Queue (VOQ) algorithms in Differentiated Services Network," in Proc. of the IEEE International Conference on SPIN, pp. 237–240, 2014.

Comparing Various Classifier Techniques for Efficient Mining of Data

Dheeraj Pal, Alok Jain, Aradhana Saxena and Vaibhav Agarwal

Abstract With recent advances in computer technology, large amounts of data could be collected and stored. But all this data becomes more useful when it is analyzed and some dependencies and correlations are detected. This can be accomplished with machine learning algorithms. WEKA (Waikato environment for knowledge analysis) is a collection of machine learning algorithms implemented in Java. WEKA consists of a large number of learning schemes for classification and regression numeric prediction. So, by using this we can find out the prediction value of dataset and the data which we stored can be seen in different forms in the form of matrix, graph, curve, tree, etc. In this paper, we are researching or comparing the results of the three classifiers, the classifiers we are using such as J48, Naïve Bayes, and preprocess the data. We compare the results which provide easy way to understand all the datasets and its condition.

Keywords Data mining · Classification · Clustering · Visualization · WEKA · Data preprocessing · Algorithms

Dheeraj Pal (✉) · Alok Jain · Vaibhav Agarwal
Computer Science Engineering, Amity University, Gwalior, MP, India
e-mail: dpal@gwa.amity.edu

Alok Jain
e-mail: ajain1@gwa.amity.edu

Vaibhav Agarwal
e-mail: vagarwal@gwa.amity.edu

Aradhana Saxena
Computer Science Engineering, RJIT, Tekanpur, Gwalior, MP, India
e-mail: aradhana298@gmail.com

1   Introduction

The changes that are constantly taking place in terms of rapid technical and technological developments affect society as a whole. Companies invested in building data warehouse that contains millions of records and attributes. They cannot produce sufficient output due to lack of knowledge, lack of staffs and appropriate tools. In recent years, there has been increasing interest in the use of data mining [1, 2] to investigate scientific questions with in the details of company employees. An area of enquiry about the details of employees, which is influenced by many qualitative attributes such as income of employee, age, marital status, sex, education, etc. [3, 4]. Overall, by this information, we can predict that how many of the employees get the minimum salary, maximum salary, mean and the standard deviation for the salary, and we can calculate how many of the employees are single, married, divorced, and their education like high school, graduation, college, etc. For this, we are applying many techniques to predict how many of them get how much salary and their overall status.

2   Literature Survey

Data mining, containing multiple disciplines, including machine learning, statistics, database systems, information science, visualization, and many application domains, has made great progress in the past decade [4]. Though many methods are proposed to address the issue of imbalanced classification, but still the solutions are problem dependent [5]. As the previous papers had using techniques like rapid miner tool, orange, and weka but most of the results come out from rapid miner tool and orange, and by using these tools they apply techniques such as k-nearest neighbour, Neural network, Fuzzy, Genetic, and others in the environment [3]. This paper describes the comparative analysis and performance of classification techniques [3]. H.2.8 [Database Management]: Database Applications—Data Mining, Spatial databases, and GIS; H.5.1 [Information Interfaces]—in the paper [6] for finding the time series data.

3   Study

3.1   *Hypothesis*

The research objective was to find a very large amount of data which is collected from the local serve areas and run data mining algorithms against it. The popular data mining software is used to evaluate the data tool is weka [7] which provides a superior result [5]. Additionally, the raw data was graphed to visualize [1] the

identified patterns through visual inspection which the mining software might overlook. By mining the data, a trend was expected about the employee's information how many of them come under maximum, minimum, mean, and standard deviation.

3.2 Data Collection and Preparation

The dataset [3, 4] was stored in a table, each type of data given an own column in a table within a database. Table 1, represents the qualitative value which is gathered for performing the experiment. All types of data are stored in the same table so it is easy for the software to calculate the data. Table 1 shows the attributes; there are five attributes income, age, children, marital status, and education and a total of 10 instances.

The data cannot be directly compared to each other because they have different form of values such as income is in numeric form and marital status is in nominal [8, 9]. So, here, we are only calculating the maximum and minimum form. With the help of Table 2, it can define the attributes and the values for predicting the data.

Table 1 Data which is collected for experiment

	A	B	C	D	E
1	Income	Age	Children	Marital status	Education
2					
3	25,000	35	3	Single	Highschool
4	15,000	25	1	Married	Highschool
5	20,000	40	0	Single	Highschool
6	30,000	20	0	Divorced	Highschool
7	20,000	25	3	Divorced	College
8	70,000	60	0	Married	College
9	900,000	30	0	Married	Graduateschool
10	200,000	45	5	Married	Graduateschool
11	100,000	50	2	Divorced	College

Table 2 Collection of qualitative data

Attributes	Values
Income	15,000–900,000
Age	20–50
Marital status	Single, married, divorced
Children	0–5
Education	High school, grad. school, college
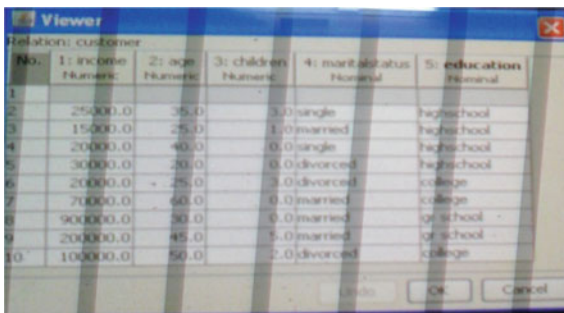
3.3 Data Preprocessing

Data shown in Fig. 1 is processed in the software; it would be quite massive to remake the table for local database. So, a java program was used which can pull data from the database and convert it into data mining format [9]. The program then compiles the result into a weka file form. The file can be read like a table which has its own column. This lets data be efficiently organized and allow for mining techniques.

Now, with the help of this table, we can easily select the attributes as it is saved in the software, i.e., in the "WEKA" tool (Fig. 2).

It can be noted from the related work that the attribute selection plays an important role to identify parameters that are important and significant for an excellent result. Where blue, red, and sky blue colour represents the number of maximum, minimum, and mean values. Blue is for maximum, red is for mean, and sky blue is for minimum by which it is easy to identify that their income, age,

Fig. 1 The below table was readable by software directly

Fig. 2 Preprocessing qualitative data
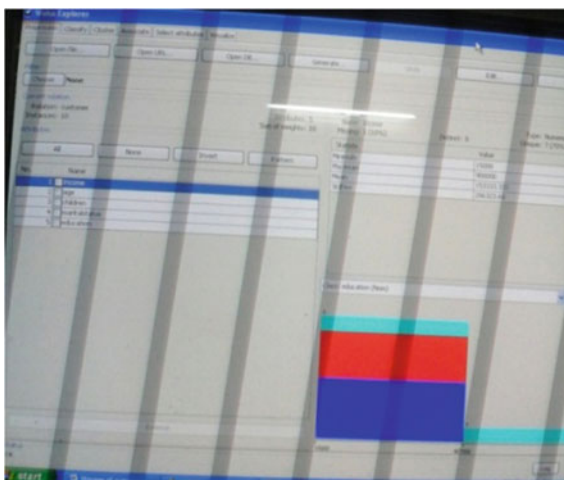
Fig. 3 Visualize the overall dataset
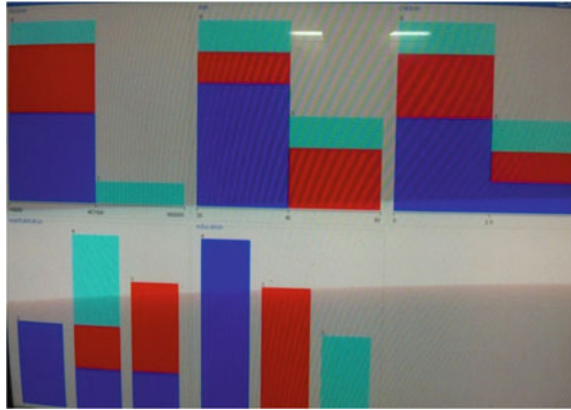
children, marital status, and education lies between the given data. In Fig. 3, we can clearly see all the dataset at a time by selecting the option visualize all the data or all the attributes [1, 6, 10].

4 Methodology of Data Mining Techniques for Classification

4.1 Classification Techniques for Prediction

In an effort to finding patterns, a variety of algorithms are used [9, 11]. There are three main algorithms that provide some form of result and those are

- Classification Rule Algorithm.
- Various Decision Tree Algorithms.
- Naive Bayes.

Many classification [9, 12] models have been proposed by researchers in machine learning, pattern recognition, and in statics. Generally, the classification techniques can follow the two step process which is used to predict the class labels for training data. In classification step, test data are used to estimate the accuracy of classification rules. There are many techniques that can be used for classification. Such techniques are decision tree, Naive Bayes rule, and many others. Figure 4 represents the classification methodology of research process that can be learned from training data which are analyzed from classification algorithm. Test data are used to classify and estimate the accuracy of classification rules.
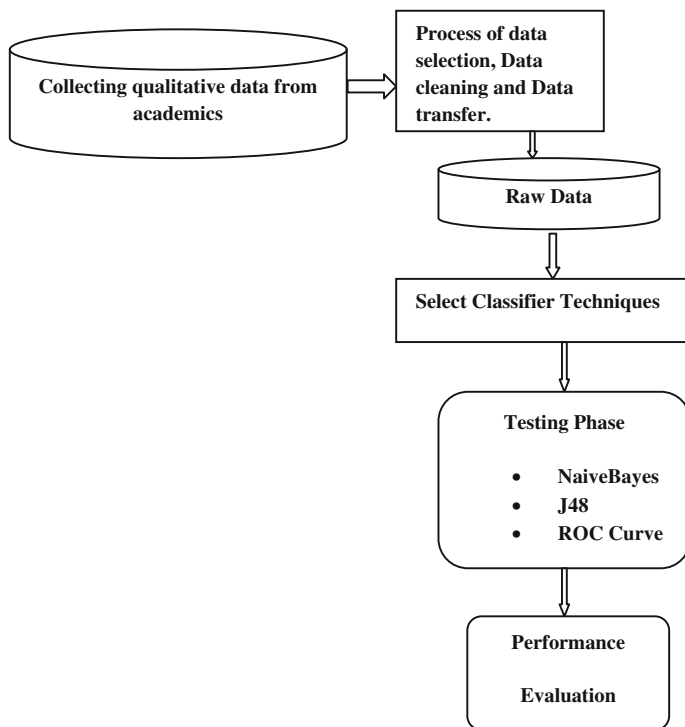
Fig. 4 Methodology of classification technique used in data mining

4.2 Flow Chart

See Fig. 1.

4.3 J48. Tree Classification Technique

J48 is a tree technique which is enhanced by ID3 algorithm. It is one of the most popular tree algorithms in tree classification technique; with the help of this tree, we can determine the number of employees between in which criteria. The objective is to reduce the impurity or uncertainty in data as much as possible (Fig. 5).

Decision tree is very effective in mining the data with the help of this tool it is easy to reduce the bulk of data, and clearly understand the details or the information with the help of this tree.

This algorithm is an extension of ID3 algorithm and possibly create small tree. It uses a divide and conquers approach. Decision tree is closely related to the rule

Fig. 5 "J48 the decision tree" has been built
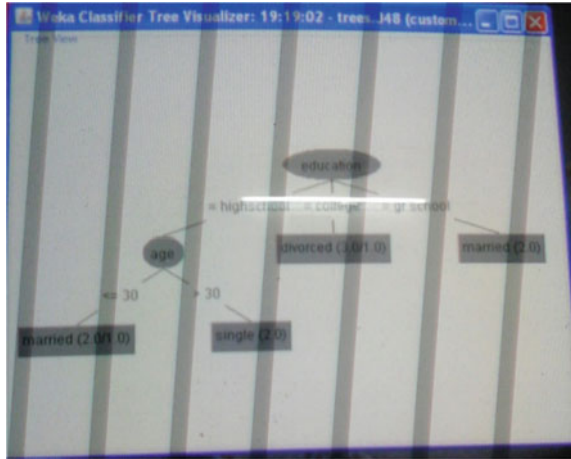
indication. A tree includes root node, i.e., "education and internal node that represents test condition (applied on attributes) a leaf node" married, single, "divorced".

4.4 Naive Bayes Classifier

In probability theory, Naive Bayes classifier checking the condition rule and it can be classified by learning phase and testing phase. Bayesian reasoning is applied to decision-making that deal with probability inference which is used to gather the knowledge of prior events by predicting events through rule base. Once the model has been trained and tested, we need to measure the performance of the model. For this purpose, we use three measures namely: precision, recall, and accuracy.

$$\text{Precision}\,(P) = tp/(tp\,+\,fp)$$

$$\text{Recall}\,(R) = tp/(tp\,+\,fn)$$

$$\text{Accuracy}\,(A) = (tp\,+\,tn)/\text{Total}$$

where tp, fp, tn, and fn are true positive, false positive, true negative, and false negative, respectively.

Figure 6 show the running information with the help of Bayes classifiers.

That while performing Naïve Bayes in this Test Mode is used percentage Split (split 66.0 % default value) (Fig. 7).

Naïve Bayes shows the attributes income, age, children, marital status, and the class is education and its attributes is high school, college, graduation. With the

Fig. 6 Naïve Bayes
classifiers

Fig. 7 Naive Bayes shows
values mean maximum,
minimum, and standard
deviation

help of this technique mean, std. deviation, weight, and precision is calculated for
each and every attribute (Figs. 8 and 9).

Table 3 shows that what proportion of test instances has been correctly and
incorrectly classified.

This is the correct accuracy rate according to the number of instances. Table 4
shows a "confusion matrix".

Fig. 8 Naïve Bayes running image
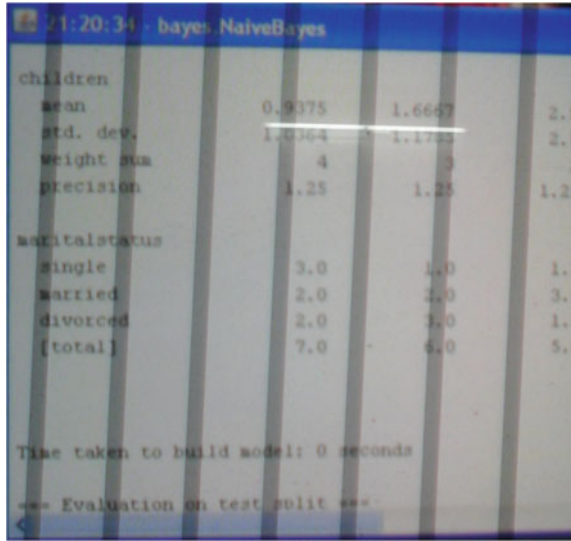
Fig. 9 Shows the "confusion matrix"
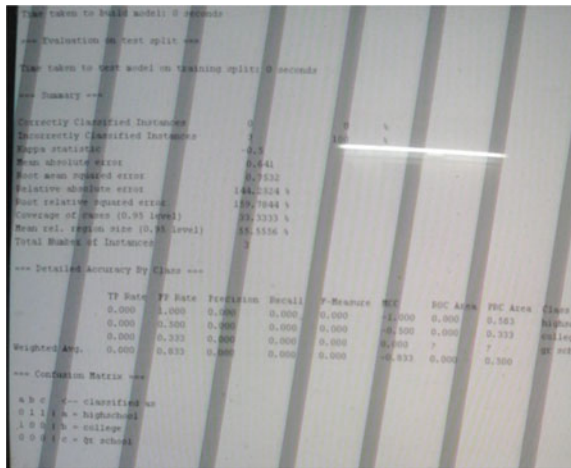
Table 3 Shows the accuracy of instances

Correctly Classified Instances 8 88.8889 %	
Incorrectly classified Instances 1 11.1111 %	

Table 4 Confusion matrix

a	b	c	Classified as
4	0	0	a = high school
1	2	0	b = college
0	0	0	c = Graduate school

4.5 Receiver Operating Characteristics (ROC) Curve

This curve can be described as follows: there are two curves that can be seen together with different attributes and class and it is easy to find out the maximum and minimum instances of the values. Table 2 represents first graph its *X*-axis shows education (nominal values) in which high school, school and graduation, and its *Y*-axis shows income (numeric values) while in next Table 4 shows its *X*-axis represent the marital status (nominal value) *Y*-axis represent the age (numeric value).The jitter of graph first shows education of the candidates. While the jitter of graph two represents children. After all, at the bottom of the graph shows the confusion matrix in both form, i.e., in minimum and in maximum form or we can also say that it also shows according to the cost and benefit (Fig. 10).

The accuracy also shown in Table 5 predicted by the table (a) and by table (b) classification accuracy in prediction (a) is 60.6666 % and classification accuracy for prediction (b) is 30.3333 % shown in Table 6 and the result can also be different according to the population, target value, and score threshold.

With the help of this algorithm, we can see the graph and compare their results at a time. Both graphs can be opened and we can easily change their axis (Tables 7 and 8).

Fig. 10 ROC curve of J48 tree
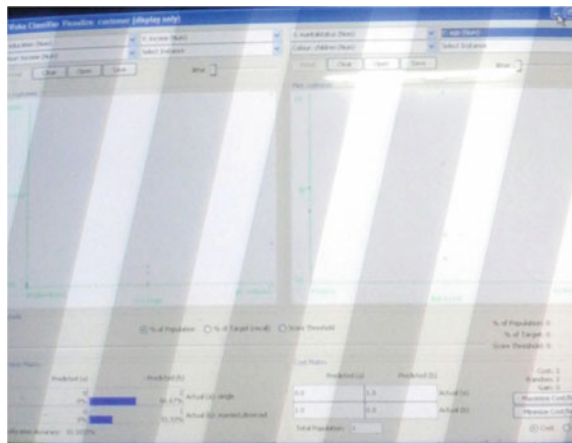
Table 5 Represents graph.1

Axis	Attributes	Colour (marital status)
X	Education	Blue–single
		Red–married
		Black–divorced
Y	Income	

Table 6 Represents graph.2

Axis	Attributes	Colour (education)
X	Marital status	Blue–high school
		Red–college
		yellow–graduate school
Y	Income	

Table 7 Prediction of table (**a**)

Threshold value	0.75
Percentage of population	10
Percentage of target	25
Classification accuracy	60 %

Table 8 Prediction of table (**b**)

Threshold value	1
Percentage of population	10
Percentage of target	20
Classification accuracy	30 %

5 Conclusion and Future Work

In this paper, the research work compares the result of all the algorithms with each other Naïve Bayes, J48 algorithm, ROC curve, preprocessing the data set. So, it is found that the result of J48 tree and the ROC curve is far better or easy to understand compare to Naïve Bayes rule. This study is very helpful to identify the ratio of given or the collected data it can easily calculate the maximum, minimum, mean, and standard deviation of the data. In future, the result of Naïve Bayes multi functional can be improved because while performing the Naïve Bayes it does not make any effective difference by which result can be differentiated. So, in future there is scope that many other data mining algorithms can be added to it.

Method	Model representation	Model evaluation
J48 classifier	Tree, matrixform	Very compact, easy to analyze
Naïve Bayes	Confusion matrix	Posterior probability
Pre-process data	Explicit and indicate all the data	Visualize all the data in bar form
ROC curve	Construction phase, curve, graph	Classify accuracy, error, threshold values.

References

1. Eleonora Brtka, Vladimir Brtka, Visnja Ognjenovic and Ivana Berkovic, "The data visualization technique in e-learning system", IEEE 10th Jubilee International Symposium on Intelligent Systems and Informatics (September 20–22, 2012).
2. P. Nevlud, M. Bures, L. Kapicak and J. Zdralek, Anomaly based Network Intrusion Detection Methods Advances in Electrical and Electronic Engineering, pp. 468–474, (2013).
3. M. Mayilvaganan, D. Kalpanadevi "Comparison of Classification Techniques for predicting the performance Of Students Academic Environment", 2014 India, coimbatore International Conference on Communication and Network Technologies (ICCNT).
4. S. Ummugulthum Natchiar, Dr. S. Baulkani, "Customer Relationship Management Classification using Data Mining Techniques" International Conference on Science, Engineering and Management Research (ICSEMR 2014).
5. Sabri Serkan Güllüoğlu, "Segmenting Customers With Data Mining Techniques", ISBN: 978-1-4799-6376-8/15/©(2015) IEEE.
6. Patricia Morreale, Steve Holtz, Allan Goncalves, "Data Mining and Analysis of Large Scale Time Series Network Data", 2013 27th International Conference on Advanced Information Networking and Applications Workshops.
7. Darshana Parikh, Priyanka Tirkha, "Data Mining & Data Stream Mining—Open Source Tools" International Journal of Innovative Research in Science, Engineering and Technology, Vol. 2, Issue (10, October 2013).
8. Charles A. Fowler and Robert J. Hammell Converting PCAPs into Weka Mineable Data copyright 2014 IEEE SNPD 2014, (June 30–July 2, 2014), Las Vegas, USA.
9. Swasti Singhal, Monika Jena, A Study on WEKA Tool for Data Preprocessing, Classification and Clustering, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-2, Issue-6, (May 2014).
10. C. M. Velu, K. R. Kashwan, "Visual Data Mining Techniques for Classification of Diabetic Patients", Maharashtra, INDIA, 2013 3rd IEEE International Advance Computing Conference (IACC).
11. D. Rajeswara rao, Vidyullata Pellakuri, SathishTallam, T. Ramya "Harika Performance Analysis of Classification Algorithms using healthcare dataset" (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2), 2015.
12. Manisha Girotra, Kanika Nagpal, Saloni Minocha, Neha Sharma, "Comparative Survey on Association Rule Mining Algorithms" International Journal of Computer Applications (0975–8887) Volume 84–No (10, December 2013).

Fingerprint Recognition System by Termination Points Using Cascade-Forward Backpropagation Neural Network

Annu Agarwal, Ajay Kumar Sharma and Sarika Khandelwal

Abstract Fingerprint authentication belongs to one of the oldest biometric systems. This paper defines a new approach for fingerprint recognition. In this paper only termination points of minutiae are used for authentication. This system matches only the fingerprint image with database image when there is 100 % match or more than 90 %. Finally, the neural network approach is applied for measurement of neural network performance. The false accept rate and false reject rate are also defined.

Keywords Biometric · Fingerprint · Cascade-forward backpropagation · Neural network · Matlab

1 Introduction

Biometric recognition is the most oldest and important technology for personal identification [1]. Biometric offers more security than any other traditional method of personal recognition, which is first introduced by Alphonse Bertillon [2]. Biometric authentication is based on physiological and behavioral characteristics. Fingerprint recognition comes in physiological characteristics. The uniqueness of the fingerprint lies on its own characteristic, i.e., minutiae which are unique to everyone [3]. The identical twin does not have the same fingerprint image pattern. The proposed system focuses on termination points that are sufficient for fingerprint authentication. The basic idea of the system is to extract the termination points and

Annu Agarwal (✉) · A.K. Sharma · Sarika Khandelwal
Computer Science & Engineering, Geetanjali Institute of Technical Studies,
Udaipur, India
e-mail: agarwalannu04@gmail.com

A.K. Sharma
e-mail: profsharmaak@gmail.com

Sarika Khandelwal
e-mail: sarikakhandelwal@gmail.com

save it in the database; then the authentication process is applied and the neural network performance is measured. The advantage of using neural network is fast and easy computation. Before the storage of fingerprint image, it is also preprocessed to remove the noise and enhance the image quality. So the database is made with good quality termination points of fingerprint image.

2 Related Work

In fingerprint matching, various methods are used for good authentication results. Various researches are carried out which are defined as follows:

Mohammed and Nyongesa [4] define fingerprint classification system using fuzzy neural network. This method gave good results.

Hsieh and Shing [5] proposed a different method for fingerprint recognition. They used ridge bifurcation for matching process. Their experimental results define that fingerprint minutiae are reliable and robust.

Karu and Jain [6] define fingerprint classification system. In this fingerprint is classified into five classes, i.e., right loop, left loop, arch, tented arch, and whorl. In this paper, singular point extraction is defined. The given experiment is invariant to rotation, translation, and small amounts of scale changes, which gives accuracy of 85.4 % for five classes and 91 % for four classes.

Lu et al. [7] made an algorithm for minutiae extraction. This algorithm is made for improvement performance in fingerprint authentication. Using this algorithm for minutiae extraction, the overall performance of automatic fingerprint identification is increased, and in this method removal of spurious minutiae is done in post-processing.

Annapoorani and Caroline [8] define survey on fingerprint matching using FPR. In this method, the preprocessing steps are used and then the matching process is applied. Usually, a technique called minutiae matching is used to handle automatic fingerprint recognition with a computer system.

Jain et al. [9] trained a multilayer feedforward network using a quick propagation training algorithm. In this experiment, the neural network has 20 neurons in one hidden layer and 192 neurons in input layer. The multilayer feedforward method is used for quick propagation training algorithm. Here, the accuracy is 86.4 % using five classes and 92.1 % using four classes.

Chatterjee et al. [10] define fingerprint identification and verification system by minutiae extraction using artificial neural network. In this, a new method for fingerprint identification technology by minutiae feature extraction using backpropagation algorithm is defined. For an input image, the local ridge orientation is estimated and the region of interest is located. Then a feature extractor finds minutia features such as ridge end, bifurcation, short ridge, and spur from the input fingerprint images and the digital values of these features are applied to input of the neural network for training purpose. For fingerprint recognition, the verification part

of the system identifies the fingerprint-based training performance of the network. Finally, experimental result shows 95 % using artificial neural network.

Kashyap and Yadav [11] define fingerprint matching using neural network. Here, Levenberg–Marquardt backpropagation (LMBP) algorithm is used for training purpose because it is the fastest technique for complex data sets and gives better performance in such situation. Input image is trained by trainlm () function for producing different result sets like performance plot, regression, etc.

Singla and Sharma [12] define fingerprint authentication using artificial neural network. In this approach, a multistage method is used. The whole preprocessing operation is used, and the segmentation of fingerprint is done followed by normalization and thinning. The crossing number is used to extract the minutiae features. This method tackles the variations in the fingerprint images.

Various researches and experiments are used in fingerprint recognition to increase efficiency. In Sect. 3 Methodology is discussed; in Sect. 4 Experiments and Results of this method is defined; and in the Sect. 5 the Conclusion of whole process is explained.

3 Methodology

In this proposed system various stages are defined as follows:

In step I, the image is loaded. The CASIA database is used which is available online or we can acknowledge by "Portions of the research in this paper use the CASIA Fingerprint collected by the Chinese Academy of Sciences' Institute of Automation (CASIA)" [13]. After loading image the rotation of the image is defined. The meaning of rotation is to define that how the image should be rotated in the Matlab. In Matlab the rotation is done in anticlockwise direction using imrotate () function. To rotate an image in clockwise direction the negative (−) sign is used before defining the rotation angle.

In the step II and step III of experiments, the normalization and histogram equalization are applied as preprocessing steps. These preprocessing steps are essential for any pattern recognition.

In the step IV, the Gabor filter is applied for image enhancement. Image enhancement is important because this process improves the image quality. Using the image enhancement process, the digitally stored image can be made lighter or darker [14]. Gabor filters are used which have both frequency-selective and orientation-selective properties and after filtering we get the enhanced image.

In step V, the minor orientation of the image is checked by Hough transformation and after checking the minor rotation, the previous without rotation-oriented image is used for the next steps.

In step VI the binarization is applied which transforms the 8-bit gray image into the 1-bit image.

In the step VII the thinning is applied. Thinning is necessary for any image to remove the redundant pixels and reduction of ridge pattern thickness into a single line [3]. It clearly shows the ridge termination and ridge bifurcation.

3.1 Minutiae Representation and Matching

In step VIII the minutiae extraction is done. In our experiment green color is used for representation of bifurcation and the red is used for ridge termination. For finding minutiae points, first the region of interest is defined which is suppressed into the minutiae. Then the false minutiae points should be removed, and after this the marking is done. In this system the centroid (center of mass) and valid termination points are found and stored in the database.

Then the popup menu is shown in which the different options are visible which are as follows: add to database, compare with database, clear database, and exit option. So for storage the user choose add to database option, and for comparison the compare with database option is used. The user should also delete the whole database and make the new one. When the compare with database option is chosen, the matching percentage with database entry name is defined in the popup menu and the neural network performance is shown.

4 Experiments and Results

All the experiments and results are shown in this section. Figure 1 shows the representation of the ridge termination and ridge bifurcation.

After extraction, the popup menu should be visible to the user and in the Matlab window, the termination points are shown. These visible termination points show that the proposed system works correctly. The flexibility is given to the user and the database is made according to the user requirements.

In the compare with database option, the given image is compared with stored database image. After clicking the compare with database option, the matching percent with the database entry name is defined. In our system the matched with database option is only shown when there is 100 % matching or more than 90 %; otherwise it shows imposter. If the system does not match, then the imposter popup menu is defined with percentage. For measuring the neural network performance, the mean squared error is defined.

Figure 2 shows match percentage with the database entry. The figure shows that the fingerprint image matched with database entry 6 with 100 % matching percent.

False Accept Rate = (No. of Imposter/Total No. Fingerprints) × 100
False Accepts Rate = (1/50) × 100 = 2 %
False Reject Rate = (No. of Genuine/Total No. Fingerprints) × 100
False Reject Rate = (0/50) × 100 = 0.
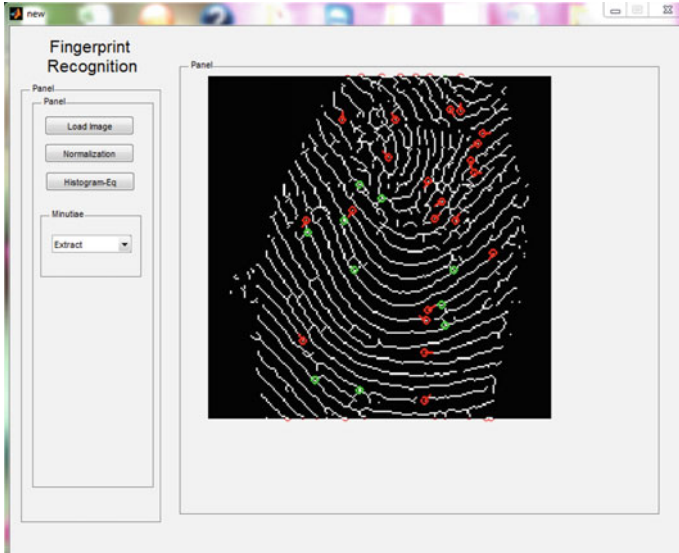
Fig. 1 Representation of ridge termination and bifurcation

Fig. 2 Matching percentage with database entry name
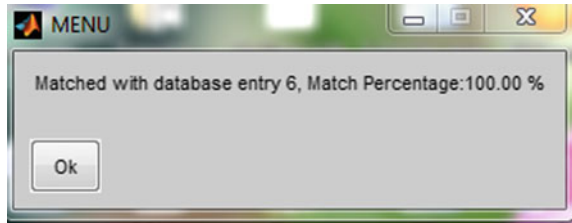
4.1 Neural Network Performance

In this experiment, a two-layer cascade-forward backpropagation neural network has been used and neural network tool of Matlab 2012 is used. In neural network, three kinds of different processes are defined which are as follows:

Training The training is used for adjustment of the network according to the error. The process is very important in neural network. In the training process, the training algorithm is shown. In our experiment, Levenberg–Marquardt backpropagation (LMBP) algorithm is used for training. It is the fastest algorithm in neural network. *Validation* This is mainly used to measure the network generalization. *Testing* It is the main process in the neural network. Using testing process we check for the final solution. The performance of neural network is measured on the basis of mean squared error. Mean squared error is the average squared difference
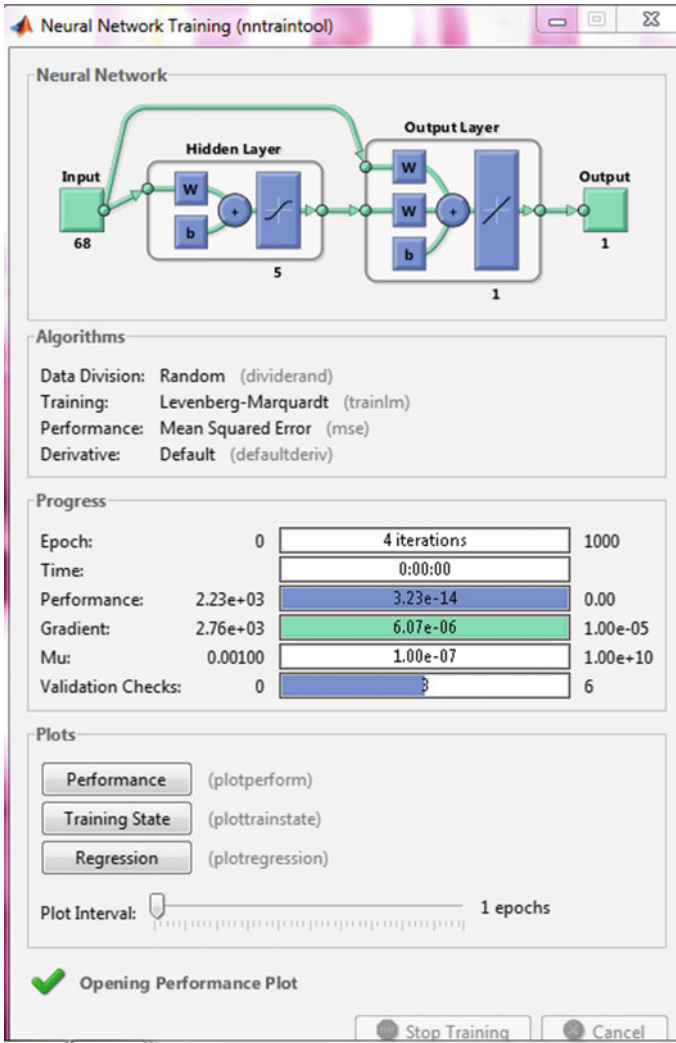
Fig. 3 Neural network training result

between the output and the target values. Lower values of mean squared errors are considered as better one than higher values.

In this experiment, we have taken 1000 epochs and the training was completed in four iterations. The network is trained through cascade-forward backpropagation algorithm. Here, 68 neurons are taken in input layer and the 5 neurons are taken in hidden neurons. In the output layer the neuron should be 1. Figure 3 shows the training process of neural network and Fig. 4 shows mean squared error of testing
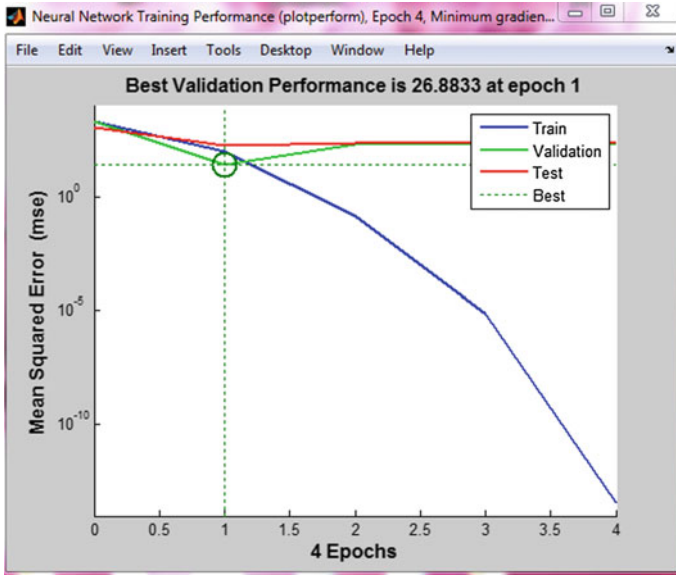
Fig. 4 Neural network training performance

Fig. 5 Imposter matching percentage with database entry name
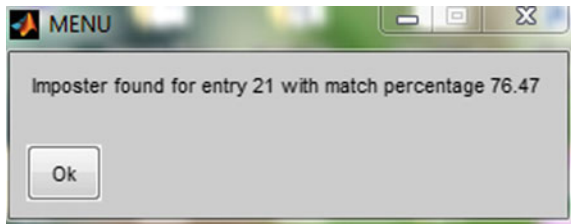
process. The value of mse is low which is good and the value is 26.8833 at the epoch 1.

The validation, training, and testing curves are also shown in Fig. 4.

When the imposter fingerprint image is matched, the imposter popup menu is shown after the termination extraction. Here, we only use termination points for matching process, but we represent both ridge termination and bifurcation points of minutiae. Figure 5 shows the imposter matching percentage with database entry name. In this figure the imposter fingerprint image is matched with database entry 17 and the matching percentage is 76.47.

Like for genuine match, for imposter of the neural network architecture, epoch, the training algorithm name, the value of gradient, and validation checks are

defined. Here, 68 neurons are taken in input layer and the 5 neurons are taken in hidden neurons as shown. In the output layer the neuron should be 1. From the above figure the total number of iterations to train the network is 4 and the performance of neural network is defined in terms of mean squared error, the value of gradient, validation checks, etc.

Figure 6 shows the mean squared error for imposter fingerprint image. For an imposter image the value of mean squared error is 443.1561 at epoch 1.

5   Conclusions

In this work it has been shown that the fingerprint image is only authenticate when there is 100 % match or more than 90 %. When this condition is not fulfilled, then the popup menu shows the imposter message with matching percentage. This experiment is applied to 50 fingerprint images and measured the neural network performance. In our experiment, Levenberg–Marquardt backpropagation (LMBP) algorithm is used for training. From saved fingerprint images we get the lower mean squared error and 100 % matching result. In this, maximum four iterations are taken to train the network and the given lower mean squared error defines the good performance of neural network. In this experiment, we also check for imposter fingerprint image. Imposter fingerprint image can be defined by the image that is not saved in database. In future the work can be done with ridge bifurcation to make it the automate process.

References

1. D. Maltonie, D. Maio, A. K. Jain, S. Prabhakar: Handbook of Fingerprint Recognition, Springer, New York (2003).
2. Anil K. Jain, Arun Ross and Salil Prabhakar: An Introduction to Biometric Recognition. In: IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 1, January (2004).
3. R. Lakshaman, S. Selvaperumal, C. H. Mun: Intergrated Fingerprint Recognition using Image Morphology and Neural Network. In: International Journal Of Advanced Studies In Computer Science And Engg., Vol 3, Issue 1 (2014).
4. Mohamed. S. M, Nyongesa H.: Automatic Fingerprint Classification System using Fuzzy Neural Techniques. In: IEEE International Conference on Artificial Neural Networks, Vol. 1, pp. 358–362 (2002).
5. Ching Tang Hsieh, Chia Shing: Humanoid Fingerprint Recognition based on Fuzzy Neural Network. In: International Conference on Circuit, Systems, Signal and Telecommunications, pp. 85–90 (2007).
6. K. Karu, A. Jain: Fingerprint classification. In: Pattern Recognition Vol. 29, pp. 389–404 (1996).
7. Haiping Lu, Xudong Jiang, Wei-Yun Yau: Effective and Efficient Fingerprint Image Post Processing. In: International Conference On Control, Automation, Robotics And Vision, Vol. 2, pp. 985–989 (2002).
8. Annapoorani. D., Caroline Viola Stella Mery: A Survey Based on Fingerprint Recognition Minutiae Fingerprint Matching using FPR. International Journal Of Science And Research, Vol. 3 Issue 10 (2014).
9. Anil Jain, Salil Prabhakar, Lin Hong: A Multichannel Approach to Fingerprint Classification. In: IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 21, No. 4, pp. 348–359 (1999).
10. Atanu Chatterjee, Shuvankar Mandal, G. M. Atiqur Rahaman, Abu Shamim Mohammad Arif: Fingerprint Identification and Verification System by Minutiae Extraction Using Artificial Neural Network. In: JCIT, Vol. 1, Issue 1 (2010).
11. Kalpna Kashyap, Meenakshi Yadav: Fingerprint Matching Using Neural Network Training. In: International Journal of Engineering And Computer Science, Vol. 2, Issue 6, pp. 2041–2044 (2013).
12. Neeraj Singla, Sughanda Sharma: Biometric Fingerprint Identification Using Artificial Neural Network. In: International Journal of Advanced Research In Computer Science & Technology, Vol. 2, Issue 1 (2014).
13. Biometric Ideal Test. CASIA Fingerprint, http://biometrics.idealtest.org.
14. R.M. Mandi, S.S. Lokhande: Rotation Invariant Fingerprint Identification System. In: International Journal of Electronics Communication and Computer Technology, Vol. 2 Issue 4 (2012).

On the Dynamic Maintenance of Spanning Tree

Isha Singh, Bharti Sharma and Awadhesh Kumar Singh

Abstract The paper presents a centralized heuristic algorithm for the secure and dynamic maintenance of spanning tree in wireless networks. Initially, we construct the minimum spanning tree that models the given network. Later, in order to reflect the topological dynamics in secure manner, we reorganize the minimum spanning tree. The resulting logical structure is a spanning tree; however, it may not be minimum spanning tree. Our findings have been substantiated with simulation results.

Keywords Wireless networks · Spanning tree · Dynamic maintenance

1 Introduction

The use of spanning tree is an established approach to model various static and dynamic networks. It is the induced tree that spans over all the nodes of the network graph under consideration. An induced tree is termed minimum spanning tree (MST) as long as the sum of its edge weights is minimum among all such induced trees. In the literature, there exist many excellent algorithms to compute MST, e.g., GHS algorithm [1], Chin–Ting [2], Gafni [3], Awerbuch algorithm [4], Garay–Kutten–Peleg algorithm [5], Kutten–Peleg algorithm [6], Elkin [7], and Peleg–Rabinovich [8]. Further details and few more algorithms can be found in [9]. However, if the total number of nodes or their positions in the network undergo some change, then the distribution of nodes and links in the network gets amended. Consequently, the corresponding spanning tree is required to be rearranged.

Isha Singh (✉) · Bharti Sharma · A.K. Singh
National Institute of Technology, Kurukshetra, India
e-mail: isha.05.92@gmail.com

Bharti Sharma
e-mail: bharti_kanhiya@yahoo.co.in

A.K. Singh
e-mail: aksingh@nitkkr.ac.in

In wireless networks, particularly involving mobile or sensor nodes, this phenomenon is quite common. Therefore, the algorithm for frequent reorganization of spanning tree, preferably, needs to be local, quick, and message efficient. As the topological changes are frequent, from the practical point of view, in most of the wireless network applications, it is more preferable to have a simple and light method for the dynamic maintenance of spanning tree, rather than the spanning tree being minimum spanning tree (MST), because the dynamic maintenance of MST is heavy [10–13]. Motivated by such observation, we propose a centralized heuristic algorithm that uses the information available and partially reorganizes the existing logical structure to reflect topological changes and ensure connectivity; however, after each topological change, we do not recompute MST. Consequently, the resulting logical structure is spanning tree, though, not necessarily minimum spanning tree. Our dynamic maintenance procedure is simple and easy to implement; also, it avoids the need of frequent MST recomputation.

2 The Basic Idea

We consider a wireless network modeled as an undirected connected graph $G = (V, E)$, where V is the set of vertices (nodes) and E is the set of edges (communication links) between them. Each edge $e \in E$ has nonzero weight w, and each node has unique ID. Any two nodes are called neighbors if they are one hop away from each other and communicate directly. Also, we assume that despite multipath effect and varying channel conditions the message propagation between neighbor nodes is FIFO. We aim to collect the entire graph information at one node and use this information to create minimum spanning tree. The entire graph information gets converged at a single node, which is not fixed a priori, called central node.

We assume an initiator node $v \in V$, which performs breadth first search (BFS) in the beginning. Initiator is an arbitrary node in the system. The BFS procedure outputs the BFS tree with v as root. In BFS tree, there are two types of nodes, namely, leaf and non-leaf nodes. The leaf node has single edge connecting parent and non-leaf node has two or more edges that connect it to its neighbors. Assume that there are total N vertices and E edges in the graph. The value of E is upper bounded by N^2 when each vertex is connected to every other vertex. Each node is aware of its neighbors and the weight associated with the edges connecting them.

3 Message Types and Data Structures

3.1 Messages

1. *Make_Me_Parent*: It is used by a node to request some other node to become its child.

2. *ACCEPT*: It is sent by a node, which accepts to become child, to the sender of *Make_Me_Parent* message.
3. *REJECT*: It is sent by a node, which has already become child of some other node, in response to *Make_Me_Parent* message.
4. *E_Msg*: It is sent by $node_i$ to its parent. It is edge information massage containing Id of all neighbors for $node_i$, the weights associated with the edges connecting them, and all other *E_Msg* messages received from its BFS neighbors. It also contains Id of its source node.
5. *MST_Info_Msg*: It is the message containing MST information.
6. *Graph_Info_Msg*: It is the message containing complete graph information.

3.2 Data Structures

Each $node_i$ maintains the following data structures:

1. *has_parent$_i$*: Boolean variable indicating whether the $node_i$ has parent.
2. *allNeighborList$_i$*: It is the list containing Id of all neighbors (BFS and non-BFS) of $node_i$.
3. *countInfoEdge$_i$*: It contains the number of edges through which $node_i$ has received *E_Msg*. It is initialized to zero and incremented on reception of each *E_Msg*.
4. *countBFSedge$_i$*: It is a variable that contains the number of BFS neighbors of $node_i$.
5. *BFS_NeighborList$_i$*: It is the list containing Id of all BFS neighbor nodes of $node_i$.
6. *Array_list$_i$[]*: Each $node_i$ maintains a 1-D array having number of elements equal to *countBFSedge$_i$*; *Array_list$_i$[j]* = 1, in case, *E_Msg* received from node j; *Array_list$_i$[j]* = 0, otherwise. *Array_list$_i$[j]* is initialized to 0 for all j.
7. *Reject_Count$_i$*: Each $node_i$ maintains this variable to count the number of *REJECT* messages received.

4 The Algorithm

We present the algorithm in event action form. First, the initiator starts the BFS protocol by sending *Make_Me_Parent* to its neighbors.

At Ordinary Node

```
Event  1:  on  receiving  Make_Me_Parent  message:  if
     has_parent = 1 then reply REJECT else reply ACCEPT;
     Forward Make_Me_Parent to its other neighbors.
Event  2:  on  receiving  REJECT  message:  increment
     Reject_Count; If Reject_Count = |allNeighborList| -
     1, send E_Msg to parent.
Event 3: on receiving ACCEPT message: Put the sender Id
     in BFS_NeighborList.
Event 4: on receiving E_Msg from node j:
    Set Array_list[j] ← 1;
    countInfoEdge ← countInfoEdge + 1;
    If (countInfoEdge = = countBFSedge - 1)
        Then scan Array_list[] to find node k for which
      Array_list[k] = 0;
        append its own E_Msg message to other received
      E_Msg messages and forward this message to node k.
    Else If (countInfoEdge < countBFSedge - 1)
          Then store the message;
        Else if (countInfoEdge == countBFSedge) then
            If (node_i . id < E_Msg.sender.id){
          /* the entire graph information has converged
          at node_i */
            calculate the MST using Kruskal's or Prim's
          algorithm;
             send  MST_Info_Msg  on  newly  computed  MST
edges.}
```

The Working of Algorithm

The nodes that have received some *ACCEPT* messages are non-leaf nodes; however, the nodes, which do not receive *ACCEPT* message, become leaf nodes. Thus, each node is inherently aware of its status as leaf or non-leaf node. The leaf nodes send edge information message to their respective parent nodes. If a non-leaf node has total e edges, then the node would wait for the arrival of edge information messages on each of its $e - 1$ BFS tree edges. Once it has received edge information messages on its $e - 1$ edges, the node appends its own edge information message to the received messages and forwards the combined message on the remaining eth edge. Finally, there would be a single node in the system that would receive edge information messages on all of its BFS edges. As this node contains the entire graph information, we call it 'central' node, henceforth. Afterwards, the central node computes MST using Kruskal's or Prim's algorithm and disseminate the MST information on newly computed MST edges.

5 The Dynamic Maintenance of Spanning Tree

The wireless networks undergo topological changes due to mobility and crash of nodes. The movement of a node from one location to another is equivalent to the situation when the node crashed at one location recovers at another. Thus, in our illustration we call such node as moved node or leave node. Similarly, a crashed node that is unable to recover and a moved node that does not join the system again are two equivalent situations. Also, a node that is outside the system may join the network at any location. We call such node as new node. Now, we present the data structures, messages, and procedures required to handle the node leave and join events. The topological changes may partition the network graph into multiple subgraphs. The subgraph containing the central node is referred as central subgraph (CSG, henceforth) and the subgraph not containing the central node is called other subgraph (OSG, henceforth).

5.1 Common Data Structures

ANN: The set of nodes that were neighbors of moved node in the MST. We call them affected neighbor nodes (ANN). In fact, the movement of a node creates as many subgraphs as the number of ANNs. In fact, each ANN is the representative of its subgraph and it is used to identify the same.

Parent: This variable is maintained at every node. It stores Id of the sender of *LN_Msg* message.

Child: The MST neighbors of a node to whom it sends *LN_Msg* message.

countNewEdges: This variable is maintained at central node. It stores the number of newly constructed edges about which new edge information has been received. It is initialized to 0.

connectedToCSG: It is a Boolean variable maintained at each node, initialized to 0. If the node's subgraph has got a path to CSG, it is set to 1 else it is set to 0.

Each node maintains the following lists:

My_Req_List: It is the list of OSG nodes to which the node has sent the *OSG_Req_Msg* message to know whether their subgraphs have got a path to CSG.

OSG_Req_List: It is the list of OSG nodes from which the node has received *OSG_Req_Msg* message.

AlreadySentReplyList: The list that contains ANN Id's of the OSGs to which *OSG_Reply_Msg* message has been send.

5.2 Messages

1. *LN_Msg*: It contains Id of moved node and ANN node.
2. *Connect_Ask_Msg*: The node sends this message to its parent to get permission to connect with the other subgraph. It contains the requester's Id *requesterID* and the edge between two nodes belonging to different subgraphs.
3. *OSG_Req_Msg*: It is the request by a node to OSG node to know whether it has a path to CSG.
4. *Connect_CSG_Success*: A node sends this message to its *Child* node to grant permission to connect to OSG that may be CSG. It contains *requesterID* of the node that has found a path either to CSG or to OSG that has found a path to CSG. As a result, a new edge between the requester node and the CSG (or the OSG, as the case may be) node will be formed. Also, this message is used by a node to inform its other *Child* nodes about the newly found path to CSG.
5. *OSG_Reply_Msg*: The node sends this message to the members of *OSG_Req_List* once it finds a path to CSG.
6. *Edge_Permitted_Msg*: When a node gets permission from its ANN to create a new spanning tree edge, it sends this message to the OSG node that is connected to CSG. In fact, it contains the new edge information that is to be submitted to central node.
7. *Not_Required_Msg*: The node sends this message to the nodes in *My_Req_List* on receiving *Connect_CSG_Success* message.
8. *ST_Edge_Msg*: It is the information regarding new spanning edges. It originates from the central node and propagated in whole graph. It marks the end of node leave operation when it is received by a node.

6   The Procedures

The node currently taking an action or executing a procedure is referred as current_node.

Connect_Central(central_node, moved_node, current_node): If any BFS neighbor node *x* of the current_node belongs to CSG, the procedure returns *x*.

Connect_OSG(current_node, moved_node): The procedure returns ANN Id of all neighbor nodes that do not belong to CSG.

Procedure Node_Leave()

When a node crashes or leaves its current position, its neighbor nodes get affected as they do not receive beacons from it within time out. The ANNs, other than the ANN belonging to CSG, initiate reconnection with the central node by sending *LN_Msg* message to its MST neighbors.

Procedure Node_Join()

When a new node N_i appears in the neighborhood of some node N_j, procedure *Node_Join* is executed at node N_j. However, if the entrant node N_k is moved node, procedure *Node_Join* will not initiate until procedure *Node_Leave* has terminated. The termination is marked by the reception of *ST_Edge_Msg* at node N_j. In other words, we assume that the executions of *Node_Leave* and *Node_Join* procedures are never interleaved. The reason is as follows. When a new node N_i joins a graph, the MST is in place. Hence, if N_i attempts to join as neighbor of node N_j, then node N_j sends new edge information on MST edges. However, due to the movement of a node, some MST edges may disappear. Thus, when a moved node N_k attempts to join as neighbor of node N_j, the MST may not be in place. Therefore, before executing *Node_Join* procedure, node N_j waits for procedure *Node_Leave* to terminate. In fact, after the termination of *Node_Leave* procedure, node N_k becomes as good as new node; thus, it can be treated accordingly. Conceptually, the node leave procedure involves the following two steps.

1. Send new node's table info containing its edge weights to neighbors to the central node.
2. Send new spanning edge info connecting the new node to the spanning tree already in place throughout the ST.

At ANN

```
Event 1: If timeout expires
moved_node = node for which timeout has expired
if current_node ∈ CSG then skip()
else    if    Connect_Central(central_node,    moved_node,
current_node)
  {
  return x;
  send Connect_CSG_Success to its MST neighbors;
  } else {
      requestedSubgraph    =    Connect_OSG(current_node,
moved_node);
      send OSG_Req_Msg to each OSG ∈ requestedSubgraph;
      send  requestedSubgraph  with  LN_Msg  to  all  MST
neighbors except moved_node
      }
```

At Ordinary Node

```
Event 2: On receiving LN_Msg:
  Parent = LN_Msg.sender.id;
  if      (Connect_Central(central_node,      moved_node,
current_node)
   send Connect_Ask_Msg to Parent
   else
   send OSG_Req_Msg to OSGs to which it has not been sent
yet;
   append remaining OSG list to LN_Msg and send it to the
MST neighbors other than sender.
Event 3: On receiving OSG_Req_Msg:
  If connectedToCSG = = 1
  send OSG_Reply_Msg to the sender of OSG_Req_Msg
  else
   store the sender Id in OSG_Req_List
```

7 The Simulation Results

7.1 The Simulation of Node Leave Procedure

The simulation experiment has been performed for node leave procedure by varying the number of nodes and edges in the graph randomly. Accordingly, the number of edges per node also varies. The crash message overhead is the number of additional messages exchanged due to crash of a node. The total message is the number of messages required to construct BFS tree and disseminate MST information in the graph. The crash message overhead depends upon the network topology. For instance, if the number of edges is more, the crash message overhead would be less; however, more messages would be required to construct BFS tree and disseminate MST information in the graph. On the other hand, if the number of neighbors of crashed node (i.e., ANN) is more, the crash message overhead would be more. Thus, we have specified node density in order to maintain average number of neighbors per each node between 1.2 and 10. Also, to compute crash message overhead, the nodes having at least three neighbors are the candidate crashed nodes.

The following Fig. 1 shows that the crash message overhead per node (shown in red) is increasing at a rate slower than the number of edges per node, i.e., edge density (shown in blue) of the graph. Also, the slope of the plot representing the crash message overhead per edge (shown in green) is negative. Thus, the crash message overhead decreases with increase in edge density.
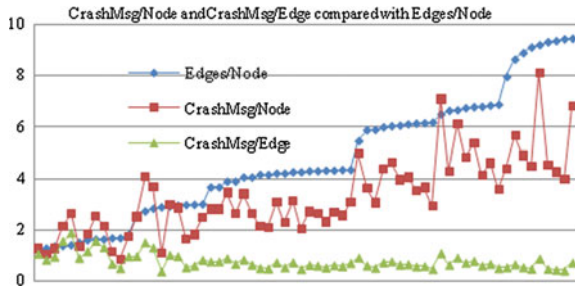
Fig. 1 Crash message overhead

Table 1 Join message overhead

Sr. No.	No. of nodes	No. of edges	Join message overhead	Total message
1	17	31	18	103
2	25	51	26	147
3	50	121	51	124
4	65	155	66	89
5	82	200	83	179
6	101	241	102	153
7	151	361	152	286
8	201	481	202	367
9	401	1019	402	434

7.2 The Simulation of Node Join Procedure

In the following Table 1, in all nine cases, the join message overhead is just one more than the number of nodes in the graph. Moreover, it is independent of the number of edges in the graph.

8 Conclusion

A dynamic maintenance method for spanning tree was presented in this paper. The simulation results confirm that it has low message overhead because the major computation work is local in the approach. Also, the approach is easily understandable and it is simple to implement.

References

1. Gallager, R., Humblet, P., Spira, P.: A distributed algorithm for minimum weight spanning trees. ACM Trans. Programming Languages and Systems. 5, 1, 66–77 (1983).
2. Chin, F., Ting, H.: An almost linear time and $O(n \log n + e)$ messages distributed algorithm for minimum weight spanning trees. In: 26th IEEE Symp. Foundations of Computer Science. pp. 257–266 (1985).
3. Gafni, E.: Improvements in the time complexity of two message-optimal election algorithms. In: 4th ACM Symp. Principles of Distributed Computing. pp. 175–185 (1985).
4. Awerbuch, B.: Optimal distributed algorithms for minimum weight spanning tree, counting, leader election, and related problems. In: 19th ACM Symp. Theory of Computing, pp. 230–240 (1987).
5. Garay, J., Kutten, S., Peleg, D.: A sublinear time distributed algorithm for minimum weight spanning trees. SIAM J. Comput. 27, 302–316 (1998).
6. Kutten, S., Peleg, D.: Fast distributed construction of k-dominating sets and applications. J. Algorithms, 28, 40–66 (1998).
7. Elkin, M.: A faster distributed protocol for constructing minimum spanning tree. In: ACM-SIAM Symp. Discrete Algorithms. pp. 352–361 (2004).
8. Peleg, D., Rabinovich, V.: A near-tight lower bound on the time complexity of distributed MST construction. In: Proc. of the 40th IEEE Symp. on Foundations of Computer Science, pp. 253–261 (1999).
9. Tel, G.: Introduction to Distributed Algorithms. 2/e, Cambridge Univ. Press, New York, pp. 209–210 (2000).
10. Dynia1, M., Korzeniowski, M., Kutylowski, J.: Competitive maintenance of minimum spanning trees in dynamic graphs. In: Proc. Jan van Leeuwen et al. (Eds.): SOFSEM 2007, LNCS 4362, pp. 260–271 (2007).
11. Flocchini, P., Mesa Enriquez, T., Pagli, L., Prencipe, G., Santoro, N.: Distributed minimum spanning tree maintenance for transient node failures. IEEE Transactions on Computers, 61, 3, 408–414 (2012).
12. Cattaneo, G., Faruolo, P., Petrillo, U., Italiano, G.: Maintaining dynamic minimum spanning trees: an experimental study. Discrete Applied Mathematics, 158, 404–425 (2010).
13. Awerbuch, B., Cidon, I., Kutten, S.: Optimal maintenance of a spanning tree. Journal of the ACM, 55, 4, 18:1–18:45 (2008).

Internet of Things: Future Vision

Sushma Satpute and Bharat Singh Deora

Abstract The increase in the communication devices and their adoption by modern world gives rise to Internet of Things (IoT), which also covers sensors and actuators that are blended seamlessly with the environment around us. As IoT covers variety of enabling of various device technologies like sensors, communication devices, and smart phones, hence, next revolution will be transformation of present Internet to fully integrated future Internet. In this paper, we presented proposed architecture of cloud-based IoT.

Keywords IoT · Internet · Cloud data storage

1 Introduction

Today, the world is dependent on computers and, therefore, the Internet. Everybody almost wholly dependent on machines for information. Even digital copies of old knowledge are maintained. However, the problem is that people have limited time, and hence are not able to capture, maintain, and update data about things in the real world properly. If machines can present the exact data without human help, then it will reduce waste of time, help in recovering lost data, and also we would know when modification in data is required. This requirement is partially fulfilled by IoT [1]. IoT will be in future consists of transfer of information and connectivity between human to human, machine to human, and machine to machine, that too without human interference [1]. IoT has evolved from convergence of wireless technologies, micro-electromechanical systems and the Internet. Based on the study report by Ericsson, in 2003, more than 6.3 billion people from all over the

S. Satpute (✉)
Pacific University, Udaipur, India
e-mail: sushma.satpute@gmail.com

B.S. Deora
JRN Rajasthan Vidyapeeth University, Udaipur, India
e-mail: bsdrvu@yahoo.com

world are connected to each other via digital devices and nearly 500 million devices are recognized over Internet (mostly PCs and a few smart phones). Another study says that today in 2015 every person has his own one or two digital devices on an average, and this number is increasing very fast [2]. By 2020, population will grow about 7.6 and 50 billion more devices will be connected to the Internet [1]. Technology is advancing making devices more small in size, yet making them more powerful and these devices will have capability of getting attached to other devices easily. Think of something like washing machine is accessed via Internet and also able to generate a report visible on smart phone. Hence, in future we will see all things right from thermos to water purifier and storage tanks will be on Internet, as manufacturers add their products with its description, price, etc., also operating them via internet. All these although cloud technology may handle the challenges generated by increasing device connectivity still are in the beginning phase. As per recent observations the existing systems do not cover the integrity in terms of communications and management of devices [3]. Although many observers look at cloud computing as a solution, cloud computing can be part of IoT.

2   Difference in Cloud Data and IoT

Is there any difference terms IoT and cloud computing? Both apparently similar working, but as we are going to deal with enormous data we need more improved cloud computing techniques and increased integrity, interoperability in cloud computing will be one of the components of IoT. IoT cannot be differentiated from cloud computing.

3   Definition

We define our IoT as "interconnection of sensors and actuators provides the ability to share information across platforms through a unified framework developing a common operating picture for enabling a innovative project, which is achieved by seamless large scale sensing data analytics and information representation using cutting edge ubiquitous sensing and cloud computing" [4].

Proposed IoT consists of three main components, which are as follows:

(1) *Hardware component* hardware components are mainly communication hardware components;
(2) *Middleware component* These components contain data analysis processes and storage of data as per requirement;
(3) *Presentation component* These components include visualization and report generating tools and can be widely accessed at various platforms and can be used for different applications and at various locations.

Fig. 1 Conceptual framework of IoT with cloud computing
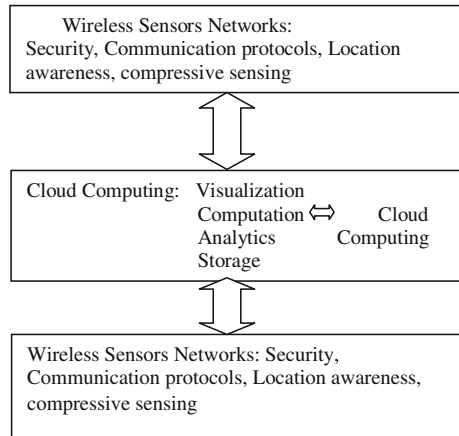
4 Conceptual Framework of IoT

Conceptual framework consists of three layers: hardware, middleware, and application or presentation layer.

  I. Hardware layer
     Hardware layer consists of networking devices which include data reading devices like RFID and also components of wireless sensor network. This layer decides communication protocols, security algorithms, data of location awareness, compressive sensing, and quality of service.
 II. Middleware layer
     This layer consists of data storage mechanism along with accessing data mechanism, data security, etc. Here cloud computing covers all these issues. Cloud computing through SaaS software as service, PaaS platform as service, and IaaS infrastructure as a service gives visualization of data, computation of data, and analysis and storage of data.
III. Presentation layer
     This layer consists of application layer, where data outcomes are utilized in applications like surveillance. It is used in critical infrastructure monitoring, environment monitoring, and health monitoring and transportation systems.

   This can be graphically represented in Fig. 1.

5 Cloud Centered Internet of Things

The term Internet of Things consists of terms Internet and things. Internet involves Internet services while things mean the real-life object. A combined framework with cloud at the center gives flexibility of dividing association costs in the most

logical manner and also highly scalable. Sensors integrate the data stored on cloud and communicate devices via Internet. Analysts find the relevant information and convert it into knowledge using various data mining techniques. Graphics designer gives variety of visualization tools for interpreting the results.

All these services are offered by cloud computing. Cloud platforms like Microsoft azure and Manjrasoft Aneka control all processes like storage of data, analysis, and interpretation of results.

6 Architecture

Hence, the proposed architecture with cloud as a center system will have connection with each and every individual machine whether it is a small printer or a big capacity server at any corner. Similarly, it will take care of data backup.

As shown in Fig. 2 the bottom layer consists of various end users. End users communicate the system via devices like RFID reader or sensors. Thus real-time object is transferred into digital. The data is put on to Internet which means devices getting read and data is stored on cloud. This layer consists of network protocols and communication and networking devices. The data from first layer is stored on cloud in the second layer. This layer is cloud centered. This layer consists of communication protocols, transporting the data, and after that analysis process. Here data is analyzed and processed by cloud computing. Various data mining algorithms are applied here. The IoT control unit also exists in the same layer. IoT unit integrates with cloud computing. Hence, the controlling of various end user devices is done by IoT. All the processes right from the first to the third layer are managed centrally by a IoT control unit. The last layer is to visualize or interpret the results. Here various visualization tools like advanced graphics tools and many more softwares interpret the results in efficient and understandable manner. The

Fig. 2 A new proposed architecture for IoT
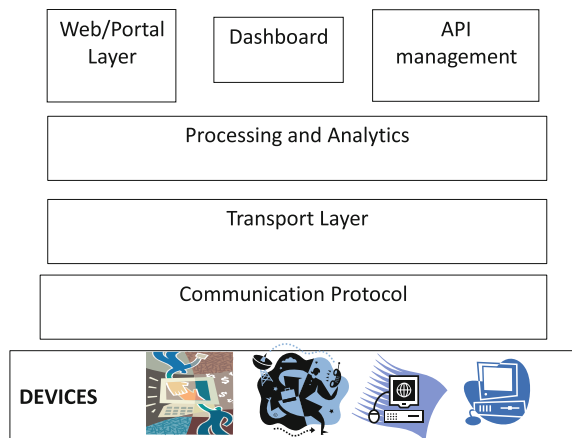
results can be viewed on Web or through any portal, or can be on dashboard or displayed on some management interface device.

The heterogeneity of data is, however, managed by software like OpenStack. OpenStack is actually group of open-source software packages designed for building public and private clouds [5]. Also on cloud individual user can access the data by dedicated virtual machine instance, and hence possible for multiple jobs to share single operating system. The process of job scheduling on cloud can be done using various strategies. A resource allocation strategy uses two pools—core nodes and accelerator nodes—and then dynamically adjusts the size of each pool to reduce cost or improve utilization. Also, to provide good performance while guaranteeing fairness in shared heterogeneous cluster, we propose progress share as a new fairness metric to redefine the share of a job in heterogeneous environment [6].

We can enable data protection to the archive process instead of back up of data which is big and changing. One approach may be cache storage of data. Copies can be made to multiple tape devices. This increases performance access for processing, but also high protection and long-term durability. To handle big data on cloud various data mining techniques are available. Amazon EMR offers best solutions to data mining. EMR is an on-demand service, which can be classified as a category of SaaS (Software as a Service) and PaaS (Platform as a Service) solutions, depending on the implementation by the user. EMR offers flexible resources, programmability, payment according to the standard CC principle only used resources, geographically dislocated EC2 infrastructure, and in most cases an increased level of security [7].

7   Conclusion

The proposed architecture is more flexible and open, and users from different platforms can interact in IoT framework. Still the challenges include privacy, participatory sensing, and data analytics. Another challenge is management of big data and processing of number of things. Also, there are wireless sensor network challenges like security, time constraint proper protocols, and quality of service. At the end we can say that connecting different objects, making and integrating them to make a big system comprising the small modules which can be easily integrated with any subsystem or object, and making them smarter will be an ideal IoT. Here proposed framework allows networking, computation, storage, and visualization themes separate thereby allowing independent growth in every sector, but complementing each other in a shred environment. Thus we can provide a better solution to current system via IoT which will connect each one with everyone and will remove the need of costly hardware and software requirements and in future will prove as a new big step in IT.

References

1. The Internet of Things–Industry Report, U.S. Research Published by Raymond James & Associates. © 2014 Raymond James & Associates, Inc., member New York. Technology & Communications January 24, 2014.
2. The-us-digital-consumer-report.html http://www.nielsen.com/us/en/insights/reports/2014/.
3. http://www.iot-a.eu/public/public-documents/copy_of_d1.2.
4. Internet of Things (IoT): A vision, architectural elements, and future directions Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, Marimuthu Palaniswami, Department of Electrical and Electronic Engineering, The University of Melbourne, Vic - 3010, Australia, Department of Computing and Information Systems, The University of Melbourne, Vic - 3010, Australia.
5. Steve Crago, Kyle Dunn, Patrick Eads, Lorin Hochstein, Dong-In Kang, Mikyung Kang, Devendra Modium, Karandeep Singh, Jinwoo Suh, John Paul: "Heterogeneous cloud computing" in 2011 IEEE International Conference on Cluster Computing 978-0-7695-4516-5/11 $26.00 © 2011 IEEE DOI:10.1109/CLUSTER.2011.49.
6. Gunho Lee, Byung-Gon Chun, Randy H. Katz University of California, Berkeley, Yahoo! Research" Heterogeneity-Aware Resource Allocation and Scheduling in the Cloud" in Proceedings of HotCloud, 2011 - static.usenix.org.
7. Robert Vrbić University Vitez, Travnik, Bosnia and Herzegovina "Data Mining and Cloud Computing" Journal of Information Technology and Applications. source// www.jita-au.com.

Design and Analysis of Energy Efficient OPAMP for Rectifier in MicroScale Energy Harvesting (Solar Energy)

Vijendra K. Maurya, R.M. Mehra and Anu Mehra

Abstract The physical process by which the energy is gathered from the surrounding environment is called energy harvesting. For several micro-scale electronic systems, electrical energy harvested from the sunlight proves to be an attractive and feasible solution. The paper presents the energy efficient design of OPAMP which is in conjunction with the one of the most important integral component of electrical interface known as rectifier, an AC–DC converter. A high gain OPAMP is proposed to harvest the maximum electrical energy obtained from the AC–DC energy converter, since the electrical signal obtained lies within few millivolts as a result it needs the amplification to provide startup voltage for the high end electronic applications. In this work a detailed analysis has been done between two designs of OPAMP one is high gain OPAMP design and other is design of OPAMP with current buffer compensation technique targeted to maximize the power extraction from the converters. Both the designs are implemented in 180 nm technology, whereas high gain design OPAMP operates at 1.2 V and another design of OPAMP operates at 1.8 V. The gain reported by the high gain OPAMP design is 95.41 dB with the power efficiency of 188 nW and current buffer compensation technique OPAMP is 90.71 dB with the power efficiency of 244 nW.

V.K. Maurya · Anu Mehra (✉)
ASET, Amity University, Noida, India
e-mail: anumehra2004@yahoo.com

R.M. Mehra
University of Delhi, New Delhi, India

Fig. 1 Block diagram of
micro-scale harvesting system
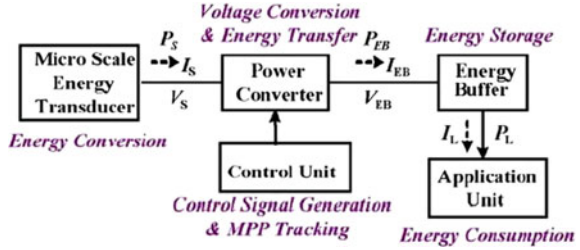
1 Introduction

To overcome from the challenges that rose due to the power supply requirements,
harvesting of energy can be done from the environmental energy sources such as
sunlight, thermal, fuel, vibration, etc. The paper is mainly focused on the solar
energy; the sunlight obtained is converted into usable electrical energy to power up
the electronic equipments. Many ideas have been explored to use the energy har-
vested from the environment and applied to the macro-level projects such as solar
farm, wind mills but the design issue comes when the micro-level harvesting system
designed which suffers from the small size constraint which demands the minia-
turized energy transducers responsible to generate the equivalent electrical quantity
from the physical quantity, as a result the output obtained ranges within few mil-
lvolts. So it is needed to design the subsystems of harvesting which ensures the
extraction of maximum power at the same time design also proves to be energy
efficient [1–8].

The micro-scale energy harvesting system consists of five blocks named as
transducer, power converter, control unit, buffer, and application unit. The trans-
ducers capture the ambient energy from the surrounding and transform into elec-
trical energy and then store in buffer for example rechargeable battery or capacitors.
The control unit maximizes the overall efficiency of the system. The role of power
converter is to obtain the maximum amount of power as much as possible from the
transducer and deliver it to the output. The block diagram is shown in the Fig. 1.

2 Design of OPAMP

2.1 Proposed Design of High Gain OPAMP

A fully differential topology has been proposed in order to attain the high gain and
good range of bandwidth. The OPAMP gets its supply from the energy which is
harvested from the environment; OPAMP is designed to operate in subthreshold
region which ensures the low power consumption.

Design of input stage, deals with the design of complementary stage, comprises of N-differential pair (NMOS_2 and MOS_4) and P-differential pair (PMOS_3 and PMOS_4). The transistors with current biasing were employed in the circuit to ensure the constant flow of current in the differential stage. The transistors (PMOS_1, PMOS_2, NMOS_1, and NMOS_2) used in the circuit are responsible to keep the g_m at constant value. The remaining transistors employed in the cascode stage are utilized as summing stage. The transistors PMOS_7, PMOS_8, NMOS_7, and NMOS_8 are connected in the circuit to avoid null value of the current at the output.

If the PMOS_5, NM7, NMOS_6, and NM6 are not connected in the circuit then the output will depend on g_m. Design of output stage, the output stage consists of push–pull pair NMOS_15 and PMOS_15, the driver circuit comprises of NMOS_3, NMOS_4, PMOS_3, PMOS_4, NMOS_13, and PMOS_13 and MOS operating as current generator comprises of NMOS_12 and NMOS_17. The structure of the driver is symmetrical such that the aspect ratios of transistors have kept same.

When the input and output design are cascaded, it forms the complete design of OPAMP which comprises of three stages: first stage is differential pair and constant g_m circuit design, second is summing stage, and third stage is to obtain large output swing of electrical voltage for which the push–pull configuration is proposed.

2.2 Design of OPAMP with Current Buffer Compensation Technique

In this section two stages CMOS OPAMP with current buffer compensation is presented, designed to operate at low voltage supply of 1.8 V which results the operation of OPAMP in weak inversion region. The circuit is realized in 180 nm CMOS technology, the weak inversion region allows the operation of OPAMP in at low bias current and low voltage. When the current is scaled it results in decreased power consumption but at the same time the compromises are made with the dynamic characteristics and similarly when the voltage is scaled it becomes difficult to attain the transistors in saturation region, in order to achieve the proper characteristics.

In the current buffer compensation technique the feed forward path is removed from the output from the first stage to the OPAMP output. The feed through compensation technique is involved in the use of buffer to break the bidirectional path through compensation capacitor but unfortunately zero result is obtained in the circuit which is in right half plane, which makes the condition unsuitable since to achieve the stability of OPAMP it is required that the poles must be in left half plane.

The design is targeted to achieve low power consumption and to operate at low voltage. The circuit presented in the Fig. 5 consists of cascode stages which involve in voltage to current conversion stage and current to voltage conversion stage. The

first is the differential stage that comprises of NMOS_1, NMOS_2, PMOS_1, and PMOS_2 which take the differential input voltage and transform to the differential currents and then applied to the current mirror load to recover back the voltage. The current source of the 1 nA is used in the design which ensures the constant current flow in the differential circuit. The second stage is formed by the common source MOSFET which transforms the second stage input voltage to the current. The current sink load is used at the output of the configuration which is responsible to transform the current to voltage at the output.

The C_c is the compensation capacitor used in the design which couples the output of the preceding stage to the output stage. The stability is achieved using the compensation capacitor.

2.3 Figures and Tables

2.3.1 Micro-Scale Harvesting System

Figure 1.

2.3.2 Design of Proposed High Gain OPAMP

The OPAMP is implemented in 180 nm CMOS technology at the circuit level, using Tanner EDA tool packed with (S-Edit, T-Edit, and W-Edit) where the schematic is designed using S-EDIT tool, the simulation is done using T-SPICE and waveform obtained from the W-EDIT tool. The circuit operates at the voltage of 1.2 V (Fig. 2).
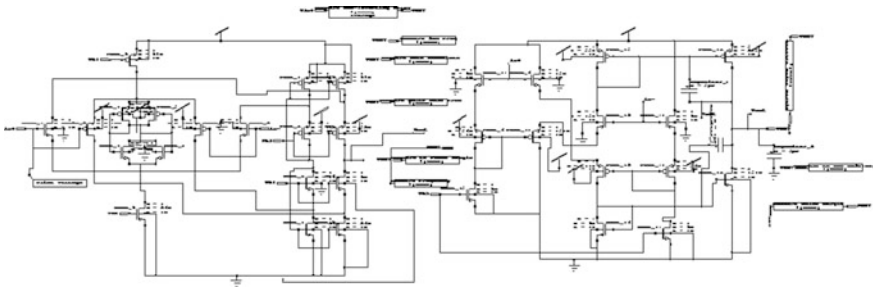
Fig. 2 Schematic of proposed high gain OPAMP
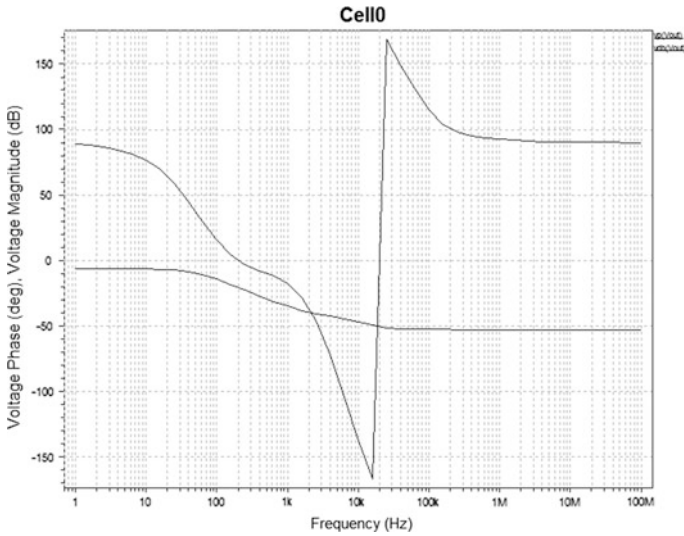
Fig. 3 AC analysis of proposed OPAMP

2.3.3 Simulation Waveform of Proposed Design

See Figs. 3 and 4.

2.3.4 Design of Current Buffer Compensation OPAMP

The OPAMP is realized in 180 nm CMOS technology, using Tanner EDA tool where the schematic is designed using S-EDIT tool, the simulation is done using T-SPICE and waveform obtained from the W-EDIT tool. The circuit operates at the voltage of 1.8 V (Fig. 5).

2.3.5 Simulation Waveform of Current Buffer Compensation OPAMP

See Figs. 6 and 7.
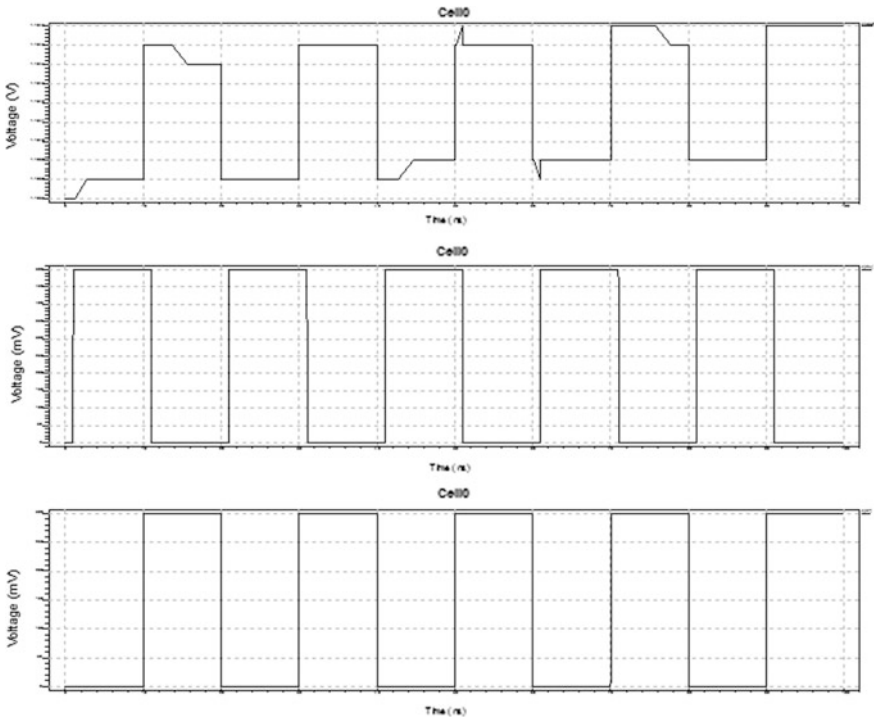
Fig. 4 DC analysis of proposed OPAMP
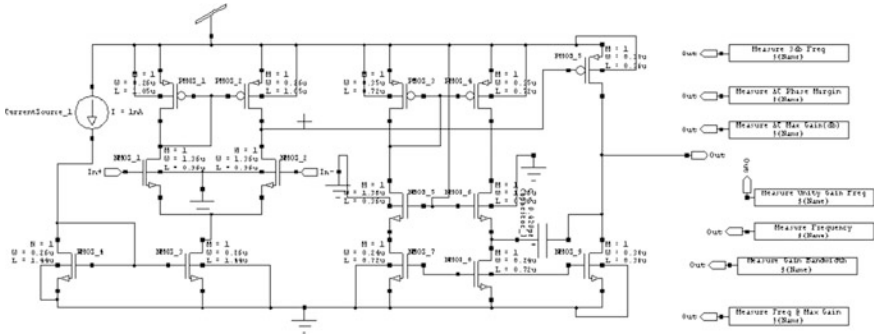
Fig. 5 Schematic of current buffer compensation OPAMP
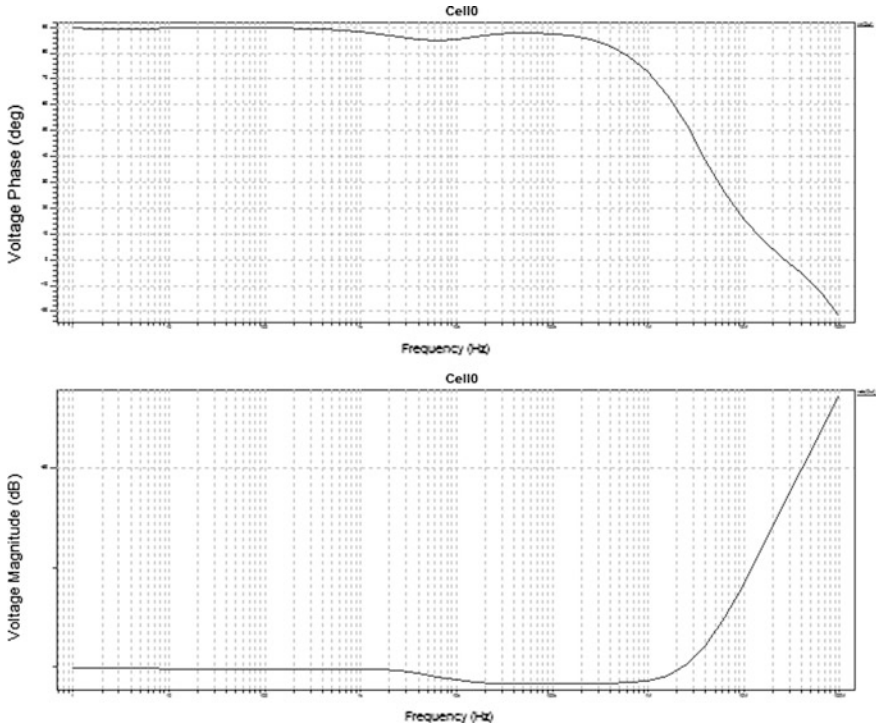
Fig. 6 AC analysis of current buffer compensation OPAMP
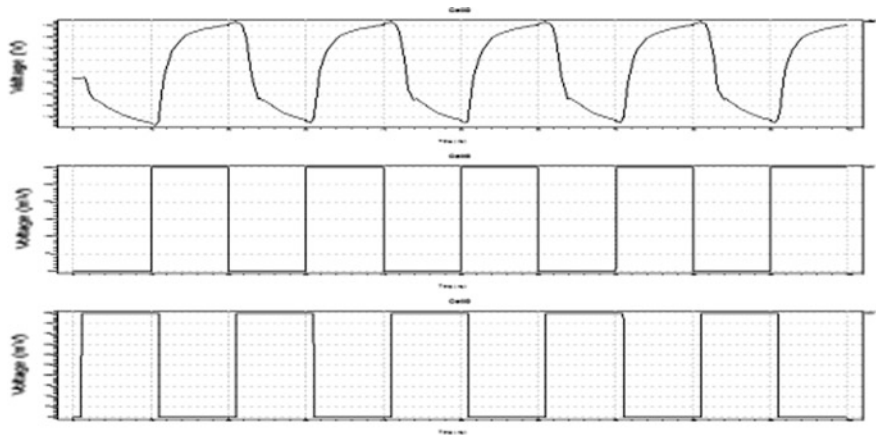
Fig. 7 DC analysis of current buffer compensation OPAMP
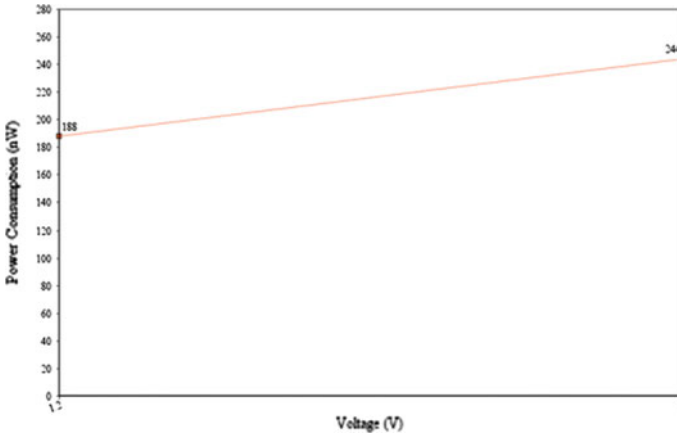
Fig. 8 Graph between supply voltage and power consumption
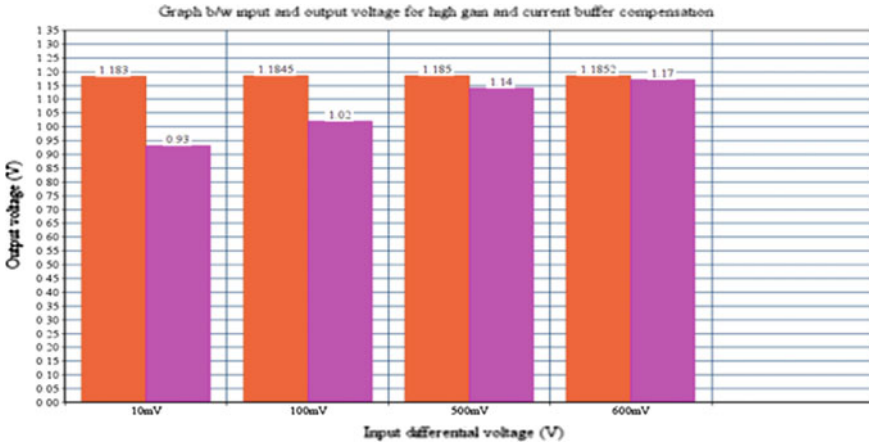
Fig. 9 Performance analysis between high gain OPAMP and current buffer compensation

2.3.6 Graphs

See Figs. 8 and 9.

2.3.7 Performance Analysis Table between Proposed High Gain OPAMP and Current Buffer Compensation OPAMP

S. No.	Non inverting input voltage (In+)	Inverting input voltage (In−) (mv)	Proposed design of OPAMP (Vout)	Current buffer compensation OPAMP (Vout)	Proposed design of OPAMP (Gain) (dB)	Current buffer compensation OPAMP (Gain) (dB)	Power consumption proposed design (Watt)	Power consumption current buffer (Watt)
1	30 mV	20	1.1830	933 mV	95.41	90.71	2.1×10^{-7}	6.6×10^{-6}
2	500 mV	300	1.1845	1.02 V	33.16	32.58	2.9×10^{-6}	1.3×10^{-5}
3	500 mV	400	1.1847	1.15 V	49.24	48.54	1.3×10^{-6}	7.49×10^{-5}
4	1 V	500	1.1850	1.14 V	17.25	16.48	3.8×10^{-6}	1.6×10^{-5}
5	1.2 V	600	1.1852	1.17 V	13.56	13.35	4.0×10^{-6}	1.7×10^{-5}

S. No.	Parameters	Proposed high gain OPAMP	Current buffer compensation OPAMP
1	CMOS technology	180 nm	180 nm
2	Operating Voltage	1.2 V	1.8 V
3	Gain bandwidth product	100 kHz	2.5 MHz
4	3 dB frequency	1.531 Hz	1.44×10^4 MHz
5	Unity gain frequency	2.8×10^5 Hz	3.87×10^5 Hz
6	AC gain	2.13 dB	6.4576 dB
7	Gain	95.41 dB	90.71 dB
8	Common mode gain	77.08 dB	72.32 dB
9	Settling time	0.66 s	69 μs
10	Offset Voltage	11.5 mV	1.0 V
11	Power consumption	188 nW	244 nW

2.4 Formulas

Design formula for the proposed OPAMP

$$(W/L) = 2 \times I_{\text{ds}}/\left(K' \times V_{\text{deff}}^2\right) \tag{1}$$

$$(W/L)_{\text{NMOS\_4, NMOS\_3}} = (W/L)_{\text{NMOS\_9}}/2 \tag{2}$$

$$(W/L)_{\text{PMOS\_3, PMOS\_4}} = 2.5^*(W/L)_{\text{NMOS 4}} \tag{3}$$

$$(W/L)_{\text{PMOS\_9}} = 2.5^*(W/L)_{\text{NMOS\_9}} \tag{4}$$

$$(W/L)_{\text{NMOS 9}} = 2 \times I_{\text{SS}}/\left(K' \times V_{\text{deff}}^2\right) \tag{5}$$

$$(W/L)_{\text{PMOS 1, PMOS 2, NMOS3, NMOS 4}} = 3^*(W/L)_{\text{PMOS 3}} \qquad (6)$$

$$(W/L)_{\text{NMOS 1, NMOS 2}} = 3^*(W/L)_{\text{NMOS 4}} \qquad (7)$$

$$(W/L)_{\text{PMOS 5,NM 7}} = 3.5^*(W/L)_{\text{PMOS 9}} \qquad (8)$$

$$(W/L)_{\text{NMOS 5, NM 6}} = (W/L)_{\text{PMOS}}5/2.5 \qquad (9)$$

2.5 Program Code

2.5.1 Program Code for Proposed Design of OPAMP

```
CCapacitor_1 N_18 Vout  2p
CCapacitor_2 N_22 Vout  1p
CCapacitor_3 Vout Gnd  2p
MNMOS_1 N_4 N_4 N_10 Gnd NMOS W=46u L=1u AS=41.4p PS=93.8u AD=41.4p PD=93.8u
MNMOS_2 N_4 N_4 N_10 Gnd NMOS W=46u L=1u AS=41.4p PS=93.8u AD=41.4p PD=93.8u
MNMOS_3 N_6 VIN- N_5 Gnd NMOS W=10u L=1u AS=9p PS=21.8u AD=9p PD=21.8u
MNMOS_4 N_9 Vin+ N_5 Gnd NMOS W=18u L=1u AS=16.2p PS=37.8u AD=16.2p PD=37.8u
MNMOS_5 N_8 Vb4 N_3 Gnd NMOS W=18u L=1u AS=16.2p PS=37.8u AD=16.2p PD=37.8u
MNMOS_6 Vdd N_12 N_13 Gnd NMOS W=42u L=1u AS=37.8p PS=85.8u AD=37.8p PD=85.8u
MNMOS_7 N_3 N_8 Gnd Gnd NMOS W=36u L=1u AS=32.4p PS=73.8u AD=32.4p PD=73.8u
MNMOS_8 N_1 N_8 Gnd Gnd NMOS W=36u L=1u AS=32.4p PS=73.8u AD=32.4p PD=73.8u
MNMOS_9 N_10 Vb2 Gnd Gnd NMOS W=36u L=1u AS=32.4p PS=73.8u AD=32.4p PD=73.8u
MNMOS_10 VIN- Vin- Gnd Gnd NMOS W=600n L=1u AS=540f PS=3u AD=540f PD=3u
MNM6 Vout Vb4 N_1 N_7 NMOS W=9u L=1u AS=8.1p PS=19.8u AD=8.1p PD=19.8u
MNMOS_16 N_22 N_22 Gnd Gnd NMOS W=9u L=1u AS=8.1p PS=19.8u AD=8.1p PD=19.8u
MNMOS_17 N_21 N_17 Gnd Gnd NMOS W=9u L=1u AS=8.1p PS=19.8u AD=8.1p PD=19.8u
MNM7 Vout Vb3 N_9 Vdd PMOS W=23u L=1u AS=20.7p PS=47.8u AD=20.7p PD=47.8u
MPMOS_10 VIN- Vin- Vdd Vdd PMOS W=2.5u L=1u AS=2.25p PS=6.8u AD=2.25p PD=6.8u
MPMOS_11 Gnd Vin+ N_16 Vdd PMOS W=127u L=1u AS=114.3p PS=255.8u AD=114.3p PD=255.8u
MPMOS_1 N_4 N_4 N_2 Vdd PMOS W=35u L=1u AS=31.5p PS=71.8u AD=31.5p PD=71.8u
MPMOS_2 N_4 N_4 N_2 Vdd PMOS W=36u L=1u AS=32.4p PS=73.8u AD=32.4p PD=73.8u
MPMOS_6 Vin+ Vin+ N_14 Vdd PMOS W=26u L=1u AS=23.4p PS=53.8u AD=23.4p PD=53.8u
MPMOS_7 N_6 N_8 Vdd Vdd PMOS W=90u L=1u AS=81p PS=181.8u AD=81p PD=181.8u
MPMOS_8 N_9 N_8 Vdd Vdd PMOS W=90u L=1u AS=81p PS=181.8u AD=81p PD=181.8u
MPMOS_9 N_2 Vb1 Vdd Vdd PMOS W=90u L=1u AS=81p PS=181.8u AD=81p PD=181.8u
.param Vpwr=1.2V
.include "C:\Users \Documents\Tanner EDA\Tanner Tools v14.1\L-Edit and LVS\LVS\SPR_Core\hp05.md"
.tran 1n 100n
Vin+ in+ Gnd PULSE (0 0.5 1n 20p 20p 10n 20n) ROUND=20n
Vdd Vdd Gnd 1.2
Vin- in- Gnd PULSE (0 0.1 10n 20p 20p 10n 20n) ROUND=20n
vb1 Vb1 Gnd 0.4
vb2 Vb2 Gnd 0.4
vb3 Vb3 Gnd 0.4
vb4 Vb4 Gnd 0.4
vb5 Vb5 Gnd 0.4
.print noise
.print tran v(in-) v(in+) v(VOUT)
.op
.power
.probe
.end
```

2.5.2 Program Code for Current Buffer Compensation OPAMP

```
CCapacitor_1 N_15 Out  420f
MNMOS_1 N_1 In+ N_3 Gnd NMOS W=1.36u L=360n AS=1.224p PS=4.52u AD=1.224p PD=4.52u
MNMOS_2 N_4 In- N_3 Gnd NMOS W=1.36u L=360n AS=1.224p PS=4.52u AD=1.224p PD=4.52u
MNMOS_3 N_3 N_7 Gnd Gnd NMOS W=260n L=1.44u AS=234f PS=2.32u AD=234f PD=2.32u
MNMOS_4 N_7 N_7 Gnd Gnd NMOS W=260n L=1.44u AS=234f PS=2.32u AD=234f PD=2.32u
MNMOS_5 N_10 Vdd N_13 Gnd NMOS W=1.36u L=360n AS=1.224p PS=4.52u AD=1.224p PD=4.52u
MNMOS_6 N_11 Vdd N_15 Gnd NMOS W=1.36u L=360n AS=1.224p PS=4.52u AD=1.224p PD=4.52u
MNMOS_7 N_13 N_16 Gnd Gnd NMOS W=240n L=720n AS=216f PS=2.28u AD=216f PD=2.28u
MPMOS_2 N_4 N_1 Vdd Vdd PMOS W=260n L=1.05u AS=234f PS=2.32u AD=234f PD=2.32u
MPMOS_3 N_10 N_10 Vdd Vdd PMOS W=350n L=720n AS=315f PS=2.5u AD=315f PD=2.5u
MPMOS_4 N_11 N_10 Vdd Vdd PMOS W=350n L=720n AS=315f PS=2.5u AD=315f PD=2.5u
MPMOS_5 Out N_4 Vdd Vdd PMOS W=350n L=360n AS=315f PS=2.5u AD=315f PD=2.5u
ICurrentSource_1 Vdd N_7  DC 5u
.param Vpwr=1.8V
.include "C:\Users \Documents\Tanner EDA\Tanner Tools v14.1\L-Edit and LVS\LVS\SPR_Core\hp05.md"
.tran 1n 100n
VIn+ In+ Gnd PULSE (0 0.5 1n 20p 20p 10n 20n) ROUND=20n
Vdd Vdd Gnd 1.8
VIn- In- Gnd PULSE (0 0.3 10n 20p 20p 10n 20n) ROUND=20n
.print noise
.print tran v(In+) v(In-) v(Out)
.op
.power
.probe
.end
```

3 Conclusion

For the amplification of the electrical signals obtained from the energy converter, the proposed OPAMP is best suited since it operates at the voltage of 1.2 V. From the simulation the experimental result obtained from the proposed OPAMP signifies the low power consumption along with the high value of gain observed as 95.41 dB, although the bandwidth calculated from the frequency response from the design is low as a fact with the lesser output current the dynamic range is affected. At the same time the frequency response of the current buffer compensation technique is wide. The common mode gain of the proposed design is better than current buffer compensation which signifies reduced noise, the gain obtained is 77.08 dB. The offset voltage parameter of the proposed design is 11.5 mV which indicates that when no input is applied the output of OPAMP remains very low as ideally remains zero. The proposed OPAMP result indicates that the power extracted from the rectifier is more than the current buffer compensation OPAMP.

References

1. K.S Thanga Pandian and V Ramkumar: Design and analysis of an Energy Efficient OPAMP Based Rectifier in Piezoelectric Energy Harvesting System, International Journal of Modelling and Simulation of Engineering Research, vol 01, issue 2 (2015).
2. S. Falgini and M. Moallem: A low Power Electronic Converter with the input resistance control for Piezoelectric Energy Harvesting, IEEE/ASME, France (2014).
3. Y. Sun, N. H hieu, C.J Jeong and S.G Lee: An Integrated High Performance Active Rectifier for Piezoelectric Vibration Energy Harvesting System, IEEE trans. Power Electronics, vol. 27 (2012).
4. Rohan Dayal, Suman Dwari, and Leila Parsa: Design and implementation of a direct ac-dc boost converter for low voltage energy harvesting, IEEE Trans. Ind. Electron, vol. 58 (2011).
5. Suman Dwari, Leila Parsa: An Efficient AC–DC Step-Up Converter for Low-Voltage Energy Harvesting by Member, IEEE Transactions on Power Electronics, vol. 25, no. 8 (2010).
6. AA. Nasiri, S. A. Zabalawi, and G. Mandic: Indoor power harvesting using photovoltaic cells for low-power applications, IEEE Trans. Ind. Electron., vol. 56 (2008).
7. K. Lahiri, A. Raghunathan, and S. Dey: Battery-driven system design: a new frontier in low power design, Proc. IEEE International Conference on VLSI Design (2002).
8. Randall L. Geiger, Phillip E. Allen and Noel R. Strader, "VLSI Design Techniques For Analog And Digital Circuits", McGraw-Hill Inc (1990).

Design and Analysis of Various Charge Pump Schemes to Yield Solar Energy Under Various Sunlight Intensities

Anurag Paliwal, R.M. Mehra and Anu Mehra

Abstract A design of low-voltage, ultralow power, four-stage charge pump making utilization of dynamic charge transfer switch scheme targeted to minimize the loss of voltage due to threshold voltage drop and body effect implemented at circuit level implemented in 0.18 μ CMOS process proposed to harvest energy obtained from sunlight. The output stage pumping is tackled by the clocking technique of proposed charge pump instead of output stage configured with the diode. The proposed design is capable to boost solar voltage starting from 0.1 to 3.3 V with the maximum output reported by the simulation as 4.7 V. The sunlight intensity does not remain static it changes, so output voltage is analyzed under various sunlight intensities. From the result of simulation, it is reported that power consumption with the proposed design is 89 nW which is lower in comparison with the existing design. In this work, comparison is made between the various design schemes and analysis is done between the power consumption, frequency, pumping voltage, and delay.

Keywords Voltage loss · Body effect · Charge transfer switch · Switched capacitor · Low power consumption

1 Introduction

One of the most popular renewable energy sources which can be utilized from the surrounding environment to yield the energy is solar energy, since the electrical energy generated from the sunlight proves to be efficient for self-powered portable devices and biomedical implants. The output voltage obtained from individual solar cells ranges within hundreds of millivolt and required to be boosted. As the field of

Anurag Paliwal · Anu Mehra (✉)
ASET, Amity University, Noida, India
e-mail: anumehra2004@yahoo.com

R.M. Mehra
University of Delhi, New Delhi, India

electronics targets mobile communication and wireless networking which caters to operation of sensors wirelessly, this demands efficient supply of power for its operation and requires high power density than the normal energy stored in battery. The important integral part of power management block is the design of power converter. There exist several power converters amongst which is the charge pump. The charge pump consists of switching element capacitor which is responsible to store the charge and retains the stored energy whenever required also known as green energy storage circuit. The extensive application of charge pump comes due to low power consumption, efficient performance, low current driving capability, and small area. A charge pump is responsible for the generation of DC voltage which is higher in comparison to the voltage obtained from the power supply without any need of amplifier or transformer. The design of charge pump is based on capacitor, where the voltage is boosted at each stage which depends upon voltage gain of the circuit. Most of the charge pumps are based on Dickson charge pump mechanism which makes use of two phase clocks which are out of phase and consistently charge during half of each clock cycle [1–10].

The performance of charge pump depends upon various characteristics such as range of input voltage, output voltage range, and value of internal capacitors, output current, and output frequency.

2 Design of Charge Pump Circuit

2.1 Design of Dickson Charge Pump

The design of Dickson charge pump deals with the utilization of pumping capacitors and NMOS connected diode, since there exists the body effect which is associated with the NMOS and threshold voltage as well as large loss in voltage which causes degradation in voltage gain per stage of charge pump. With the increase in number of stages the output voltage degrades due to the body effect. Hence the output voltage has the nonlinear relationship with number of stages which in turn degrades the efficiency of pumping.

In this section, conventional four-stage Dickson charge pump circuit is designed, drain and gate terminals of NMOS (NMOS_1–NMOS_5) have been connected together which acts as diode, so that the charge can be pumped unidirectionally. The two pumping clock signals namely phi1 and phi2 tend to remain in out of phase with respect to each other, the amplitude of two clock signals is tied to supply voltage V_{DD} and the capacitance of the coupling capacitors C_1–C_4 is kept the same as C in the design. The clock charges the capacitor alternately such that the charges are pushed upwards by the two clock signals through MOSFET; as a result this raises the voltage of the node. The fluctuation in the voltage at each node is denoted by ΔV given in the following equation.

When the clock signal phi1 rises from high to low and phi2 (inverted phi1 signal) when goes from low to high, at node 1 voltage settles to $V_1 + \Delta V$ and at node 2 voltage is given by V_2, where the steady state voltage at node 1 is denoted by V_1 and V_2. The MOSFET NMOS_1 and NMOS_3 both are reverse biased, such that the charges accumulated are pushed from node 1 to node 2 through the NMOS_2 and finally the voltage difference existing between nodes 1 and 2 is the threshold voltage defined for the NMOS_2. However, certain drawbacks exist with this design such as threshold voltage of MOSFET (NMOS_1–NMOS_5) which causes degradation in the output voltage (V_{out}) of $\sum_k^n V_{\text{th}}(V_K)$, such that the k-stage pumping gain is degraded by $V_{\text{th}}(V_K)$.

2.2 Design of Pelliconi Cascade Charge Pump Circuit

The Pelliconi charge pump architecture is proposed in order to eliminate the drawbacks discovered in Dickson charge pump. The Pelliconi architecture proposed against the Dickson charge pump is due to the following reason: by connecting the source of MOS to the substrate, the body effect can be controlled. In this cascade architecture the source/substrate has different voltages, as a result the transistor connected in each stage is well isolated with respect to the wells of other stages. With the individual wells existing each stage is responsible to confine the leakage current within their respective stage. Another reason for the proposal of the architecture is high voltage gain, it can reach up to $0.9 * V_{\text{DD}}$ output. The circuit operates with two nonoverlapping clocks; its simple clocking scheme makes the circuit interesting. The architecture proposes the voltage gain 2.6 times that of Dickson charge pump. The number of stages required to pump the voltage is reduced with this architecture, which in turn reduces the parasitic capacitance introduced in the circuit. Another benefit comes from the design capacitance of the last, which is similar to the other stage and eliminates the need for output capacitor.

2.3 Design of Proposed Dynamic Charge Transfer Switch Charge Pump

A new design concept of charge pump has been proposed in this section which makes utilization of dual symmetrical branches and transfer transistor is PMOSFET employed at each stage. A four-stage charge pump is designed to eliminate the problems identified in the existing charge pump, which were voltage loss and body effect. The diode connected MOSFET configuration obtained from the NMOS_1, NMOS_3, NMOS_5, and NMOS_7 is used to develop initial voltage as shown in Fig. 4. The NMOS charge transfer switches NMOS_2, NMOS_4, NMOS_6, and NMOS_8 are employed in the design to control the switches. During the charge

transfer process the switches NMOS_2, NMOS_4, NMOS_6, and NMOS_8 are not driven into off state. The NMOS and PMOS are employed in the circuit to drive the switches dynamically in on/off state in order to reduce the voltage loss problem. On an account to obtain high voltage from the preceding stage backward control scheme is used. With the motive to obtain high gain both the signals CLK1 and CLK2 are out of phase but amplitudes are tied to V_{DD}. The output voltage of the proposed circuit can be represented by the equation,

$$\Delta V = V_{clk} \frac{C_{pump}}{C_{pump} + C_{par}} - \frac{I_0}{f(C_{pump} + C_{par})} \quad (1)$$

where V_{clk} is the amplitude of the clock signal, C_{pump} is the pumping capacitance, C_{par} is the parasitic capacitance existing at each node, I_0 is the output current, and f denotes the clocking frequency. For the condition where the CLK2 is low and CLK1 is high is the same as V_2 while the voltage existing at node 3 is more than the $2\Delta V$ of the voltage at node 1, if

$$2\Delta V > V_{tp} \text{ and } 2\Delta V > V_{tn}(V_2) \quad (2)$$

By the voltage existing at node 3, NMOS_4 and PMOS_2 tend to remain in on state and NMOS_10 in off state, during this phase gate to source voltage is zero.

For the condition where CLK1 falls low and CLK2 rises high, voltage existing at node 1 is V_1. The voltages existing at the node 2 and 3 are greater than the $2\Delta V$,

$$2\Delta V > V_{tn}(V_1) \quad (3)$$

where V_{tn} denotes the threshold voltage of NMOS.

NMOS_4 is free from the control of node 3, during this period NMOS_10 is switched on and PMOS_2 is switched off in order to achieve successful operation of design both Eqs. 1 and 2 are desired. During the short transition time the charges at the gate node of CTS can be injected in the well, when the CLK1 or CLK2 transits from high to low.

The pumping circuitry transistor aspect ratio 0.14/0.18 μ has been kept the same throughout the circuitry and the amplitude of the clock is kept the same as the supply voltage; the pumping capacitance kept at 0.1 pF; smoothing capacitor with the value of 0.1 pF, and supply voltage ranges from 1.8 to 3.3 V.

The proposed design of charge pump possesses high gain characteristic with the elimination of parasitic capacitance, in order to lower the power consumption.

2.4 Figures and Tables

2.4.1 Design of Dickson Charge Pump

Schematic has been designed using TANNER EDA tool, with the help of S-EDIT tool schematic is designed and implemented in 0.18 µ CMOS process, with the help of T-SPICE the script of design is obtained and simulated, hence the waveform is obtained in W-EDIT (Figs. 1 and 2).
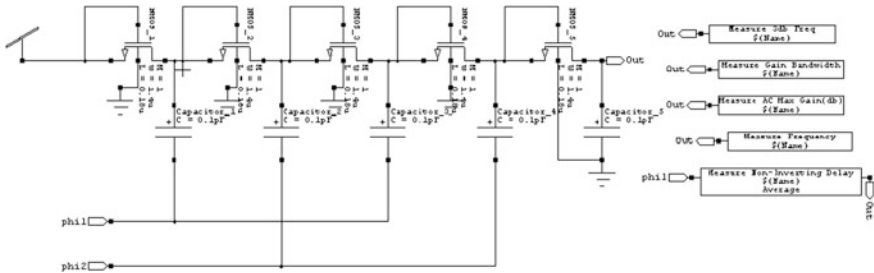
Fig. 1 Schematic of Dickson charge pump
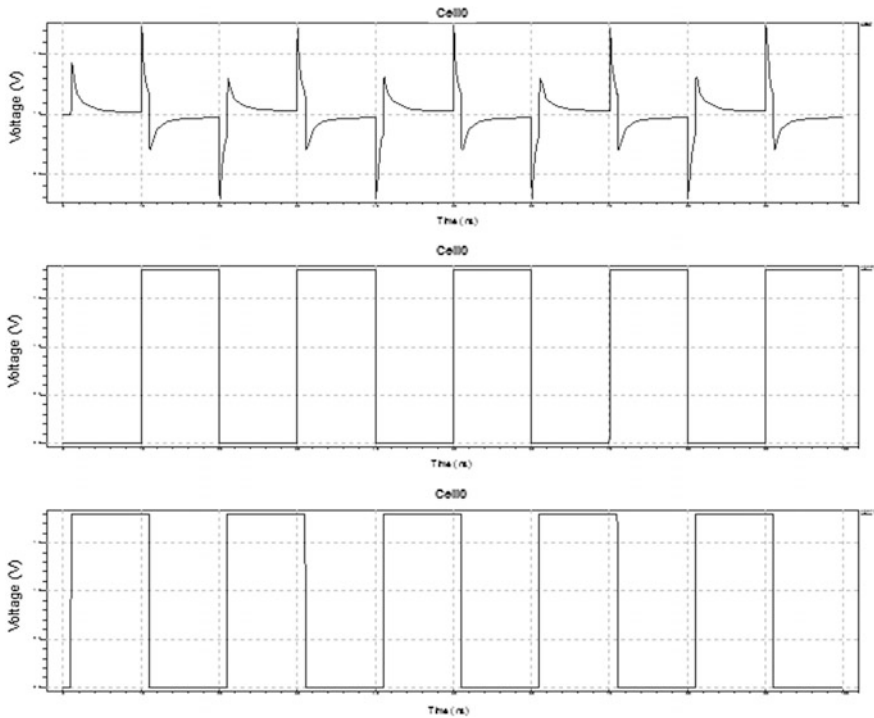
Fig. 2 Simulation waveform of Dickson charge pump

2.4.2 Design of Pelliconi Charge Pump

See Figs. 3 and 4.
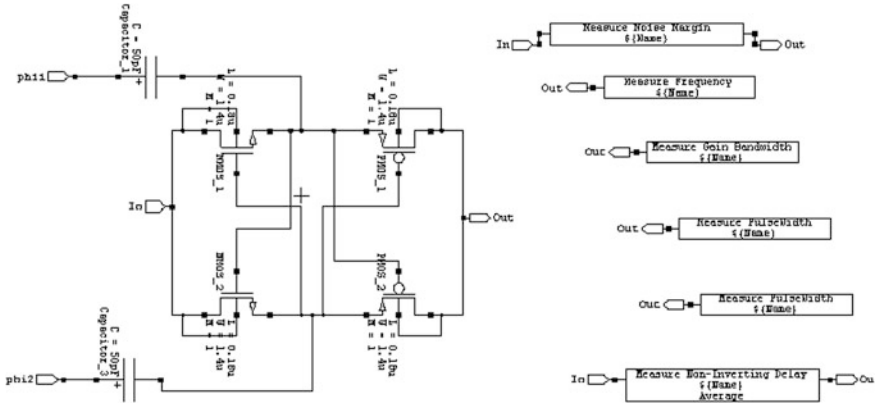
Fig. 3 Schematic design of Pelliconi charge pump
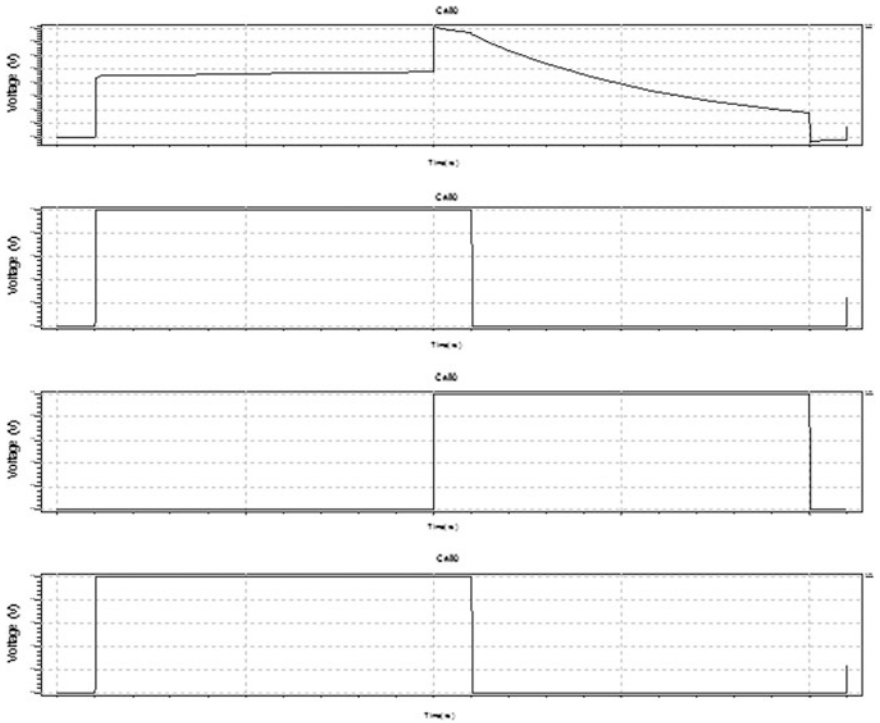
Fig. 4 Simulation waveform of Pelliconi charge pump

2.4.3 Design of Proposed Dynamic Charge Transfer Switch Charge Pump

See Figs. 5 and 6.
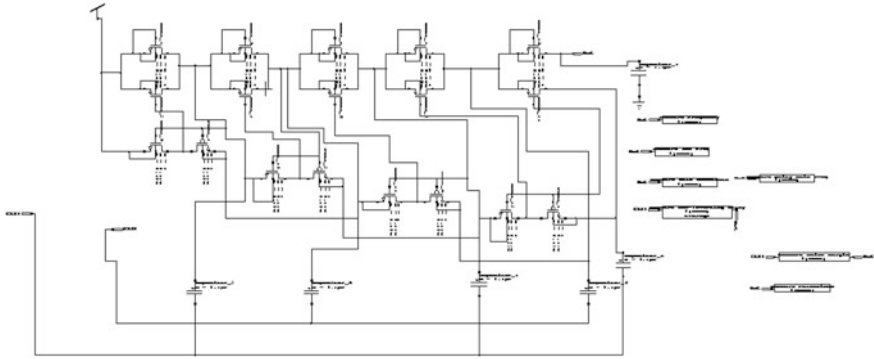
Fig. 5 Schematic of dynamic charge transfer switch

Fig. 6 Output waveform of dynamic charge transfer switch

2.4.4 Performance Analysis Table Between Various Charge Pump Architectures

S. No	V_{IN} (input voltage) (V)	V_{OUT} proposed charge pump (V)	V_{OUT} Dickson charge pump (V)	V_{OUT} Pelliconi charge pump (V)	Power consumption proposed design (W)	Power consumption Dickson pump (W)	Power consumption Pelliconi pump (W)
1	0.1	1.84	1.3	1.5	8.9×10^{-8}	3.8×10^{-5}	1.52×10^{-2}
2	1	2.1	2.1	2.2	6.5×10^{-6}	4.2×10^{-5}	1.57×10^{-2}
3	1.8	2.8	2.9	2.9	5.5×10^{-5}	4.6×10^{-5}	1.6×10^{-2}
4	2.0	3.1	3.2	3.1	6.9×10^{-5}	5.9×10^{-5}	2.0×10^{-2}
5	2.5	3.5	3.4	3.6	1.0×10^{-4}	9.4×10^{-5}	2.7×10^{-2}
6	3.0	3.8	3.7	4.0	1.6×10^{-4}	1.3×10^{-4}	3.2×10^{-3}
7	3.3	4.7	3.9	4.0	1.9×10^{-4}	1.5×10^{-5}	4.1×10^{-3}

S. No	Parameters	Proposed charge pump	Dickson charge pump	Pelliconi charge pump
1	Frequency	4.9 MHz	5.0 MHz	5.0 MHz
2	Delay	1.0×10^{-8} s	5.9×10^{-9} s	3.9×10^{-9} s
3	3 dB frequency	1.487 Hz	1.560 Hz	1.586 Hz
4	Gain	58.24 dB	51.29 dB	54.16 dB

2.4.5 Graphs

See Figs. 7 and 8.

Fig. 7 Relationship between input voltage and output voltage of various charge pump architectures
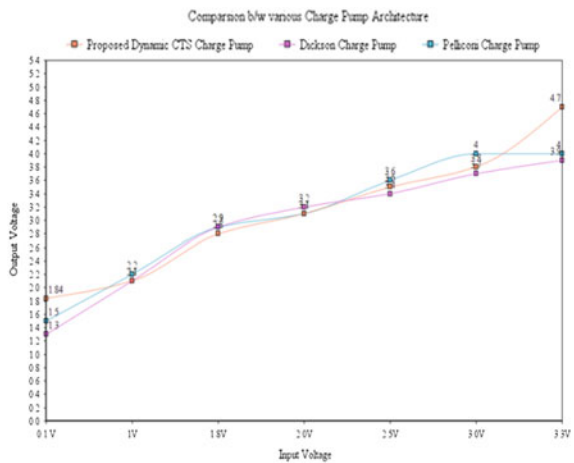
Fig. 8 Graph between input
frequency and gain of various
charge pump architectures
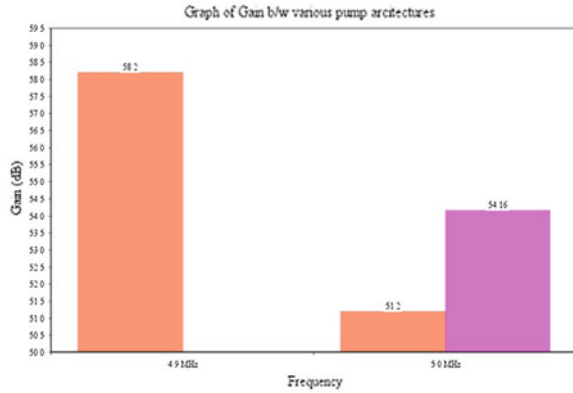
2.5   Program Code

2.5.1   Proposed Dynamic Charge Transfer Switch Charge Pump

```
CCapacitor_5 N_15 CLK1 100f
CCapacitor_6 N_13 CLK2 100f
CCapacitor_1 Out Gnd 100f
CCapacitor_2 N_1 CLK1 100f
CCapacitor_3 N_10 CLK2 100f
CCapacitor_4 N_11 CLK1 100f
MNMOS_3 N_10 N_1 N_1 N_1 NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MNMOS_4 N_10 N_4 N_1 N_1 NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MNMOS_5 N_11 N_10 N_10 N_10 NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MNMOS_6 N_11 N_6 N_10 N_10 NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MNMOS_7 N_13 N_11 N_11 N_11 NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MNMOS_13 Out N_13 N_13 N_13 NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MNMOS_14 N_15 N_14 N_13 N_13 NMOS W=140n L=180n AS=126f
PS=2.08u AD=126f PD=2.08u
MNMOS_1 N_1 Vdd Vdd Vdd NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MNMOS_2 N_1 N_2 Vdd Vdd NMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MPMOS_1 N_2 N_1 N_10 N_10 PMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
```

MPMOS_2 N_4 N_10 N_11 N_11 PMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MPMOS_3 N_6 N_11 N_13 N_13 PMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
MPMOS_4 N_8 N_14 N_15 N_15 PMOS W=140n L=180n AS=126f PS=2.08u
AD=126f PD=2.08u
.param Vpwr=1.8 V
.include "C:\Users\ Documents\Tanner EDA\Tanner Tools v14.1\L-Edit and LVS
\LVS\SPR_Core\hp05.md"
.tran 1n 100n
Vdd Vdd Gnd 1.8
VCLK1 CLK1 Gnd PULSE (0 1.8 1n 20p 20p 10n 20n) ROUND=20n
VCLK2 CLK2 Gnd PULSE (0 1.8 10n 20p 20p 10n 20n) ROUND=20n
.print tran v(CLK1) v(CLK2) v(Out)
.power
.end

2.5.2   Dickson Charge Pump

CCapacitor_1 N_1 phi1 100f
CCapacitor_2 N_2 phi2 100f
CCapacitor_3 N_3 phi1 100f
CCapacitor_4 N_4 phi2 100f
CCapacitor_5 Out Gnd 100f
MNMOS_1 N_1 Vdd Vdd Gnd NMOS W=1.4u L=180n AS=1.26p PS=4.6u
AD=1.26p PD=4.6u
MNMOS_2 N_2 N_1 N_1 Gnd NMOS W=1.4u L=180n AS=1.26p PS=4.6u
AD=1.26p PD=4.6u
MNMOS_3 N_3 N_2 N_2 Gnd NMOS W=1.4u L=180n AS=1.26p PS=4.6u
AD=1.26p PD=4.6u
MNMOS_4 N_4 N_3 N_3 Gnd NMOS W=1.4u L=180n AS=1.26p PS=4.6u
AD=1.26p PD=4.6u
MNMOS_5 Out N_4 N_4 Gnd NMOS W=1.4u L=180n AS=1.26p PS=4.6u
AD=1.26p PD=4.6u
.include "C:\Users\ Documents\Tanner EDA\Tanner Tools v14.1\L-Edit and LVS
\LVS\SPR_Core\hp05.md"
.tran 1n 100n
Vdd Vdd Gnd 1
Vphi1 phi1 Gnd PULSE (0 1.8 1n 20p 20p 10n 20n) ROUND=20n
Vphi2 phi2 Gnd PULSE (0 1.8 10n 20p 20p 10n 20n) ROUND=20n
.print tran v(phi1) v(phi2) v(Out)
.power
.end

2.5.3 Pelliconi Charge Pump

```
CCapacitor_1 phi1 N_1 50p
CCapacitor_3 phi2 N_5 50p
MNMOS_1 In N_5 N_1 In NMOS W=1.4u L=180n AS=1.26p PS=4.6u AD=1.26p
PD=4.6u
MNMOS_2 In N_1 N_5 In NMOS W=1.4u L=180n AS=1.26p PS=4.6u AD=1.26p
PD=4.6u
MPMOS_1 Out N_5 N_1 Out PMOS W=1.4u L=180n AS=1.26p PS=4.6u
AD=1.26p PD=4.6u
MPMOS_2 Out N_1 N_5 Out PMOS W=1.4u L=180n AS=1.26p PS=4.6u
AD=1.26p PD=4.6u
.include "C:\Users\Documents\Tanner EDA\Tanner Tools v14.1\L-Edit and LVS
\LVS\SPR_Core\hp05.md"
.tran 1n 50n
Vphi1 phi1 Gnd PULSE (0 2.5 1n 20p 20p 10n 20n) ROUND=20n
Vdd Vdd Gnd 1.8
Vphi2 phi2 Gnd PULSE (0 2.5 10n 20p 20p 10n 20n) ROUND=20n
VIn In Gnd PULSE (0 2.5 1n 20p 20p 10n 20n) ROUND=20n
.print tran v(phi1) v(phi2)v(In) v(Out)
.power
.end
```

3 Conclusion

An enhanced design technique is utilized for the charge pump, CTS charge pump from the simulation is reported so that the power consumption obtained is 89 nW which is far better than the other charge pump showcased in the paper. The voltage gain observed with the proposed design is 58.24 dB which is high in comparison to the other design architecture. The design is able to function at very low voltage giving better pump voltage having least parasitic capacitance which is also a reason that gives good power efficiency. The simulation results showcase the maximum voltage that can be obtained from the proposed design as 4.7 V, which is better in comparison to others. The simulation has been carried under the same frequency range of 5 MHz. The minimum delay is offered by the Pelliconi charge pump. The designed charge pump is suitable for harvesting the solar energy capable to generate high voltage with respect to low voltage obtained from the individual solar cell.

References

1. Hye-Im Jeong, Jung-Woong Park, Ho-Yong Choi, and Nam-Soo Kim: High Performance Charge Converter with Integrated CMOS Feedback Circuit, Transaction on Electrical and Electronic Materials, ISSN: 1229–7607, March (2014).
2. Sung-Dae Yeo, Young-Jin Jang, Kyoung-Kun Lee, Seong-Jong Kim and Seong-Kweon Kim: Charge Pump Circuit design for Low Input Voltage, International Journal of Control and Automation, Vol. 7, No. 5, pp. 259–268 (2014).
3. Kashyap K. Patel, Nilesh D. Patel, Kruti P. Thakore: Charge pump, Loop Filter and VCO for PLL using 0.18 micron technology, IOSR journal, ISSN No.: 2319–4197, Volume 2, Issue 4, pp. 21–25 (2013).
4. Charge Pump in 0.18 micron CMOS process, International Journal of Scientific and Engineering Research, Volume 4, Issue 8, August-2013 International Journal of Scientific & Engineering Research, ISSN 2229–5518, Volume 4, Issue 8 (2013).
5. Jyoti Gupta, Ankur Sangal and HEmalata Verma: High Speed CMOS Charge Pump Circuit for PLL Applications Using 90 nm CMOS Technology, Middle-East Journal of Scientific Research 12, ISSN 1990-9233 (2012).
6. V. Sujatha, R.S.D. Wahida: High Performance Charge Pump PLL with Low Current Mismatch, IJCSI, ISSN (Online):1694-0814, Vol. 9, Issue 1, No 2 (2012).
7. Feng Peng, Li Yunlong, and Wu Nanjian: A high efficiency charge pump circuit for low power application, Journal of Semiconductors, Volume 31, January (2010).
8. J. Carlson: A 20 mV input boost converter with efficient digital control for thermoelectric energy harvesting, IEEE J. Solid-State Circuits, vol. 45, no. 4, pp. 741–750 (2010).
9. H. Shao, C-Y. Tsui and W. H. Ki, A micro power management system and maximum output power control for solar energy harvesting applications, ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED), 298–303 (2007).
10. G. Palumobo, D. Pappalardo, Charge Pump Circuits with Only Capacitive Loads: Optimized Design, IEEE Transactions on Circuits and Systems, 53(2), 128–213 (2006).

Performance Evaluation of Speech Synthesis Techniques for English Language

Sangramsing N. Kayte, Monica Mundada, Santosh Gaikwad and Bharti Gawali

Abstract The conversion of text to synthetic production of speech is known as text-to-speech synthesis (TTS). This can be achieved by the method of concatenative speech synthesis (CSS) and hidden Markov model techniques. Quality is the important paradigm for the artificial speech produced. The study involves the comparative analysis for quality of speech synthesis using hidden Markov model and unit selection approach. The quality of synthesized speech is evaluated with the two methods, i.e., subjective measurement using mean opinion score and objective measurement based on mean square score and peak signal-to-noise ratio (PSNR). Mel-frequency cepstral coefficient features are also extracted for synthesized speech. The experimental analysis shows that unit selection method results in better synthesized voice than hidden Markov model.

Keywords TTS · MOS · HMM · Unit selection · Mean · Variance · MSE · PSNR

1 Introduction

A speech synthesis system is a computer-based system that produce speech automatically, with the conversion steps of grapheme-to-phoneme transcription of the sentences with the inclusion of prosodic features. The synthetic speech is generated

S.N. Kayte (✉) · Monica Mundada · Santosh Gaikwad · Bharti Gawali
Department of Computer Science and Information Technology,
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India
e-mail: bsangramsing@gmail.com

Monica Mundada
e-mail: monicamundada5@gmail.com

Santosh Gaikwad
e-mail: santosh.gaikwadcsit@gmail.com

Bharti Gawali
e-mail: bharti_rokade@yahoo.co.in

with the available phones and prosodic features from training speech database [1, 2]. The speech units are classified into phonemes, diaphones, and syllables. The output of speech synthesis system depends on the size of the speech units involved in the execution of the method. The designing of text-to-speech system is organized into two parts: front end and back end. In the front-end module, initially the input text comprising of symbols like numbers and abbreviations are transformed into the equivalent words. This process is defined as the text normalization, preprocessing, or tokenization. The text into prosodic units like phrases, clauses, and sentences are assigned with the phonetic transcriptions. The back-end phase produces the synthesis of the particular speech with the use of output provided from the front end. The symbolic representations from first step are converted into sound speech and the pitch contour, phoneme durations and prosody are incorporated into the synthesized speech.

The paper is structured in five sections. The techniques of speech synthesis are described in Sect. 2. Database for synthesis system is explained in Sect. 3. Section 4 explains speech quality measurement. Section 5 is dedicated with experimental analysis followed by conclusion.

2 Concatenate Synthesis

Concatenate speech synthesis is a method where speech is generated by concatenating speech units one after the other as per the requirement. There are three different types of concatenate speech synthesis. They are domain specific synthesis, diphone synthesis, and unit selection synthesis [2]. The focus of the paper is unit selection synthesis.

In this method, the database is built up with all phones present in the particular language. The design of such database includes well-labeled phones with high-quality utterances. The synthesized speech output signal is generated with the concatenated parts from the database [2]. The output speech produced in this method has greater impact on intonation, way of speaking style, and emotions associated with the speech. Also, the construction of large database corpus produces the high quality of synthesized speech. The USS (unit selection synthesis) extracts the prosodic and spectral part from input speech signal during the training part. In synthesis part, the analysis of text is done and prosody is incorporated with the use of algorithm and artificial speech is produced [2]. In USS, initially the text is converted into phones of the particular segment. Then the phones are assigned the labels like vowels, semivowels, and consonants. With the help of acoustic trees the IDs are generated for the given input. At the final step, the speech is synthesized with the USS algorithm with the incorporation of the needed prosody elements.

3 Hidden Markov Model-Based Speech Synthesis

Hidden Markov model synthesis is also called statistical parametric synthesis of speech. The significance of this method allows the variation in voice easily. In this method, the speech is synthesized on the basis of the parameters extracted from the recorded utterances. The HTS system, the context-dependent hidden Markov model generates the excitation parameters [3]. Thus they are treated as input for the speech waveforms in the later stage. In HMM-based speech synthesis there is no need for large database corpus and the quality of voice maintained [2]. In the two stages of HMM execution part initially the spectrum and excitation parameters are extracted from speech database and modeled by context-dependent hidden Markov model. In the next stage, context-dependent hidden Markov model is concatenated according to the text to be synthesized. Then spectrum and excitation parameters are generated from the hidden Markov model using a speech parameter generation algorithm. Both the techniques are implemented using Festival framework [4].

4 Speech Quality Measurement

Speech quality measurement is the level of audible and perceptible level of the output speech. There are two methods for performing this relative task.

4.1 Subjective Quality Measure

The quality associated with the produced speech depends on the paradigms of subjective perception and judgment process. In this method, the personal assessment is done from the individual with the help of mean opinion score (MOS) test. The subject evaluating the sentence assigns the grades depending on the quality of speech with respect to 5-point scale. In which the grade 1 is assigned to the least or unsatisfactory speech and the grade 5 to the excellent speech quality. In this the system is trained with the analysis performance report of each perceptual listener involved in the test [5, 6].

4.2 Objective Quality Measure

This method computes for the automated real-time quality measurement. The perceptual listener goes through the computational algorithm of the system. Real-time quality monitoring is gained only with the objective speech quality measurement. This method proves more reliable and accurate as compared to

subjective listening experiments. In the objective quality measure mean square error (MSE) and peak signal-to-noise ratio (PSNR) techniques were used.

(a) Mean Square Error (MSE)

The mean squared error (MSE) is defined as the average of the squares of the errors, that is, the difference between the estimator value and the estimated value. The difference occurs because of randomness or because the estimator does not account for information that could produce a more accurate estimation of speech synthesis [7].

(b) Peak Signal-to-Noise Ratio (PSNR)

PSNR is measured in terms of the logarithmic decibel scale. The use of PSNR is to measure the quality of reconstruction of signal and image. The signal in this case is the original utterance, and the noise is the error introduced during the process of speech synthesis [8]. Peak signal-to-noise ratio (PSNR) is measured as the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the quality of its representation.

4.3 Signal-Based Quality Measure

In the signal-based quality measure the perceptual evaluation of speech quality (PESQ) technique [9–11] is followed. The current version P.862 of PESQ algorithm is highly reliable [12]. For this experiment we proposed MFCC features for the signal-based quality measure. Mel-frequency cepstral coefficients (MFCC) technique is robust and dynamic technique for speech feature extraction. The fundamental frequency, prosodic, energy variation in the syllable and many other features are studied with MFCC feature set. For the quality measure we extracted 13 features from synthesized speech and original speech file.

5 Speech Database

The speech database collected for this experiment includes the sentences from philosophy and short stories. The sentences were recorded by male and female speakers. Male speaker was with South Indian accent and female voice was with normal accent. The male and female both were from academic field and practiced the session. The recording was done in noise-free environment. The speech signal was sampled at 16 kHz. The set of 30 sentences were synthesized using unit selection and hidden Markov model. Noise-free lab environment with multimedia laptop speaker was used to play these utterances to the postgraduate students. The students were of age group 22–25, with no speech synthesis experience.

6 Experimental Analysis

Analysis of Mean Opinion Score (MOS)

MOS is calculated for subjective quality measurement. It is calculated for the synthesized speech using the unit selection synthesis and HMM approach. It was counseled to the listeners that they have to score between 01 and 05 (Excellent—05; Very good—04; Good—03; Satisfactory—02; Not understandable—01) for understandability. The mean of the scores given by each individual subject for ten sentences of the unit selection approach is shown in Table 1. The details of MOS score obtained from HMM speech synthesis method for ten sentences are shown in Table 2.

Table 1 Unit selection speech synthesis of the scores given by each subject for each synthesis system

Sentence	Subject									
	1	2	3	4	5	6	7	8	9	10
1	5	5	5	5	4	4	5	4	4	5
2	5	5	4	5	5	4	5	4	4	5
3	4	4	5	4	3	3	4	2	5	4
4	5	4	4	5	4	4	5	5	5	5
5	5	5	5	5	4	4	5	3	3	5
6	5	4	5	5	5	4	5	4	4	5
7	4	5	4	4	4	4	4	4	4	4
8	4	4	5	4	4	5	4	5	5	4
9	5	3	5	5	3	4	5	3	5	5
10	5	5	4	5	4	4	5	4	4	5

Table 2 HMM-based speech synthesis of the scores given by each subject for each synthesis system

Sentence	Subject									
	1	2	3	4	5	6	7	8	9	10
1	4	5	5	5	5	4	5	4	5	3
2	3	5	4	4	3	4	3	3	3	4
3	5	4	4	4	4	3	4	4	3	5
4	5	4	4	4	3	4	4	5	4	3
5	3	4	5	5	5	3	4	4	5	4
6	2	4	4	3	4	2	4	5	4	4
7	3	5	5	2	5	1	5	3	3	5
8	4	4	4	1	4	2	5	4	2	3
9	4	3	3	3	5	3	4	5	2	4
10	5	4	4	1	2	2	4	4	4	5

Table 3 Mean and variance of the scores obtained across the subjects from unit selection and HMM approach

Subject	Unit selection method		HMM synthesis approach	
	Mean score	Variance	Mean score	Variance
1	4.56	0.25	3.90	1.19
2	4.23	0.52	2.73	1.71
3	4.03	0.79	2.56	1.35
4	4.56	0.25	2.80	1.61
5	4.10	0.43	2.46	0.947
6	4.03	0.37	3.10	1.05
7	4.56	0.25	2.80	1.68
8	3.96	0.72	2.33	1.26
9	4.16	0.62	2.73	1.37
10	4.63	0.24	2.63	1.48

The mean and variance of the score obtained according to the subject using the two experimental approaches, i.e., unit selection and HMM-based speech synthesis approach are shown in Table 3.

It is observed that from Tables 3 and 4 mean scores increase with the increase in the syllable coverage.

(a) PSNR and MSE Quality Measure

The PSNR and MSE methods were used for subjective quality measure of speech synthesis based on hidden Markov model and unit selection approach. Table 4 represents the MSE and PSNR values for unit selection-based speech synthesis. HMM-based speech synthesis using MSSE and PSNR is shown in Table 5.

Table 4 MSE and PSNR values for unit selection-based speech synthesis

Sr. No	Original speech file	Synthesized file	MSE	PSNR
1	A1	a1	7.94	3.30
2	A2	a2	4.57	6.72
3	A3	a3	1.02	3.21
4	A4	a4	3.70	4.20
5	A5	a5	7.61	2.57
6	A6	a6	5.32	1.26
7	A7	a7	8.06	7.56
8	A8	a8	7.20	1.29
9	A9	a9	9.25	3.24
10	A10	a10	7.01	4.08
Average			6.168	3.743
Quality (100-Average)			93.83	96.26

Table 5 MSE and PSNR values for hidden Markov model speech synthesis

Sr. No	Original speech file	Synthesized file	MSE	PSNR
1	B1	b1	9.15	7.315
2	B2	b2	8.38	6.24
3	B3	b3	13.5	5.25
4	B4	b4	10.4	8.40
5	B5	b5	9.26	8.76
6	B6	b6	9.38	9.42
7	B7	b7	10.10	9.05
8	B8	b8	9.63	6.56
9	B9	b9	10.42	8.49
10	B10	b10	12.40	7.44
Average			10.26	7.69
Quality (100-Average)			89.73	92.31

Table 6 Comparative result of unit and HMM speech synthesis

Sr. No	Approach of synthesis	MFCC mean (%)	MFCC STD (%)	MFCC var (%)	MSE (%)	PSNR (%)
1	HMM	80	80	80	89.73	92.31
2	Unit selection	90	90	80	93.83	96.26

Table 6 shows the comparative performance of both unit and HMM for accent recognition using MFCC, MSE, and PSNR techniques.

From Table 6, it is observed that the unit selection-based accent identification gives better performance than HMM-based speech synthesis.

(b) Mel-frequency Cepstral Coefficients

The signal-based synthesis quality measure is experimented for unit selection and hidden Markov model-based speech synthesis. For the performance variation, we calculated the mean, standard deviation, and variance as a statistical measure. The details of sentences and label used for the unit selection-based speech synthesis are described in Table 7.

The MFCC mean-based performance of unit selection-based synthesis is shown in Table 8. Table 9 represents the details of standard deviation of MFCC for unit selection speech synthesis.

The sentences used for hidden Markov model-based synthesis using MFCC-based method is shown in Table 10. The details of performance of MFCC mean and standard deviation for hidden Markov model-based speech synthesis are shown in Tables 11 and 12, respectively.

Table 7 Sentences and label used for unit selection-based speech synthesis

Sr. No	The original sentence	Label used for original speech file	Label used for synthesis speech file
1	Will we ever forget it	A1	a1
2	There was a change now	A2	a2
3	I had faith in them	A3	a3
4	She turned in at the hotel	A4	a4
5	We'll have to watch our chances	A5	a5
6	It was a curious coincidence	A6	a6
7	There was nothing on the rock	A7	a7
8	I have no idea, replied Philip	A8	a8
9	Anyway, no one saw her like that	A9	a9
10	Surely I will excuse you, she cried	A10	a10

Table 8 The performance of MFCC mean-based unit selection speech synthesis

Synthesized speech										
Original speech signal		**a1**	**a2**	**a3**	**a4**	**a5**	**a6**	**a7**	**a8**	**a9**	**a10**
	A1	**0.120**	3.242	2.31	1.392	3.42	3.008	2.983	5.094	7.01	1.234
	A2	1.281	**0.009**	2.111	2.453	7.632	1.90	4.02	1.223	8.01	3.04
	A3	1.453	3.21	**0.080**	3.25	1.99	2.843	3.921	2.963	2.093	6.70
	A4	3.02	1.230	2.564	**0.899**	2.786	5.453	1.672	1.981	2.67	3.45
	A5	2.40	1.450	5.432	2.932	**0.021**	3.921	6.05	4.675	7.00	3.674
	A6	3.896	8.09	3.983	2.732	3.674	**0.673**	2.843	5.03	3.894	4.92
	A7	1.893	1.563	2.03	2.100	4.92	3.67	**1.460**	1.273	3.521	7.38
	A8	2.932	3.674	2.732	3.721	3.567	2.732	3.876	**0.783**	2.673	3.643
	A9	1.776	2.732	4.332	5.893	2.783	3.874	2.743	4.87	**1.091**	3.021
	A10	2.873	2.983	1.873	1.90	2.763	1.563	4.02	1.788	2.032	4.328

Table 9 The performance of MFCC STD-based unit selection speech synthesis

Synthesized speech										
Original speech signal		**a1**	**a2**	**a3**	**a4**	**a5**	**a6**	**a7**	**a8**	**a9**	**a10**
	A1	**0.456**	1.200	1.892	3.902	3.872	2.893	5.783	3.872	4.500	2.673
	A2	1.231	**0.632**	2.762	2.090	1.988	2.763	1.235	1.6532	4.673	6.011
	A3	5.632	3.982	**1.050**	1.928	2.782	2.782	1.892	1.292	3.020	5.873
	A4	3.092	2.093	4.092	**1.837**	4.932	3.091	5.781	3.982	2.983	2.983
	A5	1.882	1.0291	3.0281	3.091	2.872	**1.092**	4.092	3.982	3.982	5.021
	A6	3.091	4.873	3.982	2.983	1.022	**0.932**	1.829	2.893	4.092	4.093
	A7	2.983	3.092	2.993	4.984	2.831	8.011	**1.920**	1.892	3.001	3.092
	A8	5.011	3.921	4.984	5.092	2.931	4.982	1.092	**0.982**	1.778	4.832
	A9	2.938	5.011	2.932	6.091	2.983	1.921	2.932	4.632	**1.092**	1.920
	A10	3.092	1.821	1.9082	3.921	4.921	3.842	4.983	4.530	2.321	**0.210**

Table 10 Sentences and label used hidden Markov model speech synthesis

Sr. No	The original sentence	Label used for original speech file	Label used for synthesized speech file
1	Gad, do I remember it	B1	b1
2	I can see that knife now	B2	b2
3	They robbed me a few years later	B3	b3
4	Now, you understand	B4	b4
5	He caught himself with a jerk	B5	b5
6	How does your wager look now	B6	b6
7	It won't be for sale	B7	b7
8	Now it was missing from the wall	B8	b8
9	It is the nearest refuge	B9	b9
10	He can care for himself	B10	b10

Table 11 The performance of MFCC mean-based hidden Markov model speech synthesis

Synthesized speech											
Original speech signal		b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
	B1	**0.234**	3.781	5.155	6.280	2.662	5.442	3.601	5.432	3.970	11.950
	B2	5.227	**0.200**	8.700	6.191	1.7100	5.327	5.465	8.932	2.242	6.126
	B3	1.900	3.815	**0.210**	9.044	3.123	1.090	2.120	3.030	5.445	9.580
	B4	5.559	0.934	1.980	**0.936**	1.2315	1.780	2.090	2.050	6.318	12.272
	B5	2.980	3.800	3.178	2.153	**0.119**	2.130	2.150	3.092	2.339	23.011
	B6	2.051	9.1400	**0.221**	1.050	1.781	3.873	1.363	1.030	3.335	16.09
	B7	2.463	4.990	5.900	2.191	2.130	2.172	**0.181**	1.192	2.991	8.700
	B8	4.839	6.566	2.550	2.781	1.630	1.152	0.800	**0.500**	5.344	9.811
	B9	3.992	7.502	3.300	2.050	1.980	1.050	1.262	2.111	**0.300**	7.020
	B10	7.793	6.176	3.528	2.128	6.512	7.900	3.512	1.393	**0.810**	1.23

Table 12 The performance of MFCC STD-based hidden Markov model speech synthesis

Synthesized speech											
Original speech signal		b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
	B1	**0.110**	1.221	2.119	1.290	1.890	1.233	2.030	3.01	1.178	5.020
	B2	2.120	**1.20**	1.900	3.402	7.680	8.900	11.02	2.678	2.343	7.890
	B3	0.900	1.123	**2.342**	3.784	3.134	4.030	2.178	9.01	2.05	5.030
	B4	2.564	3.100	2.870	**0.900**	1.760	2.345	2.403	3.050	2.870	3.435
	B5	1.890	1.450	2.123	**0.200**	5.40	2.656	1.934	2.999	7.030	9.01
	B6	0.890	1.543	2.212	3.210	3.521	**0.321**	2.986	1.776	2.832	3.02
	B7	2.320	1.564	2.220	5.040	3.442	5.021	**0.210**	3.450	3.887	7.00
	B8	1.747	2.030	5.020	2.456	3.022	**0.884**	3.007	4.302	2.345	1.284
	B9	2.336	3.020	3.040	3.121	3.998	7.060	4.990	3.2020	**1.02**	1.998
	B10	1.987	2.030	2.0440	2.312	3.220	1.228	6.070	3.009	4.006	**0.320**

7 Conclusion

The quality of speech synthesis is experimented using MOS score, MSE, PSNR, MFCC-based techniques for hidden Markov model (HMM) and unit selection synthesis (USS) approach. The MFCC-based method is evaluated using the mean, standard deviation, and variance. For all the estimated methods the unit selection method gives a better performance than hidden Markov model techniques as the database is small.

References

1. Mohammed Waseem, C.N Sujatha, "Speech Synthesis System for Indian Accent using Festvox", International Journal of Scientific Engineering and Technology Research, ISSN 2319-8885 Vol. 03, Issue. 34 November-2014, Pages: 6903–6911.
2. Sangramsing Kayte, Kavita Waghmare, Dr. Bharti Gawali "Marathi Speech Synthesis: A review" International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 3 Issue: 6, 3708–3711.
3. T. Yoshimura, K. Tokuda, T. Masuko, T. Kobayashi and T. Kitamura, "Simultaneous Modeling of Spectrum, Pitch and Duration in HMM-Based Speech Synthesis" In Proc. of ICASSP 2000, vol 3, pp. 1315–1318, June 2000.
4. A. Black, P. Taylor, and R. Caley, "The Festival Speech Synthesis System. System documentation Edition 1.4, for Festival Version 1.4.3 27th December 2002.
5. Series P: Telephone Transmission Quality "Methods for objective and subjective assessment of quality" Methods for Subjective Determination of Transmission Quality ITU-T Recommendation P.800.
6. ITU-T P.830, Subjective performance assessment of telephone-band and wideband digital codecs.
7. Lehmann, E. L.; Casella, George. "Theory of Point Estimation (2nd ed.). New York: Springer. ISBN 0-387-98502-6. MR 1639875.
8. Huynh-Thu, Q.; Ghanbari, M. (2008). "Scope of validity of PSNR in image/video quality assessment". Electronics Letters 44 (13): 800. doi:10.1049/el:20080522.
9. SR Quackenbush, TP Barnwell, MA Clements, Objective Measures of Speech Quality (Prentice-Hall, New York, NY, USA, 1988).
10. AW Rix, MP Hollier, AP Hekstra, JG Beerends, PESQ, the new ITU standard for objective measurement of perceived speech quality—part 1: time alignment. Journal of the Audio Engineering Society 50, 755–764 (2002).
11. JG Beerends, AP Hekstra, AW Rix, MP Hollier, PESQ, the new ITU standard for objective measurement of perceived speech quality—part II: perceptual model. Journal of the Audio Engineering Society 50, 765–778 (2002).
12. ITU-T P.862, Perceptual evaluation of speech quality: an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech. codecs 2001.

Systematization of Reliable Network Topologies Using Graph Operators

A. Joshi and V. Subedha

Abstract The aim of this paper is to study and compare the reliability of networks using Wiener index. The computer communication using electronic messaging has increased in recent years. The calculation of the overall reliability of the networks becomes an important problem. This paper presents the topology invariant which calculates the reliability of the newly constructed network using graph operations tensor product and Cartesian product in Topology theory. The simulated experimentation of the proposed topology invariant for the new topologies have been done and compared with existing topologies.

1 Introduction

The communication networks play an important place in day-to-day life. It is very important to construct networks with high reliability. For successful construction of such networks, we used graphs and its operators. Thus we need to introduce the new network topology so that the reliability is improved. The analysis of the reliability focuses on following factors as

 (i) Constructing new topologies using graph operators [7].
 (ii) Estimation of the reliability of the newly constructed topologies using Reliability Wiener Index [16] in terms formula constructed using distance (hops).

A. Joshi (✉)
Department of Computer Science and Engineering, St. Peter's University, Chennai, India
e-mail: joshiseeni@gmail.com

V. Subedha
Department of Computer Science and Engineering, Panimalar Institute of Technology, Chennai, India
e-mail: subedha@gmail.com

(iii) Comparing the reliabilities of the new network topologies and existing topologies [10].

Reliability is measured using different methods. Overall network reliability is calculated using probability of every connected node in the network. Reliability is efficient if it is calculated from the source to sink. Instead of one sink we can consider many sink say K called K-sink reliability, which is the probability that all nodes are communicated to all sinks.

Many algorithms and techniques available in the literature for the computation of reliability measures using probability but they are applicable for networks with limited size. Construction and modification of communication networks with high reliability is a difficult task. This problem can be solved using graphs with n nodes [13]. In Kumar et al. [1] developed a genetic algorithm to design large size computer networks. Dengiz et al. [2] focused on large size communication network design using genetic algorithm, but standardized it to the all sink design problem to give an effective methodology.

This paper is intent to present the new topologies from existing topology by applying Topology operators and comparing the reliability of old and new topologies.

2 Reliability Wiener Index

Now we are interested in the study of construction of networks with high reliability. While constructing such type of networks, we are not considering weights to edges. There are different measures available in the literature for weighted topologies. The minimum total distance implies maximum total reliability. The calculation of reliability is considered in terms of sum of distance between connected nodes. The sum of distance between every pair of connected node is defined as Wiener index.

The Wiener index $W(G)$ of a topology G with vertex set V with n nodes is defined as the sum of distances between all connected nodes of G [15].

$$W(G) = \frac{1}{2} \sum \sum d_{ij}$$

where dij is the distance between the nodes i and j.

The out-reliability $R_1^+(i)$ of a vertex i in a digraph G of n nodes by [12]

$$R_1^+(i) = \sum_{j=1}^{n} \bar{F}_{ij}$$

where $\bar{F}_{ij}$ number of the hops in the most reliable path from i to j.

The out-reliability Wiener index [3] of G is defined as

$$W_{R_1^+}(G) = \sum_{v \in V(G)} R_1^+(v)$$

The out-reliability Wiener index of G is a measure of reliability.

3 Tensor Product Topology

3.1 Construction of Topology

It is identified that the tensor product [5] of C_3 (ring topology with three nodes) and C_n (Ring topology with n nodes) is a connected Topology. The four-regular Topology $C_3 \wedge C_n$ is a connected topology [9].
  The construction of a tensor product of C_3 and C_n contains two parts

 (i) Construction of a cycle with $3n$ vertices.
(ii) Drawing lines with in the cycle.

Algorithm
    Input: Ring topology with 3 vertices and Ring topology with n vertices.
    Output: A 4-regular topology with 3n nodes.
Begin
for i =1 to 3
    for j = 1 to n

  $V = \{w_i^{\,j}\}$

  $E_1 = \left\{(w_1^1, w_2^n) \cup (w_2^n, w_3^{n-1}) \cup (w_3^{n-1}, w_1^n)\right\}$

    for k = n to 3

  $E_2 = \left\{(wi^k, w_{i+1}^{k+1}) \cup (w_{i+1}^{k\,1}, w_{i+2}^{k\,2}) \cup (w_{i+2}^{k\,2}, w_i^{k\,1})\right\}$
  $E_3 = \left\{(w_1^{\,2}, w_2^1) \cup (w_3^n, w_1^1)\right\}$
  $E_4 = \left\{(w_1^{\,1}, w_2^2) \cup (w_2^{\,2}, w_3^3)\right\}$

    for k = 2 to n-2

  $E_5 = \left\{(wi^k, w_{i+1}^{k+1}) \cup (w_{i+1}^{k+1}, w_{i+2}^{k+2}) \cup (w_{i+2}^{k+2}, w_i^{k+1})\right\}$
$E_6 = \left\{(w_2^n, w_3^1) \cup (w_3^1, w_1^n) \cup (w_1^n, w_2^1) \cup (w_2^1, w_3^2) \cup (w_3^2, w_1^1)\right\}$
    $E = E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6$
End

Fig. 1 Tensor product topology of C_3 and C_4
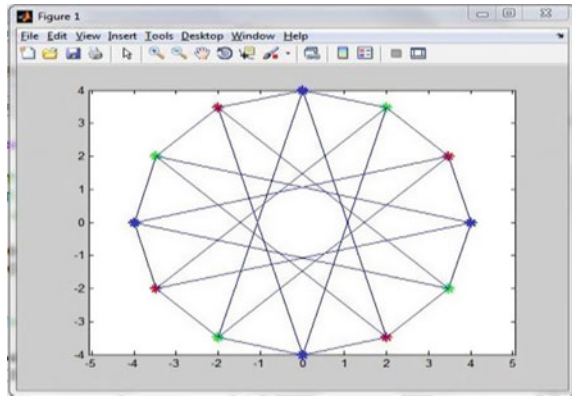
See Fig. 1.

4   Reliability Comparisons

See Fig. 2.

5   Cartesian Product of Bus Topology P_m and Ring Topology C_n

The Cartesian product of bus topology P_m and ring topology C_n is a structured web topology which is a combination of m ring topologies and n bus topologies in the form of web [6]. The Cartesian product of bus topology P_4 and ring topology C_7 is a topology consist of 4 ring topologies with seven nodes and seven bus topologies with four nodes.

Fig. 2 Comparison of reliability C_n and tensor product topology
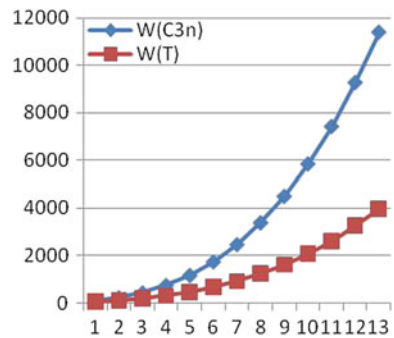
5.1 Pseudocode

```
m=input('Enter m: ');
%P n=input('Enter n: ');
%C c=ones(m,n+1);
s=ones(m,n+1);
t=linspace(0,2*pi,n+1);
for i=1:m
  s(i,1:n+1)=i*sin(t);
  c(i,1:n+1)=
  i*cos(t); hold on,
  %plotting points
  plot(s(i,1:n+1),c(i,1:n+1),'r*','MarkerSize',10);
  pause(1);
  %Drawing circles
  plot(s(i,1:n+1),c(i,1:n+1),'Color','g','MarkerSize',30);
  pause(1);
  hold
off end
%Drawing Lines
for i=1:m-1
  for j=1:n
   a1=[s(i,j) s(i+1,j)];
   b1=[c(i,j)
   c(i+1,j)]; hold on,
   line(a1,b1,'LineWidth',1,'Color','b');
   pause(1);
   hold off
   end
end
```

See Figs. 3 and 4.

Fig. 3 Webtopology constructed from P_5 and C_{20} i.e., m = 5 and n = 20
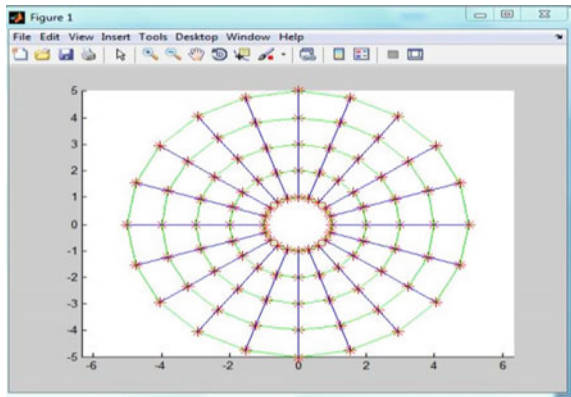
Fig. 4 Comparison chart
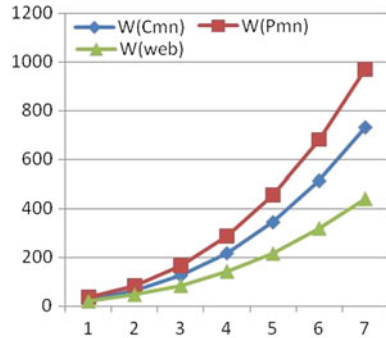
6 Conclusion

We presented a new methodology for constructing topologies with existing topologies. The generalization developed in this paper makes more reliable and flexible method since the construction of topologies is not restricted to the selection of an integer n. The reliability of newly constructed topology in Table 1 shows that the tensor product topology is having three times higher performance than the ring topology when $n >= 26$ and this shows the convergence property in mathematics. Similarly, Table 2 shows that the structured web topology performance is higher than the path and ring topologies.

Table 1 Reliability for different values of n

N	$W(C_{3n})$ (1)	$W(C_3 \wedge C_n)$ (2)	(1)/ (2)
3	90	63	1.43
4	216	120	1.8
5	420	195	2.15
6	729	315	2.3
7	1155	462	2.5
26	59,319	20,085	2.9534
100	3,375,000	1,126,200	2.9968
100,000	3.3750e + 015	1.1250e + 015	3.0000
1,000,000	3.3750e + 018	1.1250e + 018	3.0000

Table 2 Reliability for different values of n and m

m	n	$W(C_{mn})$	$W(P_{mn})$	W (web)
2	3	27	35	21
2	4	64	84	48
2	5	125	165	85
2	6	216	286	144
2	7	343	455	217
2	8	512	680	320
2	9	729	969	441
3	3	90	120	63
3	4	216	286	136
3	5	420	560	235
3	6	729	969	387
3	7	1155	1540	616

References

1. Kumar. A, R.M. Pathak and Y.P. Gupta, "Genetic-Algorithm Based Reliability Optimization for Computer Network Expansion", IEEE Transaction on Reliability, Vol. 44, No.1 pp 63–72,1995.
2. B. Dengiz, F. Altiparmak and A.E. Smith, "Efficient Optimization of All-Terminal Reliable Networks, Using an Evolutionary Approach", IEEE Transaction on Reliability, Vol. 46, No. 1 pp 18–26 1997.
3. Juan Alberto Rodriguez-Velazquez et al, "On Reliability indices of Communication Networks," Computers and Mathematics with Applications 58, pp 1433–1440, 2009.
4. A. Andrey Dobrtnin "Formula for calculating the Wiener Index of Catacondensed Benzenoid Topologys" J. Chem. Comput. Sci. 38, pp. 811–814, (1998).
5. J. Baskar Babujee and R. Jagadesh, "On tensor Product of Topologies", Acta Ciencia Indica, Vol 32 M, No. 3, pp. 1319–1321, (2006).
6. J. Baskar Babujee and A. Joshi, "Wiener Number Sequence for Sequence of Planar Topologys" International Journal of Applied Mathematical Analysis and Applications, Vol 4, No. 1, pp. 25–32 2009.
7. J.A. Bondy and U.S.R. Murty, "Topology Theory with Application", Macmillan Press Ltd 1976.
8. A.A. Dobrynin, R.C. Entringer and Gutman, I., "Wiener index of Trees: Theory and Applications", Acta Applications Mathematicae 66, pp. 211–249, 2001.
9. A. Joshi and J. Baskar Babujee, "Wiener Index for Tensor Product of Topologys", Proceedings of the Indian Conference on Intelligent Systems, Allied Publishers, pp. 113–114, 2007.
10. A. Joshi "Reliability Measurement and Enhancement Using Wiener Index of Communication Networks" IEEE Xplore digital library, ISBN No: 978-1-4673-51416, DOI:10.1109/INCOSET. 2012. 6513934, (2012), pp 365–368, 2012.
11. A.R. Majeed, J. A Hussein, "Weighted Network Reliability and Modeling," IEEE 5th IMCSSD, 2008.
12. Mou Dasgupta and G.P. Biswas, "Reliability Measurement and Enhancement of the Communication Networks" International Journal of Computer Applications vol 1-no 9, PP 18–25, 2010.

13. K.R. Parthasarathy, "Basic Topology Theory", Tata McGraw-Hill Publishing Company limited, New Delhi, 1994.
14. M. Randic, "On Characterization of Molecular Branching", J. Am. Chem. Soc. 97, 6609,1975.
15. N. Trinajstic, "Chemical Topology Theory", CRC Press, Boca Raton, 2nd revised Ed,. 1983.
16. H. Wiener, "Structural determination of paraffin boiling points", J. Amer. Chem. Soc. 69, pp. 17–20 1947.

Network Performance Analysis of Startup Buffering for Live Streaming in P2P VOD Systems for Mesh-Based Topology

Nemi Chand Barwar and Bhadada Rajesh

Abstract This paper explores mesh-based clustering for different start video streaming in P2P systems and estimates the performance of noncluster and clustered models. These models are based on mesh-based topology of P2P streaming consisting of peer join/leave. A new approach by way of "clustering" peers is proposed to tackle P2P VOD streaming. The proposed models were simulated and verified using OMNET++ V.4. A clustered model for video streaming is proposed and simulated to consider the performance of network under startup buffering for frame loss, startup delay, and end-to-end delay parameters. The results obtained from simulations are compared for both noncluster versus cluster models. The results show the impact of startup buffering on both models is also bounded due to time limits of release buffer and playing buffer under the proposed models, which causes reduction in wait time to view video improving the overall VOD system performance. The proposed model is also able to provide missing parts (of video) to late viewers, which gives the facilities of both live and stored streaming from user's point of view, therefore it serves to be functionally hybrid and is most useful.

Keywords Peer-to-peer (P2P) · Video streaming · Video on demand (VOD)

N.C. Barwar (✉)
Faculty of Engineering, Department of Computer Science & Engineering,
M.B.M Engineering College, J N V University, Jodhpur, India
e-mail: ncbarwar@yahoo.com; ncbarwar@jnvu.edu.in

Bhadada Rajesh
Faculty of Engineering, Department of Electronics & Communication Engineering,
M.B.M Engineering College, J N V University, Jodhpur, India
e-mail: rajesh_bhadada@rediffmail.com

1 Introduction

P2P technology has been able to establish itself as an effective and flexible method for video streaming in dynamic and heterogeneous environments. Over the past few years, P2P-based technique had begun succeeding over the Internet in various applications such as file sharing, file download, and streaming transmission. Successful examples include PPLive and Coolstreaming [1]. By continued research a "cache-and-relay" method has come up, where client caches the recently played data and then keeps that copy of content to serve others. Some researchers have even used this concept for the P2P streaming system [2–4]. So recently, many systems use P2P technology and VOD services in a bundle to reduce the cost of video transmission effectively [4, 5]. Unfortunately, supporting interactive operations such as pause, fast forward, and rewind functions [4, 6, 7] are still at a bottleneck in development.

In a P2P network, data is shared among the group of peers and required data is provided by a peer by means of searching through submitting queries to neighbors or to directory server. When the requisite data are traced, the peer downloads the data directly from the other peer's computer. As in this arrangement data are selectively replicated among peers, which allow sharing of data by a large community at low cost, as dedicated servers are not needed. Each peer shares its local storage and capacity to cater multimedia distribution service to other peers. The multicast approach has been attempted to minimize server loading and network bandwidth. In a multicast stream, a group of the clients can receive the stream at a time. Compared with unicast approach, it can significantly optimize the server load. However, popularity aspect of video shatters the prospectus of multicasting.

P2P VOD is an emerging service for users to select a video or different part of an available video. P2P VOD can bring substantial reduction in server loading [8, 9]. Many P2P VOD versions have been tried such as PPLive, Joost, GridCast, PPStream, UUSee, etc., however, their architectures have broadly been grouped as either tree-based or mesh-based structures [9, 10].

2 Background

Architecture of P2P VOD streaming is based on tree or mesh structure.

(A) **Tree-Based Systems**: In this type, peers build up a tree structure at application layer and original server is treated as root. Other clients in the network have only one parent and may have several children. Streaming server at root encodes new packets and sends them to all its children, while similarly each node gets packets from the parent and forwards them to their children. This mechanism is a push approach and here no requests and control packets are sent to the parent nodes. Delay is based on the depth of the tree where number of hops of a packet has to travel. Tree-based system lacks in strength as it takes

more time for construction of optimal tree and needs more knowledge of available network resources including bandwidth. Further dynamic changes in network makes task of maintenance complex and hence the reorganization process needs more time. Moreover, when clients offer small buffers or nodes move from the top of the tree (low playback delay) to the bottom (long playback delay), this may generate interruption at video playbacks.

(B) **Mesh-Based Systems**: Under heterogeneous networks, clients keep joining and leaving the network frequently, needing the tree to be repaired or reorganized very often. Under such condition, mesh-based systems can better handle the networks and the unstable user behavior. Here a client keeps searching for possible sources in the network to download needed video segments. To choose right peers, a client has to know where the needed segments are stored. It can be done by the exchange of information between the clients through buffer maps (bit vectors which indicate the availability of segments). With help of bit vectors, one of the sources is selected and client can request the peer to send this segment. To organize these exchanges control traffic is generated as overhead. Generally the clients have option to download the same segment out of many peers, making mesh-based systems more resistant against node failures.

Ketmaneechairat et al. [11] proposed a nonclustered mesh-based model for broadcasting in P2P system consisting of five stages. We have extended this model for video streaming by way of nonclustering and clustered model for analyzing performance of P2P systems.

3 Proposed Clustering

The strategy of proposed cluster-based system architecture and design is described as below. Whenever a new peer joins and downloads chunks from the peers under this model for different start video, a global tracker decides which cluster and node will be best to join.

3.1 System Architecture

In this approach, peers (nodes) shall be grouped according to their joining time or the chunks available with them. It will have a server, a global tracker (GT), super node (SN) or local tracker (LT), backup-node (BN), and normal nodes (NN, seed, and leech). The server shall be a special node that provides all chunks of a live video for streaming. The global tracker maintains the list of all super nodes and is known to all nodes. A super node maintains a list of all nodes in a cluster and functions as a local tracker. All such super nodes are connected with the global
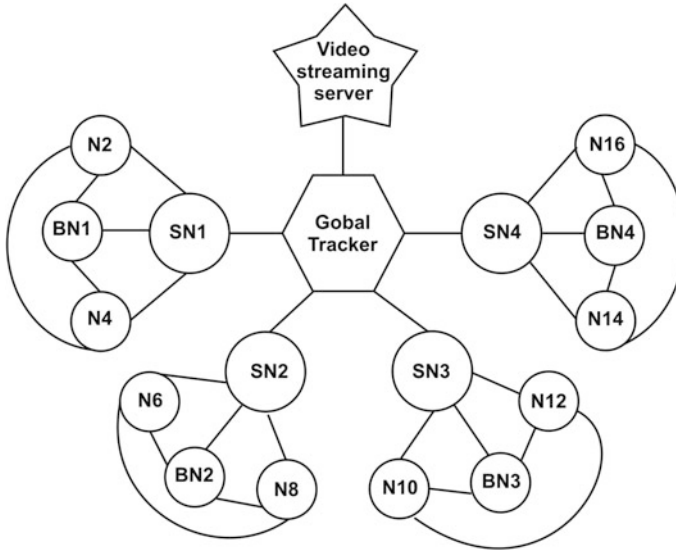
Fig. 1 Proposed clustering of nodes for video streaming

tracker to perform function of synchronizing the lists of all nodes in the cluster. The SN, NN, and BN will be used for performance of downloading (leech) and uploading (seed) of chunks (Fig. 1).

4 Simulation Setup

For the proposed clustered model, a setup was prepared to conduct simulations under clustered and nonclustered models under various parameters using discrete event network simulator, OMNET++ version 4.1. 'Oversim' is an open source frame work for implementation of P2P network in OMNET++, containing models for unstructured network (mesh network). A program in OMNET++ configured mesh-based P2P media streaming as nonclustered and clustered model for this work.

4.1 Simulation Configuration

The network sizes (varying between 25 and 500 nodes), peer churn, and number of neighbors have been considered for simulations. For simulating valid video stream 'Star Wars IV' trace file was applied. Under different clustered and nonclustered configurations, simulations were performed 5 times and average of all 5 runs was taken as the peer's output for each scenario.

The simulation of physical topology is generated using Georgia Tech Internet Topology Model (GT-ITM) tools for OMNET++ V.4 with 28 AS(backbone routers) and 28 access routers per AS in top-down mode.

A peer selects a router randomly and connects to it by a random physical link. Each peer selects a number of neighbors as per the configuration of experiment. Neighbors exchange buffer maps; each window of interest includes segmentsof 1 s each. Video file is divided into chunks. In this simulation it has been assumed that size of each chunk is equal to one second of player length.

4.2 Measured Parameters

To investigate the performance of the proposed model of P2P video streaming for mesh type noncluster model and clustered model have been carried out via simulations. Following network performance parameters were considered and analyzed.

1. *Frame loss ratio*: The ratio of the dropped frames with respect to total video frames transmission (% of frame loss in total).
2. *Startup delay*: Time interval between a peer decides to connect to a video session in mesh and begins its playback.
3. *End-to-end delay*: The time elapsed between the video frame originated from the source server and that video frame reaches the client peer.

5 Performance Analysis

To measure and assess the impact of startup buffering over frame loss ratio, startup delay, and end-to-end delay, the simulator was run under mesh-based noncluster and cluster models by varying the number of peers with the observed parameters.

5.1 Network Performance of Noncluster Model (Impact of Startup Buffering)

Figure 2 shows the impact of startup buffering over frame loss for the video streaming. Frame loss ratio is minimized when startup buffering time is kept higher despite having increased network size.

Figure 3 depicts the impact of startup buffering over startup delay for the video streaming. It is seen from the figure that startup delay is minimum for lower startup buffering period which yields better QoS for VOD to the users. However, almost negligible variation in startup delay due to growth in network size was observed.

Fig. 2 Impact of startup buffering over frame loss (noncluster)
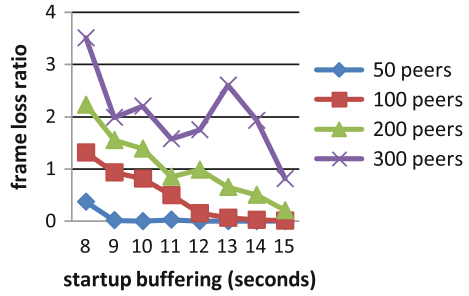
Fig. 3 Impact of startup buffering over startup delay (noncluster)
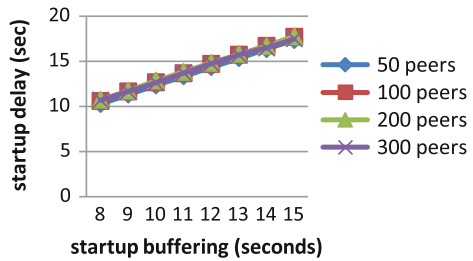
Fig. 4 Impact of startup buffering over end to end delay (noncluster)
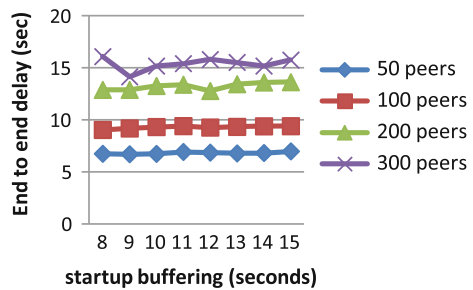
Figure 4 depicts the impact of startup buffering over end-to-end-delay for video streaming. It is observed from the figure that end-to-end delay is nearly constant by varying the startup buffering, hence it does not seem to affect the video streaming play out time at different clients. However, growth of network results in more end-to-end delay.

5.2 Network Performance of Cluster Model (Impact of Startup Buffering)

Figure 5 illustrates the impact of startup buffering over frame loss for the video streaming for the 300 peers with varying the size of clusters and also varying the

Fig. 5 Impact of startup buffering over frame loss (cluster model)
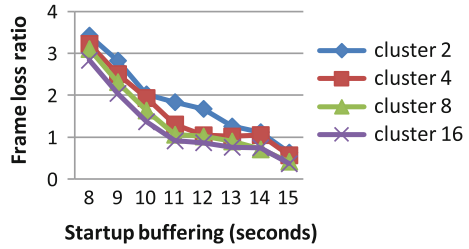
Fig. 6 Impact of startup buffering over startup delay (cluster model)
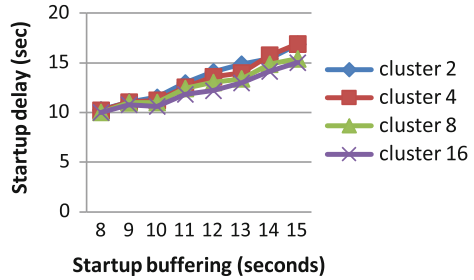
startup buffering time. Frame loss ratio significantly falls with increase in startup buffering time. However, it is almost similar to increasing the cluster sizes indicating not much dependence on the size of cluster.

Figure 6 illustrates the impact of startup buffering over startup delay for the video streaming the 300 peers with varying the size of clusters and also varying the startup buffering time. It is seen from the figure that increase in startup buffering brings in more startup delay. The startup delay can be minimized with the lower startup buffering to yield better service/delivery of video to the users. It is almost similar irrespective of the size of cluster.

Figure 7 depicts the impact of startup buffering over end-to-end delay for the video streaming the 300 peers with varying the size of clusters and also varying the startup buffering time. It is seen from the figure that end-to-end delay is nearly constant with varying the startup buffering but is somewhat less in the larger size of clusters. Thus, it does not affect much the video streaming play out time at different clients.

Fig. 7 Impact of startup buffering over end-to-end delay (cluster model)
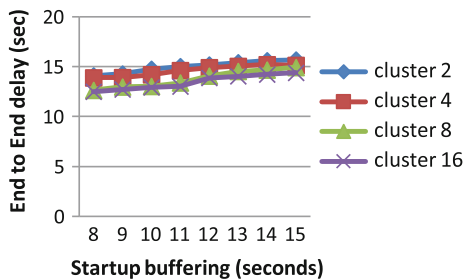
6   Conclusion

This paper presents a technique of clustering of peers in mesh-based live video streaming for P2P VOD system. On the basis of simulation results a comparative performance analysis of noncluster and cluster models is made to conclude the following: (i) that the impact of startup buffering on both models is also bounded due to time limits of release buffer and playing buffer under the proposed models, which causes reduction in wait time to view video improving overall VOD system performance. (ii) Proposed cluster model outperforms compared to the noncluster wherein the startup delay is minimized due to consideration of time bound limits of release buffer and playback buffers, which bounds the startup delay almost constant in proposed model.

Further it is also stated that the proposed model is able to provide missing parts (of video) to late viewers, which gives the facilities of both live and stored streaming from user's point of view, therefore it serves to be functionally hybrid and is most useful.

References

1. Xinyan Zang, Jiangchuan Liu, BO Li and Tak-Shing Peter Yem: Cool Streaming/ DONet: A Data-driven Overlay Network for Peer-to-Peer Live media Streaming, In: IEEE (2005).
2. Shahzad Ali, Anket Mathur and Hui Zhang,: Measurement of Commercial Peer-to-Peer Live Video Streaming, In: Proc. of ICST workshop on recent advances in Peer to Peer streaming, (2006).
3. A. Sentinelli, G. Marfia, M. Gerla, and L. Kleinrock: Will IPTV ride the peer-to-peer stream? In: Proc. IEEE Communications Magazine, In: vol. 45, no. 6, pp. 86–92, Jun. (2007).
4. Yu-Wei chen and Yu-Hao Huang: An Interactive Streaming Service over Peer-to-Peer Networks, In: Proc. of International Conference on Software and Computer Applications, (2011).
5. Xiaojun Hei, Chao Liang, Jian Liang, Yong Liu and Keith W. Ross: A Measurement Study of a Large-Scale P2P IPTV System, In: IEEE Transactions on Multimedia, Vol. 9, No. 8, Dec (2007).
6. Le Chang: Video-on-Demand streaming on Peer-to-peer Networks, In: Proc. International conference on Communications and Networking, IEEE (2007).
7. R Bhadada, N C Barwar & R Bhansali: Review of the Challenges to Remove Jitter and Packet Loss During Continuous Playback of Streamed Video Data in Video On Demand (VOD) System Receivers, In: International Journal of Computer Applications (IJCA) Special Issue on Electronics, Information and Communication Engineering, pp 3–7, Dec (2011).
8. Y. Huang, T. Z. J. Fu, John C. S. Lui, and C. Huang: Challenges, Design and Analysis of a Large-scale P2P-VoD System, In: Proc. ACM SIGCOMM'08 conference on Data communication, pp. 375–388, Aug. 17–22, (2008).
9. C. Wu and B. Li: Exploring Large-Scale Peer-to-Peer Live Streaming Topologies, In: Proc. ACM Transaction on Multimedia Computing, Communications, and Applications, vol. 4, no. 3, pp. 19–23, pp. 1–25, Aug. (2008).

10. Nazanin Magharei, Reza Rejaie and Yang Guo: Mesh or Tree: A Comparative Study of Live P2P Streaming Approaches, In: Proc. of the IEEE International Conference on Computer Communications (InfoCom07), pp 1424–1432, (2007).
11. H. Ketmaneechairat, P. Oothongsap, and A. Mingkhwan,: Smart Buffer Management for Different Start Video Broadcasting, In: Proc. of ACM ICIS'09, pp. 615–619, Nov. 24–26, (2009).

Cryptanalysis of Image Encryption Algorithms Based on Pixels Shuffling and Bits Shuffling

Pankaj Kumar Sharma, Aditya Kumar and Musheer Ahmad

Abstract In this paper, cryptanalysis of a certain class of shuffling-based image encryption algorithms is presented. The class includes the image encryption algorithms which either use a combination of pixels and bits level shuffling or bits level shuffling only. The proposed chosen plaintext attack on such algorithms can recover the plaintext image without having any clue or secret key. To demonstrate the practicability and effectiveness of the attack, it is applied to break one such algorithm recently suggested by Huang et al. in [Telecommun Syst 52(2), 563–571, 2013]. The Huang et al. image encryption uses a combination of pixel-level shuffling followed by bit-level shuffling. The simulation of the attack justifies the feasibility of the proposed cryptanalysis. The attack can be easily extended to break other such image encryption algorithms as well. Further, some security issues related to these algorithms are discussed and few preferments are propounded which can make the class of encryption algorithms more robust against the proposed attack.

Keywords Cryptanalysis · Image encryption · Security · Pixel shuffling · Bits shuffling

P.K. Sharma (✉) · Aditya Kumar · Musheer Ahmad
Faculty of Engineering and Technology, Department of Computer Engineering,
Jamia Millia Islamia, New Delhi 110025, India
e-mail: sharmapankaj1992@gmail.com

Aditya Kumar
e-mail: kwikadi@live.com

Musheer Ahmad
e-mail: musheer.cse@gmail.com

1   Introduction

With the massive growth of Internet in the last few decades, the need of secure communication has risen especially because of the acclivity of eCommerce and online payments. Though many sophisticated algorithms and standards have been developed for text encryption, image encryption is still an active area of research. The reason that well-developed text encryption algorithms like AES, Blowfish, CAST5, RC4, Triple-DES, IDEA incur high computation powers and are found inefficient when applied for image encryption [1]. The reason being that the images have bulk data capacity, high redundancy and high level of correlation amongst the neighbouring pixels. To void the impact of high redundancy and pixels correlation, image encryption algorithms [2–6] are designed which explore the sources that can effectively generate cryptographically strong pseudorandom sequences. Statistically, the usage of chaotic maps in encryption algorithms is extensively utilised to realise secure image-based communication. Chaotic maps are ideal for generating pseudorandom sequences. These sequences could be used to carry out the shuffling of image pixels in order to nullify the correlation existing among the neighbouring pixel of any plain image. These chaotic random sequences are also processed to perform the necessary confusions and diffusions of the plain image to develop a strong image cryptosystem [7, 8]. The randomness property inherent in chaotic maps makes them suitable for image encryption techniques.

Chaotic systems have many important properties such as the sensitive dependence on initial conditions and system parameters, pseudorandom property, non-periodicity and topological transitivity, etc. These properties meet some requirements such as diffusion and mixing in the sense of cryptography [8]. Therefore, chaotic cryptosystems have proved to be more reliable and practical for security applications. Chaos-based cryptographic algorithm was first proposed around 1989 in [9]. In recent past, many image encryption algorithms were proposed in the literature to meet the demand of fast, secure and reliable image encryption so that sensitive image data can be transferred securely and privately. A number of classical encryption algorithms were proposed based on Josephus traversing, Arnold transform and chaos map. The studies have been done to design improved chaos-based image encryption schemes. Mainly, there are two stages in chaos-based image cryptosystems and they follow the confusion–diffusion architecture. During the confusion stage, pixel permutation is done where the pixels are securely scrambled over the entire image without disturbing the gray value of pixels and thus the image becomes unrecognisable. Pixel permutation is carried out by extracting the permutation sequences from the chaotic system [6]. To improve the security, the second stage of the encryption process aims at changing the value of each pixel to make it robust against attacks. During the diffusion stage, the pixel values are modified sequentially with secure keystream generated from chaotic systems. The chaotic behaviour is controlled by the initial conditions and control parameters which are derived from the initial secret key. The whole confusion–diffusion process repeats for a number of times to achieve a satisfactory level of

security. However, many such encryption schemes suffer from serious security weaknesses which make them susceptible to cryptographic attacks. As a result, few image encryption algorithms have been successfully broken by cryptanalysts under various attacks [10–17].

In this paper, the security of this class of image encryption algorithms follows chaos-based pixel shuffling or bits shuffling or a combination of both pixel and bit shuffling to encrypt the images. Here, we provide detailed analysis of such an image encryption algorithm recently proposed by Huang et al. in [5]. The algorithm makes use of Chua chaotic system for the generation of random sequences.

The rest of the paper is further fragmented into the following sections. Section 2 provides a briefing of Huang et al. image encryption algorithm. Section 3 provides a chosen plaintext attack (CPA) on Huang et al. algorithm with some suggested measures to improve the resistance of image encryption algorithms. Finally, the concluding remark of the work is made in Sect. 4.

2 Review of Huang et al. Algorithm

In this section, we discuss the Huang et al. scheme along with its strong and weak points, which make it vulnerable to our chosen plaintext attack. The Huang et al. scheme makes use of Chua chaotic system [7] for generation of three sequences X, Y and Z. The size of sequences X and Y are said to be equal to that of the image to be encrypted; the paper does not suggest anything about sequence Z. It is assumed that the size of Z sequence should be $k \times M \times N$ where k is the number of bits representing each pixel and $M \times N$ is the number of pixels in the image.

The processing steps of the Huang et al. encryption algorithm are as follows:

Step 1   Initial values of $x0$, $y0$ and $z0$, along with other parameters (α, β, γ, $m0$, $m1$) and the grey-level matrix $A_{M \times N}$ for the plain image are taken.

Step 2   Chaotic sequences X, Y and Z are obtained by iterating the Chua map.

Step 3   Column indexing and shuffling are performed using the matrix $X1$, obtained from the chaotic sequence X for all columns of matrix $A_{M \times N}$ to obtain matrix $A_{e1(M \times N)}$.

Step 4   After column shuffling, a similar row indexing and shuffling is performed on $A_{e1(M \times N)}$ using chaotic sequence Y to obtain another matrix $A_{e2(M \times N)}$.

Step 5   Transform matrix $A_{e2(M \times N)}$ to $A_{e2(1 \times MN)}$.

Step 6   Perform bits indexing and shuffling on $A_{e2(1 \times MN)}$ (binary) using Z chaotic sequence to obtain vector $A_{e21(1 \times MN)}$.

Step 7   After dimensional transformation, encrypted image matrix $A_{e(M \times N)}$ is obtained from $A_{e21(1 \times MN)}$.

The decryption algorithm is simply the reverse of the encryption process. It is to be noted that the authors have not provided much information about the binary and bit conversions and bits shuffling. We have assumed that the conversion is carried out by converting the vector $A_{e2(1 \times MN)}$ to $A_{e2(1 \times MN)}$ (binary) by converting each

pixel to its equivalent binary representation. The main advantages of this encryption technique are astronomic key space and almost no correlation amongst the adjoining pixels. Based upon the statistical analysis provided by the authors of the paper, the system also provides high pixel alteration, measured and justified against NPCR and UACI parameters [3, 6]. Another noteworthy key feature of the system is immense cohesion between the initial values and the sequence used for doing the shuffling making the scheme robust against statistical attacks.

3 Proposed Cryptanalysis

Though the image cryptosystem suggested by Huang et al. possesses multiple advantages that a modern cryptographic scheme must have. However, a careful analysis of the scheme shows that the algorithm is vulnerable to our proposed chosen plaintext attack. An inherent flaw that makes it vulnerable is using the same initial values for (re)generating the sequence during rows, columns and bits shufflings every time a different image is encrypted. Though, due to the enormous size of the key space, it is almost impossible to guess the sequence. But, a careful analysis of the algorithm reveals that the bits reshuffling will always be the same and the final position of any particular bit will always be the same. For example, no matter what, the bit (whether 0 or 1) at the first position in vector $A_{e2(1 \times MN)}$ (binary) will always appear at a certain position as the keys are not dependent upon the initial image. Using this inherent vulnerability of the system, we proposed a chosen plaintext (CPA) attack against the Huang et al. system. According to Auguste Kerckhoffs axioms: "*A cryptographic system should be secure even if everything about the system, except the key, is public knowledge*" [13, 14]. Hence, everything about the encryption algorithm including its implementation is public except the secret key which is private. In other words, the attacker has a temporary access to the encryption or decryption machines. The goal of an attacker is to recover the secret key or the plain image without having any prior knowledge of the key used. In CPA attack, the attacker can have temporary access to an encryption machine and choose some specially designed plaintext images to generate corresponding ciphertext images. The CPA attack is indeed simple in nature and does not require any knowledge of the initial value of the secret key used. Due to the independent nature of the encryption system under scrutiny to pending plain image information, the mapping will remain the same regardless of image used. The attack basically finds the mapping of bit values at different positions in the original image and the encrypted image. The algorithmic steps of the proposed CPA attack are as follow:

1. Let $Z_{M \times N}$ be the encrypted image.
2. Set $K = [128, 64, 32, 16, 8, 4, 2, 1]$. The common property amongst all the elements of vector K is that all of them consist of a single bit as 1 and the rest of them are 0. For example, 128_{10} in binary is 10000000_2.
3. Let B be a matrix having all grey values zero of size $M \times N$.

4. Let A be a empty binary matrix of size $8 \times M \times N$.
5. For each value of i in range $(1, M)$, repeat steps 6–15.
6. For each value of j in range $(1, N)$, repeat steps 7–15.
7. For each $k \in K$, repeat steps 8–15.
8. $C = B$.
9. $C[i][j] = k$.
10. $C_Z = Huang\_Encrypt(C)$; here Huang et al. algorithm is used as a black-box.
12. $[m, n]$ = non-zero pixel position in C_Z;, only 1 bit in C_Z is non-zero.
13. $pz = C_Z[x][y]$; i.e. value of the exclusive non-zero pixel value of encrypted image.
14. Let z be the position of pz in K; Since only a single bit was 1, $pz \in K$.
15. $A[i][j][k] = Z[x][y][z]$;
16. Convert image A (binary) to grey-image $A_{M \times N}$.
17. Declare $A_{M \times N}$ as recovered image.

In this attack, the custom-built gray images having same dimensions as the original image are used. The main property of these custom-built images is that each of these specially chosen images held has only one bit of any one pixel of the images is 1 and all the remaining bits are 0. The image encryption is, eventually, a bits shuffling algorithm, the encrypted image would also consist of one bit value as 1 and rest as 0. For example, if we initially keep the first bit of the first pixel of the first row as 1 and rest as 0. Let us call this location as I(0, 0, 0), where first index represents the row (r), second index represents the column (c) and the third index represents the bit position (b) in the pixel at (r, c) having the value 1. In an encrypted image, we will have only a single bit position with value 1, let the position of that bit be I(m, n, o). So, we can say that the plain image bit position (0, 0, 0) is mapped to the final encrypted image bit position (m, n, o). In a similar fashion, we can apply the proposed CPA attack to reveal the secret mappings or the final shuffling positions of the plain image bits, going left to right and top to bottom to final I(M, N, 8) without the secret key used in Huang et al. cryptosystem during encryption. Using this mapping, we can build a reverse map that can be used to recover the original image from the received encrypted image generated from Huang et al. scheme. The reverse map helps us to know the final shuffled position of the bits of plain images. By unveiling the final shuffled position, the bits are de-shuffled back to their original positions to recover the original image. As an example, a plain image of size 3×3 is shown in Fig. 1 and its corresponding encrypted image obtained with Huang et al. scheme is depicted in Fig. 2. The respective mapping and reverse mapping extracted using the proposed CPA attack

Fig. 1 Plain image of size 3×3 in decimal and binary

$$\begin{bmatrix} 102 & 57 & 43 \\ 69 & 29 & 93 \\ 21 & 167 & 82 \end{bmatrix} \quad \begin{bmatrix} 01100110 & 00111001 & 00101011 \\ 01000101 & 00011101 & 01011101 \\ 00010101 & 10100111 & 01010010 \end{bmatrix}$$

Fig. 2 Encrypted image using Huang et al. scheme of size 3×3 in decimal and binary

$$\begin{bmatrix} 128 & 247 & 227 \\ 238 & 135 & 3 \\ 227 & 64 & 232 \end{bmatrix} \quad \begin{bmatrix} 10000000 & 11110111 & 11100011 \\ 11101110 & 10000111 & 00000011 \\ 11100011 & 01000000 & 11101000 \end{bmatrix}$$

Table 1 Mapping of bit positions in the original image (I) and the encrypted image (I') using proposed attack

I(m,n,k)	I'(m,n,k)	I(m,n,k)	I'(m,n,k)	I(m,n,k)	I'(m,n,k)	I(m,n,k)	I'(m,n,k)
(1, 1, 1)	(2, 3, 3)	(1, 3, 3)	(1, 3, 1)	(2, 2, 5)	(2, 1, 2)	(3, 1, 7)	(3, 1, 5)
(1, 1, 2)	(2, 3, 7)	(1, 3, 4)	(3, 2, 6)	(2, 2, 6)	(3, 1, 3)	(3, 1, 8)	(2, 2, 1)
(1, 1, 3)	(1, 2, 3)	(1, 3, 5)	(1, 3, 3)	(2, 2, 7)	(2, 2, 2)	(3, 2, 1)	(1, 1, 1)
(1, 1, 4)	(1, 1, 2)	(1, 3, 6)	(3, 2, 5)	(2, 2, 8)	(1, 2, 4)	(3, 2, 2)	(3, 3, 4)
(1, 1, 5)	(1, 1, 5)	(1, 3, 7)	(1, 3, 2)	(2, 3, 1)	(2, 1, 4)	(3, 2, 3)	(2, 2, 8)
(1, 1, 6)	(3, 3, 3)	(1, 3, 8)	(1, 2, 8)	(2, 3, 2)	(3, 1, 1)	(3, 2, 4)	(1, 2, 5)
(1, 1, 7)	(2, 1, 1)	(2, 1, 1)	(3, 3, 7)	(2, 3, 3)	(2, 2, 4)	(3, 2, 5)	(3, 3, 6)
(1, 1, 8)	(3, 1, 4)	(2, 1, 2)	(2, 1, 6)	(2, 3, 4)	(3, 1, 8)	(3, 2, 6)	(2, 1, 5)
(1, 2, 1)	(1, 1, 4)	(2, 1, 3)	(3, 2, 8)	(2, 3, 5)	(1, 3, 7)	(3, 2, 7)	(1, 2, 6)
(1, 2, 2)	(1, 1, 7)	(2, 1, 4)	(1, 3, 4)	(2, 3, 6)	(3, 2, 2)	(3, 2, 8)	(3, 3, 5)
(1, 2, 3)	(3, 3, 2)	(2, 1, 5)	(3, 2, 4)	(2, 3, 7)	(2, 1, 8)	(3, 3, 1)	(2, 3, 2)
(1, 2, 4)	(2, 1, 3)	(2, 1, 6)	(2, 2, 6)	(2, 3, 8)	(1, 2, 7)	(3, 3, 2)	(1, 2, 1)
(1, 2, 5)	(3, 1, 2)	(2, 1, 7)	(3, 1, 6)	(3, 1, 1)	(2, 3, 6)	(3, 3, 3)	(3, 2, 1)
(1, 2, 6)	(2, 2, 3)	(2, 1, 8)	(2, 3, 8)	(3, 1, 2)	(2, 3, 4)	(3, 3, 4)	(2, 1, 7)
(1, 2, 7)	(2, 3, 1)	(2, 2, 1)	(1, 1, 3)	(3, 1, 3)	(2, 3, 5)	(3, 3, 5)	(3, 2, 7)
(1, 2, 8)	(1, 3, 8)	(2, 2, 2)	(1, 1, 6)	(3, 1, 4)	(1, 2, 2)	(3, 3, 6)	(2, 2, 5)
(1, 3, 1)	(1, 3, 5)	(2, 2, 3)	(1, 1, 8)	(3, 1, 5)	(3, 3, 8)	(3, 3, 7)	(3, 1, 7)
(1, 3, 2)	(3, 2, 3)	(2, 2, 4)	(3, 3, 1)	(3, 1, 6)	(2, 2, 7)	(3, 3, 8)	(1, 3, 6)

is listed in Tables 1 and 2, for the 3×3 dimensional image, where each pixel is represented with 8 bits. Using the generated reverse mapping given in Table 2, the original image, shown in Fig. 3, is recovered.

The presented CPA attack requires $8 \times M \times N$ plain images to completely recover the original image. The running time of the attack will be approximately equal of the order of $O(MN)$ for $M \times N$ gray images. The attack takes not more than 10 min to recover a 256×256 image. Assuming, if each round of the Huang et al. encryption takes 1 ms to complete, the brute force attack is likely to take around 8.1177×10^{21} years to exhaust all possible keys. The running time of the proposed cryptanalysis is insignificant over the time taken by brute force attack.

As discussed earlier, the class of algorithms that we aimed to cryptanalyse posseses multiple advantage. The major reason that makes this class of encryption schemes vulnerable to our CPA is using the same keys for any distinct plain image and the initial keys not being dependent upon the pending plain image. As a result,

Table 2 Reverse mapping of bit positions in the original image (I) and the encrypted image (I')
using the proposed attack

I(m,n,k)	I'(m,n, k)	I(m,n,k)	I'(m,n, k)	I(m,n,k)	I'(m,n, k)	I(m,n,k)	I'(m,n, k)
(1, 1, 1)	(3, 2, 1)	(1, 3, 3)	(1, 3, 5)	(2, 2, 5)	(3, 3, 6)	(3, 1, 7)	(3, 3, 7)
(1, 1, 2)	(1, 1, 4)	(1, 3, 4)	(2, 1, 4)	(2, 2, 6)	(2, 1, 6)	(3, 1, 8)	(2, 3, 4)
(1, 1, 3)	(2, 2, 1)	(1, 3, 5)	(1, 3, 1)	(2, 2, 7)	(3, 1, 6)	(3, 2, 1)	(3, 3, 3)
(1, 1, 4)	(1, 2, 1)	(1, 3, 6)	(3, 3, 8)	(2, 2, 8)	(3, 2, 3)	(3, 2, 2)	(2, 3, 6)
(1, 1, 5)	(1, 1, 5)	(1, 3, 7)	(2, 3, 5)	(2, 3, 1)	(1, 2, 7)	(3, 2, 3)	(1, 3, 2)
(1, 1, 6)	(2, 2, 2)	(1, 3, 8)	(1, 2, 8)	(2, 3, 2)	(3, 3, 1)	(3, 2, 4)	(2, 1, 5)
(1, 1, 7)	(1, 2, 2)	(2, 1, 1)	(1, 1, 7)	(2, 3, 3)	(1, 1, 1)	(3, 2, 5)	(1, 3, 6)
(1, 1, 8)	(2, 2, 3)	(2, 1, 2)	(2, 2, 5)	(2, 3, 4)	(3, 1, 2)	(3, 2, 6)	(1, 3, 4)
(1, 2, 1)	(3, 3, 2)	(2, 1, 3)	(1, 2, 4)	(2, 3, 5)	(3, 1, 3)	(3, 2, 7)	(3, 3, 5)
(1, 2, 2)	(3, 1, 4)	(2, 1, 4)	(2, 3, 1)	(2, 3, 6)	(3, 1, 1)	(3, 2, 8)	(2, 1, 3)
(1, 2, 3)	(1, 1, 3)	(2, 1, 5)	(3, 2, 6)	(2, 3, 7)	(1, 1, 2)	(3, 3, 1)	(2, 2, 4)
(1, 2, 4)	(2, 2, 8)	(2, 1, 6)	(2, 1, 2)	(2, 3, 8)	(2, 1, 8)	(3, 3, 2)	(1, 2, 3)
(1, 2, 5)	(3, 2, 4)	(2, 1, 7)	(3, 3, 4)	(3, 1, 1)	(2, 3, 2)	(3, 3, 3)	(1, 1, 6)
(1, 2, 6)	(3, 2, 7)	(2, 1, 8)	(2, 3, 7)	(3, 1, 2)	(1, 2, 5)	(3, 3, 4)	(3, 2, 2)
(1, 2, 7)	(2, 3, 8)	(2, 2, 1)	(3, 1, 8)	(3, 1, 3)	(2, 2, 6)	(3, 3, 5)	(3, 2, 8)
(1, 2, 8)	(1, 3, 8)	(2, 2, 2)	(2, 2, 7)	(3, 1, 4)	(1, 1, 8)	(3, 3, 6)	(3, 2, 5)
(1, 3, 1)	(1, 3, 3)	(2, 2, 3)	(1, 2, 6)	(3, 1, 5)	(3, 1, 7)	(3, 3, 7)	(2, 1, 1)
(1, 3, 2)	(1, 3, 7)	(2, 2, 4)	(2, 3, 3)	(3, 1, 6)	(2, 1, 7)	(3, 3, 8)	(3, 1, 5)

Fig. 3 Recovered image in decimal and binary using proposed cryptanalysis

$$\begin{bmatrix} 102 & 57 & 43 \\ 69 & 29 & 93 \\ 21 & 167 & 82 \end{bmatrix} \qquad \begin{bmatrix} 01100110 & 00111001 & 00101011 \\ 01000101 & 00011101 & 01011101 \\ 00010101 & 10100111 & 01010010 \end{bmatrix}$$

this class of image encryption algorithms has no plain-image sensitivity, i.e. a minute change in the plain image does not result a drastic avalanche in the encrypted content. These encryption schemes could be made to become more robust and resistant to the proposed and other types of attacks by using the following suggested measures:

1. Deploying a proper keys rotation policy which rotate the usage of keys after varying the interval of time.
2. Making initial conditions and keys dependent and sensitive to minor change in the plain image information. This could be done by taking the hash code of the plain image and deriving the initial keys.
3. Employing concealing of pixels information by using a XOR or some other operation before actually employing shuffling.

4   Conclusion

This paper analyses a particular class of image encryption which uses bit- or pixel-level shuffling for obtaining the required confusion in the encrypted images. Though, the shuffling is based upon the pseudo-random sequences generated from chaotic functions. The vulnerability among this class of encryption scheme is being explored in this paper and a chosen plaintext attack is given for such an image encryption proposed recently by Huang et al. The plain image is recovered successfully without any prior knowledge of the secret key. The outcomes of the theoretical and simulation analyses of the proposed cryptanalysis validate the feasibility of the attack. Hence, the image cryptosystem suggested by Huang et al. is completely insecure and unsuitable for encrypting sensitive digital images. Few suggestions are provided so as to make these image encryption schemes more resistant to proposed attack that can also inherit all their inherent security merits of cryptosystem.

References

1. Menezes, A.J., Oorschot, P.C.V., Vanstone S.A.: Handbook of Applied Cryptography, CRC Press (1997).
2. Lian, S., Sun, J., Wang, Z.: A block cipher based on a suitable use of the chaotic standard map. Chaos, Solitons & Fractals 26(1) 117–129, (2005).
3. Wong, K-W., Kwok, B.S., Law, W-S.: A fast image encryption scheme based on chaotic standard map. Physics Letters A 372(15) 2645–2652, (2008).
4. Ahmad, M., Ahmad, T.: Securing multimedia color imagery using multiple high dimensional chaos based hybrid keys. Int. J. Commun. Netw. Distrib. Syst. 12(1), 113–128 (2014).
5. Huang, C. K., Liao, C.W., Hsu, S.L., Jeng, Y.C.: Implementation of gray image encryption with pixel shuffling and gray-level encryption by single chaotic system. Telecommunication Systems 52(2) 563–571, (2013).
6. Zhang, Y., Xiao, D.: An image encryption scheme based on rotation matrix bit-level permutation and block diffusion. Commun Nonlinear Sci Numer Simulat 19(1) 74–82, (2014).
7. Yang, T., Wu, C.W., Chua, L.O.: Cryptography based on chaotic systems. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications. 44(5) 469–472, (1997).
8. Baptista, M. S.: Cryptography with chaos. Physics Letters A 240(1) 50–54, (1998).
9. Matthews R.: On the derivation of a chaotic encryption algorithm. Cryptologia 13 29–42, (1989).
10. Rhouma, R., Belghith, S.: Cryptanalysis of a new image encryption algorithm based on hyper-chaos. Physics Letters A 372(38), 5973–5978, (2008).
11. Rhouma, R., Belghith, S.: Cryptanalysis of a spatiotemporal chaotic cryptosystem. Chaos, Solitons & Fractals 41(4) 1718–1722, (2009).
12. Solak, E., Rhouma, R., Belghith, S.: Cryptanalysis of a multi-chaotic systems based image cryptosystem. Optics Communications 283(2), 232–236, (2010).
13. Özkaynak, F., Özer, A.B., Yavuz, S.: Cryptanalysis of a novel image encryption scheme based on improved hyperchaotic sequences. Optics Communications 285(2), 4946–4948, (2012).
14. Ahmad, M.: Cryptanalysis of chaos based secure satellite imagery cryptosystem. In: Contemporary Computing. Springer Berlin Heidelberg, 81–91, (2011).

15. Sharma, P.K., Ahmad, M., Khan, P.M.: Cryptanalysis of image encryption algorithm based on pixel shuffling and chaotic S-box transformation. In: Security in Computing and Communication. Springer-Verlag Berlin Heidelberg, CCIS 467, 173–181, (2014).
16. Ahmad, M., Ahmad, F.: Cryptanalysis of Image Encryption Based on Permutation-Substitution Using Chaotic Map and Latin Square Image Cipher. In: Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications. Springer International Publishing Switzerland, AISC 327, 481–488 (2015).
17. Ahmad, M., Khan, I.R., Alam, S.: Cryptanalysis of Image Encryption Algorithm Based on Fractional-Order Lorenz-Like Chaotic System. In: Emerging ICT for Bridging the Future. Springer International Publishing Switzerland, AISC 338, 381–388 (2015).

Circular-Shape Slotted Microstrip Antenna

Sumanpreet Kaur Sidhu and Jagtar Singh Sivia

Abstract In this paper, a novel single-feed circular-shape slotted microstrip antenna is proposed for C-band (4–8 GHz) applications. Antenna is designed on FR4 glass epoxy material with dielectric constant 4.4 and height 1.56 mm. Antenna is fed by a coaxial probe feed. Ansoft-based HFSS software is used for the simulation of proposed antenna. Vector Network Analyzer (VNA) is used for measuring the Return Loss (RL) and Voltage Standing Wave Ratio (VSWR) of the proposed antenna. The measured results such as gain, bandwidth, and return loss confirm the validity of this design and show a good agreement with simulated results. The proposed antenna shows a bandwidth of 191 MHz at 4.34 GHz, 190 MHz at 5.9 GHz and 106 MHz at 6.51 GHz frequency.

Keywords Circular-shape slot · Microstrip antenna · Gain · Bandwidth · Return loss

1 Introduction

In today's fast time, wireless technology has proved a blessing in the communication field. Antenna is one of the most essential elements in wireless communication. According to the IEEE standard definitions, the antenna or aerial is defined as "a means of radiating or receiving radio waves" [1]. In other words, antenna can be said as a metallic device meant for transmitting information. This information is in the form of electromagnetic waves.

The most commonly used antennas in the microwave region are microstrip patch antennas. These antennas are same as the capacitor in which there is a dielectric

S.K. Sidhu (✉) · J.S. Sivia
Yadavindra College of Engineering, Punjabi University, Guru Kashi Campus,
Talwandi Sabo, Bathinda, Punjab, India
e-mail: ersumansidhu@gmail.com

J.S. Sivia
e-mail: Jagtarsivian@yahoo.com

Fig. 1 Schematic
configuration of the proposed
compact circular-shaped slot
antenna
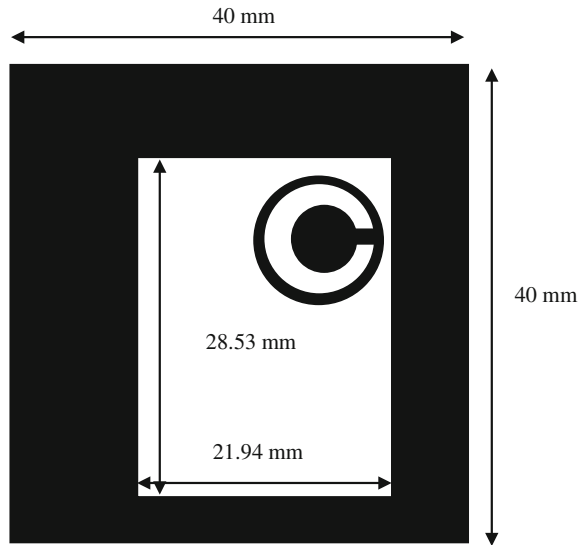
region enclosed between two plates. In microstrip antenna, we have a dielectric substrate in between ground and patch. In simple form it consists of a patch (generally circular or rectangular) on the top of the substrate [2]. These are the most widely used antennas due to their simplicity and ease of fabrication.

Due to their compact size and simplicity these antennas find a number of applications like in vehicle-based satellite link antennas [3], global positioning systems (GPS) [4], radar for missiles and telemetry [3], and mobile handheld radios or communication devices [4].

The word 'Fractal' has been derived from a Latin word called 'Fractus' which means broken. In fractal antenna, we divide the whole antenna into smaller copies of itself which means the antenna is split into parts which leads to increase in length of the antenna. In other words, the perimeter of antenna increases hence more space is available for transmitting or receiving information [5]. These antennas give two main advantages. First is we can design small scaled antennas [6] and second is it can operate at a number of frequencies [7].

In this paper a single-feed circular-shape slotted microstrip antenna is proposed. The geometry of the proposed design is illustrated in Fig. 1. The antenna prototype with an overall size of 40 mm × 40 mm × 1.56 mm achieves good impedance matching, gain, and radiation patterns over the entire operating bands. The antenna is practically fabricated and results are measured.

2 Design of Antenna

Design of proposed antenna starts with rectangular patch geometry whose dimensions are calculated using Eqs. 1–4. The size of the ground plane is 40.0 mm × 40.0 mm. The proposed antenna resonates at 3.2 GHz frequency.

Patch width of proposed antenna is calculated by using transmission line model equation as given in [8–10]

$$W = \frac{C_o}{2f_r} \sqrt{\frac{2}{\varepsilon_r + 1}} \tag{1}$$

$\varepsilon_r = 4.4$ (FR4_epoxy)
$f_r = 3.2$ GHz
$C_o = 3 \times 10^8$ m/s

With this equation the width becomes as, $W = 28.53$ mm.

Effective dielectric constant of material is calculated using equation as given in [8–10]:

$$\varepsilon_{\text{reff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \left[1 + 12 \frac{h}{W} \right]^{\frac{-1}{2}} \tag{2}$$

Effective length of patch is calculated using equations as given in [8–10]:

$$\frac{\Delta L}{h} = 0.412 \frac{(\varepsilon_{\text{reff}} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\varepsilon_{\text{reff}} - 0.258)\left(\frac{W}{h} + 0.8\right)} \tag{3}$$

$$L = \frac{C_o}{2f_r \sqrt{\varepsilon_{\text{reff}}}} - 2\Delta L \tag{4}$$

Its value becomes 21.94 mm.

The geometry for the proposed antenna with circular-shaped slot is as shown in Fig. 1.

Design steps of proposed antenna are:

Step 1 Design starts with base rectangular patch geometry with dimensions 21.94 mm × 28.53 mm is known as zeroth iteration geometry of proposed antenna.

Step 2 A circle with radius 5.4 mm is subtracted from the first quadrant of the base patch geometry to get the first iteration geometry as shown in Fig. 2a.

Step 3 A circle with radius of 4.8 mm, concentric with circle of first iteration, is united to get second iteration geometry as shown in Fig. 2b.

Step 4 A circle of radius 3 mm, concentric with circle of second iteration, is subtracted to get third iteration as shown in Fig. 2c.
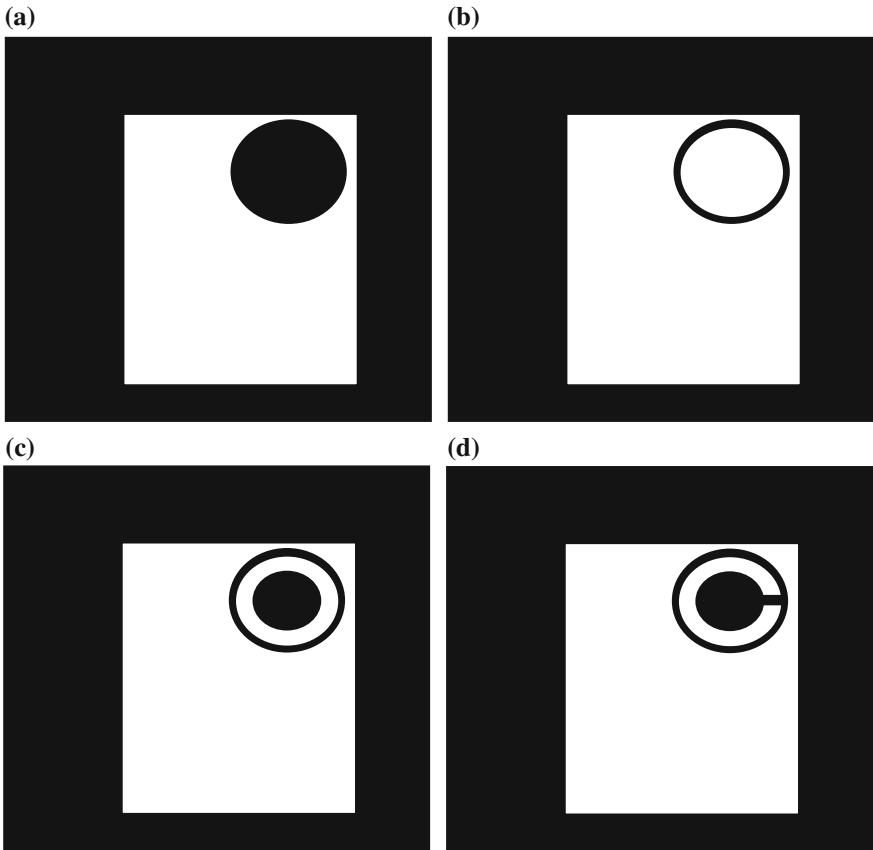
Fig. 2 Various iterations of proposed antenna. **a** First iteration. **b** Second iteration. **c** Third iteration. **d** Fourth iteration

Step 5   A rectangular slot with dimensions 3 mm × 1 mm is subtracted so as to connect the smallest circle to the biggest one to form fourth iteration as shown in Fig. 2d.

The fabricated geometry of proposed antenna is shown in Fig. 3.

3   Results

Simulated and measured return loss versus frequency plot for third iteration is shown in Fig. 3. Return losses are less than −10 dB at 4.34, 5.9, and 6.51 GHz. Thus the proposed antenna works at these three frequencies as shown in Fig. 4.
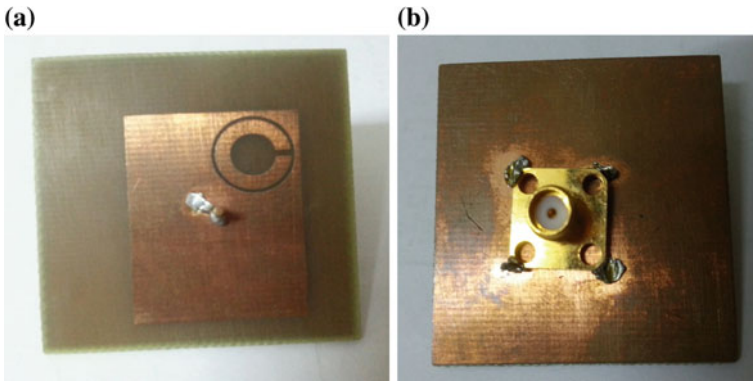
Fig. 3 Fabricated geometry of proposed antenna: **a** Front view. **b** Back view
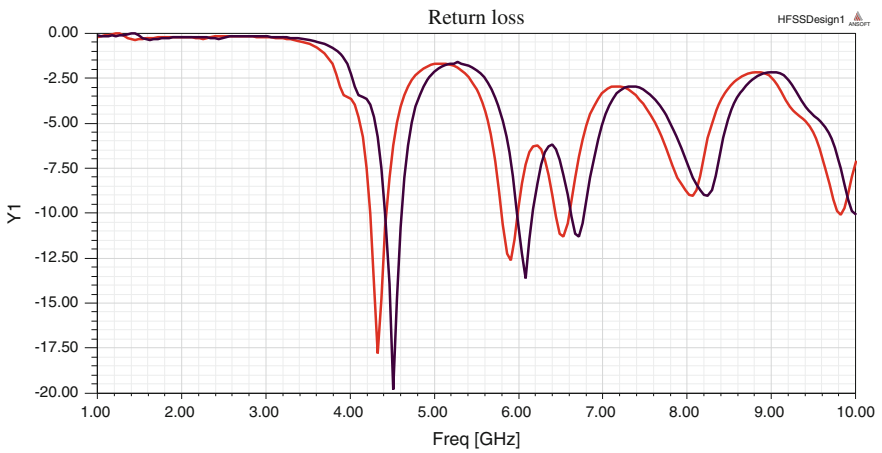
Fig. 4 Comparison of simulated and measured Return Losses of proposed antenna

Table 1 Performance parameters of proposed antenna

Frequency (GHz)	Simulated return loss (dB)	Bandwidth (MHz)	Gain (dB)	Measured return loss (dB)
4.34	−17.3915	191	4.6589	−18.5671
5.9	−12.6853	190	1.5234	−12.9631
6.51	−10.5740	106	6.5177	−10.6244

A Comparison of simulated and measured Return Losses of proposed antenna are shown are shown Fig. 4 which shows that simulted and measured results are good agreement with each other. Different performance parameters of antenna such as resonant frequencies, Return Losses (Simulated and Measured), Bandwidth and Gain
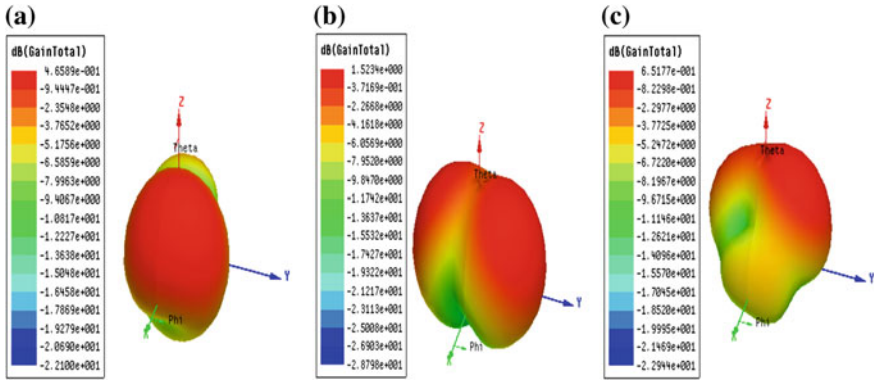
Fig. 5 3-D radiation pattern of circular-shape slotted microstrip antenna at frequency. **a** 4.34 GHz, **b** 5.9 GHz, **c** 6.51 GHz
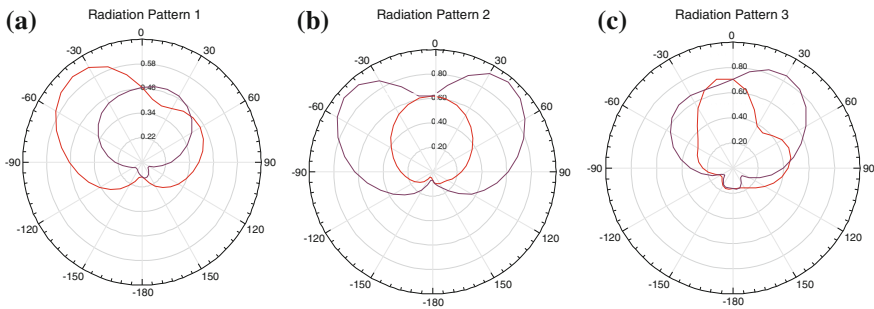
Fig. 6 2-D radiation pattern of circular-shape slotted microstrip antenna at frequency. **a** 4.34 GHz, **b** 5.9 GHz, **c** 6.51 GHz

are shown in Table 1. Three dimensional radiation pattern of proposed antenna at 4.34, 5.9, and 6.51 GHz are shown in Fig. 5a–c, respectively. Proposed antenna has a gain of 4.6589 dB at 4.34 GHz, 1.5234 dB at 5.9 GHz, and 6.5177 dB at 6.51 GHz.

2-D radiation pattern of proposed antenna at these frequencies are also shown in Fig. 6a–c, respectively.

4 Conclusion

In this paper a circular-shape slotted antenna is designed. The proposed design works at three frequency bands 4.24–4.43, 6.45–6.56, and 5.79–5.98 GHz. It is suitable for the C-Band applications such as Fixed Satellite Service (FSS) and Broadcast Satellite Service (BSS). Further when the results are compared with the fabricated design, they are similar and comparable.

References

1. Constantine, A. Balanis, "Antenna Theory Analysis and Design" 2nd Edition, Wiley India (p.) Ltd., 2007.
2. 6x9 Handbook/Antenna Engineering Handbook/Volakis/147574-5/ Chapter 7.
3. Mailloux, R.J., et al, "microstrip antenna technology", IEEE Trans. Antennas and Propagation, Vol. AP-29, January 1981, pp. 2–24.
4. James, R.J., et al, "Some recent developments in microstrip antenna design", IEEE Trans. Antennas and Propagation, Vol. AP-29, January 1981, pp. 124–128.
5. Puente C., Romeu J., POUSX R., Garcia, Benitez F., "Fractal Multiband Antenna Based on the Sierpinski Gasket", Electronics Letters, 1996, Vol. 32, No. 1, pp. 1–2.
6. Gianvinono J., Rahmat-Samii Y., "Fractal Antennas: A Novel Antenna Miniaturization Technique, and Applications", IEEE Antennas & Propagation.
7. Werner D.H., Werner P.L., "Frequency independent Feauhlrcs of Self-Similar Fractal Antennas", Radio Science, 1996, Vol. 31, No. 6, pp. 1331–1343.
8. Sivia J.S., Singh A. and Kamal T.S., "Design of Sierpinski Carpet Fractal Antenna using Artificial Neural Networks", International Journal of Computer Applications, pp. 5–10, Vol. 68, No. 8, April 2013.
9. Singh A.P. and Singh J., "On the design of rectangular microstrip antenna using artificial neural network", Institute of Engineer Journal—Electronics and Telecommunication, Vol. 90, pp. 20–25, Sept 2009.
10. Singh A.P., Sivia J.S. and Kamal T.S., "Estimatin of Feed position of rectangular Microstrip Antenna using Artificial Neural Antenna", Institute of Engineer Journal—Electronics and Telecommunication, Vol. 90, pp. 20–25, Sept 2010.

Energy Efficient Data Aggregation Technique Using Load Shifting Policy for Wireless Sensor Network

Samarth Anavatti, Sumedha Sirsikar and Manoj Chandak

Abstract Data redundancy is quite common in wireless sensor networks (WSNs) where nodes are deployed densely. The reason behind such deployment is to achieve reliability from communication failure. Communication failure happens when particular node transmitting data fails. In WSN there is no other way than keeping redundant nodes to solve communication failure problem. If redundant nodes are available then at the time of node failure the data of failed node can be recovered from its redundant nodes. Though we can achieve reliability through redundant nodes presence of redundant nodes will generate more number of redundant packets which will consume more energy of network. Because nodes which are densely deployed will sense the same information and send it to sink node and sink will waste its energy in processing redundant data and also redundancy will generate heavy traffic in network. Hence, there is a need to trade-off between energy conservation and reliability. To do this trade-off we need to find optimization point of redundancy in WSN. So that reliability and energy conservation both will be maintained. In this paper we have used clustering-based load shifting policy (LSP) to eliminate redundancy up to an adequate level to achieve optimization point of redundancy. Data aggregation eliminates redundancy from WSN. We are performing data aggregation at two levels and at the same time we are keeping redundant nodes up to 50 % to achieve reliability. In this paper, we have done comparison of traditional cluster-based data aggregation with our LSP-based data aggregation. Simulation result shows that LSPDA has lesser average energy consumption and longer lifetime than traditional cluster-based data aggregation method.

Samarth Anavatti (✉) · Sumedha Sirsikar
Maharashtra Institute of Technology, Pune, Maharashtra, India
e-mail: anavattisamarth@gmail.com

Sumedha Sirsikar
e-mail: sumedha.sirsikar@mitpune.edu.in

Manoj Chandak
Shri Ramdeobaba College of Engineering & Management,
Nagpur, Maharashtra, India
e-mail: chandakmb@gmail.com

1 Introduction

In wireless sensor networks nodes have limited capacity and lifetime. As the communication happens without any connection, chances of failure are more in WSN. Usually sensor nodes are deployed to sense the information about particular activity and if failure occurs that information will be lost. It is very hard to re-collect the lost data in wireless environment. The failure might be because of node failure or link failure, here as it is wireless we are considering node failure only because no fix links will be there in between two nodes.

The node failure problem can be solved by dense node deployment topology. In this topology nodes are deployed closer to each other, all redundant nodes will sense the same information about an activity and send it to sink. Now sink has multiple copies of same information. This will improve reliability because if one node fails, sink will receive same information from other nodes. Though this kind of node deployment improves reliability it will result in data redundancy at sink node and sink will waste most of its energy in processing redundant data. This redundancy will generate traffic in the network also.

Data aggregation is the process which eliminates the redundancy from network. According to data aggregation, all nodes will send data to one intermediate node rather than sink and this intermediate node performs aggregation and sends aggregated result to sink node. In aggregation process all redundant packets are merged together and one packet as a result is forwarded to sink node. Aggregation process saves energy of entire network and prolongs the lifetime of WSN [1].

The problem with data aggregation approach is one intermediate node is performing aggregation all the time and its energy gets consumed faster than other nodes which results in aggregator node dies quite early. This problem is also known as hotspot problem. To solve this problem there is a need to balance the traffic load.

Proposed system uses clustering based-LSP. LSP helps in keeping the redundancy up to adequate level in order to maintain reliability and clustering helps in balancing the traffic load. Energy efficiency is achieved by performing aggregation at two levels. The proposed system will balance trade-off between energy conservation and reliability.

The paper is organized as follows: Sect. 2 describes the related work. Section 3 includes problem statement. Sections 4 and 5 describe network model and proposed system. Section 6 shows performance evaluation with simulation results and Sect. 7 describes conclusion.

2 Related Work

This section describes the related work done in order to improve the performance of data aggregation in WSN so that lifetime of WSN will be improved.

SDRE: SVM-based data redundancy elimination for data aggregation in WSN. This technique makes use of support vector machine (SVM) and locality sensitive hashing (LSH) for redundancy elimination. SVM is used for detecting outliers and the nodes with high redundancy. Redundancy of nodes is detected by using LSH code. Each node, when it gets data, generates LSH code for it and sends this LSH code to aggregator node. Aggregator node makes a table and finds redundancy count of each node and if redundancy count is more than particular threshold, then that node is not considered in data aggregation process. In this way redundancy is eliminated to achieve larger lifetime [3].

DBST: Dynamically balanced spanning tree. This technique solves the hotspot problem. The reason behind this hotspot problem is that, only one node performs the role of aggregation for all sessions, thus its energy goes on decreasing continuously. This makes a node die early and this disturbs the network. DBST creates new routing tree for each session with new aggregator node and solves hotspot problem by balancing the traffic load [2].

BHCDA: Bandwidth efficient heterogeneity aware cluster-based data aggregation. This technique works on heterogeneous nodes. Geographical area is divided into grids. Clustering approach is used for routing and data aggregation. Redundancy is eliminated by performing intracluster and intercluster aggregation. This approach saves bandwidth by eliminating redundancy [4].

REDD: Redundancy eliminated data dissemination. It divides geographical area into grids and performs intracluster and intercluster redundancy elimination as BHCDA. One extra thing done by REDD is, it validates data before performing aggregation. It makes use of context aware system, this system has a rule engine which checks data sensed by nodes against rules. Data which satisfies rules are considered as valid otherwise invalid. In this way this technique assures accurate data aggregation [5].

EERDAT: Energy efficient reliable data aggregation technique. This technique dynamically changes the size of the cluster by adding or removing forwarding nodes. To add or remove nodes from the cluster is determined by calculating loss ratio. Loss ratio calculation is done by a node outside the cluster called coordinator node (CN). If loss ratio is more, packet loss is more and hence more number of nodes are added so as to reduce packet loss. If loss ratio is very less, forwarding nodes are removed so that energy consumption will be less. In this way this technique balances trade-off between reliability and energy conservation [6].

We have briefly discussed all the above techniques in our survey paper [7]. Particular issue, either lifetime or traffic load, is addressed in all the above techniques. Proposed system's focus is on solving all these issues of data aggregation.

3   Problem Statement

Data redundancy provides reliability. In densely deployed WSN nodes will generate redundant packets. This redundancy provides reliability but at the same time it will generate heavy traffic on the network. Most of the nodes will waste their energy in forwarding packets towards sink and energy of sink is wasted in processing redundant packets. This leads to extra energy consumption and which ultimately results into decrease in overall lifetime of WSN. Data aggregation eliminates redundancy and saves energy of node. At the same time data aggregation suffers from hotspot problem. Hence there is need to balance the load of an aggregator node.

In this paper clustering-based LSP for data aggregation is proposed. The proposed system will give better results for parameters such as lifetime and average energy consumption.

4   Network Model

Network model shows all components for our proposed system. Let the network model be M

$$M = \{N, C, \text{CH}, \text{SH}, \text{BS}\}$$

 where

N: Number of nodes, C: Cluster of sensor nodes, CH: Cluster head, SH: Sector head, BS: Base station.

In this model geographical area is divided into various sectors. A number of sensor nodes are deployed in each region to sense information about agriculture field. Sensor nodes are deployed in such a way that there are some redundant nodes for every node placed. The reason behind such placement of nodes is to achieve reliability from node failure. If any node fails we can collect information about it from its redundant nodes. It will improve reliability of our proposed system. Multiple clusters are formed in each sector. Cluster head is elected on the basis of energy and degree of a node. Each sector contains one sector head (SH) that combines the information coming from all CH's and send it to base station (BS). Network model is shown in Fig. 1 [8].

5   Proposed System

Proposed clustering-based LSP for data aggregation scheme is explained with the help of state diagram as shown in Fig. 2.

Fig. 1 Network model for clustering-based multilevel data aggregation
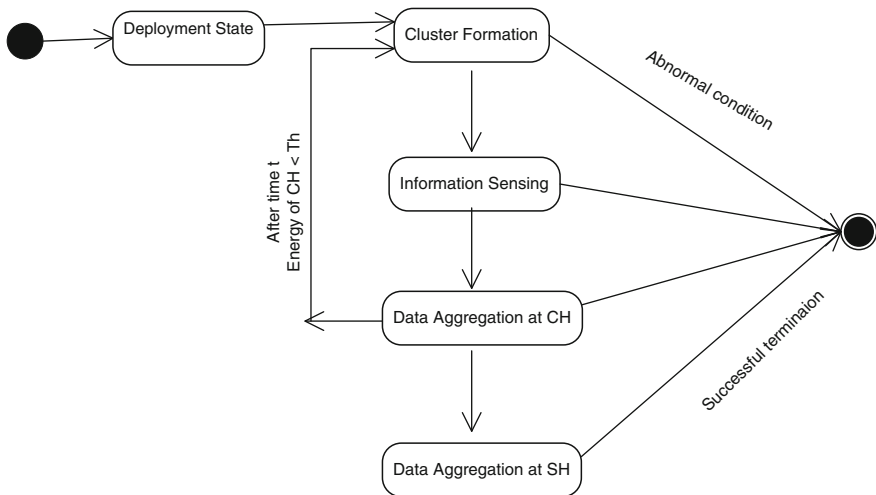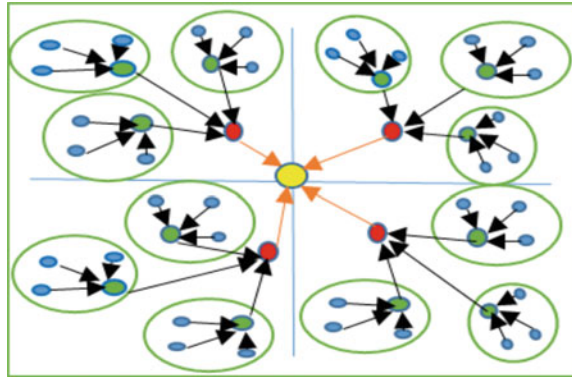
Fig. 2 State diagram of LSPDA

To eliminate redundancy in WSN, there should be some mechanism that enhances the lifetime of sensor networks. On the contrary, redundancy sustains reliability. Therefore there is need to have some mechanism which maintains trade-off between energy conservation and reliability. The proposed system, i.e., clustering-based LSP will perform data aggregation at two levels, one at cluster head and another at sector head node. Within the cluster redundancy count of each node is calculated. Based on this redundancy count only 50 % of nodes are allowed to send data. When 50 % nodes are sending data then remaining 50 % nodes are in sleep mode.

Fig. 3 System architecture of LSPDA
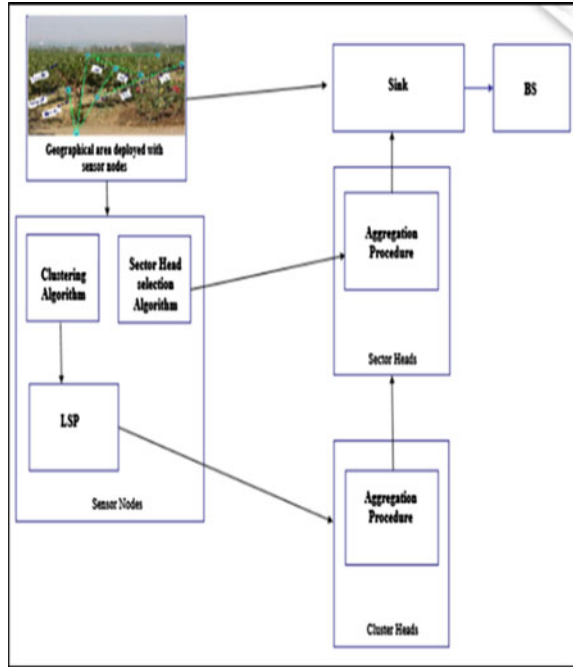
In next round, nodes those are sending data currently are forced to enter into sleep mode. This is called as load shifting policy (LSP). Due to this LSP, number of nodes participating in aggregation will be reduced; this will result into less traffic and less energy consumption. The distance between nodes are calculated and if two nodes are in sensing range of each other, they are considered as redundant nodes. Here system is not eliminating all redundant nodes. The purpose is to keep some data redundancy to conserve reliability. Aggregation will be done at CH. SH will aggregate packets coming from all CHs to send resulting packets to BS. This system will try to maintain trade-off between energy efficiency and reliability and also it balances the traffic load (Fig. 3).

5.1 Deployment and Sink Selection

In this phase all the nodes are deployed in geographical area of size 2000×2000 m^2. This geographical area is divided into four sectors. The node which is at the center of the whole region will act as a sink node. Sector head (SH) is elected for each sector (Fig. 4).

Fig. 4 Node deployment
scenario
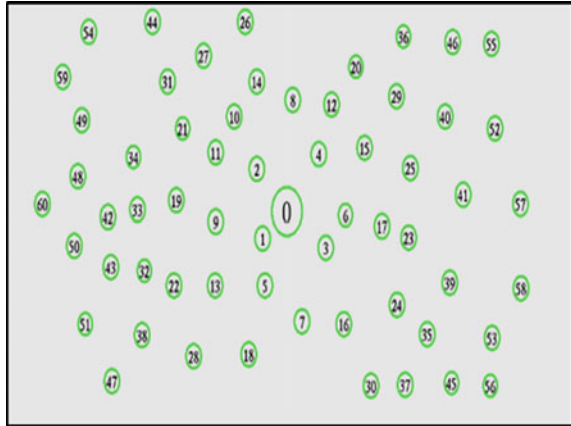
5.2 Sector Head Selection

For each sector, there will be one sector head. This SH is one hop away from the
sink. The algorithm used for sector head selection is as follows:

```
For each sector:
    If n be the nodes and n € Sector I
        For all n
            If (distance(node I, sink)) is lesser than all other nodes from same sector
             Node I will become SH
```

5.3 Cluster Head Election and Cluster Formation

This phase generates multiple clusters by electing CH. Each sector is virtually
divided into four subsectors. In each subsector the node with highest number of
neighbors are elected as CH. Initially CH is elected based upon its degree. After few
rounds when the energy of CH will be reduced than any other node, the node with
highest residual energy will become CH.

Steps for CH selection are

Each sector is logically further divided into 4 regions.
Nodes with highest degree are identified from each region of a sector and
elected as CH.
All nodes coming under CHs range will form a cluster.

Fig. 5 Scenario after sink, SH and CH election, and redundancy calculation
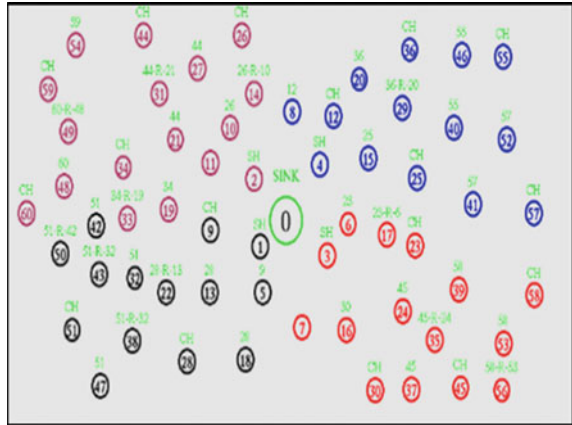
5.4 Calculating Redundancy Count (RC)

This phase calculates redundancy count of each node inside the cluster. If two nodes are coming in each other's sensing range, they will have the same data value. For each node, a number of redundant nodes are calculated. This node count is termed as RC count. Figure 5 shows cluster formation and redundant nodes for each CH.

```
1. Begin
        initialization
        int rc=0;                          //redundancy
count
        boolean rcflag = 1;                //redundancy
check flag for all nodes
2. if rcflag of node(i) =1;
        rcflag=0;                          //node i's
redundant nodes are checked
        loop
        if(dist (node (i),node 1) < 200)
        rc++
        rcflag of node (i) = 0
        goto loop
        else
     i++ ,
   goto step 2
3. stop
```

5.5 Apply Load Shift Policy (LSP)

A total number of redundant nodes are computed after calculating RC value for nodes. Considering only one node instead of all redundant nodes, will surely

improve the network lifetime but reliability may be compromised. Accordingly, to maintain trade-off between reliability and energy conservation, 50 % of redundant nodes are considered in aggregation process. The remaining 50 % nodes are kept in sleep mode to save energy because even though they are not taking part in aggregation process their energy is wasted in receiving packets. After one round, active node is forced to sleep and sleep nodes are awaken for data collection. In this way, the load of data transmission is rotated among redundant nodes. This policy will not only balance the energy consumption among the nodes inside the cluster but also improve the network lifetime.

5.6 Aggregation at CH and SH

After applying LSP the eligible nodes will send data to CH and it will perform aggregation of packets. Aggregation is the process of combining the packets coming from different sensor nodes and sending one packet to the next level. In this way if 4 eligible nodes are sending 10 packets each, CH will receive 40 packets and it will send only 10 packets to next level. This will help to reduce the traffic on the network. All CHs from particular sector will send their aggregated packets to their respective SH. SH will perform aggregation of packets and will send resulting packets to sink. Energy of sink node will be saved as it has very less number of packets to process.

6 Performance Evaluation

NS2 simulator is used to evaluate performance of proposed system [9]. WSN of size 2000 m × 2000 m s is created by deploying the sensors randomly. Initially 100 nodes are deployed randomly in the agricultural field. Sink node is assumed at the center of the field. Table 1 summarizes the simulation parameters and their values are used in the simulation.

6.1 Performance Metrics

A. Average Energy Consumption (AEC)

It is the total amount of energy consumed by nodes in transmitting and receiving the data packets to the total number of participating nodes. If the total energy consumption increases, lifetime of sensor network decreases.

$$AEC = \frac{\sum \text{Energy consumed of participating node}}{\text{Total number of participating nodes}}$$

Table 1 Simulation parameters

Parameter	Value
Network area	2000 m × 2000 m
Number of nodes	60, 70, 80, 90, 100
Position of base station	1000, 1000
Number of sensor nodes	120
Transmission radio range	250, 300, 350 m
Maximum buffer size (packet)	100
Initial energy	100 J
Data rates	50, 100, 150, 200, 250 bytes/s
Node distribution	Random
Directional antenna	Omni directional
Energy model	Radio energy model
Interval for sending packet	3 s

Life Time Lifetime of WSN is defined as the total time till the last node dies.

$$\text{Lifetime} = \frac{\text{Avg. Initial Energy} * \text{time } t \text{ s}}{\text{Avg. Energy consumed at time } t \text{ s}}$$

B. Result Analysis (Based on Data Rate)

In this simulation initially 100 nodes are considered. The data rate is varied from 50 to 250 bytes/s, in the step of 50.

Figure 6a shows graph of packet delivery ratio (PDR) versus data rate. It is observed from the graph that, as the data rate is increased, PDR goes on decreasing.

In Fig. 6b, graph of AEC versus data rate is shown. It is observed from the graph that average residual energy is more when data rate is 200 bytes/s but it is lowest for 150 bytes/s.

Figure 6c shows the graph of total energy consumption (TEC) versus data rate. From the graph, it is observed that for data rate of 150 bytes/s total energy consumption is highest and for 200 it is lowest.

C. Comparison with Traditional Aggregation Technique

In this section we have shown the comparison of performance of our proposed system (LSPDA) with traditional clustering-based data aggregation (CBDA) and with clustering but without aggregation (WA) for average energy consumption and lifetime.

Based on the graph, number of nodes versus average energy consumption is shown in Fig. 7a. It is observed that average energy consumption is more if aggregation is not performed (i.e., WA). The graph also shows that proposed LSPDA has better results for average energy consumption than CBDA.

Similarly, number of nodes (*X* axis) versus lifetime (*Y* axis) is shown in Fig. 7b. It is observed that the lifetime achieved from LSPDA is more than other two CBDA and WA. Hence LSPDA prolongs the lifetime of WSN.
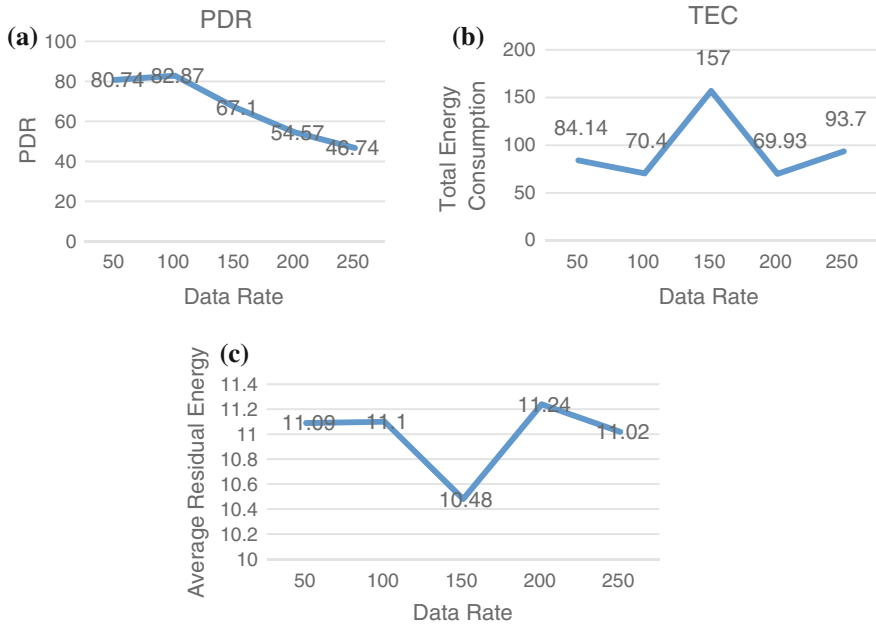
Fig. 6 **a** PDR versus data rate. **b** TEC versus data rate. **c** PDR versus ARE
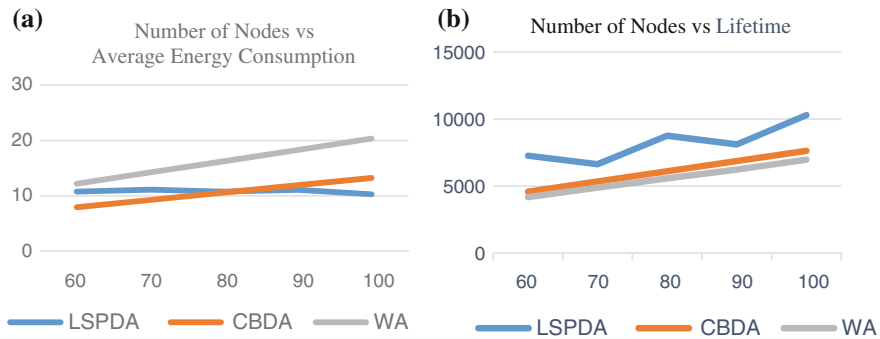
Fig. 7 **a** Comparison of techniques. **b** Comparison of techniques based on AEC on lifetime of WSN

7 Conclusion

Data aggregation process reduces redundancy from WSN and minimizes the overhead of redundant processing at sink node. But the intermediate nodes performing aggregation suffers from the problem of redundant processing. Our system removes redundant nodes from participating in aggregation process with the help of

LSP and it improves the performance of data aggregation. System also prolongs the lifetime of WSN and adequate redundancy removal maintains reliability too. Simulation result proves our claim that proposed system has more WSN lifetime than traditional aggregation techniques. Our system will work well in agricultural environment, because once you deploy the nodes they will work for long time.

In future we will try to achieve accurate results of aggregation by eliminating outlier values due to environmental changes.

References

1. Nandini Patil, P. Patil: Data Aggregation in Wireless Sensor Network, IEEE International Conference on Computational Intelligence and Computing Research, ISBN: 97881 8371 3627.
2. Avid Avokh, Ghasem Patil: Dynamic Balanced Spanning Tree (DBST) for Data Aggregation in Wireless Sensor Networks, 5th International Symposium on Telecommunications (IST2010), 978-1-4244-8185-9/10 © 2010 IEEE (2010).
3. Prakash Patil, Umakant Kulkarni: SVM based Data Redundancy Elimination for Data Aggregation in Wireless Sensor Networks, 978-1-4673-6217-7/13 © IEEE (2013).
4. Dnyaneshwar Mantri, Neeli Rashmi Prasad, Ram jee Prasad: BHCDA: Bandwidth Efficient Heterogeneity aware Cluster based Data Aggregation for Wireless Sensor Network, ICRTIT 2011, 978-1-4673-6217-7/13 IEEE (2013).
5. Sumalatha Ramachandran, Aswin Kumar Gopi, Giridara Varma Elumalai, Murugesan Chellapa: REDD: Redundancy Eliminated Data Dissemination in Cluster Based Mobile Sinks, ICRTIT 2011, 978-1-4577-0590-8/11 © IEEE (2011).
6. Basavaraj S. Mathapati, Siddarama. R. Patil: Energy Efficient Reliable Data Aggregation Technique for Wireless Sensor Networks, International Conference on Computing Sciences, 978-0-7695-4817-3/12 © IEEE (2012).
7. Samarth Anavatti, Sumedha Sirsikar: Issues in Data Aggregation Methods in WSN: A Survey, 4th International Conference on Advances in Computing, Communication and Control (ICAC3'2015), 1877–0509 © Procedia Computer Science, Elsevier B.V., Mumbai-400050, Maharashtra, India, (April 03–04, 2015).
8. Samarth Anavatti, Sumedha Sirsikar: Energy Efficient Multilevel Data Aggregation technique for adequate redundancy removal in WSN, 4th Post Graduate Conference for Information Technology (iPGCon-2015), Amrutvahini College of Engineering, Sangamner, (24–25 March 2015).
9. Network Simulator: http://www.isi.edu/nsam/ns.

Retrieving Instructional Video Content from Speech and Text Information

Ashwini Y. Kothawade and Dipak R. Patil

Abstract The interest of today's generation to learn from video lectures is becoming popular due to its considerable advantages and easy availability than classroom learning. To involve into this, many institutes and organizations are using this method for teaching and learning. An enormous amount of data is generated in video lecturing form. To extract the desired information from the desired video from this vast video information available on internet becomes difficult. In this paper, we have used techniques for automatically retrieving the information from video files to collect it as a metadata for those files. For efficient retrieval of text from videos we use the OCR (Optical Character Recognition) tool to extract text from slides and ASR (Automatic Speech Recognition) tool for recognizing information from speech given by the speaker. First, we do segmentation and classification of video frames for identifying the key frames. Then the OCR and ASR tool is used for extracting the information from video slides and audio speech respectively. The collected data can be stored as a metadata for the file. Finally, the search can be made more efficient by applying clustering and ontology concept.

Keywords OCR · ASR · Video content retrieval · Instructional videos · e-Learning · Tele-lecture · Tesseract OCR · Video lectures indexing

1 Introduction

Instructional videos are used for various purposes to gain information like news bulletins, food recipes, training for new technology, to learn something that may be related to academics or hobbies and interest. Nowadays, the Internet provides vast amount of information about these things in video form. The user can get appropriate knowledge by watching and listening to videos. Sometimes, problems may

A.Y. Kothawade (✉) · D.R. Patil
Department of Information Technology, Amruvahini College of Engineering, Sangamner, India
e-mail: aykothawade@kkwagh.edu.in

arise in providing the exact input query through browser for the desired information or the user may want to see only a part of the video in which the concept is explained, without watching the whole video comprising 1–2 h. For these things to work, the user must know previously the content of every video lecture.

The content can be easily accessible through the metadata attached to the file. However, the metadata attached has very less information, which the creator inserts manually and which does not give the exact content of the videos. For example, the video has its title, date of creation, or author name attached as a metadata to a file, but that is not giving the extact content information of the videos. So there is need for a system that can access the content of the videos, make it as a metadata for the video file and return the exact result to the users' query.

All instructional videos mainly have two kinds of sources of information; either we get the information from the audio or the information is available in the video text form. To extract this information different techniques can be used. In our paper, we have used OCR for retrieving text from video and ASR for recognizing the audio of the video file.

The feature extraction technique can also be used by developers but as the extraction is based on features such as color, pixel intensity, etc. [1] it does not give appropriate results in the case of instructional videos, because video lectures have homologous features between frames with many frames having similar content. Hence, these techniques used in multimedia content retrieval cannot be used for instructional video retrieval. In traditional systems, the search for video lectures is provided based on the metadata linked to it, which is inserted manually by the creator of the video. There are many disadvantages identified due to this manual insertion technology, as only a limited amount of data can be provided with each video file such as its title, creation date, its type, size, etc., which are insufficient for large size videos covering many concepts. When the user fires the query for getting the data, the data may not be available in the metadata but the video may contain some information related to it. At this time, the search becomes inefficient due to the limited metadata information. To insert more information in metadata using this method is time- and cost-consuming.

Hence, for increasing the efficiency of the search, a more advanced technique is required, which collects the data from video files automatically and treats it as a metadata. In this paper, we use the ASR and OCR techniques for retrieving the content, which provides greater accuracy over metadata formation. The search will become more efficient than the existing one, which extracts the content by identifying texts on the video lecture file and analyzing the words spoken by the speaker [2]. The two major sources of videocontent are one who speaks and the other is from the slideshow content or handwritten text written on the board. Existing technologies use OCR for slide text retrieval and ASR for retrieving the audio information from the video lectures. They provide more relevant data by assuming that the slides contain the important topics with larger font that needs to be extracted and frequently used words by the speaker are stopwords that have to be removed. Many technologies and research on this provide results that may be variable in accuracy for text extraction.

In this paper, we study the automatic generation of metadata from video files based on the content information. The content information from the video files are extracted by optical character recognition tool; we use Tesseract OCR and audio content retrieved by the automatic speech recognition method. While extracting information using these tools, some challenges may have to be faced. Nowadays, video recording is done based on multi-scene format in which the frame may contain multiple scenes at a time, like a news bulletin in which a part of the monitor is covered by the speaker while at the top or bottom the news is displayed.

A challenge that may occur in OCR system is due to variation in creation, entering texts, handwritten slides, text in the form of graphs, tables, etc. Speech analysis, which is an important content of the information to be extracted also has some challenges such as the speaker's fluency of speech, his pronunciation, and background noise. For reducing this noise preprocessing is needed before performing extraction. The extracted texts by these both methods will generate a large amount of data, which needs to be reduced. The keywords are extracted further by calculating the term frequency inverse document frequency (TFIDF) score. For providing the linked videos of the given search, ranked keywords are clustered for different video files by k-means algorithm with similarity between terms calculated by the Euclidian distance algorithm. The different instructional video formats are as shown in Fig. 1.

In our system, we have taken five video lecture files [3] related to 'c' language. The result is shown for all the operations of the video file processing linked to this video. The results from proposed techniques are more efficient than the existing search techniques as it is more accurate in less time.

Fig. 1 Different types of scenes and different presentation Formats in instructional videos

2 Literature Survey

2.1 Video Content Retrieval

Madhav Gitte, Harshal Bawaskar, Sourabh Sethi, and Ajinkya Shinde developed a content-based video retrieval system that uses multimodal features to extract videos from multimedia warehouse. The system is not complex but it extracts information based only on feature extraction algorithm and then although the clustering method is applied it does not give efficient results in the case of video lectures, as the content in video lectures may be more and to extract the information by feature extraction may also return nontext blocks (like images) [4].

Leeuwisl et al. [5] focused on English speech recognition in Translanguage English database (TED) corpus. He developed the language model by lecture transcript. The training dataset provided has to be inserted manually, so it takes time and it is also difficult to be extended or optimized. The word error rate returned by this was nearly 40 %.

Hürst et al. [6] proposed a method that extracts audio information from video lectures by developing the different ASR method. Large vocabulary automatic speech recognition (LVASR) is used which handles all kinds of audio signals as low quality signal, noisy signal, fast speaker and word detection, which improves word accuracy over other methods for audio retrieval. But in this system, the ranking to be given to words identified is difficult and the method does not collect any information from the slides in the video, which contain most important information and the method is also applicable only for German lecture videos. Xiao et al. suggested the new technique to be used with the ASR, which is more efficient than the traditional HMM-based ASR system. The author has developed an IR system which identifies the subword units into the valid word. Initially, instead of decoding the speech into word, it is decoded first in subword units [7].

Matthew Cooper compared the results from both the OCR and ASR techniques. He shows that there is much difference in the texts extracted from ASR and OCR. Also, he has proved that the OCR technique gives more accurate results for text extraction than the ASR technique. Because the ASR has many sources of error like pronunciation of the speaker, phonetic errors, etc., the errors that occur in this will directly affect the retrieval performance which does not happen in OCR [8].

Hauptmann et al. show a different video retrieval technique for the multimedia video data, which generates and combines all the metadata. The authors also show the results for retrieval by using all methods and compare them. The authors prove that the result is more efficient when retrieval is done by using enhanced OCR with preprocessing done effectively, the probabilistic model is used, and if the speech recognition algorithm is also used with it [9]. Tiecheng Liu and John R. Kender developed a new software tool for content retrieval of video lectures based on its low level features. The system especially works on real-time content retrieval and

key frame extraction from videos which have handwritten slide scenes. The authors used a rule-based model for detecting the boundaries between frames. As the handwritten slides are there for extraction so it has a substantial amount of frames to be generated for each event, which has more similarity in content. So it becomes difficult to identify the boundaries or shots between frames rather than to find in other scene type like already prepared slides or notes. However, retrieval is only done with the content and no further processing is done for efficiency of the search [10].

2.2 Automated Tagging

In [11], the author has extracted information from video lectures by speech recognition and slide transition using the OCR tool which gives accuracy for audio retrieval at 45 %. Also, in [12], the author has used MPEG-7 as a metadata framework for inserting collaborative information from multimedia video files. Then annotation and tagging are provided based on this information, which is extracted from the vocabulary defined in linked data. The annotation done by this is time-consuming and cost inefficient.

In our paper, we first carry out segmentation as a preprocessing task and then apply OCR and ASR methods for creating the metadata. The TFIDF score is calculated and based on the keyword frequency at which the clusters are formed. While searching for the video file, initially the clusters are identified and then all the video files within the cluster are returned as a result.

3 Proposed Model

3.1 Automated Tagging

The objective of the system is to provide efficient search for video lectures through the browser which can be done by adding more relevant metadata with the lecture video files. The first step is segmentation, then text extraction is done, and after that the clusters are formed.

The architecture of the proposed system is shown in Fig. 2.

As shown in the architectural diagram, metadata extraction is done from visual as well as audio resources of the video lectures using OCR and ASR techniques. For forming clusters the frequency of the extracted words by ignoring stopwords is calculated. TFIDF score is calculated and the word with the highest TFIDF will form the cluster of that video file.
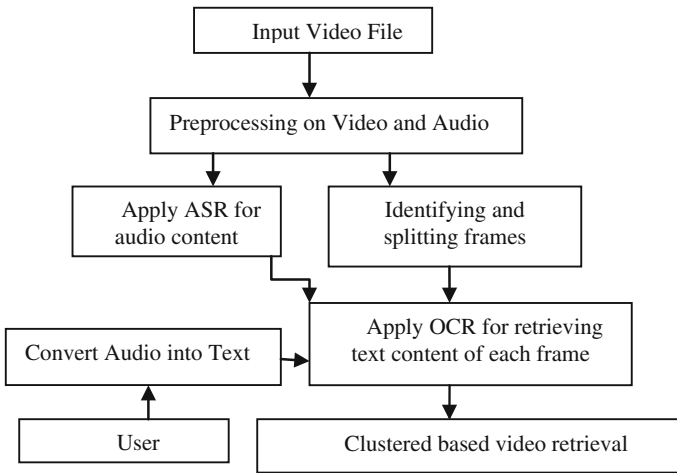
Fig. 2 Architecture of proposed system

3.2 Implementation Modules

The modules involved in the proposed system are as described below.

3.2.1 Slide Segmentation

Segmentation is done to identify the key frames. When the frames are separated from the lecture video files, many of the frames we get are similar in content due to monotonous scene view.

The segmentation is to be done by title or subtitle on the slide. The segmentation method is used for the video lecture files, which identifies the difference among frames by connected component (CC) analysis. The connected component analysis forms the group of similar intensity connected pixels. The segmentation method consists of the following steps:

- The time interval from 3 to 5 s is provided for analysis of frames. The frames coming after the given time period are considered as key frames while the others are discarded, assuming that the frames coming within interval are monotonous. Sometimes, the same frame is displayed for a longer period of time; to reduce duplication we have to increase the time interval of video segmentation.
- Actually, similarity detection between the content of the frames is done by connected component analysis. Initially the number of CCs is calculated for the first two key frames. This calculated number of CC will act as a threshold value for segmentation. The further segmentation is done only when the number of CCs exceeds this threshold value.

- In the next step of segmentation, the title and content region are first defined. The title region is identified by its font size. If the text covers more than 20 % of the frame height then it is identified as the title region. Any small change in the title region may cause slide transition.
- Again the threshold value is calculated and slide transition occurs when the difference among object regions exceeds this threshold value.

3.2.2 OCR Technique

Optical character recognition retrieves the text information from images and converts it into editable text. In this paper, we have used the Tesseract OCR tool, which is an open source and platform-independent tool for extracting data from video slides. For Tesseract the image needs to be converted into binary format [13]. For obtaining effective results by OCR tool, we need to do some preprocessing task on the key frames. The preprocessing task identifies the key frames from the video files.

The steps to be followed by the OCR tool are as follow [13]:

- The first step is adaptive thresholding, which converts input image into binary format.
- The next step is connected component analysis, which can be used to detect character outlines.
- Lines and words are analyzed within a fixed area or equivalent text size.
- Character outlines are organized into words by two passes. In the first pass, the word is recognized by text and is passed to an adaptive classifier. In the second pass, the adaptive classifier will have training dataset provided which can be used to resolve various issues for text extraction from images.

3.2.3 ASR Technique

The automatic speech recognition (ASR) technique extracts speech or voice from multimedia files and converts it into meaningful textual information. Speech is one of the most important carriers of information in video lectures. Therefore, it is of distinct advantage that this information can be applied for automatic lecture video indexing. Unfortunately, the ASR technique is still under development and does not provide the efficient results expected. The Word Error Rates returned by the existing systems are not as expected. ASR is aimed to enable computers to recognize speech characters without human intervention.

Automatic Speech Recognition model mostly uses the probabilistic approach for identifying the original word. When the word or word sequence is pronounced, its score is calculated using acoustic properties of phonemes for matching the word with the speech signal [14]. The ASR model has the following steps:

- Preprocessing
- Feature Extraction
- Decoding
- Postprocessing.

Preprocessing is done for removing unnecessary sounds in speech like background noise, door closing, voice, etc. The high pass filtering method can be used for reducing this noise and for identifying the speech and nonspeech segmentation. Signal energy-based algorithms can be used for identifying the start of speech segments; using this algorithm the speech segment can be easily detected when it crosses the given signal threshold value. As there may be some small energy signals with pauses between the words, the algorithm must be enhanced by time windowing.

For extracting the features, the acoustics observations have to be extracted from a time frame of uniform length of 25 ms. From this time frame a multidimensional acoustic vector is calculated. The human ear can respond to nonuniform frequency bands. Band-pass filtering can be used for nonuniform frequency bands by defining the frequencies in Mel scale. By discrete cosine transformations, the spectrums created are correlated. As the first coefficient carries the most significance, they are selected to form the feature vector. The resulting features are called Mel cepstra, for which further processing is done by cepstral mean subtraction. The vector created will be of high dimensionality. To project it into lower dimension, algorithms as principle component analysis are used.

Decoding is the process of matching the sequence of words with the acoustic represented by the feature vector. The prerequisite for decoding is availability of the dictionary, which has words to be spoken with its phoneme sequence. Three information sources must be available for decoding:

- An acoustic model with an HMM for each unit.
- A dictionary with a list of words and phonemes.
- A language model with word or word sequences.

For improving the recognition accuracy, rescoring is done by higher order language model. The triagram model-based rescoring is done in this system.

3.2.4 Database Creation

After extracting the text from audio and video, the collected texts are compared with the dictionary for getting the proper words or keywords. We have used the ELD dictionary. The words that are found in the dictionary after comparison are considered as valid words and are added into the database. The databases created are collected from OCR, ASR, and also from the metadata which is manually created for the video files. This database information will act as a metadata for those

Table 1 Created database after extraction

V_id	Image_id	Extracted texts (OCR and ASR)
1	11	Main, structure, lines, basic
1	12	Compile, data types
1	13	Performance, workers, basic
1	14	C, binary, data, type, function, print
2	22	Data, binary, assembly, language
2	23	Used, binary, *C* etc.
2	24	Integrity, method, function,
2	25	NUM, designing, UC
3	33	Language, code, scan
3	34	Operators, variables, for
3	35	While, for, *C*, object, variable
4	95	Pointer, structure, coding etc.
4	96	Longest, for, each, while
5	100	Learning, major, object, oriented

video files. The redundancies are checked between the words collected for avoiding wastage of storage space.

Table 1 gives an overview of some values from results of the OCR and ASR extraction for the five video files [3].

3.2.5 Clustering for Efficient Search

The databases collected from the above sources are large in size, in which all characters and words are present, including insignificant words like stopwords. The stopword removal technique is used which gathers only the important characters. The search is related to all the video files for the same query for which the dataset collection has to be arranged efficiently. For efficient finding of video files the clustering methodology is used. Clustering is done based on the frequency of the terms. The TFIDF score is calculated, which identifies keywords. These important words are again clustered for different video files for returning the result with all its related videos. Here, we use k-means clustering algorithm, which forms the clusters based on term frequency. The k-means clustering algorithm is more efficient than any other algorithm for text mining from images [15].

The formula for calculating TFIDF can be given as

$$\text{Tfidf}_{\text{vid-level}}(\text{kw}) = \text{tfidf}_g * w_g + 1/N(\text{tfidf}_{\text{ocr}} * 1/n\,\text{type})\sum_i w_i + \text{tfidf}_{\text{asr}} * w_{\text{asr}}$$

(1)

The algorithm of k-means clustering is as follows:

Algorithm k-means (input video extracted text linked to videos)

Step (1) Select the text with the highest TFIDF in one video.

Step (2) Search the other videos linked to this text with TFIDF greater than the threshold.

Step (3) Repeat Steps 2 and 3 until we get the same text with the highest TFIDF in all the video files.

4 Result Analysis

The result analysis of the system is done for the five video lectures of 'C' language, four of the object-oriented concept, and five from the data structures from the nptel. The result is conducted for different search keys like keyword, title of the video, audio clip of any video, or the video clip of any video. The precision and recall rates are calculated from the results shown in Table 2.

From the above values, the graph is generated for precision, recall, and time as shown in Fig 3.

Table 2 Result analysis

Search key	Precision	Recall	Time (in ms)
Keyword	0.78	0.67	1.1
Title	0.96	0.85	0.52
Audio clip	0.84	0.79	1
Video clip	0.74	0.65	1.3

Fig. 3 Result comparison between existing and proposed system by considering time for search as a parameter
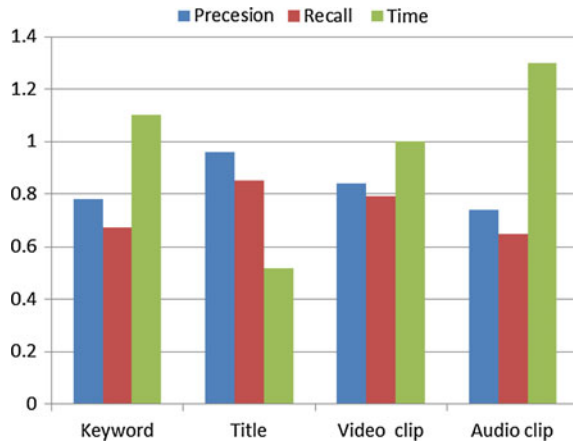
5 Conclusion

In this paper, we have developed an efficient information retrieval method from video lecture files. We have used Tesseract OCR with ASR technique with pre-processing methods, which reduces the word error rate considerably. After each stage of extraction the word is corrected by comparing with the dictionary. The clustering method used after data extraction, which is based on ontology of the terms, provides much related and desired metadata for the file and linked information for the videos. So video lecture browsing becomes efficient. This metadata created can be treated as indexing to the video file that is automatically created. As the results show, the proposed method also provides more accuracy with least search time. The future scope of this system will be to retrieve information for handwritten text which can be combined with the audio text.

References

1. Duc Phuong Nguyen, Martin Guggisberg, Helmar Appendix: Springer-Author Discount Burkhart,: Multimedia Information and Mobile-Learning, Eighth IEEE International Symposium on Multimedia (ISM'06), 2006.
2. Haojin Yang and Christoph Meinel,: Content Based Lecture Video Retrieval Using Speech and Video Text Information, IEEE Transactions On Learning Technologies, vol. 7, no. 2, April–June 2014.
3. P.P. Chakraborty, Programming and Data Structure (Video), C Programming-I, Indian Institute of Technology, Kharagpur. Available at http://nptel.ac.in/course.php?disciplineId=106.
4. Madhav Gitte, Harshal Bawaskar, Sourabh Sethi, Ajinkya Shinde,: Content based video retrieval system, International Journal of Research in Engineering and Technology, volume: 03 issue: 06 | June 2014.
5. E. Leeuwis, M. Federico, and M. Cettolo,: Language modelling and transcription of the ted corpus lectures, in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., 2003, pp. 232–235.
6. Wolfgang Hürst, Thorsten Kreuzer, Marc Wiesenhütter, A qualitative study towards using large vocabulary automatic speech recognition to index recorded presentations for search and access over the web:, ICWI, page 135–143, IADIS, 2002.
7. Xiaoqiang Xiao, Jasha Droppo and Alex Acero,: Information retrieval methods for automatic speech recognition, IEEE international conference on Acoustic Speech and Signal Processing, 2010.
8. Matthew Cooper,: Presentation Video Retrieval using Automatically Recovered Slide and Spoken Text, Multimedia content and mobile devices, SPIE proceedings, vol. 8667, 2013.
9. Alexander G. Hauptmann, Rong Jin, and Tobun D. Ng,: Video Retrieval using Speech and Image Information, Electronic Imaging Conference (EI'03), Storage Retrieval for Multimedia Databases, Santa Clara, CA, January 20–24, 2003.
10. Tiecheng Liu and John R. Kender,: Rule-based semantic summarization of instructional videos, IEEE International conference on Image Processing, vol 1, 2002.
11. Chirag Patel, Atul Patel, Dharmendra Patel, Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study, International Journal of Computer Applications (0975–8887) volume 55, no. 10, October 2012.

12. Vijaya Kumar Kamabathula, Sridhar Iyer, Automated Tagging To Enable Fine-Grained Browsing of Lecture Videos, IEEE International Conference on Technology for Education (T4E), 2011.
13. Harald Sack, Jörg Waitelonis,: Integrating Social Tagging and Document Annotation for Content-Based Search in Multimedia Data, Proceedings of the 1st Semantic Authoring and Annotation Workshop (SAAW'06), Athens (GA), USA, (November 2006) ISSN 1613-0073.
14. Dong Yu, Li Deng,: Automatic Speech Recognition, Springer Signal and communication technology, 2012.
15. Hong Liu, and Xiaohong Yu,: Application Research of k-means Clustering Algorithm in Image Retrieval System, Proceedings of the Second Symposium International Computer Science and Computational Technology, 2009.

Performance Improvement of SLM-Based MC-CDMA System using MIMO Technique

Madhvi Jangalwa and Vrinda Tokekar

Abstract This paper presents performance of SeLective Mapping (SLM)-based Multicarrier Code Division Multiple Access (MC-CDMA) system with MIMO technique. MC-CDMA fulfills the requirement of the forthcoming generation of wireless communication and offers high data rate with additional benefit of low Inter Symbol Interference (ISI). However, high Peak to Average Power Ratio (PAPR) is the major limitation of the MC-CDMA system. This high PAPR reduces efficiency of High Power Amplifier (HPA) and radio frequency components at the base station. To reduce PAPR, SLM is a simple approach, in which input data sequences are multiplied by M_p phase sequences before multicarrier modulation, and then the one with the minimal value of PAPR is chosen for transmission. In this work MIMO technique is used with SLM-based MC-CDMA downlink transmission system, which not only reduces PAPR but improves the BER performance. The results show that the PAPR of the proposed scheme is lower than the SLM-based MC-CDMA system. In addition, BER performance of proposed scheme outperforms SLM-based MC-CDMA system. A comparative study of PAPR and BER performance shows that low PAPR can be achieved by the proposed scheme with improved BER.

Keywords SLM · MC-CDMA · PAPR · HPA · MIMO · BER

1 Introduction

MC-CDMA is a suitable choice for forthcoming generation of wireless communication as it exploits high data rate, high spectral efficiency, with additional benefit of immunity against narrow band interference. However, high PAPR is one of the

Madhvi Jangalwa (✉) · Vrinda Tokekar
Institute of Engineering and Technology, Devi Ahilya University, Indore, India
e-mail: mjangalwa@yahoo.com

Vrinda Tokekar
e-mail: vrindatokekar@yahoo.com

challenging issues in the MC-CDMA system. MC-CDMA is a hybrid technique which combines Orthogonal Frequency Division Multiplexing (OFDM) and Code Division Multiple Access (CDMA) techniques [1, 2]. MC-CDMA takes the advantages and disadvantages of OFDM techniques. It takes the advantages of OFDM technique in terms of high data rate, high spectral efficiency, and provides frequency diversity. In contrast, it takes the disadvantages of OFDM technique in terms of high PAPR, which increases nonlinear distortion in communication channels and reduces efficiency of High Power Amplifier (HPA) and radio frequency components at the base station [3].

For reduction of PAPR a number of methods have been proposed. These suggested methods include clipping, clipping and filtering, SLM, Partial Transmit Sequence (PTS), SLM without side information, and many other techniques [3–9]. Clipping of OFDM signal is a PAPR reduction technique which reduces PAPR, but with this technique BER increases due to signal distortion [3, 4]. In [4], it is shown that SLM and PTS are distortionless which reduces PAPR; however, these methods have greater computational complexity and additional side information is required for recovery of the signal at receiver. A new SLM with low complexity is presented in [5], where to determine the optimal stage is slightly critical. An improved SLM method proposed in [6] reduces PAPR but this method is based on correlation analysis among the alternative symbols. In [7] low complexity PAPR reduction technique without side information is presented. Although the technique proposed in [8] does not require side information having more computational complexity. In [9] a new modified SLM without side information is proposed, which has low computational complexity but the PAPR reduction and BER performance is same as compared with the modified SLM scheme (with CMs). Further, Multiple Input Multiple Output (MIMO) technique provides higher data rate without extended bandwidth and smaller error rate. In MIMO technique multiple transmitting antennas and receiving antennas are used for transmitting and receiving the signal. To achieve higher bit rate, multiple independent information signals are simultaneously transmitted through multiple antennas and to achieve the smaller error rate, the same information signals are transmitted and/or received through multiple antennas [2, 10].

In this paper, SLM-based MC-CDMA with MIMO technique is proposed for downlink transmission system to achieve low PAPR with improved BER performance. In the proposed technique for phase rotation M_p different phase sequences are generated. These phase sequences are multiplied with spreaded signal. The resultant generated sequences are multicarrier modulated to get the phase rotated MC-CDMA signals. After that, the lowest PAPR valued signal is selected and encoded by STBC encoder for transmission.

The remainder of this paper is organized as follows: System description is described in Sect. 2. MIMO technique is given in Sect. 3. Results and discussion are presented in Sect. 4. Finally, Sect. 5 concludes the paper.
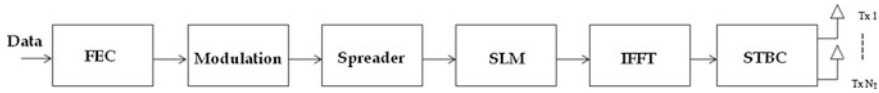
Fig. 1 SLM-based MC-CDMA transmitter with MIMO

2 System Description

Consider a downlink SLM-based MC-CDMA system with MIMO technique as shown in Fig. 1. U users data are randomly generated, encoded, and modulated. These narrow band modulated data are converted into wideband by *Walsh-Hadamard* (*W-H*) orthogonal code. These spreaded data of U users are added and then SLM operation is performed on this spreaded data. In SLM, M_p different phase sequences are generated. These phase sequences are multiplied by spreaded data. These resultant data are modulated by multicarrier modulation, and then the one with the minimal value of PAPR is chosen for STBC encoder. This lowest PAPR valued signal is given as

$$s(t) = \sum_{j=1}^{M} \sum_{u=1}^{U} a_\mathrm{u} d_\mathrm{u}(j) c_\mathrm{u}(t - jT_\mathrm{s}) \tag{1}$$

where the number of the active users are U, j is the number of the information symbol, length of the information symbol is L_M, amplitude of uth user is a_u, jth binary data sequence of uth user is $d_\mathrm{u}(j)$, and c_u is the spreading code of uth user [11].

The continuous-time PAPR of MC-CDMA signal is represented as the ratio of peak instantaneous power and average power. It can be represented as follows:

$$\mathrm{PAPR} = \frac{\max |s(t)|^2}{E\left(|s(t)|^2\right)} \tag{2}$$

where the numerator represents peak power and denominator represents average power of the MC-CDMA signal respectively.

3 MIMO Technique

The lowest PAPR valued signal is encoded by STBC encoder. In this work, STBC based on Alamouti scheme is considered for MIMO technique. For a given time slot $\tau = t$, two symbols c_1 and c_2 and at time slot $\tau = t + T$ two symbols c_1^* and $-c_2^*$ are transmitted concurrently from the two antennas. The encoded pattern of antenna one is s_1 and of antenna two is s_2, which is given as

$$s_1 = c_1(\tau = t); -c_2^*(\tau = t + T) \tag{3}$$

$$s_2 = c_2(\tau = t); c_1^*(\tau = t + T) \tag{4}$$

where * denotes complex conjugate operation. Two time slot t and $t + T$ are required to transmit two symbols [10]. The encoded matrix for two transmit antennas is given as

$$C_2 = \begin{bmatrix} c_1 & c_2 \\ -c_2^* & c_1^* \end{bmatrix} \tag{5}$$

4 Results and Discussions

The proposed method for SLM-based MC-CDMA system is simulated for *16QAM* modulation. Reed Solomen and convolution code is used for channel coding with 3/4 coding rate. *Walsh-Hadamard* code is used for spreading the user data, spreading code length L is 8, number of subcarrier Nc is equal to the length of spreading code, cyclic prefix is taken as ¼, and the total number of active users are U.

Figure 2 compares the results of PAPR for the proposed method, i.e., SLM-based MC-CDMA with MIMO scheme and SLM-based MC-CDMA scheme.
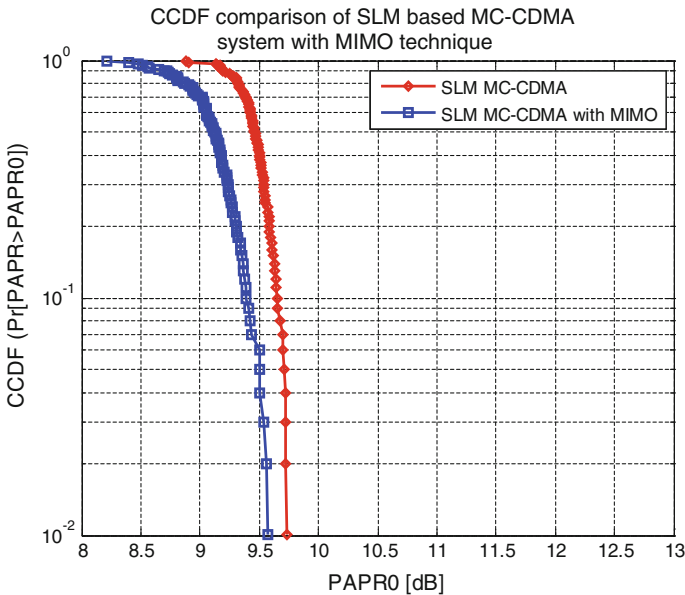
Fig. 2 CCDF of the PAPR for 16QAM, coding rate = 3/4, $N_t = 2$, $N_r = 4$, $U = 8$
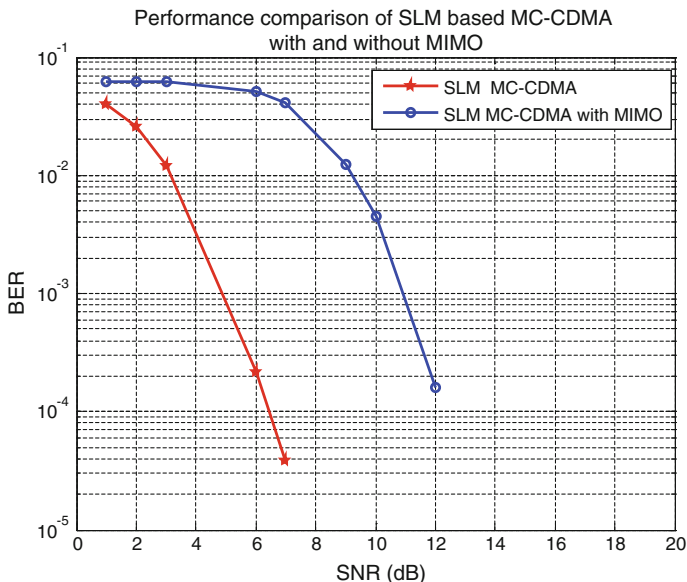
Fig. 3 BER versus SNR for 16QAM, coding rate = 3/4, $N_t = 2$, $N_r = 4$, $U = 8$

The value of PAPR0, at CCDF = 10^{-2} is 9.567 dB and 9.731 dB corresponding to the proposed scheme and SLM-based MC-CDMA scheme. The proposed scheme reduces PAPR compared to SLM-based MC-CDMA system.

BER performance of the proposed method and SLM-based MC-CDMA scheme for antenna configuration $N_t = 2$ and $N_r = 4$ is shown in Fig. 3. From the results it is seen that at SNR = 7 dB, the BER is found to be 3.851×10^{-5} using the proposed scheme (SLM-based MC-CDMA with MIMO scheme), while 4.086×10^{-2} by SLM-based MC-CDMA system. From the results it is clearly visible that the proposed method outperforms SLM-based MC-CDMA system by exploiting diversity gain. The reason is that with MIMO technique, diversity gain increases, which reduces bit error rate. Therefore, this proposed technique not only improves the BER performance but also improves PAPR (shown in Fig. 2).

In Fig. 4 comparison is shown among the transmitting antennas ($N_t = 2$) and receiving antennas ($N_r = 1$, 2 and 4) for the proposed method. From the results fast decay is observed in BER with increasing receiving antennas.

For *16QAM* modulation technique, BER versus SNR performance with different number of active users ($U = 8$, 6 and 4) is shown in Fig. 5. It is observed that BER performance rapidly gets poor. This is due to the fact that with increasing the number of users, Multiple Access Interference (MAI) increases, which increases BER.
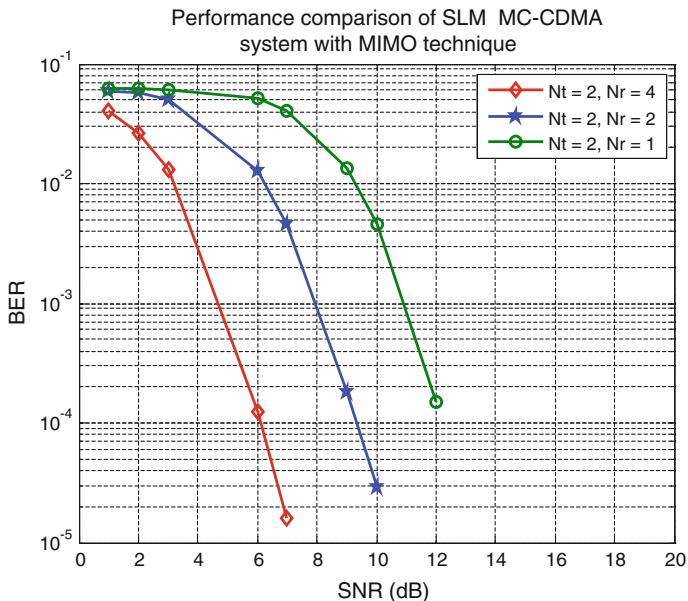
Fig. 4 BER versus SNR for 16QAM, coding rate = 3/4, $U = 8$
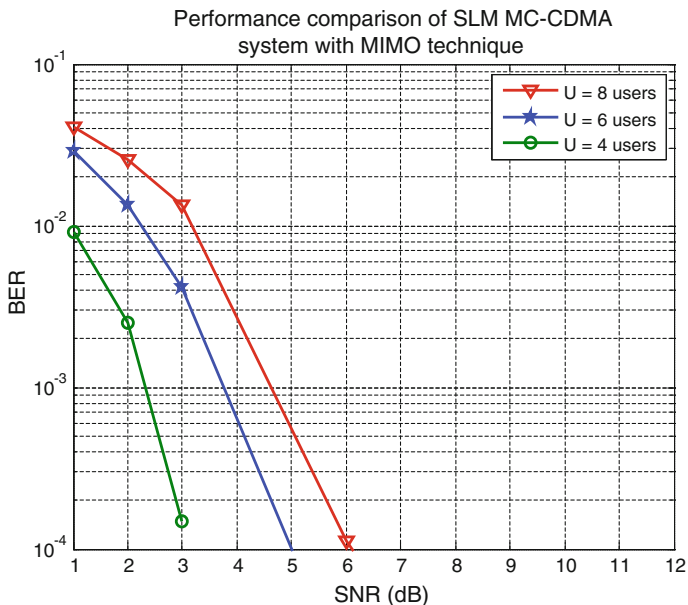
Fig. 5 BER versus SNR for 16QAM, coding rate = 3/4, $N_t = 2$, $N_r = 4$, $U = 8, 6, 4$

5 Conclusion

In this paper, SLM-based MC-CDMA system with MIMO technique is proposed for achieving low PAPR with improved BER. The results show that the proposed method reduces PAPR with a significant improvement in BER performance in comparison with SLM-based MC-CDMA system. BER performance of the proposed method is far better than the SLM-based MC-CDMA system by exploiting diversity gain. Performance of system also improves by increasing receiving antennas. From the results it is observed that there is fast decay in BER as the number of active user increases. By increasing the number of users, MAI increases which increases BER.

References

1. Hara, S., Prasad, R.: Overview of Multicarrier CDMA. IEEE Communications Magazine 35, 126–133 (1997).
2. Hwang, T., Yang, C., Wu, G., Li, S., Li, G.Y.: OFDM and Its Wireless Applications: A Survey. IEEE Transactions on Vehicular Technology 58(4) 1673–1694 (2009).
3. Rahmatallah, Y., Mohan, S.: Peak to Average Power Ratio Reduction in OFDM Systems: A Survey and Taxonomy. IEEE Communications Surveys & Tutorials 15(4) 1567–1592 (2013).
4. Jiang, T., Wu, Y.: An Overview: Peak to Average Power Ratio Reduction Techniques for OFDM Signals. IEEE Transactions on Broadcasting 54(2) 257–268 (2008).
5. Lim, D.W., No, J.S., Lim, C.W., Chung, H.: A New SLM OFDM Scheme with Low Complexity for PAPR Reduction. IEEE Signal Processing Letters 12 (2) 93–96 (2005).
6. Chang, P., Xiao, Y., Dan, L., Li, S.: Improved SLM for PAPR Reduction in OFDM Systems. 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications 1–5 (2007).
7. Eom, S.S., Nam, H., Ko, Y.C.: Low Complexity PAPR Reduction Scheme Without Side Information for OFDM Systems. IEEE Transactions on Signal Processing 60(7) 3657–3669 (2012).
8. Jiang, T., Ni, C., Guan, L.: A Novel Phase Offset SLM Scheme for PAPR Reduction in Almouti MIMO-OFDM Systems Without Side Information, IEEE Signal Processing Letters 20(4) 383–386 (2013).
9. Ji, J., Ran, G.: A new Modified SLM Scheme for Wireless OFDM Systems without Side Information. IEEE Signal Processing Letters 20(11) 1090–1093 (2013).
10. Mietzner, J., Schober, R., Lampe, L., Gerstacker, W.H., Hoeher, P.A.: Multiple-Antenna Techniques for Wireless Communications – A Comprehensive Literature Survey. IEEE Communications Surveys and Tutorials 11(2) 87–105 (2009).
11. Jangalwa, M., Tokekar, V.: Performance Improvement of MC-CDMA System over MIMO Channel using PIC Receiver. IETE National Journal of Innovation & Research 2(1) 20–24 (2014).

SNR Improvement for Evoked Potential Estimation Using Wavelet Transform Averaging Technique

M.L. Shailesh and Anand Jatti

Abstract Evoked potential (EP) comprises giving stimulus to the subject and record the response of the brain. Here, background noise electroencephalogram (EEG) is to be removed to see the response for the stimulus being given. In this paper, wavelet transform has been used to extract the responses and also to improve the signal to noise ratio (SNR). Wavelet transform averaging technique of estimation improves the SNR by a large amount in almost many sweeps of EP. The two different wavelet transforms such as Daubechies wavelet transform and Biorthogonal wavelet transform have been used to improve the SNR. SNR comparison is made with the conventional ensemble averaging technique. In this paper, Visual Evoked Potential (VEP) signals have been considered for analysis.

Keywords Wavelet transform · Biorthogonal · Daubechies · Evoked potential · Ensemble averaging · DWT · SNR

1 Introduction

An evoked potential or evoked response is an electrical potential recorded from the scalp of a human or other animal by giving proper kind of stimulus, and is distinct from the normal potentials produced as electroencephalography (EEG), electromyography (EMG), or other electrophysiological recording techniques.

In any clinic, usually evoked potentials (EP) [1] are recorded by giving the proper type of stimulus, depending on the kind of problems being considered. Normally, EP recoding is a noninvasive method, for obtaining the information or data for the purpose of analysis or diagnosis. The EP signal is originated from the

M.L. Shailesh (✉)
Pacific University, Udaipur, Rajasthan, India
e-mail: shailesh.ml@gmail.com

Anand Jatti
R.V. College of Engineering, Bangalore, Karnataka, India
e-mail: anandjatti@yahoo.com

central nervous system. If the strength of the EP signal is very faint or no signal appears across the electrodes output [2] but only background signal appears, i.e., EEG, then the physician recognizes this as a problem with regard to the corresponding nerve under test process. This is the key due to which EP signal plays a vital role in recognizing problems associated with the central nervous system. Visual evoked potentials (VEP) are recorded by giving stimulus such as flash of light or checker board [3]. The subject takes a little time to adjust to the test environment and later responds properly for each stimulus. After recording, it should be aligned to make different sweeps of data using computer. These sweeps of data are used for extracting the Evoked Potential signal.

2   Methods

The following two methods have been implemented in this paper.

2.1   Ensemble Averaging Technique

Generally, signal averaging is a technique [4] for separating a repetitive signal from noise without introducing much signal distortion. Ensemble signal averaging sums a set of time epochs of the signal together with the addition of random noise signal. If the epochs are properly adjusted, the mathematical analysis is as follows:

The input waveform eeg_ep(t) has a signal portion of interest S(t) and a noise portion N(t). Then

$$\text{eeg\_ep}(t) = S(t) + N(t) \tag{1}$$

Let eeg_ep(t) be sampled at every T seconds, where T is sampling interval. The value of any sample point in the time epoch [4] ($i = 1, 2, \ldots, n$) is the sum of the noise component and the signal component.

$$\text{eeg\_ep}(iT) = S(iT) + N(iT) \tag{2}$$

Each sample point is stored in memory. The value stored in memory location i after m repetitions is

$$\sum_{k=1}^{m} \text{eeg\_ep}(iT) = \sum_{k=1}^{m} s(iT) + \sum_{k=1}^{m} N(iT) \tag{3}$$

The recorded signal for sample point i is the same at each response if the signal is stable and the sweeps are aligned together perfectly. Then

$$\sum_{k=1}^{m} S(iT) = m\, S(iT) \tag{4}$$

The assumptions for this development are that the signal and noise are uncorrelated and that the noise is random with a zero mean. After many repetitions, $N(iT)$ has an rms value of σn.

$$\sum_{k=1}^{m} N(iT) = \sqrt{(m\, \sigma n^2)} = \sqrt{(m)}\sigma n \tag{5}$$

Taking the ratio of Eqs. (4) and (5) gives the SNR after m repetitions as

$$\text{SNRm} = \sqrt{m}\,\text{SNR} \tag{6}$$

Therefore, ensemble signal averaging technique improves the SNR by a factor of $\sqrt{m}$

$$\text{SNRm} = \sqrt{m} * \text{SNR} \tag{7}$$

where m is the number of sweeps.

2.2 Algorithm for Ensemble Averaging Technique

1. Consider different sweeps of data and assign it to different variables.
2. Add all the arrays or variable contents and assign it to new variable.
3. Compute the ensemble average by dividing it by the number of sweeps being considered.
4. Compute the output SNR for each sweep and store it in an array. Plot the SNR array elements.

2.3 Wavelet Transform Averaging Technique

Following are the two wavelet transforms used in this paper:

1. Daubechies Wavelet Transform(db14).
2. Biorthogonal wavelet Transform(Bior 4.4).

Fig. 1 Decomposition of EP signal
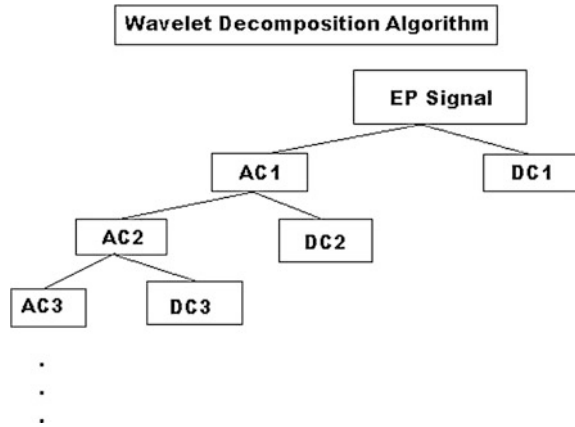
2.4 Algorithm for Wavelet Transform Averaging Technique

1. First decompose the signal using [5] discrete wavelet transform as shown Fig. 1.
2. Apply the Hard threshold for the wavelet coefficients.
3. Wavelet used is that where maximum SNR is obtained.
4. Reconstruct the signal using inverse discrete wavelet transform as shown in Fig. 2.
5. Repeat Steps 1–4 for different sweeps.
6. The results so obtained for each single sweep are to be ensemble averaged.
7. Plot the result.
8. Calculate the SNR.

Fig. 2 Reconstruction of EP signal
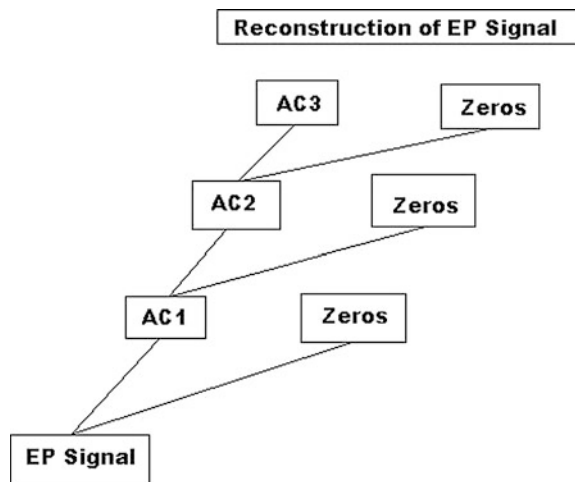
3   Results

3.1   Interpretation of Results

In this paper, simulated data have been taken for analysis as shown in Fig. 3, which shows the original evoked potential signal, which is a simulated signal. Figure 4 shows three different sweeps of data taken at sweep no. 11, sweep no. 21, and
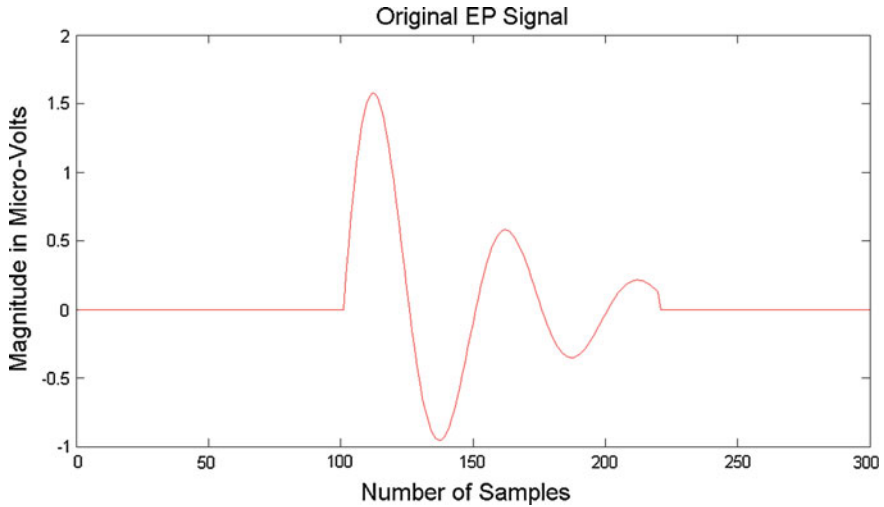
Fig. 3   Original EP signal
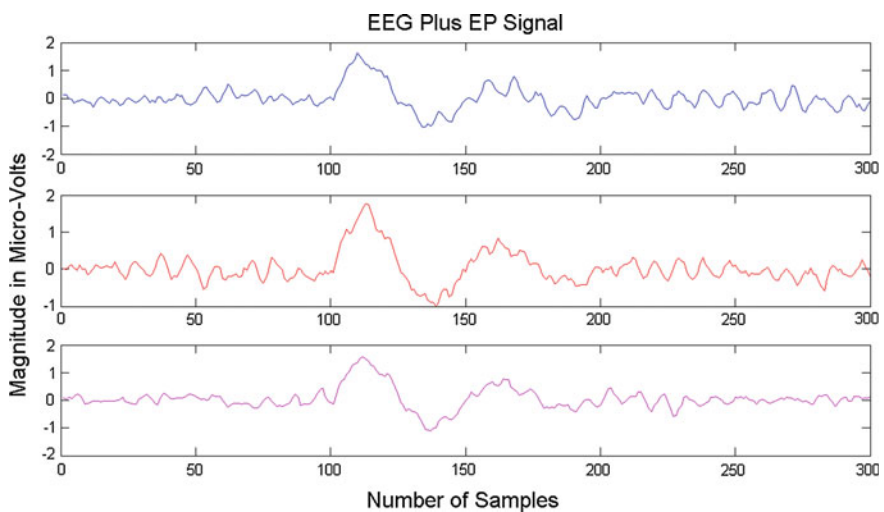
Fig. 4   EEG plus EP signal

Table 1 Ensemble average technique of SNR table for EP

Data	Number of sweeps	Ensemble averaging technique SNR (dB)
#1	10	17.1
#2	20	18.63
#3	40	19.48
#4	60	21.28

sweep no. 59 respectively. One sweep contains 300 samples, 60 such sweeps of data have been taken for analysis. Only three sweeps of data are shown in Fig. 4. The simulated signal contains EP and EEG signal and both signals are added to form the contaminated signal. For the ensemble averaging technique, 60 such sweeps of data have been considered for obtaining the output and also to calculate SNR values. Table 1 shows SNR values for different number of sweeps for obtaining the output.

Figure 5 shows the output waveform of ensemble averaging technique. The figure gives a description of the repetitive signals almost averaged to highlight EP signal. The strength of the noise signal which is an EEG (back ground signal) reduces as more number of sweeps is considered. Table 1 shows that as more number of sweeps are considered, SNR improves by a factor of almost the square root of number of sweeps.

In wavelet transform averaging technique of estimating EP signal, the same data have been processed to improve the SNR. Three levels of decomposition are performed for each wavelet transform. Hard thresholding is used for each decomposed signal, since EP signal is low frequency signal and background signal is an EEG
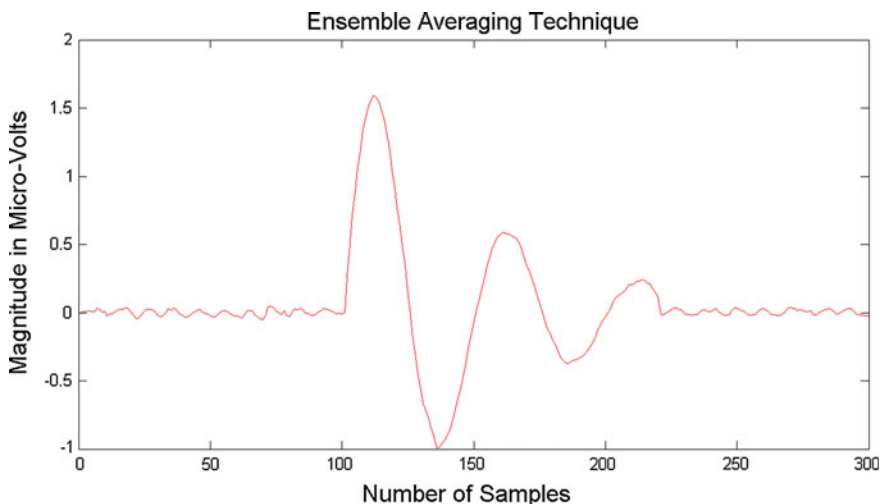
Fig. 5 Ensemble average technique output
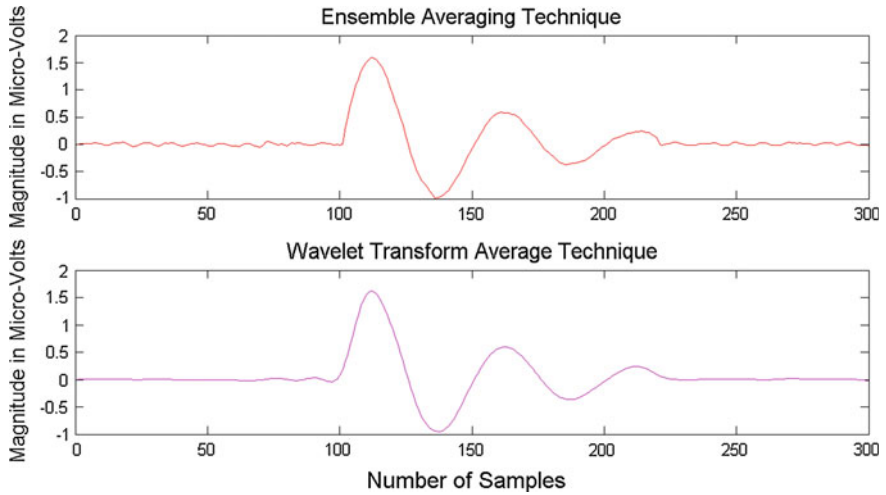
Fig. 6 Output Waveform of Ensemble Averaging Technique and Wavelet Transform Averaging Technique (Daubechies (db14) Wavelet Transform)

signal, which is a high frequency signal. In this method smooth curve is obtained since high frequency components are removed in the process.

The wavelet transforms used are Daubechies and Biorthogonal WT. These wavelet transforms (WT) have been chosen based on the best SNR obtained. Figure 6 shows the output plots of ensemble averaging technique and wavelet transform outputs averaging technique. Figure 7 shows the plots of original EP
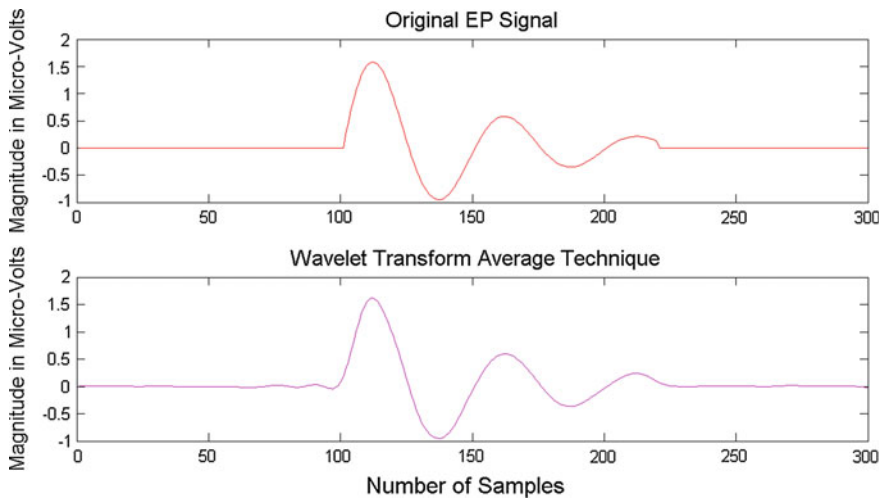
Fig. 7 Output Waveform of Original EP Signal and Wavelet Transform Averaging Technique (Daubechies (db14) Wavelet Transform)
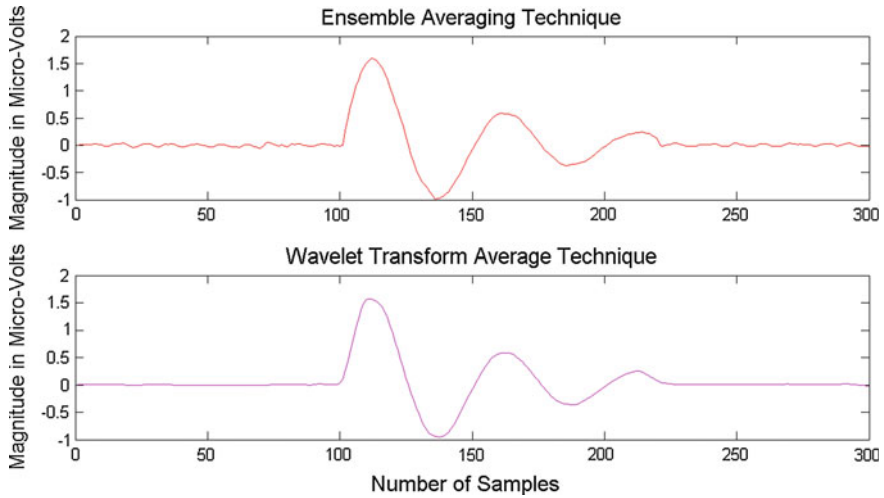
Fig. 8 Output Waveform of Ensemble Averaging Technique and Wavelet Transform Averaging Technique (Biorthogonal (bior4.4) Wavelet Transform)

Signal and wavelet transform outputs averaging technique. The wavelet transform used is Daubechies(db14). It is evident that the output of the proposed method does not contain much noise; almost all small noise peaks are eliminated. Similarly, Fig. 8 shows the output plots of ensemble averaging technique and wavelet transform outputs averaging technique. Figure 9 shows the plots of original EP
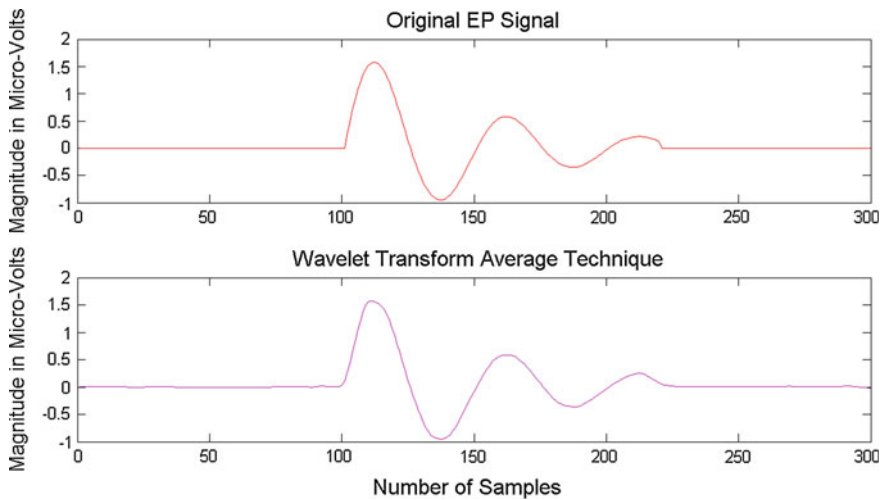
Fig. 9 Output Waveform of Original EP Signal and Wavelet Transform Averaging Technique (Biorthogonal (bior4.4) Wavelet Transform)

Table 2 Wavelet transform average technique of SNR table for EP

Ensemble average technique SNR in dB	Daubechies wavelet transform average technique SNR in dB	Biorthogonal wavelet transform average technique SNR in dB
21.28	27.25	27.9557

Signal and wavelet transform outputs averaging technique. The wavelet transform used is Biorthogonal (bior4.4) It is evident that the output of the proposed method does not contain much noise, almost all small noise peaks are eliminated. Here, more sweeps are required to estimate the EP signal. The output shown in Fig. 9 is better than that in Fig. 7, even in terms of output SNR. The output SNR values are tabulated in Table 2. This signifies the maximum SNR is obtained if the corresponding Biorthogonal wavelet transform is used.

4   Conclusion

In the ensemble signal averaging technique, it improves the SNR by a factor of $\sqrt{m}$. where m is the number of sweeps. It is evident that some noise is present in the output. In this paper Daubechies wavelet and Biorthogonal wavelet transform improves SNR in the results obtained and is more suitable for EP signal estimation.

In this paper it has been shown that the wavelet transform averaging techniques can be useful in the estimation of evoked potential in terms of quality of the signal and obviously, signal to noise ratios. To emphasize once again Biorthogonal wavelet transform performs better than that of any other wavelet transforms to improve SNR.

References

1. Zhisng Wang, Alaxander Maier David A. Leopold, Nikos K. Logothetis, Hualou Liang, "Single – Trial Evoked potential estimation using wavelets, Computers in Biology and Medicine 37, 2007 Pages 463–473.
2. Geeta Kaushik, Dr. H.P. Sinha, "Biomedical Signal Analysis through Wavelets: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 9, September 2012 Pages 422–428.
3. Mohd Zuki Yusoff, Nidal Karnel, Ahmad Fadzil Mohd Hani, "Estimation of Visual Evoked Potentials using a signal subspace Approach", International Conference on Intelligent and Advanced Systems 2007. Pages 1157–1162.
4. Willis. J. Tomkins Bio Medical Digital Signal Processing, PHI Publications.
5. Stephane Mallat A wavelet tour of signal processing, PHI Publications.

A GIS and Agent-Based Model to Simulate Fire Emergency Response

Mainak Bandyopadhyay and Varun Singh

Abstract In this paper a computing model for Fire Emergency Response is discussed. The Fire Emergency Response (FER) is considered here as a complex heterogeneous system. The FER system is composed of Fire Incident, Fire Station, Fire Emergency Vehicle, and Road Network components. Agent-based Modeling (ABM) is used to model the properties and behavior of each of the components. To incorporate the spatial properties and spatial operations in relation to modeling the behavior of components Geographical Information System (GIS) is used. The simulation of the model will provide the results of dynamical behavior of components through interaction with other components. The FER system model is implemented in GAMA 1.5.1 a GIS and agent modeling platform for simulation.

Keywords Fire emergency response · Agent-based modeling and simulation · Agent · Complex system · Geographical information system

1 Introduction

As the notion of smart cities is gaining popularity among policy makers in transitional countries, a better and smart emergency management infrastructure as part of smart cities is considered a basic component. Here, development of a computing model for fire emergency response (FER) is discussed. From a system point of view, FER can be divided into various behaviorally independent components like Fire Incident, Fire Emergency Vehicle, Fire Station, and Road Network, which trigger the activity of other components in some point of time or space. It can be regarded as a complex heterogeneous system. Agent-based modeling

Mainak Bandyopadhyay (✉)
GIS Cell, MNNIT Allahabad, Allahabad, India
e-mail: Ermainak@gmail.com

Varun Singh
Department of Civil Engineering, MNNIT Allahabad, Allahabad, India
e-mail: Vsingh.ce@gmail.com

(ABM) represents various constituent components of a complex system as Agent and Resources. Being a micro simulation approach ABMS captures both local and global system dynamics in a heterogeneous complex system [1]. While modeling a system as Agents the dynamics of the system is captured through interaction between the ABM entities. While modeling the FER as ABM, the components of the FER are modeled as ABM Entities, i.e., Agents and Resources. It is inevitable to accurately and completely model and simulate FER without including the spatial dimension. The locational information and related operations are incorporated in ABM using GIS. The GIS-based ABM is implemented in GAMA 1.5.1 for simulation.

2 Related Background

The modeling and simulation of complex heterogeneous systems requires identification of basic components of the system and modeling them as individual and autonomous entities. The overall dynamics of the system at a particular instance of time depends on the collective interaction of various components of the system according to their behavior.

2.1 Agent-Based Modeling and Simulation

Agent-based modeling represents various constituent components of a complex system as agent. A typical agent is a computing system situated in an environment that has autonomy, reactivity, proactiveness, and social ability as its essential characteristics [2]. The autonomy allows the agent to act independently without the direct direction of other agents; the behavior of the agent depends on the decision rules incorporated in the agent. The perceiving environment of an agent consists of a subset of ABM entities with which the agent interacts during simulation according to its behavior rules. An agent consists of properties and behavior rules through which it interacts with the entities in the perceiving environment and takes appropriate actions to achieve its objective.

Properties of Agent. The properties of an agent are of two types: physical and mental. The mental properties of an agent represent the perception of the agent and are updated through interaction with other agents in the perceiving environment, whereas the physical property represents the state of an agent at a particular instance of time. The physical property is updated based on the mental property and the physical properties of the agent [3]. It must be noted that only the physical properties of agents are available for accessing by other agents in their perceiving environment.

Behavior of Agent. The dynamics of the agent is divided into the following three actions which ensures the reactivity and proactivity characteristics of agents:

- *Agent_Action*: This action accesses the physical properties of the agent and based on the rules updates the physical properties, thus changing the state of the agent. This action ensures the proactivity of agents.
- *Agent_Percept*: The percept action accesses, analyzes, and manipulates the physical property of the entities in the perceiving environment and updates its mental property.
- *Agent_Perform*: The perform action accesses the mental property and physical property of the agent itself and updates the physical properties; thus it changes the state of agent.

The reactivity of agent is achieved by sequential execution of Agent_Percept and Agent_Perform actions [4].

3 Suitability of ABM for FER Modeling and Simulation

The main components in FER system are Fire Emergency Station, Fire Emergency Vehicles, Road Network, and Fire Incident. Each of these components is mostly independent in their behavior but influence or triggers each other's actions at some point of time or space. The various major interactions between the components in FER system are as follows:

- The information related to Fire Incident is communicated to the Fire Emergency Station at a particular time after which the Fire Emergency Station takes decision to dispatch Fire Vehicles.
- The Fire Emergency Station dispatches stationed Fire Vehicles.
- The Fire Vehicles moves toward destination according to the speed profile of Road Segment.
- The influence of Fire Incident starts decreasing once the Fire Vehicles reaches the Fire Incident Location.
- The Fire Vehicle recharges itself once it reaches the Fire Emergency Station or Resource Location.

3.1 Designing of Agents for FER

The various physical and mental properties of FER components are briefly provided in Table 1. The details of various actions, perceiving environment, accessing properties, and updated properties of Fire Incident, Fire Station, and Fire Emergency Vehicle are provided in Tables 2, 3, and 4 respectively. In the following sections detailed description of the FER components are provided.

Table 1 Properties of agents

Agent	Physical properties	Mental properties
Fire Incident	Location HeatReleaseRate (Q) RiskMass FireAlert FlashoverPoint	Attend MaximumHeatReleaseRate FireGrowthCoefficient EffectiveHeatofCombustion
Fire Emergency Vehicle	Location CurrentWaterCapacity Flow rate Status	Ready DispatchTime IncidentLocation StationLocation ResourceLocation VehicleSpeed MaximumWaterCapacity
Fire Station	Location TimeDispatch NumberofFEVDispatch	TimeofReceivingInfo PreparationTime DistancetoFireIncident DistancePerceptionforMultipleFEVsending FireSeverity

Table 2 Behavior of Fire Incident

Action	Type	Perceiving environment	Accessing properties	Update properties
Fire_Alert	Agent_Percept	Fire Incident	FireAlert	Attend
Onset	Agent_Perform	Fire Incident	HeatReleaseRate(Q), MaximumHeatReleaseRate FireGrowthCoefficient FlashoverPoint EffectiveHeatofCombustion RiskMass Attend	HeatReleaseRate (Q), RiskMass FlashoverPoint
Arrived	Agent_Percept	Fire Incident Fire Vehicle	Location(FI), Location (FV)	Attend
Offset	Agent_Perform	Fire Incident	Attend	HeatReleaseRate (Q)

Fire Incident Agent. In the Fire Emergency Response System, Fire Incident is the most important triggering component. Figure 1 shows the typical behavior or lifecycle of a Fire Incident Agent.

The growth phase of a fire starts after the ignition of combustible materials until it spreads on the entire material. The growth phase is crucial as it determines the rate at which fire is growing/spreading. The growth of a fire is described using t^2–fire growth model [5].

Table 3 Behavior of Fire Station Agent

Action	Type	Perceiving environment	Accessing properties	Update properties
Get_Fire_Info	Agent_Percept	Fire Incident	FireAlert	DistancetoFireIncident FireSeverity TimeofReceivingInfo
Dispatch_Decision	Agent_Perform	Fire Station	DistancetoFireIncident FireSeverity PreparationTime TimeofReceivingInfo DistancePerceptionfor -MultipleFEVsending	TimeDispatch NumberofFEVDispatch

$$Q = \alpha t^2 \tag{1}$$

where,

Q is the heat release rate (kW)

α is the fire growth coefficient (kW/s^2)

t is time (s)

Apart from t^2–fire growth model, which is specifically used for describing growth phase, the heat release rate can be measured using Eq. (2) [6].

$$Q = \dot{m}\Delta H_{eff} \tag{2}$$

where

$\dot{m}$ is mass loss rate (Kg/s)

ΔH_{eff} is the effective heat of combustion (KJ/Kg)

If unattended, the fire slowly decays with decrease in heat release rate and ultimately disappears. There is no well-established law to define the decay phase.

Fire Station Agent. The decisions regarding dispatching of resources or FEV is generally taken in the Fire Station. As the main control center it receives subjective information about fire incident which consists of location, locality, and materials/substance on fire, and processes the information to dispatch either single or multiple FEVs to the incident location.

Fire Emergency Vehicle Agent. Fire Emergency Vehicle unit carries resources from the Fire Station to the Fire Incident location. Water is the main resource used by the Fire Department to tackle Fire Incidents. Each FEV has some maximum capacity of water that it can carry. While carrying out emergency activities the volume of water decreases according to the flow rate from the equipment through which water is delivered to the fire, i.e., nozzle. When the resources are exhausted carrying out emergency activity the FEV returns to the Fire Station or Resource Location for recharging. Dijkstra's Algorithm on weighted road segment network models the route selection behavior of FEV drivers.

Table 4 Behavior of Fire Vehicle Agent

Action	Type	Perceiving environment	Accessing properties	Update properties
Receive_Dispatch_Message	Agent_Percept	Fire Station	TimeDispatch	DispatchTime IncidentLocation
Dispatch	Agent_Percept	Fire Vehicle	DispatchTime MaximumWaterCapacity CurrentWaterCapacity	Ready
Speed_Update	Agent_Percept	Road	Speed	VehicleSpeed
Move_To_Incident	Agent_Perform	Fire Vehicle	VehicleSpeed Ready IncidentLocation	Status Location
Use_Resource	Agent_Perform	Fire Vehicle	Location IncidentLocation MaximumWaterCapacity Flow Rate	CurrentWaterCapacity
Recharge_Alert	Agent_Percept	Fire Vehicle	CurrentWaterCapacity	Ready
Change_Incident_Location	Agent_Percept	Fire Incident	HeatReleaseRate (Q)	IncidentLocation
Return	Agent_Perform	Fire Vehicle	StationLocation VehicleSpeed	Location Status
Return_for_Resources	Agent_Perform	Fire Vehicle	ResourceLocation CurrentWaterCapacity VehicleSpeed	Location
Recharge	Agent_Perform	Fire Vehicle	Location ResourceLocation MaximumWaterCapacity	CurrentWaterCapacity
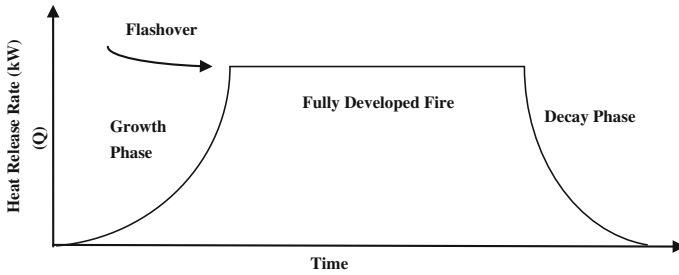
Fig. 1 The typical lifecycle of a fire

Road Segment Resource. The Road Segment is a Resource Entity whose properties are accessed by agents. The Resource Entity does not have any behavior. The properties of Road Segment are Location, Length, Width, and Speed.

4 Implemented Model

Implementation is done in GAMA 1.5.1, a GIS-based Agent modeling and Simulation platform [7]. The platform provides support for GIS operations along with ABMS. As a case study agent-based model is developed for the Allahabad fire emergency service. The geospatial data for Road segments and Fire Station is imported in GAMA, whereas the location of Fire Incident and Fire Emergency Vehicle is initialized and updated during simulation. The various parameters related to FER are initialized at the beginning of the simulation as shown in Fig. 2. During
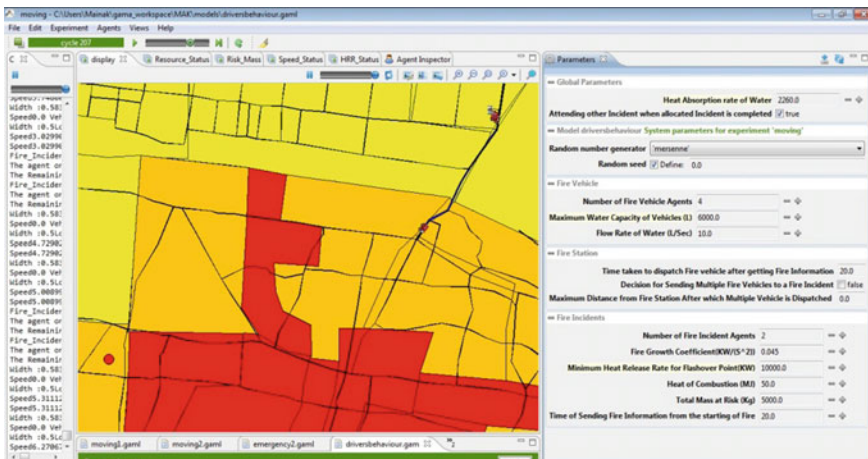
Fig. 2 GIS based agent model for FER and initialization of agent properties
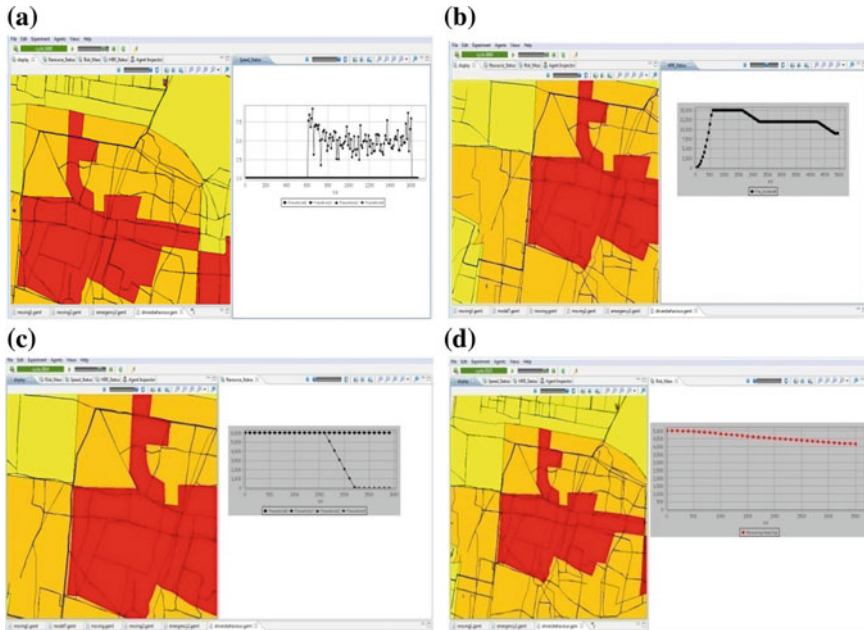
Fig. 3 The status of various agent properties due to interaction between agents during simulation. **a** The speed variation of Fire Emergency Vehicle Agent, provides the departure time and reaching time at the incident location. **b** The variation in Heat Release Rate of Fire Incident due to interaction with Fire Emergency Vehicle Agent. **c** The variation in Water Capacity of Fire Emergency Vehicle Agent due to interaction with Fire Incident Agent. **d** The variation in Risk Mass of Fire Incident Agent

simulation the agents interact with each other and the values of parameters represent dynamics of the agents at a particular instant of time as shown in Fig. 3.

5 Conclusion

This paper considers the FER system as a complex heterogeneous system. To model FER various independent components are identified as Fire Incident, Fire Emergency Vehicle, Fire Station, and Road Segments. To represent the properties and dynamics of each component and interactions between the components, each component is modeled as ABM entities. The locational information and related operations are incorporated in ABM using GIS. The simulation system developed here captures various dynamics of FER and in the future will help to determine the performance of FER.

References

1. Birkin, M., Wu, B.: A review of microsimulation and hybrid agent-based approaches. Agent-based models of geographical systems, Springer Netherlands, 51–68 (2012).
2. Wooldridge, M., Jennings, N.R.: Intelligent agents: Theory and Practice. Knowledge Engineering Review, 10(2), 115–152 (1995).
3. Sterling, L., Taveter, K.: The Art of Agent-Oriented Modeling. The MIT Press Cambridge, Massachusetts London, England (2009).
4. Bandyopadhyay, M., Singh, V.: Formalization of Entities for Agent Based Simulation Using Situation Calculus: A Specific Case Study of Fire Emergency Response. In IEEE *International Conference on Computational Intelligence and Communication Networks (CICN),* 1188–1194 (2014).
5. Bukowski, R.W., Hurley, M. J.: Fire Hazard Analysis Techniques. NFPA, Fire Protection Handbook, Section 3, Chapter 7, 121–134 (2003).
6. Karlsson, B., Quintiere, J. G.: Enclosure fire dynamics (2000).
7. https://code.google.com/p/gama-platform/downloads/detail?name=GAMAv151%20-%20doc.pdf.

Analyzing Link Stability and Throughput in Ad Hoc Network for Ricean Channel by Varying Pause Time

Bineet Kumar Joshi and Bansidhar Joshi

Abstract The growth and usage of wireless communication networks in the last decade is quite significant and when compared with other new technologies, it is vast. The cutting edge of a wireless network is the capability of a wireless node to remain in contact with the whole world while remain mobile. However, the continuous movement of nodes in Ad hoc networks also affects the performance of the network. This paper studies the effects of pause time in the link connectivity and throughput of mobile nodes in the Ricean fading channel. In this work pause time and Ricean constant (k) have been varied to study their effects on performance of mobile nodes connectivity. It has been observed that the throughput and link duration increases with increase in both pause time as well as Ricean constant.

Keywords Mobile ad hoc networks · Pause time · Throughput · Link duration · Ricean channel · Ricean factor K

1 Introduction

A Mobile Ad hoc network (MANET) can be defined as a self-configuring, infrastructure less network of mobile devices not connects by wires [1]. Particularly, central controller is not present; it implies that there is no device which represents a base station in a cellular network. Mobile nodes in MANETs can be located in random way and can move in any direction at any given time. Due to this network's topology and connection between nodes changes unpredictably. Mobility of devices can vary drastically, depending on network's purpose and structure.

B.K. Joshi (✉)
Faculty of Science & Tecnology, ICFAI University, Dehradun 248197, India
e-mail: bineetjoshi@gmail.com

B. Joshi
Department of Computer Science, JIIT University, Noida 201306, India
e-mail: bansidhar11@gmail.com

Normally, high node movements normally results in low link capacity and vice versa. Data exchanged between nodes of ad hoc networks occurs through wireless transceivers.

Due to this, wireless channel model consider as an important building block of Ad hoc networks [2]. The Ricean fading channel model is used when the channels have a strong dominant component among multipath signals [3]. Other than the dominant component, a large numbers of reflected and scattered waves are also received by antenna.

This paper analyzes the efficiency of channel by varying the pause time and Ricean constant. Generally a channel should perform better when nodes moves less; this paper will try to analyze the combine effects of Ricean constant and pause time on the performance of Ad hoc networks.

Rest of the paper is as follows: Section 2 introduces Mobile Ad hoc networks (MANETS). Section 3 describes Ricean channel and performance metrics used to find the channels efficiency. In Sect. 4, different simulation parameters and results obtained are discussed. Conclusions are given in Sect. 5.

2 Mobile Ad Hoc Networks

Mobile Ad hoc technology for dynamic wireless networks has been deployed in military since 1970s and due to the advances in wireless communications, commercial interest in such networks has recently grown. A Mobile Ad hoc network is a collection of wireless mobile nodes that have the capacity of communicating with each other without the use of any centralized administration. [4] Hosts in MANETs are not bound to any central controlling techniques. This offers unlimited mobile movements and total connectivity to the users, but the responsibility of network management lies entirely on the network's nodes. As the nodes of an ad hoc network have limited transmission range, more than one hop may be required for data exchange with another node across the network. Here each mobile node has to act as a host and a router, to forward data packets to other devices present in the network that cannot reached directly. Here every node participates in routing process which allows it to find multi hop routes to all the nodes of the network. This technology of MANET is also known as infrastructure less networking, because mobile nodes automatically establish routes by their own to form network among them self while remaining mobile. This happens instantaneously, and uses multi hop routing method to transmit information (Fig. 1).

Mobile ad hoc networks provide us an efficient way for establishing communications in situations where terrains are not favorable for normal transmission methods, such as army missions, and other critical situations.

Wireless Adhoc Network
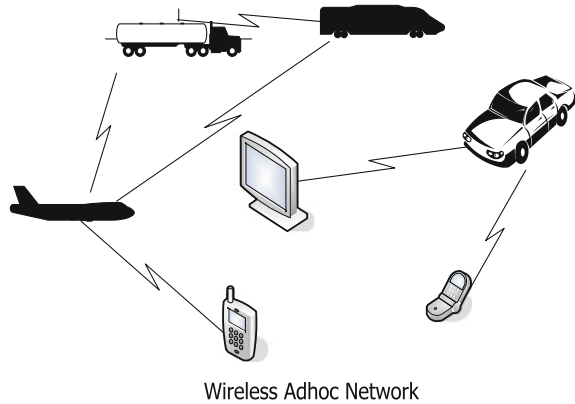
3 Ricean Channel Model and Performance Metrics

3.1 *Ricaen Channel Model*

Mobile devices forming Ad hoc networks contact each other through wireless transceivers and due to this wireless channel models are considered as a crucial ingredient in Ad hoc networks.

Ricean channel model is a small-scale fading or multipath fading model. It can model even a small-scale variation of the radio signal intensity according to a random variable with Ricean distribution [5]. This model predicts variations of the signal intensity over very short distance.

Probability density function for this model can be written as

$$P_{\text{Rice}}(r) \equiv \frac{r}{\sigma^2} e^{-\frac{(r^2 + s^2)}{2\sigma^2}} I_0 \left(\frac{r^2}{\sigma^2} \right) \tag{1}$$

Here, *s* represents the peak amplitude of dominant signal, Io is modified Bessel's function of the first kind and zero-order [6, 7].

It is given as

$$K \equiv 10 \ \log \frac{s^2}{2\sigma^2} \tag{2}$$

K represents the Rice factor [6] and identifies the Rice distribution.

3.2 Performance Parameters

Although various studies have done to check the throughput and link connectivity under various channel model [7, 8]. In none of them, combine effects of pause time and Ricean constant on the throughput and link connectivity is checked. Pause time is the time duration for which all nodes hold the same positions at waypoints. It is the time duration for which mode node halts at a position and them move again to another randomly. Throughput can be defined as the ratio between the total numbers of sent packets versus total number of received packets. Link duration [9], is the average time that a connection between a given pair of devices holds without being disconnected. Stability of the link between nodes is measured efficiently by link.

4 Results

4.1 Simulation Parameters

Due to the nature of an ad hoc network, simulation modeling becomes a valuable mean for understanding the functioning of ad hoc network. The presence of an extensive set of correct implementations of models, protocols and algorithms will dramatically shorten our start-up time, since we will have to worry only about the implementation of our algorithms and of possible modifications/extensions to existing modules While fields testing is important to understand the real efficiency of cellular networks, simulation gives an atmosphere to carry the testing and check the results [10]. Following parameter was chosen for this simulation.

QualNet is used for simulation study. Random Waypoint mobility model is used for node movements [11]. Time for simulation was 120 s and the scenarios had been formed in 30 m × 30 m area. In Random Waypoint nodes may be initially put in location of the testing field, and then they can move in any direction by choosing direction between [0, 2∏]. Their movement continues fix time and this is repeated for predefined time. Pause time has been increases with a step size of 10 s. Data rate was 2 Mbps. Number of devices remains constant during the experiment. Wireless channel examined was Ricean channel model with different Rice constant. Value of K was changed from 4 to 7 increasing one at a time. Dynamic source routing protocol was used for the present work.

4.2 Result Analysis

Figure 2 clearly shows that throughput increases steadily as pause time is increased from 10 to 50 s for all four values of Ricean constant K, i.e., from 4 to 7. This is

Fig. 2 Throughput of mobile
nodes in Ricean
communication channel
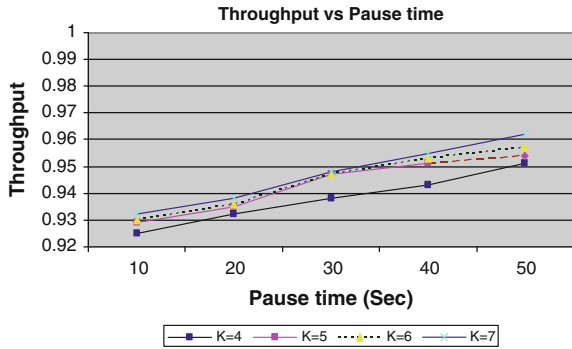
Fig. 3 Link duration of
mobile nodes in Ricean
communication channel
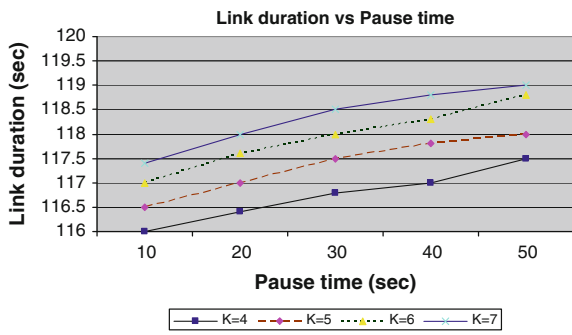
because as the pause time is increases mobility of nodes deceases and they received
better signal power and thus throughput increases. The throughput also increases as
the value of Ricean constant increases. It happened because Ricean constant value
(*K*) is directly preoperational to line of sight (dominant) component of signal, so
increase in value of *K* leads to increase in value of line of sight. It means signals are
correctly received; hence, receiver received signals greater than the threshold value
most of the times.

Figure 2 shows that the Link duration increases constantly as the pause time
increases, this is expected as the value of pause time increases nodes remains
stationary for more duration of time period, i.e., from 10 to 50 s and they receive
greater signal strength, which leads to greater link duration (Fig. 3).

5 Conclusion

By analyzing the result, we can say that with the increase in pause time, both
throughput and link duration for the Ricean channel has been increases. With the
increase in pause time, random movement of mobile nodes decreases. Due to this
nodes stays at one position for more duration, this reduces the fading effects and

allow more packets to be reached destination correctly and hence throughput increases. Similarly less movement of mobile nodes causes greater link duration. Continuous increase in Ricean constant (K) results in continuous increase in presence of dominant signal component which also helps in increasing both throughput and link duration. In future, we would like to analyze the effect pause time in other channel models.

References

1. M. Frodigh, P. Johansson, and P. Larsson; Wireless Ad Hoc Networking: The Art of Networking Without a Network, Ericsson Review, 4, 248–263 (2000).
2. H.L.Bertoni; Radio Propagation for Modern Wireless Systems, Prentice Hall PTR, NJ (2000).
3. R. Agarwal and J. Cioffi; Capacity of Fading Broadcast Channels with Limited-Rate Feedback, Forty-Fourth Annual Allerton Conference, pp 336–344, UIUC, Illinois, Sept 27–29, (2006).
4. Charles E. Perkins; Ad hoc Networking, Addison-Wedey (2001).
5. J. Mullar, and H. Huang; Impact of Mutipath Fading in Wireless Ad Hoc Network, Pe-Wasun'05, October 10–13, Montreal, Canada (2005).
6. T. S. Rappaport; Wireless Communications; Principles and Practice, Prentice-Hall (1996).
7. B. Joshi and N. Mishra, "Link Analysis of Mobile Nodes in Ad Hoc Network Using Ricean Fading Channel Model, I. J. Mod. and Opt 3, 84–86, (2013).
8. B. Joshi; A Comprehensive Study of Performance of MANET in Ricean Channel Model, An Analytical Approach, LAP Lambert Academic Publishing (2013).
9. S. Xu, K. Blackmore, and H. Jones; Movement and Linkage Analysis of Mobile Nodes in Ad Hoc Networks, Infocom, pp. 201–204, (2005).
10. L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, and H. Yu; Advances in network simulation, IEEE Computer, 33, 59–67 (2000).
11. H. Bai, H. Aerospace, and M. Atiquizzaman; Error Modeling Scheme For Fading Channels in a Wireless Communication: A Survey, IEEE Com. Surveys, 5, no. 2 (2003).

A Review on Pixel-Based Binarization of Gray Images

Ankit Shrivastava and Devesh Kumar Srivastava

Abstract This paper presents a review study on binarization of gray images. Binarization is a technique by which an image is converted into bits. It is an important step in most document image analysis systems. Since a digital image is a set of pixels. Many binarization techniques have a definite intensity value for each pixel. A gray image is just an image which has each pixel of same intensity. That means there is not much difference in color or value information of pixels. Usually, a picture in black and white is considered as gray image in which black has least intensity and white have highest.

Keywords Binarization · Images · Restoration · Segmentation, etc

1 Introduction

For last few years there has been a continued and significant work in the area of gray image binarization. Some techniques use bi-level images which decrease the processing cost of gray-level image information. Binary form is the most common form of data storage [1]. Data in the form of text is extremely rare. Binary files especially images are used for several reasons. Reaction time for binary data is quite less as compared to data in normal character form. Conversion of a larger integer requires some time although the time taken in one conversion is not significant but the accumulated time of multiple. Conversions might result in very slow processing [2]. Applications such as computer games which require a big amount of data to be processed in real time would come to a standstill if there is no binarization of data. Also, the size of a binarized image is much smaller than the actual image, hence

A. Shrivastava (✉) · D.K. Srivastava
Department of Computer Science & Engineering, SCIT, Manipal University Jaipur,
Jaipur, Rajasthan, India
e-mail: ankitshrivastava82@gmail.com

D.K. Srivastava
e-mail: devesh988@yahoo.com
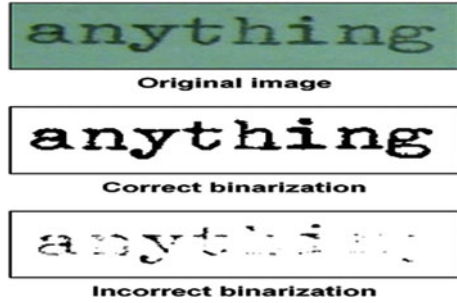
lesser requirement of storage space. Also the processing time of smaller image is faster. For example, a DVD would not be able to hold a movie if it were to be stored in its original form, hence it is stored in binary form [3] (Fig. 1).

2 Literature Review

A comprehensive review on analysis of image binarization for degraded documents. The research is based on the image binarization of degraded documents. It is challenging to successfully improve the quality of such documents because of the variation in the foreground and background text of different document images. The various parameters that are involved when studying the features of these document images are: stroke brightness, stroke connection, stroke width, and document background [4]. The various steps involved in the binarization of a documented image are preprocessing, contrast image construction, edge detection, text stroke edge pixel detection (Otsu's method/Canny's Edge Detector), local threshold estimation, and the state of the art method. The next step after binarization of an image is its restoration. Restoration is done by degrading the image using prior knowledge of the degraded phenomenon. Once degraded, the inverse is applied to the degraded image in order to extract the original image. Adaptive degraded document image binarization, unlike the standard techniques, is a newer approach to the binarization and improvement of degraded document images [5]. The given method does not utilize any prior parameter tuning by the user and can deal with degraded images which have been degraded either due to shadows, nonuniform illumination, low contrast, or large signal-dependent noise, or by smear and strain. The steps involved in restoring such images are: a preprocessing technique using a low pass filter; an approximation of foreground areas; a background area calculation by interpolating adjacent background intensities, and a threshold calculation by combining the calculated background surface [6]. The above step is done while incorporating image up sampling. A postprocessing step is used to improve and preserve the stroke connectivity and the quality of text regions. This study deals with the highly challenging task of segmenting text from badly degraded document

Fig. 2 Primary local window
and secondary local window
of Feng's technique of
binarization
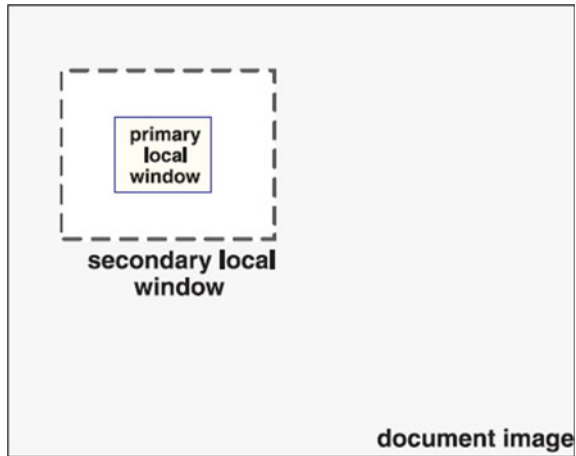
images due the very high variation in the background and foreground text of various document images. It is done using a technique known as adaptive image contrast. It is an amalgamation of the local image gradient and the local image contrast which in turn is tolerant to background and text variations caused due to the degradations [7]. The steps undertaken in this technique are: construction of adaptive contrast map for an input of degraded document image. Now follows the binarization of the contrast map and is then pooled with Canny's Edge map in order to identify the text stroke edge pixels. Segmentation of the document text by a local threshold which is estimated based on the intensities of spotted text stroke edge pixels in a localized window. The proposed method does not involve any complexities and hence is simple and robust. The following study does not focus on any one method for the image segmentation but presents a depth analysis of multiple methods that are optimum for the task stated. In a camera-based recognition system, there are generally three steps that are involved; target region detection; character segmentation; and character restoration. The work mainly focuses on the third step which is character restoration [8]. Multiple methods involving character restoration are used (Fig. 2).

3 Techniques for Binarization

We are in the process of studying and analyzing some techniques of image binarization, focusing on Feng's Technique, T.R. Singh's Method, and Eikvil Technique and we are also trying to design and develop a new technique for better results. There are different parameters on which any binarization technique can be considered as efficient. In many cases, parameters for quality, time, and efficiency are taken into account. Time taken in process of binarization of an image plays an

important role. On the basis of time taken in process, we can classify the different techniques. The technique which takes the least amount of time is considered as an optimal technique [9]. Here, it will be a motive to reduce the time taken in technique to be evolved. The target of this project is to develop an algorithm for binarization of grayscale image which is efficient in quality and also in matter of time. Beside this, to compare the sample obtained image after binarization using three-mentioned techniques and to find out comparisons. This work primarily includes the study of binarization of gray images done using three techniques. They are: Feng's Technique, T.R. Singh Technique, and Eikvil Technique. Feng technique mainly focuses on the binarization of grayscale document images, taking input from mainly PDA's, mobile phones, faxing machines, auto notes taking, and OCRs. The proposed method reduces the common problems of low-quality document images, such as nonuniform illumination, undesirable shadows, and random noise. T.R. Singh proposed a method of binarizing images by a method called automatic binarization which involves changing a gray image into a binary image without standard deviation which is used in Sauvola's technique and Niblack's [10]. This technique is expressed as

$$T(x, y) = m(x, y)\left[1 + k\left(\frac{\partial(x, y)}{1 - \partial(x, y)} - 1\right)\right]$$

where $\delta(x, y) = I(x, y)—m(x, y)$ is the local mean deviation and k belongs to [0, 1]. The higher value of k makes the threshold value lower and vice versa. It uses the local mean. In the Eikvil technique, a page is a collection of components as text, images, and background. The hitches caused by brightness, noise, and many other sources are discussed. Two novel techniques define a local threshold for each and every pixel. The performance calculation of the technique uses sample images, calculation metrics for binarization of documented images, and a weight-based grade procedure for final result performance.

4   Methodology for Binarization

Binarization, which converts grayscale images into bits using two stages [11]. This is because of two-level information which greatly reduces the processing load and the algorithm complexity. Binarization is a complex stage, since any error in this stage will pass on to the successive ones. An important binarization approach is a gray-value thresholding and this technique could be further categorized into local and global thresholding. Each and every pixel of the input image is compared with the threshold value and then the pixels are divided into two categories "foreground and background." The threshold is a major parameter in binarization and selection of a threshold value is the main task [12].

A threshold $T(x, y)$ is a value such that

$$b(x,y) = \begin{cases} 0 & if \ I(x,y) \leq T(x,y) \\ 1 & otherwise \end{cases}$$

where $b(x, y)$ is binarized image and $I(x, y)$ could be 0 and 1 depending upon the intensity of a pixel (x, y) of the image I.

In this paper, we only focus on binarization of grayscale document images with the help of local thresholding technique, generally color documents do not lose much information when converted to grayscale images in context with the differentiation between page foreground and background. Severely degraded or noisy images can be improved drastically using these techniques but unfortunately is slow due to the computational complexity of min, max, and mean from the local region of each pixel in the image [13]. The differences can be seen in Figs. 3 and 4. The project is divided into basically three steps. Basically, here, the whole process is categorized into three major parts. They are preprocessing, processing, and binarization which can be seen in Fig. 5.

In the pre-processing stage, there are several substeps. First of all steps, which comes in the preprocessing step is selecting original image to be tested. That image is then digitally acquired in the second step of the preprocessing stage, i.e., the image acquisition step. Third step is image optimization. This step of re-processing is further divided into several small processes such as grayscale conversion, contrast

Fig. 3 Original images

Fig. 4 Binarized images

HAGWANT DAS, THE THEN RULER OF
AND LATER ON IT WAS MADE CAPITAL
DHO SINGH, BROTHER OF RAJA MAN
MADHO SINGH WAS DIWAN IN THE
IF MUGHAL EMPEROR AKBAR (A.D 1556-1605)
MAINS AT BHANGARH CONSIST OF FORT
TEWAYS, MARKETS. HAVELIS. TEMPLES,
CE, CHHATRIS, TOMB, ETC IMPORTANT
ARE GOPINATH, SOMESVARA, KESHAV RAI
GLA DEVI, ALL IN NAGARA STYLE OF TEMPLE
TURE. THE ROYAL PALACE IS SAID TO HAVE
TOREYS BUT NOW ONLY FOUR STOREYS
WHOLE TOWNSHIP WAS PROTECTED BY
CCESSIVE FORTIFICATIONS. THE OUTERMOST
TION IS PROVIDED WITH FIVE GATEWAYS
RTH TO SOUTH- AJMERI, LAHORI, HANUMAN
RI AND DELHI GATE.

Fig. 5 Steps of overall process
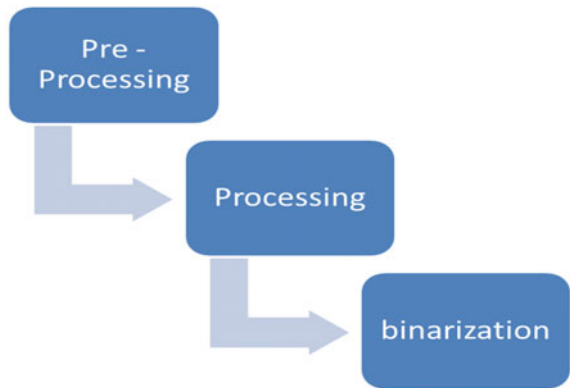
optimization, smoothing, refining edges, etc. [14]. After image optimization, the next step is histogram equalization. In this step, image is optimized by using the histogram feature. The fifth step is obtaining the sample. After this step, the processing stage starts. The process is shown in Fig. 6.

Fig. 6 Steps involved in preprocessing stage
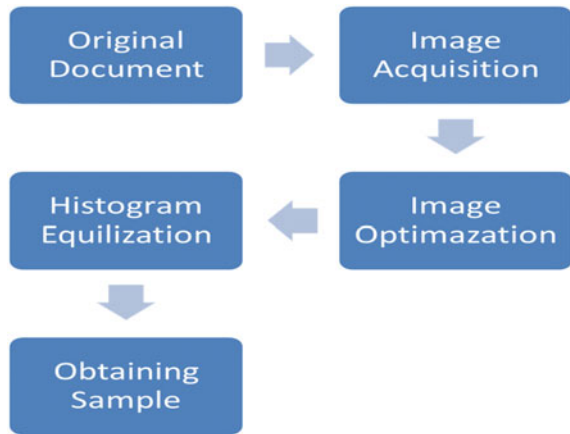
5 Conclusion and Future Scope

This study has revealed that document analysis, conversion, understanding, and management in general are very complex issues and still very far from having global, uniform, and clear-cut solutions. Any algorithm does not work well for most types of images but some work suggested that better performance could be gained by choosing some efficient algorithms for document images. We have explained algorithms that have local features and works on the ghost mechanism. Many algorithms need some preprocessing steps to work with the document image. By applying this algorithm which is based on local region, Otsu's technique and watershed model, results obtained are without ghost; so, further iterations and additions with other techniques can be derived out with better results in future. Binarization of images is a very progressive field today and in the future. There are a lot of improvements going in the field of binarization. Future is waiting for multidimensional imaging media. For that, it is required to convert images and document images in binarized form. Documents and images which have nonuniform brightness, needs local threshold on the small level. The local threshold varies according to various image conditions.

References

1. Meng-Ling Feng and Yap-Peng Tan, "Contrast Adaptive Binarization Of Low Quality Document Images" The Institute of Electronics, Information and Communication Engineers (IEICE) Electronic express, Volume 1, Issue No. 16, November 2004, Page 501–506.
2. Chien-Hsing Chou, Wen-HsiungLin and FuChang, "A binarization method with learning-built rules for document images produced by cameras", Pattern Recognition 43, 2010, 1518–1530.

3. P. Subashini and N. Sridevi, "An Optimal Binarization Algorithm Based on Particle Swarm Optimization", International Journal of Soft Computing and Engineering (IJSCE), Volume-1, Issue-4, September 2011.

4. S.S. Bedi and Rati Khandelwal, "International Journal of Advanced Research in Computer and Communication Engineering", International Journal of Soft Computing and Engineering (IJSCE), Vol. 2, Issue 3, March 2013.

5. O. Imocha Singh, O. James, Tejmani Sinam and T. Romen Singh, "Local Contrast and Mean based Thresholding Technique in Image Binarization", International Journal of Computer Applications, Volume 51– No. 6, August 2012.

6. Rajesh K. Bawa and Ganesh K. Sethi, "A Binarization Technique For Extraction Of Devanagari Text From Camera Based Images", Signal & Image Processing: An International Journal (SIPIJ), Vol. 5, No. 2, April 2014.

7. Youngwoo Yoon, Kyu-Dae Ban, Hosub Yoon, Jaeyeon Lee and Jaehong Kim, "Best Combination of Binarization Methods for License Plate Character Segmentation", Electronics and Telecommunications Research Institute (ETRI) Journal, Volume 35, Number 3, June 2013.

8. Ntogas, Nikolaos, Ventzas, Dimitrios, "A Binarization Algorithm For Historical Manuscripts", 12th WSEAS International Conference on COMMUNICATIONS, Heraklion, Greece, July 23–25, 2008.

9. Geetanjali Thakur, "A Comprehensive Review On Analysis Of Image Binarization For Degraded Documents, International Journal of Advance Research In Science And Engineering (IJARSE), Vol. No.3, Issue No.7, July 2014 ISSN-2319-8354(E), Page 325.

10. Bolan Su, Shijian Lu, and Chew Lim Tan, "Robust Document Image Binarization Technique for Degraded Document Images", IEEE Transactions On Image Processing, Vol. 22, No. 4, April 2013.

11. Aroop Mukherjee and Soumen Kanrar, "Enhancement of Image Resolution by Binarization", International Journal of Computer Applications, Volume 10– No. 10, November 2010.

12. B. Gatos, I. Pratikakis and S.J. Perantonis, "Adaptive degraded document image binarization", Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research "Demokritos", 153 10 Athens, Greece, September 2005.

13. Chirag Patel, Dr. Atul Patel and Dr. Dipti Shah, "Threshold Based Image BinarizationTechnique for Number Plate Segmentation", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, July 2013, Page 108–114.

14. T. Romen Singh, Sudipta Roy and Kh. Manglem Singh, "Histogram Domain Adaptive Power Law Applications in Image Enhancement Technique", International Journal of Computer Science and Information Technologies(IJCSIT), Vol. 5, Issue 3, 2014 0.

Mobility- and Energy-Conscious Clustering Protocol for Wireless Networks

Abhinav Singh and Awadhesh Kumar Singh

Abstract In this paper, we present a distributed clustering protocol for mobile wireless sensor networks. A large majority of research in clustering and routing algorithms for WSNs assume a static network and hence are rendered inefficient in cases of highly mobile sensor networks, which is an aspect addressed here. MECP is an energy-efficient, mobility-aware protocol and utilizes information about the movement of sensor nodes and residual energy as attributes in network formation. It also provides a mechanism for fault tolerance to decrease packet data loss in case of cluster head failures.

Keywords Wireless sensor network · Clustering · Mobility · Energy efficient · Distributed

1 Introduction

Wireless sensor networks (WSNs) consist of a large number of sensor nodes that are densely deployed in a region of interest and connected through wireless links to collect data about a target or event, and cater a variety of sensing and monitoring applications [1]. In many applications that belong to marine environments, wildlife tracking and protection, and various other such activities, the mobile sensors are more effective as compared to their static counterparts. However, the sensors being energy-constrained nodes, the mobile sensors are more prone to crash due to battery exhaustion as mobility incurs more computation overhead and thus it is energy intensive. Hence, the traditional WSN protocols [2, 3] are not suitable for

A. Singh (✉) · A.K. Singh
Department of Computer Engineering, National Institute of Technology,
Kurukshetra, Haryana, India
e-mail: abhinavsingh282@gmail.com

A.K. Singh
e-mail: aksingh@nitkkr.ac.in

deployment in the environment where sensors are mobile. Therefore, the mobility-aware protocols are the more preferred option.

In the literature, a number of approaches have been proposed to handle mobility in the wireless scenario [4, 5]. Clustering is a popular approach to handle mobility and improve scalability in distributed computing systems [6, 7]. In clustering, the network is partitioned into nonoverlapped regions and the activities of nodes belonging to each cluster are coordinated by a distinct node called a cluster head (CH). Though, CH is responsible for efficient communication and data dissemination, its failure may lead to the discontinuity of the application. Therefore, for long-running applications, fault tolerance is an utmost desirable feature. Second, the movement of cluster head may also impact the application in a similar way. Therefore, in this paper, we propose an energy-efficient distributed clustering protocol that is fault tolerant and also handles mobility in WSNs.

2   System Model

The WSN consists of a unique set of N nodes connected with wireless links. The nodes which are used in the communication range of a node N_i are called the neighbors of N_i. The nodes which are assumed to be mobile and can trigger the topological changes. It is also assumed that each sensor node is capable and compatible of sensing its velocity, e.g., via an accelerometer or any other sensing hardware embedded in it.

3   Related Work

A large number of clustering protocols for sensor nodes have been proposed in the literature. However, due to limitation of space, some representative protocols are being reviewed in the following paragraphs.

The LEACH-Mobile protocol proposed by Kim Do-Seong and Yeong-Jee Chung [8] supports sensor nodes mobility in WSN by adding membership declaration to LEACH protocol. The idea behind this membership declaration is to confirm the inclusion of sensor nodes in a specific cluster during the steady-state phase. The CH sends "data request" message to its members, and receives the data sent back from them. The SN with minimum mobility is elected as the cluster head. The LEACH-Mobile outperforms LEACH in terms of packet loss in mobility centric environment. However, it is not traffic and mobility adaptive protocol and there is higher energy waste in idle listening and overhearing of this protocol [9].

The HEED protocol [10] is a distributed clustering protocol for long-lived ad hoc sensor network in which the main parameter for the cluster head selection is residual energy levels and leads to a prolonged life of nodes.

4 The MECP Concept

4.1 Clustering Assumptions

In our protocol, each node takes decisions on the basis of two parameters, namely residual energy and relative velocity with respect to its neighbors. The intracluster communication costs are also considered for the clustering process in order to increase the efficiency of energy consumption. For example, cost can be a function of distance from CH or node density of cluster. Generally, a node has several transmission power levels, where the higher power levels can cover greater distances for transmission. We reserve the lower power levels for intracluster communication between a normal node and a CH to reduce communication costs. Further, we reserve the higher transmission power levels for intercluster communication. Because intercluster communication with low power levels may lead to link failure or may render unidirectional links. If the power level used for intracluster communication is fixed, then the cost of communication for a node can be determined by the node degree (D). For instance, to create dense clusters, the cost of communication with a CH is set proportional to 1/D so that dense clusters have lower cost of communication. Alternatively, for load-balanced clustering, the cost of communication is set proportional to D so that dense clusters have a higher cost of communication. Consequently, a new node would prefer to join a cluster where the cost of communication is lower. In case, multiple power levels are allowed for intracluster communication, we define a cost factor (communication cost factor) CCF where

$$CCF = \frac{\sum_{i=1}^{m} minPower_i}{m} \tag{1}$$

where, $minPower_i$ is the minimum required power level for communication between a node and its ith neighbor. We also define a velocity factor (VF) as

$$VF = \begin{cases} \frac{1}{Va}, & Va \geq 1 \\ 1, & Va < 1 \end{cases} \tag{2}$$

Va is the average of relative velocities between node and its neighbors defined as.

$$Va = \frac{\sum_{i=1}^{m} \text{Relative velocity of Node and neighbour } i}{m} \tag{3}$$

4.2   The Working of MECP

The protocol is executed at every node and requires X number of iterations denoted by X_{it}. The probability of a node becoming a cluster head is denoted by CH_{prob}

$$CH_{prob} = K \times \frac{E_{res}}{E_{max}} \times VF \qquad (4)$$

where K is the percentage of nodes that become cluster heads (e.g., 10 %) initially. E_{res} is the estimated amount of energy remaining, and E_{max}, the maximum energy stored by the battery. The CH_{prob} is restricted in the range $[P_{min}, 1]$ to allow efficient termination of protocol and will be explained soon. The protocol introduces fault tolerance by allowing each CH to select an Assistant CH (ACH) from within its cluster. In case of CH failure, due to reasons like physical damage, depletion of energy, or movement of CH out of communication range, etc., the cluster members suffer data packet loss. After timeout, the ACH assumes the role of CH. In such case, the cluster members resend the data packet to the new CH. The new CH possesses all the updated routing information that the previous CH had. Therefore, it successfully avoids any further loss of data packet and renders application continuity.

The initialization phase starts with initializing the values of L_{adj} using neighbor discovery (Fig. 1). Afterward, each node broadcasts its cost and velocity to its neighbors and calculates its CH_{prob} (Fig. 2).

Fig. 1 Neighbor discovery
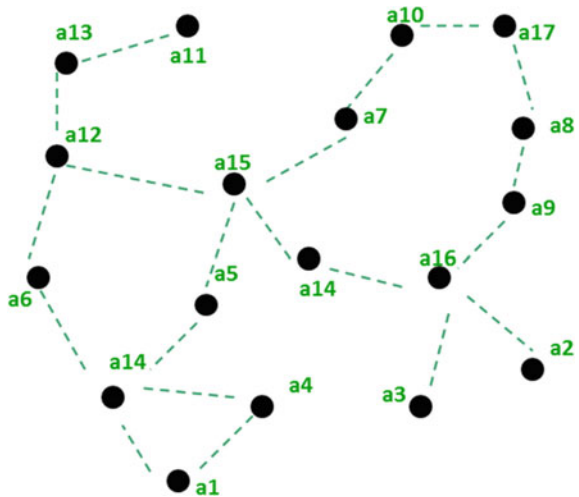
Fig. 2 Compute CH$_{prob}$ and cost
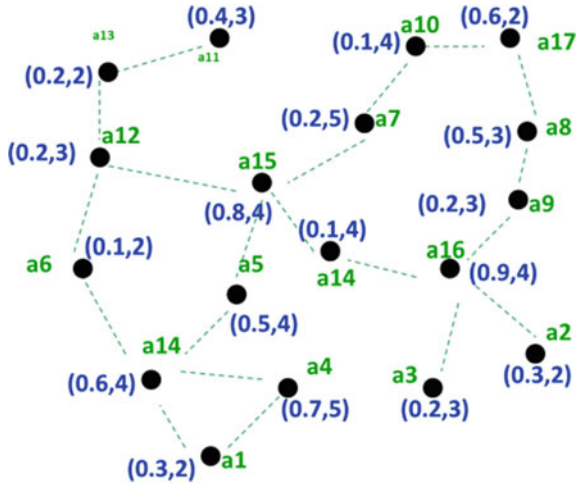
Fig. 3 Cluster head election
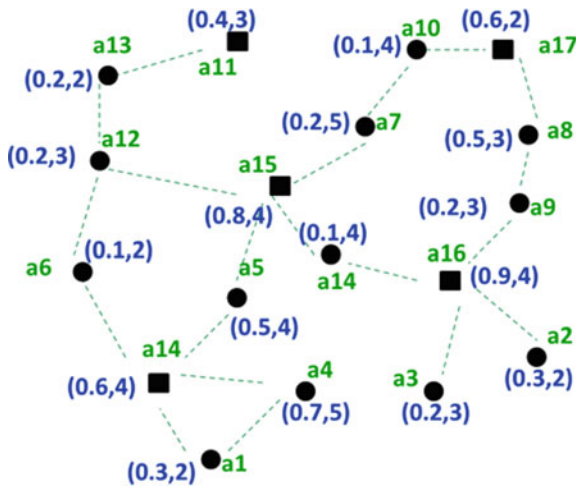
In the second phase, every node searches for a CH or a tentative CH and in the absence of one, the nodes become a tentative CH with a probability CH$_{prob}$. At the end of each iteration, the value of CH$_{prob}$ doubles and the process continues till the value of CH$_{prob}$ becomes ≥ 1. In the third phase, a node, either becomes a Final CH and broadcasts *Declare_FinalCH* message to its neighbors or it becomes a member of a cluster (Figs. 3 and 4).

Fig. 4 Normal nodes select their CH
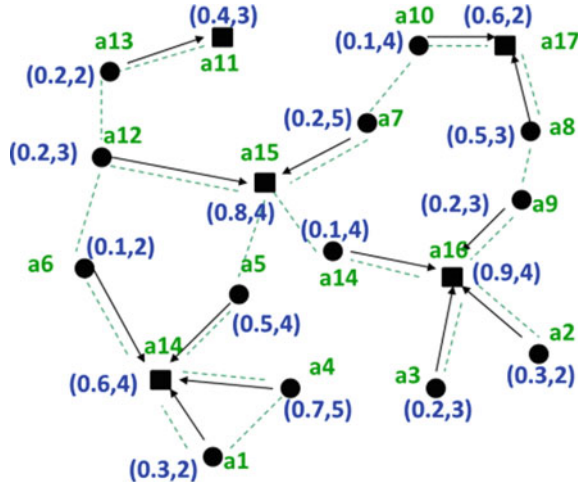
Fig. 5 Selection of ACH by each CH
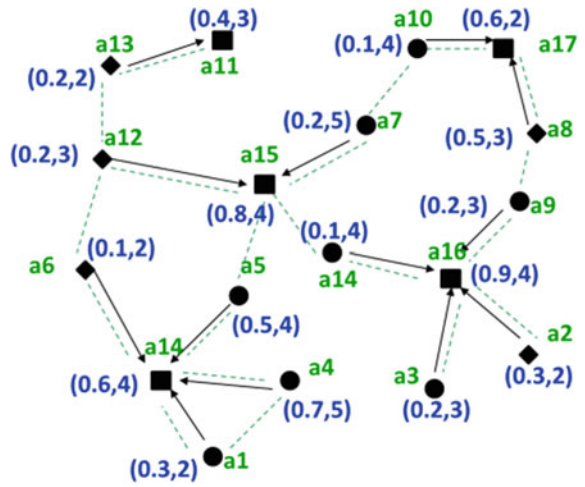
In the last phase, each CH selects its ACH and broadcasts the ACH ID to its member nodes (Fig. 5). In case of failure of CH during protocol operation, ACH takes charge of CH. In case, a node moves out of its cluster during protocol operation, it requests to join another cluster in its transmission range.

The protocol is carried out as explained by the pseudocode below.

4.3   Pseudo Code

Phase I   Initialization
```
1)  L_adj ← Add all nodes in communication range to list
2)  Compute Cost of communication for each node in L_adj  and
    broadcast cost, my_velocity
```
3) $CH_{prob} = K \times \dfrac{Eres}{Emax} \times VF$ //Initialise value of CH_{prob}
```
4)  Set final_CH = False
5)  Set Assistant_CH = False
```

Phase II
```
    Repeat
1)  L_CH ←All nodes in L_adj which are final CH or Tentative CH
2)  If L_CH  ≠ Φ                         // List is not empty
    {       my_CH = least_cost(L_CH)      //select least cost node as CH
            If(my_ch = =node_ID)
            {       If(CH_prob = 1)
                    {       Declare_FinalCH(NodeID,Cost)
                            Set final_CH ← True}
                    Else
                    {       Declare_Tentative_CH(NodeID,Cost) }
            }
    }
    ElseIf (CH_prob = = 1)
    {       Set my_CH = nodeID
            Declare_FinalCH(NodeID,Cost)
            Set Final_CH = true}
    Elseif (Random (0,1) ≤ CH_prob)
    {       Declare_Tentative_CH(NodeID,Cost) }

3)  CH_prev = CH_prob
4)  CH_prob = min (CH_prob *2, 1)
5)  Until CH_prev = 1
```

Phase III   Termination
```
    1)  If (final_CH = = False )
```

```
        {   L_FCH ← All Final Cluster Heads in L_CH
            1a)      If(L_FCH ≠Φ̄)
            {        My_CH = least_cost(L_FCH)
                     Declare_join(Cluster_head_ID,Node_ID,cost) }
    1b)     Else
            {        Set Final_CH = True
                     L_Members  ← All nodes that have elected to
                     join my cluster.
                     Declare_FinalCH(NodeID,Cost) }
        }

    2) Else

        {
            Set Final_CH = True
            L_Members  ← All nodes that have elected to join my
cluster.
            Declare_FinalCH(NodeID,Cost)
        }

    proc recieve_message (recieve declaration of formation of
final CH or tentative CH)
    {    Lch<-- save  nodeID along with identifier final or
         tentative }

    After completion of clustering,
    Proc select_ACH(NodeID,cost)
    {        if(node is a CH)
             {L_Members  ← All nodes that have elected to join my
                    cluster.
             My_ach= least_cost(L_Members )
             Declare_ACH(NodeID of ACH,cost);}

             Else
             {proc recieve_message  (recieve  declaration  of
             formation of final CH or tentative CH)}
    }
```

Phase IV During T_p

```
    If(packet data is lost when sending data to CH)
    {
1           Resend data to ACH
2           If(packet data is lost when sending data to
            ACH)
3           {
4           L_FCH ← All CHs in neighbourhood
5           My_CH= least_cost(L_FCH)
6           Declare_join(Cluster_head_ID,Node_ID,cost)
7           My_ACH ← Receive ID of ACH node from CH
8           }

    }
    proc recieve_message (declaration of new ACH by CH )
    {   my_ACH = NodeID of ACH recieved;}
```

4.4 Proof of Correctness

Lemma i *MECP terminates in $X_{iter} = O(1)$ iterations*

Proof The worst-case scenario is when a node has very low residual energy and very high Va. In this case, the CH_{prob} is equal to P_{min}. Now since the CH_{prob} doubles every iteration and terminates when it finally reaches a value ≥ 1, therefore we have

$$P_{min} * 2^{X_{iter}-1} \geq 1 \quad \text{and} \quad X_{iter} \leq \left\lceil \log_2 \frac{1}{P_{min}} \right\rceil + 1$$

Therefore, $X_{iter} \approx O(1)$.

Since essentially this translates into $\leq \left\lceil \log_2 \frac{1}{CH_{prob}} \right\rceil + 1$, a node with higher E_{res} and lower Va will terminate its MECP execution much faster than other nodes and hence this will allow low energy or highly mobile nodes to join its cluster.

Lemma ii *Any node that wishes to join the WSN will do so by the end of MECP.*

Proof Let us assume that the node is not a part of the WSN and hence is not a CH or a regular node by the termination of MECP. This means line 1 of Phase III is satisfied while 1a is not, this means that 1b shall execute and the node becomes a CH which is a contradiction.

Lemma iii *No node can be a part of more than one cluster by the end of MECP.*

Proof Let us assume a node is part of two clusters. This means that line 1 and 1a of phase III must have been executed after which the node becomes part of one cluster and end MECP execution. There is no provision for it to execute lines 1 and 1a again in the same cycle which contradicts our assumption.

5 Intercluster Communication

The CHs aggregate data over a round and send the aggregated data in that round in single transmission. However, in certain situations, the need of intercluster communication may arise. The intercluster communication in MECP is multihop through various CHs. Nevertheless, in the mid of intercluster communication, a CH may move beyond the range of transmission of other CHs incurring packet data loss because it does not reach the sink and the data for the entire round from a particular cluster can be lost. Though, an appropriate recovery mechanism may help in recovery of lost data, it can also be avoided by using gateway nodes that are also called guard node as proposed in DEMC [11] protocol. Guard nodes are intermediate nodes that help in transmission of data, in case, two CHs are not within the transmission range of each other.

6 Conclusion

We presented a multiphase distributed clustering protocol that is energy efficient and also effectively handles mobility of nodes during application execution. Out of the three phases of clustering, only the second phase may involve multiple iterations nevertheless that are bounded by parameter $P_{\min}$. This feature makes the process of reclustering lightweight in addition to significant reduction in the latency involved in clustering. Furthermore, the clustering-related decisions in MECP are based primarily on local information and therefore MECP suffers limited message overhead.

References

1. Misra, S., Reisslein, M., Xue, G.: A survey on multimedia streaming in wireless sensor networks. IEEE Commun. Surv. Tutor., Vol. 10, Iss. 4, pp. 18–39 (2008).
2. Yuan, Y., Chen, M., Kwon, T.:A novel cluster-based cooperative MIMO scheme for multi-hop wireless sensor networks. EURASIP J. Wirel. Commun. Netw., Vol. 2, pp. 38–47 (2006).
3. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.:An application specific protocol architecture for wireless micro sensor networks. IEEE Trans. Wirel. Commun., Vol. 1, Iss. 4, pp. 660–670 (2002).
4. Karim, L., Nasser, N.: Energy Efficient and Fault Tolerant Routing Protocol for Mobile Sensor Network. Proc. of IEEE ICC, pp. 1–5 (2011).
5. Nasser, N., Al-Yatama, A., Saleh, K.: Zone based routing protocol with mobility consideration for wireless sensor networks. Telecommunication Systems, Springer, Vol. 52, Iss. 4, pp. 2541–2560 (2012).
6. Abbasia A.A., Younis M.: A survey on clustering algorithms for wireless sensor networks. Computer Communications, Network Coverage and Routing Schemes for Wireless Sensor Networks, Volume 30, Issues 14–15, pp. 2697–2994 (2007).
7. Liu X.: A Survey on Clustering Routing Protocols in Wireless Sensor Networks. Sensors (14248220), Vol. 12, Iss. 8, Pp. 11113–11153 (2012).
8. Kim, D.S., Chung, Y.J. :Self-organization routing protocol supporting mobile nodes for wireless sensor network. Proc. First Int. Multi-Symp. on Computer and Computational Sciences, Hangzhou, China, pp. 622–626, (2006).
9. Lofty S., Padmavati : A Survey on Mobility Based Protocols in WSNs. Proc. of Int. Conf. on Computing, Communication & Manufacturing, pp 12–15, (2014).
10. Younis, O., Fahmy, S.: Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy Efficient Approach. IEEE Transactions on Mobile Computing, Vol. 3, Iss. 4, pg 366–379 (2004).
11. Ali S., Madani, S.: Distributed Efficient Multi Hop Clustering Protocol for Mobile Sensor Networks. The International Arab Journal of Information technology, Vol. 8, Iss. 3, pp 302–309 (2011).

RE^{2}R—Reliable Energy Efficient Routing for UWSNs

Khan Gulista, Gola Kamal Kumar and Rathore Rahul

Abstract Wireless sensor networks are used in each and every area of our daily life. It helps us in our life in a straightforward and cost-effective manner. WSNs are usually deployed at unreachable places to collect information about surroundings. As we know three-fourths of the earth is covered with water. So, for information underwater, we need underwater sensor networks. Compared to earthy networks underwater, sensor networks have several limitations of low bandwidth, high propagation delay, and low transmission power. Somewhat similar to earthy sensor nodes, a crucial problem with UWSNs is finding an efficient route between a source and a destination. We reviewed several routing algorithms based on location and locationless algorithms and protocols and we came to know of the several challenges existing in designing energy efficient routing protocols. Considering the several challenges overall, we tried to find a reliable and energy efficient routing protocol for under water sensor networks. Through simulation study using NS-2 simulator, we showed a significant improvement in terms of data delivery ratio, node dead ratio, and energy consumption.

Keywords UWSNs · Routing · Energy consumption · Dijkstra algorithm · Network lifetime

Khan Gulista (✉) · G.K. Kumar · Rathore Rahul
Department of Computer Science and Engineering, Teerthanker Mahaveer University, Moradabad, India
e-mail: gulista.khan@gmail.com

G.K. Kumar
e-mail: kkgolaa1503@gmail.com

Rathore Rahul
e-mail: rahulrathore1512@gmail.com

1 Introduction

WSNs generally work on the principle of multipath propagation. An on-board processor is fitted in the sensor nodes, which helps to make nodes intelligent, so that continuous communication can take place even in the worse environmental situation. Every node in the sensor network is responsible for the fusion of data. The amount of data transmitted between sensor nodes and the base station can be reduced in the process of data fusion. Nowadays underwater sensing is one of the most important areas of sensor networks. Underwater sensing comprises detection of the underwater environment, such as underwater flora, monitoring marine creature, and chemical waste. It is differing from earthly WSNs in the context of their message broadcasts and node activity in the network. In [1] it is said that in underwater sensor networks, radio and optical signals do not show good performance. Radio waves have low frequencies varying between 300 and 3 MHz at long distances, it further has limitations of high antennas and high transmission powers to use radio signals. Optical signals are exaggerated from scattering and need high accuracy in directing the fine laser beams [2]. In contrast, acoustic signals performed better in UWSNs; acoustic signals are considered suitable for communication underwater. The service of sound signals carries out many unique challenges in underwater network system, such as bulky broadcast delay of acoustic signals of 15 km/s, limited bandwidth of <100 kHz, and high bit error rates because of the exciting features of the underwater network [3].

In our previous research, we proposed a routing protocol named energy efficient routing protocol—clustering algorithm [4]. This was using dynamic structure of network. Nodes are movable with water current. It used the concept of super node as discussed in [1]. In that algorithm we also assumed that the super node has very large and varying transmission power. Super nodes are placed at lower depths and at various places under water. Clusters are formed by super nodes at lower depths and then cluster heads are selected. Data transfer from source to sink contains various steps discussed in [4].

In the proposed paper, an architecture for reliable data transfer is presented along with finding tier head selection algorithm, finding the shortest path in the tiers, and finally a reliable energy efficient algorithm.

In the following sections some routing algorithms are reviewed, RE^{2}R architecture is proposed, and the algorithm is explained.

2 Literature Review

In this section, we analyze the previous routing protocols related to our proposed work. These protocols are divided into two categories called position-dependent routing protocols and position-free routing protocols.

2.1 Position-Dependent Routing Protocols

Here, we analyze the routing protocols that are dependent on the position of sensor nodes.

In [5], VBF (vector-based forwarding) protocol, the node that wants to send data makes a path for itself to the destination node. The nodes in the region nearby the calculated path only have permission to contribute to data packet promoting. The area within the specified address of the path is known as a routing tube. Thus, VBF has an organized flooding mechanism in which the data packets cannot send outside the routing tube. It has some boundaries in the situation of scarce density; it is also possible that there will be no node in routing tube that will affect the promotion.

In [6], HHVBF (hop-by-hop VBF) is also a path form, like in VBF, but the vector formulated here is computed on a per-hop basis. This means at every node a new vector is computed toward the destination node. Each time compilation of the path at individual hop can overcome the problem of scarce density, however, this depends on the essential estimate of the location dependency.

FBR (focused beam routing) protocol [7] uses diverse broadcast powers to choose the node that will be further responsible for data transmission. For the data transmission process, a sender sends a ready-to-send message packet with an assured broadcast power. If the sender receives a clear-to-send message packet, the node can send the data packet. If no clear-to-send message packet is received, the sender should increase the broadcast power. This process is repeated until a clear-to-send message packet is received. This protocol increases the delay due to clear-to-send message packet and Ready-to-send message packet.

2.2 Position-Free Routing Protocols

In [1] the authors propose MRP (multilayer routing protocol for UWSN). This protocol assumes two things. First, the protocol works on acoustic signals because under water we cannot use radio signals that do not allow to set up large antennas underwater. Also, this protocol assumed a special node called super node consisting of varying transmission power and abundant energy. In this, a limited number of nodes is involved during transmission of data from source node to destination node which reduces the energy consumption.

In [8], the author proposed a location free routing algorithm named depth-based routing. This algorithm uses the concept of depth locations of the sensor nodes for data forwarding. The sensor node that wants to send data to sink node prepares a packet, includes its depth measurement, and forwards the packet on the way to sink. The nodes on the way to sink compare their depths included in the packet with the depth of the sender node and the nodes having smaller depths parameter can further forward the packet toward the sink.

In [9] the authors proposed the hop-by-hop dynamic addressing algorithm. This routing protocol works according to the procedure to assign an identity to each node called a hop Id. Hop Id is based on node count route from source to sink. The node Id is allotted with the help of messages by the destination nodes where sink is placed at the surface. Every node that receives the message from sensor assigns a hop Id. Hop Id is incremented from origin to sink. Sender nodes choose a node having least node Id for its data transmission. The main disadvantages of this algorithm are that it suffers from delay.

3   Challenges in Designing UWSNs

Some challenges and future scopes are presented in [3]. Of all the presented challenges and future directions, we found some challenges to work on.

- Available data rates in UWSNs are extremely low, so routing overhead should be low as possible.
- Most the algorithms have fixed packet length data transfer. For efficient utilization of channel variable data packet should be used.
- Network should send data over multiple small steps to develop energy efficient routing protocol.
- Routing algorithm should be designed on the latest information received.
- Protocols can give better results if load is distributed over multiple nodes.
- Algorithm should be designed to give a node movement model.

4   Proposed Architecture

The architecture of RE^{2}R is presented. In this architectural arrangement sensor nodes are present, super nodes are shown in orange. Tiers formation is done by super node having abundant energy and varying transmission power. The super node broadcasts a message with some transmission power. The sensor nodes that listen to the message will form Tier 1. Likewise, a tier is formed. Again super node increases its transmission power and broadcasts a message with Tier Id 2. The nodes listening to this message consider themselves in Tier 2. In this way, all tier partitioning is done. Then a tier head is selected based on the algorithm proposed. Sensor nodes sense data regularly at any time nodes collect data from the surroundings; they send it to their respective tier heads. This will send data to the super node and then super node sends data to the receiving base station (Fig. 1).
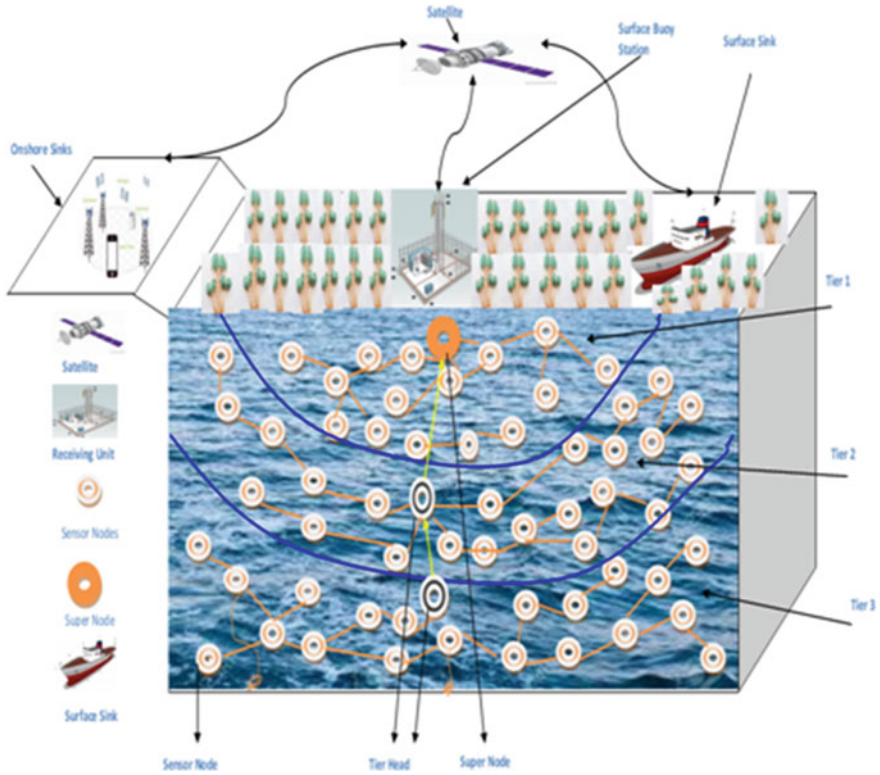
Fig. 1 Proposed architecture of RE²R

5 Proposed Algorithm, i.e., RE²R

Terminology used:

a. Dynamic routing algorithm used to form a chain between tier head and all nodes in tier.
b. Threshold Energy.

5.1 Architecture Formation Phase

In this phase some devices for receiving data are placed on the water surface. In this algorithm we used one super node having abundant transmission power and energy. Base station is fixed at some location. Sensor nodes are distributed under water in 3D. Super node is fixed at some lower depth near to water surface.

5.2 Tier Formation

Tiers are formed of sensor nodes to reduce the energy consumption. Each tier is constituted by a super node at lower depth. First, the super node broadcasts a tier formation message with some transmission power having Tier Id 1. The number of nodes that received message will form Tier 1. For Tier 2 super node increases transmission power and sends message with Tier Id 2. The nodes that listen to the broadcast message can considered themselves into Tier 2. Also if Tier 1 node listens to second message it cannot change their Id once assigned. Same as the second, super node increases its transmission power and sends message having Tier Id 3 to form Tier 3.

5.3 Tier Head Selection

For the first tier, tier head is the super node. For the remaining the tiers head can be selected based on the broadcasting method as given in [1].

5.4 Finding the Best Path

After the tier head selection, the shortest paths are selected around the heads using dynamic Dijkstra algorithm.

5.5 Data Forwarding Phase

In this phase each node senses and collects data from surroundings and sends it to its tier data. Tier head will have aggregated data and sends to other nodes in the process toward the receiving device.

5.6 Super Node

a. Super node is at the lowest depth and sends aggregated data to the receiving device at water surface.
b. After sending data it sends an acknowledgment message with the maximum power.

5.7 Receiving Device

The receiving device receives aggregated data from the network of super nodes and sensor nodes, and sends this data for analyzing at base station for end user queries.

5.8 Base Station

Base Station collects data from receiving devices and analyzes it for end user queries.

6 Results

Energy model proposed in [8] is used for energy consumption. This section includes the results for energy consumption and network lifetime.

6.1 Network Lifetime

Network lifetime is less increased in RE²R. We found this works more efficiently as its cost is low compared to EERU [4] or any other technique (Fig. 2).

6.2 Energy Consumption

Energy consumption is decreased in RE²R. Analogous to network lifetime we found this work more efficient because its cost is low compared to EERU or any other technique (Fig. 3).

Fig. 2 Network lifetime against DBR, MRP, and EERU-CA
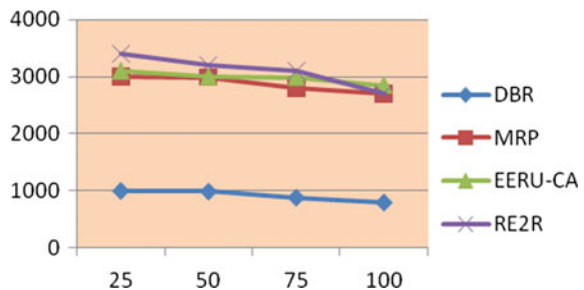
Fig. 3 Energy consumption against DBR, MRP, and EERU-CA [4]
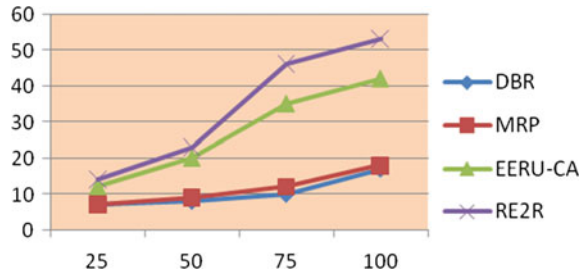
7   Conclusion

This paper presents a (RE^{2}R) reliable energy efficient algorithm for underwater sensor networks. In this super node assumes to have abundant energy and variable transmission power. Also, this technique divides the network into tiers. Among the tiers we proposed a technique to select the cluster head. Then shortest path finding algorithm runs to find a path. After getting a suitable tier head, path communication starts over this. The energy model used in [8] is used for energy consumption. Besides having a smaller improvement over EERU-CA [4] it could be assumed more reliable because its cost is much less than EERU-CA [4].

References

1. Abdul Wahid, "MRP: A Localization free Multi layered routing protocol for Underwater Wireless Sensor Networks". Volume 77, Issue 4, pp 2997–3012, August 2014.
2. Akylidiz, I. F., Pompili, D., & Melodia, T. Underwater acoustic sensor networks: Research challenges. Ad Hoc Networks, 3(3), 257–279, (2005).
3. Chen, K., Zhou, Y., & He, J. (2009). A localization scheme for underwater wireless sensor networks, International Journal of Advanced Science and Technology, 4, 9–16.
4. Khan Gulista, Gola Kamal Kumar, Ali Wajid, (2015), "Energy efficient routing algorithm for UWSNs-CA", published in second international Conference on Advances in Computing and Communication Engineering. 2015.
5. Xie, P., C ui, J. H., & Lao, L. (May 2006). VBF: Vector-based forwarding protocol f or underwater sensor, networks. In Proceedings of the IFIP networking conference. Coimbra, Portugal (pp. 1216–1221).
6. Nicolaou, N., See, A., Xie, P., C ui, J. H., & Maggiorini, D. (2007). Improving the robustness of location-based routing for underwater sensor networks. In Proceedings of the IEEE OCEANS'07, Europe, June, 2007, pp. 1–6, doi:10.1109/OCEANSE.2007.4302470.
7. Jornet, J. M., Stojanovic, M., & Zorzi, M. (Sept 2008). Focused beam routing protocol for underwater, acoustic networks. In Proceedings of the Mobicom WUWNet'08, San Francisco, C A, USA, pp. 75–82, doi:10.1145/ 1410107.1410121.
8. Yan, H., Shi, Z., & Cui, J. H. (May 2008). DBR: Depth-based routing for underwater sensor networks. In Proceedings of the IFIP networking'08 conference, Singapore, pp. 72–86.
9. Ayaz, M., & Abdullah, A. (Dec 2009). Hop-by-Hop dynamic addressing based (H2-DAB) r outing protocol for underwater wireless sensor networks. In Proceedings of 2009 International conference on Information and Multimedia Technology ICIMT'09, Jeju Island, Korea, pp. 436–441.

Mitigation Techniques for Gray Hole and Black Hole Attacks in Wireless Mesh Network

Geetanjali Rathee and Hemraj Saini

Abstract Wireless mesh network (WMN) is considered as a key technology in today's networking era. In this, security is determined as a vital constraint. Several approaches have been proposed to provide secure communication in WMN but communication security possibilities always exist and are very hard to maintain. Black hole and gray hole are the two major attacks at the network layer in WMN. RAODV, IDSAODV, and RIDAODV are some security approaches against these attacks. These approaches can immune the communication at a certain rate. We have proposed a cache-based secure AODV routing protocol, i.e., SDAODV in which instead of using the RREQ messages, security is provided by using the last sequence number of each packet. Using this approach, we have improved the network throughput at a certain rate. The performance of the proposed approach is evaluated by showing the throughput graph.

Keywords Wireless mesh network · Black hole · Gray hole · Authentication · SDAODV · Security

1 Introduction

A wireless mesh network [1] can be defined as a synonym of ad hoc networks which have the capability of self-organizing and self-healing the network topology. Due to much more connectivity than traditional network, wireless mesh network provides better internet services. A WMN is made up of two types of nodes, mesh

Geetanjali Rathee (✉) · Hemraj Saini
Computer Science Department, Jaypee University of Information Technology,
Waknaghat, Solan, Himachal Pradesh, India
e-mail: geetanjali.rathee123@gmail.com

Hemraj Saini
e-mail: hemraj1977@yahoo.co.in
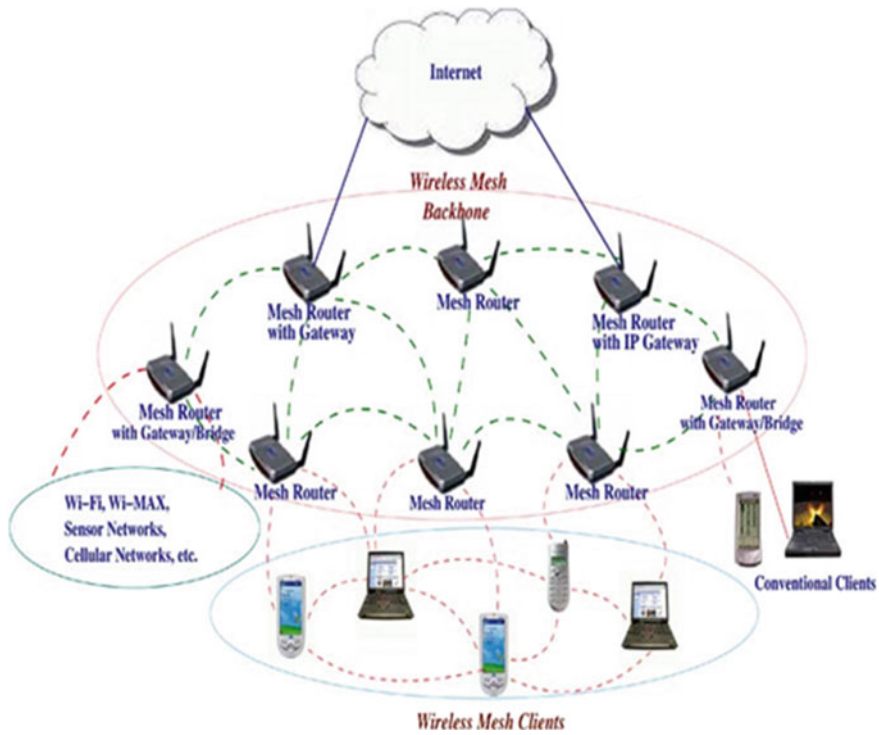
Fig. 1 Key distribution and inter-cluster communication

client and mesh router (as shown in Fig. 1). Mesh routers are generally stationary and are used to provide the services to mesh clients. Mesh clients are generally mobile in nature and utilize the services from the internet. As different organizations use wireless mesh networks because of their tremendous features, it is required to provide security [2] to mesh clients. The self-organizing nature of wireless mesh network leads to various types of attacks [3]. The goal of multi-hop network is to provide the services with the security solutions. The security of a system can be broadly divided into four parts: (i) availability, (ii) confidentiality, (iii) integrity, and (iv) authentication [4]. Due to its broadcasting nature, WMN is highly liable to security attacks. An intruder or malicious node may drop the packets partially or fully to perform various attacks, i.e., black hole attack and gray hole attack [5]. In this paper, we briefly discuss about each of the attacks with their security mechanism.

The paper is organized as follows. In Sect. 2 we discuss about various attacks with their security mechanisms and in Sect. 3 an algorithm is proposed to remove the drawbacks that occurred in previous approaches. The comparative study will be analyzed in Sect. 4 and finally the paper is concluded in Sect. 5.

2   Related Work

Gray hole and black hole are the special cases of DOS attacks which are also referred as selective forwarding bout. A gray hole attack is the one where a node drops the selected packets which it has to forward along the path. In this, an attacker may simply hijack a mesh or legitimate node in the network. A number of detection and prevention approaches have been proposed for both gray hole and black hole attacks. In this section let us discuss some of the detecting and preventing techniques for both the attacks.

2.1   Mechanisms to Detect Gray Hole Attacks

Karlof et al. [6] proposed a multi-hop forwarding approach. Instead of defining attacks, the author has proposed the way to detect gray hole. Chemas [7] also applied the same procedure used by Karlof. It detects the malicious behavior by monitoring their neighbors and collaborating with the cluster nodes. In order to detect the malicious behaviors, some of the intermediate nodes are selected as checkpoints whose task is to generate an acknowledgment for each packet received along the forwarding path. The major drawback of this approach is overhead. An another gray hole detection approach is a watchdog and path rater [8]. In this scenario Matri et al. used watchdog to observe its neighbors and to identify a malicious node, after that path rater measures the corresponding results to select the reliable path for packet delivery.

   To detect the collaborative attacks, the authors have proposed a BSMR [9] technique. It is a multicast routing protocol to capture the colluding adverse behavior. But the major drawback of this approach is that it uses a static detection threshold and makes it difficult to identify the intruder at different malicious dropping rates. Another mechanism to prevent gray hole attack is Markov game. Shila et al. defined a framework of non-cooperativeness between legitimate or malicious nodes. It detects the attack by checking the throughput of the network. If a node is genuine, then it will simply maximize the throughput by minimizing packet loss while attacker's aim is to minimize the throughput by dropping the packets. Above we have discussed some detection techniques for gray hole attacks. Now let us focus on some black hole detection schemes.

   The only difference between a black hole and a gray hole attack is that instead of discarding the whole packets, gray hole attack discards selective packets coming from the source node.

2.2 Mechanisms to Detect Black Hole Attacks

Jiltsn [10] proposed a path-based detection algorithm. The aim of the node is to detect its neighboring node instead of detecting all the nodes in the current path. Patcha et al. [11] proposed a watchdog where nodes of the network are divided into three parts: (i) trusted node, (ii) watchdog nodes, and (iii) ordinary nodes. The task of watchdog is to detect the behavior of its neighboring nodes, whether they are behaving as normal nodes or malicious nodes. Dneg et al. [12] has proposed a malicious node detecting technique by categorizing the nodes as strong or weak nodes. Strong nodes are those having well computing power or radio range. But the major drawback of this approach is that, if the attack is put on strong nodes then the algorithm fails to work.

The major drawbacks with the previous approaches are:

a. As multipath forwarding approach is the first detection theory of detecting various attacks, but it is not able to detect the attackers or intruders.
b. Chemas increases the overhead inside the network and consumes a large number of network resources.
c. Watchdog and path rater not only increase the throughput, but the watchdog has its own limitations also.
d. The major drawback of BSMR is that it is unable to detect the attacks at different places.
e. The path-based algorithm which detects a black hole attack is able to perceive only its succeeding nodes in the current path.
f. Watchdog methods increase the load inside the network to check the behavior of malicious or legitimate nodes.

The major drawback of the path rater approach is that if an attack is encountered on strong nodes then the algorithm fails to work.

So far we have discussed various detection algorithms for black hole and gray hole attacks. Now there are several researchers who have proposed several techniques to remove these attacks, i.e., RAODV [13], IDSAODV [14], and RID-AODV [15].

2.3 RAODV

The RAODV follows an AODV process in reverse order, i.e., by accepting a RREQ from source node as shown in Fig. 2. Destination node will send a RREQ in reverse order, i.e., R-RREQ. The simulation results of author improve the network performance, i.e., throughput, packet delivery ratio, and so on.

Fig. 2 RAODV diagram
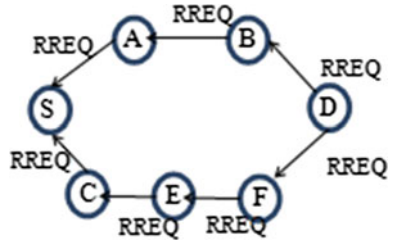
Fig. 3 IDSAODV diagram
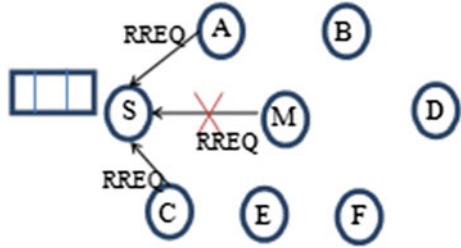
2.4 IDSAODV

Dokuror et al. has proposed another light weight routing protocol to prevent the black hole attack. In this the author assumed that during the route discovery phase, the first RREP message is always from an intruder and is declared as a black hole attack while the second RREQ message is from destination node. In this, cache mechanism is used to count the sent RREP during the route discovery process as shown in Fig. 3.

2.5 RIDAODV

Another black hole solution is proposed by Shree et al. which is a combination of RAODV and IDAODV protocols as shown in Fig. 4.

Fig. 4 RIDAODV diagram
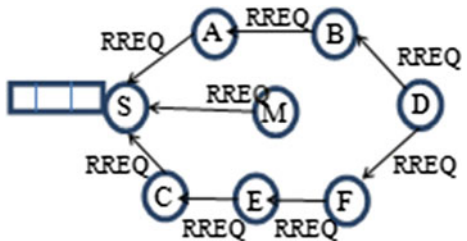
But the major drawback of all these approaches is load on the entire network and as a result the throughput decreases. Section 3 shows a novel approach, i.e., SDAODV.

3 Proposed System

To remove all the previous drawbacks a new approach, i.e., SAODV is proposed against black hole attack. The proposed approach is an advanced version of IDSAODV. SDAODV approach uses a cache mechanism as defined in IDSAODV but instead of counting or storing the RREQ messages in the cache, it stores the sequence number of packets. Each packet has a unique sequence number to remove the duplication in the message. To implement this solution, each node has a cache which is used to store the last sequence number of received packets. The cache is updated regularly during rising or transmitting the packets. A destination node initiates a RREP message having last packet's sequence number after receiving the RREQ message (as shown in Fig. 5).

Abbreviations used in the proposed solution are shown in Table 1.

3.1 Working of SAODV

Initially, source route request message (Srreq) is broadcasted to the neighbor N in the network. Whenever an intermediate node has a route R to destination D, the

Fig. 5 SAODV diagram
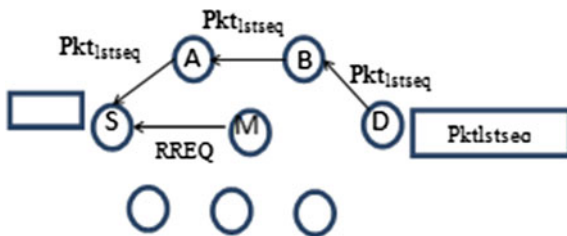
Table 1 Comparative analysis of previous proposed protocols

Abbreviations	Terms
Srreq	Source route request
R	Route
RREP	Route reply
Srreq	Source route request

Table 2 Measured parameters

Abbreviations	Terms
Surface size	500 * 500
Node group	20, 25
Black hole range	0 up to 5
Starting time	450 s
Stop timing	390 s

intermediate node sends the RREP message having the last packet's sequence number of the destination node. The path followed by RREQ message from destination node D to source node S is {D-B-A-S}.

The proposed approach is simulated on NS-2. The requirements during simulation are shown in Table 2.

To analyze the network performance, throughput is taken as an important parameter and is measured over RAODV, IDSAODV, RIDAODV, and SDAODV routing protocols. As shown in Table 2 the SDAODV routing protocol measures the highest throughput rate against the black hole attack.

3.2 Performance Analysis

In order to prove the efficiency of proposed work and to reduce the above mentioned drawbacks (i.e., storage overhead and network load, the technique is compared with RAODV, IDSAODV, and RIDAODV protocols against the throughput parameter. In this to reduce the storage and computation overhead at each node, a cache is used which stores only the last sequence number of the packets to avoid message duplication. The throughput achieved by implementing individual protocols is depicted in Table 3.

Table 3 Comparative values of all the protocols against throughput parameter

Number of nodes	Throughput			
	RAODV	IDSAODV	RIDAODV	SOAODV
10	60	68	85	92
20	62	69	86	93
30	64	75	88	94
50	67	78	90	96

4   Results

Figure 6a, b show the RAODV and IDSAODV graphs where maximum throughput achieved is 67 and 78 % as black hole nodes increase in the network.

Figure 6a, b show the throughput graphs of RIDAODV and SDAODV where performance during the increase in black hole nodes is 90 and 96 %. So the corresponding graphs show that proposed SDAODV is a more preferable protocol against the black hole attack in WMN.

Fig. 6 a Throughput of RAODV and IDSAODV. **b** Throughput of RIDAODV and SDAODV
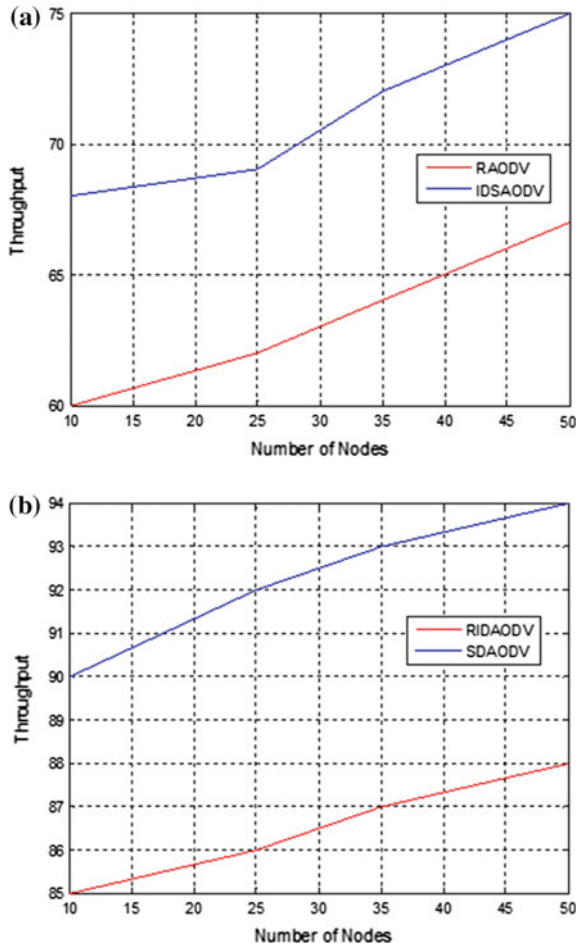
5 Conclusion

In this paper, we have proposed a cache-based secure AODV routing protocol in which security is provided by using the last sequence number of packets. Source node S receives the RREQ messages by rejecting the first RREQ message. As black hole node sends the RREQ reply to the source node, source node simply rejects the RREQ as there is no last sequence packet in the request. The aim of proposed technique is to provide the security inside the network with improved throughput parameter. We have simulated the approach over NS2 simulator against throughput parameter and have shown the corresponding graphs achieved by all the approaches. SDAODV provides the highest throughput rate. The future plan of our work is to reduce the number of computation and storage overhead done by each router to provide the security.

References

1. Akyildiz, Ian F., and Xudong Wang: A survey on wireless mesh networks. In: Communications Magazine, IEEE 43(9) (2005).
2. Redwan, Hassen, and Ki-Hyung Kim: Survey of security requirements, attacks and network integration in wireless mesh networks. In: IEEE conference on New Technologies, Mobility and Security, NTMS (2008).
3. Zhang, Yanchao, and Yuguang Fang: ARSA: an attack-resilient security architecture for multihop wireless mesh networks. In: IEEE Journal on Selected Areas in Communications, 24 (10) (2006).
4. Baradeli, Nagaraj B., and Mr SP Srikanth: Architecture to Achieving Security in WMN (2004).
5. Cai, Jiwen, et al.: An adaptive approach to detecting black and gray hole attacks in ad hoc network. In: 24th IEEE International Conference on Advanced Information Networking and Applications (AINA), (2010).
6. C. Karlof, D. Wagner.: Secure routing in wireless sensor networks: attacks and countermeasures. In: Elsevier's Ad Hoc J., vol 2(3), pp. 293–315 Sept (2003).
7. B Xiao, B. Yu, and C. Gao.: Chemas: identifying suspect nodes in selective forwarding attacks. In: Journal of Parallel Distributed Comput. Vol 67 pp. 1218–1230.
8. S. Marti, T. J. Giuli, K. Lai and M. Baker.: Mitigating Routing behaviour in mobile ad-hoc networks. In: Proc. International conference on mobile computing and networking, Boston, MA (2009).
9. R. Curtmola and C. Nita-Rotaru.: BSMR: Byzantine-resilient secure multicast routing in multi-hop wireless networks. In: Proc. Sensor, Mesh and Ad Hoc Communications and Networks, June (2007).
10. JiItsn Cai, Ping Yi, Jialin Chen, Zhiyang Wang, Ning Liu.: An adaptive approach to detecting black and Gray hole attacks in Adhoc networks. In: 24th IEEE International Conference on Advanced Information networking and application, pp. 775–891 (2010).
11. A. Patcha, A. Mishra.: Collaborative Security architecture of black hole attack prevention in mobile ad hoc networks [C]. In: Radio and Wireless Conference, pp. 75–78 (2003).
12. Hongmei Deng, Itsi Li, and Dharma P. Agrawal.: Routing Security in Wireless Ad Hoc Network. In: IEEE Communications Magazine, 40(10), pp. 70–75 October (2002).

13. C. Kim, E. Talipov and B. Ahn.: A Reverse AODV Routing Protocol in Ad Hoc Mobile Networks. In: Proceeding from EUC'06: The 2006 International Conference on Emerging Directions in Embedded and Ubiquitous Computing, Seoul, 1(4), pp. 522–531 August (2006).
14. S. Dokurer, Y. M. Erten and E. A. Can.: Performance Analysis of Ad-Hoc Networks under Black Hole Attacks. In: Proceeding from SECON'07: IEEE Southeast Conference, Richmond, pp. 22–25 March (2007).
15. Shree, Om, and Francis J. Ogwu.: A Proposal for Mitigating Multiple Black-Hole Attack in Wireless Mesh Networks." (2013).

Investigation for Land Use and Land Cover Change Detection Using GIS

Prakash Shivpuje, Nilesh Deshmukh, Parag Bhalchandra, Santosh Khamitkar, Sakharam Lokhande, Vijay Jondhale and Vijay Bahuguna

Abstract For better development and good space planning, it is required to understand the recent changes that have occurred in the surrounding environment. The Change Detection method is usually deployed for this. The underlying research study did the same for Nanded city. The criterion used was to understand land use and land cover change variables as a token of change detection. The results helped us to understand urban spreading out that took place in the time zones 1973–1992 and 1992–2014. Our approach was using Remote Sensing Technology. Landsat images as MSS, TM, and LC were collected and their analysis was carried out. This paper can work as a role model for GIS-based LU/LC change detection of any city in India.

Keywords Geographic information system (GIS) · Multispectral scanner (MSS) · Thematic mapper (TM) · Change detection · LU/LC · Bands

Prakash Shivpuje (✉) · Nilesh Deshmukh · Parag Bhalchandra ·
Santosh Khamitkar · Sakharam Lokhande · Vijay Jondhale
School of Computational Sciences, S.R.T.M. University, Nanded 431606, MS, India
e-mail: nndvcl@gmail.com

Nilesh Deshmukh
e-mail: nileshkad@yahoo.com

Parag Bhalchandra
e-mail: srtmun.parag@gmail.com

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

Sakharam Lokhande
e-mail: lokhande_sana@rediff.com

Vijay Jondhale
e-mail: vijay.jondhale@gmail.com

Vijay Bahuguna
Department of Geography, DBS PG College, Dehradun 248001, UK, India
e-mail: vijaybahugunadbs@gmail.com

1 Introduction

To understand urban spreading and expansion of human habitats, the remaining land for future development, proper planning of a city, the side effects of human developments on natural resources, etc., the Change Detection approach is popularly practiced. It is a timely phenomenon that gives a good idea for proper management of land resources [1, 2]. It is usually done using GIS tools. Using GIS, we take satellite images, store them, decode them, and then process them to understand the changes detected so far [3, 4]. Two common variables are used for doing this activity: LC and LU. Land Cover (LC) is the total land available on the earth's surface. Land Use (LU) is the total land used for development including natural resources and habitats. To understand LC–LU ratios, we take satellite photographs for different years and then compare them using Change Detection approach [5, 6].

2 Study Background

2.1 Area Selected for Study

Nanded Tahsil is situated in Nanded district in Maharashtra state and is located between 18°57′ and 19°21′ North latitudes, 77°9′ and 77°24′ East longitudes by covering an area of 335.79 km^2 Nanded Tahsil is bounded by Ardhapur on the north, Purna Tahsil on the west of the Parbhani district, Mudkhed Tahsil on the east, and Loha Tahsil in the south.

2.2 Data Acquisition in Terms of Satellite Images

We have taken the help of Land-Sat imagery. Land-Sat supplies good quality, sufficient resolution imagery. These images can be visible, infrared, thematic, thermal, and panchromatic. These images can be taken by multispectral sensor, thematic mapper, and enhanced thematic mapper plus sensors. There are 30 satellites for Land-Sat imagery applications. Satellite L1-4 gives MSS sensor images of 60 m pixel resolution. Satellite L4-5 gives thermal sensor images of 30 m pixel resolution. Satellite L7 gives enhanced thematic images (plus) and thermal images of 30 and 60 m resolutions respectively. All these images have screen size of 185 by 185 km. The images for our study were downloaded from glcf.umd.edu [7].

2.3 Research Methodology

Figure 1 shows systematic work methodology for data collection, processing, and understanding changes in the desired parameters.

The various blocks are described below:

(1) *The Data Sources and Collection* All these images are downloaded from glcf. umd.edu and USGS
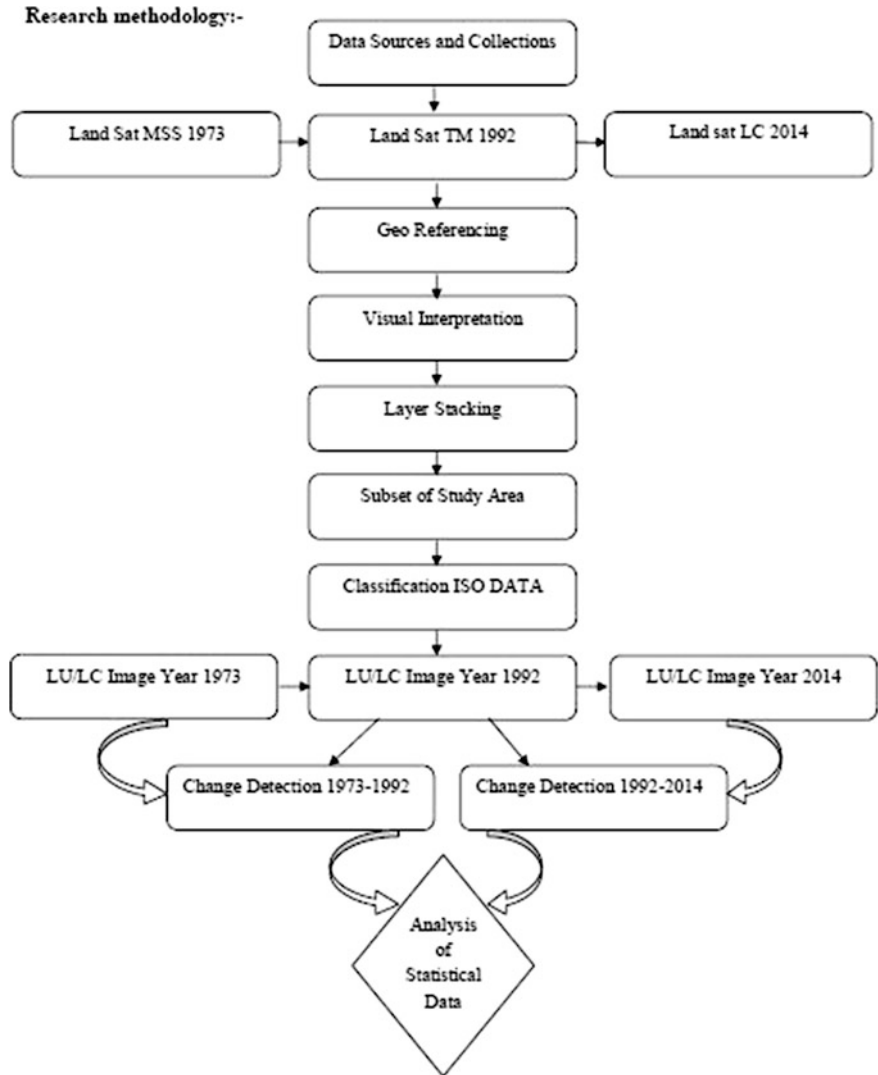
Fig. 1 Research methodology for analysis

    (a) 1973: MSS (Multispectral Scanner)
    (b) 1992: TM (Thematic Mapper)
    (c) 2014: LC

(2) *Geo Referencing* Most GIS projects have need of georeferencing some raster data. Georeferencing is the method of transmission real-world coordinates to each pixel of the raster.

(3) *Visual Interpretation* The visual interpretation of satellite images is a composite process. It includes the meaning of the image content but also goes beyond what can be seen on the image in order to identify spatial and scenery patterns. This progression can be approximately separated into two levels:

    (a) The identification of objects such as streets, fields, rivers, etc. The quality of recognition depends on the capability in image interpretation and visual perception.

    (b) Accurate explanation can be ascertained through conclusions (from previously recognized objects) of situations, recovery, etc. Subject specific information and expertise are critical.

(4) *Layer Composition* The images are available in different bands so we have to combine the image into one band and create a stack so all layers get into one.

(5) *Extracting Desired Area* After layer stacking we have to extract the area of interest from the given layer stacking.

(6) *ISO DATA for Classification* In this method we have selected some sample pixel point and from those points we started doing classifications.

(7) *LU/LC Maps* Now we have generated LU/LC maps, and from these maps we come to know the classification of various classes.

(8) *Change Detection* Here we have actually obtained differences in various maps of different classes'.

3 Results and Discussion

After analysis using the above research methodology, we prepare a table for LU and LC for the years 1973, 1992, and 2014 (Table 1).

Table 1 Change detection values for study area

Sr. No	LULC type	1973	1992	2014
1	Agricultural land	103,383	111,505	136,762
2	Water body	10,129	7839	9542
3	Vegetation	130,011	136,638	79,186
4	Wasteland	119,800	94,687	66,957
5	Settlement	71,403	84,057	142,269

Fig. 2 a Status in 1973.
b Status in 1992 and c Status
in 2014
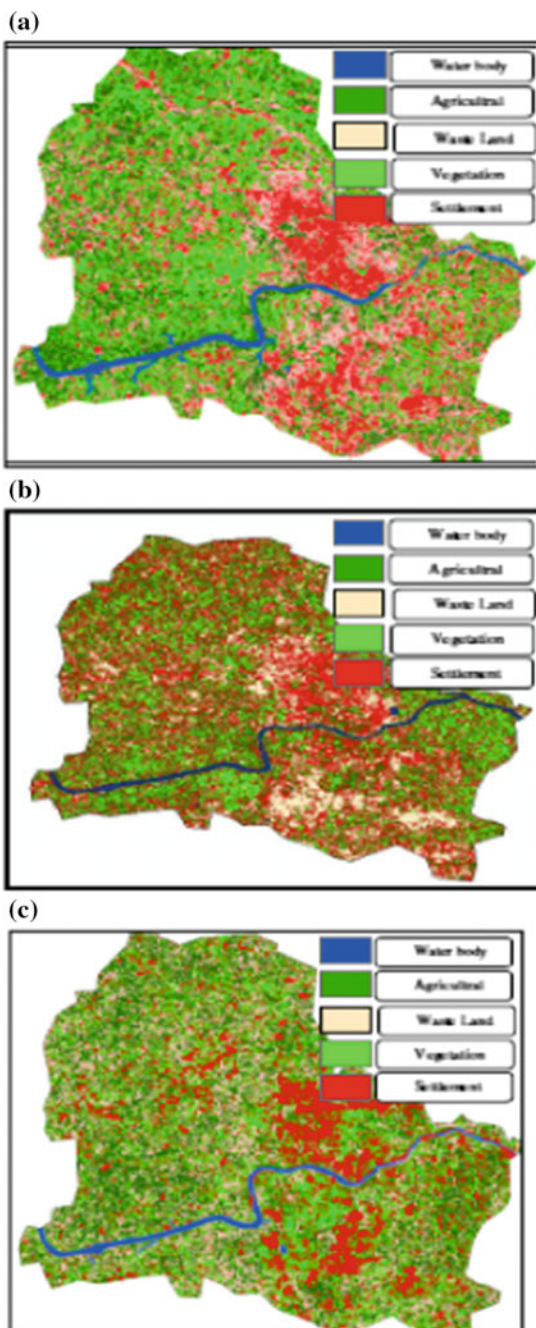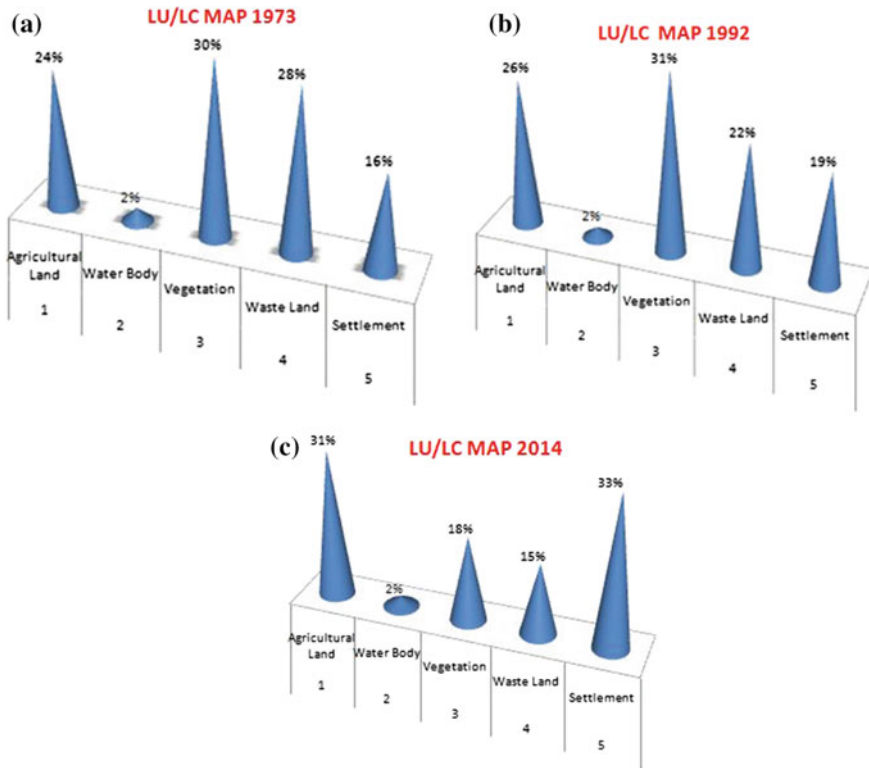
Fig. 3 **a** Status in 1973. **b** Status in 1992 and **c** Status in 2014

It is evident that he settlement pace is increasing whereas wasteland is decreasing. This eventually highlights the high pace of development. Using the Arc-GIS software, we have plotted these changes as shown in Fig. 2a–c.

Figure 3 shows a graphical representation of LU/LC maps.

4 Conclusion

Our study has found change detection analysis of Nanded city using the GIS approach. The parameters used were LC and LU. Satellite images were collected, processed, and analyzed to understand changes in the agricultural area, vegetation area, water body area, settlement area and wasteland. Our approach can be used as a freelancing logistical aid for government for better developing and planning. Our study can also be used for land surveys. From the above analysis, it is observed that there is increase in the built up area mainly because of the growing demand for residential spaces (Personal and Commercial). We come to know that the position in 1973 for agricultural land was 24 %

which increased by 2 % in 1992 and by 7 % in 2014; for vegetation land position in 1973 it was 30 % and rapidly reduced in 2014 by 12 %; for wasteland position in 1973 it was 28 % which was reduced in 1992 by 6 and 13 % in 2014.

5 Future Work

1. Implementation of Markov Transition Model for future change detection.
2. Implementation of change detection by calculating area in SQM.
3. Comparing the best accuracy using both models.

References

1. Change Detection Mapping: Using Remote Sensing and GIS Technology–A Case Study of Achanakmar-Amarkantak Biosphere Reserve, Central India, IJRSG, 2319-3484.
2. Land Use/Land Cover Changes Detection & Urban Sprawl Analysis, IJASRT, 2249-9954.
3. Land Use Land Cover Mapping, Change Detection and Conflict Analysis of Nagzira-Navegaon Corridor, Central India Using Geospatial Technology, ISSN 2277-9051.
4. Investigations on Change Detection in Chandpur (M.S.) Watershed area by using RS & GIS, ISSN: 2321-3124.
5. Tiwari Kuldeep, Khanduri Kamlesh, Land Use/ Land cover change detection in Doon valley (Dehradun Tehsil), Uttarakhand: using GIS& Remote Sensing Technique, IJGG, volume 2, no 1, 2011. [11].
6. Land use and land cover change detection through Remote Sensing approach: A case study of Kodaikanal Taluka, Tamil Nadu, ISSN 0976-4380.
7. Web resource at http://www.glcf.umd.edu.

Effectiveness of Computer-Based Instructional Visualization and Instructional Strategies on e-Learning Environment

Sanju Saha and Santoshi Halder

Abstract Learning with visualization is a new trend for the teaching and learning environment. However, in this study the question is do all types of visualization and strategies equally affect achieving various learning objectives? How computer generated questions with and without feedback strategies affect achievement of learning objective? To investigate the effectiveness of different types of visualization and strategies, researchers developed three different types of instructional modules (static, animated and interactive) and two types of instructional strategies (question and feedback). A total of 540 students were selected to conduct the study with specific matching criteria. MANOVA was done to find out group differences in different conditions. The results showed a momentous mean difference in different conditions i.e., in interactive visualization condition students performed better than animated and static condition; besides, question and feedback conditions were more effective than no strategies and only question conditions with respect to various learning outcome. The result is discussed critically from several theoretical focal points.

Keywords Interactive instructional visualization · Instructional strategies · Method of instruction · Learning objective · e-Learning

Sanju Saha (✉)
University of Calcutta, Alipore Campus, 1 Reformatory Street, Kolkata 700027, India
e-mail: sanju_saha@yahoo.co.in

Santoshi Halder
Department of Education, University of Calcutta, Alipore Campus, 1 Reformatory Street, Kolkata 700027, India
e-mail: santoshi_halder@yahoo.com

1   Introduction

In the Indian education system, one of the major problems is enormous student–teacher ratio [1]. In developed countries this ratio stands at 11.4 and in India, it is as high as 22.0 (The United Nations Educational, Scientific and Cultural Organization, UNESCO 2009). It is even lower in Commonwealth Independent States (CIS) 10.9, Western Asia 15.3, and Latin America 16.6 [2]. Research ensures that in India, even today 20 % of the schools have just one teacher [1].

Various published reports, such as Annual Status of Education Report (ASER), the Programme for International Students Assessment (PISA) and the Quality Education Study (QES) state that in teaching and learning process especially in science learning, the drop-out rate is 5–10 % in the past few years [1]. Nevertheless, compared with 74 other countries, students of India are ranked second to last at 73rd position, just above Kyrgyzstan in respect of drop-out rate in the teaching and learning [1]. Now the question that arises is why this vast difference ratio in respect of global orientation? Is there some gap or erroneous way being followed in the teaching learning process?

To handle the above-mentioned situation over the past, various researchers have raised their voice and encouraged to utilize various modern technologies in the present teaching and learning process. However, in the present situation computer-based visual learning environment is a growing trend encouraging individualized learning environment [2–4]. Now, the important question is, does various kinds of visualization equally affect achieving various learning objectives?

In addition, in the present situation especially in India questions, with or without feedback strategies are mainly used in traditional lecture-based classroom environments which mainly focus on print material and teacher-based approaches. However, feedback is a continuous process until the learner achieves the learning objective; feedback needs to be provided repeatedly. Nonetheless, teacher-based feedback is limited to provide repeatedly as well as individually at the individual pace of the students. Taking this point into consideration, in the present study major focus has been given to explore the effectiveness of computer-based question with or without feedback.

To unveil the mentioned situation this study is undertaken to find the effectiveness of various instructional visualization (static, animated, and interactive) and various instructional strategies (no question, question and question plus feedback) to achieve various learning objectives (factual, conceptual and rules, and principal knowledge).

1.1   Instructional Visualization: Research Overview

A different option to apprehend the present study related to animations, explored by a number of researchers [5], regards manipulation features on the animation visual. Interactivity provides the possibility to stop, rewind, and restart, gradual down, or

alter instructional materials. Interactivity, from the view of reminiscence needs, must lead to less cognitive load and should make stronger comprehension [6]. However, manipulation also offers the learner the opportunity to replay a component, thus shaping the display. A learner can adapt the display velocity to his/her process free of cognitive processing. The basic stage of learner-control involving the rate of presentation of multimedia slides has been studied by various researchers [7, 8]. These experiences confirms the skills of this stage of learner-manipulation.

1.2 Questioning and Feedback Strategies in Teaching and Learning Environment

Regarding the effectiveness of instructional strategies previous findings on feedback strategy are contradictory. Authors [9] found that knowledge of correct response (KCR) and knowledge of response (KOR) feedback is more effective than no feedback. Besides, researcher [9] gives more emphasis on the prestructural attribute of feedback in the sense that they have given more importance to unfold the purpose of feedback and its function for establishing effective connection between feedback strategies and the learner. However, researcher identified feedback as a reinforcing mechanical process that acts as stimulus and increases the future probability of the response which it follows. In addition, it has also been explained in previous research that question strategies function as a behavioral mechanism which in the prior stage informs the learner about the expectation of their learning [10]. Nevertheless, question functions as a recall mechanism that helps to retrieve information from short-term memory (STM) to long-term memory LTM [10]. Several empirical studies have been conducted in the field of question and no question strategies. Researcher [9] found that in recall perspective question is more effective than no question condition. Cameron and Dwyer [11] found that self-questioning strategy is significantly effective strategy for the long-term retention of material vis-a-vis lecture method.

However, a majority of research focused on the functional structure of question and feedback strategies. This study aims to unveil the effect of feedback and questing strategies in interactive instructional visualization on the specific knowledge domain, viz, factual, conceptual and rules, and principle.

2 Objectives

(a) To investigate the effect of instructional visualizations (static, animated, and interactive) on students' learning of different educational objectives in a CBI environment.
(b) To investigate the effect of varied instructional strategies (no strategy, questions, and questions plus feedback) on student learning in a CBI environment.

3   Research Tools

General Information Schedule: This comprises of students' demographic information.

Computer Proficiency Test: This test was prepared for initial screening of students to assess students' computer proficiency on three dimensions namely: basic knowledge about computer, usability of computer, and use of computer.

Prior Knowledge Test: This test originally developed by Dwyer [12] was restandardized and validated by researchers in the Indian context consisting of 36 multiple-choice questions on the subject physiology.

Criterion Measures Test: The three criterion tests adapted were originally developed by Dwyer [12] to measure different learning objectives. Each consists of 20 multiple-choice questions. There was no time limit for test completion. Each test item was worth 1 point for a total of 20 possible points per test. The subdimensions of the test are identification test, terminology test, and rules and principle test.

3.1   Reliability and Validity of All Three Criterion Tests and Computer Proficiency Test

The KR 20 result indicated high reliability for the three criterion tests (0.86 for identification test, 0.81 for terminology test, and 0.85 for comprehension test) and 0.85 for computer proficiency test indicating all the above 0.80, which is a satisfactory level of reliability [13].

4   Development of Instructional Module and Learning Material

Instructional content material used in the study was adapted from a color-coded, paper-based booklet originally developed by [12] regarding the human heart divided into five units; (1) the heart structure; (2) the veins and arteries; (3) the valves of the heart; (4) the blood flow through the heart; and (5) the phases of the heart cycle. This content is chosen after discussing and analyzing with subject expert as it allows the evaluation of different levels of learning objectives.

4.1   Illustration of Development Instructional Module

For the purpose of the study the following three separate instructional modules were developed by the researchers:
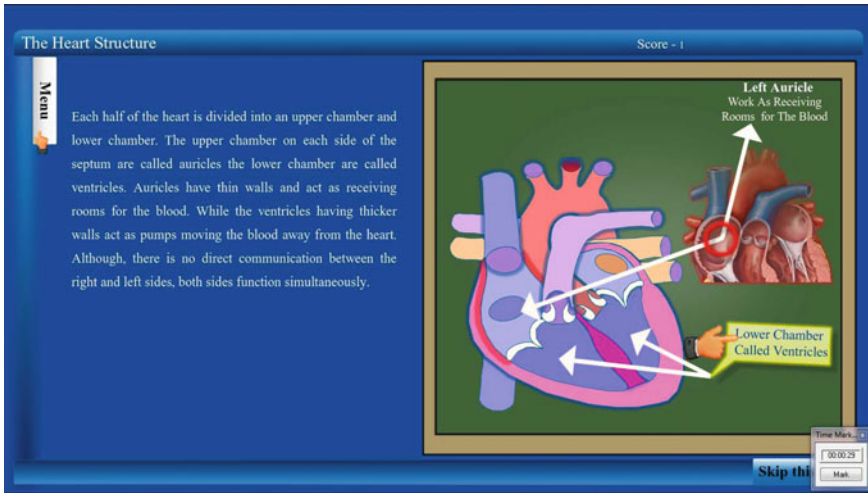
Fig. 1 The screenshot of interactive visual condition

Static Visualization Condition: In this treatment group participants were provided static image complementing with corresponding text. The static picture was provided in combination with line drawing along with some realistic picture. Figure 1 presents the screenshot of only static visualization condition.

Animated Visualization Condition: In this treatment condition visual representation of instructional material are in animation form. Additionally, there is play, pause button sequence to control the animated segment.

Interactive Visualization Condition: In this condition participants were provided with some interactive learner control features for manipulation of each graphical representation as well as some program-based action sequence and programmed generated instruction. Participant needs to follow as per program generated instruction by click, drag, and manipulation (Fig. 1).

Question Condition In this condition various related questions were provided along with visualization.

Question Plus Feedback Condition Akin to static condition also provided computer generated feedback (Fig. 2).

5 Sample

This study was conducted on the secondary level students of Central Board of Secondary Schools (CBSC) of Kolkata. Out of 630 students 540 were selected based on the matching criteria. A majority of them belonged to lower-middle-class families and their age ranged from 13 to 15 years (mean age 14.26 years and SD = 1.75). A pilot study was conducted on 25 % of the total samples and all the tools were finalized after that.

Fig. 2 The screenshot of Questioning + Feedback Frame

6 Results Analyses and Interpretation

6.1 First Phase: Preliminary Data Analysis Physiology Prior Knowledge Test (Covariate)

A variance of analysis was conducted in respect of prior knowledge test scores to determine if there was a significant difference among the treated groups on the prior knowledge test. The result of the ANOVA analysis indicating interaction between three dependent variables and prior knowledge test found no significant differences among the treatment groups on the test see (Table 1) score $F = 36$, 1.28, $p = 0.17$. The result indicated that the participants were approximately equal in their prior knowledge of the content material used in the study and therefore any results of treatment effects would not be attributed to the difference in participants' prior knowledge.

Table 1 Tests of between-subject effects (prior knowledge test and three criterion tests $N = 540$)

Source	df	Mean square	F	Sig.
Identification	18	3.46	0.56	0.91
Terminology	17	13.24	2.16	0.06
Comprehension	17	7.09	1.15	0.32
Identification * Terminology * Comprehension	36	7.90	1.28	0.17

Table 2 MANOVA Pillai trace test interactive visualization and instructional strategies

Independent variable	Value	F	df	Sig.	$\eta\rho^2$
Instructional visualization	0.535	63.481	6	0.000*	0.268
Instructional strategies	0.507	59.018	6	0.000*	0.254

6.2　Second Phase: MANOVA Group Comparison

In order to test the three-null hypothesis, the following analyses were conducted with all test items as well as the subset of difficult test items in each of the three criterion tests examined by question level and feedback level: (a) two-way MANCOVAs; (b) follow-up tests of the between-subjects effects; and (c) post hoc pair-wise comparisons.

From Table 2, it is found that two independent variables (instructional visualization and instructional strategies) depicted a significant main effect on three criterion test results (identification terminology and comprehension) (Pillai trace = 0.53, F (6, 63.48) and $\rho = 0.00 > 0.05$). After that the main effect of instructional strategies was noted with Pillai trace = 0.50, F (6, 59.01) and $\rho = 0.00 > 0.05$.

Follow-up by test between subjects effects (univariate ANOVA) in instructional visualization and instructional strategies. Following Huck recommendation regarding significance MANOVA, subsequent exploratory analysis was conducted to further examine the differences. The results of the exploratory follow-up analysis using univariate ANOVA are presented in Table 3.

Follow-up tests of between-subjects effects presented in Table 3 found significant differences in achievement among students who received different types of instructional visualization on the three-criterion test.

Follow-up tests of between-subjects effects presented in Table 3 found significant differences in achievement among students who received different type of instructional Strategies on the three-criterion test.

Table 3 Test between subject effects of instructional visualization and instructional strategies on three-criterion test

Experimental group	Test by treatment	df	F	Sig.	Partial eta square
Instructional visualization	Identification	2	142.37	0.00[a]	0.35
	Terminology	2	94.43	0.00[a]	0.26
	Comprehension	2	44.72	0.00[a]	0.14
Instructional strategies	Identification	2	16.061	0.000[a]	0.056
	Terminology	2	107.511	0.000[a]	0.286
	Comprehension	2	98.470	0.000[a]	0.268

Note [a]Mean difference significance and each of the criterion tests contains 20 items

Adjusted means of type of instructional visualization and instructional strategies on three-criterion test. It was found from the adjusted mean difference that students who used interactive visualization outperformed the static and animated visualization group with (mean = 12.74) and students who used animated visualization outperformed (mean = 9.67) the static visualization (mean = 8.55) in identification test, which measures factual knowledge. In terminology test which measures conceptual knowledge interactive visualization outperformed the static and animated visualization group with (mean = 13.15) and students who used animated visualization outperformed (mean = 11.91) the static visualization (mean = 9.27). In comprehension test which measure rules and principal knowledge interactive visualization outperformed the static and animated visualization group with (mean = 13.66) and students who used animated visualization outperformed (mean = 12.86) the static visualization (mean = 11.05).

It was found from the adjusted mean difference that students who used question + feedback strategies outperformed the no question and question strategies group with (mean = 11.20) and students who used question strategies outperformed (mean = 10.41) than no strategies (mean = 9.34) in identification test, which measure factual knowledge. In terminology test which measures conceptual knowledge question + feedback group outperformed the no question and question strategies group with (mean = 13.88) and students who used question strategies outperformed (mean = 11.49) the no strategies (mean = 8.96). In comprehension test which measure rules and principal knowledge question + feedback group outperformed the no question and question strategies group with (mean = 14.93) and students who used question strategies outperformed (mean = 12.18) the no strategies (mean = 10.46).

7   Significance of the Study

The major conception of 'learning by doing' is the active participation of learner in the overall instructional phase. This research proves that interactivity in learning material enhance learning achievement of students by engaging them throughout the learning process. However, one of the practical contributions from the study results in providing a pathway and reinforces to use interactivity-based learning material in the classroom environment.

A growing phenomenon is the use of interactive instruction and computer application in educational environment, still there may be questions regarding inclusion of new elements of what types of features are more effective. On a practical level, this research finding provides a significant road map for instructional designer that virtual manipulation features in interactive visualization activating motor cortex and inclusion of dragging features rather than pre-programmed animation can enhance teaching learning.

Question and feedback strategies in the study are used as additional/complementary to interactive onscreen visual instruction. The results of this study provide a deeper view for instructional designer regarding question and feedback strategies for effective visual onscreen teaching learning process.

References

1. Mukhopadhyay, M.: Indian education: overview, issues and IT intervention, (2004).
2. Govindaraju, R., Venkatesan, S.: A Study on School Drop-outs in Rural Settings. J Psychology. 1, 1, 47–53 (2010).
3. Guemide, B. and Benachaiba, C.: Exploiting ICT and E-Learning in Teachers Professional Development in Algeria: the Case of English Secondary School Teachers. Turkish Online Journal of Distance Education-TOJDE, 13, 1302-6488 (2012).
4. Halder, S., Saha, S., and Das, S.: Computer based Self-pacing Instructional Design Approach in Learning with Respect to Gender as a Variable, In Mandal et al (Eds), Information Systems Design and Intelligent Applications, 37–47, India: Springer doi:10.1007/978-81-322-2247-7_5 (2015).
5. Schwan, S., Riempp, R.: The cognitive benefits of interactive videos: learning to tie nautical knots. Learning and Instruction. 14, 293–305 (2004).
6. Lowe, D.: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision. 60, 91–110 (2004).
7. Boucheix, J., Lowe, R., Putri, D., Groff, J.: Cueing animations: Dynamic signaling aids information extraction and comprehension. Learning and Instruction. 25, 71–84 (2013).
8. Mayer, R., Chandler, P.: When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages?. Journal of Educational Psychology. 93, 390–397 (2001).
9. Agina, A., Kommers, P., Steehouder, M.: The effect of the nonhuman external regulator's answer-until-correct (AUC) versus knowledge-of-result (KR) task feedback on children's behavioral regulation during learning tasks. Computers in Human Behavior. 27, 1710–1723 (2011).
10. Terzis, V., Moridis, C., Economides, A.: The effect of emotional feedback on behavioral intention to use computer based assessment. Computers & Education. 59, 710–721 (2012).
11. Kornell, N.: Attempting to answer a meaningful question enhances subsequent learning even when feedback is delayed. Journal of Experimental Psychology: Learning, Memory, and Cognition. 40, 106–114 (2014).
12. Dwyer, F.: Strategies for improving visual learning. Learning Services, State College, PA. (1978).
13. Anastasi, A.: Psychological testing. Macmillan, New York (1988).

Test Case Reduction Using Decision Table for Requirements Specifications

Avinash Gupta, Anshu Gupta and Dharmender Singh Kushwaha

Abstract Majority of the software development cost is incurred in software testing. Most often, the testing of the software is carried out after the code has been prepared and the test cases are obtained from the code. This approach may work well but shall not guarantee that all the requirements are incorporated in the code and that each of the critical paths has been tested. In early phases of software development, decision table is used for test case generation for functional testing. This paper proposed a technique for automated test case optimization generated through decision table. In this paper a framework for test case generation from decision table generated form SRS, and an algorithm for decision table optimization, has been proposed.

1 Introduction

Software testing is a crucial job of software development life cycle that ensures software quality and precisely impacts the development cost and progress of the software. Numerous test case generation techniques have been proposed by researchers for different programming paradigms such as object-oriented, procedural, component-based, web applications, aspect, and database. It is crucial to find the most efficient and effective test design procedure that is possible to automate and apply. An important research problem, in this context, is to have a relevant technique

Avinash Gupta (✉) · Anshu Gupta · D.S. Kushwaha
Department of Computer Science and Engineering, MNNIT Allahabad,
Allahabad, India
e-mail: avinashg.mnnit@gmail.com

Anshu Gupta
e-mail: 14anshug@gmail.com

D.S. Kushwaha
e-mail: dsk@mnnit.ac.in

for the generation of test cases that would minimize both the testing time and effort without sacrificing the thoroughness of testing based on the user requirements.

Though decision table is an old technique for testing, these are increasingly being used by IT analysts and developers to manage, integrate, and execute complex logic more effectively.

2 Related Work

Testing process is an essential part of the entire software development and requires higher than 50 % of the net cost incurred on the total development of the application [1]. It also becomes necessary to make an early estimation of the testing effort [2] as later the errors detected, higher the cost to rectify it. Hence requirement analysis and specification is recognized as an important phase of software development and has to be handled very carefully [3]. Jamoussi [4] proposes software systems like process-control systems, require huge amount of testing to measure their reliability as according to the reliability requirements enforces on them. Hence, it is necessary to enhance the test generation process to slice the certification time. Zhou et al. [1] suggested that the important task during software testing is the generation of relevant test data. Many techniques have been suggested to automate such task. Keiji et al. [5] suggested that appropriate test cases generation is costly and difficult, often for testing those software systems whose input is structurally complex. DeMilli et al. [6] suggested a different technique for automatic generation of test data, which is based on mutation analysis and creates test data that approximates relative adequacy. Korel [7] suggests a different method of generation of test data based on execution of the code under test, dynamic data flow analysis, and function minimization methods. Engström et al. [8] proposed a method for automatic generation of unit test data for branch coverage using symbolic execution, mixed integer linear programming, and execution trees. Melvin et al. [9] propose test automation effort estimation and test execution framework for the test case selection on the basis a controlled natural language, uses a manual coverage and automated test case generation technique for effort estimation [10].

Wang et al. [11] use induction of ordinal classification rules, which assign objects to preference ordered decision classes. These relationships are used by decision rules induced from converted corresponding decision tables. Authors in [12–15] have proposed various methodologies based on requirements and design phase that aim at reducing test effort along with handling change management.

3 Proposed Work

The proposed approach aims at designing the test cases after the requirements and definition phase of the SDLC model. It takes the natural language requirements as input and extracts the necessary conditions/actions from these requirements.

The conditions/actions are used to design the decision tables which are further used to determine the rules and generate test cases. An algorithm has been proposed so that these test cases are further optimized and can be applied on the code for unit testing, thus reducing the overall regression test effort.

3.1 Proposed Framework

The proposed approach takes the natural language requirements [16] as input and extracts the necessary conditions/actions from these requirements.

The conditions/actions are used to design the decision tables which are further used to identify the rules and generate test cases. After that the data which are tested is determined and the generated test cases are applied on the code for unit testing which is generated in the implementation phase and the test output is generated. The proposed work is represented through the diagram shown in Fig. 1.

Steps involved in the proposed work are elaborated here.

Step 1: Functional Requirements Analysis and Condition/Action Determination: This step has been divided into 2 sub steps that are:

Step 1.1: Functional Requirements Analysis.
Step 1.2: Condition/Action Determination:

Step 2: Input Generation for Rule Deployment and Testing: This step has been divided into 3 sub steps that are:

Step 2.1: Create Decision Table: We map the conditions and actions into spreadsheet to generate the decision tables which tells us what action needs to be performed when a certain condition is met.
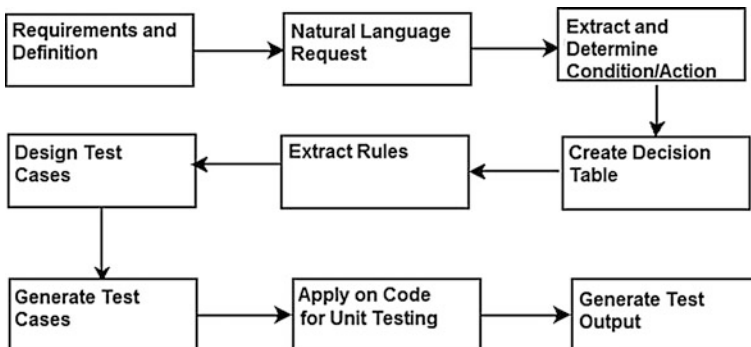
Fig. 1 Idea of the proposed work

Step 2.2: Create Corresponding Rules: Using the decision table formed in the previous step, we create the rules by identifying the various objects related to the requirements specification and the relationship between these objects. The rules are presented in the spreadsheet in the form acceptable by the OpenRules tool used in the Step 3.

Step 2.3: Define Test Data:

Step 3: Rules Deployment and Testing Using OpenRules: This step has been divided into two sub steps that are:

Step 3.1: Rules Deployment and Testing using OpenRules: All the results of the previous three steps, namely decision tables, rules and the test data are passed to the OpenRules tool [17] for their deployment. Now, the tool will deploy these decision tables and rules and use the defined test data to verify the decision tables. The verification is done by checking whether the desired/expected output is obtained or not.

Step 3.2: Generation of Test Output: The test output is generated after the rules deployment and testing using the test data.

3.2 Algorithm—Decision Table Optimization

1: Reading the decision table from a .txt file
2: Adding an extra flag column
3: Traversing each row
  Each row is picked and is checked with each other row of decision table to see whether this row can be merged or not.
4: Checking condition of optimization
  For row x and y,
  If (action(x) == action(y))
    If (difference in condition in row entries(x, y) == 1)
      Then goto step 5
5: Merge the two rows and make the entry in the differing condition as "don't care condition" and mark the entry in flag field as "t" and goto step 2 (to pick other row) else goto step 6
6: Repeat steps 4 & 5 for each non-merged row remaining in the table. If no other row merges with this row, then goto step 2 and start the above process of optimization again with new decision table till there is no chance of optimization left.
7: Percentage optimization is calculated from the original decision table and optimized decision table.
  %optimization= (no of row in original table - new table) / (in original table) *100

At least one test case shall be written for one entry of the decision table. The quantum of optimization achieved is elaborated here.

Let number of test case (TC) in original table before optimization = TC_B
Let number of test case in optimized table = TC_O
Let reduction in test case = TC_R
Therefore percentage reduction in test case $TC_R = (TC_B - TC_O)/TC_B * 100$

3.3 Comparison

In order to establish the results obtained from open source rule engines that proposes framework for generation of test suites, decision table optimization algorithm is also proposed. A decision table is created and then through the proposed decision table optimization algorithm, the entries of the tables are analyzed and reduced to remove the redundant conditions that are present resulting in a minimum size of test cases required to test any application, without compromising on the test cases to be generated ensuring complete code coverage. This validates the proposed approach (Tables 1 and 2).

This allows testers to make an early estimation of errors and thus, reducing the overall testing cost and time. Moreover, for this type of method since the testing is done using the requirement specification it does not require the tester to have the knowledge of coding or programming logic.

Table 1 Test cases (sample) for code-based approach

Sr. No	Emp_Age	Emp_Yos	Emp_Gender	Emp_MaritalStatus	Leave
1	>= 18 years <22		Female	Unmarried	12 CL, 5ML
2	>= 22 years <45	>= 0 years <10	Female	Unmarried	12 CL, 10EL, 5ML
3	>= 22 years <45	>= 10 years <20	Female	Unmarried	12 CL, 15EL, 10ML
4	>= 22 years <45	>= 20 years	Female	Unmarried	12 CL, 15EL, 15ML
5	>= 45 years <60	>= 0 years <10	Female	Unmarried	12 CL, 15EL, 10ML
6	>= 45 years <60	>= 10 years <20	Female	Unmarried	12CL, 15EL, 15ML
7	>= 45 years <60	>= 20 years	Female	Unmarried	12CL, 15EL, 30ML
8	>= 60 years		Female	Unmarried	12CL, 15EL, 30ML
9	>= 18 years <22		Female	Married	12 CL, 15ML
10	>= 22 years <45	>= 0 years <10	Female	Married	12 CL, 10EL, 15ML
11	>= 22 years <45	>= 10 years	Female	Married	12 CL, 15EL, 30ML
12	>= 45 years		Female	Married	12CL, 15EL, 30ML

Table 2 Optimized test cases (sample) for code proposed approach

Sr. No	Emp_Age	Emp_Yos	Emp_Gender	Emp_MaritalStatus	Leave
1	>= 18 years <22		Female	Married	12 CL, 15ML
2	>= 22 years <45	>= 0 years <10	Female	Married	12 CL, 10EL, 15ML
3	>= 22 years <45	>= 10 years	Female	Married	12 CL, 15EL, 30ML
4	>= 45 years		Female	Married	12CL, 15EL, 30ML
5	>= 18 years <22				12 CL, 5ML

Having optimized the table, percentage optimization, i.e., reduction in test cases is calculated from the original decision table and optimized decision table as:

Let number of test case (TC) in original table before optimization = TC_B
Let number of test case in optimized table = TC_O
Let reduction in test case = TC_R
Therefore percentage reduction in test case $TC_R = (TC_B - TC_O)/TC_B * 100$
$= (28 - 12)/28 * 100$
$= 57.14 \%$

The proposed work is useful in those scenarios where there is a need to make an early estimation of errors and thereby reducing the testing effort. It also helps to reduce redundancy that might creep in the test cases when generated manually and thus, reducing the size of test cases. Since this approach uses the decision table for test case generation method, minimum hitherto complete set of test cases required to test the application is achieved. Here, the test cases are generated by using the requirements specification hence, making it possible to generate the test case without being bothered of the implementation coding. Only for the sake of completeness, the code is illustrated here.

4 Performance Analysis

For the performance analysis, the approach is implemented on few programs so as to ascertain the performance of the proposed approach. After comparing the number of test cases required after optimization (TC_O) with the number of test cases in actual table before optimization (TC_B), it is found that the proposed test case optimization is able to lower the size of test cases on an average by 27.8 % (Table 3).

Table 3 Analysis of test case reduction

Sr. No.	Project	Number of test case (TC) in actual table before optimization (TC_B)	Number of test case in optimized table (TC_O)	Reduction in test case (TC_R) %
01	Branch coverage	10	5	50
02	Determine employee leave	28	12	57.14
03	Leap year	6	6	0
04	Patient therapy	14	10	40
05	Business customer module	38	34	11.7
06	Loan prequalification	28	22	27.2
07	Income tax calculator	27	22	22.7

5 Conclusion

Most often, the testing of the software is carried out after the code is ready. This is because the test cases are obtained from the code. In the proposed framework, the rules are kept in the form of decision tables. Each of such row demands a test case to be written. This greatly lowers the chances of test case omission. For the performance analysis, the approach is implemented on few programs so as to ascertain the performance of the proposed approach. It is found that the proposed test case optimization is able to lower the size of test cases on an average by 27.8 %.

References

1. Z. Zhou, B. Scholz, and G. Denaro. Automated software testing and analysis: Techniques, practices and tools. In System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference, IEEE, 2007.
2. Aprna Tripathi, D.S. Kushwaha and Prof. Arun Misra. 'ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India-Vol II'. Advances in Intelligent Systems and Computing (2014).
3. Sharma, Ashish, Manu Vardhan, and Dharmender Singh Kushwaha. 'A Versatile Approach for the Estimation of Software Development Effort Based On SRS Document'. Int. J. Soft. Eng. Knowl. Eng., Volume 24, Issue 01, February 2014, pp. 1–42, doi:10.1142/S0218194014500016.
4. A. Jamoussi. An automated tool for efficiently generating a massive number of random test cases. In High-Assurance Systems Engineering Workshop, 1997, pages 104–107. IEEE, 1997.
5. Keiji Uetsuki; Tohru Matsuodani; Kazuhiko Tsuda, "An efficient software testing method by decision table verification", Inderscience International Journal of Computer Applications in Technology (IJCAT), Vol. 46, No. 1, 2013.
6. R. DeMilli and A. J. Offutt. Constraint-based automatic test data generation. Software Engineering, IEEE Transactions on, 17(9):900–910, 1991.
7. B. Korel. Automated software test data generation. Software Engineering, IEEE Transactions on, 16(8):870–879, 1990.
8. E. Engström and P. Runeson. A qualitative survey of regression testing practices. In Product-Focused Software Process Improvement, pages 3–16. Springer, 2010.
9. Melvin Philips, Nikhil Pawar, Nitesh Joshi, Sanket Khandebharad, Sunil Deshmukh & Kailash Tambe, "Automated Test Case Generation Using Multiple Modelling Techniques", International Journal of Science and Research, Volume 3 Issue 3, March 2014.
10. Ajitha Rajan, Michael W Whalen, Mats P. E. "Heimdahl Model Validation Using Automatically Generated Requirements-Based Test", 10th IEEE-High Assurance Systems Engineering Symposium, 2007.
11. Hailiang Wang, Mingtian Zhou & Kun She, "Induction of ordinal classification rules from decision tables with unknown monotonicity", European Journal of Operational Research, 242 (2015) 172–181, http://dx.doi.org/10.1016/j.ejor.2014.09.034, 2015.
12. D.S. Kushwaha and A.K Misra, "Software Test Effort Estimation", ACM SIGSOFT Software Engineering Notes, Vol. 33, No. 3, May 2008.
13. Aprna Tripathi and Dharmender Singh Kushwaha, "Package Level Coupling: A New Metric for Object-Oriented Software", 1st IEEE International Conference on Recent Trends in Computer Science and Engineering, Patna, Feb 8–9, 2014.

14. Prateek Khurana, Aprna Tripathi and Dharmender Singh Kushwaha, "Change Impact Analysis and its Regression Test Effort Estimation", 3rd IEEE International Advance Computing Conference (IACC-2013), Ghaziabad, February 2013.
15. Rajat Swapnil, Aprna Tripathi and Dharmender Singh Kushwaha, "Software Change Validation Using Class Diagram and SRS", 3rd IEEE International Advance Computing Conference (IACC-2013), Ghaziabad, February 2013.
16. Sharma, Ashish, and Dharmender Singh Kushwaha. 'Natural Language Based Component Extraction from Requirement Engineering Document and Its Complexity Analysis'. SIGSOFT Softw. Eng. Notes 36.1 (2011).
17. OpenRules® architecture and related components http://openrules.com/architecture.htm.

Resource Management Using ANN-PSO Techniques in Cloud Environment

Narander Kumar and Pooja Patel

Abstract In the cloud environment, multiple requests are coming from the client on the datacenters. We have to assign the resources to all the requests. In this paper, the main objective is to find out suitable mapping between requests and resources. To do this we are using the artificial neural network (ANN) with the PSO algorithm. In this algorithm input layer (client) sends request with some requirement. According to requirements we calculate the resources cost on the behalf of the three clusters namely high, medium, and low. Since ANN supports the parallel processing, so we can process all the requests whether they belong to high, medium, and low, hence we optimize the processing time and cost also. PSO algorithm works on the hidden layer as a scheduler. Since particle swarm ptimization (PSO) algorithm supports fast convergence and time constraints, etc. Therefore both techniques minimize the cost and increase the availability and the reliability as well as results show improved performance.

Keywords ANN · PSO · Resource mapping · Swarm intelligence · Velocity

1 Introduction

In the current scenario, increasing the workload day by day in the cloud computing environment due to use of internet is to make a backbone of our daily life. There are required such technique which manages the resources as well as other parameters like cost, reliability–availability, load balancing, taken less time consuming. If one data center has taken more time then user may switch to other data center, so we can say this is a challenge in business perspectives. To solve the above discussed issues

Narander Kumar (✉)
Department of Computer Science, BBA University (A Central University), Lucknow, India
e-mail: nk_iet@yahoo.co.in

Pooja Patel
Department of Computer Science, Banasthali University, Vanasthali, Rajasthan, India
e-mail: ppptl17@gmail.com

we proposed a mechanism using artificial neural network (ANN) technique with particle swarn optimization algorithm. ANN-PSO is a technique for solving the problem which is different in nature as well as same type of requests that come to the data centers to avail the services since ANN supports high processing capability our proposed system provides the services. One main important feature of ANN structures is it working in dynamic environment also. In this paper we use an ANN structure with three layer concept, i.e., input, hidden, and output. On the input layer we make cluster of requests, i.e., high, medium, and low. At the hidden layer the PSO algorithm has work to find the resources as per requirements of the coming requests. Our main aim is to provide minimum cost to incoming requests/user for availing the services of data centers which are beneficial to consumer and service provider.

The organization of this research paper is as follow: Sect. 1 presents the introduction. Related work has been given in Sect. 2. In Sect. 3 we described the particle swarm optimization technique. Sect. 4 gives the mathematical formulation. Working example has been given in Sect. 5. Sect. 6 describes the results and discussion. Conclusions are given in Sect. 7.

2 Related Work

There are many more techniques available in the literature so there isone of the first works which provides cost-aware scheduling algorithm that has been applied on PSO that works in cloud computing environment has been discussed in [1]. Working on total cost of resources minimizes by using the fitness function and the main objective is to minimize the use of resources as well as decreasing performance overhead has been described in [2]. The application of selective algorithm has been applied on the efficient resources provisioning as working on the min–min and max–min algorithm is used for the allocation of the resources and used task scheduling algorithm [3, 4]. The Ipso algorithm is to improve the task scheduling in cloud computing environment and main work is to describe using PSO with the simulated annealing algorithm. IPSO is used to achieve the optimal virtual machine with the help of using resources scheduling in cloud computing environment to find out the effective solution to solve complexity on the computation of the virtual machine in [5, 6, 7]. The heuristics algorithm for managing the resources and allocation of the resource with energy aware in data centers some parameters of the quality of services. The objective is to minimize the cost, and having high potential to improvement of energy efficiency under the dynamic workload has been described in [8]. BPSO, i.e, modified butterfly particle swarm optimization obtained the optimal solution and doing the comparative study between ant colony and honey bee algorithm gives the optimal result in normal load condition and implemented the simplest round robin algorithm that works in idle in [9]. Deadline limits

that on that basis to find out the score. That score helps to manage the less execution time and reduces the failure rate in application of workflow and using simulation that has been done on the cloudsim toolkit is given by [10]. To manage the load in a cloud computing environment on this work enhancement of the load balancing algorithm with ant colony algorithm has to be done[11]. Financial plans and timeliness parameters are considered to achieve the fault tolerant load balanced scheduling algorithm in computational grids for the user satisfaction that working in heterogeneous dynamic cloud environment discussed in [7, 12–14].

As above extensive review of the literature, we find that there is a lack of such technology which incorporates the ANN and the priority-based technique in cloud environment. In this paper we propose such system model which manages the resources using ANN and priority queue technique in cloud computing.

3 Particle Swarm Optimization

The concepts of PSO are that taking from the concept of swarm intelligence analyzing the behaviors of organisms like bees, ants, fish, and other species. Our main motive is a dynamic model for the cloud computing. PSO is a fast convergent algorithm, so it helps in management of resources and assigning the resources to the requests. PSO is a latest evolutionary algorithm that helps for the optimization and management of resources. In the proposed model we are using the PSO as scheduling algorithm in ANN that helps in solving the different issues of cloud computing environment. PSO algorithm efficiently works in D-dimension space. The equations for calculating the positions and velocity for each resource are as follows:

$$V_i[t+1] = W_i V_i(t) + C_1 \operatorname{rand}(.)(P_i - X_i) + C_2 \operatorname{Rand}(.)(P_g - X_i)$$
$$X_i[t+1] = X_i[t] + V_i[t+1]$$

where V_i is the velocity of resources, W_i is the inertia of weight that helps to find the optimum solution, X_i is the position of the resources, P_i is the local best position, P_g is the global best position, Rand and rand normalized uniform distribution function which ranges between 0 and 1.

4 Mathematical Formulation

Initially we apply the clustering algorithm to categorize the request according to their resources cost. Resources cost is calculated by the requirement of the users. When users send the requests to the datacenters then it will send all parameter requirements on the basis of the requirement we calculate for the resources cost.
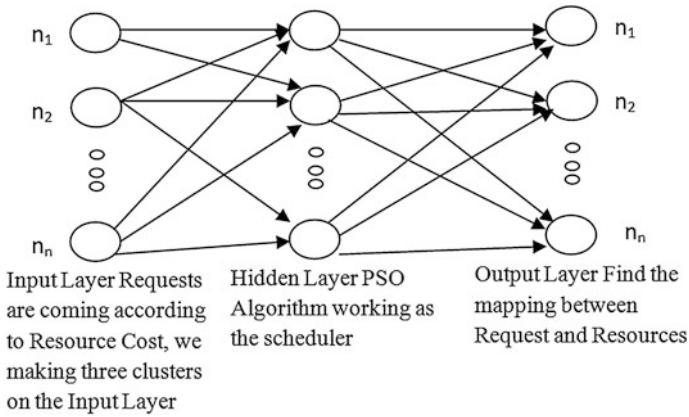
Fig. 1 System model

According to the resources cost we make a cluster. In this paper the cluster is categorized into three types high cost, medium cost, and low cost (Fig. 1).

Now we assign these clusters in three queue. In this paper we use the artificial neural network as a scheduler to allocate the jobs or request to the resources (Tables 1 and 2).

Each request for resources their cost is denoted as RESOURCES_COST

$$\$RESOURCES\_COST = \$COST(COMPUTE) + \$COST(NETWORK) + \$COST(STORAGE) \tag{1}$$

Table 1 Instances types taken from Amazon EC2

Instance type	Disk (GB)	Memory (GB)	No. of cores	$Hour
M1.SMALL	17	160	1	0.85
M1.LARGE	7.5	850	2	0.34
M2.2XLARGE	34.2	850	4	1
C1.MEDIUM	1.7	350	2	0.14
CC1.4XLARGE	23	1690	8	1.6
CG1.XXLARGE	23	1690	8	2.1

Source aws.amazon.com/ec2/instance.types

Table 2 Instances types

Instances type	Usage (h/day)	Disk (GB)	Instances	Memory (GB)
Linux on t1.micro	24	32	1	34
t2.micro	24	15	2	64

where

COMPUTE$(C) = [COST [VMTYPE] × hrs]
NETWORK$(C) = [COST [PER HOUR] × hrs]
STORAGE$(C) = [COST [PER MONTH × STORAGE SIZE] × MOUNTH_HRS]

4.1 Algorithm

```
Initially assign high=0; medium=0; low=0
For (i=1; I ≤ request; i++)
     If ($RESOURCES_COST is high)
     Cluster_1=request;
End for
For (i=1; I ≤ request; i++)
     If ($RESOURCES_COST is medium)
     Cluster_2=request;
End for
For (i=1; I ≤ request; i++)
     If ($RESOURCES_COST is low)
     Cluster_3=request;
End for
```

Now these clusters assign as an input node in ANN. Input—we assign request of cluster on input node and we use three nodes called as neuron. Output—on the output layer mapping is performed between the request and resources. At input node when assigning request and applying some weight we find from the REQUEST_COST and applying bias

$$\text{hidden}_j = g\left(\sum_{i=1}^{l} \text{weight}_{ji} x_{\text{input}\,i} + \text{bias}_j\right), \quad 1 \leq j \leq p \tag{2}$$

```
  where, The weight_{ji} = $REQUEST_COST;
     x_{input} = 1, if request comes;
          = 0, if request not comes;
     bias_j = for the adjustment of weight;
          p = number of request;
```

Now at the hidden layer a scheduler algorithm is working for assigning the resources to the request. In this algorithm PSO working at the hidden node it's main objective is to find the minimizing cost.

4.2   Particle Swarm Optimization Technique

PSO is working on the hidden layer working as a scheduling algorithm

$$Y_{\text{Koutput}} = F\left(\sum_{j=1}^{q} \text{weight}_{kj} \times \text{hidden}_j + \text{bias}_k\right), 1 \le k \le m \tag{3}$$

where f is described as a sigmoid function (i.e, is a nonlinear function), weight_{kj} denotes the connection weights between the hidden nodes and output node, respectively, $\text{weight}_{kj} = \text{fitness(function)}$, $\text{bias}_k =$ helps to adjust the weight so that we can find out the optimal solution.

$$\begin{aligned}\text{Fitness(function)} =&\alpha\, \text{COST}_{\text{totalx}}(x) + (1 - \alpha)\$\text{RESOURCES\_COST} \\ &+ (1 - 2\alpha)\text{MAKESPAN}_{\text{totalx}}(x) \quad 1 \le \alpha \le 1\end{aligned} \tag{4}$$

α a weight given to total cost
$1 - \alpha$ weight given to $RESOURCES_COST
$1 - 2\alpha$ weight given to the makespan

$$\text{COST}_{\text{total}}(R_i) = \text{cost}_{\text{exei}}(R_i) + \text{cost}_{\text{transi}}(R_i) \quad 1 \le i \le x \tag{5}$$

Here we considered the $\text{cost}_{\text{total}}(R_i)$ is defined as the total request processing cost mapping with resources:

$$\text{COST}_{\text{totalx}}(x) = \sum_{i=1}^{x} \text{COST}_{\text{totalx}}(R_i) \tag{6}$$

and
    $\text{MAKESPAN}_{\text{totalx}}(x) =$ finish time of the Request $-$ start time of the first request.
    Here according to the problem the PSO algorithm works on the hidden layer. So here particles refer to the mapping between the request and resources. Particles in the search process update themselves by observing the known the best positions. one best known position called as local best position which works as the individual best known position in terms of their fitness value reached so far by the particle itself. Another best known position known as global best position is the best position in entire population (means that in data center having request and resources are processing). Here velocity is referred as a speed of finding the mapping between the request and resources.
    Input = The processing request according to their $RESOURCES _COST
    Output = on the output layer we get the global best position of particles (resources mapping with request).

Algorithm:

```
1.  Set iteration generation I=1
2.  Set the input on ANN applying on the cluster i.e.
    high, medium, and low.
3.  Set weights as a resources cost between the input
    and hidden layer.
4. Now PSO algorithm is work as a scheduling algorithm
    on the hidden layer.
5.  In this algorithm we take the random value of the
    velocity and particles in D- dimension.
3.  Now checking the request in the queue, the request
    is available or not.
4.   For  each  resource  mapping  to  the  request  we
    calculate the fitness value using fitness function.
5.  Now check  the  fitness  function  value. If fitness
    function  value  is  less  than  or  equal  to  the
    expected  fitness  function  value  than  assign  the
    value to request. If not than go to step six.
6.  I=I+1 repeat the step 4 until not get the expected
    fitness function value.
8.  Update the local best position.
9.  each  resource  mapping  to  the  request  update  the
    position and velocity.
10. return  the  global  best  position  i.e.  it  find  the
    best resources mapping with request.
```

Now the output layer of each request is assigned to the resources and gets the feasible solution.

5 Working Example

When a client sends the request with some required need parameters, according to that we estimated $resources cost. For example we had taken some requirements by that we are estimating the cost. Customers can start with Amazon EC2 having following services: 750 h of EC2 running Linux t2.micro instances usage and 15 GB for data processing.

According to that we calculated the t1.micro instances type resources cost on monthly basis, i.e., $14.64 and t2.micro instances type resources cost on year basis, i.e., $600.56. Now that resources cost acts as the weight in between the hidden and input layer. Request comes then we take the value of x=1 otherwise zero and bias value is considered 1. Now it is processing on the hidden layer. On hidden layer particle swarm optimization is working as a scheduler. In PSO algorithm initially choose particles position (mapping between request and resources) and velocity (bandwidth) randomly. Now check whether the stopping criteria is satisfied or not if not satisfied then we calculate the fitness value.

Fitness function value = total processing cost of resources + $(1 - \alpha)$ resources cost according to customer demand + $(1 - 2\alpha)$ Total Makespan time
Fitness function value = $17.56 + $14.56 + 24 h/day

Then check the fitness value with the expected fitness value. If not then it is repeated otherwise, it updates local best position update and updates the global best position and velocity in result it provides the mapping between request and resources.

6   Results and Discussion

In this paper we propose the model of ANN with PSO to find an optimal solution for mapping of resources according to the request of users, and increased the reliability, availability of resources, and less time consuming. When the request has come to the datacenters, first it chooses clusters, i.e., high, medium, and low on the basis of resources cost as the input layer that are given as input to the hidden layer. On the hidden layer we consider various types of parameters that are used to calculate the fitness function. By this fitness function we find out the fitness value that plays a major role in PSO algorithm for mapping the resources with request. ANN is having high parallel processing capacity, so we can handle the overload problem also by improving the algorithm. PSO is also a technique of ANN and having the fast convergent properties. Therefore proposed mechanism has reached the goals that are increase reliability and availability, minimize the time and minimize the cost, and handle overloaded request for the resources that fulfill by the proposed ANN with PSO model.

7   Conclusion and Future Prospects

This resources management scheduling problem has been suggested in many of the literatures. However, these models have some limitations like multiple requests for the same type of resources. So we are using the model ANN-PSO to solve the limitations, i.e., Improve the reliability and availability of resources, and less time consuming. It is important to develop such a model that works as dynamic as well as adaptive algorithm to improve feasible solution of the resource management. When client sends a request to the data center then on the basis of resources cost we find different three clusters based on request priority, i.e., high, medium, and low. That cluster work acts on the input layer further processing on the hidden layer. On the hidden layer PSO scheduler algorithm is used to assign the resources to the request. That mechanism works well and maximizes profits, both customers and the provider. As a future perspective, we improve further ANN-PSO algorithm by using stored previous information of resources taken from the cloud information

services, training it to predict the load coming in future and on the other way, we trained information of resources that information uses to make cluster of those resources that are mostly used or in other words client mostly request for those resources, with this making faster to respond, and less time consuming, and perform less computing. On the other way it may attempt to implement the algorithm by compromising the Quality of Service referred on the SLA as well as different types of computing environment using the other techniques like the operations research techniques.

References

1. Gang Zhao: Cost—Aware Scheduling Algorithm Based on PSO in Cloud Computing Environment. International Journal of Grid and Distributed Computing, Vol. 7, No. 1 (2014), pp. 33–42 http://dx.doi.org/10.14257/ijgdc.2014.7.1.04.
2. Andrew J. Younge, Gregor von Laszewski, Lizhe Wang, Sonia Lopez-Alarcon, Warren Carithers: Efficient Resource Management for Cloud Computing Environments. Bloomington, In USA, Rochester, NY, USA, Email: slaeec@rit.edu, wrc@cs.rit.edu, fajy4490, laszewski, lizhe.wangg@gmail.com.
3. Mayanka Katyal and Atul Mishra: Application of Selective Algorithm for Effective Resource Provisioning In Cloud Computing Environment, IJCCSA, Vol. 4, No. 1, February (2014).
4. Kai Wu: A tunable work flow scheduling algorithm based on the particle swarm optimization for cloud Computing, http://scholarworks.sjsu.edu/etd_project, San José State University SJSU April (2014).
5. Shaobin Zhan, Hongying Huo: Improved PSO-based Task Scheduling Algorithm in Cloud Computing, Journal of Information & Computational Science 9: 13 (2012) 3821–3829, http://www.joics.com.
6. Anisaara Nadaph and Prof. Vikas Maral: cloud computing—partitioning algorithm and the load balancing algorithm, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol. 4, No. 5, October (2014).
7. P. Keerthika, P. Suresh: A Budget and Deadline Constrained Fault Tolerant Load Balanced Scheduling Algorithm for Computational Grids, World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol: 9, No: 2, (2015).
8. Anton Beloglazov, Jemal Abawajy, Rajkumar Buyya: Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing, Future Generation Computer. Systems 28 (2012) 755–768.
9. Aashish Kumar Bohre, Dr. Ganga Agnihotri, Dr. Manisha Dubey, Jitendra Singh Bhadoriya: A Novel method to find optimal solution based on modified butterfly particle swarm optimization, International Journal of Soft Computing, Mathematics and Control (IJSCMC), Vol. 3, No. 4, November (2014).
10. Ranjit Singh and Sarbjeet Singh: Score based deadline constrained workflows scheduling algorithm for the cloud systems, International Journal on Cloud Computing: Services and Architecture (IJCCSA), Vol. 3, No. 6, December (2013).
11. Anisaara Nadaph and Prof. Vikas Maral: Continental Division of load and balanced ant family (BAF) algorithm for the load balancing on public cloud, International Journal on Cybernetics & Informatics (IJCI) Vol. 3, No. 5, October (2014).
12. Hao Yuan, Changbing Li, Maokang Du: Optimal Virtual Machine Resources Scheduling Based on Improved Particle Swarm Optimization in Cloud Computing, Journal of Software, VOL. 9, NO. 3, March (2014).

13. Eun-Kyu Byun, Jin-Soo Kim, Yang-Suk Kee, Ewa Deelman, Karan Vahi, Gaurang Mehta: Efficient Resource Capacity Estimate of Workflow Applications for Provisioning Resources, Information Sciences Institute, USC ƒyskee, deelman, vahi, gmehta*g*@isi.edu.
14. Suraj Pandey, Linlin Wu, Siddeswara Mayura Guru, Rajkumar Buyya: A Particle Swarm Optimization-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments, Cloud Computing and Distributed Systems Laboratory, CSIRO Tasmanian ICT Centre, {spandey, linwu, raj}@csse.unimelb.edu.au, siddeswara. guru@csiro.au.

Congestion Control in Heterogeneous Wireless Sensor Networks for High-Quality Data Transmission

Kakelli Anil Kumar, Addepalli V.N. Krishna and K. Shahu Chatrapati

Abstract Heterogeneous wireless sensor network (HTWSN) is the most preferable and demanding technology for military applications because of low cost and high performance in terms of high-quality data transmission with low end-to-end delay. HTWSN can be established with variable-configured sensor nodes for detection and monitoring the complex and multitasking events efficiently within the network. But congestion is the most serious issue which may cause high packet loss; increase the number of retransmissions, frequent link failure, and node failure; lower the network life time by increasing the energy consumption; and lower the throughput. Most of existing congestion control protocols are developed for homogeneous wireless sensor network which may not help to achieve high throughput for HTWSN. So we have proposed a new congestion control protocol (CCP) for HTWSNs which can estimate the congestion control degree (CCD) at each node prior to identify the future congested nodes in the network. Accordingly, CCP can enable its load balancing technique effectively and balance the data traffic between the future congested nodes and source node to achieve high-quality data transmission in HTWSN.

Keywords Heterogeneous wireless sensor networks · Homogeneous wireless sensor networks · Congestion control protocol · Congestion control degree · Congestion control notification

K. Anil Kumar (✉)
Indore Institute of Science and Technology, RGPV, Indore, MP, India
e-mail: anilsekumar@gmail.com

A.V.N. Krishna
Navodaya Institute of Technology, Raichur, KA, India
e-mail: hari_avn@rediffmail.com

K. Shahu Chatrapati
COE, Manthani, Jawaharlal Nehru Technological University,
Hyderabad, TS, India
e-mail: shahujntu@gmail.com

1 Introduction

Wireless sensor network (WSN) [1] consists of spatially distributed autonomous tiny wireless sensor nodes used for sensing, processing and communication. WSN is having crucial applications like military, environmental, agriculture, wildlife monitoring, and artificial intelligence. WSN is the most suitable technology for military applications for boarder line monitoring, sensing, monitoring, and communication about the enemy objects moment in the sensitive areas. WSN can be deployable in two ways, homogeneous and heterogeneous [2]. The traditional WSN is known as homogeneous WSN (HOWSN). It can be easily deployable with large number of similar low-power and low-cost sensor nodes (LPSNs). After the HOWSN deployment, all the LPSNs form the ground network will be ready to use for sensing and transmitting the event's information to destination node (DN). HOWSN has many advantages such as low cost, easily deployable, node configuration similarity, and availability of more number of multiple paths. Simultaneously, homogenous WSN have many limitations such as sensor nodes in the network cannot handle the complex, huge multitasking events because of low computational power, limited energy, low memory, and low transmission power. Due these limitations homogeneous WSN may not be a preferable technology for military applications. The second type of WSN deployment is heterogeneous WSN (HTWSN) the most preferable and demanding technology in today's world for many important applications such as military, agriculture, environmental, and artificial intelligence. Because HTWSN has few high-power sensor nodes (HPSNs), these nodes are used to sense the high-quality images and video of the moving enemy target objects for long distances in the targeted area. Simultaneously the sensed data which is in large size will be processed and transmitted to DN efficiently. The network that can be established with few number of HPSNs and large number of LPSNs and deployed in a targeted area is known as heterogeneous wireless sensor networks (HTWSN) [3].

1.1 Classification of Data Traffic and Transmission Techniques in HTWSN

Generally the data traffic in HTWSN is classified into four types, event-based, query-based, continuous, and hybrid-based [4]. In event-based data transmission, data transmission can takes place through the HTWSN when the event has occurred. In query-based data transmission, the DN node forwards the several on-demand queries to SN, according to that, SN generates the response and forwards them to the DN through HTWSN [5]. The continuous data transmission is where the SN can transmit the data continuously to DN for regular time interval. Hybrid data transmission can be either event-based and continuous or event- and query-based on user demanding strategy. The data transmission techniques in

HTWSN are mainly classified into three types, one-to-one (1-to-1), many-to-one (N to 1) [6], many-to-few (N to X). In 1-to-1 approach, the data transmission can take place between single SN and single DN in the network. This approach can result high quality data transmission (HDT), and low performance when multitasking is needed. In N-to-1 approach, the data transmission takes place between multiple SNs and single DN. This approach gives better performance than 1-to-1, but results high transmission overhead at DN or at neighbor nodes due to bottleneck data transmission. In N-to-1 approach, the data traffic will become highly convergent and leading to the formation of upstream data traffic congestion at DN or its neighbor nodes. This approach results high congestion, quick node failure, and low network life time. The third type of approach N-to-X is many-to-few where the data transmission can take place between many SNs and few DNs in HTWSN. This approach is highly helpful for minimization of effects of bottleneck data transmission at DN, due to the existence of multiple DNs. Establishment of HTWSN with multiple DNs is the best approach for achieving HDT. The major challenge of this approach is high cost and effective routing establishment from SNs to DNs.

2 Congestion in HTWSN

The multi path routing in HTWSN [7] can discover multiple routing paths between SN to DN for transmission of high-quality data. It is having many advantages such as efficient load balancing, efficient bandwidth utilization, low end-to-end delay with high fault tolerance over single path routing. Single path routing can discover only single routing path between SN to DN [8]. There are two types of types of congestion mainly occur in HTWSN, one is node level and the other is link level [6]. In HOWSN, the congestion is mainly occurs at the last hop nodes of the network, which are the neighbored nodes of DN or at DNs due to heavy convergent traffic from upstream nodes. In HTWSN, the congestion is mainly occurs at the nodes of first hop of SNs due to variable node configurations among them with variable link capacities. The data traffic received by first hop intermediate nodes INs from SNs is represented by upper data traffic (UDT) of first hop INs and the data traffic released by first hop INs to second hop INs is known as lower data traffic (LDT). If the difference between UDT and LDT is high and beyond the node's buffer capacity (BC) results node-level congestion. Once the node-level congestion [9] exits and continues for longer duration, it results many adverse effects in HTWSN such as increasing of pack loss rate, higher latency, huge energy consumption, frequent node, or link failure, lowering the networks life time and performance, wastage of network resources by retransmission of data or route discovery, and poor quality of service (QoS). The link-level congestion is formed because of several nodes in the network are trying to share the same wireless channel simultaneously. Several MAC protocols have been proposed to minimize the link-level congestion or collision for efficient allocation and utilization of wireless channel by the multiple sensor nodes in the network.

3 Congestion Control Protocol (CCP)

Congestion control protocol CCP has several phases to control the congestion in HTWSN environment.

1. Congestion detection
2. Congestion notification
3. Congestion control

3.1 Congestion Detection

Generally congestion detection in HTWSN can be classified into two ways; one is end-to-end approach and the other is link-to-link approach. In end-to-end approach, the end-to-end nodes are responsible for congestion detection and notification for congestion control by using congestion control notification (CCN). In link-to-link approach, any node which can experience the congestion is responsible for congestion detection and notification. End-to-end approach is not preferable for HTWSN because generally end-to-end nodes are not majorly affected by congestion in HTWSN. The link-to-link approach is highly preferable for HTWSN because congested node is responsible for congestion detection and notification. So CC protocol prefers the link-level approach for detecting the congestion in HTWSN [10]. Once the congestion is experienced by any node, prior to it, the node can experience higher value of buffer-free occupancy (BFO) than threshold value defined by CCP. According to the level of BFO, the CCP generates congestion degree indicator (CDI) of the node as congestion notification. After The CDI is generated, it will be released to its upstream nodes towards the SNs. The CDI is generated with the parameters of queue length, packet transmission time or the ratio of packet transmission time over packet arrival time at congested node [11]. In link-to-link approach the congestion control messages have to be transmitted through several multi-hop nodes towards the SN which may cause to increase of unnecessary network overhead and energy consumption. HTWSN is an energy-constraint network, so it is highly essential to minimize the energy consumption [12] required for congestion detection and control message generation and forwarding to SN. EC_{CCN} is energy consumption for congestion control notification (CCN) in joules, $EC_{CDI\ ESTIMATION}$ is energy consumption for the estimation of congestion degree indicator, EC_{CCN} is energy consumption for transmission of CCN, CGN is congested node, SN is source node, PST is total packet service time of node, BCO is buffer capacity occupancy

$$EC_{CCN} = (\text{No. of Hops between CGN to SN}) \\ * [EC_{CDIESTIMATION} + EC_{CCNTransmission}] \qquad (1)$$

$$BFO_{Tn} = Total\ node's\ Memory - BCO_{Tn} \quad where\ 1 \geq n \leq N \qquad (2)$$

$$BCO_{Tn} = PST * Packet\ size \quad where\ 1 \geq n \leq N \qquad (3)$$

$$PST = \sum_{1}^{N}(T_{LastBitSent}\ of\ Packet - T_{FirstBitReceive}\ of\ Packet) \qquad (4)$$

3.2 Congestion Control Degree and Notification (CCD/CCN)

The CCD can be estimated by the CGN, when its BFO value reaches the predefined threshold value α_T. The CCP can be considered standard α_T is 25 %. If BFO of any node is equal to α_T, which indicates the node is getting congested in specific interval of time T [13]. The congestion control notifications CCN are of various types, NO CCN as NCCN, Threshold CCN as CCN_T and high threshold CCN as CCN_{HT}. CCD is congestion control degree can estimated with following expression

$$CCD = (BFO/BCO * 100) \qquad (5)$$

where $CCD < \alpha_T => NCCN, CCD = \alpha_T => Generate\ CCN_T, CCD > \alpha_T => Generate$ CCN_{HT}

3.3 Congestion Control

Congestion minimization is the crucial task of CC protocol. CCP can apply the congestion control mechanism only based on the CCN acknowledgment generated by CGN. If NCCN is generated by the node indicates that the node's buffer memory is free, then no congestion will be formed at the node for particular time duration. If the node generates CCN_T which indicates the node is becoming congested in specified time duration. After CCN_T generation by the node in the network, the CCP will transmit the acknowledgment to all upstream nodes towards the SN, until the data transmission is continuous towards the CGN. Meanwhile the BFO of CGN helps to continue the data transmission without any data loss. Once the CCN_T is received by SN, it can immediately hold the data transmission and estimate the time of hold for data transmission. CCN_{HT} is considered as the high level congestion notification, when it is generated and released to upstream nodes, each upstream node in the routing path can hold the data transmission immediately and save the data packets in its buffer memory and transmit the CCN_{HT} to its upstream nodes. Likewise, CCN_{HT} will receive through multiple upstream hop nodes in route towards the SN. Once SN receives CCN_{HT}, it can immediately hold the data

transmission and estimates that all the hop nodes also hold the data transmission up to CGN. SN and all upstream nodes of CGN can hold the data transmission up to the receiving of cleared congestion notification CCCN by CGN. CCCN can be generated and released by CGN only when CCD is less than α_T. First hop upstream node of CGN receives CCCN from CGN and transmits the data which is stored in buffer memory and estimates its CCD. If the CCD is less than α_T then it forwards the CCCN to second hop upstream node. Similarly all the upstream nodes forward the CCCN among them and finally it reaches SN. Once the SN receives CCCN, then it resumes data delivery towards DN through the routing path. CCP protocol can effectively minimize the congestion for achieving the high-quality data transmission in HTWSN.

4   Result Analysis

All our experiments are conducted in GloMoSim [14] simulator with simulation parameters as follows: Simulation area: 1000 * 1000 m^2; number of nodes: 50–500; types of deployment: random, simulation time; 300–600 s; battery capacity: 2400 mAh; propagation limit: −65 dBm; propagation path loss: free-space; temperature: 290.0 K; radio type: radio acc-noise; radio frequency: 2.4 GHz; radio bandwidth: 512 Kbps to 1 Mbps; radio-Rx-type: SNR-bounded; radio-Rx-SNR-threshold: 10.0 dBm; radio-TX-power: 15.0 dBm (Single source) 10.0 dBm (Multiple source); radio-RX-sensitivity: −91.0 dBm; radio-RX-threshold: −81 dBm; MAC-protocol; SMAC; routing protocol: CCP. From Fig 1, the packet delivery ratio of CCP is high with increase of the packet arrival rate in HTWSN. CCP has given almost 40 % higher than multi-constrained QoS multipath routing MCMP [15]

Fig. 1 Packet delivery ratio of CC protocol in comparison to MCMP and EQSR real-time and non-real-time data transmission
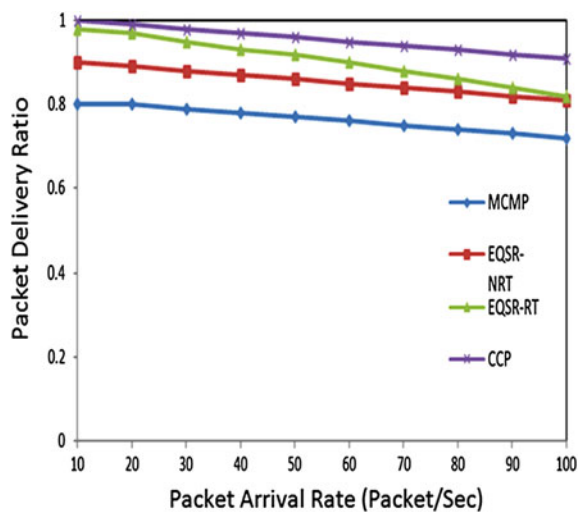
Fig. 2 Energy of consumption of CCP, EQSR and MCMP in joules/sec with packet arrival rate (Packet/sec)
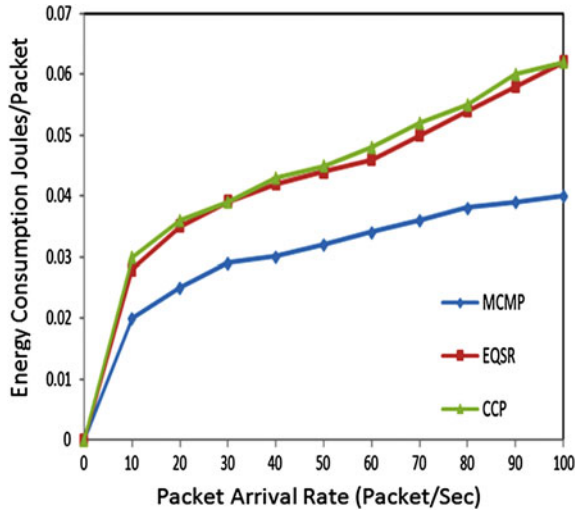
protocol and 30 % higher than energy-efficient and QoS aware multipath routing EQSR protocol [12] for real-time and non-real-time data under heavy traffic conditions. From Fig 2, the energy consumption of CCP is at satisfaction level as compared to MCMP and EQSN. It is observed that the CCP protocol consumed only 5–10 % excess energy consumption than MCMP and EQSR under heavy traffic conditions in HTWSN given high-quality data transmission. From Fig 3, the average end-to-end delay of CCP is very low compares with the MCMP and EQSR real-time

Fig. 3 Average end-to-end delay in seconds with node failure probability of CCP
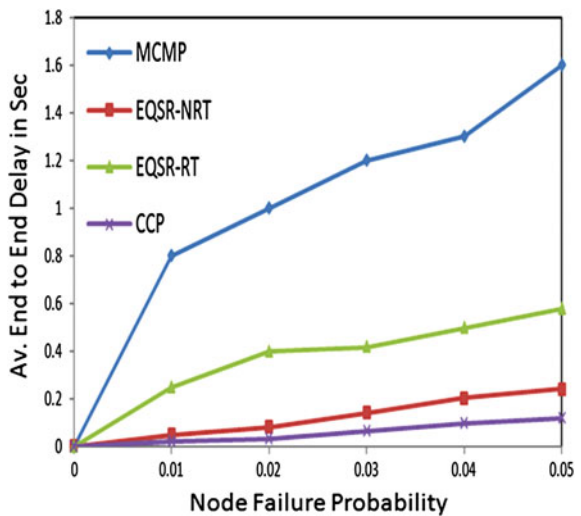
and non-real-time data. The node failure probability is very low compared with CCP, i.e., almost it is negligible. The CCP results indicate that the protocol is efficient in controlling the congestion with minimum energy consumption.

5   Conclusion

CCP is one of the most efficient congestion control protocol for heterogeneous wireless sensor networks under heavy traffic conditions. In HTWSN, the data traffic can be variable due to variable data rates, but CCP has given satisfactory performance even with lower bandwidths. CCP can control the congestion before its existence at the node in the network. CCP has been implemented with pre-congestion estimation and control mechanism, due to that the nodes in active routing path can be protected efficiently from congestion.

6   Future Work

CCP is the most efficient congestion control protocol but its limitation is that, its energy consumption rate has slightly increased up to 5–10 % as compared with existing congestion protocols like EQSR and MCMP under heavy traffic loads. So our future work is to overcome the limitation to make CCP as an energy-efficient protocol.

References

1. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci.: Wireless sensor networks: A survey. Int. J. on Computer Networks. vol. 38, pp. 393–422 (2002).
2. V. Tickoo.: A Comparison Study of Congestion Control Protocols in WBAN. Int. J. on innovations and advancements in computer science. vol. 4, no. 6, pp. 121–127 (2015).
3. K. Anil Kumar.: IMCC protocol in heterogeneous wireless sensor network for high quality data transmission in military applications. In: 1st IEEE International Conference on Parallel, Distributed and Grid Computing, pp. 339–343 (2010).
4. N. Assad, B. Elbhiri, M. A. Faqihi, M. Ouadou, D. Aboutajdine.: Analysis of the Deployment Quality for Intrusion Detection in Wireless Sensor Networks. J. of Computers, Hindawi, vol. 2015 (2015).
5. Mohamed Mubarak, T., Sattar, S. A., Appa Rao, G., Sajitha, M.: Intrusion detection: An energy efficient approach in heterogeneous WSN. In: International Conference on Emerging Trends in Electrical and Computer Technology, pp. 1092–1096, (2011).
6. C.-T. Ee, R. Bajcsy.: Congestion control and fairness for many-to one routing in sensor networks. In: Proceedings of 2nd International conference on Embedded networked sensor systems, ACM, pp 148–161 (2004).

7. M. M. Monowar, M. O. Rahman, C. S. H. C. S. Hong.: Multipath Congestion Control for Heterogeneous Traffic in Wireless Sensor Network. In: 10th Int. Conference on Adv. Communication Technologies, vol. 3 (2008).
8. Chonggang Wang, Kazem Sohraby, Victor Lawrence, Bo Li, Yueming Hu:. Priority-based Congestion Control in Wireless Sensor Networks. In: IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, vol 1 (2006).
9. Md. Ahsanul Hoque, Nazrul Islam, Sajjad Waheed, Abu Sayed Siddique:. Priority based Congestion Control Mechanism in Multipath Wireless Sensor Network. Int. Research J. on Global Journal of Computer Science and Technology, volume 14, issue 5 (2014).
10. Wan, C.-Y., Eisenman, S. B., Campbell, A. T.: CODA: Congestion Detection and Avoidance in Sensor Networks. In: Proceedings of ACM SenSys, volume 5–7, pp 266–279 (2003).
11. Wang, C., Li, B, Sohraby, K., Daneshmand, M., Hu, Y.: Upstream Congestion Control in Wireless Sensor Networks Through Cross-Layer Optimization. IEEE J. on Selected Areas in Communications, vol. 25, no. 4, pp 786–795 (2007).
12. J. Ben-Othman, B. Yahya.: Energy efficient and QoS based routing protocol for wireless sensor networks. Elsevier, J. on Parallel Distrib. Computing, vol. 70, no. 8, pp. 849–857 (2010).
13. A. Sharif, V. M. Potdar, J. D. Rathnayaka.: ERCTP: End-to-End Reliable and Congestion Aware Transport Layer Protocol for Heterogeneous WSN. Scalable Computer. Pract. Exp., vol. 11, no. 4, pp. 359–371 (2010).
14. X. Zeng, R. Bagrodia, M. Gerla.: GloMoSim: a library for parallel simulation of large-scale wireless networks. In: Proceedings. Twelfth Work. Parallel Distributed Simulation, pp. 154–161 (1998).
15. X. Huang, Y. Fang.: Multi constrained QoS multipath routing in wireless sensor networks. J. of wireless networks, volume 14(4), pp 465–478 (2008).

Robust Data Model for Enhanced Anomaly Detection

R. Ravinder Reddy, Y. Ramadevi and K.V.N. Sunitha

Abstract As the volume of network usage increases, inexorably, the proportions of threats are also increasing. Various approaches to anomaly detection are currently being in use with each one has its own merits and demerits. Anomaly detection is the process of analyzing the users data either normal or anomaly, most of the records are normal records only. When analyzing these imbalanced types of datasets with machine learning algorithms the performance degradation is high and cannot predict the class label accurately. In this paper, we proposed a hybrid approach to address these problems. Here we combine the class balancing and rough set theory (RST). This approach enhances the anomaly detection rate and empirical results show that considerable performance improvements.

Keywords Anomaly detection · Imbalanced data · Rough sets · Classification · Intrusion detection

1 Introduction

Recent research, mostly focusing on detecting unknown attacks as well as the existing known attacks is the anomaly-based network intrusion detection. Most of the times, intrusion detection problem is treated as a classification problem [1, 2]. In the process of detecting the anomaly in the system, to classify the data as normal

R. Ravinder Reddy (✉) · Y. Ramadevi
Department of Computer Science and Engineering, Chaitanya Bharathi
Institute of Technology, Hyderabad 500075, Telengana, India
e-mail: ravindra_rkk@cbit.ac.in

Y. Ramadevi
e-mail: yrd@cbit.ac.in

K.V.N. Sunitha
B.V. Raju Institute of Technology for Women,
Bachupally, Hyderabad 500090, Telengana, India
e-mail: k.v.n.sunitha@gmail.com

and anomaly needs a good classification approach. Most of the anomaly detection techniques use classification techniques from the machine learning and data mining. Classification is a supervised machine learning technique. There is a major issue relevant to supervised classification, i.e., class balancing. Because anomalous instances are rare compared to normal instances in the data. The anomaly data is an imbalanced class distribution [3]. The improper distribution of these training data often makes the task of learning more challenging. To address these challenges we proposed a new robust data model.

Robust data selection is a demanding approach for analysis of anomaly detection. The anomaly-based intrusion detection system has become more dependent on learning methods, especially on classifications schemes. For the classification problem, the records should be identically independent distribution is required. Probability distribution should be balanced among the classes and are important. To make the classification more accurate and effective, more robust approaches are required. Data selection and type of input for the classification techniques is very effective on the anomaly detection rate. Classifier accuracy directly depends on the type of input. The input data selection is a key aspect for the anomaly-based network intrusion detection system [4]. The need to build an effective anomaly model robust data approaches is required. In this regard, we combine different data boosting techniques. Preprocessing has considerable impact on the accuracy and capability of supervised anomaly detection.

Accuracy of the anomaly detection depends on the quality and size of the input data used to train the model and its distribution of records. For this purpose, here, we are considering the unbalanced data for anomaly detection, it affects the classifier quality, for that we increased the number of records of the minority class for balancing the given dataset. Balancing the class is very important, it has wide application including image, intrusion detection, etc. Class balancing is an important aspect for improving the quality of the data, which is given to classifier. It will improve accuracy for the anomaly detection. Class balancing will increase the size of the dataset.

Rough set theory [5] is used for reducing the dimensionality of the feature vector. Feature vector size is also a problem for classification, many of the features are not involved in the process of classifying, to remove these features we need to adopt the feature selection technique, here we used rough set approach. In the feature selection process, RST approach produces better results when compared with the traditional principle component approach (PCA). Because PCA needs lot of space and computational time is required for the computation of Eigen vectors. It is difficult when the data size increases.

The remaining topics are organized as follows, in Sect. 2 briefs outline of the related concepts, in Sect. 3 discusses the details of implementation and results are discussed in Sect. 4. Conclusion is in Sect. 5.

2 Related Work

2.1 Intrusion Detection

Intruder tries to break the CIA triangle to penetrate into the system, to get unauthorized access of the system resources. Users need to ensure that authenticated and authorized entities are able to reliably access the secured information. Most of the times these principles are violated by users intentionally and some times without knowingly, the prevention tools can not stop these activities fully. To protect the system securely, we need another layer of protection called the intrusion detection system. Intrusion is an attempt to access the system resources in an unauthorized way to modify or destroy the resources from outsiders or may be the insiders. So intrusion detection system is the second wall of the protection of the system. The firewall will only filter packets. Basically, based on the behavior of intruders it divides two aspects

1. Misuse detection
2. Anomaly detection

2.2 Imbalanced Data

Supervised learning techniques like classification assumed that the training data is balanced and well-distributed. Fewer datasets shows the proper results not all, after the years of research found that class of interest is having very few records and affects the system performance. Most of the real-world datasets are class imbalanced; very few records are representing the main class of interest. This is known as the class imbalanced problem [6–8].

Mainly, class imbalanced problem can be solved using the following methods:

1. Oversampling
2. Under-sampling
3. Threshold moving
4. Ensemble technique

Oversampling and under-sampling are used to increase and decrease the number of tuples in the training set. Oversampling works by resampling the minority classes, so that the resulting training set contains equal number of class samples.

Anomaly detection mainly concerns on the user's abnormal behavior in the system. When the user behavior deviates from the normal, we can say the anomaly. For analysis of this, we need to analyze the user data in a proper manner. For this, very few anomaly records will be available in the system; we need to oversample these records for the analysis of anomaly behavior. Once balanced the dataset it will enhance the classifiers performance.

2.3   Rough Set Theory

Rough set Theory was introduced by Pawlak, it is an extension of conventional set theory [5, 9]. It is used to represent the un-précised and vague data using the approximations called lower and upper approximations. Feature selection property of rough set theory helps in finding reduct of the oversampled dataset. In this way it not only reduces the size of dataset, but also improves the classifier performance.

We adopt the rough set theory to define the necessity of features. Among the existing feature selection techniques rough set approach has significant advantages. Because, it uses the heuristics in the feature selection process. The main purpose of rough sets introduced here for the feature selection.

2.4   Dataset

To evaluate any system we need a benchmark input and compare the results. Fortunately for evaluation of the intrusion detection system we have used The "HTTP dataset CSIC 2010" [10] contains thousands of web requests automatically generated. It can be used for the testing of web attack protection systems.

The main motivation behind this is current problem in web attack detection is the lack of publicly available datasets to test WAFs (web application firewalls). Most of the intrusion detection systems use the DARPA dataset [11, 12]. However, it has been criticized by the IDS community [13]. Regarding web traffic, it is not appropriate for web attack detection. Existing datasets are out of date and do not include many of the attacks. The problem of data privacy is also a concern in the generation of publicly available datasets and is probably one of the reasons why most of the available HTTP datasets do not target real web applications. Because of these reasons, we decided to use the HTTP dataset CSIC 2010.

The HTTP dataset CSIC 2010 contains 36,000 normal requests and more than 25,000 anomalous requests. Each record is labeled as normal or anomalous and the dataset includes attacks such as SQL injection, buffer overflow, information gathering, and files disclosure, CRLF injection, XSS, server side include, parameter tampering, and so on.

3   Methodologies

In this method, we address the two issues regarding anomaly detection. They are, feature selection and balancing the dataset, here we mainly focused on class balancing. Anomaly datasets are imbalanced in class, while using these types of data to train the machine learning techniques like classification, it does not perform well compared to the normal distributed data. Balance the dataset [6, 7] using the data

mining technique, it will improve the prediction rate. In this approach, we increase rare class data using oversampling technique. Here the proposed approach will address these issues.

Algorithm: Hybrid Data sampling
Input: HTTP CSIC dataset
Output: The anomaly prediction rate
1. Preprocess dataset to the required format.
2. Apply the rough set feature selection.
3. Prepare the new dataset with the obtained feature set.
4. Apply the data sampling approach to balance the class label.
5. Redistribute the data tuples.
6. Apply SVM classifier to the refined dataset.

In the process of balancing the dataset, it may increase the size of the dataset in this approach, it will consume system resources to avoid this problem we are applying the rough set approach for reducing the dimensionality of the dataset. The rough set approach enormously reduces the data size without affecting the classifier accuracy.

In this approach, we used smote algorithm for balancing the dataset by over-sampling the minority class. SMOTE (synthetic minority oversampling technique) is used to generate synthetic samples of minority class in order to balance the dataset. SMOTE algorithm [14, 15] considers the minority class instances and oversamples it by generating synthetic examples joining all of the k minority class nearest neighbors. The value of k depends upon the amount oversampling to be done. The process begins by selecting some point yi and determining its nearest neighbor's yi1 to yik. Random numbers from r1 to rk are generated by randomized interpolation of the selected nearest neighbors.

Synthetic samples of minority class can be generated as follows:

1. Consider the minority class feature vector and calculate the difference between with its nearest neighbors.
2. Multiply the difference by a random number between 0 and 1, and add it to the feature vector under consideration.
3. This causes the selection of a random point along the line segment between two specific features.

Once the data sampling is completed we need to re distribute the records. Distribution of samples is also an important issue in the classification process. For the dimensionality reduction, we used rough set approach, in this we used Johnson's reduct, Johnson's algorithm is a dynamic reduct [16] computation algorithm. The process of reduct generation starts with an empty set, RS. Iteratively, each conditional attribute in the discernibility matrix is evaluated based on a heuristic measure and the highest heuristic valued attribute is to be added to the RS and deletes the same from the original discernibility matrix. The algorithm ends

when all the clauses are removed from the discernibility matrix. Pseudo code for Johnson's reduct generation is given below.

Algorithm: Johnson Reduct (Ca, fD)
Input: Ca, the set of conditional attributes,
fD is the Discernibility function.
Output: RS, The minimal reduct set
1. RS ← Ø; bestca = 0;
2. While (discernibility function, fD is not empty)
3. For each c Ca that appears in fD
4. h = heuristic (c)
5. If (h > bestca) then
6. bestca = h;
7. bestAttribute ← c
8. RS ← RS U bestAttribute
9. fD ← removeClauses (fD, bestAttribute)
10. Return RS

The reduct generated by the Johnson's algorithm may not be optimal, still research is going on to find an optimal feature set for a given dataset. Here the HTTP dataset CSIC 2010 is used, it contains the following conditional features. The decision attribute is normal or anomaly.

{index, method, url, protocol, userAgent, pragma, cacheControl, accept, acceptEncoding, acceptCharset, acceptLanguage, host, connection, contentLength, contentType, cookie, payload, label}

Applying the rough set feature selection the 17 conditional features are reduced to 8 features and they are as follows:

{cookie, payload, index, url, contenetLength, method, host, contentType}

In Table 1 we compare the computational time for the rough set model, in Fig. 1 shows that there is huge difference for both the models.

Table 1 Time comparison for the approaches

Technique	Time taken for classify	Time taken to classify with reduct
SVM classifier	**2593**	**1782**

Fig. 1 Computational time comparison
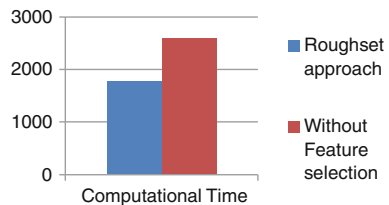
Table 2 Comparison of results with balanced and rough set reduct datasets

Measure	Unbalanced	Balanced	Balanced and rough set reduct
Time	2593	2780	1782
Accuracy	99.44	99.85	99.80
Precision	0.994	0.998	0.998
FP rate	0.005	0.002	0.002
Recall	0.994	0.998	0.998
F-measure	0.994	0.998	0.998

Fig. 2 Performance measures
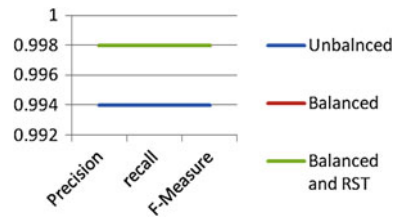
4 Result Analysis

In this model we used the HTTP dataset CSIC 2010, it contains millions of records labeled as normal and anomaly. In this method, we applied the SMOTE algorithm to oversampling the data among the normal and anomaly records. After this randomize the record to distribute the oversampled records throughout the dataset. Once we obtain the optimal feature vector we applied this data to the SVM classifier and calculated the results. We used RBF kernel for evaluation of the results. It performs well on the balanced data and produce good results compared with the other balancing methods.

The empirical results show that using the balancing and rough set reduct improves the classifier accuracy as well as reduces the computational time. As shown in Table 2 false positive rate is also decreased, we compute the precision, recall, and F-measure and those results shows that the robust approach is performing well compared to the earlier methods. Figure 2 shows that the hybrid approach performs well compared with the unbalanced dataset.

5 Conclusion and Future Enhancements

Classifying the anomaly accurately with the available dataset may not possible in all the cases. Here we reduce computational time using rough set-based feature selection, second we balance the classes, using oversampling the data for predicting

the anomaly classes properly. This approach enhances the performance of the system. The experimental results show the performance enhancement with this hybrid approach.

In the future, genetic and fuzzy algorithm may used to get better records using the fitness function and membership values. Feature vector size may also decrease based on the availability of the optimal feature selection algorithms. Still lot of research is going on for finding optimal feature subset. Using these techniques may achieve the better anomaly detection model. The other classification techniques can be used on well-distributed and balanced data.

References

1. Lee, W., Stolfo, S., Chan, P., Eskin, E., Fan, W., Miller, M., Hershkop, S., & Zhang, J. (2001). Real time data mining-based intrusion detection. In: DARPA information survivability conference & exposition II, 2001, DISCEX'01, Proceedings (Vol. 1, pp. 89–100).
2. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," ACM Computing Surveys, vol. 41, no. 3, pp. 15:1–15:58, September 2009.
3. M. V. Joshi, R. C. Agarwal, and V. Kumar, "Mining needle in a haystack: classifying rare classes via two-phase rule induction," in Proc. of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2001, pp. 293–298.
4. P. N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. Addison-Wesley, 2005.
5. Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, Dordrecht, MA, 1991.
6. Nitesh V. Chawla Chapter on data mining for imbalanced datasets: An overview, Springer.
7. Nitesh V. Chawla, Nathalie Japkowicz, "Data Mining for Imbalanced Datasets: An Overview" A journal on special issue on learning from imbalanced datasets, volume 6, Issue 1 pp: 853–857.
8. Han, Jiawei, Micheline Kamber, and Jian Pei. "Classification", Data Mining, 2012.
9. Pawlak Z: Rough Sets and Intelligent Data Analysis, Information Sciences, 2002, 147:1–12.
10. http://iec.csic.es/dataset/.
11. R. P. Lippmann, D. J. Fried, I. Graf, J. W. Haines, K. Kendall, D. McClung, D. Webber, S. Webster, D. Wyschograd, R. Cunninghan, and M. Zissman. Evaluating Intrusion Detection Systems: The 1998 DARPA offline intrusion detection evaluation. In Proc. of DARPA Information Survivability Conference and Exposition (DISCEX00), Hilton Head, South Carolina, January 2527. IEEE Computer Society Press, Los Alamitos, CA, 1226 (2000).
12. R. Lippmann, J. W. Haines, D. J. Fried, J. Korba and K. Das. The 1999 DARPA OffLine Intrusion Detection Evaluation. In Proc. Recent Advances in Intrusion Detection (RAID2000). H. Debar, L. Me, and S. F. Wu, Eds. Springer-Verlag, New York, NY, 162182 (2000).
13. J. McHugh. Testing Intrusion Detection Systems: A Critique of the 1998 and 1999 DARPA Intrusion Detection System Evaluations as Performed by Lincoln Laboratory. In Proc. of ACM Transactions on Information and System Security (TISSEC) 3(4), pp. 262294 (2000).
14. Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Oversampling Technique. Journal of Artificial Intelligence Research, 16:321–357.
15. Enislay Ramentol, Yaile Caballero, A journal on SMOTE-RSB, 23 December 2009.
16. Jan G. Bazan, Marcin Szczuka, "The rough set exploration system (2005)" Transactions on Rough Sets III, Springer.

Database Retrieval-Based Digital Watermarking for Educational Institutions

T. Sridevi and S. Sameen Fatima

Abstract Multimedia broadcast monitoring is one of the major challenges and has drawn the attention of several researchers. Database retrieval technique is an application used to provide security for the sensitive images. Image authentication is provided by adding a watermark of the given string to the sensitive images. The security for image is provided by encoding the given image with the key which is provided by the user and is stored in the database as a numeric data. User is provided with a unique id which is sent to the receiver. Receiver uses this unique id for retrieving the image. Digital image watermarking is one among the several methods that hides information. Two factors are considered during watermarking, imperceptibility, and robustness. Measure of imperceptibility is PSNR and robustness is NC. Genetic algorithm is employed in the proposed watermarking algorithm.

Keywords Broadcast monitoring · Encoding · Decoding · Unique id · Watermarking · PSNR · NC · Genetic algorithm

1 Introduction

Nowadays because of information age, the distribution of multimedia has changed its pattern. Copies of digital data are created and spread throughout the internet. During the transmission of the data steganography and cryptography techniques can

---

T. Sridevi (✉)
Department of CSE, Chaitanya Bharathi Institute of Technology, Gandipet,
Hyderabad, Telangana, India
e-mail: sritumula@gmail.com

S. Sameen Fatima
Department of CSE, Osmania University, Hyderabad, Telangana, India

only guarantee for authenticity, integrity, and confidentiality. The existing techniques do not give protection, security techniques [1] that are founded on stenography and cryptography for data transmission only gives guarantee for authenticity, integrity, and data confidentiality.

Nowadays computing systems have wide range of processors, communication networks, and information repositories that play an important role in many aspects of our day-to-day life. The tremendous increase of internet users and adoptability nature of e-commerce have affected the organizations more liable to mischievous attacks. Several security measures have been recommended to control these attacks, so that integrity, authenticity, and availability of resources can be guaranteed.

With the always expanding development of media applications, security is a vital issue in correspondence and capacity of pictures, and encryption is one the approaches to guarantee security. Picture encryption procedures attempt to change over unique picture to another picture that is difficult to comprehend and also to keep the picture classified between clients, in other word, it is vital that no one could get to know the content without a key for decryption. Embedding domain is one of the methods of categorizing the digital watermarking algorithms. Embedding domain can be spatial domain or frequency domain. Discrete wavelet transform (DWT) [2] domain will provide security. To achieve perceptual invisibility and to achieve robustness against attacks, the watermark is inserted in the middle frequencies of the DWT domain. The embedded watermark should not alter the cover image quality and also should be robust against attacks. Copyright protection can be achieved using watermark techniques [3]. In literature there are number of watermarking schemes. In [4] genetic algorithm's population is initially generated randomly. The fitness of the population is calculated in every generation. Based on the criteria the chromosomes with better fitness value are selected from the present population and next generation population is generated. The remaining paper is arranged as follows. Section 2 describes the methodology, Sect. 3 discusses about the metrics used. Section 4 reveals the results to show the compressed ratio of watermark patterns, and Sect. 5 concludes the work and future work is also discussed.

2 Digital Image Watermarking

With the tremendous growth of technology, security is one of the most important problems. By increasing the network security there will not be any guarantee that images can also be secured from the intruders. These intruders can understand and

deduce the information as the system configuration is easily understood. Information gathered by the intruders can be easily altered to forge a person or a system or an application for attacks. This problem can be addressed by storing the encoded information in the database. Since the information is encoded it becomes difficult to the intruder to perceive and interpret. There are the components in this recommended method. (i) Image Encoding and Decoding (ii) Image Database and Storage (iii) Watermarking Scheme for images

(i) Image Encoding and Decoding

To transfer the image securely through the transmission media, image is encoded so that no illegal user can decode the image. There are numerous applications for image encoding which fall like military communication, medical imaging, etc. Figure 1 shows the flow chart for encoding and decoding.
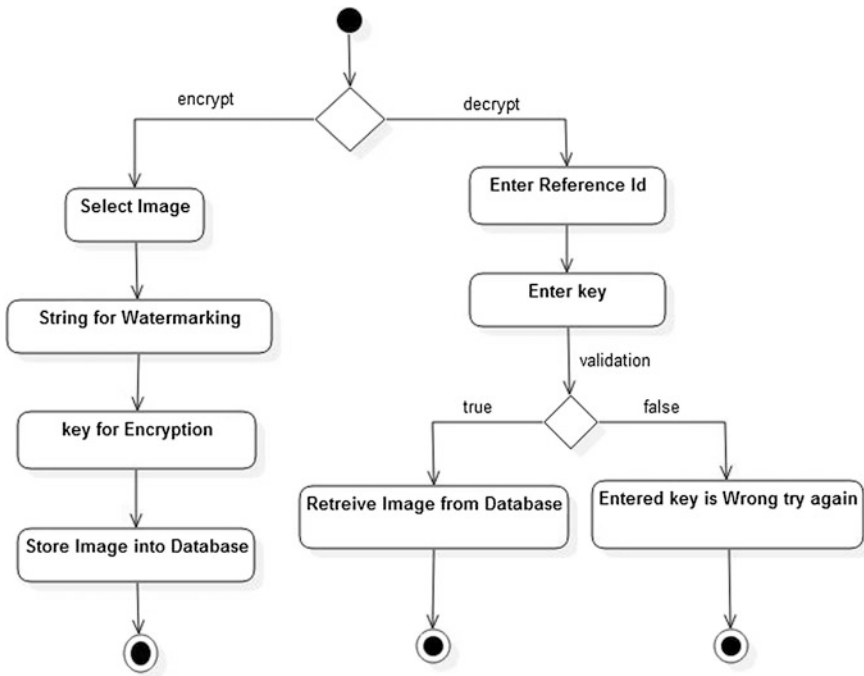
Fig. 1 Flow chart for encoding and decoding

The key generation algorithm for encoding

```
begin

        for a := 0 to I_L

        step 1:image converted to bytes

    step 2:bytes stored  into a 1D array as byte
array

        end for

end

begin

    for b=0 to A_L step i do

        key  is  encrypted    in  the  hash  function
(array[b])

            end for

end
```

where I_L is image length and A_L is array length.
The algorithm used for decryption is

```
begin

    while((reference id AND key)!=null)

    Add (reference id and key)   into byte array

  end while

end

begin

  for k=0 to A_L step i do

    decode the array[k] with the key

    end for

end
```

Here the procedure of encoding and decoding provides two key securities to the images.

(ii) Image Database and Storage

The encoded image obtained from image encoding procedure is stored in database for easy access to the image. In database images can be stored in two different ways.
First one is using BLOB and the second is using a one-dimensional array

Binary Large Objects (BLOB)

Images are huge volumes of data and SQL server enables customers with a unique data type called BLOB. A BLOB is a binary string of varying length up to 2,147,483, and 647 characters long. Analogous to previous binary types, BLOB strings are not associated with a code page and they do not embrace character data. The length for BLOB is usually given in bytes with one of the suffixes such as K, M, or G is given which are multiples of 1024, 1024 * 1024, and 1024 * 1024 * 1024, respectively. By default two gigabytes of length is provided to BLOB.

One-dimensional Array

In this procedure image is converted into bytes and stored as 1D array. These bytes are modified during the encoding process.

Storing the Image in Database

The encoded image is converted into one-dimensional byte array and is stored in the MySQL database. The procedure for storing the image in database is as given below.

```
Begin

    Step i: encrypted image is converted into one
dimensional array.

    Step ii: Add encrypt key in the one dimensional
array at a position.

        Step iii: Store the one dimensional array in the
database.

 end
```

Sender will send both encryption key and reference id to the receiver for decoding of the image.
The encoded image data are retrieved from the database based on the reference id and key given by the user. The decoding procedure is as follows:

```
begin

    if (reference id)

    Image is decoded from database

end
```

The procedure for extracting the decrypt key from the array is as follows:

```
Begin

    If(decryptkey= =storedkey)

    Decode the image

end
```

(iii) Watermarking Scheme for images

A watermarking system is divided into two distinct steps. Embedding process and Extraction process.

Embedding Process

1. Image of size m * n is taken as an input.
2. One-level DWT is applied to the original Image.
3. Watermark of size m_w * n_w is taken as an input.
4. Watermark is permuted.
5. Singular matrix of watermark is embedded in the singular matrix of original image using proposed genetic algorithm.
6. Inverse DWT is applied.
7. Watermarked Image is obtained.

Extraction Process

1. Watermarked Image is taken as a source image.
2. One-level DWT is applied to the watermarked frame.
3. Singular values of watermark are extracted from the embedded watermarked frame.
4. Inverse DWT is applied to secret frame by replacing the extracted bits in the watermark frame sub-band.
5. Inverse permute is applied.
6. Extracted watermark is obtained.

Attacks on Watermarking

Watermarked image gets distorted because of attacks. Attacks can be categorized as intentional or unintentional attacks. The watermarking algorithm is evaluated based on the robustness against these attacks. Signal processing attacks and geometric attacks are two broad way of classifying the attacks. Compression of image, addition of noise like Gaussian or salt and pepper noise, gamma correction,

filtering, brightness, sharpening, histogram equalization, averaging, collusion, printing, and scanning are all categorized as signal processing attacks. StirMark benchmark is a well-known evaluation tool for watermarking robustness and divides attacks into the subsequent nine categories. They are signal enhancement, compression, scaling, cropping, shearing, rotation, linear transformations, other geometric transformations, and random geometric distortions. In the case of signal scaling, cropping, shearing, rotation, linear transformations, and other geometric transformations, the attacked images prevail with and without jpeg 90 % quality factor compression. Petitcolas-recommended PSNR should have a minimum value of 38 dB.

3 Experimental Results

Images from database[5] are considered for experimentation. Images are low, medium and high contrast (Table 1).

Imperceptibility and robustness are two important parameters affecting the watermarking algorithm. To measure these two parameters, the corresponding metrics are: PSNR for imperceptibility and NC for robustness.

1. Peak signal-to-noise ratio (PSNR)
2. Normalized correlation (NC).

PSNR (peak signal-to-noise ratio) [6] is for measuring the imperceptibility and NC normalized coefficient) is to measure the robustness PSNR is usually expressed in terms of the logarithmic decimal scale. The PSNR is most commonly used as a measure of quality of reconstruction of lossy compression codec's (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression.

Table 1 Processing times of JPEG images for different sizes

Images	Size in pixels	PSNR	NC	Retrieval time (ms)
Lena (jpg)	100 × 100	38.69	0.96	144
	200 × 200	38.21	0.96	149
	300 × 300	37.83	0.96	156
Boat (BMP)	100 × 100	38.73	0.95	151
	200 × 200	38.37	0.95	172
	300 × 300	37.91	0.95	191
Peppers (PNG)	100 × 100	38.82	0.97	153
	200 × 200	38.43	0.97	163
	300 × 300	38.17	0.97	170

4 Conclusions and Future Work

Securing image using encryption and database techniques is a really interesting subject that is used to secure confidential images. In the current world there intruders who always try steal the privacy or break the security of multimedia. To overcome this type of attacks mainly on images this application provides a two-level security to confidential images. This system provides the flexibility to the user to choose any type of image for securing. This system unlike the previous systems which provides either keys or watermarking for security this system provides two keys for security along with watermarking for better authorization or ownership on the images.

In this paper, digital image watermarking using genetic algorithm is presented. Watermark is embedded into the suitable locations of an image. DWT is applied to the original image followed by wavelet decomposition. Watermark is embedded and fitness is evaluated. In this work to select new positions selection is used. The role of fitness function proposed is used to ensure the reliability through optimization.

Genetic algorithm is an optimization technique used to get an optimal solution by achieving the robustness and imperceptibility genetic algorithm is used to find the positions for embedding the watermark in the image. Optimization techniques are used to maximize the values of PSNR.

It has been observed experimentally that the performance of the scheme is satisfactory for several images after applying attacks like contrast increasing, contrast decreasing, noise, rotation, and Jpeg compression. The images are robust to all attacks used in the experiment. StirMark benchmark attacks are also applied on image.

DWT-genetic algorithm. From the experiment, a conclusion has been made that embedding the watermark using proposed method has high PSNR and NC which means that the algorithm is robust and imperceptible.

In this proposed method the unique id is sent between sender and receiver which has to be confidential. Transfer of this unique id should be handled in a proper way. This aspect can be extended for the next work. Second aspect is instead of taking watermark embedding strength randomly if it is obtained by any algorithm based on characteristics of watermark can increase robustness. Third aspect is the number of DWT level decompositions can be further increased. Fourth aspect to be considered is other optimization techniques such as particle swarm optimization can also be used along with genetic algorithm to obtain the maximum fidelity and robustness.

References

1. T. Sridevi, B. Krishnaveni, Y. Ramadevi and V. Vijaya kumar "A video watermarking algorithm for MPEG videos" Proc. ACM conf. Amrita ACM-W Celebration on Women in Computing in India, 2010.
2. Rakesh Kumar1 and Savita Chaudhary: Video Watermarking Using Wavelet Transformation: International Journal of Computer Trends and Technology (IJCTT)-Volume 4 Issue 5-May-2013.

3. S S Bedi, Rakesh Ahuja and Himanshu Agarwal. Article: Copyright Protection using Video Watermarking based on Wavelet Transformation in Multiband: International Journal of Computer Applications 66(8):1–5, March 2013.
4. Chin-Shiuh Shieh, Hsiang-Cheh Huang, Feng-Hsing Wang, and Jeng-Shyang Pan.: Genetic watermarking based on transform-domain techniques: Pattern recognition, 37(3):555–565, 2004.
5. USC-SIPI Image Database.
6. Sanjana Sinha, Prajnat Bardhan, and Swarnali Pramanick,: Digital Video Watermarking using Discrete Wavelet Transform and Principal Component Analysis: International Journal of Wisdom Based Computing, Vol. 1 (2), August 2011.

Bibliography

7. Alka N Potkar and Saniya M. Ansari, Article: Review on Digital Video Watermarking Technique: International Journal of Computer Applications 106(11):13–12, November 2014.
8. Deepak Sharma: Classification of Image Watermarking Schemes: Center for Advanced Computer Studies, University of Louisiana at Lafayette.
9. Mr Mohan A Chimanna, Prof. S.R. Khot: Digital Video Watermarking Techniques for Secure Multimedia Creation and Deliver: Vol. 3, Issue 2, March-April 2013, pp. 839–844.
10. Nitin A. Shelke and Dr. P.N. Chatu: A Survey on Various Digital Video Watermarking Schemes: International Journal of Computer Science & Engineering Technology (IJCTT) - Volume 4 Issue 12–Dec 2013.
11. Saik. A.M Al-Taweel: Robust video Watermarking Base on 3D-DWT Domain: International Journal OD computer applications 2009.

Conceptual Modeling Through Fuzzy Logic for Spatial Database

Narander Kumar, Ram Singar Verma and Jitendra Kurmi

Abstract Fuzzy logic shows a degree of vagueness or uncertainty that is expressed with crisp spatial objects. It supports spatial database objects which can precisely determine shape and boundaries. The uses of fuzzy logic with spatial database for conceptual modeling handles uncertain data types in fuzzy object-oriented database. Fuzzy approaches have been extensively applied for modeling different databases to explicitly represent and manipulate the imprecise and uncertain data precisely. In this paper, a new conceptual modeling approach has been developed and studied on the case of traffic modeling and route identification problem.

Keywords Fuzzy object-oriented database (FOODB) · IF$_2$O model · EER model · Aggregation · Specialization · Inheritance · UFO model

1 Introduction

Most of the time classical database models are not accurate in representation and manipulation of uncertain data that may be found in real-world and engineering application. Fuzzy logic has proposed in [1] various classical data models to make them capable of dealing with uncertain and imprecise data in information. Due to rapid development in the field of computing power, it brought many opportunities for database in emerging fields like CAD/CAM, multimedia, geographical information system (GIS) applications and spatial database. The characteristics of these applications are required for modeling, manipulation of complex object and

Narander Kumar (✉) · R.S. Verma · Jitendra Kurmi
Department of Computer Science, BBA University (A Central University),
Lucknow, India
e-mail: nk_iet@yahoo.co.in

R.S. Verma
e-mail: singar_ram@yahoo.co.in

Jitendra Kurmi
e-mail: Jitendrakurmi458@gmail.com

semantic relationship for handling the complex entities. The relational database and their fuzzy extensions are not sufficient to deal with complex object needed by the above applications. Such objects are well defined by modeling and represented to use object-oriented modeling techniques. This paper has been organized as follows. The review of fuzzy logic integration to different conceptual database modeling and other issues is described in Sect. 2. Section 3 deals with the new developed approaches of fuzzy's conceptual data modeling for spatial database. The route identification of modeling and traffic system using fuzzy conceptual spatial data modeling has been developed and described in Sect. 4. The conclusions as well as future work have been discussed in Sect. 5.

2   Related Work

The next generation of the development of data modeling in database is concerned with object-oriented modeling and their fuzzy extensions. This section presents the latest review on the different approaches, regarding modeling and representation of imprecise data and uncertain information in fuzzy-object oriented database. A data model (NF2) that is used to represent, manipulate complex and uncertain data using extended nested relation data model is presented in [2]. The extended algebra and SQL query languages are defined in data modeling defined the structured query language and advanced algebra. To represent complex relationship between the objects and attributes in NF2, data model is a challenging task. Some of the object-oriented database features are not supported by NF2 data models like class hierarchy, inheritance, super/subclass, and encapsulation in order to achieve the attribute of complex data model as well as complex relationship among objects. Further, the research is preceded with the development of conceptual data models and object-oriented data models. Insufficient information like null values are where incomplete schema and their object can be differentiated in [3] with the help of imprecise and uncertain information in OOD model.

The concept of using range of attribute values to represent the set of allowed values for an attributed of given class in [4]. It depends on the inclusion of actual attributes values of the given objective into the range of the attributes for the classes the degree of an object of the membership in a class can be calculated. Subsequently there is object-oriented based fuzzy data model and the extensions graph-based object data model is discussed in [5]. Linguistic qualifier presents strength which is associated with occurrence of relationship and object of class. Thus fuzzy classes and fuzzy class hierarchy are modeled in object-oriented database.

The graph-based operation has been proposed in [6], which deals with selecting and browsing such fuzzy object-oriented database that are used to control crisp and fuzzy information. To represent fuzziness and uncertainty by means of fuzzy set theory and generalized fuzzy sets as well as connective fuzzy sets through an uncertainty and fuzziness OOD model is proposed in [7]. The fuzzy hierarchies deal

with the concept of partial inheritance and multiple inheritances. Fuzzy object-oriented database model has been proposed in [8] that uses fuzzy attributed values which contain certain factor and SQL type data manipulation language. A concept based on possibility theory to represent vagueness and uncertainty in class hierarchies is proposed in [9] and defined fuzzy range of subclass attributes to define restriction on that of the superclass attributed. The inclusion of fuzzy range of their attributes depends on degree of inclusion of subclass in superclass. There are some major notions in OODB such as objects, classes, object-class relationship, sub/superclass, and multiple inheritance extended in fuzzy environment is proposed in [10]. Object data management group object model has been proposed in [11] for subsequent development of the fuzzy object-oriented database (FOODB). The concept of object-oriented database modeling techniques is proposed with the help of "Level 2 Fuzzy set" to represent with uniform and advantageous representation for perfect and imperfect real-world information [12].

The addition of fuzzy types to fuzzy objects-oriented database and to manage vague structure in [13, 14]. The class of OODB can be used to represent fuzzy type and describe the mechanism of the instantiation and inheritance can be modeled using new type of object-oriented database. The comparison of complex object in fuzzy context is developed [15]. The fuzzy relationship in object models has been investigated in [16, 17]. A fuzzy intelligent architecture has been proposed in [18] based on uncertain object-oriented data models. If-then rules are used to define the knowledge and possibility theory used to represent vagueness and uncertainty.

A simple theoretic model in [19] has been proposed to understand the fuzzy objects for easier analysis and specification of integrated computation by refereeing the object-oriented approach. This approach is proposed by Lee et al. [20] for OOM based on fuzzy logic to formulate imprecise requirements along with four directions: fuzzy class, fuzzy rules, fuzzy class relationship, and fuzzy association among classes. The fuzzy rules, rules with linguistic terms are used to define the relationship among attributes. Some of the special fuzzy object database [21, 22] proposed, and fuzzy probabilistic object-based have been proposed [23]. Also these fuzzy objects-oriented databases have been applied with different areas such as geographical information system and multimedia system proposed [24, 25].

A prototype of fuzzy object-oriented database has been implemented using VERSANT and VISUAL C++ [26] in fuzzy database. The nested fuzzy SQL query has been proposed [27] in fuzzy data. The several types of nested fuzzy queries have been extended to process unnested techniques.

3 Fuzzy Conceptual Spatial Data Modeling

The overall objective of the proposed models is to develop spatial temporal conceptual database model. Various kinds of data types and constructs are available for such a modeling.

Fig. 1 Point, line, and region data

Point data   Line data   Region data
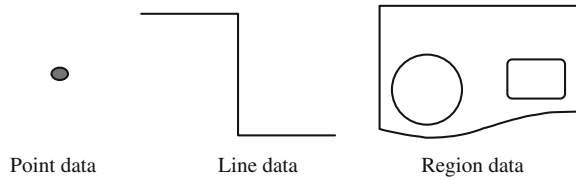
3.1 Spatial Data Types

As far as the spatial conceptual database modeling in concerned, the data types are point, line, and field region. Point is a data which considers the position only, but no focus on shape, size, or other spatial properties. In line data, the length and shape are considered, but area factor is not considered. Often road and river are represented by lines. Field is data which varies continuously from one place to another place. The examples of field data are terrain, pollution cul, soil types, etc. Region data is considered as a geographical object, which focuses on size and shape of interest for example, a state or country (Fig. 1).

These types of data items are represented by different diagrammatic representation. There are a lot of uncertainties available in geographical data type. Hence, for uncertainty representation fuzzy logic has been applied.

3.2 Uncertainty Issues in Spatial Modeling

Different uncertainty issues are as follows:

1. The information about objects may consist of uncertainty, which includes 1. missing data, 2. uncertain data 3. geostatic 4. multidimensional uncertainties, and many others.
2. The region boundaries are uncertain in its nature, for this fuzzy logic approach is applied to decide on whether a particular region is included in the specified area or not.

Another issue of the uncertainty is with the querying of spatial data types, for example, "Find all the large in area A1?"

A proposal has been given to handle all the above uncertainty issues.

Fuzzy integration with various data types:

(a) *Point data* It is defined by point G. The location of G is represented by the coordinates (x, y) in the spatial region and represented by $G(x, y)$ where x and y are the latitudes and longitudes, respectively.

Fig. 2 Fuzzy line data

$(x_1, \mu_{G'}(x_1), y_1, \mu_{G'}(y_1))$

$(x_2, \mu_{H'}(x_2), y_2, \mu_{H'}(y_2))$
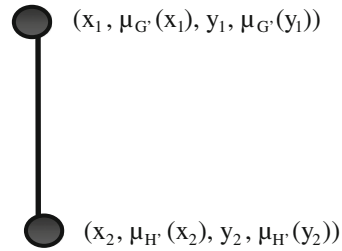
The fuzzy representation of the above expression would be as follows:

$$G' = G[(x, \mu_G(x))(y, \mu_G(y))].$$

Here $\mu_G(x)$ and $\mu_G(y)$ are the membership degrees, respectively, and

$$0 \leq \mu_{G'}(x) \leq 1 \text{ and } 0 \leq \mu_{G'}(y) \leq 1.$$

(b) *Fuzzy line data* The line data is expressed generally by LINE (G, H) where $G = G(x_1, y_1)$ and $H = H(x_2, y_2)$
In the fuzzy representation,

$$G' = G[(x_1, \mu_A(x_1)), (y_1, \mu_A(y_1))] \text{ and } H' = H[(x_2, \mu_A(x_2)), (y_2, \mu_A(y_2))]$$

Now the fuzzy representation of LINE is as follows:
F-LINE $= ((G', H'), \mu_L(x))$ where $\mu_L(x)$ is the degree which the inclusion of line in a particular region $0 \leq \mu_L(x) \leq 1$ (Fig. 2)

(c) *Fuzzy region data* A region data is represented by $R(a_1, a_2,\ldots, a_n)$ where $a_1, a_2,\ldots, a_n$ are the points which decide the geographical region associated with these points. Now for fuzzy representation, all the $a_1, a_2,\ldots, a_n$ are the geographical points represented by

$$a'_1 = a_1[(x_1, \mu_{a1}(x_1)), (y_1, \mu_{a1}(y_1))]$$
$$a'_2 = a_2[(x_2, \mu_{a2}(x_2)), (y_2, \mu_{a2}(y_2))]$$
$$\cdots$$
$$\cdots$$
$$A'_n = a_n[(x_n, \mu_{an}(x_n)), (y_n, \mu_{an}(y_n))]$$

Also, a conjunction can be defined including multiple regions with fuzzy membership degrees

$$R' = (R_1, \mu_R(R_1)), (R_2, \mu_R(R_2)), (R_3, \mu_R(R_3))\ldots(R_n, \mu_R(R_n))$$

A graphical representation is as follow (Fig. 3)

Fig. 3 Fuzzy region data

$$(a'_1, \mu_R(a_1))$$

$$(a'_5, \mu_R(a_5))$$

$$(a'_4, \mu_R(a_4))$$

$$(a'_3, \mu_R(a_3))$$

$$(a'_2, \mu_R(a_2))$$

where $a'_n = a_n [(x_n, \mu_R(x_n)), (y_n, \mu_R(y_n))]$
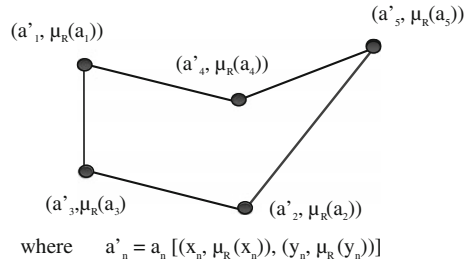
4   Route Identification Modeling in Traffic System Using Fuzzy Conceptual Spatial Data Modeling

This modeling example includes three basic steps:

1. Identification of region
2. Identification of roads (Line Data)
3. Identification of junctions (Point Data)

The selection of route would depend on interest area whether someone wants to go and the distance costs. The area of interest would be varying and it would be shown by the fuzzy membership degree values. The membership degree of the above point data which is making line data are decided on basis of distance factor. The route would be identified by these membership degrees. An example of route selection in the above as follows the route identification between a_1 to a_{15} (Fig. 4)

These possible routes

Route 1

$$\{(a_1, \mu_R(a_1)), (a_2, \mu_R(a_2)), (a_3, \mu_R(a_3)), (a_9, \mu_R(a_9)),$$
$$(a_8, \mu_R(a)), (a_{12}, \mu_R(a_{12})), (a_{11}, \mu_R(a_{11})), (a_{15}, \mu_R(a_{15}))\};$$

Route 2

$$\{(a_1, \mu_R(a_1)), (a_2, \mu_R(a_2)), (a_3, \mu_R(a_3)), (a_4, \mu_R(a_4)),$$
$$(a_8, \mu_R(a_8)), (a_{12}, \mu_R(a_{12})), (a_{11}, \mu_R(a_{11})), (a_{15}, \mu_R(a_{15}))\};$$

Route 3

$$\{(a_1, \mu_R(1)), (a_2, \mu_R(a_2)), (a_3, \mu_R(a_3)), (a_9, \mu_R(a_9)),$$
$$(a_8, \mu_R(a_8)), (a_{10}, \mu_R(a_{10})), (a_{11}, \mu_R(a_{11})), (a_{15}, \mu_R(a_{15}))\};$$

Fig. 4 Schematic diagram

where μ is degree of intrusion in the route
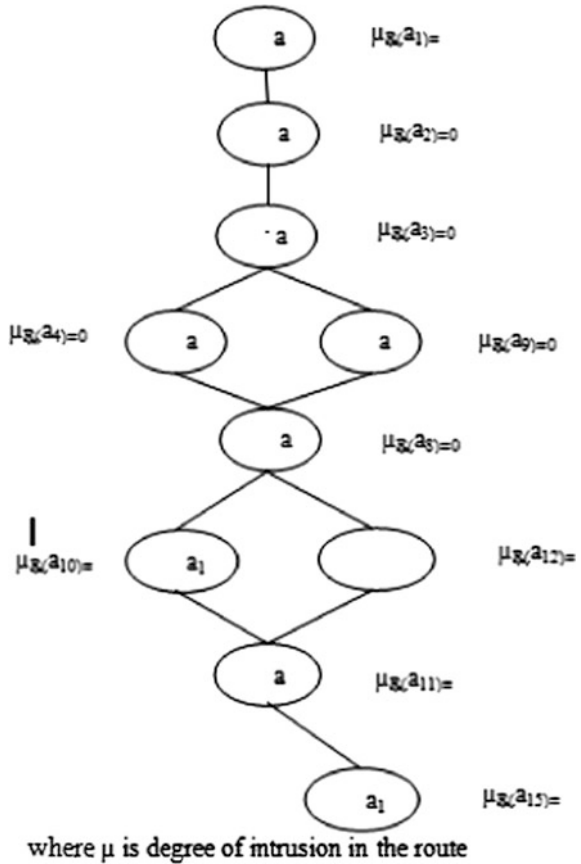
Route 4

$$\{(a_1, \mu_R(a_1)), (a_2, \mu_R(a_2)), (a_3, \mu_R(a_3)), (a_4, \mu_R(a_4)),$$
$$(a_8, \mu_R(a_8)), (a_{10}, \mu_R(a_{10})), (a_{11}, \mu_R(a_{11})), (a_{15}, \mu_R(a_{15}))\};$$

Following schematic diagram including all the four routes above discussed the solution is as follows. The identified route is $a_1, a_2, a_3, a_4, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{15}$.

5 Conclusion

A conceptual modeling using fuzzy logic has been proposed to deal with uncertainty for data type. The fuzzy spatial feature introduced fuzzy data types like fuzzy points, fuzzy line, and fuzzy region to use in spatial database:the characteristic of designing of conceptual modeling of spatial database using fuzzy logic. Spatial data modeling

has lot of inexactness in itself. In order to deal with the inexactness and to represent it precisely a fuzzy-based conceptual modeling approach has been developed in this paper. The application of proposal is carried out in the route identification problem of traffic network. A review has been carried out on the issues related to application of fuzzy logic in modeling and development of different databases also.

References

1. Zadeh, L. A, Fuzzy Sets, Information & Control, Vol. 8 1965 338–353.
2. Yazici, A., Soysal, A., Buckles, B. P. and Petry, F. E., Uncertainty in a nested relational database model, Data &Knowledge Engineering, 30 (3) 1999 275–301.
3. Zicari, R. and Milano, P., Incomplete information in object-oriented databases, ACM SIGMOD Record, 19 (3) 1990 5–16.
4. George, R., Srikanth, R., Petry, F. E. and Buckles, B. P., Uncertainty management issues in the object-oriented data model, IEEE Transactions on Fuzzy Systems, 4 (2) 1996 179–192.
5. Bordogna, G., Pasi, G. and Luearella, D., A fuzzy object-oriented data model for managing vague and uncertain information, International Journal of Intelligent Systems, 14 1999 623–651.
6. Bordogna G. and Pasi G., Graph-based interaction in a fuzzy object oriented database, International Journal of Intelligent Systems, vol. 16 2001 821–841.
7. Gyseghem, N. V. and de Caluwe, R., Imprecision and uncertainty in UFO database model, Journal of the American Society for Information Science, 49 (3) 1998 236–252.
8. Umano M., Imada T., Hatono I., and Tamura H., Fuzzy object-oriented databases and implementation of its SQL-type data manipulation language, in Proceedings of the 7th IEEE International Conference on Fuzzy Systems, vol. 2 1998 1344–1349.
9. Dubois D., Prade, H. and Rossazza J. P., Vagueness, typicality, and uncertainty in class hierarchies, International Journal of Intelligent Systems, vol. 6 1991 167–183.
10. Ma Z. M., Zhang W. J., and Ma W. Y., Extending object-oriented databases for fuzzy information modelling, Information Systems, vol. 29 2004 421–435.
11. Cross V., Caluwe R., and van Gyseghem N., A perspective from the fuzzy object data management group (FODMG), in Proceedings of the 6th IEEE International Conference on Fuzzy Systems, vol. 2 1997 721–728.
12. de Tré G. and de Caluwe R., Level-2 fuzzy sets and their usefulness in object-oriented database modelling, Fuzzy Sets and Systems, vol. 140, 2003, 29–49.
13. Marín N., Medina J. M., Pons O., Sánchez D., and Vila M. A., Complex object comparison in a fuzzy context, Information and Software Technology, vol. 45 2003 431–444.
14. Marín N., Pons O., and Vila M. A., A strategy for adding fuzzy types to an object oriented database system, International Journal of Intelligent Systems, Vol. 16 2001 863–880.
15. Marín N., Vila M. A., and Pons O., Fuzzy types: A new concept of type for managing vague structures, International Journal of Intelligent Systems, Vol. 15 2000 1061–1085.
16. Cross V., Fuzzy extensions for relationships in a generalized object model, International Journal of Intelligent Systems, Vol. 16 2001 843–861.
17. Cross V., Defining fuzzy relationships in object models: Abstraction and interpretation, Fuzzy Sets and Systems, Vol. 140 2003 5–27.
18. Ndouse T. D., Intelligent systems modelling with reusable fuzzy objects, International Journal of Intelligent Systems, Vol. 12 1997 137–152.
19. Akiyama Y. and Higuchi K., A Simple Theoretic Model to understand fuzzy objects, IEEE International Conference on Systems, Man & Cybernetics, 1998 20–40.
20. Lee J., Xue N. L., Hsu K. H., and Yang S. J. H., Modelling imprecise requirements with fuzzy objects, Information Sciences, Vol. 118 1999 101–119.

21. Koyuncu M. and Yazici A., IFOOD: an intelligent fuzzy object-oriented database architecture, IEEE Transactions on Knowledge and Data Engineering, Vol. 15 2003 1137–1154.
22. Yazici A. and Koyuncu M., Fuzzy object-oriented database modelling coupled with fuzzy logic, Fuzzy Sets and Systems, Vol. 89 1997 pp. 1–26.
23. Cao T. H. and Rossiter J. M., A deductive probabilistic and fuzzy object-oriented database language, Fuzzy Sets and Systems, Vol. 140 2003 129–150.
24. Cross V. and Firat A., Fuzzy objects for geographical information systems, Fuzzy Sets and systems, Vol. 1132000 19–36.
25. Majumdar A. K., Bhattacharya I., and Saha A. K., An object-oriented fuzzy data model for similarity detection in image databases, IEEE Transactions on Knowledge and Data Engineering, vol. 14 2002 1186–1189.
26. Firat A., Cross V., Lee T. C. Fuzzy set Theory in object oriented databases: A prototype implementation using VERSANT ODBMS & VISUAL C++, IEEE Conference of North American Fuzzy Information Society, 20–21 Aug 1998 146–150.
27. Yang Q., Zhang, W., Wu J., Yu C. Na Kajima. H. Rishe N. D. Efficient Processing of nested fuzzy SQL Queries in fuzzy databases, IEEE Transaction on Knowledge and Data Engineering, 13(6) Nov/Dec. 2001 884–901.

Performance Analysis of Clustering Algorithm in Sensing Microblog for Smart Cities

Sandip Modha and Khushbu Joshi

Abstract Smart city is an aspiration of the various stakeholders of the city. We strongly believe that social media can be one of the real-time data sources, which help stakeholder to realize this dream. In this paper, we have analyzed the real-time data provided by Twitter in order to empower citizens by keeping them updated about what is happening around the city. We have implemented various clustering algorithms like k-means, Hierarchical agglomerative, LDA topic modeling on Twitter stream and reported results with purity 0.476, normal mutual information (NMI) 0.3835, and F-measure 0.54. We conclude that HA-ward outperforms K-means and LDA substantially. We also conclude that results are not impressive and need to design separate feature based clustering algorithm. We have identified various tasks to mine microblog in the ambit of smart city such as event detection, geo-tagging, city clustering based upon the user activity on ground.

Keywords Microblog analysis · Clustering · Smart city · Topic modeling

1 Introduction

Today, smart city is one of the buzzwords in the world including developed and developing countries. The rationale behind the smart city is to provide state-of-the-art services to their citizens to enable them to take informed decision. The smart city is a multidisciplinary research area involving many engineering disciplines. We firmly believe that social media can be one of the real-time data

Sandip Modha (✉)
Information Retrieval and Language Processing Lab, DA-IICT,
Gandhinagar, India
e-mail: sjmodha@gmail.com

Khushbu Joshi
Department of Electronics and Communication, LDRP-ITR,
Gandhinagar, India
e-mail: skhushi86@gmail.com

sources that might contain city's dynamics. We have chosen Twitter as microblog for this experiment.

Today, social media users post millions of messages called post (or tweets in case of Twitter) on microblogging about their personal lives, politics, sports event, controversial event, emergency such as earthquake, accident, fire, etc. It is noteworthy that some incidents get reported in social media prior to news TV channel. (e.g., Michael Jackson's death) [1]. It is impossible to keep track of the entire post due to its massive volume. To tackle this issue, we have tried to cluster these tweets using existing clustering algorithm inspired by approach taken by news aggregating services like Google.

Twitter is one of the popular microblogging social network websites having 302 million [2] active users (out of 500 million) posting 400 million multilingual tweets (or post, Twitter message) every day. Due to 140 characters limitation, tweets often contain noisy text, URLs, tags, and Twitter names [3]. Twitter user often use informal language or native language written in roman script, nonstandard abbreviations (e.g., 2mro, Tomroo, Dat). Therefore, Twitter poses serious challenge to the research community for mining as compared to mining text documents.

Given a random Twitter stream, we have applied k-means, hierarchical agglomerative, and LDA topic modeling algorithms to perform clustering based on textual features. We report the result in terms of purity, normal mutual information (NMI), and F-measure. The rest of the paper is organized as follows. In Sect. 2, we have discussed related work, in Sect. 3 we formally describe our problem, in Sect. 4 we discuss Twitter data corpus preparation and cleaning process. In Sects. 5 and 6 we describe the experiment and metric to analyze clusters. In Sect. 7, we discuss the role of social media to achieve the dream of smart city and future research direction. This paper is our first attempt to mine or to sense social media to empower a city. We will discuss various research tracks for this relatively unexplored research area.

2 Related Work

Twitter is one of the popular and fastest growing social networks. Due to the limited length (140 characters) of the tweet, it poses difficult challenges to research community for mining the text, at the same time it also gives rich social network property. Event detection and summarization are the major research areas where research community has focused. Clustering is the first task before one can proceed to any of the above tasks.

TweetMotif [4] is the first attempt toward the tweet clustering. Unsupervised approach has been applied to cluster the twitter messages. However, they have not quantified the results.

Rosa et al. [5] performed classification on military short messages using k-nearest neighbor, support vector machines, and naïve Bayes.

Our work is partially similar to Rosa et al. [5]. The main difference between the two is (i) We have used hierarchical agglomerative clustering algorithm with

different linkage to cluster tweet (ii) we have concluded that HA-ward outperform *K*-mean and topic modeling algorithm in terms of purity, NMI, and *F*-measure.

3 Problem Definition

In this section, we formally describe the problem. Given a random Twitter stream, T is a set of tweets $T = \{T_1, T_2, \ldots, T_n\}$. H is a set of all hashtags $H = \{H_1, H_2, \ldots, H_n\}$. In Twitter, hashtag represents the event or topic. Hashtags are the keywords which start with prefix # and appear anywhere in the tweet. We will aggregate all tweets hashtag wise. These aggregate tweets are the gold-standard label for our clustering algorithm. In this experiment, we have considered each tweet as a document.

Let us describe problem formally. F is a clustering algorithm, T is set of Tweet, C is set of our predefined hash-ag wise class and Ω is a set of cluster produced by clustering algorithm $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. We want to compute assignment $F\colon T \rightarrow \{\omega_1, \omega_2, \ldots, \omega_n\}$. We have compared cluster generated by existing algorithm with C. We report the result in terms of purity, NMI, *F*-measure.

4 Twitter Data Corpus

4.1 Real-Time Tweet Extraction

Sample tweets are available to researchers and practitioners through public APIs at no cost. For this experiment, we have downloaded tweets during 23/1/2015 to 25/2/2015 using Twitter REST API. We have identified popular hashtag trending on a particular day. Table 1 shows the detail description of the hashtag.

Table 1 List of hashtags and their description

Sr. No.	Hashtag name	Types of hashtag
1	#Teresa	Religious controversy
2	#CWC15	Sports/cricket
3	#Bigdata	Academic/commercial
4	#AAPvision4Delhi	Political campaign
5	#WhereISRahul	Political controversy

4.2 Tweet Cleaning and Corpus Preparation

Twitter Rest API returns the tweet with 16 features including text, username, time etc. We are only interested in text feature. We have stored all tweets inside MongoDB database. Tweets in nature are very noisy. We need to clean and pre-process the tweets before creating the corpus. Following are the list of steps to perform tweet cleaning:

- Convert text into uppercase
- Removal of username or mention start with @
- Removal of hashtag
- Stop word and blank space removal
- URL and special character removal
- Word stemming

After the cleaning phase, some tweets become blank which we have removed from the corpus. We treat each tweet as a document.

5 Experiment Detail

We have conducted experiments using R language and chose to store tweets in MongoDB database. R is the most preferred statistical tool to analyze massive data. Because of its non-proprietary nature, and its availability under GNU General Public License, we find it most suited to perform our experiments on this test-bed (Table 2).

Once corpus is prepared and converted in a vector (e.g., DTM/TDM). We can apply various clustering algorithms on our corpus. In our experiments, we have done clustering using k-means, hierarchical agglomerative, and LDA topic modelling (Table 3).

Table 2 List of packages used in experiment [6]

Sr. No.	Name of package	Description of package
1	TwitteR	Used to extract tweet using twitter's restful API
2	StreamR	Allow us to extract tweet using twitter streaming API
3	TM	For performing basic text mining task
4	Topicmodel	Implement topic modelling algorithm
5	Ggplot2	An implementation of the grammar of graphics
6	Rmongodb	R driver for MongoDB
9	ROAuth	R interface for Oauth
10	NLP	Natural language processing infrastructure

Table 3 Top keywords for each cluster generated *k*-means

Cluster	Top word
1	thanks jan donating 2015 data big rahul india gandhi will
2	teresa great love things amp can happy think amaya debate
3	mother teresa hindu adds criticism religious indian leaders statement tens
4	rss mother teresa chief bhagwat mohan remarks chiefs dont must
5	india trending number position 2015 is feb gmt tue now rank

Table 4 Top keywords for each topic generated by LDA

Topic	Top word
1	India, trending, position, number, u0645u0646, world, cup, Pakistan
2	Trending, india data, rahul, now, ist1, leave, great
3	Teresa, mother, rss, rahul, Gandhi, bhagwat, statement, twitter
4	Thanks, jan, donating, 2015, 1000, united, kumar, 100
5	Big, terersa, data, mother, Indian, good, leader, aap

5.1 Latent Dirichlet Allocation (LDA) Model

Topic modeling [7] is a form of text mining, a way of identifying patterns in a corpus. Topic models [7] are a suite of algorithms that uncovers the hidden thematic structure in document collections. These algorithms help us develop new ways to search, browse and summarize large archives of texts. We have used latent Dirichlet allocation (LDA) [7] as an example of a topic model (Table 4).

6 Results and Cluster Analysis

In order to measure the performance of clustering algorithm we have identified the following hashtags (Table 1). We believe that hashtag provides some insight of topics of the tweet. We have considered this as gold standard set. We measure our clustering algorithm result $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ with these hashtag classes of $C = \{c_1, c_2, \ldots, c_n\}$ and report the result in various metric(e.g., purity, NMI, F-measure). Table 5 show our results in terms of purity, NMI, and F-measure. In the next section we will define each metric in detail (Table 6).

Table 5 Result on tweet dataset

Sr. No.	Algorithm	Purity	NMI	F-measure
1	k-means	0.3704	0.2894	0.40
2	Topic modeling	0.3896	0.1491	0.3877
3	HA-complete link	0.22	0.057	0.34
4	HA-single link	0.2056	0.0109	0.33
5	HA-average link	0.2056	0.0109	0.33
6	HA-ward	0.476	0.3835	0.54

Table 6 Hashtag wise precision, recall, F1 value of cluster using LDA model

Sr. No.	Hashtag	Precision	Recall	F1
1	#Teresa	0.494	0.488	0.491
2	#CWC15	0.372	0.388	0.380
3	#Bigdata	0.337	0.353	0.345
4	#AAPvision4Delhi	0.31	0.27	0.29
5	#WhereISRahhul	0.58	0.62	0.60

An experiment was performed with same hashtag listed in Table 1. In this experiment, we have taken 2500 random tweets from each hashtag. The results are as below.

Our results are broadly in line with previous researchers' results in case of k-means and topic modeling algorithm but to the best of our knowledge, no one has reported result on hierarchical agglomerative with different linkage. In our experiment, hierarchical agglomerative with ward method had reported best result in terms of NMI, purity, and F-measure.

6.1 Purity

To compute purity [8], each cluster is assigned to the class which is most frequent in the cluster, and then the accuracy of this assignment is measured by counting the number of correctly assigned tweet and dividing by N. Bad clustering have purity values close to 0, a perfect clustering has a purity of 1. High purity is easy to achieve when the number of clusters is large in particular, purity is 1 if each document gets its own cluster. Thus, we cannot use purity to trade off the quality of the clustering against the number of cluster [8]

Formally,

$$\text{Purity}(\Omega, C) = \frac{1}{N} \sum_k \max |\omega_k \cap c_j|. \tag{1}$$

where $\Omega = \{\omega_1, \omega_2, \ldots, \omega_K\}$ is the set of clusters and $C = \{c_1, c_2, \ldots, c_J\}$ is the set of classes. We interpret ω_k as the set of tweets in cluster k and c_j as the set of tweets in class j. N is equal to number of tweets.

6.2 Normal Mutual Information

As a more information theoretic measure of cluster quality, we also evaluate normalized mutual information (NMI) defined as follows [8].

$$\text{NMI}(\Omega, C) = \frac{I(\Omega; C)}{[H(\Omega) + H(C)]/2}. \tag{2}$$

where I is mutual information and defined by

$$I(\Omega, C) = \sum_k \sum_j P(\omega_k \cap c_j) \log \frac{P(\omega_k \cap c_j)}{P(\omega_k)P(c_j)}. \tag{3}$$

$$= \sum_k \sum_j \frac{|\omega_k \cap c_j|}{N} \log \frac{N|\omega_k \cap c_j|}{|\omega_k||c_j|}. \tag{4}$$

where $P(\omega_k)$, $P(c_j)$, and $P(\omega_k \cap c_j)$ are the probabilities of a document being in cluster ω_k, class c_j, and in the intersection of ω_k and c_j, respectively [8] where H is entropy

$$H(\Omega) = -\sum_k P(\omega_k) \log P(\omega_k). \tag{5}$$

$$= -\sum_k \frac{|\omega_k|}{N} \log \frac{|\omega_k|}{N}. \tag{6}$$

NMI is always a number between 0 and 1 and will be 1 if the clustering results exactly match the category labels while 0 if the two sets are independent.

6.3 F-Measure

Another frequently used external clustering evaluation measure is commonly referred to as clustering accuracy. The calculation of this accuracy is inspired by the information retrieval metric of F-measure. The formula for this clustering F-measure as described in [9] is shown below.

Let N be the number of tweet, C the set of classes, Ω is the set of clusters, and n_{ij} be the number of members of class $c_i \in C$ that are elements of cluster $\omega_j \in \Omega$

$$F(C, \Omega) = \sum_{c1 \in C} \frac{|c1|}{N} \max \left\{ F\left(c_i, \omega_j\right) \right\} \text{ where } \omega_j \in \Omega. \tag{7}$$

where

$$F\left(c_i, \omega_j\right) = \frac{2 * R\left(c_i, \omega_j\right) * P\left(c_i, \omega_j\right)}{R\left(c_i, \omega_j\right) + P\left(c_i, \omega_j\right)}. \tag{8}$$

$$R\left(c_i, \omega_j\right) = \frac{N_{i,j}}{C_{i,j}}. \tag{9}$$

$$P\left(c_i, \omega_j\right) = \frac{N_{i,j}}{\omega_{i,j}}. \tag{10}$$

6.4   Challenges

Since we want to perform tweet clustering, we consider each tweet as a document. During the cleaning process, some tweets became blank (tweet length is zero) so it is next to impossible to perform LDA on such blank tweets. We also face problems with the text of tweets. Some tweets are written in native language (e.g., Hindi, Gujarati). Some tweets are transliterated content which means they are written in native language using roman script.

6.5   Result Discussion and Conclusion

Results have been displayed in Tables 4 and 5. From the results we can conclude that classical clustering algorithm performs poorly on tweet due to short text (140 characters), noisiness and absence of the context. We are getting purity in the range of 0.20–0.47, which is not up to the mark. We can conclude that there is a need for feature-based clustering algorithm.

7   Future Work

We are working in the broad problem of smart city by exploiting the social media and this paper (tweet clustering) is the first step toward this big problem. In future, we would like to develop native clustering algorithm base on spatial, temporal,

linguistic feature of the tweet. There are broadly two types of social media (a) tradition social media(like Twitter/Facebook) and location-based social media (LBSN), e.g., Foursquare, Instagram etc. we want sense social media to transform city into smart city. Social media can be one of the real-time data sources, which continuously generate massive data and exploiting this real-time data by city's stakeholder can enable them to take smart decision [10]. Today Twitter provides city level trend in multilingual language. We have listed following research area that can be explored in future

1. City level community discovery
2. Summarization of local trend in multilingual language
3. Local event detection in multilingual/transliterated tweet.

Location added with social network bridges the gap between the physical world and digital world [11]. Examples of LBSNs which are widely used are Foursquare and Instagram. With LBSNs, we are able to understand users and locations which can explore the relationship between them. The tasks that can be further explored are (i) estimate user similarity (ii) finding local experts in a region (iii) location recommendations (iv) itinerary planning (v) location–activity recommender (vi) place-clustering.

References

1. Shamanth Kumar Fred Morstatter Huan Liu 2014: A book on Twitter Data Analytic published Springer 2014.
2. Wikipedia http://en.wikipedia.org/wiki/Twitter.
3. Ranganath R., Jurafsky, D., and McFarland. D. 2009: It's Not You, it's me: Detecting Flirting and its Misperception in Speed-Dates. In Proceedings of Conference on Empirical Methods in Natural Language Processing.
4. O'Connor B., Krieger, M., and Ahn, D. 2010: TweetMotif: Exploratory Search and Topic Summarization for Twitter. In Proceedings of Fourth International AAAI Conference on Weblogs and Social Media.
5. Kevin Dela Rosa, Rushin Shah, Bo Lin, Anatole Gershman, Robert Frederking: Topical clusterings of tweets in Proceeding of SWSM'10, July 28, 2011, Beijing, China.
6. The Comprehensive R Archive Network https://cran.r-project.org/ for R-Package.
7. Blei D. M., Ng, A. Y., and Jordan, M. I. 2003: Latent Dirichlet Allocation. J. Mach. Learn. Res. 3 (Jan. 2003) 993–1022.
8. Christopher D. Manning, Hinrich Schütze, and Prabhakar Raghavan: A book on Introduction to information retrieval Cambridge University Press.
9. Benjamin C. M. Fung, Ke Wang, and Martin Ester. 2003: Hierarchical document clustering using frequent itemsets. In Proc. of the SIAM International Conference on Data Mining.
10. A White paper by department of Business innovation & skill UK government Oct-2013.
11. Yu Zheng 2014: Urban Computing: Concepts, Methodologies, and Applications ACM Transactions on Intelligent Systems and Technology, Vol. 5, No. 3, Article 38, Publication date: September 2014.

Offloading for Application Optimization Using Mobile Cloud Computing

Chinu Singla and Sakshi Kaushal

Abstract Mobile applications are increasingly becoming ubiquitous which provides rich functionalities and services to mobile users. But despite having some technological advancements, smart mobile devices (SMDs) are still minimum potential computing devices enforced by battery life, storage capacity, and network bandwidth which obstructs the possible execution of computational intensive applications. Thus, mobile cloud computing is employed to optimize the computationally intensive applications which use computation offloading techniques for increasing SMD's capabilities. The aim of this paper is to emphasize the specific issues related to mobile cloud computing, offloading, and thus concluding the paper by analyzing the challenges that have not been met perfectly and charts a roadmap towards this direction.

Keywords Mobile cloud computing · Application offloading · Distributed systems · Smart mobile devices

1 Introduction

The latest development in mobile computing has altered user preferences for computing. As users expect to operate computationally intensive applications on their smart mobile devices (smartphone's, personal computers, tablets) but smart mobile devices have certain limitations such as battery life, storage capacity, CPU potential, etc, due to which the user faces so many problems while running large applications. Examples of such applications comprise image processors [1, 2],

Chinu Singla (✉) · Sakshi Kaushal
Computer Science and Engineering, UIET, Panjab University,
Chandigarh, India
e-mail: cheenusingla10@gmail.com

Sakshi Kaushal
e-mail: sakshi@pu.ac.in

video processing [3], and online gaming. Such applications need battery lifetime, more computing power on resource constrained SMDs [4]. To overcome these limitations, the concept of mobile cloud computing has been introduced. Besides, its computational offloading is a good solution to enhance the performance of the mobile application.

Cloud computing (CC) directs to manipulating, configuring and accessing the application via the Internet. It provides online infrastructure, data storage, and applications. Cloud computing provides various services and resources to SMDs such as software as a service (SaaS), platform as a service (PaaS), infrastructure as service (IaaS), testing as a service (TaaS), etc., through different service provider applications in an on-demand fashion and users can access these resources through centralized cloud server. With the support of cloud computing and the eruption of mobile applications, mobile cloud computing (MCC) is introduced, which is defined as the combination of cloud computing into the mobile environment. MCC provides ubiquitous and computationally intensive mobile applications by attaining the benefits of cloud data centers. It brings new facilities and services to mobile users by taking the benefits of CC and thus reduces the resource limitations of SMDs. For increasing computing potentials of SMDs and to eliminate resource limitations, MCC utilizes application offloading techniques [5, 6]. In application offloading, computationally intensive applications are offloaded by migrating intensive applications to remote server nodes. The terms "surrogate computing" or "cyber foraging" are also used to define computation offloading. Computation offloading methods improves the performance of mobile systems by saving energy and by reducing execution cost. Offloading can be performed either in the static environment (constant network bandwidth, server loads, and connection status) or in the dynamic environment (by changing network bandwidth, server loads, and connection status). For offloading intensive applications partitioning of the applications is done and it may be static or dynamic. In static application partitioning, various components are partitioned either at compile time or at runtime in a static manner, i.e., at development time, whereas in dynamic application partitioning, the applications are separated in a dynamic manner at runtime. Static partitioning is valid if and only if all the performance parameters are predicted in advance but this approach minimizes the overall performance because the servers speed may vary. So currently offloading techniques employ dynamic partitioning techniques to improve the performance of SMDs. But the major challenging issue in dynamic partitioning is the additional consumption on mobile devices during the establishment and management of distributed applications at runtime. This paper examines different frameworks for computational offloading at runtime by discussing the critical aspects, approaches used and related issues to it.

The remaining of the paper is as follows: Sect. 2 explains the concept of mobile cloud computing and computation offloading. Section 3 analyzes the different application partitioning frameworks. Section 4 concludes the paper by discussing the future work.

2 Background

This section demonstrates the basic concepts of MCC and also describes the concept and existing work related to computation offloading in MCC.

2.1 Mobile Cloud Computing

It is the current practical distributed computing model where both data processing and data storage occurs outside of mobile device. It is a computing paradigm which provides various outspread services to smart mobile devices. It shortens the development and execution cost of the mobile devices and enable users to access a variety of cloud services in an on-demand basis. Its main objective is to increase the computing potential of SMDs. Smartphone is a compressed mobile computing device which combines computing power of handheld equipments such as PDAs and capabilities of common cellular mobile phones. MCC provides various augmentation methods such as storage augmentation, screen augmentation, energy augmentation, and application processing augmentation to alleviate resource limitations in SMDs [7]. Three major components of MCC model are SMDs, computational cloud, and Internet wireless technology. Various wireless network protocols such as Wi-Fi, LTE, 3G, etc, can be used by SMDs so that one can easily access the services of computational clouds. Key issues occurs in mobile cloud computing are end user issues, privacy and security issues, application and services level issues, data management, context awareness issues, and operational issues. These issues highlight a certain set of challenges in MCC as depicted in Fig. 1.

MCC uses various cloud services which provide cloud processing services and online storage for extending the processing capabilities of SMDs [6]. The following section explains the concept of computation offloading which is a solution to reduce the limitations of MCC and enhances mobile systems potentiality by transferring heavy computation to rich resourceful datacenters.
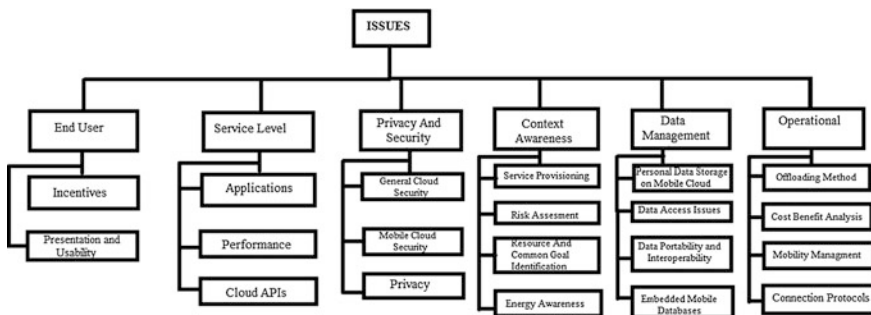
Fig. 1 Issues in mobile cloud computing

2.2 Computation Offloading

MCC utilizes the computing power of more resourceful cloud servers by offloading the computational intensive applications to cloud datacenters [8–10]. Nowadays, many computational offloading-based mechanisms have been introduced for outsourcing large applications partially or fully to remote servers [6, 11–13]. The latest computation offloading frameworks pinpoints the configuration of distributed processing of applications at runtime. Application partitioning is the most important step of offloading process and it can mainly be classified as static and dynamic.

2.2.1 Static and Dynamic Partitioning

When the offloading decision is static, it means partitioning of program is done during its development. Static partitioning involves low overhead and here the intensive application components are partitioned statically either at runtime or at compile time. This method is efficient only if all the factors are accurately known in advance. Prediction algorithms include history-based prediction [14, 15], fuzzy control [16], and probabilistic prediction [17]. It involves onetime partitioning of the application, i.e., during its design time. Here the distribution of workload occurs only once between remote servers and SMDs.

When the offloading decision is dynamic, it adapts to different runtime conditions. In dynamic offloading partitioning is done during execution of the program. This approach involves high overhead as compared to static partitioning approach because here program deals with the runtime conditions (network bandwidth, server speed). Distributed application processing frameworks is entrenched at runtime in an ad hoc manner where elastic applications are partitioned dynamically at runtime [18]. The establishment of distributed application frameworks at runtime is an energy consuming and a resource intensive method [13]. This partitioning method employs two steps, which includes application profiling and solving [11]. The mechanism where computational intensive components are identified is called application profiling, whereas application solving is the method where intensive components are separated for offloading process.

2.2.2 Algorithm for Application Partitioning and Offloading Process

Offloading Algorithm Begin:

1: Elastic mobile applications execute on SMDs and utilize the application profiling mechanism.
2: Then it evaluates availability and requirement of resources.
3: If there are insufficient resources then application solving mechanism is initialized to separate the computationally intensive applications at runtime. Go to step 5.

4: Else application executes on local SMDs. Go to step 6.

5: After partitioning, application is migrated to cloud server node for remote processing.

6: Upon successful execution of remote components of the application, the result is returned to the main application and running on SMDs.

7: Exit.

2.2.3 Various Entities Affecting Offloading Process

There are various parameters that affect the process of offloading as discussed below.

(a) Smartphone efficiency: If a smartphone has powerful processor than that of a cloud server, then taking an offloading decision is a complex process.

(b) Nature of application: If an application requires inbuilt resources, then offloading method is not beneficial.

(c) Cloud provider: The cloud provider must have the availability of powerful resources otherwise; there will not be a benefit to offload the application to the cloud.

3 Analysis of Application Partitioning Frameworks

Current offloading frameworks offload computationally intensive components at various granularity levels such as at module, class, object, bundle, thread, and method levels. At the module level, whole module is migrated for offloading process [12, 13]. In the class level partitioning, application is partitioned into classes [9, 10]. At the object level, an object is partitioned for application offloading. At the thread level, a thread is offloaded for remote processing mechanism of the applications [19]. In the same manner, bundle level granularity represents groups of classes that are partitioned [20, 21]. Migration support indicates the support level required for migrating the intensive applications to remote clouds. Various migration patterns are VM-based migration, entire application-based migration, and application partitioning-based application offloading. The main objective for application offloading of the application is to improve performance, reducing memory constraints, saving energy, and updating the application dynamically. MISCO algorithm [22] employs static partitioning which partitions the application into two functions: map and reduce. Mirror server framework [23] employed VM-based migration; the designed VM template is called mirror and the server which is responsible for its management and development is called mirror server. The critical aspects of mirror server are that it is not basically sketched for data processing due to that limited services which can be achieved from it. Elastic clone cloud [24] uses thread-based granularity at application level. A major

disadvantage of this framework is that it requires a heavy traffic synchronization mechanism. In this architecture, VM-based migration has been done because it requires more resources and an additional support from operating system. In [18] a middleware framework has been proposed for the establishment of distributed applications dynamically. Its main critical aspects are more resource consumption due to dynamic partitioning of application.

Mobile assistance using infrastructure (MAUI) [25] is a dynamic partitioning-based framework which concentrates on saving energy techniques for SMDs. It uses application proxy method at application level. The main aspects of this framework are that it requires more efforts for the development of individual method of the application. This architecture is based on method level granularity so difficulty in maintaining consistency. Elastic application model [6] is a middleware framework which implements distribution processing elastic applications at application layer. Cuervo et al. [25] and Zhang et al. [6] employed method level granularity and requires more overhead because such frameworks involves application profiling and partitioning mechanisms at runtime [26]. Investigate VM deployment and management in simulation environment can be investigated using cloudsim. Such models need feasible profiling application method which is resource starvation and a time-consuming method [27]. Considered active service migration (ASM) approach which reduces energy consumption cost (ECC) and turnaround time by deploying coarse level granularity which present the lightweight nature of proposed framework. The limitations of such frameworks are that it involves additional overhead of application offloading and more complications occur in the deployment and management of distributed applications at runtime. Shiraz et al. [28] included SaaS with IaaS for eliminating ECC during component migration at runtime such as binary code migration cost and active data state migration cost. The critical aspects of this proposed framework is that it lacks in maintaining consistency between local SMDs and cloud server node.

Since there have not been sufficient work done in order to resolve all issues; such as availability of resources, homogeneous and consistent distributed platform, security and privacy issues so there is a scope to develop a lightweight model which overcomes these issues and optimizes and enhances the overall performance of applications in MCC.

4   Conclusion and Future Scope

This paper explains the procedure of mobile cloud computing, computation offloading and analyzes current distributed application processing frameworks. In this paper, we have explored the challenges and issues to optimal and lightweight distributed applications frameworks for MCC. Further, a lightweight optimized framework is needed to be devised which will enhance the performance of application and reduces developmental efforts by overcoming the challenges related to distributed processing of applications in mobile cloud computing.

Our future work will therefore focus on proposing lightweight methods to carry out the offloaded computation tasks and to design an optimal algorithm to enhance the performance of mobile systems by investigating the system performance in a real environment.

References

1. M. Kristensen and N. Bouvin, "Developing cyber foraging applications for portable devices," in *Portable Information Devices, 2008 and the 2008 7th IEEE Conference on Polymers and Adhesives in Microelectronics and Photonics. PORTABLE-POLYTRONIC 2008. 2nd IEEE International Interdisciplinary Conference on*. IEEE, 2008, pp. 1–6.
2. J. Porras, O. Riva, and M. Kristensen, "Dynamic resource management and cyber foraging," *Middleware for Network Eccentric and Mobile Applications*, vol. 1, p. 349, 2009.
3. B. Chun and P. Maniatis, "Augmented smartphone applications through clone cloud execution," in *Proc. 8th Workshop on Hot Topics in Operating Systems (HotOS), Monte Verita, Switzerland*, 2009.
4. M. Sharifi, S. Kafaie, and O. Kashefi, "A survey and taxonomy of cyber foraging of mobile devices," *IEEE Commun. Surveys Tuts.*, 2011.
5. G. Huerta-Canepa and D. Lee, "A virtual cloud computing provider for mobile devices," in Proc. *1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*. ACM, 2010, p. 6.
6. X. Zhang, A. Kunjithapatham, S. Jeong, and S. Gibbs, "Towards an elastic application model for augmenting the computing capabilities of mobile devices with cloud computing," *Mobile Networks and Applications*, vol. 16, no. 3, pp. 270–284, 2011.
7. S. Abolfazli, Z. Sanaei, and A. Gani, "Mobile cloud computing: A review on smartphone augmentation approaches," in *Proc. 1st International Conference on Computing, Information Systems and Communications, 2012*.
8. Sanaei Z, Abolfazli S, Gani A, Buyya R. Heterogeneity in mobile cloud computing: Taxonomy and open challenges. IEEE Communications Surveys and Tutorials 2014; 16 (1):369–92.
9. Shiraz M, Ahmed E, Gani A, Han Q. Investigation on Runtime Partitioning of Elastic Mobile Applications for Mobile Cloud Computing. Journal of Supercomputing 2014; 67(1):84–103.
10. Abolfazli S, Sanaei Z, Ahmed E, Gani A, Buyya R. Cloud-based Augmentation for Mobile Devices: Motivation, Taxonomies, and Open Issues. IEEE Communications Surveys and Tutorials 2014; 16(1):337–68.
11. Cuervo E, Balasubramanian A, Cho DK, Wolman A, Saroiu S, Chandra R, Bahlx P. "MAUI: Making Smartphones Last Longer with Code Offload" MobiSys'10. June 15–18, 2010.
12. Hung, S. H., Shih, C. S., Shieh, J. P., Lee, C. P., & Huang, Y. H. (2012). Executing mobile applications on the cloud: Framework and issues. *Computers & Mathematics with Applications*, *63*(2), 573–587.
13. Shiraz, M., Gani, A., Khokhar, R. H., & Buyya, R. (2013). A review on distributed application processing frameworks in smart mobile devices for mobile cloud computing. *Communications Surveys & Tutorials, IEEE*, *15*(3), 1294–1313.
14. Gurun S, Krintz C, Wolski R (2004) NWSLite: a lightweight prediction utility for mobile devices. In: International conference on mobile systems, applications, and services, pp. 2–11.
15. Huerta-Canepa G, Lee D (2008) An adaptable application offloading scheme based on application behavior. In: International conference on advanced information networking and applications—workshops, pp. 387–392.

16. Gu X, Nahrstedt K, Messer A, Greenberg I, Milojicic D (2003) Adaptive offloading inference for delivering applications in pervasive computing environments. In: IEEE international conference on pervasive computing and communications, pp. 107–114.
17. Rong P, Pedram M (2003) Extending the lifetime of a network of battery-powered mobile devices by remote processing: a markovian decision-based approach. In: Conference on design automation, pp. 906–911.
18. A. Messer, I. Greenberg, P. Bernadat, D. Milojicic, D. Chen, T. Giul and X. Gu, "Towards a distributed platform for resource-constrained devices," in *Distributed Computing Systems, 2002. Proceedings of the 22nd International Conference on*. IEEE, 2002, pp. 43–51.
19. Goyal S, Carter J. A Light weight Secure Cyber Foraging Infrastructure for Resource—Constrained Devices, WMCSA 2004 6th IEEE Workshop. IEEE Publisher; 2–3 December 2004.
20. Giurgiu I, Riva O, Juric D, Krivulev I, Alonso G. Calling the cloud: enabling mobile phones as interfaces to cloud applications. In: Proceedings of the ACM/IFIP/ USENIX 10th international conference on middleware (Middleware'09), Urbana, IL, USA. Springer-Verlag; 2009. pp. 83–102.
21. Giurgiu I, Riva O, Alonso G. Dynamic software deployment from clouds to mobile devices. In: Proceedings of the 13th international middleware conference (Middleware'12), Montreal, QC, Canada. Springer-Verlag; 2012. pp. 394–414.
22. Dou, V. Kalogeraki, D. Gunopulos, T. Mielikainen, and V. Tuulos, "Misco: a map reduce framework for mobile systems," in *Proc. 3rd International Conference on Pervasive Technologies Related to Assistive Environments*. ACM, 2010, p. 32.
23. Zhao, Z. Xu, C. Chi, S. Zhu, and G. Cao, "Mirroring smartphones for good: A feasibility study," *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pp. 26–38, 2012.
24. Chun, S. Ihm, P. Maniatis, M. Naik, and A. Patti, "Clone cloud: elastic execution between mobile device and cloud," in *Proc. sixth conference on Computer systems*, 2011, pp. 301–314.
25. E. Cuervo, A. Balasubramanian, D. Cho, A. Wolman, S. Saroiu, R. Chandra, and P. Bahl, "Maui: making smartphones last longer with code offload," in *Proc. 8th international conference on Mobile systems, applications, and services*. ACM, 2010, pp. 49–62.
26. Shiraz M, Abolfazli S, Sanaei Z, Gani A (2012) A study on virtual machine deployment for application outsourcing in mobile cloud computing. J Supercomput 63(3):946–964.
27. Shiraz, M., Gani, A.: A lightweight active service migration framework for computational offloading in mobile cloud computing. J. Supercomput. **68**(2), 978–995 (2014). doi:10.1016/j.jnca.2014.04.009.
28. Shiraz, M., Gani, A., Shamim, A., Khan, S., & Ahmad, R. W. (2015). Energy Efficient Computational Offloading Framework for Mobile Cloud Computing. *Journal of Grid Computing*, *13*(1), 1–18.

Integration of Color and MDLEP as a Feature Vector in Image Indexing and Retrieval System

L. Koteswara Rao, D. Venakata Rao and P. Rohini

Abstract In the process of image retrieval, more information can be extracted by combining two or more features. Feature vectors based on local patterns are very popular in deriving the local information present in an image. Majority of these methods is mainly based on encoding the variation in gray scale values of center pixel and its neighboring elements. The center pixel is assigned a value which is reflected in a histogram. LBP operator became the first of its kind where the intensity value of center pixel is treated as threshold to capture the information by comparing with other neighbors. However, the information in directions is not explored in the method. The DLEPs are proposed to code the edge information mainly in four directions. The performance of directional local extrema patterns can be improved by taking the magnitude into consideration. In this paper, we propose a new feature vector for an image retrieval system by combing color and MDLEP. The results showed a significant improvement in terms of precision and recall.

Keywords Color · Texture · LBP · MDLEP · Retrieval · Precision · Recall

1 Introduction

Many images are being created across the globe as a result of significant expansion of digital media and the internet related fields. The whole purpose of generation cannot be met without a planned mechanism to handle the data. Earlier, annotation

L. Koteswara Rao (✉) · P. Rohini
Faculty of Science and Technology, IFHE, Hyderabad, India
e-mail: kots.lkr@ifheindia.org

P. Rohini
e-mail: rohini10@ifheindia.org

D. Venakata Rao
Narsaraopet Institute of Technology, Guntur, India
e-mail: dv2002in@yahoo.com

used to be done manually to index and search images from database. This way of annotation fails when the database size is large and depends on the perceptive subjectivity of the annotator. To address some of the deficiencies, a method that extracts the visual information called CBIR came into existence. Visual data such as shape, texture, and color become significant in creating the feature vector in CBIR. Formation of feature vector is the most vital stage in a retrieval framework [1–3].

Features like color and texture are extensively used in creating a feature vector to retrieve images. The most useful property of an image is the color because human eye is more sensitive to different colors in the nature. Several techniques were proposed to explore color content from the image. Texture is another property that characterizes the image according to the repetitive arrangement of pixels. Significant contributions were made to classify and part the image. In the work [4], a method to classify the texture was reported. Arivazhagan et al. [5] introduced wavelet transform to classify and segment an image. WPF and GMM were mentioned in [6] to classify the texture and segmentation. Gabor wavelets-based technique was reported in [7] to classify the texture.

1.1 Our Contribution

Existing magnitude directional local pattern extrema (MDLEP) extracts directional and magnitude data of edges as per the minima or maxima in vertical, horizontal, diagonal, anti-diagonal directions of image. In the paper, a new method is proposed that combines color feature and MDLEP to improve the output of existing MDLEP. The paper is arranged as follows. Different kinds of local patterns that are related to the work are reviewed in Sect. 2. Section 3 mentions proposed approach for image retrieval. Section 4 shows results and the discussions and conclusions are provided in Sect. 5.

1.2 Review of Related Work

An idea based on local binary pattern was explained by Ojala et al. [8], and the principle of LBP was applied in facial recognition and related areas as explained in [9]. Even though it was used in many areas, LBP has demerit of rotational variance in identifying the texture in an image. Zhang et al. [10] developed local derivative pattern by taking nth order local binary pattern. Subramanyam et al. [11] created an operator called directional local extrema pattern as a descriptor in texture analysis and classification. Reddy et al. enhanced the DLEP by taking magnitude in to consideration [12]. The MDLEP is different from the available local binary patterns and modifications in taking out directional data.

Table 1 Results for various categories of the database (in Table 1)

Category	Existing MDLEP	MDLEP+ color feature	Category	MDLEP	MDLEP+ color feature
Africans	61.3	64.4	Africans	39.25	43.7
Beach	51.25	53.7	Beach	33.82	37.5
Building	57.85	62.6	Building	31.96	36.6
Buses	94.4	98.4	Buses	73.57	77.9
Dinosaur	97.85	99.1	Dinosaur	90.28	94.5
Elephant	48.9	64.8	Elephant	30.53	34.7
Flower	89.1	93.5	Flower	69.32	77.8
Horse	66.2	79.4	Horse	36.16	45.4
Mountain	39.4	48.5	Mountain	29.35	34.1
Food	75.35	90	Food	45.3	43.5
Average precision (%)	68.16	75.44	Average recall (%)	47.954	52.57

2 Different Types of Local Patterns

2.1 Local Binary Pattern

Ojala et al. [8] introduced LBP operator. In LBP, gray scale value of center pixel is assumed as maximum level, difference in the value of center pixel and surrounding neighbors is considered to label a 0 or 1. Same procedure is followed till all elements around the center pixel get covered in the process.

$$\text{LBP}_{X,Y} = \sum_{p=0}^{p-1} y(x_p - x_c)2^p, y(c) = \begin{cases} 1 & c \geq 0 \\ 0 & c < 0 \end{cases} \tag{1}$$

where x_c is gray value of center pixel, x_p represents the intensity value of X equally spanned pixels on a circle of radius Y. For example, for the image given below, the pattern is 01110110.

22	34	49
14	27	65
71	58	16

2.2 Local Directional Pattern (LDP)

It is based on the LBP which uses edge information of neighboring pixels to code texture of image. It labels an 8-bit code to every pixel in image.

Value of 1 or 0 is coded based on the existence of an edge.

$$LDP_n = \sum_{i=1}^{8} g_i(m_i - m_k) * 2^i, g_i(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{2}$$

2.3 Directional Local Extrema Patterns (DLEP)

Principle of LBP was utilized by Subrahmanyam et al. [11] to design a novel descriptor named DLEP. In this method, two neighboring pixel intensities of one direction are compared to the value of center pixel to code 0 or 1. It describes the structure of local texture based on center pixel's extrema. Maxima and minima values of four directions can be obtained by calculating the difference between the center element and all neighbors.

The calculation is mentioned in Eq. 3.

$$M'(x_i) = M(x_c) - M(x_i); i = 1, 2, \ldots, 8 \tag{3}$$

The local extrema calculation is done according to the equations below.

$$\hat{M}_\beta(xc) = Y_3(M'(x_i) * M'(x_{j+4})); j = (1 + \beta/45)$$
$$\forall_\beta = 0°, 45°, 90°, 135° \tag{4}$$

$$Y_3(M'(x_j), M'(x_{j+4})) = \begin{cases} 1 & M'(x_j) * M'(x_{j+4}) \geq 0 \\ 0 & \text{else} \end{cases} \tag{5}$$

The DLEP is computed as ($\beta = 0°$, $45°$, $90°$ and $135°$) follows:

$$DLEP(M(x_c))|\beta = \{M_\beta(x_c); \hat{M}_\beta(x_i); \hat{M}_\beta(x_z); \ldots \hat{M}_\beta(x_8) \tag{6}$$

The details about DLEP are given in Fig. 1 and consequently, the image changed DLEP output having the values 0 to 511.

In the next level, DLEP, image is denoted by getting a histogram as per the eqn. mentioned in (7).

$$H_{DLEP|\beta}(l) = \sum_{m=1}^{Z_1} \sum_{n=1}^{Z_2} Y_2(DLEP(m, n)|_\alpha, \ell);$$
$$\ell \in (0, 511) \tag{7}$$

here the Z_1 and Z_2 are the dimensions. Collection of DLEP data for a pixel at the center in the procedure for calculation of DLEP for center pixel marked in blue
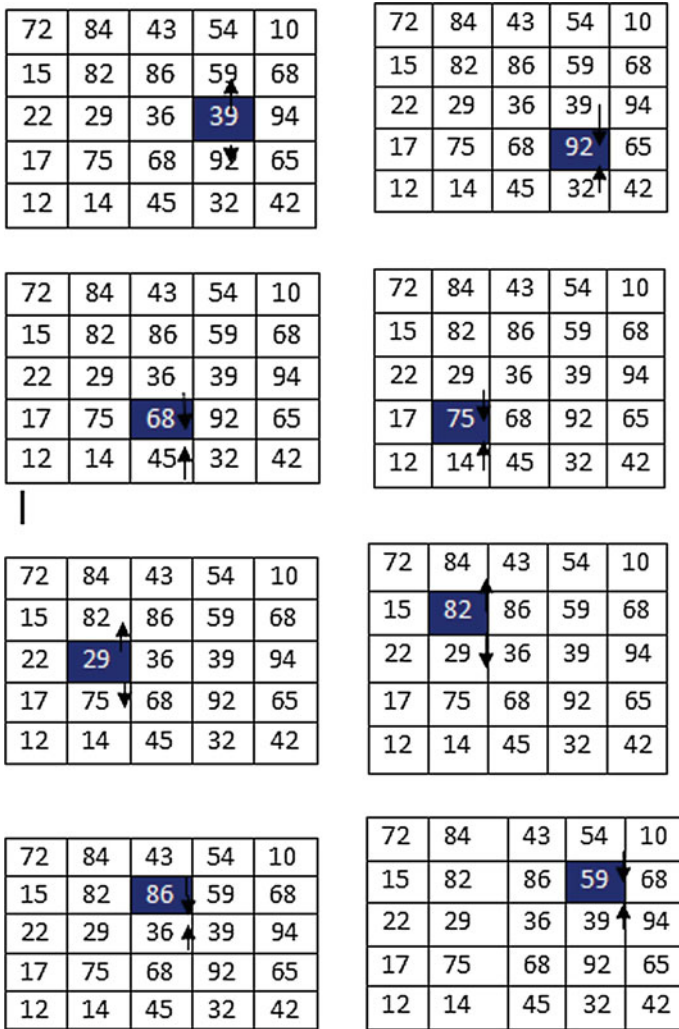
Fig. 1 Description of DLEP for 5 × 5 size

color is given Fig. 1. Information in directions is taken out using local difference between center pixel and neighbors.

For example, DLEP of 90° direction for pixel highlighted with blue color is given in Fig. 2. For a pixel of value 36, two neighbor pixels are moving away in values. Hence, pattern is assigned 1. Similarly, remaining bits of DLEP are collected making the final outcome as 110011110. In this way, the DLEPs are computed in 0°, 45°, and 135° directions.

	$0_{(27)}$	$1_{(29)}$	$2_{(80)}$	$3_{(87)}$	$4_{(88)}$	$5_{(13)}$	$6_{(78)}$	$7_{(85)}$	$8_{(63)}$	DLEP
$P\,(0^0)$	0	0	0	0	1	1	0	1	0	26
$Q(45^0)$	1	0	0	1	1	1	1	0	1	317
$R(90^0)$	1	1	0	0	1	1	1	1	0	415
$S(135^0)$	1	1	0	0	0	1	1	1	0	398

Fig. 2 Calculation of DLEP in 90°

2.4 *Magnitude Directional Local Extrema Patterns (MDLEP)*

Reddy and Reddy [12] presented a technique to increase performance by collecting magnitudes of the local patterns. MDLEP is collected according to the equation below.

$$\hat{I}_{M\beta}(xc) = Y_3(I'(x_i) * I'(x_{j+4})); j = (1 + \beta/45)$$
$$\forall_\beta = 0°, 45°, 90°, 135° \tag{8}$$

$$Y_4(I'(x_j), I'(x_{j+4})) = \begin{cases} 1 & abs(I'(x_j)) + abs(I'(x_{j+4})) \geq \text{Thrs} \\ 0 & \text{else} \end{cases} \tag{9}$$
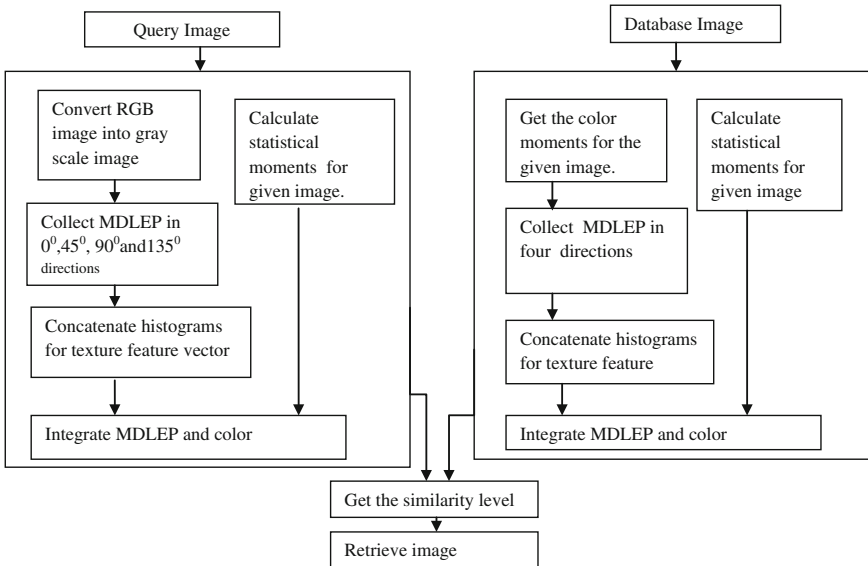
Fig. 3 Block diagram of CMDLEP for image retrieval

$$\text{Thrs} = \frac{1}{Z_1 \times Z_2} \left(\sum_{b=1}^{z_1} \sum_{c=1}^{z_2} abs\left(I'(x_j)|_{(b,c)}\right) + abs\left(I'(x_{j+4})|_{(b,c)}\right) \right) \qquad (10)$$

The MDLEP in 0°, 45°, 90°, and 135° directions is defined as

$$\text{MDLEP}(I(x_c))|\beta = \left\{ \hat{I}_{M\beta}(x_c); \hat{I}_{M\beta}(x_1); \hat{I}_{M\beta}(x_2); \ldots \hat{I}_{M\beta}(x_8) \right\} \qquad (11)$$

Subsequent to the MDLEP calculation, entire image is shown by a histogram according to Eq. (8)

3 Proposed CMDLEP System

The framework of the proposed method is shown in Fig. 3 above.

Algorithm

1. Calculate color moments for given image and convert RGB into gray scale image.
2. Compute the local extrema in 0°, 45°, 90°, and 135° directions.
3. Compute the MDLEP information in all four directions as per Step 2.
4. Get histogram of MDLEP obtained from Step 3 and join to create feature vector
5. Combine the two features to form a feature vector used in the process of retrieval

Query matching

After the feature extraction, the feature vector of query image is created. Similarly, feature vectors of all images from database are created. To identify a relevant image to query image, distance between query image and repository images is considered.

4 Experimental Results

Capability of proposed method is tested with Corel-1 k database [13]. Precision (P_r) and recall (R_e) values are calculated based on the equations below. The results are given in Figs. 4 and 5.

$$P_r = \frac{\text{Number of relevant images retrieved}}{\text{Number of images retrieved}}$$

$$R_e = \frac{\text{Number of relevant images retrieved}}{\text{Number of relevant images in the database}}$$
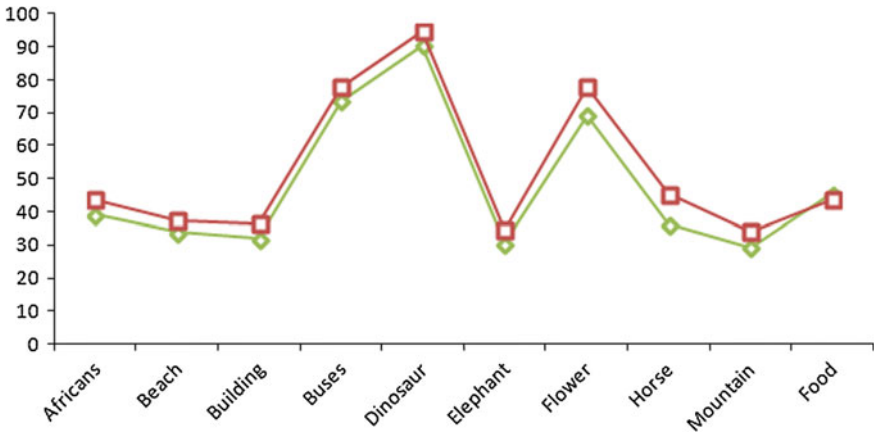
Fig. 4 Average precision of MDLEP (*green*) and CMDLEP (*red*)
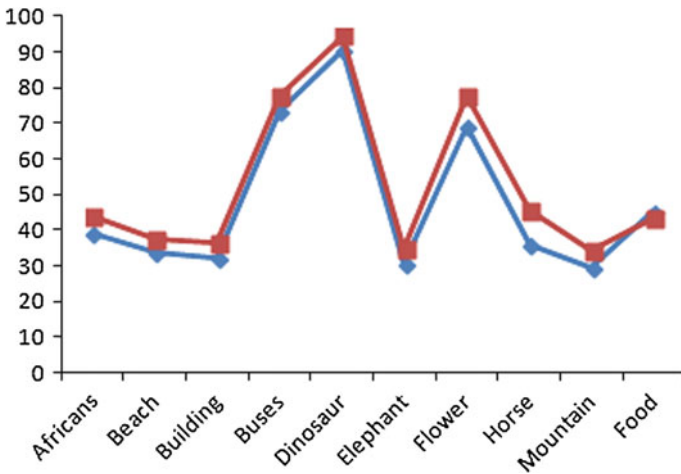
Fig. 5 Average recall of MDLEP (*red*) and CMDLEP (*blue*)

5 Conclusion

It is apparent that our method is outperforming the recent MDLEP in terms of precision and recall. Combination of color and MDLEP is exploring more information existing in an image when compared to LBP and similar local pattern-based methods.

References

1. R. Datta, D. Joshi J. Li and J. Wang, "Image Retrieval-Ideas, influences and trends of the new age", ACM Computing surveys, vol. 40, no. 2, pp. 1–60 2008
2. AM Smeulders, M Worring, S Santini, A Gupta & R Jain, "content based Image retrieval at the end of early years" IEEE Transactions aon PAMI 22(12), pp 1349–1380, 2000
3. Y Rui, T S Huang & S F Chang, "Image retrieval: current techniques, promising directions& open issues. Journal of visual communications & Image representation 10(4): pp 39–62, 1999.
4. R M Haralick, K Shanmugam & I Dinstein, "Texture features for image classification", IEEE transactions on system, man and cybernetics vol. smc-8, pp 610–621, 1973
5. S Arivazhagan and L Ganesan, Texture classification using wavelet transform(1513–1521) vol. 24, issue 9–10, June 2003.
6. Soo Chang Kim, Tae Jin Kang, Texture classification and segmentation using wavelet packet frame and Gaussian mixture model vol 40, issue 4, April 2007, 1207–1221 elsevier
7. S Arivazhagan and L Ganesan, 'Texture classification using Gabor wavelets based rotation invariant features, vol 27, ISSUE 16, December 2006 (1976–82)
8. Ojala T, Pietikainen M, Harwood D (1996) A comparative study of texture measures with classification based on feature distributions. J Pattern Recognition 29(1):51–5
9. A Hadid, T Ahonen and M Pietikainen, "Face analysis using local binary patterns," in handbook of Texture analysis, Imperial college press London 2008, pp 347–373
10. Zhang B, Gao Y, Zhao S, Liu J (2010) Local derivative pattern versus local binary pattern: Face Recognition with higher-order local pattern descriptor, IEEE Trans Image Process 19 (2):533–544
11. Subrahmanyam Murala, R.P. Maheswari, R. Balasubramanian Directional local extrema pattern: a new descriptor for content based image retrieval (2012)
12. Reddy et al. Content based image indexing and retrieval using directional local extrema and magnitude patterns, International journal of electronics and communication 68(2014) 637–643
13. http://wang.ist.psu.edu/docs/related/

Reversible Data Hiding Through Hamming Code Using Dual Image

Biswapati Jana, Debasis Giri and Shyamal Kumar Mondal

Abstract In this paper, we propose a new dual-image based reversible data hiding scheme through Hamming code (RDHHC) using shared secret keys. We collect a block of seven pixels from cover image and copy it into two arrays. We then adjust redundant LSB bits using odd parity such that any error creation encountered are recovered at the receiver end. Before data embedding, we first complement bit at the position of shared secret key of the block. After that we embed secret message bit by error creation in any position of the block except key position. The receiver detects and corrects the error using Hamming error correcting code. Two shared secret keys κ and ξ help us to perform these data hiding and recovery operations. We distribute two stego pixel blocks among dual image based on the bit pattern of shared secret key ξ of length 128 bits. Finally, we compare our scheme with other state-of-the-art methods and obtain reasonably better performance in terms of security and quality.

Keywords Reversible data hiding · Steganography · Hamming code · Least significant bit · Dual image

B. Jana (✉)
Department of Computer Science, Vidyasagar University, Midnapore 721102,
West Bengal, India
e-mail: biswapati.jana@mail.vidyasagar.ac.in

D. Giri
Department of Computer Science and Engineering, Haldia Institute of Technology,
Haldia 721657, West Bengal, India
e-mail: debasisgiri@hotmail.com

S.K. Mondal
Department of Applied Mathematics with Oceanology and Computer Programming,
Vidyasagar University, Midnapore 721102, West Bengal, India
e-mail: shyamal260180@yahoo.com

1   Introduction

The data hiding protect multimedia content by concealment of embedding message through eavesdroppers until the secret message is extracted. The goal of data hiding is to ensure embedded data extraction and original cover image reconstruction. Recently, Kim et al. [1] proposed data hiding using Hamming code (DHHC) to hide data into halftone image. Chang et al. [2] proposed a high payload steganographic scheme based on (7, 4) Hamming code for digital images. Ma et al. [3] proposed a scheme to improve Kim's scheme [1] by changing pair of pixel rather than individually which sacrifice data hiding capacity reduced by half. Lien et al. [4] proposed a dispersed data hiding scheme using Hamming code (DDHHC) through space filling curve decomposition. Pixels consist in each block randomly and distributed uniformly all over the cover image. Using Hamming code Kim et al. achieve good qualities stegos and their MPSNR is 48.20 dB and payload 0.86 bpp and Lien et al. achieve 29.66 dB of MPSNR for embedding 65, 536 bits. So far, no researcher has considered reversible data hiding using Hamming code. In this paper, we propose dual-image based reversible data hiding scheme using Hamming code (RDHHC).

Motivation:

In this paper, we introduce a new dual-image based reversible data hiding scheme through Hamming code. Our motivation of this research are stated as follows:

- Our motivation is to communicate any arbitrary length of secret message through Hamming code. The scheme is designed in such a manner that receiver can easily find the end of secret message without knowing the length of secret message. The data extraction process process will be stopped when receiver will find no error continuously in the stego images, as because data are embedded through error creation caused by tamper.
- Another motivation is to enhance security in data hiding schemes. We use two shared secret keys κ and ξ to enhance security. Also we update κ for the next block using the formula, $\kappa_{i+1} = ((\kappa_i \times \omega) \mod 7) + 1$, where ω is the data embedding position. The second shared secret key ξ is used to distribute stego pixel block among dual image.
- Another aim of this work is reversibility. Hamming error correcting codes are used to detect the error position that means message embedding position within stego images. After retrieving the secret data bits, receiver can correct the error by complementing that bits. Here, we use dual stego images SM′ and SA′. The redundant bit 1, 2, and 4 positions are used in SM′ which are modified during odd parity adjustment and other remaining bits position are unchanged and in second stego image SA′, bit positions 3, 5, 6, and 7 are used as redundant bit which are modified during odd parity adjustment and remaining bit positions are unchanged. Thus, after extraction the secret message from dual stego images, we combine 3, 5, 6, and 7 positions from SM′ and 1, 2, and 4 positions from SA′ and rearrange to recover cover image that is to say reversibility.

The rest of the paper is organized as follows. Our proposed method is discussed in Sect. 2. Experimental results with comparisons are given in Sect. 3. Finally, some conclusions are given in Sect. 4.

2   Proposed Reversible Data Hiding Through Hamming Code (RDHHC)

We introduce a dual-image based reversible data hiding scheme using Hamming code. The schematic diagram of data embedding stage is depicted in Fig. 1 and data extraction and reconstruction of cover image is shown in Fig. 2.

2.1   Embedding Stage in RDHHC

We partition the cover image into blocks of seven consecutive pixels. Then we convert pixel block to LSB block and copies it into two arrays M and A. We then apply odd parity to adjust redundant bits of both arrays separately. The redundant bits r1, r2, and r3 of M array are adjusted based on the number of 1's present in the bit position of M which is shown in Tab 1 of Fig. 1. For example, the r1 bit is set to 1 if the number of 1's present in the positions 3, 5, and 7 of M array is even. The redundant bits r1, r2, r3 and r4 of A array are adjusted and update in 3, 5, 6 and 7 bit positions of A depending on the number of 1 present in the bit position shown in Tab 2 of Fig. 1. The r1 is set to 1 if the number of 1 present in the 1 and 2 index positions of A is even and update r1 at the index position 3 of the A array. In this way, we adjust all redundant bits in both the images.

After that, we complement the bit at the position of κ (here it is 4 in Fig. 1) in the first row of M then embed secret data bit (here it is 1) by error creation in any position except the key position. So, the positions 3, 5, 6, and 7 are the suitable location for data embedding where error creation is possible (here we choose 3). Now, the key ω is the data embedding position of M that is 3 here and that key is used as key for A during data embedding in the array A. Then we do the same operation for key insertion and data embedding on A to embed next data bit. After creating error for the key ω, data bit is embed by error creation and 2, 4, 6, and 7 positions are the suitable position (here we choose 2). Now, the value of ω is 2 and for next pixel block the key κ is updated by $\kappa_{i+1} = ((\kappa_i \times \omega) \bmod 7) + 1$. The updated key κ is now 2 and is used for next block. We continue the embedding process and modify the pixels value accordingly and generate two stego images: Stego Major (SM) and Stego Auxiliary (SA). Finally, depending upon the key ξ of bit length 128 bits, the modified pixel blocks are distributed among two final stego images: SM′ and SA′. If ($\xi = 1$) then selected pixel from SM is stored on
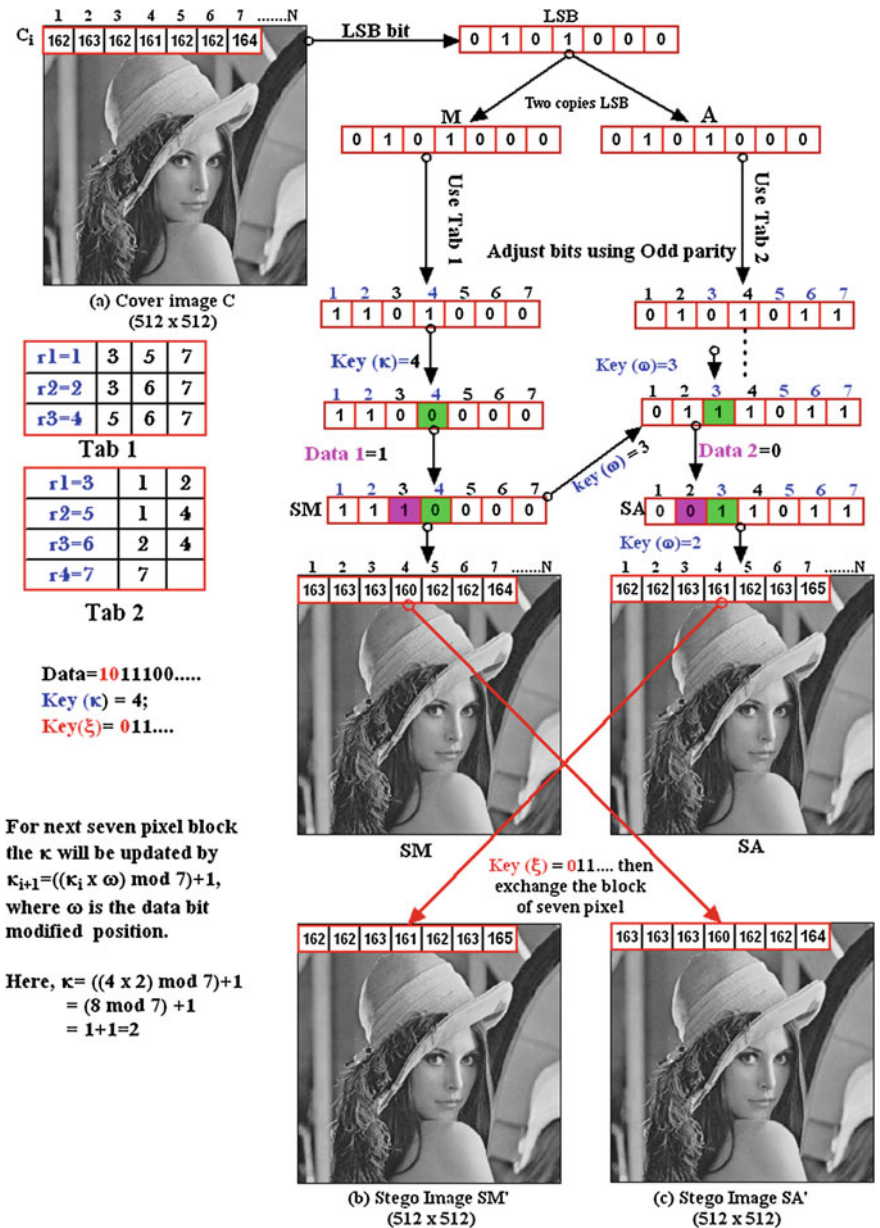
Fig. 1 Schematic diagram of data embedding stage in RDHHC

SM′ and pixel from SA is stored on SA′ otherwise pixel from SM is stored on SA′ and pixel from SA is stored on SM′. In this way, we distribute the modified pixel among two stego images.
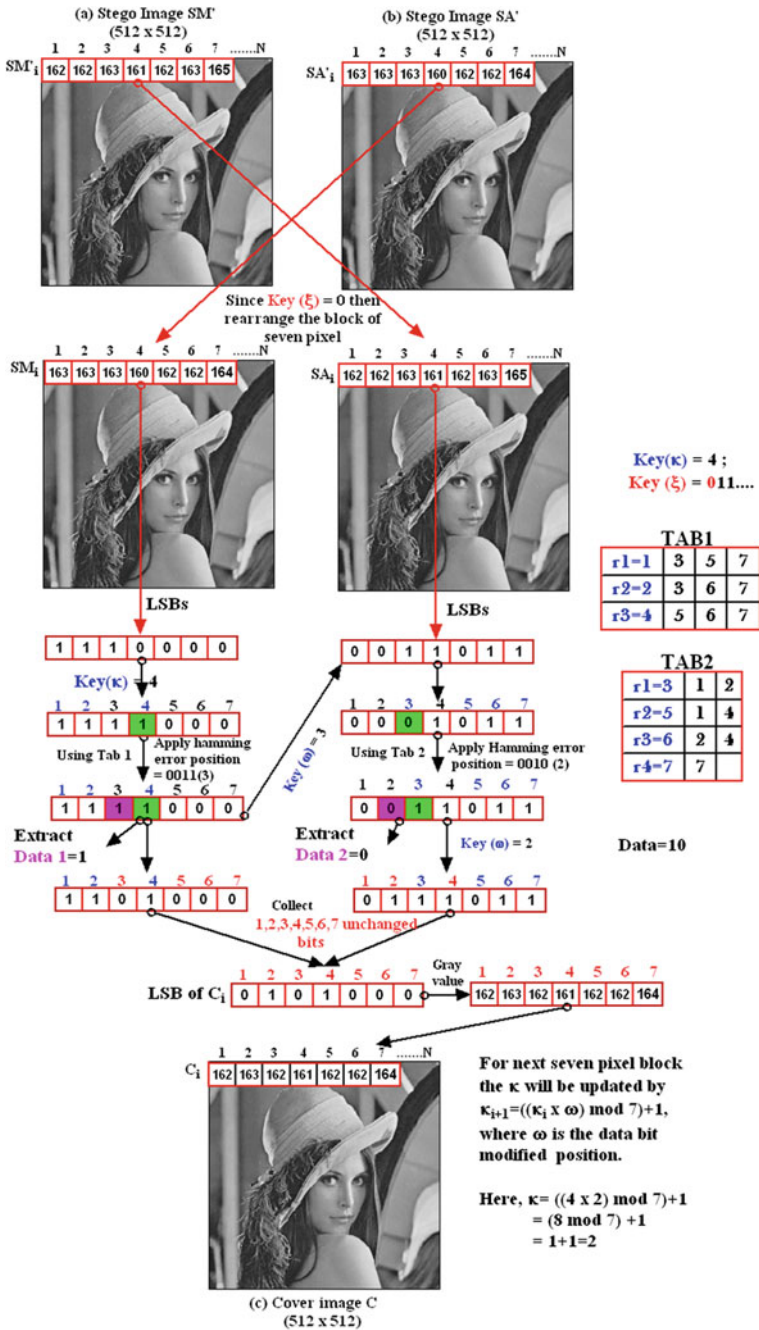
Fig. 2 Schematic diagram of data extraction and image recovery stage in RDHHC

2.2 Extraction and Image Recovery Stage in RDHHC

After receiving dual stego images SM′ and SA′, we first apply the key ξ (128 bits) to rearrange the pixel blocks as shown in Fig. 2. If ($\xi = 1$) then selected pixel block from SM′ is stored on SM_i and pixel block from SA′ is stored on SA_i otherwise pixel from SM′ is stored on SA_i and pixel block from SA′ is stored on SM_i. Now, we collect LSBs of a block of seven pixels from both stego images SM_i and SA_i. To extract the secret data, we first complement the bit at the position of secret key κ of the first row of SM_i. After that we apply the Hamming error correcting code to find the error position. In case of SM_i stego image consider the redundant bit as mentioned in Tab 1 of Fig. 2. The error position of SM_i is the data embedding position and that position is the key ω for SA_i during extraction. We complement the bit at the κ position in SA_i then find the error through Hamming code using redundant bit mentioned in Tab 2 of Fig. 2. Then again the key ω is updated by the error location of SA_i. Now the key κ is updated for the next row using formula $\kappa_{i+1} = ((\kappa_i \times \omega) \mod 7) + 1$, where, ω is the error position of SA_i (here it is 3 in Fig. 2). After extract the message bit, complement the corresponding position to generate Hamming adjusted cover image. We continue this extraction process for all secret data from dual image. The extraction process is to be stopped when receiver found no error in any pixel block. That means no error is created during data embedding process. As a result, we can send any arbitrary length of data using this scheme. After that from both the Hamming adjusted cover image, we collect the bits from the position 3, 5, 6, and 7 from SM_i and bits from the position 1, 2, and 4 from SA_i to construct original image. This scheme extract secret data and recover cover image successfully.

3 Experimental Result and Comparison

The scheme is tested through MATLAB Version 7.6.0.324 (R2008a). For our experiment we have used (512 x 512) gray scale original image shown in Fig. 3. The dual stego images are generated after embedding maximum secret data bits shown in Fig. 4. The qualities of stego images are measured using mean square error (MSE) and peak signal-to-noise ratio (PSNR) as per Eq. (1) and Eq. (2) respectively.

$$MSE = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} [X(i,j) - Y(i,j)]^2}{(M \times N)} \tag{1}$$

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

where M and N are the height and width of the image and $X(i,j)$ is the pixel value of the original image and $Y(i,j)$ is the pixel value of the stego image. Higher the PSNR
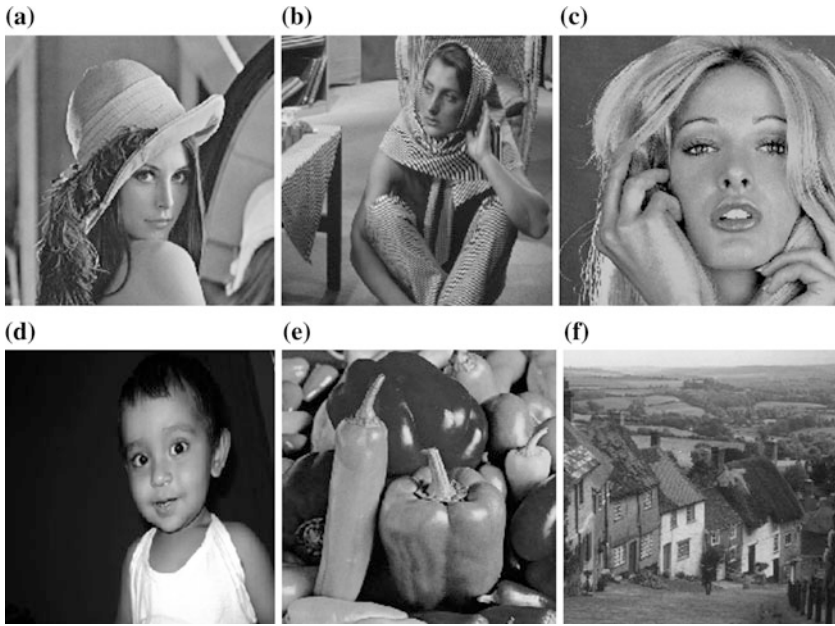
Fig. 3 Standard cover images with (512 × 512) pixel. **a** Lena (512 × 512). **b** Barbara (512 × 512). **c** Tiffany (512 × 512). **d** Mrittika (512 × 512). **e** Peppers (512 × 512). **f** Goldhill (512 × 512)
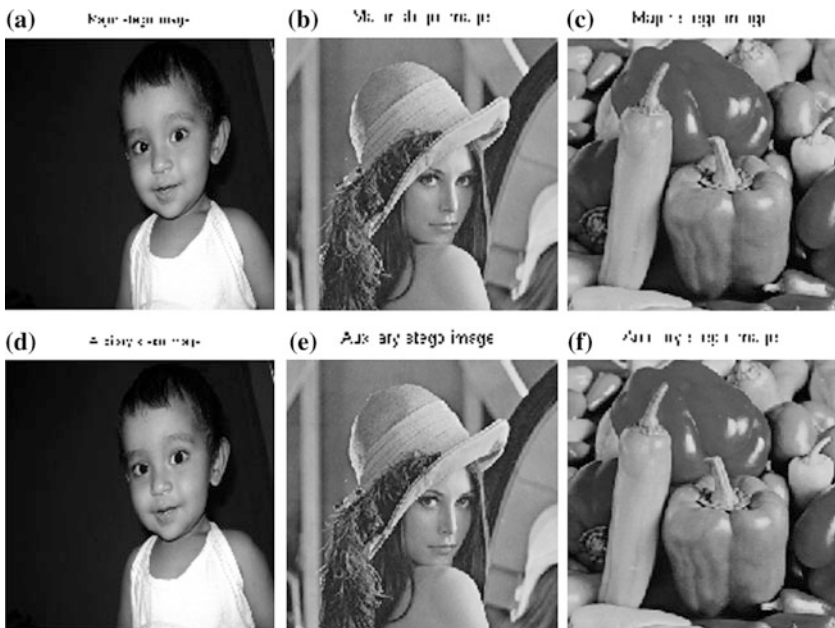
Fig. 4 Dual stego images with (512 × 512) pixel. **a** Mrittika Stego Major (SM). **b** Lena Stego Major (SM). **c** Peppers Stego Major (SM). **d** Mrittika Stego Major (SM). **e** Lena Stego Major (SM). **a** Peppers Stego Major (SM)

better the image quality and lower PSNR implies poor image quality. Table 1 shows the image quality with different embedding capacity. PSNR (C and SM) and PSNR (C and SA) represent the quality of the first and second stego image, respectively, while Avg. PSNR is the average quality of the stego image. From the Table 1 average PSNR of our RDHHC scheme is greater than 53 dB and the maximum embedding capacity (P) is $2 \times (512 \times 512/7)$ that is 74, 898 bits. The payload is measured by $B = P/(2 \times M \times N)$ bpp, where P is the total embedding capacity of two stego images. The payload in RDHHC scheme is 0.14 bpp.

Our principle is to keep message data as short as possible while using steganography and Kirchhoff's principle say that everyone knows the algorithm but the secrecy depends on the secret key. So, to enhance the security on data hiding we use two different shared secret key during embedding and extraction stage. We compare our RDHHC scheme with other existing data hiding method shown in Table 2. Three sets of secret data bits 65, 536 bits, 16, 384 bits, and 4, 096 bits are

Table 1 PSNR of dual image with embedding capacity

PSNR (dB) with data embedding capacity (bits)				
Original image C	Secret data (bits)	PSNR (C and SM)	PSNR (C and SA)	Avg. PSNR
Barbara	18,720	57.84	57.88	54.27
	37,520	54.84	54.84	
	64,800	52.47	52.47	
	74,752	51.86	51.96	
Lena	18,720	57.55	57.64	53.96
	37,520	54.27	54.34	
	64,800	52.18	52.29	
	74,752	51.85	51.56	
Peppers	18,720	57.47	57.50	53.94
	37,520	54.22	54.20	
	64,800	52.15	52.25	
	73,728	51.84	51.94	
Mrittika	18,720	57.27	57.70	54.02
	37,520	54.20	54.45	
	64,800	52.63	52.74	
	73,728	51.65	51.94	
Tiffany	18,720	57.56	57.59	53.63
	37,520	54.29	54.28	
	64,800	52.21	52.31	
	74,752	51.88	51.97	
Goldhill	18,720	57.11	57.90	53.73
	37,520	54.28	54.29	
	64,800	52.17	52.26	
	74,752	51.86	51.97	

Table 2 Comparison with the MPSNR of DHHC [4] and DDHHC [3] scheme with PSNR of RDHHC

Embedded bits	65,536 bits			16,384 bits			4096 bits		
Scheme	DHHC	DDHHC	RDHHC	DHHC	DDHHC	RDHHC	DHHC	DDHHC	RDHHC
Lena	25.71	30.57	52.20	32.03	38.60	56.09	38.12	44.71	64.68
Barbara	25.69	30.38	52.21	32.03	38.57	56.86	38.14	44.77	64.09
Tiffany	24.71	27.15	52.11	31.72	36.24	56.73	38.04	42.91	63.81
Pepper	25.60	29.90	52.12	31.98	38.02	56.16	38.10	44.19	64.39
Gold hill	25.66	30.30	52.20	32.02	38.66	56.94	38.13	44.96	64.71

used to compare with Kim et al.'s [1] (DHHC) and Lien et al.'s [4] (DDHHC) scheme. With the same embedding capacity with 16, 384 bits the MPSNR of DHHC is 24.06 dB less and DDHHC is 17.49 dB less than our proposed RDHHC. In terms of security, our scheme is superior with respect to other existing schemes, because we use two shared secret keys, κ and ξ. In this scheme, the maximum possible number of keys are $(512 \times (512/7))$ that is 37,449 for κ (values between 1 to 7) and possible number of combination of secret key ξ of length 128 bits are 2^{128}. The scheme become robust against various attacks due to the employment of these shared secret keys.

4   Conclusion

In this paper a dual-image based reversible data hiding scheme using Hamming code (RDHHC) is introduced. Here, shared secret keys are used to embed confidential message among dual image. Dual image is required to extract message and recover the original image with the help of shared secret keys. In RDHHC scheme, PSNR is better than that of the other existing Hamming code based schemes, which implies that in terms of visual quality, our scheme is better. Also any arbitrary length of secret message can be communicated through our RDHHC scheme. We use two shared secret keys which are used in both data embedding and extraction stage in every block and change accordingly, which guarantees security and it is hard to break due to huge number of possibilities of shared secret keys. In this proposed RDHHC method, Hamming code based data hiding achieves reversibility which demands the originality of our research work.

References

1. Kim C., Shin D. and Shin D., Data hiding in a halftone image using Hamming code (15,11), ACIIDS 2011, LNAI 6592, pages. 372−381. Springer Verlag Berlin Heidelberg (2011).
2. Chang CC, Kieu TD, Chou YC, A high payload steganographic scheme based on (7, 4) Hamming code for digital images, Int Symp Elect Commerce Security: pp. 1621, (2008).
3. Ma Z., Li F., Zhang X., Data Hiding in Halftone Images Based on Hamming Code and Slave Pixels, Journal of Shanghai University(Nature Science) vol. 19, No. 2, pages. 111−115, (2013).
4. Brian K. Lien, Shan-Kang Chen, Wei-Sheng Wang, kuan-Pang King, Dispersed Data Hiding Using Hamming Code with Recovery Capability, Advances in Intelligent Systems and Computing Springer International Publishing Switzerland, vol. 329, pages. 179−187, (2015).

A Hybrid Fault Tolerant Scheduler for Computational Grid Environment

Ram Mohan Rao Kovvur and S. Ramachandram

Abstract A computational grid is an environment for achieving better performance and throughput by pooling geographically distributed heterogeneous resources dynamically depending on their availability, capability, performance, and cost and user quality of self-service requirement. Fault tolerant grid scheduling is a significant concern for computational grid systems. The handling of failures can happen either before or after scheduling tasks on grid resources. Generally, there are two approaches used for the handling of failures namely, post-active fault-tolerant approach and pro-active fault tolerant approach. Recently, a fault tolerant scheduler for grids proposed uses pro-active approach in selecting resources by computing scheduling indicator. However, this study did not considered failure of node while the task is being executed. Thus in our study, we incorporates post-active fault tolerant approach to the exiting study, i.e., migrating of task to another node in the event of failure of node while the task is being executed. We constructed a hybrid fault tolerant grid scheduler using GridSim 4.2 toolkit. We demonstrated that our proposed fault tolerant scheduler shows better results in terms of success rate in comparison with the existing fault tolerant scheduler.

Keywords Grid · Fault · Scheduler · Dynamic · Hybrid

R.M.R. Kovvur (✉)
Department of Computer Science and Engineering, Vasavi College of Engineering,
Ibrahimbagh, Hyderabad, India
e-mail: krmrao@rediffmail.com

S. Ramachandram
Department of Computer Science and Engineering, Osmania University, Hyderabad, India
e-mail: schandram@gmail.com

1 Introduction

Grid is a type of parallel and distributed System that enables the sharing, selection, and aggregation of geographically distributed resources dynamically at run time depending on their availability, capability, performance, and cost and user quality of self-service requirement [1, 2].

A computational grid is an environment that organizes dynamic distributed heterogeneous resources over multiple administrative domains with diverse protection strategy into a single computing system. This grid environment allows sharing of resources in solving distributed applications in dynamic, multi-institutional virtual organizations. Some of these applications include weather forecasting, earthquake engineering, bioinformatics, video gaming, biomedical imaging, and astrophysics, etc. [3, 4].

Grid scheduling is defined as the process of mapping Grid jobs to resources over multiple administrative domains. The scheduler has the responsibility of selecting resources and scheduling jobs in such a way that the user and application requirements are met, in terms of Makespan and cost of the resources utilized [5, 6].

Fault-tolerant scheduling is one of the vital concerns for computational grid systems. The role of fault tolerance in grid system is a "to preserve the delivery of accepted services despite the presence of faults within the environment." Faults are detected and corrected, and permanent faults are located and removed while the system continues to deliver acceptable service [7, 8].

The handling of failures can occur either before or after scheduling tasks on grid resources. There are two approaches are used for the handling of failures namely, post-active fault tolerant approach and pro-active fault tolerant approach. In the post-active approach, handling of failures can happen after scheduling of job to the grid resources. This approach does not take into consideration the failure history of the resources while scheduling resources. This means that the resource with the minimum response time is selected regardless of its failure history. Whereas, the pro-active approach, the handling of failures can happen before scheduling of job to the grid resources. This approach takes failure history of resources into consideration before scheduling. Thus, this approach reduces the failure rates within the grid and also increases the performance and throughput [2, 9, 10].

Recently, a fault tolerant scheduling system for computational grids [2] was proposed with pro-active approach in selecting resources by computing scheduling indicator. This indicator comprises of the response time and the failure rate of grid resources. Whenever a grid scheduler has tasks to schedule on grid resources, it utilizes the scheduling index to build the scheduling decisions. The key scheduling strategy of this technique is to select resources that have the lowest tendency to fail. However, this study did not considered failure of node while the task is being executed. Thus in our study, we incorporates post-active fault tolerant approach to the exiting work, i.e., reassign of task to another node in the event of failure of node while the task is being executed caused by node overload.

2 Components of the Proposed Hybrid Fault Tolerant Scheduler

The interaction between the components of the proposed system is shown in Fig. 1.

The proposed scheduling system has five major components namely (i) grid portal (ii) scheduler (iii) grid resources (iv) resource information server (RIS) (v) fault handler with pro-active and post-active approaches

- *The Grid Portal*
  A Grid portal provides an interface for a user to submit their yasks for execution that will utilize the resources and services provided by the grid.
- *The Scheduler*
  The scheduler selects the optimal resources to execute the task. The scheduling decisions of the scheduler are based on the response time and the fault rate of the grid resources.
- *Grid Resource*
  A computing grid resource GR = $\{N_1, N_2 \ldots N_m\}$ can be local or remote grid service provider. Each node *Ni* is associated with two values such as computing speed and failure frequency of the node.
- *The Resource Information Server (RIS)*
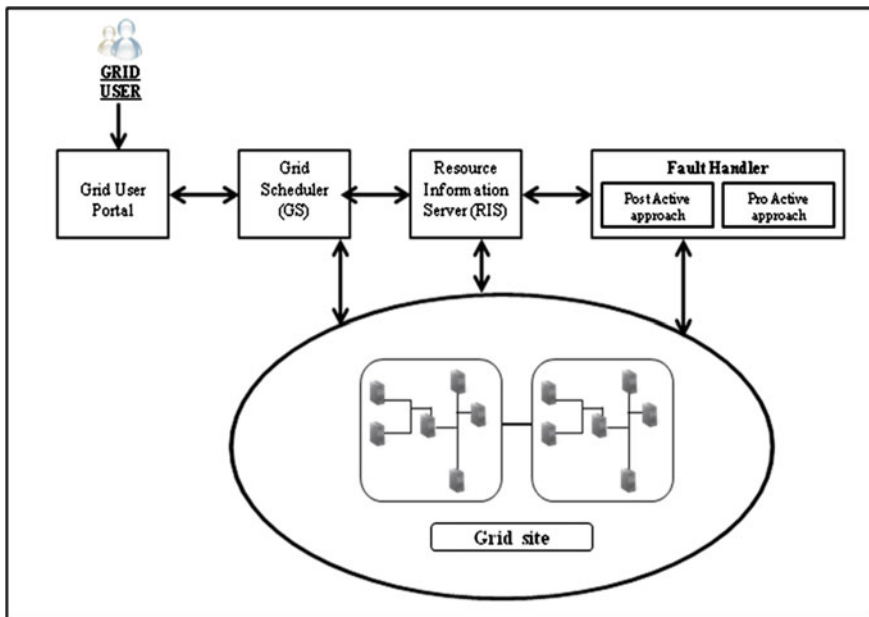
Fig. 1 Components of fault tolerant scheduling system

The resource information server contains information about all resources in the grid. The information can include computation speed, memory available, load, and so on. The RIS supplies the scheduler with the required information.

- *The Fault Handler*

  The fault handler is responsible for detecting faults of resources and estimating the rate of resources fault. It uses two fault handler approaches namely, post-active and pro-active approaches.

  - *The Pro-active approach*, the handling of failures can happen before scheduling of task to the grid resources. This approach takes failure history of resources and it is used to compute the scheduler index (SI) value.
  - *The Post-active approach*, handling of failures can happen after scheduling of task to the grid resources. This approach uses re-scheduling component, i.e., rescheduling of task to another resource in case of failure of resource while the task is being executed.

3   Hybrid Fault Tolerant Grid Scheduler Algorithm

HFTGS-scheduler procedure initially computes completion time of task on all the nodes at grid site and then computes failure rate of each node and scheduling index [2] as proposed by Mohammed Amoon. Further the nodes in the resource list is sorted in the increasing the order of completion times.

Then the task *Ti* is assigned to the first node of the resource list. In case the assigned grid node is down/fail, the task is reassigned to next available node of resource list and update grid node information with the fault handler. In case of the failure of task during execution, the executed task is sent to post-active fault handler to reassign the task to the next available node.

```
Procedure HFTGS (  )
Input:  Task queue Q [ i ]
          AvailableResourcelist  R [ j ]
Begin
        Step 1.0 For each task T i in Q[ i ]
        Begin
        Step 2.0 For each node Nj in R [j]
                    2.1 compute execution time of Ti on Nj
                    2.2 compute failure rate of Nj
                    2.3 compute schedule index (SI)
         Step 3.0 Sort the Resourceslist  R[j] in the increasing order with respect to SI
         Step 4.0 Assign task Ti on Resource Node R[j]
         Step 5.0 if (Status of Assigned Node! = 'Fail')
                    5.1 Set the Status Nj as busy
                    5.2 if (Node Ni is Failed during execution)
```

        5.2.1 Reassign task Ti on R [j+1] ;
        5.3 If (Task Ti completes execution on Nj)
              Increment Success count;
        5.4 Else
              Increment Failure Count;
    Step 6.0 Else If (Assigned Node == 'Fail')
              6.1 Reassign Task Ti on R[j+1] ;
              6.2 Goto Step 5.0
    End
  Next Task in the Queue Q[i+1]
 End

4 Simulation Environment and Results

In this section, we illustrate our simulation environment used and present the simulation results by comparing proposed hybrid fault tolerant scheduler with existing fault tolerant scheduler.

4.1 Simulation Environment

We used GridSim [3] simulator for simulating grid environment and performed extensive simulations and obtained the simulation results by computing successful schedule percent by varying resource failure rate. In our study, we considered the following parameters at the grid user with task length in million instructions (MI), task file size, and task output size in bytes and at the grid site with computing nodes $N = \{n1, n2 \ldots nm\}$ with computing speed of each nodes in million instructions per seconds (MIPS).

At grid site, there is a grid resource (GR) comprises of multiple nodes with different processing capabilities and a grid user submits several tasks to the grid scheduler. The communication within the site (among the grid nodes) is 100 Mbps. In our study, we assumed the number of grid nodes as 10 % of the tasks under consideration.

4.2 Simulation Results and Discussion

Scenario 1: Computing successful schedule percent by increasing number of tasks/gridlets with 20 % of node failures.

We evaluated the proposed and the existing fault tolerant grid scheduler with respect to the success schedule percent of grid system by increasing number of tasks/gridlets from 100, 200, 300, 400, and 500 with 20 % of node failures in the grid environment. From Fig. 2, it is noticed that the proposed scheduler performed better in terms of successful schedule percent by 0.8 % in contrast with existing scheduler.

Fig. 2 Computing success
percent by varying number of
tasks at 20 % of node failures
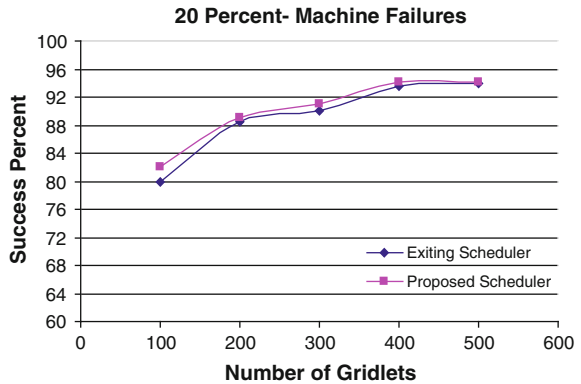
Fig. 3 Computing success
percent by varying number of
tasks at 25 % of node failures
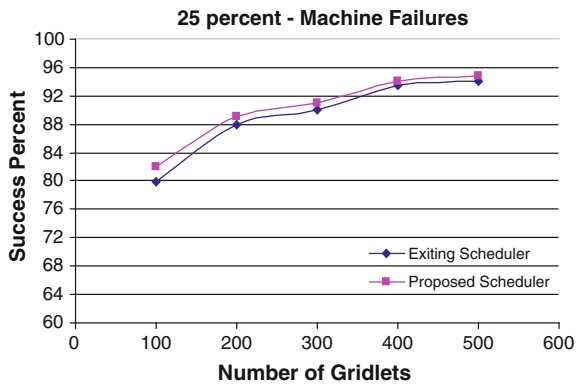
Scenario 2: Computing successful schedule percent by increasing number of tasks/gridlets with 25 % of node failures.

We evaluated the proposed and the existing fault tolerant grid scheduler with respect to the success schedule percent of grid system on an increasing number of tasks/gridlets 100, 200, 300, 400 and 500 with 25 % of machine failures in the grid environment. From Fig. 3, it is noticed that proposed scheduler performed better in terms of successful schedule percent by 1 % in contrast with existing scheduler.

Scenario 3: computing successful schedule percent by increasing number of tasks/gridlets with 33 % of machine failures.

we evaluated the proposed and the existing fault tolerant grid scheduler with respect to the success schedule percent of grid system on an increasing number of tasks/gridlets 100, 200, 300, 400, and 500 with 33 % of node failures in the grid environment. From Fig. 4 it is noticed that proposed scheduler performed better in terms of successful schedule percent by 1.2 % in contrast with existing scheduler.

Scenario 4: Computing successful schedule percent by increasing number of tasks/gridlets with 50 % of machine failures.

Fig. 4 Computing success percent by varying number of tasks at 33 % of node failures

33 Percent - Machine Failures
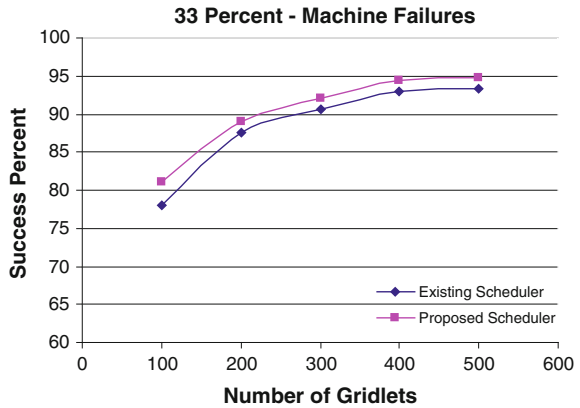
Fig. 5 Computing success percent by varying number of tasks at 50 % of node failures
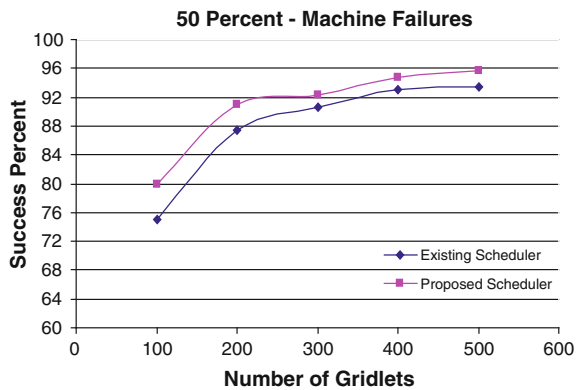
50 Percent - Machine Failures

we evaluated the proposed and the existing fault tolerant grid scheduler with respect to the success schedule percent of grid system on an increasing number of tasks/gridlets 100, 200, 300, 400, and 500 with 50 % of node failures in the grid environment. From Fig. 5, it is noticed that proposed scheduler performed better in terms of successful schedule percent by 2.8 % in contrast with existing scheduler.

5 Conclusions

In this study, fault tolerant grid scheduler is presented using pro-active and post-active approaches. The proposed fault tolerant grid scheduler initially uses pro-active approach, before scheduling of tasks to resources by computing scheduler index. Further it also uses a post-active fault tolerant approach for handling the constraint "failure of mode while the task is being executed" by migration method. This rescheduling component makes the task to migrate to another resource in event

of failure of node while the task is being executed due to node overload. This study can be extended to hierarchical scheduling scheme with local and global policies in the grid environment. Further this study may also be extended by integrating security concerns into the grid.

References

1. Maozhen Li, Brunel University, UK and Mark Baker University of Portsmouth, UK, "The Grid Core Technologies", 2005.
2. Mohammed Amoon, "A fault-tolerant scheduling system for computational grids", Journal of Computer and Electrical Engineering", 2012.
3. Rajkumar Buyya, Manzur Murshed, "GridSim: A toolkit for the modeling and simulation of distributed resource management and scheduling for Grid computing" Concurrency Computation Practical Experience, 2002, pp. 1175–1220.
4. Fran Berman, Geoffrey Fox, and Tony Hey, "The Grid: Past, Present, Future", 2003 John Wiley & Sons, Ltd.
5. J. Joseph, C. Fellenstein, "Grid Computing", Prentice Hall/IBM Press, Edition 2004.
6. Klaus Krauter, Rajkumar Buyya and Muthucumaru Maheswaran, "A taxonomy and survey of grid resource management systems for distributed computing", John Wiley & Sons Ltd., 17 Sept 2001.
7. Ram Mohan Rao Kovvur, S. Ramachandram, Vijayakumar Kadappa, A. Govardhan, "A Reliable Distributed Grid Scheduler for Mixed Tasks", PDCTA, CCIS 203, pp. 213–233, 2011.
8. J.H. Abawajy, "Fault-Tolerant Scheduling Policy for Grid Computing Systems", IPDPS'2004.
9. Paul Townend and Jei Xu, "Fault Tolerance within a Grid Environment" Proceedings of AHM2003.
10. Kovvur, R.M.R.; Ramachandram, S.; Kadappa, V.; Govardhan, A., "A distributed dynamic grid scheduler for mixed tasks," Advance Computing Conference (IACC), 2013 IEEE 3rd International pp. 110, 115, 22–23 Feb. 2013 doi:10.1109/IAdCC.2013.6514204.

Image Transmission in OSTBC MIMO-PLC Over Nakagami-*m* Distributed Background Noise

Ruchi Negi and Kanchan Sharma

Abstract In this paper MATLAB tool is used to simulate a MIMO-PLC system model using orthogonal space-time block codes (OSTBC) under the presence of additive background noise. The additive background noise is modeled using Nakagami–*m* distribution. The MIMO-PLC system model is tested with an input grayscale baboon image and BER curve is plotted for the MIMO-PLC system. Furthermore a closed form expression for the average BER of MIMO-PLC is derived theoretically by using moment generating function (MGF) based approach. The MATLAB simulated BER results are then compared with the theoretically derived results obtained from MGF based approach. Using MGF based approach a mathematical method is also presented to calculate the outage probability of PLC-MIMO for Nakagami-*m* distributed background noise.

Keywords OSTBC · PLC · Nakagami-*m* distribution · OFDM · Outage probability

1 Introduction

In power line communication (PLC) utility power lines are used for power transmission as well as data rate transmission, as existing power lines are used for data transmission so it eliminates the use of additional infrastructure required, which saves the time and cost of the data service provider and the user.

Multiple input multiple output (MIMO) antenna systems provide diversity, lower bit error rate (BER), high channel capacity and hence improve the data rate and reliability of a radio and wireless communication system, to exploit these

Ruchi Negi (✉)
Guru Gobind Singh Indraprastha University, New Delhi, India
e-mail: chitra_ait@yahoo.com

Kanchan Sharma
Indira Gandhi Delhi Technical University for Women, New Delhi, India
e-mail: joinkanchansharma@gmail.com

advantages of a MIMO in a PLC system, a lot of work is going and the researchers are shifting their focus from SISO-PLC to MIMO-PLC [1–3]. A 2×2 MIMO-OFDM system not only improves the BER relative to the existing SISO-OFDM system but is also insensitive to crosstalk [4]. In this paper orthogonal space-time block codes (OSTBC) are used for generation of MIMO streams, as OSTBC provides diversity at both transmitter and receiver side with very little add on to the transmitter and receiver circuitry A PLC channel is mainly corrupted by two types of additive noise namely additive background noise and additive impulsive noise. Background noise can be modeled by Nakagami-m distribution [5] while the impulsive noise follows Middleton distribution [6].

This paper is organized in following parts: in Sect. 2 a literature view is presented for all the work done in the past in the field of MIMO-PLC, Sect. 3 is further divided into four subsections in the first subsection OSTBC MIMO-PLC system model is given and in the second subsection the noise model for additive background noise is given, in third subsection closed form expression is derived for calculation of average BER of MIMO-PLC under the effect of background noise, and in the fourth subsection the expression for the outage probability of MIMO-PLC is derived. In Sect. 4 simulation results are discussed.

2 Literature Review

OSTBC MIMO in wireless communication is deeply discussed in [7–12]. In [13] Hashmat et al. studied MIMO communication for inhome PLC.

Choe et al. [4] studied the space-time coded and space-time frequency coded MIMO in PLC. In [3] Middleton while studying the electromagnetic interference gave an expression for additive impulsive noise which exists in PLC, while in [5] Choe et al. proposed a closed form expression in which background noise of PLC was modeled by Nakagami PDF. Further he extended his work in [14, 15] to study the error rate performance of binary transmitted and QPSK transmitted signal in PLC system under Nakagami-m distributed background noise. In [16] Mathur et al. studied the maximum likelihood decoding of QPSK signal in PLC system over Nakagami-m Additive noise and used copula approach to calculate the BER and SER.

After SISO-PLC the researcher shifted their attention to MIMO-PLC systems. In [1] Schneider et al. implemented the first real-time feasibility study of MIMO-PLC. Hashmat et al. In [2] presented two different models of background noise for Inhome MIMO-PLC namely the Omega model and Esmailian Model. In [3] Nikfar and Vinckie studied MIMO-PLC system under additive impulsive noise and presented the results for equal gain combining (EGC) and maximal ratio combining (MRC) receiver diversity scheme. In this paper PSNR of transmitted image is also calculated under the effect of additive white Gaussian noise (AWGN) and additive background noise. In [17] Jain et al. used PSNR as figure of merit for image transmission under Gaussian noise and proposed an algorithm for filtering of Gaussian noise based on the statistics of robust estimation and further in [18]

proposed a novel algorithm for filtering of Gaussian outliers based on local features of image.

In this paper a new OSTBC MIMO-PLC model is simulated in MATLAB tool and a grayscale 256×256 baboon image is transmitted to the OSTBC MIMO-PLC system using Quadrature Amplitude Modulation (16-QAM). A BER curve is plotted for the above image transmission under the effect of AWGN and additive background noise. Furthermore with the help of moment generating function (MGF) approach a theoretical expression is derived for the average BER of OSTBC MIMO-PLC under the effect of additive background noise and a comparison is made between the BER curve of the simulated OSTBC MIMO-PLC model and theoretically derived BER. Moreover in this paper an expression for outage probability is also derived using MGF approach for the OSTBC MIMO-PLC system.

3 System Model

3.1 OSTBC MIMO-PLC Model

The idea of MIMO-PLC system came from the fact that in most countries the domestic electrical wiring has three points: (1) Phase (2) Neutral (3) Phase Earth. The SISO-PLC uses phase (P) and neutral (N) ports for transmitting and receiving signals, while the MIMO-PLC exploits the phase earth (PE) point along with P and N point to create a 2×2, 2×3 MIMO or a 1×2 Single Input Multiple Output (SIMO).

OSTBC MIMO-PLC system block diagram is shown in Fig. 1. The system consists of $N_t = 2$ transmit ports and $N_r = 2$ receive ports. FEC block contains convolution encoder and interleaver. MIMO is implemented after Quadrature Amplitude Modulation (16-QAM). OSTBC is used to generate two separate MIMO
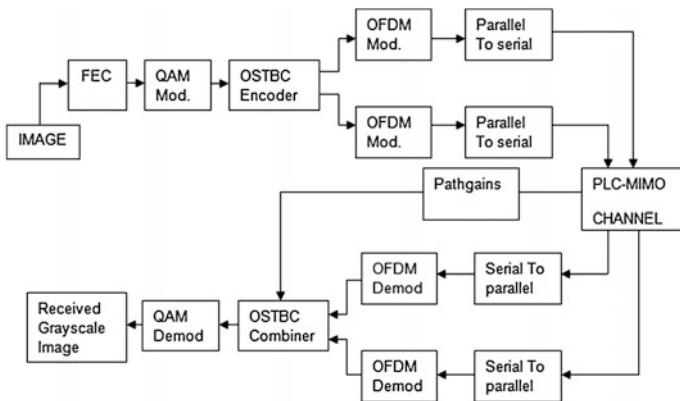
Fig. 1 Block diagram of 2×2 MIMO-PLC system

streams. Each of these two MIMO streams passes through separate OFDM modulator. In the OFDM modulator pilot carriers are added after which $N = 64$ point IFFT is used followed by a cyclic prefix addition of 25 %. After OFDM modulation the parallel MIMO stream is converted into serial stream which passes through the PLC channel.

At the receiver side after conversion from serial stream to parallel each MIMO stream passes through separate OFDM demodulator. In OFDM demodulator cyclic prefix is extracted first and then 64 point FFT is done followed by pilot carrier extraction. After OFDM demodulation the two MIMO streams are combined together in OSTBC combiner, OSTBC combiner requires the prior knowledge of pathgains so a perfect channel state information (CSI) is assumed. In Fig. 1 pathgains are shown to be connected from channel and going to the OSTBC, this connection shows that the pathgains of the channel are fed to the OSTBC combiner, as we have assumed a perfect CSI. After passing through OSTBC combiner the two separate MIMO streams are converted into single stream which passes through QAM demodulator and after that a reconstructed image is obtained.

In a MIMO-PLC system with N_t transmit ports and N_r receive ports, the ith receive signal is given by Eq. (1)

$$y_i(t) = \sum_{j=1}^{N_t} h_{ij}(t)x_j(t) + n_i(t) \tag{1}$$

where $h_{ij}(t)$ is the channel coefficient between ith received signal and jth transmit signal, $x_j(t)$ is jth transmitted signal and $n_i(t)$ is channel noise. $n_i(t)$ can be modeled as a summation of additive background noise and additive impulsive noise and AWGN. In matrix form the received signals can be written as given by Eq. (2).

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{N_r} \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{1N_t} \\ \vdots & & \vdots \\ h_{N_r 1} & \cdots & h_{N_r N_t} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdots \\ x_{N_t} \end{bmatrix} + \begin{bmatrix} n_1 \\ \cdots \\ n_{N_r} \end{bmatrix} \tag{2}$$

3.2 Background Noise Model

In this paper only effect of additive background noise and AWGN is taken into consideration and the effect of additive impulsive noise is ignored. In this paper a Nakagami-m distribution is assumed for the additive background noise. The probability density function (PDF) of Nakagami-m distribution is given by Eq. (3)

$$f_x(x) = 2\left(\frac{m}{\Omega}\right)^m \frac{x^{2m-1}}{\Gamma(m)} \exp\left(-m\frac{x^2}{\Omega}\right), \ x \geq 0; \tag{3}$$

where Γ is the gamma function, x is a random variable bounded by $x \geq 0$, Ω is the average power gain where $\Omega = E[x^2]$ and the fading parameter $m = \Omega^2 \Big/ E\left[\left(|x^2| - \Omega\right)^2\right]$.

3.3 Theoretical Average BER

For theoretical average BER calculation an MGF based approach is used assuming a Nakagami-*m* distributed background noise channel. Equation (3) is used to find the PDF of background noise (Nakagami-*m* distribution) under OSTBC MIMO. The PDF obtained is further used to calculate the MGF of background noise for OSTBC MIMO system. This MGF is used to calculate the theoretical expression of average BER for PLC-MIMO system.

Let $\overline{\gamma}$ is the average SNR, N_t is number of transmit ports, R_c is the code rate of OSTBC. Then for OSTBC MIMO the effective SNR at the output of receiver will be given by $\gamma = \overline{\gamma}/N_t * R_c$ PDF of background noise with OSTBC MIMO is given by Eq. (4)

$$f_{\text{OSTBC}}(\gamma) = \left(\frac{(mN_t R_c)^{mN_t N_r}}{\overline{\gamma}^{mN_t N_r}}\right)^{mN_t N_r} \times \frac{\gamma^{mN_t N_r - 1}}{\Gamma(mN_t N_r)} \times \exp\left(-mN_t R_c \frac{\gamma}{\overline{\gamma}}\right) \qquad (4)$$

where γ is the instantaneous SNR and $\overline{\gamma}$ is the average SNR. N_t and N_r are the number of transmit and receive ports, respectively. The average BER of M-QAM modulation scheme for Nakagami-*m* distributed background noise is given by Eq. (5) [19].

$$\text{ber} = \frac{4(\sqrt{M} - 1)}{\sqrt{M} \times \log_2(M)} \sum_{i=1}^{\sqrt{M/2}} \frac{1}{\pi} \int_0^{\pi/2} M\gamma\left(\frac{-(2i-1)^2 \times E_b \times g}{N_o \sin^2\theta}\right) d\theta \qquad (5)$$

where M is the modulation order which is 16 (16-QAM) in our case and M_γ is the MGF of the Nakagami-*m* distribution, E_b/N_o is the instantaneous SNR, i.e., γ and g is given by Eq. (6)

$$g = \frac{3}{2} \times \frac{\log_2(M)}{M - 1} \qquad (6)$$

For OSTBC MIMO the above Eq. (5) will be given by [19]

$$\text{ber} = \frac{4(\sqrt{M} - 1)}{\sqrt{M} \times \log_2(M)} \sum_{i=1}^{\sqrt{M}/2} \frac{1}{\pi} \int_0^{\pi/2} M_{\gamma\,\text{OSTBC}}\left(\frac{-(2i-1)^2 \times E_b \times g}{N_o \sin^2\theta}\right) d\theta \qquad (7)$$

$M\gamma_{OSTBC}$ is the MGF of the OSTBC-Nakagami-m distribution and E_b/N_o is the instantaneous SNR γ, N_t is number of transmit antenna. MGF of OSTBC-Nakagami-m distribution can be calculated by using Eq. (8) [19]

$$M_{\gamma\,OSTBC}(-s) = \int\limits_0^\infty f_{OSTBC}(\gamma)e^{-s\gamma} \tag{8}$$

By putting Eq. (4) in (8) the expression for MGF of OSTBC-Nakagami-m distribution reduces to Eq. (9).

$$M_{\gamma\,OSTBC}(-s) = \left(1 + \frac{s_{\bar{\gamma}}}{mN_tR_c}\right)^{-mN_tN_r} \tag{9}$$

BER of OSTBC MIMO system with M-QAM modulation can be numerically calculated by substituting Eq. (9) in Eq. (7).

Now the total average BER for MIMO-PLC using N-point DFT then can be expressed by Eq. (10)

$$ber_{total} = 1 - (1 - ber)^N \tag{10}$$

3.4   Outage Probability

Outage probability measures the performance of the diversity system operating under fading channels and is a very important measure of the quality of service (QOS) of a communication system. Outage probability is denoted by Pout and is defined as the probability that the instantaneous error rate exceeds a specified value or the combined SNR falls below a certain specified threshold z_{th} [19]. In this paper the outage probability of the MIMO-PLC system is calculated using MGF based approach. We will use the MGF of OSTBC MIMO-Nakagami-m distribution to calculate the Pout of OSTBC MIMO-PLC. Equation (11) given by [19] is used to calculate outage probability Pout, Where error term is bounded by Eq. (12) [19]. Where A, K, N are the parameters with values $A = 30$, $K = 10$, $N = 10$, z_{th} is the threshold BER, $\alpha_n = 1$ and MGF $M_{\gamma OSTBC}$ is given by Eq. (13) [19],

$$Pout(X \leq z_{th}) = \frac{2^{-K}e^{A/2}}{z_{th}} \sum_{k=0}^{K} \binom{K}{k} \sum_{n=0}^{N+k} \frac{(-1)^n}{\alpha_n} Re\left\{ \frac{M_{\gamma tOSTBC}\left(-\frac{A+2\pi j_n}{2z_{th}}\right)}{\frac{A+2\pi j_n}{2z_{th}}} \right\} + E(A,K,N) \tag{11}$$

$$E(A,K,N) = \frac{e^{-A}}{1+e^{-A}} + \frac{2^{-K}e^{A/2}}{z_{th}} \left| \sum_{k=0}^{K} (-1)^{k+N+1} \binom{K}{k} \mathrm{Re}\left\{ \frac{M_{\gamma\,\mathrm{tOSTBC}}\left(-\frac{A+2\pi j(N+k+1)}{2z_{th}} \right)}{\frac{A+2\pi j(N+k+1)}{2z_{th}}} \right\} \right|$$

(12)

$$M_{\gamma\,\mathrm{tOSTBC}}(s) = \prod_{l=1}^{L} M_{\gamma\,\mathrm{lOSTBC}}(S);$$

(13)

For $L = 1$ Eq. (13) reduces to Eq. (14) [19]. By plugging Eq. (9) in Eq. (11) Pout can be calculated

$$M_{\gamma\,\mathrm{tOSTBC}}(s) = M_{\gamma\,\mathrm{lOSTBC}}(s) = M_{\gamma\,\mathrm{OSTBC}}(s)$$

(14)

4 Results

A 256 × 256 grayscale image is passed through the 2 × 2 MIMO-PLC system, and 16-QAM modulation is used with perfect channel state information. A Nakagami-*m* distributed channel model as described in Sect. 3 is used to model the MIMO-PLC transmission system. An uncoded transmission is considered, OSTBC encoder and decoder is used at transmitter and receiver side. Figure 2a shows the original baboon
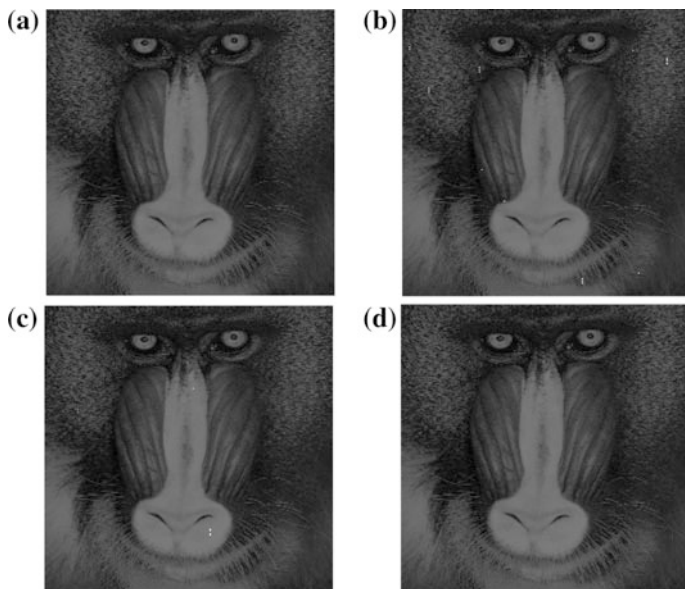
Fig. 2 **a** Original baboon grayscale image; **b** reconstructed image at $m = 1$, SNR = 16 dB; **c** reconstructed image at $m = 2$, SNR = 16 dB; **d** reconstructed image at $m = 3$, SNR = 16 dB
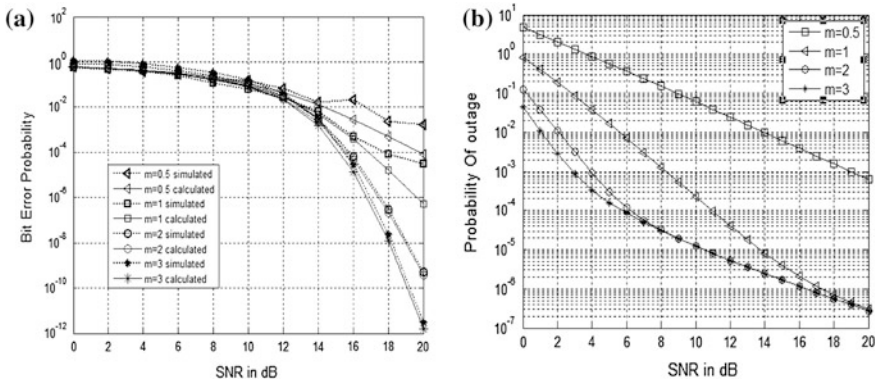
Fig. 3 **a** Theoretical and simulated BER curve for 2×2 MIMO at different values of m; **b** the outage probability of MIMO-OFDM system for different values of m

image and Fig. 2b–d shows the reconstructed baboon images at SNR of 16 dB for different values of m.

Figure 3a demonstrates the BER performance of a MIMO-PLC with uncoded transmission and 16-QAM Modulation with perfect CSI. The BER figure shows the effect of Nakagami-m distributed noise on the BER performance of MIMO-PLC. The influence of the Nakagami fading parameter m can be also be seen in the figure. When m is lower the severity of additive background noise is high and when m parameter is increased the severity of additive background noise decreases and hence increasing values of m there is an improvement in BER curve of MIMO-PLC system. For $m = 1$ BER of 1e-3 is obtained at 16 dB, and if m is increased to $m = 3$ BER of 1e-5 is obtained at 16 dB. Reconstructed error free image is obtained at 20 dB for $m = 1$, at 18 dB for $m = 2$ and at 17 dB for $m = 3$.

In Fig. 3a also the simulated BER is compared with the theoretically obtained BER. The theoretical BER is obtained from Eq. (10). It can be seen that the BER obtained from image transmission closely matches with the numerically obtained BER and with increasing m the BER shifts more and more towards left. However, a small difference can be observed between the simulated BER and theoretical BER which can be accounted by the presence of channel noise which is much more severe and channel conditions being much worse at the time of simulation as compared to the theoretical model of noise and channel condition. The complete BER results and PSNR for different values of Nakagami fading parameter m are tabulated in Table 1.

Figure 3b shows the result for the outage probability of 2×2 MIMO-PLC, the threshold BER of 1e-4 is taken which corresponds to z_{th} of 6.9 dB. It can be seen that outage probability also varies with the additive background noise parameter m. As m is increased the outage probability decreases.

Table 1 Simulated and theoretically evaluated BER and PSNR for baboon image transmission

SNR	Simulated BER and PSNR for image transmission								Theoretical BER			
	m = 0.5		*m* = 1		*m* = 2		*m* = 3		*m* = 0.5	*m* = 1	*m* = 2	*m* = 3
	BER	PSNR	BER	PSNR	BER	PSNR	BER	PSNR				
2	0.53968	13.245	0.52228	14.578	0.549796	15.557	0.54785	16.457	0.53658	0.528869	0.52486	0.5235
4	0.51768	14.630	0.491436	15.378	0.494358	18.699	0.4612	18.699	0.43539	0.422709	0.41605	0.41379
6	0.38872	15.357	0.28768	17.234	0.413456	22.540	0.40879	24.110	0.31741	0.29904	0.28942	0.28615
8	0.37136	15.879	0.22203	18.356	0.249506	26.678	0.24506	26.997	0.20412	0.18428	0.17465	0.17154
10	0.14617	18.2780	0.12141	20.336	0.110861	30.2789	0.11728	32.214	0.11016	0.092543	0.0846	0.0821
12	0.06595	22.678	0.02737	24.989	0.033339	45.689	0.03021	49.370	0.04475	0.030092	0.02334	0.02115
14	0.01737	26.768	0.00648	30.650	0.003134	55.678	0.00217	58.328	0.01286	0.004898	0.00219	0.00152
16	0.02235	30.265	0.00052	40.678	6.52E-05	60.349	2.03E-05	65.265	0.00278	0.000378	4.57E-05	1.42E-05
18	0.00233	35.743	8.54E-05	55.743	3.06E-07	62.567	1.76E-08	inf	0.00051	1.67E-05	2.14E-07	1.23E-08
20	0.00169	40.865	3.05E-05	60.895	5.12E-10	inf	2.23E-12	inf	8.54E-05	5.38E-07	3.58E-10	1.56E-12

5   Conclusion

In this paper a MIMO-PLC system employing OSTBC is discussed. It is shown that OSTBC MIMO-PLC system shows considerable performance improvement as compared to SISO-PLC system. A theoretically expression for the BER has been given which is in accordance with the results of the simulated OSTBC MIMO-PLC system model. It is shown that background noise considerably affects the performance of a MIMO-PLC system. This system model can be extended in the future to study the effect of impulsive noise, or to improve the performance of MIMO-PLC system.

References

1. Daniel Schneider, Andreas Schwager, Joachim Speidel and Altfried Dilly.: Implementation and Results of a MIMO-PLC Feasibility Study, 2011 IEEE International Symposium on Power Line Commun. And its Applications (ISPLC), pp. 54–59, April 2011.
2. Reshan Hashmat, Pascal Pagani, Thierry Chonavel, Ahmed Zeddam.: Analysis and Modeling of Background Noise for Inhome MIMO-PLC Channels, 2012 IEEE International Symposium on Power Line Commun. and its Applications (ISPLC), pp. 316–322, March 2012.
3. Babak Nikfar, A.J. Han Vinck.: Combining Technique Performance Analysis in Spatially Correlated MIMO-PLC systems, 2013 IEEE International Symposium on Power Line Commun. and its Applications (ISPLC), pp. 1–6, March 2013.
4. S. Choe, J. Yoo.: Space-time/Space-time-frequency-Coding Based MIMO-OFDM over Powerline Channels, IEEE Electronic letters, vol. 48, no. 16, (2012).
5. Youngson Kim, S. Choe, Hui-Myoung oh.: Closed-form Expression of Nakagami-Like Background Noise in Power-line Channels, IEEE transactions on power delivery, vol. 23, no. 3, pp. 1410–1412, (2008).
6. D. Middleton.: Statistical-Physical Models of Electromagnetic Interference, IEEE transactions on Electromagnetic Compatibility, vol. EMC-23, pp. 106–127, Aug. 1977.
7. S.M. Alouni.: A Simple Transmitter Diversity Scheme for Wireless Communications, IEEE Journal on Selected Area Communication, vol. 16, pp. 1451–1458, Oct. 1998.
8. V. Tarokh, H. Jafarkhani, and A. R. Calderbank.: Space-Time Block Codes from Orthogonal Designs, IEEE Trans. on Inform. Theory, vol. 45, pp. 1456–1467, July 1999.
9. V. Tarokh, H. Jafarkhani.: Space–Time Block Coding for Wireless Communications, IEEE Journal on Select. Areas Commun., vol. 17, pp. 451–460, Mar. 1999.
10. X. Li, T. Luo, G. Yue, C. Yin.: A Squaring Method to Simplify The Decoding of Orthogonal Space–Time Block Codes, IEEE Trans. Commun., vol. 49, pp. 1700–1703, Oct. 2001.
11. H. Shin, J. H. Lee.: Exact Symbol Error Probability of Orthogonal Space–Time Block Codes, in Proc. IEEE GLOBECOM'02, Taipei, Taiwan, R.O.C., 2002, pp. 1197–1201.
12. O. Tirkkonen and A. Hottinen.: Square-matrix Embeddable Space–Time Block Codes for Complex Signal Constellations, IEEE Trans. Inform. Theory, vol. 48, pp. 384–395, Feb. 2002.
13. R. Hashmat, P. Pagani, T. Chonavel.: MIMO Communication for Inhome PLC Networks: Measurement and Results up to 100 MHZ, in Proc. IEEE Intl. Symp. Power Line Commun., pp. 120–124, April 2010.
14. Y. Kim, H.M. Oh, S. Choi.: BER Performance of Binary Transmitted Signal for Power line Communication under Nakagami-like Background Noise, in Proc. Energy 2011: The First International Conference on Smart Grids, Green Communications and IT Energy Aware Technologies, pp. 126–129, May 2011.

15. Y. Kim, H.M. Oh, S. Choi.: Error Rate Performance of QPSK Transmitted Signal for Power Line Communication under Nakagami-like Background Noise, in Proc. Energy 2012: The Second International Conference on Smart Grids, Green Communications and IT Energy Aware Technologies, pp. 29–33, March 2012.
16. Ashish Mathur, Manav R. Bhatnagar, Bijaya K. Panigrahi.: Maximum Likelihood Decoding of QPSK Signal in Power Line Communication Over Nakagami-*m* Additive Noise, 2015 IEEE International Symposium on Power Line Commun. and its Applications (ISPLC), pp. 7–12, March 2015.
17. A. Jain et al.: A versatile Denoising Method for Images Contaminated with Gaussian Noise, Proc. Of (ACM ICPS) CUBE International Information Technology Conference & Exhibition, pp. 65–68, 2012.
18. A. Jain et al.: A Novel Detection and Removal Scheme for Denoising Images Corrupted with Gaussian Outliers, Proc. Of IEEE Students Conference on Engineering and Systems (SCES-2012), pp. 434–438, March 2012.
19. M.K. Simon, M.S. Alouni, Digital Communication over Fading Channels.: A Unified Approach to Performance Analysis, New York, NY: Wiley(2000).

Optimized Relay Node Based Energy-Efficient MAC Protocol for a Wireless Sensor Network

Kriti Ohri and C. Rama Krishna

Abstract Few of the most important quality of service (QOS) requirements for wireless sensor networks (WSNs) are bounded delay and energy efficiency. This paper presents a new Improved Energy-Efficient Distributed Receiver Based Cooperative Medium Access Control Protocol (Improved E^{2}DRCMAC) for WSNs which employs cooperative communication for delay sensitive and energy-efficient applications. We first recognized the limitations of cooperative communication which includes additional processing and reception steps at each node which increases energy and time delay with which the packets reach the destination. To address this issue, multi-hop cooperative communication scheme is proposed, which makes use of genetic algorithm (GA) to optimize relay nodes for delay minimization and energy savings. The Improved E^{2}DRCMAC outperforms E^{2}DRCMAC in terms of 70 % decrease in energy consumption, 45 % less end-to-end delay at higher data rates, and close to 100 % packet delivery ratio as node density starts increasing.

Keywords Cooperative communication · Routing · Message delay · Energy efficiency · Genetic algorithm · Wireless sensor networks (WSNs)

1 Introduction

WSN is a state-of-the-art technology within a wide spectrum of wireless technology which is specifically designed to sense sensor inputs. Sensor nodes are autonomous as each one of them is furnished with a microprocessor, battery, sensors, and a radio trans-receiver. WSNs do not have the luxury of abundant resources which otherwise

Kriti Ohri (✉) · C. Rama Krishna
Computer Science and Engineering, National Institute of Technical Teachers
Training and Research, Chandigarh, India
e-mail: Kriti.ohri@gmail.com

C. Rama Krishna
e-mail: rkc_97@yahoo.com

is a norm in hardwired networks. As a result of which, designing optimal procedures for minimizing the burden on sensor nodes becomes even more challenging. In high data rate applications like flood monitoring systems, fire detection systems, etc., in-time delivery of data and intelligent use of energy is very essential for the successful completion of sensing task [1]. Direct communication to the base station (BS) results in minimum delay but this is achieved at the cost of faster energy consumption. Hence, WSN resorts to cooperative communication in which each sensor node performs a dual role of transmitting its own data as well as assisting transmissions for other nodes in the network [2, 3]. However, cooperative communication may lead to additional hop count and additional receiving and processing steps which may further increase message delay and energy consumption. Hence, in order to further increase the benefits bought by cooperative scheme we add optimization process to path selection.

2   Problem Formulation

We consider the task of determining optimized paths $(P_1, P_2 \ldots, P_N)$ from each of the N cluster heads (CHs) to the BS in case direct communication fails. The cooperative nodes are selected in a way to minimize message delay and enhance energy efficiency. The metric of optimization is cost of paths between CHs and BS which is determined on the basis of two evaluation parameters, namely hops and distance (sum of distances between the nodes on the selected path). Keeping in view the importance of cooperative communication in WSNs and the issues with it, the research question can be formulated as: "Can optimized relay node selection in densely populated WSN reduce end-to-end delay, energy consumption and improve packet delivery ratio in a cooperative environment?"

3   Network Model and Assumptions

We consider a two tired multi-hop network which is divided into number of independent blocks called clusters; each cluster contains number of sensor nodes called cluster members (CMs) and a CH. The position of sensor nodes S_i ($i = 1, 2, \ldots S$) is (x_i, y_i) which is determined randomly and are deployed uniformly in an area of $M \times M \ m^2$. A BS is positioned more or less in the center of the network. Random placement of BS needs to be avoided in delay sensitive applications as it may lead to uncontrolled and irregular results [4]. Following assumptions have been considered for building up the scheme: (1) The sensor nodes remain static once deployed. (2) Sensor nodes are similar in terms of initial battery energy. (3) All sensor nodes are aware of their location details through global positioning system. (4) The sensor nodes dissipate energy as per the free space model (d^2 power loss).

4 The Proposed Optimized Cooperative Scheme: Improved E^{2}DRCMAC

This section proposes a new cooperative scheme called "Improved E^{2}DRCMAC". It assumes the standard IEEE 802.15.4 physical model which is specially designed for MAC and physical layer of wireless personal area networks suited for low-cost, low-power, and low-complexity devices [5]. The scheme supports both direct and cooperative communication. In the first scenario only direct communication is permissible between CHs and BS. If direct communication fails, cooperative communication is adopted. The energy model used for the scheme is represented in [6, 7]. We consider energy consumption in receiving and transmitting mode of a sensor node.

4.1 Clustering Phase

The proposed scheme uses location of nodes, neighboring nodes list, and distance of sensor nodes from the BS as the metric for selection of CHs. Initially when the sensor nodes S are deployed at time t_0 all of them are normal sensor nodes and become potential contenders for CH election. As and when the nodes wake up from their sleep mode, they determine their neighboring nodes which fall within their transmission range. Each potential CH as per the metric determines a value (based upon its distance to the BS) and exchanges it among its neighboring nodes. The potential CHs with minimum values then become final CHs among their neighborhood. Each CH then broadcasts a signal packet indicating that it is a CH. Except for CHs, all other nodes then decide to which CH they should belong depending upon the nearest CH in its range.

4.2 Time Slot Assignment

All transmissions from the nodes start with the beginning of each frame. The time is divided into frames, where each frame consists of mini time slots which contain the preamble and data sending slots along with update data structure and ACK as shown in Fig. 1. The first N available time slots are given to the nodes that arrive first from sleep state with data to send in their buffers. Each frame lasts up to a simulated 20 ms. In each mini time slot hopping of a packet between nodes is scheduled. It is assumed that each node attains any of the three states: sleep, active and back-off.
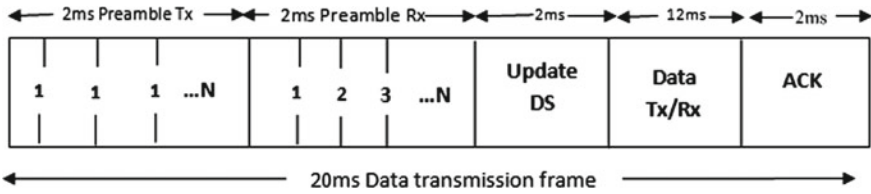
Fig. 1 Frame structure

4.3 Genetic Algorithm for Selection of Optimized Relay Nodes

GA is an evolutionary optimization search procedure that is based on natural selection and natural genetics [8]. We will now discuss in a little detail, how GA works in generating fittest paths to the BS.

1. Create initial random population of chromosomes of size *M*.
2. Calculate the fitness value of each chromosome.
3. Generate mutated child chromosomes corresponding to the parent chromosomes of the current generation.
4. Calculate fitness value of mutated child chromosomes.
5. Perform selection process by comparing parent chromosomes with corresponding child chromosomes as per the fitness function, the resultant population will then serve as parent population for the next generation.
6. If termination criteria is not achieved repeat Steps 3–5.
7. Else designate the fittest chromosome in the population.

- *Generation of initial random population of size (M):* GA starts with an initial population of solutions (generation 0) which are randomly generated. Each chromosome represents a possible path between a CH and BS. The genes in the chromosomes depicts the location of sensor node (x_i, y_i). The chromosomes in the population can be of different sizes, as each possible path may involve different number of cooperative nodes.
- *Fitness function evaluation:* Each path is evaluated as per the fitness function to determine how well it performs the task at hand. The accuracy of fitness values is measured upto four digits after the decimal point. The paths are evaluated based upon hop count and distance. To design such a fitness function we make use of weighted sum model. Through weighted sum an integrated analysis can be done which has the ability to weight and combine multiple inputs. The designed fitness function $F_j(H, D)$ is shown in Eq. (1)

$$F_j(H, D) = \frac{H}{H_{\max}} \times W_1 + \frac{D}{D_{\max}} \times W_2 \tag{1}$$

where, $j = (1, 2, 3, \dots M)$ denotes a particular path for population M, $H = n_j - 1$ denotes hop count on a particular path j (here, n_j is the number of nodes in a particular path j), W_1 and W_2 are weights assigned to hops and distance, H_{max} and D_{max} are the maximum hops and distance that can be encountered by a packet to reach the BS, and D denotes the total distance between all nodes on a particular path j as shown in Eq. (2).

$$D = \sum_{i=1}^{n_j-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \tag{2}$$

where, (x_i, y_i) is the positions of the ith node and (x_{i+1}, y_{i+1}) is the location of next node on jth selected path. Once the values of hop count and distance is known, fitness function is evaluated for the given paths in the population. Hence, the optimization problem can be stated as in Eq. (3)

$$\text{Min}[F_j(H, D)] \tag{3}$$

- *Generation of mutants:* Mutation operator changes or alters the current population by altering the gene values of the chromosomes and produces a new set of child chromosomes. Mutation induces variety into the population and avoids premature convergence on a local maximum.
- *Selection:* It is an operation for selecting parents for the next generation. We carry selection process by comparing parent chromosomes with corresponding mutated child chromosomes as per the fitness function. If the fitness value of jth child chromosome is less than the fitness value of jth parent chromosomes then:

1. Replace the jth parent chromosomes with the jth child chromosome.
2. Also, replace the fitness value of jth parent chromosome with the fitness value of jth child chromosome.
3. If the jth parent chromosome has a lesser fitness value than the jth child chromosome, then the parent chromosome is left untouched and is carried to the next generation.

Once the resultant population is created it is again subjected to mutation and selection operations until termination condition is met. Selection and mutation when taken together depicts hill climbing mechanism, which leads to continuous improvement.

- *Termination condition and result designation:* Termination criteria for GA is when maximum generation G has reached. It was seen that after the maximum generations were reached, there was no significant improvement in fitness function and hence GA was terminated. After the termination criterion is met, i.e., generations = G, best-so-far individual in the final population of
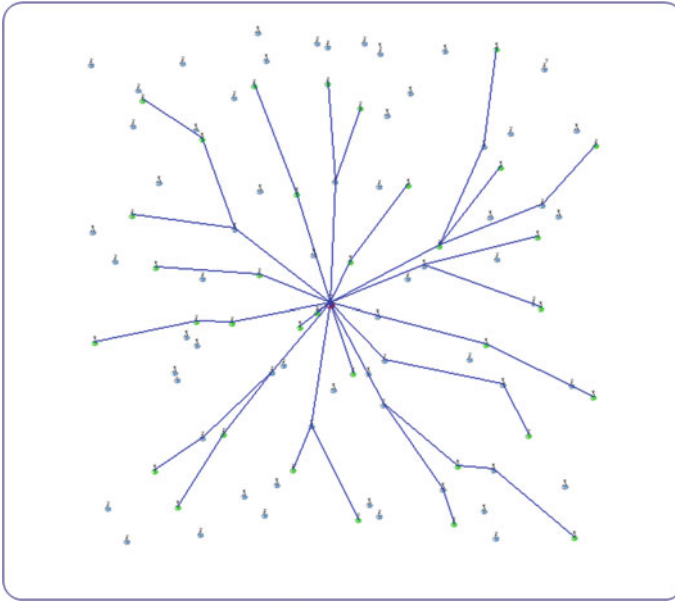
Fig. 2 Optimized routing paths

M chromosomes is selected and data is sent to the destination along the path. On the run of the algorithm, 35 CHs got elected and hence 35 optimized paths to the BS are generated as shown in Fig. 2. Blue nodes represent CMs, green nodes are CHs and red node in the center is the BS.

5   Performance Analysis

Both Improved E^2DRCMAC and E^2DRCMAC are based on the IEEE 802.15.4 network standard. The scheme represented in [9] proposes E^2DRCMAC that reduces energy consumption in the network by using cooperative communication along with overhearing avoidance. Overhearing avoidance is the process by which the redundant nodes (nodes which are not involved in communication) falling in the neighborhood between source and destination are put to sleep mode. This protocol uses channel state information (CSI) to determine the cooperative nodes for forwarding the data to the destination, in case direct communication fails. Whereas, Improved E^2DRCMAC adds optimization along with cooperative communication on 802.15.4 standard to further decrease energy consumption based on the selected features for optimization (hop count and distance).

5.1 Simulation Results

MATLAB 2012a is used as a simulation tool for the implementation of the work. The performance of Improved E^2DRCMAC is compared to E^2DRCMAC [9]. The simulation parameters are shown in Table 1. The GA parameters are as follows: population size M(10), number of generations G(20), method of selection (proportional to $Min[F_j(H, D)]$), weights assigned to evaluation factors i.e. hops and distance is 0.5 respectively.

The result in Fig. 3 shows the compared percent energy used in E^2DRCMAC and Improved E^2DRCMAC with respect to packets per second. In E^2DRCMAC energy consumption increases gradually with increase in offered load, whereas in Improved E^2DRCMAC the energy consumption is even more slower and sluggish. The main reason for this observation can be attributed to the reduced distances between CHs and BS which are provided by optimized paths. However, as the traffic load increases the energy consumption increases for both schemes but it is

Table 1 Simulation parameters

Parameter	Value
Number of sensor nodes S	100
Data	128 bytes
ACK	12 bytes
Average packets per message	6
Packet size	16 bytes
Message length	128 bytes
Area	1000 m $\times$ 1000 m
Initial energy of a node	5 J

Fig. 3 Percent energy used versus packets per second
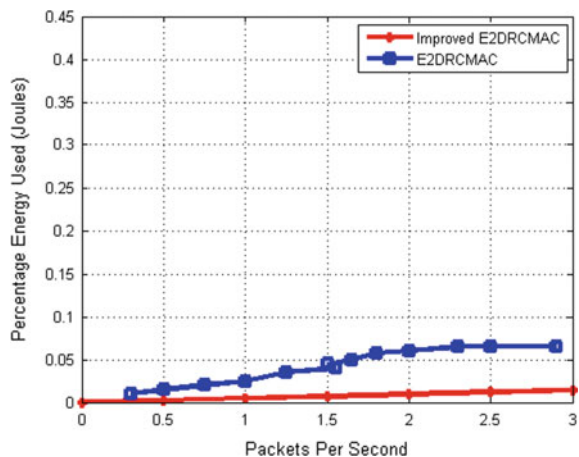
considerably small in Improved E^{2}DRCMAC. This is because as the traffic load increases the energy consumption increases but reduced distances or optimized paths between the CHs and BS in proposed scheme accounts for the overall reduction in network energy. Whereas in case of E^{2}DRCMAC both increased distances along with increase in offered load increases energy consumption of the nodes.

The result in Fig. 4 shows the compared message delay for E^{2}DRCMAC and Improved E^{2}DRCMAC with respect to packets per second.

It is clearly seen that the end-to-end delay generally increases as packets per second increases in both schemes. The increase in delay is expected as increased offered load induces more transmission delay at each intermediate hop. We observe an abrupt change in end-to-end delay in E^{2}DRCMAC from 2.5 to 2.7 packets per second, whereas in proposed scheme, the end-to-end delay linearly rises. The optimum end-to-end delay in E^{2}DRCMAC is between 0 and 2.5 packets per second, whereas proposed scheme fails to provide optimal results under same interval. However, it shows optimal results at higher data rates beyond 2.5 packets per second due to reduction in unnecessary processing and reception steps which decreases the number of hops between sender and the BS. As our approach is based on optimization technique, it gives better performance in WSN applications which produces data at high rates.

The result in Fig. 5 shows the compared packet delivery ratio *PDR* for Improved E^{2}DRCMAC and E^{2}DRCMAC with respect to node density. Node density determines the number of nodes deployed per unit area.

In the proposed scheme we observe that, as node density grows slightly *PDR* increases rapidly. It shows an improvement over E^{2}DRCMAC at lower values of node density and shows close to 100 % *PDR* beyond 0.04 nodes/m^2. The main reason for this improvement is that as node density increases, the neighborhood of a

Fig. 4 Message delay versus packets per second
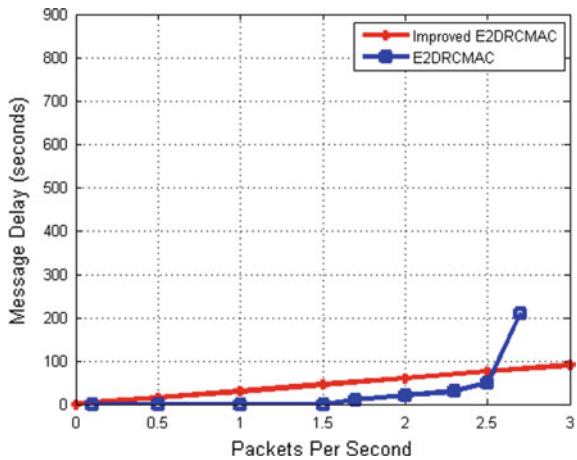
Fig. 5 Packet delivery ratio versus node density
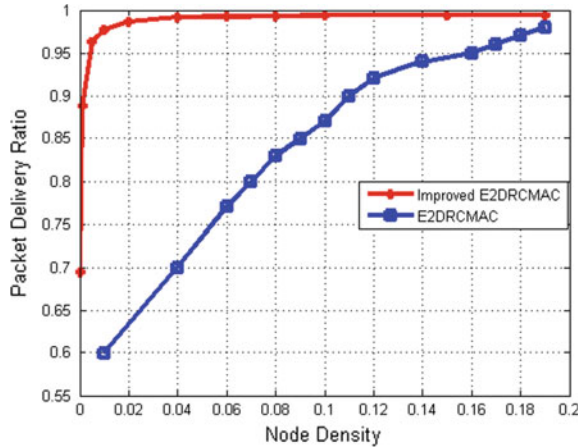
particular source also increases, and it is able to find better paths towards the BS through cooperative nodes. Whereas in case of E^2DRCMAC with increase in node density *PDR* gradually increases and attains an optimal degree of *PDR* at around 0.18 nodes/m^2 which is much later than in case of Improved E^2DRCMAC. The observation shows that the proposed scheme performs consistently well and is not much affected by node density beyond 0.1 nodes/m^2.

6 Conclusions and Future Scope

In summary, we confirmed that optimization along with cooperative communication may fulfill the timeliness and energy efficiency of WSN applications. One of the most significant factors determining in-time transport of data packets and energy consumption in WSN depends upon the path which is chosen to route the data packets to the BS. This paper proposed the Improved E^2DRCMAC whose major objective is to find optimized routes to the BS in a clustered environment based upon the relevant features selected. Reduction in unnecessary reception power and optimized routes have decreased energy consumption by 70 %, gives close to 100 % packet delivery ratio as node density starts increasing and reduces end-to-end delay by 45 % at higher data rates as compared to E^2DRCMAC. This work is an attempt to show that the benefits bought by cooperative communication can be further increased by determining optimized relay nodes in the network. Future work may be carried out by using other optimization techniques like particle swarm optimization and ant colony optimization. Moreover, additional features to the fitness function may be added for better integrated analysis.

References

1. Khan, M.F., Felemban, E.A., Qaisar, S., Ali, S.: Performance Analysis on Packet Delivery Ratio and End-to-End Delay of Different Network Topologies in Wireless Sensor Networks (WSNs). In: 9th IEEE International Conference on Mobile Ad-hoc and Sensor Networks, pp. 324–329. IEEE press, China (2013).
2. Maalej, M. and Besbes, H. and Cherif, S.: A cooperative communication protocol for saving energy consumption in WSNs. In: 3rd International Conference on Communications and Networking (ComNet). pp. 1–5. Tunisia (2012).
3. Nosratinia, A., Hunter, T.E., Hedayat, A.: Cooperative communication in wireless networks. IEEE Communications Magazine. 42, 74–80 (2004).
4. Poe, Wint Yi and Schmitt, Jens B.: Minimizing the maximum delay in wireless sensor networks by intelligent sink placement, Distributed Computer Systems Lab University of Kaiserslautern. 67655, 1–20. (2007).
5. Gutierrez, J.A., Naeve, M., Callaway, E., Bourgeois, M., Mitter, V., Heile, B.: IEEE 802.15. 4: a developing standard for low-power low-cost wireless personal area networks. IEEE Network, 15, 12–19 (2001).
6. Ghazvini, M.H.F., Vahabi, M., Rasid, M.F.A., Raja Abdullah, R.S.A.: Energy Efficiency in MAC 802.15.4 for Wireless Sensor Networks. In: 6th National Conference on Telecommunication Technologies, pp. 289–294. (2008).
7. Beiranvand, Z., Patooghy, A., Fazeli, M.: I-LEACH: An efficient routing algorithm to improve performance & to reduce energy consumption in Wireless Sensor Networks. 5th Conference on Information and Knowledge Technology, pp. 13–18. (2013).
8. Forrest, s.: Genetic algorithms: principles of natural selection applied to computation. Science. 261, 872–878 (1993).
9. Gama, S., Muddenahalli, T., Walingo, T., Takawira, F.: Energy Efficient Distributed Receiver Based Cooperative MAC for wireless sensor networks, IEEE AFRICON, pp. 1–6. (2013).

Reliable and Prioritized Data Transmission Protocol for Wireless Sensor Networks

Sambhaji Sarode, Jagdish Bakal and L.G. Malik

Abstract Wireless sensor network (WSNs) has potential to solve the problem automatically with less or no human intervention in most of the cases. Time constraint event (TCE) and non-time constraint (NTCE) event are two broad categories in heterogeneous WSNs. To address the QoS parameters in TCE and nTCE type of applications are crucial job in resource constraint multi-event WSNs. The queue scheduler performs those set of actions which are necessary to achieve the desired QoS parameters for each distinct event separately in a sensor network. A study mainly focuses on priority-based reliability with considering their different levels of priority of traffic. Decisions are taken at various intermediate nodes by managing queue. For each level of priority a separate queue is maintained. Queue management is the heart of proposed priority architecture for achieving quality of services. Due to proper queuing operation, scheduler takes the right decision at right time to attain the required reliability. Scheduler accomplishes desired goals of every application by handling their priority and deadlines.

Keywords Data transport · Priority · Wireless sensor network · Reporting rate · Scheduler · Reliability · Multi-event

Sambhaji Sarode (✉) · Jagdish Bakal · L.G. Malik
CSE Department, GHRCE, RTM Nagpur University, Nagpur, MS, India
e-mail: sambhajisarode@gmail.com

Jagdish Bakal
e-mail: bakaljw@gmail.com

L.G. Malik
e-mail: latesh.malik@raisoni.net

Sambhaji Sarode · Jagdish Bakal · L.G. Malik
MIT COE, Pune, India

Sambhaji Sarode · Jagdish Bakal · L.G. Malik
SPPU, Pune, India
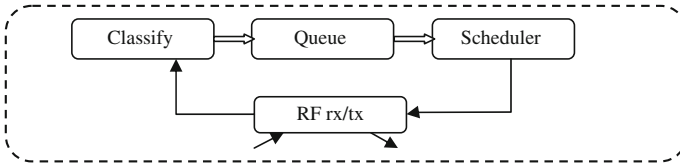
Fig. 1 Structural design of queue scheduler

1 Introduction

Nowadays, wireless sensor network [1–3] is growing rapidly in the Internet of Things (IoT) world. Contributions of sensor technologies are reflected in almost every field. Wireless technologies such as ZigBee (802.15.4), Wi-Fi (802.11), and Bluetooth are mainly used. Mainly ZigBee wireless technology is used as it consumes less energy compared to other technologies. There are many challenges in WSNs [4–7] because of having minimum energy, memory, and processing capability. In case of heterogeneous WSNs, there are more challenges as it deals with different types of traffic at a time. Handling various traffic with or without priority are challenging tasks. Making a generic protocol in WSNs is very difficult job because of having distinct QoS requirements of each one separately.

The structural design model shows the operation like classification of information, storing information in respective queue, rate-based packet scheduler, and delay-based packet scheduler as shown in Fig. 1. RF-based transceiver transfers or receives data wirelessly.

2 Related Work

The various transport layer protocols are studied such as PCS [1], CODA [3], ECODA [4], RETLP [5], ESRT [9], RT^2 [11], Performance analysis of QoS parameters [2], root cause analysis [7], etc. The study mainly focuses on examination of various factors which affects quality of services (QoS) of the sensor network.

The precedence control scheme (PCS) [1] basically addresses the reliability with considering prioritized information. PCS [1] mainly handles the priority of information using time to live (TTL) and priority identifier. The priority is assign by source node or hop node at the time of packet generation. While it is being processed at intermediate node this priori knowledge is used to take decision for transmission of higher prioritized packet to next node.

CODA [3] addresses the congestion through open loop, hop-by-hop mechanism, and closed loop mechanism. Based on the level of buffer it decides the reporting rate at each interval in order to take decision at right. ECODA [4] is an enhanced version of CODA protocol wherein more focus is given on priority-based information and

queue management. Weighted queue is designed to address the different types of data to achieve the target reliability, fairness, throughput, and delay.

RMST [8] protocol achieves the reliability using cache and noncache mode. It mostly works on reliability aspect with addressing delay issue at some extent. Fundamentally, cache helps to recover the lost packet immediately with minimum delay. It is very efficient from delay constraint network perspective.

ESRT [9] follows the centralized approach to achieve the reliability using collective approach. It continuously monitors the network state to decide the new reporting rate after every predefined decision interval to achieve the target reliability. Base station decides new frequency and broadcast to particular a region. It gives even chances to all participating node for transferring their sensed information. But it does not take care of transient traffic and congestion control effectively. RCRT [10] is another centralized approach where the base station does the job of congestion detection, rate adaptation, and rate allocation.

3 Design Considerations

Wireless sensor networks (WSNs) encompass two different types of networks: continuous and event-based. Continuous network generates traffic at fixed interval whereas event-based network generates traffic as and when event occurs and are random in nature. To predict data which is being generated at run time is difficult job. And the policy should not be generic to all different types of data. Depending upon the need of traffic, the rules should be applied with considering limitations of WSNs. Typically, information is categorized according to application domain. If the network generates data traffic of various domains then there is a need to handle it cautiously in order to justify it properly. On the basis of severity of information, the chances for early process need to be taken into consideration. For such information, the priority rules have to be set properly. So that their reliability will be assured otherwise all traffic will be treated at same level rather than their practical requirements. Though achieving 100 % reliability is the need of every application in the networking world but still not necessary to have always for every application in each case. So, in WSNs, variable reliability is a new research direction that has been explored by many authors to save the resources and increase the networks life. It initially observes each event and their tolerance level of packet loss. This helps to set the targets of each one and find the optimal solution.

While handling the data in queue, certain tasks such as preemptive and non-preemptive scheme with or without priority are to be addressed properly. The priority could be static or dynamic. Earliest deadline first algorithm with priority scheme helps to achieve the target QoS services. Network may face problems like: Problem 1: In static priority scheme, packet priority does not change in the network while it is being transmitted to its destination. In case of static priority with preemptive approach, higher priority packet preempts the lower the priority packet and acquires the resources unless it completes. Later on preempted packet resumes if

there is no other higher priority packet. Unfortunately sometime, if these low priority packets are not processed in time then unnecessarily their delay increases which in turn affect QoS parameters like throughput, performance, reliability, etc., of lower priority-based traffic. Problem 2: Keeping the static priority at large network with many hops in between source node and base station produces uneven contribution of data even if it is having higher priority. For example when a particular hop receives higher priority data from five hops away from source node and at the same time if current hop senses the data which is having same priority then there will be problem of which packet will be processed first. It is really hard to handle such a critical issue. Hence, dynamic priority would be useful approach to address this issue. Dynamic priority gets updated at every hop to distinguish the type of traffic. Here, traffic is categorized into two parts: network traffic and local traffic. Network traffic is a traffic which is generated at source node and travels via many hops in between source and base station whereas in local traffic, hop generates traffic using its own sensors. Unfortunately, if network packet and locally generated packets are having same priority then our scheme gives the first chance to network traffic as it arrives from long distance and its deadline is also considered. This proposed scheme does not allow the higher priority packets which are generated by source or hop nodes closer to base station first. Priority aware hops help to reduce delay of network traffic. Second benefit of dynamic priority-based approach with preemptive helps to acquire the required resources in time and achieves the desired goal. It is very useful for time critical applications.

4   System Model

In multi-event, the analysis has to be done on the basis of nature of application. Analysis of different types of traffic leads towards the classification of time constrained event (TCE) and non-time constrained event (NTCE). In case of time constrained event (TCE), time parameter is an important factor in some applications such as heartbeat counting, earthquake detection, fire detection, tsunami, missile launching, smoke detection, military surveillance, landslide detection, etc., whereas in non-time constrained event (NTCE), time parameter is not as important as in TCE. NTCE applications are weather monitoring, pressure, air pollution detection, water quality monitoring, motion detection, energy metering, transportation management, structural health monitoring, etc. In TCE, occurrence of event, execution of right action at right time depends upon the condition of network. Typical sensor network may consist of both TCE and NTCE traffic. Sometimes it could happen that same type of information may be sensed by many sensors and could be transferred to its upstream node. So this type of network can be categorized as cooperative. The Fig. 2 shows the flow of various activities.

Typically, source nodes sense the information with adding the appropriate tag for identification of data traffic to its upstream node. After classification of received data, the actor node takes the right decision for every event within time frame,

Fig. 2 Action flow chart of scheduler
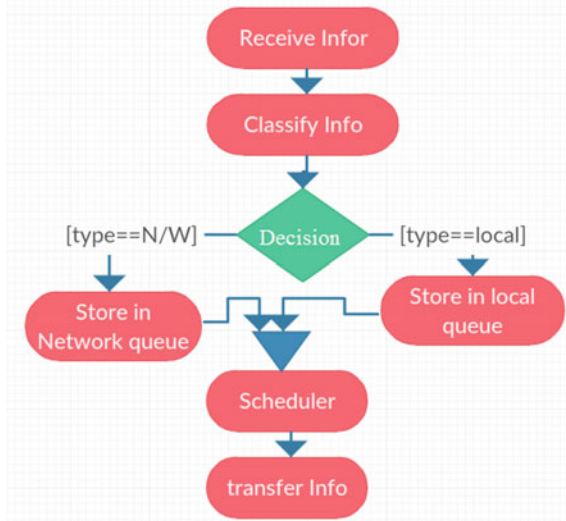
therefore, it updates new reporting rate based on desired reliability requirement, buffer condition, and network state as shown in Fig. 2. Decision could be made from received data which consists of network traffic and locally generated traffic at current node. While handling these types of traffic it is difficult to address long delay and fairness problem. Hence there is a need to design the flexible scheduler. The role of scheduler is to take the right decision at right time for transmission of data to next node based on its priority level, and TTL value with minimum delay. In this study, the more focus is given on hop-by-hop network instead of end-to-end. Processing capability has been incorporated into hop, i.e., actor node in order to reduce the load on base station and takes the decision when information is being transferred to base station. As a result, decisions are taken faster compared to traditional network.

4.1 Process Model

Typically, network consists of three major tasks in sensor network. Mainly this study focuses on processing functionality. Processing of information is a main concern for achieving the target goal set for particular context. Proposed model tries to achieve the desired reliability with considering priority issue and deadline of each traffic separately. Figure 3a, b shows classification of information, queue management, and dynamic priority. Scheduler picks the packets which has higher priority and lowest deadline. Implicit retransmission is used to achieve the desired target of each one.
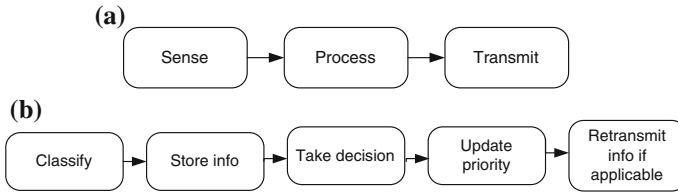
Fig. 3 a Typical state model. **b** Internal subactions of process activity

4.2 Dynamic Priority Scheduling Algorithm

Queue scheduler essentially works over the different queues. Classification of information is done prior to changes in priority. System model has some basic functions like:

1. classfiyInfo(event_type, priority_identifier, hop_count, deadline): scrutinize info
2. storeQueue(queue_type, info): stores info in respective buffer
3. updatePriority(event_type, deadline, hop_count, priority_identifier): update priority of packet considering deadline and hop count of each
4. calculateDelay(event_type): delay is calculated for each event separately
5. scheduler(): Selects priority packet from respective queue for further transmission
6. desireReliability(): Calculate desire reliability at every interval.

4.3 Scheduling Algorithm

1. receive(event Info, priority ID)

   Switch (priority)
   Case#1: (priority= =high) //Store in high priority qeue
   Case#2: (priority= =medium) //Store in medium priority queue
   Case#3: (priority= =low) //Store in low priority queue
   Case#4: (priority= =NULL) //Store in general queue

2. updatePriority(packet info, deadline) //priority is updated considering priority ID, deadline, event type, delay bound
3. scheduler(packetType) //choose the highest priority packet with earliest deadline first for transmission
4. computeDelay(delay info) //delay is computed at each hop.

The table given below illustrates the simulation setup configuration for sensor network. This particular setup has been tested at various topologies such as mesh and linear-based for 11, 21, 31, and 41 nodes. As far as this study is concerned, results

Table 1 shows the experimental setup of sensor network		
Simulator	NS2
Sensor field	500×500 m^2
MAC type	`IEEE 802.11`
Source nodes	`11`
IF queue length	20
Packet size	36 bytes
Transmit power	0.660 w
Receive power	0.395 w
Routing layer	AODV
Storage	20
Energy	50 J

are shown for 10 nodes. The desired reliability is set 100 % to each node. Every node sends 10 packets in one second. It may change as per application requirement such as 80, 70, 75 %, and so on which will save resources but this aspect is not taken into consideration. This could be explored in the further study (Table 1).

5 Simulation Results and Performance Analysis

Priority model has been tested on different topologies with different set of nodes; it has been observed that the inference of topologies and nodes are connected to hop at various levels. Priority#3 indicates highest priority whereas priority#0 indicates no priority. Reliability is set 100 % to every event in the below experiments.

Priority#3>priority#2>priority#1>priority#0

Figure 4 shows that the packet delivery ratio is good for higher priority-based traffic. The successful delivery ratio varies in between 3 to 4 %, 6 to 7 %, 8 to 15 %, and 10 to 18 % over experiments 2, 4, 6, 8, 10, and 12 s, respectively. Analysis clearly reveals that the scheduler gives more chances to higher priority events instead of lower priority events.

Figure 5 illustrates the delay comparisons at various level of priority of data. It has been observed that higher priority events transfer data in less delay compared to others. For example in first case wherein experiment time is two seconds, delay of high priority is 0.045 s, medium priority is 0.048 s, low priority is 0.049 s, and no priority is 0.052 s. In this way experiments are carried over in different times. Examination proves that the lower priority events face the problem of long delay compared to higher priority events because their waiting time in queue is more.

Figure 6 illustrates efficiency comparison over different experiments. Priority level 3, 2, 1, and no priority shows efficiency(%) in between 96 to 98, 93 to 97, 85 to 94 and 82 to 90 over experiment time 2, 4, 6, 8, 10, and 12 s, respectively. This performance analysis illustrates that higher the priority higher the efficiency because retransmission mechanism works properly to get lost immediately.

Fig. 4 Throughput
comparision
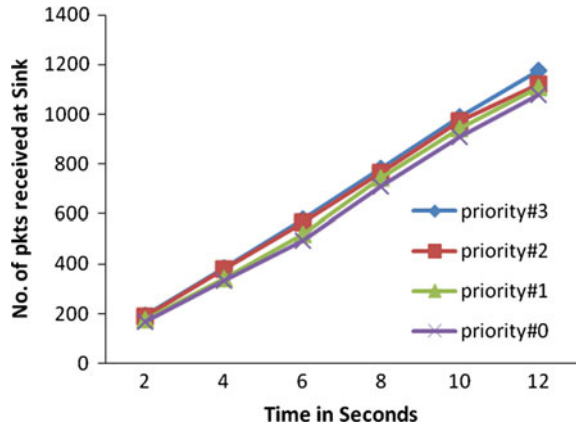
Fig. 5 Delay comparision
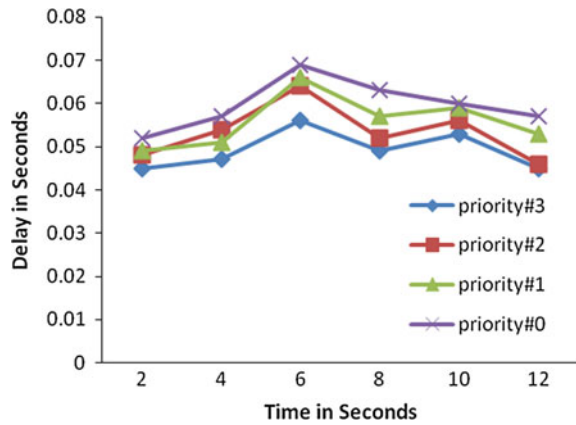
Fig. 6 Efficiency
comparision over different
time
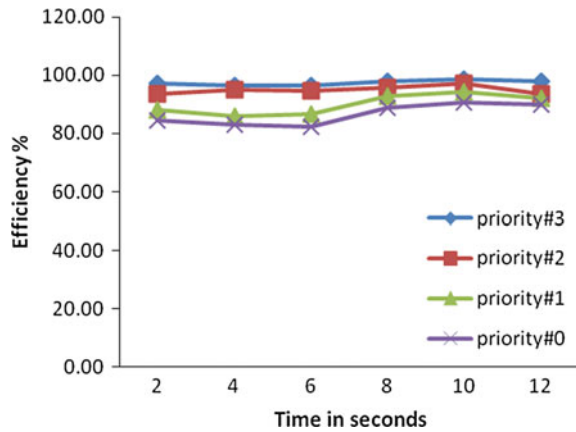
6   Conclusion and Future Enhancements

The proposed scheduler achieves priority-based reliability for heterogeneous traffic considering their priority and deadline. It has been observed that priority scheduler plays a vital role for transmission of right data at right time. Initially, information is classified then scheduler works over queue to identify accurate packet for processing. The above experimental results show that the priority-based reliability is achieved by managing queue and renewing priority time to time. The heart of the proposed scheme is queue management.

Furthermore, desired reliability can be achieved in TCE using packet splitting mechanism. Packet splitting scheme could be used to transfer data in maximum available paths in given time frame. It will not only help to achieve the target reliability within time frame but could be used to maximize the network lifetime. Fortunately, it may lead toward the even utilization of network, but dividing into number of pieces at source node and assembling them together at target location orderly is a major challenging task of priority-based data carrier scheme. This approach introduces the parallelism. As a result it may offer good throughput, desired reliability; prolong network lifetime, and less delay.

References

1. Sambhaji Sarode, Jagdish Bakal, Precedence Control Scheme for WSNs, Pervasive Computing (ICPC), 2015 International Conference on, IEEE ICPC 2015, Pages: 1–5.
2. Sambhaji Sarode, Jagdish Bakal, Performance Analysis of QoS parameters Constrained based WSNs, IEEE International Advance Computing Conference (IACC) 2015, (in press).
3. C.-Y. Wan, S. B. Eisenman, and A. T. Campbell, CODA: Congestion detection and avoidance in sensor networks, in Proc. ACM SenSys, Nov. 2003, Pages 266–279.
4. Li Qiang Tao and Feng Qi Yu, ECODA: enhanced congestion detection and avoidance for multiple class of traffic in sensor networks, IEEE Transaction, Consumer Electronics, Vol. 56, pp. 1387–1394, Aug. 2010, 0098–3063.
5. Atif Sharif, Vidyasagar Potdar, A.J.D. Rathnayaka, Priority Enabled Transport Layer Protocol for wireless sensor network, 2010 IEEE 24th International Conference on Advance Information Networking and Applications workshops.
6. Abhijeet Bagadi, Sambhaji Sarode, Jagdish Baka, A Survey of Reliable Transport Layer Protocols for Wireless Sensor Network, IJCA Proceedings on National Conference "MEDHA 2012" MEDHA(1):35–38, September 2012.
7. Vivek Deshpande, Prachi Sarode, Sambhaji Sarode, Root Cause Analysis of Congestion in Wireless Sensor Network, International Journal of Computer Applications 1(1):27–30, February 2010, Foundation of Computer Science.
8. F. Stann and J. Herdemann, "RMST: Reliable data transport in sensor networks," in Pro. 1st IEEE Workshop SNPA, Anchorage, AK, Nov. 2003, pp. 102–112.

 9. Yogesh S., O. B. Akan, Ian F. Akyildiz. ESRT: Event-to-Sink Reliable Transport in Wireless Sensor Networks. In Proc. ofMobiHoc, Annapolis, Maryland, USA, June 2003.
10. P. Jeongyeup, G. Ramesh "RCRT: Rate-Controlled Reliable Transport for Wireless Sensor Networks," in Proc, ACM SenSys, Nov. 2007.
11. V. Gungor, Ozgur B. Akan, Ian Akyildiz, "A Real Time and Reliable Transport (RT)2 Protocol for Wireless Sensor and Actor Networks IEEE/ACM Transaction on Networking, Vol. 16, No. 2, April 2008, Pages 369–370.

Caching: QoS Enabled Metadata Processing Scheme for Data Deduplication

Jyoti Malhotra, Jagdish Bakal and L.G. Malik

Abstract Increase in digital data demands a smart storage technique which provides quick storage and faster recovery of the stored data. This voluminous data needs an intelligent data science tools and methods to struggle for the space required for efficiently storing the data. Deduplication is one the emerging and widely used techniques these days for reducing the data size and then transferring it over the network thus reducing the network bandwidth. Main aim of Deduplication is to reduce storage space by allowing only unique data. In this paper, we present the review of existing Deduplication solutions and propose enhanced cache mechanism approach for efficient Deduplication over distributed system.

Keywords Backup window · Chunking · Data deduplication · Hashing · Storage

1 Introduction

In the smart and digitized era, data is increasing at a very high speed. It is an integral part of any individual and/or organization. Data comes from various sources; and then travels through the journey of its processing, integration, validation and storage. This voluminous data is limited by the lack of efficient storage

Jyoti Malhotra (✉) · Jagdish Bakal · L.G. Malik
CSE Department, GHRCE, RTM Nagpur University, Nagpur, MS, India
e-mail: jyoti.j.malhotra@gmail.com

Jagdish Bakal
e-mail: bakaljw@gmail.com

L.G. Malik
e-mail: latesh.malik@raisoni.net

Jyoti Malhotra · Jagdish Bakal · L.G. Malik
MIT COE, Pune, India

Jyoti Malhotra · Jagdish Bakal · L.G. Malik
SPPU, Pune, India

Fig. 1 Schematic
representation of data
deduplication

Deduplication
(Redundancy Removal)
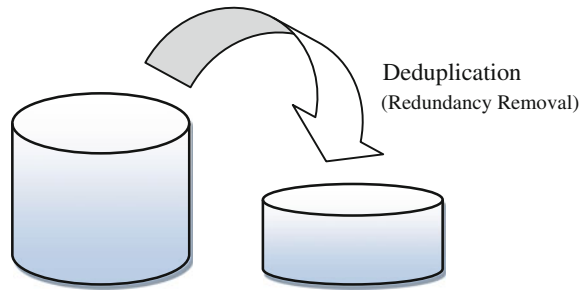
space and parameters. Data can arrive in the form of records, documents, text, image, audio, video, compressed, logs, structured or unstructured files. Storing and using this data effectively is a true challenge. This data is stored in large data pools. The majority of data is stored in relational databases. On contrary, relational databases are not enough to store this varietous and voluminous data. Wherever possible, data is shrinked or can be compressed; which can significantly improve data storage and its throughput. There are two kinds of compression- Local and Global; Local Compression is specific to certain data types, or files. Global Compression is specific to blocks, files and file systems. Techniques used to compress file systems are Archives[1] (zip/rar/tar/war), Data Deduplication, and Byte differencing.

Deduplication is a smart compression and backup technique which aims to remove duplicate data by storing unique copy of a file and generating a reference or pointer to the existing ones.

1.1 Deduplication Process

Data Deduplication (DD) helps to reduce the storage requirement for backup as shown in Fig. 1. Deduplication can be a file-grain or block-grain. In file-grain DD, similarity between two files is checked by their fingerprints or hash values; duplicate files are replaced by pointers to previously stored files. Block-grain DD is an improved step over file-grain; here rather than calculating a hash for entire file and comparing them to the hashes of existing files, block-grain chops a file into small components [1] or blocks and they are compared against existing blocks. This comparatively reduces the volume of data storage and improves storage utilization. DD can be performed either on source-side, target-side, and inline or post process [2].

DD Process can be briefly viewed in three steps as shown in Table 1.

---

[1]Archives are file containers containing one or more compressed files.

Table 1 Data deduplication process and actions

Sr. No.	Steps	Input	Action
1	Write request (backup) initiation	Data block	Returns data fingerprint or hash
2	Check data existence (check for duplicates)	Data fingerprint	Returns true or false
3	Data storage (file upload)	Data existence? True	Set pointer to existing block
		Data existence? False	Store data and update metadata

2 Related Work

The key parameters affecting the performance of Deduplication solutions are: Backup window, DD efficiency, DD throughput and metadata management.

Storing duplicate data with traditional backup method significantly increases the backup window size. Authors in [3–5] talks about combining data blocks into super blocks resulting a shrinked backup window. Techniques mentioned in [6–9] improves DD throughput to some extent.

Tan et al. [10] proposes a two level deduplication with file semantics to achieve high DD efficiency and short backup window. Authors of [11] capture fundamental relationship among the dataset versions to improve backup and restore performances.

We also experimented redundancy elimination and observed that—Variable size blocks give better deduplication ratio as compared to fixed size block.

$$\text{Deduplication ratio} = \text{Bytes In/Bytes Out}$$

Experimentation of redundancy elimination [3] was tested with 8 kB chunk size; and we were able to achieve 84.5 % R-squared on Linear Regression model. We also observed that a check on Block similarity can be enhanced with advanced similarity indices such as kulczynski measure or Jaccard Index. Metadata overhead problem can be resolved using Metadata Indexing [12] using Cuckoo filter which saves more time by reducing lookups by 25 %.

3 Challenges

An important aspect of DD performance depends on:

- Type of data
- Change rate of data, i.e., amount of data alteration/modification between two deduplication operations
- Amount of duplicate data

- System Load; affecting backup window, and
- Backup type.

  Various challenges for optimized DD process are:

- Removing duplicates with maximum DD ratio and low CPU usage
- Metadata being a backbone of the entire process; handling it is a great challenge.
- Metadata search can create bottlenecks; which can affect DD performance and increase the backup window.
- Metadata corruption can also create a problem for data recovery.

4 Proposed Mechanism

Understanding the need of scalability, throughput and above mentioned challenges, there is a need of migrating Deduplication process from an individual appliance to a clustered environment. As capacity of data is increasing, Hadoop [13] plays a vital role in storing this big data. Deduplication can be applied to structured as well as unstructured data. Hadoop stores this big data on various nodes; but it does not check for the duplicate files. Hadoop supports replication, meaning if duplicate file is sent for the backup, Hadoop will store it again along with its replicated copy for that backup instance. This increases the storage utilization eating more space.

With respect to this, a matured DD process is proposed, where DD is combined with Hadoop, Hbase and an intelligent caching mechanism as shown in Fig. 2. When users submit the files for DD; client processes the files, generates fingerprints, and maintain logs of each user. File separation unit reads the header section of each file and separates the file according to its type (text, pdf, word, image, videos, compressed) and upload files to respective clusters. Here metadata holds the log of user and files which are already stored. As per the file type, it is uploaded to respective data nodes of the cluster.

When a file is with the cluster it generates its hash value and submits it to DD unit. DD unit performs a check for an existing fingerprint in its metadata; if fingerprint is not present it will store new copy of file and if exist already then do not store the copy. Here metadata is stored in Hbase. As metadata size grows, performing a check increases database hits. In order to reduce the bottleneck of database hits, we have cached the fingerprints along with their reference count in a text file. This text file is accessed using JSON–Java Script Object Notation, widely used data-interchange format which is easy to parse and process.

After DD, data is being stored to storage units with the help of relative weighted factor/load balancing; this distributes the files according to the load of application specific nodes at that moment of time. Here cache management is achieved using <key-value> pair and cache filters such as cuckoo filter [12] and bloom filter to avoid bottlenecks in metadata lookups.
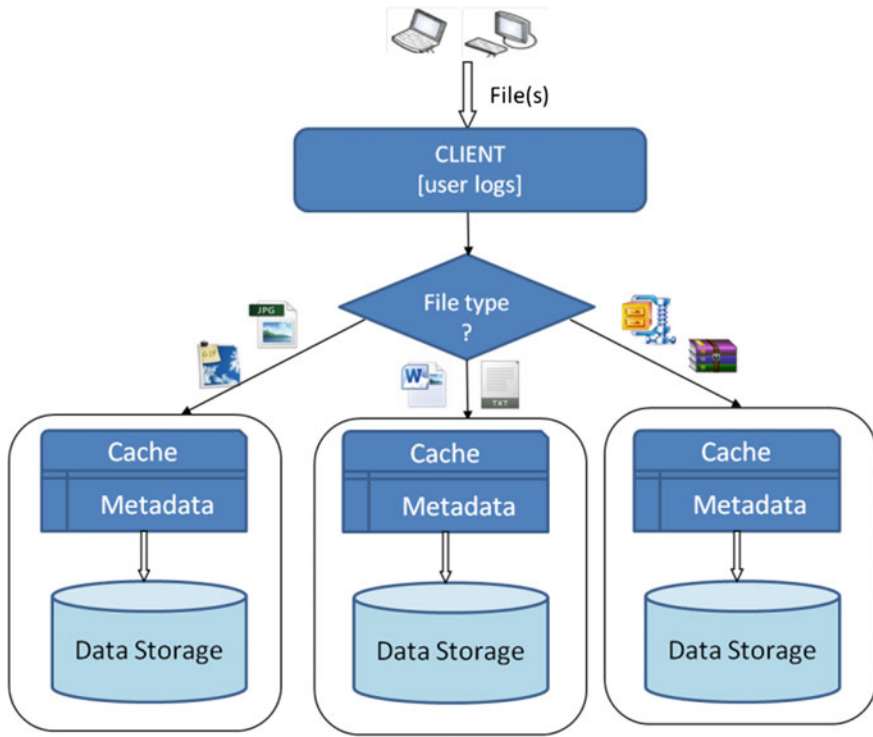
Fig. 2 Enhanced cache-based deduplication

4.1 Process Model

Let S be a set which contains Users, Input file, Hash algorithms, Hash value, Metadata, Cache, Deduplication, Upload to HDFS.

$$S = \{ U_r, I_p, H_a, M_d, C_a, D_d \text{ and } H_{upload} \}$$

where,

U_r a set of users = $\{u_1, u_2, u_3, \ldots, u_n\}$

I_p a set of input files of various types = {text, pdf, doc, zip, rar, image, audio, video, xml, etc.}

H_a a hash algorithm which generates a hash value using algorithms such as SHA, RSA and pHash [14] a perceptual hash for multimedia files

M_d a metadata which summarizes basic information about file_name, file_size, file_timestamp, file_owner, file_hashvalue, file_referencecount, file_links, and node ID where it is stored

D_d a Deduplication appliance which identifies duplicates along with Hbase records and a cache

H_upload an upload function which stores the unique copies of files on Hadoop and generating links for the duplicate files

Deduplication process as given in Algorithm#1 is applied along with above discussed process components

Input: I_p
Output: H_{upload}

1.	Maintain User_log
2.	Separate files types referring to the file header
3.	**foreach** *file from client* **do**
4.	**Transfer** files to respective *backup_node* as per their type
5.	**Run** D_d on the backup_node
6.	$fp \leftarrow$ ***Fingerprint*** *(f);*
7.	**Search** hash value *fp* of the file from metadata *M*.
8.	**If** *fp = M* **then**
9.	File is duplicate; No need to store it;
10.	Pass Reference Link and update Metadata
11.	**Else if** *fp ≠ M* **then**
12.	Provide the File to H_{upload}
13.	Update Metadata

Algorithm 1: Deduplication Process

5   Simulation Results and Performance Analysis

To evaluate the working of deduplication firstly, it was implemented using C and pthreads, on the Linux platform (Ubuntu). Here, the data in the form of text file was manually generated based on the factors such as: size and redundancy. The experiment was carried out on files of varying sizes and results were compared for the same test sets.

Figure 3 depicts a graph showing the Deduplication results; where all the files have the same content which is caused by creating multiple copies of it and saving them by different names. This graph shows that even though total file size is getting incremented with original backup; size after deduplication remains same.

This work was migrated from single Deduplication appliance to clustered Deduplication, which was further deployed on Hadoop. Table 2 depicts the experimental setup for Deduplication on Hadoop. Users' data is collected from the different Drives. The size of the data sets is varying from 3 to 20 MB.

Fig. 3 File-grain deduplication
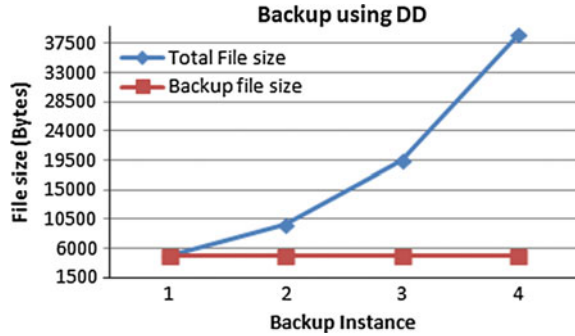
Table 2 Experimental setup

Information	Description
File size (MB)	3–50
File types	7 ('.txt', '.pdf'', doc, zip, img, audio, video)
Clusters	4
Deduplication engines	3
Platform	Ubuntu 12.04
Programming language	Java and JSON
Backend	Hbase

We perform analysis on various aspects of deduplication backup:

  (i)   CPU usage of the system.
 (ii) Deduplication ratio of the system.
(iii) Overall storage space reduction
(iv) Time to upload, modify, and delete file to and from the storage system

To know the effectiveness of our system, we applied functions to calculate time required to finish the upload or delete job. Throughput measures the time required for the deduplication process to perform. Deduplication ratio measures the ratio of input bytes to output bytes. Over all storage space reduction is also seen when storage space is reduced after applying deduplication on it.

Figure 4 shows the experimental results for uploading files on backup storage.

Here graph indicates that time taken for uploading a file using H_{upload} is less as compared to other methods [5, 8, 12]. Maximum time is required for primary data, where large amount of time is needed for breaking the file into small blocks and performing block-grain deduplication. In similarity index method, small blocks are grouped into superchunks and it performs superchunk-grain deduplication, hence it takes less time to upload a file on backup node. H_{upload} takes much less time due to file separation unit as work gets parallelize among all nodes and caching metadata into JSON enabled text file which reduces the backup window.

Fig. 4 File upload on
Hadoop
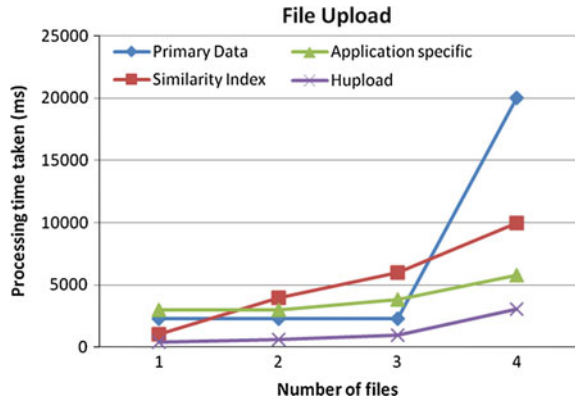
Fig. 5 File delete on Hadoop
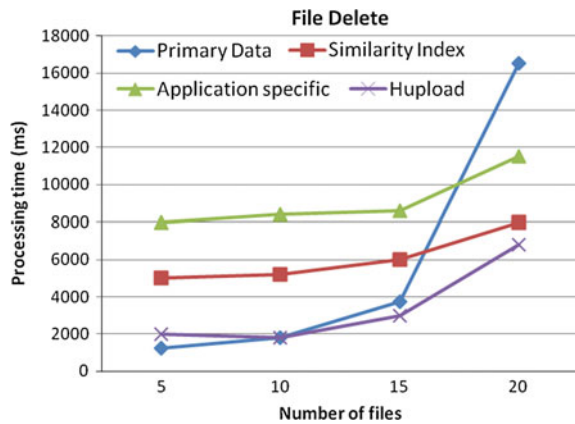
Figure 5 shows the experimental results for deleting files on backup storage. Graph indicates that H_{upload} method takes less time to check the metadata and delete the files, which in turn reduces the CPU usage time.

6   Conclusion and Future Enhancements

There are techniques whereby we can implement intelligent caching and load balancing mechanisms in order to improvise and well-architect the deduplication performance. Deduplication based on application type is good as the metadata processing is parallelized among the respective clusters. Parallelizing the work is important to avoid hash collisions in centralized metadata. To top it off, caching metadata plays an important role in Quality of Service factors. We are looking forward to enhance the work by improving cache management at system level and adding security measures at metadata level.

References

1. BingChun Chang, "A Running Time Improvement for Two Threshold Two Divisors Algortihm" 2009.
2. Nagapramod Mandagere, Pin Zhou, Mark A Smith, Sandeep Uttamchandani, "Demystifying data de-duplication", Proceedings of the ACM/IFIP/USENIX Middleware '08.
3. Jyoti Malhotra, Jagdish Bakal, "A survey and comparative study of data deduplication techniques", Pervasive Computing (ICPC), 2015 International Conference on, IEEE ICPC 2015, Pages 1–5.
4. W. Dong, F. Douglis, K. Li, H. Patterson, S. Reddy, and P. Shilane. "Tradeoffs in scalable data routing for deduplication clusters". In Proceedings of the 9th USENIX conference on File and Storage Technologies, FAST'11, pages 15–29, Berkeley, CA, USA, 2011. USENIX Association.
5. Yinjin Fu, Hong Jiang, and Nong Xiao. A Scalable Inline Cluster Deduplication Framework for Big Data Framework. Middleware '12 Proceedings of the 13th International Middleware Conference Pages 354–373.
6. David Frey, Anne-Marie Kermarrec, "Probabilistic Deduplication for Cluster-based storage system deduplication" in Proc. ACM Cloud Computing, October 2012.
7. Jin San Kong, Min Ja Kim, Wan Yeon Lee, Young Woong Ko. Two-Level Metadata Management for Data Deduplication System. IST 2013, ASTL Vol. 23, pp. 299–303.
8. Yinjin Fu, Hong Jiang, Nong Xiao, Lei Tian, Fang Liu, Lei Xu, "Application-Aware Source Deduplication for Cloud Backup Services of Personal Storage," IEEE Transactions on Parallel and Distributed Systems, 2013.
9. W. Dong, F. Douglis, K. Li, H. Patterson, S. Reddy, and P. Shilane. "Tradeoffs in scalable data routing for deduplication clusters". In Proceedings of the 9th USENIX conference on File and Storage Technologies, FAST'11, pages 15–29, Berkeley, CA, USA, 2011. USENIX Association.
10. Yujuan Tan, Hong Jiang, Dan Feng, Lei Tian, Zhichao Yan, Guohui Zhou "SAM: A Semantic-AwareMulti-Tiered Source De-duplication Framework for Cloud Backup" 2010 39th International Conference on Parallel Processing.
11. Yujuan Tan, Hong Jiang, Dan Feng, Lei Tian, Zhichao Yan "CABdedupe: A Causality-based Deduplication Performance Booster for Cloud Backup Services".
12. Archana Agrawal, Jyoti Malhotra, "Clustered Outband Deduplication on Primary Data" 2015 IEEE International Conference on Computing Communication Control and Automation.
13. https://hadoop.apache.org/.
14. http://www.phash.org/.

Design and Analysis of Grid Connected Wind/PV Hybrid System

**K. Shivarama Krishna, K. Sathish Kumar, J. Belwin Edward
and M. Balachandar**

Abstract In this article, the design and analysis of wind/PV hybrid system which is grid connected has been presented. The wind and photovoltaic sources are integrated at the DC bus and its voltage is stepped up to desired value using DC–DC boost converter. The proportional integral controller has been used to control the output power produced from wind and photovoltaic resources to achieve desired output response. A variable speed control method is implemented for permanent magnet synchronous wind generator, which is capable of extracting maximum power even if it is operated at wind speed which is lower than the rated speed. The power produced from wind/PV hybrid system is supplied to the load and in the case of excess power generation it is supplied to the grid. The proposed hybrid system is modeled in MATLAB/Simulink. The performance of the system is evaluated by not only considering the changes in wind speed and solar irradiance but also by load power variations. The simulation results show that the proposed hybrid wind/PV system is a viable option for microgrid applications.

Keywords Renewable energy sources · Permanent magnet synchronous generator · Photovoltaic · Proportional integral controller

K. Shivarama Krishna (✉) · K. Sathish Kumar · J. Belwin Edward · M. Balachandar
School of Electrical Engineering, VIT University,
Vellore 632014, Tamilnadu, India
e-mail: shivaram.k2014@vit.ac.in

K. Sathish Kumar
e-mail: kansathh21@yahoo.co.in

J. Belwin Edward
e-mail: jbelwinedward@vit.ac.in

1   Introduction

The power produced from the conventional sources like coal, oil, and gas shows more impact on environmental pollution which results in global warming. The renewable energy sources like wind, photovoltaic, hydro, biogas, and fuel cells can be used in order to meet the load demand. These renewable energy sources are dependent on environmental conditions such as solar irradiance and wind speed. The power produced from individual energy sources is not sufficient to meet the load demand. So these renewable and nonrenewable energy sources are integrated to form hybrid system that is more reliable compared to an individual source.

   This article presents design and control of wind/PV hybrid system. The article has been structured as follows. The modeling of photovoltaic system is discussed in Sect. 2. The design of wind turbine and permanent magnet synchronous generator are discussed in Sect. 3. The control strategies, power conditioning unit, and modeling of grid side converter is discussed in Sect. 4. The wind/PV hybrid system which is grid connected has been developed and its dynamic performance is explored in Sect. 5.

2   Modeling of Photovoltaic System

The photovoltaic (PV) model designed in MATLAB/Simulink environment is depicted in Fig. 1. The proposed model requires few parameters like shot circuit current, open circuit voltage, and the number of photovoltaic modules used as well as its temperature coefficient. Moreover, this model is appropriate for simulating the practical photovoltaic systems which consists of several PV modules and it represents the changes in temperature and solar irradiance that usually occur during the day [1, 2]. A 6 kW PV model is designed using the following Eqs. (1–7):
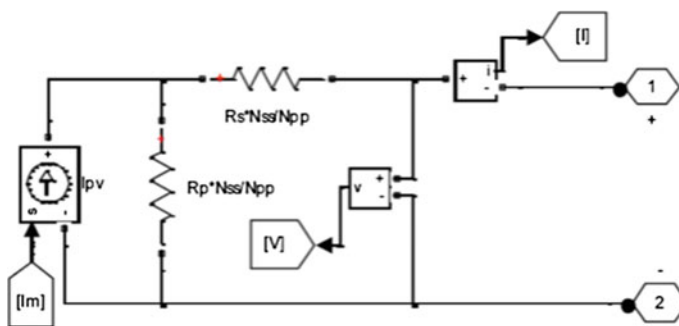
Fig. 1   Circuit-based PV model

$$I = I_{\text{PV.cell}} - \frac{I_{0.\text{cell}}\left[\exp\left(\frac{qV}{aKT}\right)-1\right]}{I_{\text{d}}} \tag{1}$$

$$I_{\text{d}} = I_{0.\text{cell}}\left\{\exp\left[\frac{qV}{A}*k*T\right] - 1\right\} \tag{2}$$

$$I = I_{\text{pv}} - I_0\left[\exp\left(\frac{V+R_{\text{s}}I}{aV_{\text{t}}}\right) - 1\right] - \frac{V+R_{\text{s}}I}{R_{\text{p}}} \tag{3}$$

$$I_{\text{PV}} = \left(I_{\text{pv,n}} + k_{\text{I}}\Delta_{\text{T}}\right)\frac{G}{G_{\text{n}}} \tag{4}$$

where $\Delta_{\text{T}} = T - T_{\text{n}}$, T is actual temperature, G denotes the device surface irradiation in W/m^2, G_{n} and T_{n} is the nominal irradiation and temperatures, respectively.

$$I_0 = I_{0,\text{n}}\left(\frac{T_{\text{n}}}{T}\right)^3 \exp\left[\frac{qE_{\text{g}}}{ak}\left(\frac{1}{T_{\text{n}}} - \frac{1}{T}\right)\right] \tag{5}$$

$$I_{\text{o,n}} = \frac{I_{\text{sc,n}}}{\exp\left(\frac{V_{\text{oc,n}}}{aV_{\text{t,n}}}\right) - 1} \tag{6}$$

$$I_0 = \frac{I_{\text{sc,n}} + K_{\text{I}}\Delta_{\text{T}}}{\exp\left(\frac{V_{\text{oc,n}} + K_{\text{V}}\Delta_{\text{T}}}{aV_{\text{t}}}\right) - 1} \tag{7}$$

The voltage and current coefficient K_{V} and K_{I} are included in Eq. (6) which results in Eq. (7).

Where I_{d}, $I_{\text{PV.cell}}$, $I_{0.\text{cell}}$, and $I_{\text{PV,n}}$ are the diode current, current produced from the incident light, diode reverse saturation current, and the current produced at the nominal condition, respectively.

I_0 and I_{pv} represent the saturation and photovoltaic current, respectively, and $V_{\text{t}} = N_{\text{s}}\,kT/q$ represents the thermal voltage of PV array.

3   Modeling of Wind Turbine

The mechanical output power P_{m} generated from the wind turbine is calculated using Eq. (8). The torque T and output power P_{m} generated from the wind turbine within the interval are [V_{min}, V_{max}], where V_{min} and V_{max} are functions of the wind speed, air pressure, turbine blade radius, and coefficients C_{P} and C_{q} [3–5]. The wind turbine is modeled in MATLAB/Simulink using following Eqs. (9–13) and it is depicted in Fig. 2.
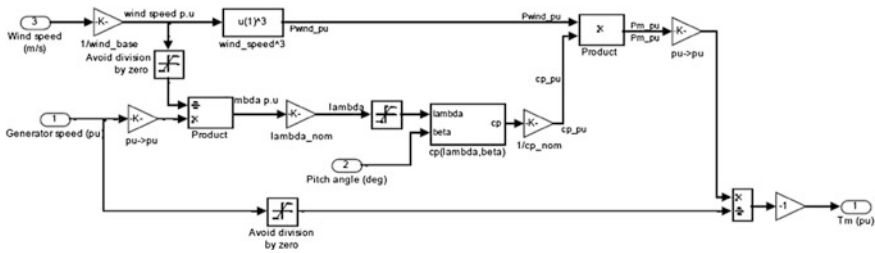
Fig. 2 The wind turbine is model in MATLAB/Simulink

$$P_\mathrm{m} = C_\mathrm{P}(\lambda, \beta) \frac{\rho A}{2} V_\mathrm{wind}^3 \tag{8}$$

$$C_\mathrm{q} = \frac{C_\mathrm{P}}{\lambda} \tag{9}$$

$$\lambda = \frac{R * \omega}{V_\mathrm{wind}} \tag{10}$$

$$T = \frac{P_\mathrm{m}}{\omega} \tag{11}$$

$$C_\mathrm{p}(\lambda, \beta) = C_1 \left(\frac{C_2}{\lambda_i} - C_3\beta + C_4 \right) e^{-\frac{C_5}{\lambda_i}} + C_6 \tag{12}$$

$$\frac{1}{\lambda_\mathrm{i}} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^3 + 1} \tag{13}$$

where

C_p	is the coefficient of performance
P_m	is the mechanical output power (W) and R denotes the radius of turbine blades (m)
A and V_wind	denotes the turbine swept area (m^2) and wind speed (m/s)
ρ	is the air density $\left(\frac{\mathrm{kg}}{\mathrm{m}^3} \right)$, λ and ω represents the tip speed ratio and angular frequency of rotational turbine, respectively
β and T	represent the blade pitch angle and torque of wind turbine, respectively

3.1 Modeling of PMSG

The synchronous generator model is expressed in direct and quadrature axis reference frame, which is given by the following Eqs. (14–15). The generator consists of permanent magnets and which has no damper winding. In order to simplify calculations, Park transformation is implemented for dynamic model of PMSG in which the synchronous reference frame is converted to direct and quadrature (d–q) rotating reference frame [6].

$$V_{ds} = -R_s i_{ds} - L_d \frac{di_{ds}}{dt} + \omega L_q i_{qs} \tag{14}$$

$$V_{qs} = -R_s i_{qs} - L_q \frac{di_{qs}}{dt} + \omega L_d i_{ds} + \omega \varphi_m \tag{15}$$

where

V_{ds} and V_{qs} denote the direct and quadrature axis voltages, respectively
i_{ds} and i_{qs} denote the direct and quadrature axis currents, respectively
R_s denotes the stator resistance, ω denotes the angular frequency of rotor
L_d and L_q denote the direct and quadrature axis inductance, respectively
φ_m denotes the flux linkage amplitude

Electrical Torque T_e is given by following Eq. (16):

$$T_e = \frac{3}{2}(P)\varphi_m i_{qs} \tag{16}$$

where P denotes the number of pole pairs of the PMSG.

3.2 Power Conditioning System

The power produced from the renewable energy sources (RES) like wind, PV, hydro, fuel cell, biomass, and microturbine is either DC/AC with different voltage and frequency levels. Power electronic interface is necessary in order to interconnect RES with grid. The input parameters like irradiation and temperature of the PV system changes according to the environmental conditions which results in unregulated DC output voltage. In case of wind energy system the output power generated is in the form of AC which is converted to DC by using uncontrolled rectifier. The DC–DC boost converter is implemented to modulate the output power produced from the wind and PV sources. The regulated output power is supplied to the load and to the grid through voltage source inverter [7–9]. The schematic diagram of DC–DC boost converter is depicted in the following Fig. 3:
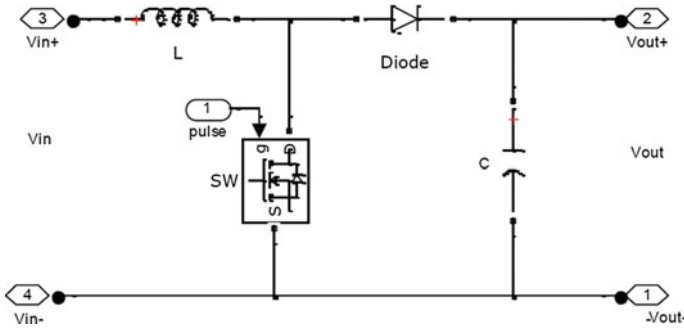
Fig. 3 Schematic diagram of DC–DC boost converter
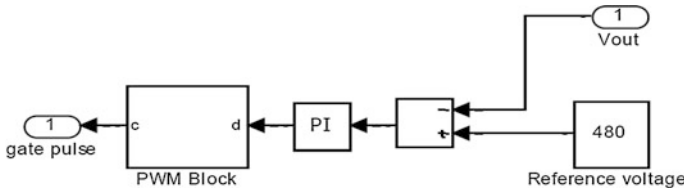
Fig. 4 Proportional integral controller for DC–DC boost converter

The proportional integral controller for DC–DC boost converter is depicted in the following Fig. 4:

3.3 Modeling of Grid Side Converter

The grid side converter mathematical modeling is presented and voltage balance across the line is obtained from Eq. (17), in which L and R denote the line reactance and resistance, respectively. The three-phase quantities are converted into two-phase quantities using d–q theory [10]. The current regulated control of grid side converter is shown in Fig. 5.
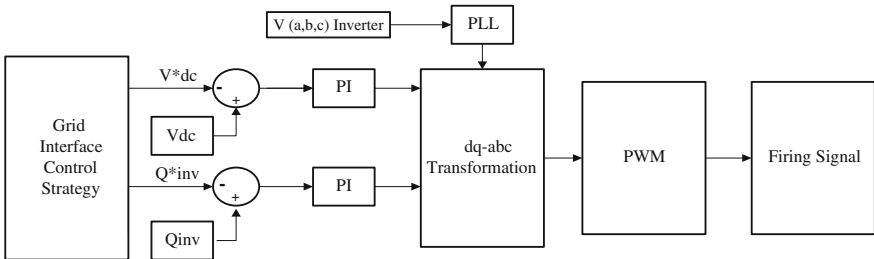
Fig. 5 Current regulated control of grid side converter

$$\begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} = R \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} + L \frac{d}{dt} \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} + \begin{bmatrix} V_{a'} \\ V_{b'} \\ V_{c'} \end{bmatrix} \tag{17}$$

The grid side converter mathematical modeling is given by the following Eqs. (18–19):

$$V_d = R i_d + L \frac{di_d}{dt} - \omega_e L i_q + V_{di} \tag{18}$$

$$V_q = R i_q + L \frac{di_q}{dt} - \omega_e L i_d + V_{qi} \tag{19}$$

where V_{di} and V_{qi} are the two-phase voltages found from $V_{a'}, V_{b'}, V_{c'}$ using d–q theory.

4 Results and Discussion

The proposed grid connected wind/PV hybrid system is modeled in MATLAB/Simulink environment is depicted in Fig. 6.

The proposed system is simulated for 1 s in which the change in generation is attained by varying the irradiance of the PV system from 900 to 600 W/m² at 0.3 s and the wind speed is changed from 6 to 8 m/s at 0.3 s. The load power variations are achieved by connecting a load 1 of 4 kW active power, 3.3 kVAR reactive power connected through breaker 1 at 0.2 s and another load 2 of 7.5 kW active power, 5.0 kVAR reactive power connected through breaker 2 at 0.5 s, the breaker is opened at 0.9 s. So from 0.2 to 0.5 s the load is 4 kW, 3.3 kVAR; from 0.5 to 0.9 s load is 11.5 kW, 8.3 kVAR and from to 0.9 to 1 s the load is 4 kW, 3.3 kVAR. The local AC load is connected to the 230 V, 50 Hz Grid. The DC link voltage, power of the hybrid system, active and reactive power distribution for the above conditions are shown in the Figs. 7, 8, 9, and 10, respectively.
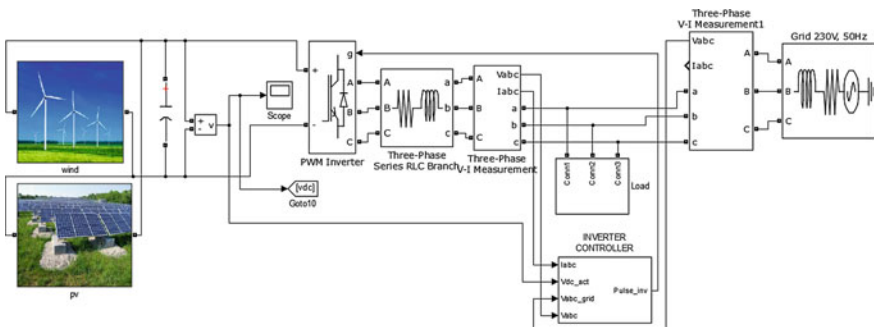
Fig. 6 MATLAB/simulink model of the proposed grid connected wind/PV hybrid system
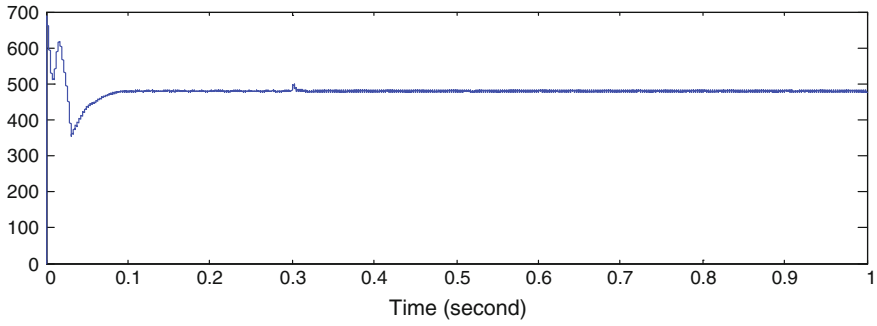
Fig. 7 DC link voltage
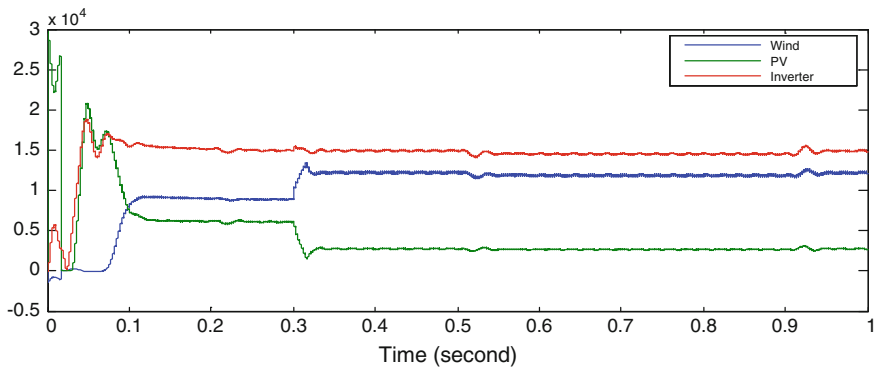
Fig. 8 Power of the hybrid system
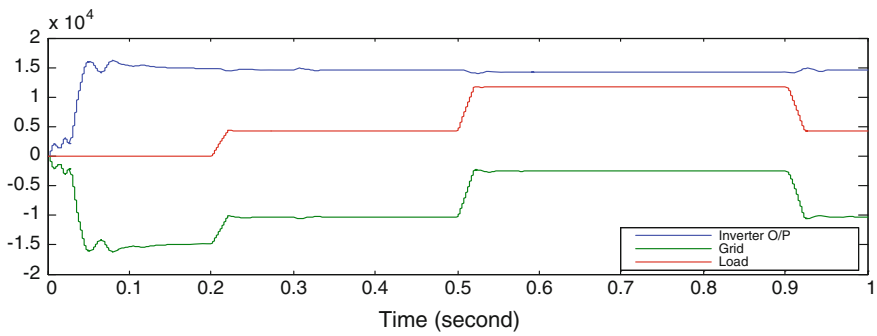
Fig. 9 Active power distribution
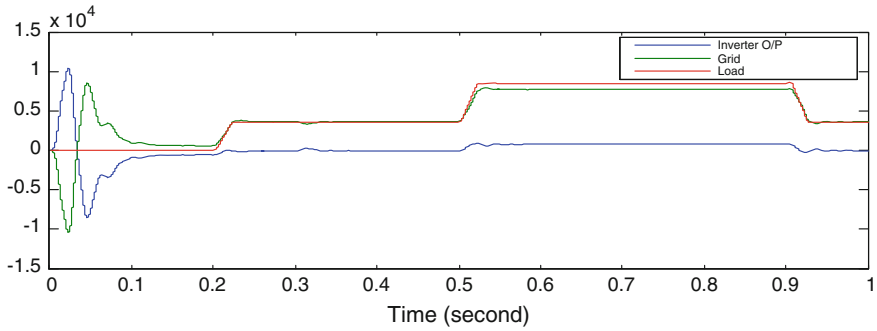
Fig. 10 Reactive power distribution
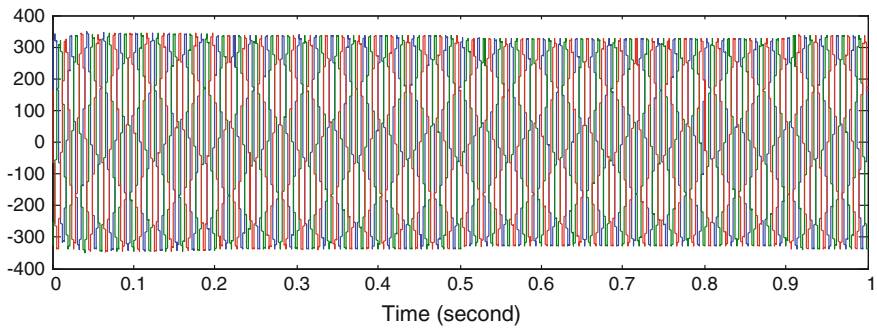
Fig. 11 Load voltage
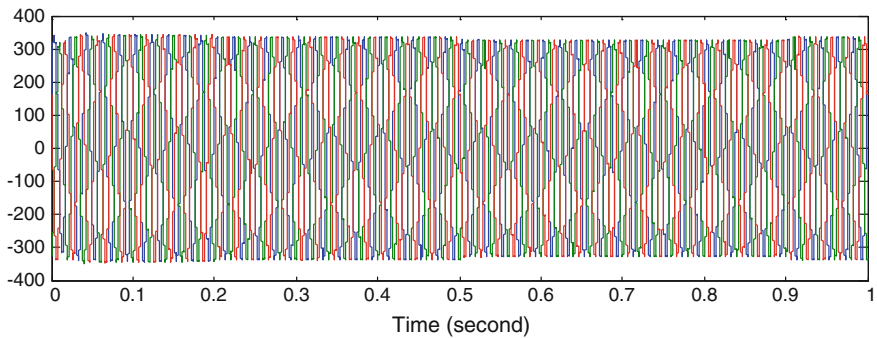
Fig. 12 Inverter voltage

The load voltage, inverter voltage, and grid voltage are shown in the Figs. 11, 12, and 13, respectively.

The load current, inverter current, and grid current are shown in the Figs. 14, 15, and 16, respectively.
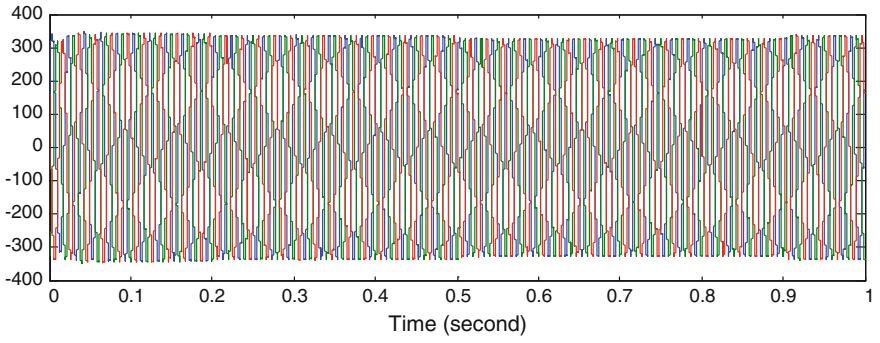
Fig. 13 Grid voltage
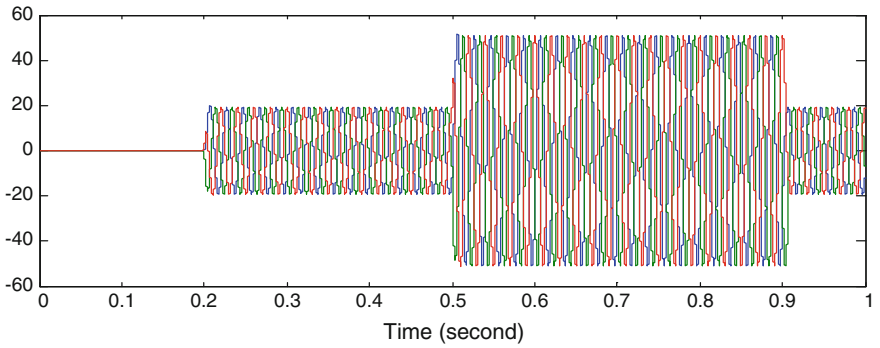
Fig. 14 Load current
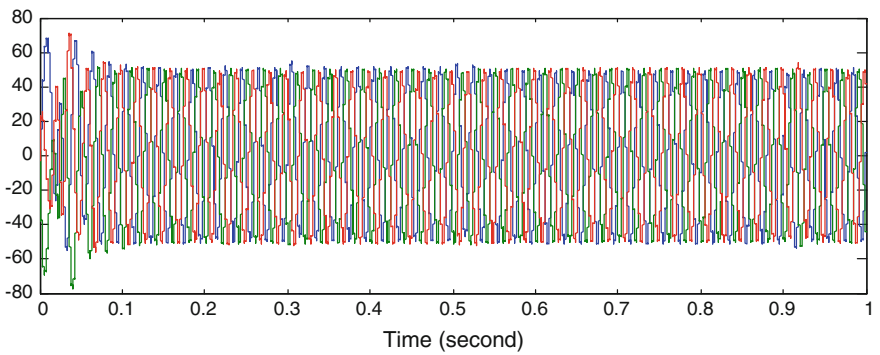
Fig. 15 Inverter current
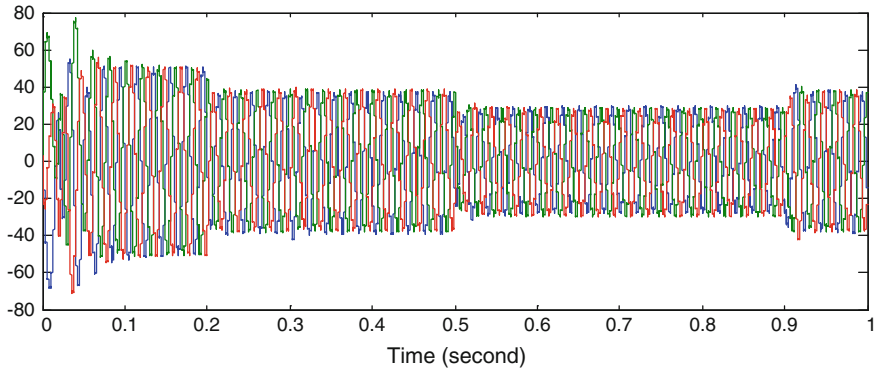
Fig. 16 Grid current

4.1 Observations

The grid current is about 50 A from 0 to 0.2 s, it is reduced to 40 A when a load of 4 kw is connected through the breaker at 0.2 s and it is reduced further to 30 A when another load of 7.5 kW is connected through the breaker at 0.5 s. Similarly, the load current is about 20 A from 0.2 to 0.5 s and it is 50 A from 0.5 to 0.9 s. The grid voltage and inverter voltage is about 325 V.

4.2 Future Scope

In this work, control strategy is developed for grid connected hybrid system. Implementation of control strategy for stand-alone hybrid system can be done to supply power to the local loads. In order to increase the reliability, battery energy storage system can be connected across the DC link.

5 Conclusion

The grid connected wind/PV hybrid system is modeled in MATLAB/Simulink environment. The wind and photovoltaic sources are integrated at the main DC bus via DC–DC boost converter. The circuit-based photovoltaic model has been considered in which the incremental conductance control method is implemented. This paper mainly focused on the dynamic behavior of the hybrid system under change in generation and load conditions. The simulation results show that the proposed wind/PV hybrid system is a viable option for microgrid applications.

References

1. Sungwoo Bae and Alexis Kwasinski, "Dynamic Modeling and Operation Strategy for a Microgrid with Wind and Photovoltaic Resources", IEEE Transactions on Smart Grid, Volume: 3, No: 4, Pages 1867–1876, December 2012.
2. K. T. Tan, P. L. So, Y. C. Chu, and K. H. Kwan, "Modeling, Control and Simulation of a Photovoltaic Power System for Grid-connected and Standalone Applications", IEEE- IPEC, Pages 608–613, (2010).
3. Jitendra Kasera, Ankit Chaplot and Jai Kumar Maherchandani, "Modeling and Simulation of Wind-PV Hybrid Power System using Matlab/Simulink", IEEE Conference on Electrical, Electronics and Computer Science, Pages 1–4, (2012).
4. D. Kastha, S. N. Bhadra, S. Banerjee, "Wind Electrical Systems," Oxford University Press, New Delhi, (2009).
5. Chen Wang, Liming Wang, Libao Shi and Yixin Ni, "A Survey on Wind Power Technologies in Power Systems", IEEE Conference Publications, Pages 1–6, (2007).
6. Alejandro Rolan, Alvaro Luna, Gerardo Vazquez and Gustavo Azevedo, "Modeling of a Variable Speed Wind Turbine with a Permanent Magnet Synchronous Generator", IEEE International Symposium on Industrial Electronics, Pages 734–739, July (2009).
7. B. M Hasaneen and Adel A. Elbaset Mohammed, "Design and Simulation of DC/DC Boost Converter", IEEE Conference Publications, Pages 335–340, (2008).
8. Frede Blaabjerg and Zhe Chen, "Power Electronics for Modern Wind Turbines", A Publication in the Morgan & Claypool Publishers' Series. IEEE 1547, "IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems". IEEE Standards Coordinating Committee 21 on Fuel Cells, Photovoltaics, Dispersed Generation, and Energy Storage, (2006).
9. Fei Ding, Peng Li, Bibin Huang, Fei Gao, Chengdi Ding and Chengshan Wang, "Modeling and Simulation of Grid-connected Hybrid Photovoltaic/Battery Distributed Generation System", China International Conference on Electricity Distribution, Pages 1–10, (2010).
10. R. Pena, J. C. Clare, G. M. Asher, "Doubly fed induction generator using back to back PWM converters and its application to variable speed wind energy generation," in Proc. IEE Electric Power Applications, Pages 231–241, May (1996).

Requirements Prioritization: Survey and Analysis

Sita Devulapalli, Akhil Khare and O.R.S. Rao

Abstract Survey on requirements prioritization to identify the practices related to requirements prioritization among the software development organizations and to understand association of requirements prioritization' effects on software deliveries and resources is designed and conducted for projects/products in different domains across organizations. The results are analyzed for identifying areas that needed attention in terms of requirements prioritization. The survey and analysis enable understanding the need for requirements prioritization for stable and smooth release cycles. A multi-level framework utilizing the concepts of ABC analysis is suggested as a method for prioritization, for predictable and stable releases.

Keywords Requirements prioritization · Multi-level framework · Stable release cycles

1 Introduction

Software solutions being built for various medium to small size organizations to enable them to leverage software solutions for their businesses are often taken up by startup or small companies. While large companies offer generic solutions as products surviving through years and provide customization for specific business needs, there is a good mix of new customized solution offerings developed a new by companies as well as customized solutions on generalized solutions meeting the needs of IT enablement of business. Similar to off-the-shelf products' initial versions, the development of software starts as a solution development and continues to

Sita Devulapalli (✉) · Akhil Khare · O.R.S. Rao
ICFAI University Jharkhand, Ranchi, Jharkhand, India
e-mail: sitadpalli@yahoo.co.in

Akhil Khare
e-mail: khareakhil@gmail.com

O.R.S. Rao
e-mail: orsrao.icfai@gmail.com

undergo enhancements and fixes, thereby evolving into business-specific products. They are certainly not one time buys and live through versions of modifications till scaling of business demands a new solution or simpler and new technology-based solutions are needed. And the cycle of new product solutions begins. Through these cycles, requirements are gathered, analyzed, refined, and prioritized as per client's business needs, technology changes, and resource needs.

In this cycle of product solutions development, often there is less clarity on requirements in the initial stages and requirements change frequently in nature and scope. Changing business needs during the development phase also results in changes in requirements. Chasing the changes in requirements often results in increased development efforts, over worked teams, and extended release dates. In order to understand the requirements handling process during the software development, a survey is designed to gather current methods, difficulties faced, and solutions adopted.

The survey is structured around parameters like products domain, maturity of the products, development process variations, and requirements handling modes. The survey is designed based on the author's industry experience in software products development. The objective of the survey is to identify practices related to requirements' prioritization among software development organizations and to understand association of requirements prioritization' effects on software deliveries and resources. The survey is conducted to understand the effectiveness of the current processes and to identify requirement's prioritization needs for enabling planned deliverables with reduced uncertainties.

The respondents participated in this survey range from organizations that are long term, enterprise products players to relatively new and single product/custom software players. The domains are related to engineering fields to commerce applications to gaming solutions. Some of the products have been under continuous enhancement and maintenance for years.

Different processes—waterfall, iterative, agile—are followed across the organizations. The products developed are typically used by large customer base of the clients for specific applications on different platforms and devices. Products undergo modifications to meet further requirements of the clients, often changing requirements as the development progresses. Providing the customers with ever enhancing products is made possible by successive releases of products at varied intervals, ranging from few weeks to few months to 1 or 2 years.

The fundamental questions that need to be addressed are—What is to be made available in the next release? How to manage the requirements under expanding client needs, cost and time implications? Will prioritization of requirements and planning releases help to streamline the project deliveries to client's satisfaction without overworking the teams or missing time to market deadlines?

This paper describes the survey conducted to bring out information about the domains and applications, process of development, how requirements are handled currently—in Sect. 2. The survey questions are prepared based on the author's experience with product development. The nature of responses and analysis of significant responses is presented in Sects. 3 and 4. Improvements that are feasible and a framework that can help simplify the process of requirements prioritization are discussed further in Sect. 5.

2 Survey, Respondents, and Organizations

The organizations of respondents varied from large (>200 employees) to medium (25–200 employees) sized to small (<25 employees), with local and global presence of the products. Some of the organizations have multiple product lines, while some have single product lines. The survey responses are gathered from 53 respondents belonging to 20 organizations. The respondents are involved in business analysis, project management and product development.

The survey questionnaire is divided into three parts. The first part elicited data related to the domains of the project, nature of the project, the role played by the respondent, the stage the product is in, and release cycle durations with 12 questions.

The part II focused on the process followed for development and gathered information on what process is followed for development, how the requirements are collected and analyzed, problem areas like over work or over-runs on time with 10 questions. Part III focused on how the requirements are handled across the projects and has 20 questions, covering collection of requirements, prioritization methods used, areas of problems, and current solutions adopted. Responses to part II and part III are presented in the following section.

3 Nature of Responses on Processes

The survey has 3 parts and 42 questions, overall, and notable points are discussed here. The domains of applications developed varied from engineering to consumer applications, across manufacturing, telecom, and finance to e-commerce. The product's life cycle stage varied from less than 2 years to greater than 10 years. The applications are typically enterprise applications, web applications, and mobile applications working across devices and platforms, used in multiple countries and are mostly three-tier applications. The processes followed are waterfall, iterative, agile, and agile being the predominant process. Classification of respondents' data is presented in Table 1.

Table 1 Development process, complexity

Size of org	Respondents following			Respondents working with products complexity		
	Waterfall/iterative/agile	Iterative/agile	Agile	3 tier/n tier	2 tier	Single tier
Large	7	2	8	15	1	1
Medium	2	3	21	16	8	2
Small		3	7	7	1	2
Total	9	8	36	38	10	5

No specific method is used for requirement's prioritization. The focus has been mostly on the customer demand for specific features. Relevant parameters considered for requirements prioritization are—business value (BV), availability of resources (AR), difficulty of implementation (DI), and impact on existing customers (IC) without weight to the parameters, mostly. The responses collated with regard to current requirements selection criteria, problems faced, and solutions adopted are presented in Tables 2, 3 and 4.

Table 2 Current processes

Features/requirements to be implemented for next release		Problem areas with current process of feature selection for upcoming release		Circumventing the problems with current process of feature selection	
Based on customer needs	36	Estimation-time resources	14	Client management/meetings	10
Time to market	5	Lack of prioritization wrt complexity, time	12	Discussions with stakeholders	10
No preference	6	Requirement clarity/changes	12	Do nothing	5
Enhancements	1	Dependencies—other modules, new tech	6	Extra time and hard work	6
Impact analysis	5			Estimate/extend/analyze	9
		No response	9	No response	13

Table 3 Additional time needs and replanning

Response	Teams working for release under pressure and for long hours in a day	Some of the team members over worked during releases	Abandoning features being implemented for a release and restart on new features
Often	19	18	2
Rarely	7	3	24
Sometimes	21	23	26
Very often	5	8	

Table 4 Resource availability, impacts, rework of resources' bandwidth

Response	Right resources availability is an issue for meeting release schedules	Abandoning features during release due to realized impacts on existing customers	Analyzing the impacts on core structure/architecture/data model, of features to be implemented a priori	Reworking of resource (time, personnel, S/W, H/W) estimates for the features during the development cycle for a release
Often	20	5	21	23
Rarely	11	26	1	6
Sometimes	16	17	11	22
Very often	5	2	19	1

A clear and systematic approach was not apparent from the responses for the question—How does the respondent choose features/requirements to be implemented for next release. It is largely based on customer needs alone. Changes in the requirements often resulted in extending the dates for releases.

Analysis of the responses to the question—"What are the problem areas you see with your current process of feature selection for upcoming release?" narrows down the problem areas with the current process followed to analysis, estimation, and planning.

The question—How do you circumvent the problems with your current process of feature selection?—indicated to typical solutions being followed like over work, extended releases, and attempts to convince clients. Often the teams worked under pressure and for long hours in order to meet requirements for release.

Response to the question—How often do you have teams working for release under pressure and for long hours in a day?—is given in Table 3.

Lack of clear-cut requirements analysis prioritization resulted in teams working for additional time often and also in replanning the releases by abandoning some features and adding new features into the release.

Analyzing and taking into account resources availability impacts on existing customers are two areas that seem to be only partially considered for defining requirements for upcoming releases. Impacts of new requirements on existing product structures is another area that seem to be not taken into account by everyone. Table 4 gives classification of data on these three aspects and resulting rework of resource bandwidth for releases.

4   Responses to the Survey on Requirements Prioritization

Requirement collection across the organizations appears to be through all channels available—marketing, existing clients, executive direction, and development team. Development teams and planning teams are providing the assessment/analysis mostly. Simple classification of requirements into three groups of—must have, good to have and need not have—appears to be the familiar method followed for requirements prioritization for product releases. Tables 5, 6, and 7 present the data for requirements collection, and classification.

Table 5 Requirements classification

Weights are associated with parameters considered for prioritization		A multi stage prioritization scheme is useful for requirements prioritization		Working out prioritization exactly for each requirement for product releases	
Most often	5	Most often	27	Most useful	21
No weights	10	Not used/never	2	Not useful	1
Not often	11	Not often	13	Not useful often	9
Often	23	Always	8	Often useful	20

Table 6 Requirements prioritization

Requirement prioritization/response	Most often	Always	Not always
It is essential to know how much important each requirement is when compared to other for prioritization	25	21	4
It is sufficient to know relative importance of requirements for prioritization rather than "how much more important"	28	13	9
Cost—value ratio for requirements is the best indicator of priority	21	4	25
Classifying into—1. must have 2. GOOD to have 3. can live without—groups	29	13	8

Table 7 Requirements prioritization

Response	Ranking of requirements (in sequence of priority) based on a parameter is sufficient for prioritization	Numerical assignment of priority (grouping by assigning priority 1, 2,3, …) to requirements is sufficient	Requirements prioritization provides traceability along the product life cycle for improved quality of the product
Most often	27	26	22
Always	7	6	16
Not always	16	19	13

Association of weight to prioritization parameters for requirements is used in some organizations. Similarly, exact prioritization for requirements finds favor to a good extent. A multi-stage prioritization method is expected to be useful to a large extent as is evident from data in Table 5.

Relative importance and quantification of relative importance appear to be important for prioritization. Perception about cost–value ratio for prioritization has a mixed response. Table 6 indicates the respondents' preferences.

Ranking of requirements based on a preferred parameter, numerical assignment of priority for prioritization of requirements do find a favor by many, though not by all participants. Prioritization of requirements added improved quality as seen in the Table 7.

5 Analysis and a Framework for Prioritization of Requirements

Responses to the survey indicate a need for focus on requirements prioritization for planning releases systematically, with controlled changes during the course of release cycle. The methods being used appear to be relative ranking and grouping

into—must have, good to have, and need not have. Utilization of weighted parameters for requirements prioritization, adopting multi-level prioritization finds a place in practice, though not by all. Lack of appropriate requirement prioritization methods, process often appears to have resulted in teams working under pressure, extended release dates, and dropped features.

The survey covered large companies with mature products releasing successive versions of products with longer release cycle, as well as midsize to small companies working on specific project based product versions with less maturity and shorter duration release cycles. Across this range of organizations, requirements analysis and prioritization for products/projects first versions, as well as successive versions is an area that needed attention and systematic methods to be adopted for stable, successful, and smooth deliverables in a predictable manner. Taking the nature of products/projects and the process prior to development as constraints, requirements prioritization for the purpose of predictable releases of products is analyzed. The following baseline is suggested for the requirements prioritization.

The purpose of getting a set of requirements implemented for the next release (time bound) is to maximize the business value of the release for the most valued customers. A strict ordering of requirements may not be the need. Need is more for a near-optimal sets of requirements. Since a release is always timed to meet customers expected needs, the following additional constraints are considered for prioritization of requirements

1. Time/duration—minimum time required for development.
2. Nature of development needed for the requirements.
3. Resources—knowledgeable in domain/technology/skill.
4. Uncertainties—changes due to expanded/extended scope.
5. Impacts on existing customers and existing product modules.

Based on the above considerations, the following framework [1] is suggested for simple and effective prioritization at multiple levels enabling implicit weight application for relevant parameters for the requirements, which enables flexible planning through the development cycle. The framework provides visualization for the changes in requirements during the release cycle and acts as a easy communicator to the involved stakeholders including testing team members.

The framework is defined as five sets based on most used parameters in the sequence of priority determination. Each set is defined by three classes/bins defined by % value of the respective set parameters. Requirements are grouped into the classes in the sets in the process of prioritization. The percentage bands may vary from industry to industry and organization to organization to some extent.

Prioritization sets—S1–S5 and classes/bins—A, B, C within are described in Table 8.

The framework is applied in a layered approach through the sets. The order of preference emerges for the requirements set through the filtering process. Not all

Table 8 Framework—sets, classes

Sets	Classes/bins—A, B, C
S1. Business value (BV) in conjunction with customer base (CB)	A: 20 % of CB with 70 % BV
	B: 30 % of CB with 25 % BV
	C: 50 % of CB with 5 % BV
S2. Requirements applicability with respect to product, where UW: user interface, BI: business logic, CP: core	A: 70 % UW, 30 % BI, 0 % CP
	B: 50 % UW, 40 % BI, 10 % CP
	C: 30 % UW, 50 % BI, 20 % CP
S3. Implementation cost, where MI: marginal implementation, NI: new implementation, IR: impact recovery	A: 70 % MI, 25 % NI, 5 % IR
	B: 50 % MI, 40 % NI, 10 % IR
	C: 30 % MI, 50 % NI, 20 % IR
S4. Time requirement, where L: 8–16 person weeks, M: 4–8 person weeks, S: 2–4 person weeks	A: 10 % L, 20 % M, 70 % S
	B: 15 % L, 25 %M, 60 % S
	C: 20 % L, 30 % M, 50 % S
S5. Resource requirement, where RC: core aware, RI: industry aware, RT: technology aware	A: 10 % RC, 20 % RI, 70 % RT
	B: 15 % RC, 25 % RI, 60 % RT
	C:20 % RC, 30 % RI, 50 % RT

sets may be required to be used. When all sets are used for classification, we will arrive at 243 bins of requirements. Based on the constraints and release theme, the bins can be selected in the order of preference for the releases.

6   Conclusion

The survey conducted across organizations developing software products—first version to multiple versions, brings out the lack of systematic methods usage for requirements prioritization. It also brought out the associated problem areas and difficulties in achieving successful software product deliveries. Prioritization of requirements based on parameters relevant to the product development a priori and during the development cycle facilitates stable and predictable deliveries with less resource allocation uncertainties. The framework suggested enables simple and effective methodology for requirements prioritization for successive releases under dynamic changes and leads to better understanding and planning of releases. It helps build traceability and visualize effects of plan changes and helps in informed quality planning.

Reference

1. Sita Devulapalli, Akhil Khare: "A Framework for Requirement Prioritization for Software Products" IUJ Journal of Management, Vol 2, No. 1, May 2014.

Bibliography

2. Frank Moisiadis.: THE FUNDAMENTALS OF PRIORITISING REQUIREMENTS. Systems Engineering, Test & Evaluation Conference, Sydney, Australia, October 2002.
3. Karlsson J, Ryan K.: A cost-value approach for prioritizing requirements. IEEE Software 1997;14(5):67–74.
4. Wiegers KE.: Software Requirements, Microsoft Press (1999): Redmont, DC.
5. Karlsson J, Wohlin C, Regnell B: An evaluation of methods for prioritizing software requirements, Inform. Software Technol. 1998, 39(14–15): 939-947.
6. Alan M. Davis, Ed Yourdon, Ann S. Zweig: Requirements Management Made Easy. 39–947. www.omni-vista.com.
7. Regnell, B., H¨ost, M., Natt och Dag, J., Beremark, P., Hjelm, T.: An Industrial Case Study on Distributed Prioritization in Market-Driven Requirements Engineering for Packaged Software. Requirements Engineering 2001, vol 6, no 1, pp 51–62.
8. Laura Lehtola, and Marjo Kauppinen: Suitability of Requirements Prioritization Methods for Market-driven Software Product Development. Software Process Improvement Practice 2006; 11: 7–19.

GPU Acceleration of MoM for Computation of Performance Parameters of Strip Dipole Antenna

Hemlata Soni, Pushtivardhan Soni and Pradeep Chhawcharia

Abstract This paper presents the use of general purpose graphics processing units (GPUs) computing to accelerate impedance matrix assembly phase of one of the popular computational electromagnetic method, namely method of moments (MoM). MoM is widely used computational electromagnetic (CEM) technique for solving electromagnetic problems governed by an electric field integral equation (EFIE), and ideally suited for radiation and scattering problems. To validate accuracy, radiation analysis of strip dipole antenna using standard Rao-Wilton-Glisson (RWG) basis and weighting functions which is a good trade off between accuracy and complexity, is considered for the serial and parallel implementations. Performance parameters of strip dipole antenna are computed as postprocessing part of the simulation process.

Keywords CEM · GPGPU · MoM · CUDA · Strip dipole · Radiation

1 Introduction

With the continued and rapid growth in computer science, modeling and numerical simulation has grown exceedingly as a tool for understanding and analyzing complex problems in electromagnetics [1]. Most of the complex problems of antenna engineering are handled by the CEM techniques such as finite element method (FEM), finite-different time-domain (FDTD) method, method of moments

Hemlata Soni (✉) · Pradeep Chhawcharia
Department of Electronics and Communication Engineering, Techno India
NJR Institute of Technology, Udaipur, Rajasthan, India
e-mail: hemlatasoni90@gmail.com

Pradeep Chhawcharia
e-mail: pradch123@gmail.com

Pushtivardhan Soni
Department of Electrical Engineering, IISc, Bangalore, Karnataka, India
e-mail: pushtivardhan@gmail.com

(MoM), physical optics (PO), and geometrical optics (GO) [2]; where the MoM is ideally suited for scattering and radiation analysis [3].

The MoM provides an approach to reduce functional equation into matrix equation [4]. The task of obtaining solution for the functional equation is now reduced to a sequence of two subtasks. First is to assemble a matrix corresponding to the problem and secondly, solving this matrix equation by means of numerical methods.

As the demand for repeated simulations along with higher accuracy in EM design flow goes up, the time required for computation becomes critical and there is a constant drive for "faster, better, and cost effective" solution for these increasingly growing complex problems [1]. Although, various field solvers such as FEKO and momentum that use MoM in their core are available but the need for high performance computing (HPC) in CEM, is still prominent, in present [1] and one of the approaches to achieve this is parallelism [5]. Parallelism is probably the only key unleashing the performance from the modern hardware consisting massive parallel processing cores. One of the fields where the move towards massive parallelism is the field of general purpose GPGPU computing [6].

Among the solution phases of MoM—assembling impedance matrix, excitation vector, and solving matrix equation, impedance matrix assembly is the most compute intensive phase and involves massive data-based parallelism; computation of each matrix element requires execution of a common program with unique data set. Therefore the impedance matrix assembly phase is subjected to the GPU acceleration [7] using CUDA (compute unified device architecture) that supports single instruction, multiple data (SIMD) paradigm.

For computing purpose and to validate the performance results, strip dipole antenna is considered as a radiating body and the discussion of the results will be based on this problem.

2   Theoretical Aspects of MoM

A linear problem of any discipline can be solved (approximated) with the application of MoM [8]. However, the naming convention may differ, but the underlying concept remains the same. The general MoM-based approximation entails the following steps (Table 1):

Table 1 A general outline of solving a linear problem via method of moments	1. Formulating the problem in a form as functional equation
	2. Choosing appropriate basis and weighting functions
	3. Assembling the impedance matrix [Z]
	4. Assembling the excitation vector [V]
	5. Solving for current vector [I] by solving the matrix equation $[Z][I] = [V]$
	6. Assembling the approximate solution

3 Impedance Matrix Assembly Process

In MoM, the unknown quantity (usually the surface current) is discretized spatially through meshing. An appropriate functional dependence also known as basis is then defined over the elements, which describes spatial variation of unknown quantity over the element. The unknown is determined by the application of the method to the patchwork of elements which approximates the original geometry, where the accuracy of the solution is related to the degree of discretization known as mesh size. Although, simple rectangular patch functions can be used, but the high approximation error makes it unsuitable for arbitrary surfaces. In this presented paper, commonly used triangular patch functions, i.e., Rao-Wilton-Glisson (RWG) function, are considered for the radiation analysis [9]. In RWG patchwork, a common edge between couples of triangles serves as a basic element known as RWG edge element. An edge element comprises two triangles, labeled as positive and negative triangle. Each triangle has three vertices; two of them are shared by the common edge and third is known as free vertex.

A most popular form which uses electric and magnetic potentials to derive impedance matrix element is given by

$$Z_{mn} = l_m \left[\frac{j\omega}{2} \left(\overrightarrow{A_{mn}^+} \cdot \overrightarrow{\rho_m^{c+}} + \overrightarrow{A_{mn}^-} \cdot \overrightarrow{\rho_m^{c-}} \right) - \left(\phi_{mn}^+ - \phi_{mn}^- \right) \right] \tag{1}$$

Where the constitutive equations are

$$\overrightarrow{A_{mn}^\pm} = \frac{\mu}{4\pi} \left[\frac{l_n}{2A_n^+} \int_{T_n^+} \overrightarrow{\rho_n^+}(\vec{r}) G^\pm \left(\overrightarrow{r_m^{c+}}, \vec{r} \right) \mathrm{d}S + \frac{l_n}{2A_n^-} \int_{T_n^-} \overrightarrow{\rho_n^-}(\vec{r}) G^\pm \left(\overrightarrow{r_m^{c-}}, \vec{r} \right) \mathrm{d}S \right] \tag{2}$$

$$\phi_{mn}^\pm = -\frac{1}{4\pi j\omega\epsilon} \left[\frac{l_n}{A_n^+} \int_{T_n^+} G^\pm \left(\overrightarrow{r_m^{c+}}, \vec{r} \right) \mathrm{d}S + \frac{l_n}{A_n^-} \int_{T_n^-} G^\pm \left(\overrightarrow{r_m^{c-}}, \vec{r} \right) \mathrm{d}S \right] \tag{3}$$

$$G^\pm \left(\overrightarrow{r_m^{c+}}, \vec{r} \right) = \frac{e^{-jk \left\| \overrightarrow{r_m^{c+}} - \vec{r} \right\|}}{\left\| \overrightarrow{r_m^{c+}} - \vec{r} \right\|} \tag{4}$$

where l_n is the length of the edge, $T_n^\pm$ is the positive and negative triangles associated with the edge, $\vec{r}$ and $\vec{\rho}$ are the position vectors of a point on the surface with respect to origin and free vertex respectively, the superscript c denotes the centroid point on the triangle, A_n^+ and A_n^- are the areas of positive and negative triangles, respectively.

Fig. 1 Primary triangle
subdivided into nine
subtriangles
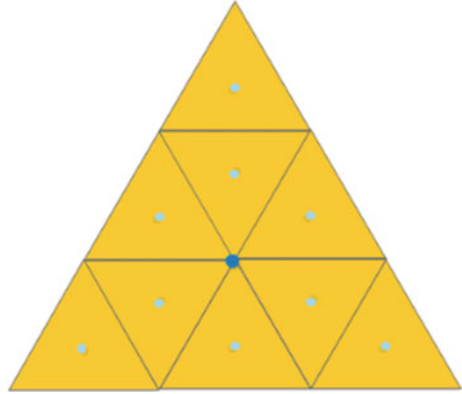
The surface integration operations of the above formulae are implemented using quadrature integration method. A triangle is divided into 9 subtriangles as shown in Fig. 1, over which the integration is assumed to be constant.

The integral of a function $f(\vec{r})$ over triangle is then approximated as (5)

$$\int_{r \in \text{Triangle}} f(r)\, dS = \frac{\text{AreaTotal}}{9} \sum_{i=1}^{9} f(r_i^{\text{center}}) \tag{5}$$

4 Data-Level Parallel Implementation

The formulation used for computing matrix element Z_{mn} involves the integrations over the triangles T_n^+ and T_n^- associated with source edge n, where the integrand (i.e., Green's function) is defined over the triangle pair T_m^+ and T_m^- associated with the observation edge m. Therefore, four integrations are required to calculate one matrix element.

For the purpose of parallel implementation, an edge-pair based approach is used for computing each matrix element. An edge-pair based approach suggests computing all four integrations associated with source and observation edges n and m, for each matrix elements. In this method, each matrix element can be computed independently by processing only its edge m and n data; this makes the matrix assembly a data-parallel task. In other words, a single set of instruction (or program) for multiple data objects, which is a SIMD style, for which GPUs is well suited.

In CUDA setting, the impedance matrix is assembled by launching CUDA threads that shares a common kernel program [10].

```
MoM_ImpedanceMatrix_GPU   <<<dimGrid,   dimBlock   >>>
(Edge_dev, f_dev, epsilon_dev, mu_dev, Z_dev);
```

where dimGrid and dimBlock specify the dimension of the grid (in blocks) and the dimension of the blocks (in threads). Each thread is responsible for computing unique element of the matrix [11]. A thread identifies the edge data by its block (blockIdx) and thread (threadIdx) identifiers as

```
unsigned int m = blockIdx.x*TILE_WIDTH + threadIdx.x; //ob-
servation edge
unsigned int  n =  blockIdx.y*TILE_WIDTH  +  threadIdx.y;
//source edge
```

5   Validation and Performance Results

For the purpose for verification of the results, surface current density is considered as the quantity of interest, and the reference values of current density is computed with FEKO (a commercial CEM software package), [12] (Figs. 2, 3, 4 and Tables 2, 3, 4).

6   Conclusion

The current density results show a close agreement with the reference values generated with a wide used commercial electromagnetic simulator FEKO. The GPU-implementation outperforms the CPU implementation, which reflects the power of parallelism. Moreover, due to the widespread availability of GPUs in

Fig. 2 Plot showing magnitude of current density computed with CPU and GPU-implementation with respect to spatial coordinates (frequency = 75 MHz)
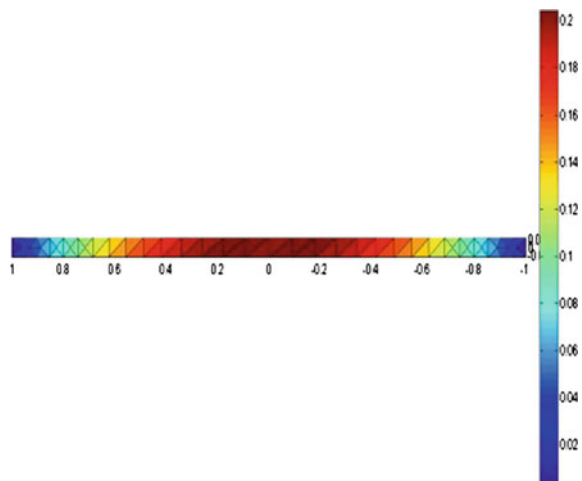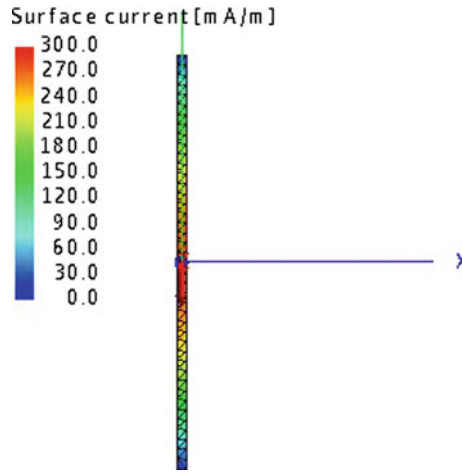
Fig. 3 Plot showing magnitude of current density computed using FEKO with respect to spatial coordinates (frequency = 75 MHz)
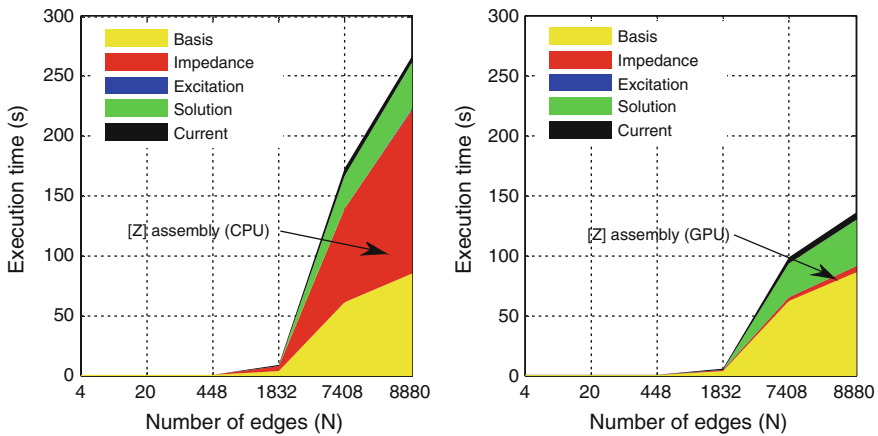
Fig. 4 Execution time for various phases of MoM

Table 2 Specifications of the computing system 1 used for testing

Specifications	CPU (HOST)	GPU (DEVICE)
	Intel Core i7-4510U	NVIDIA GeForce GT 850M
Architecture	Haswell	Maxwell GM107
Cores	2	640
Clock frequency	3.1 GHz	901 MHz
Memory	8.00 GB	4 GB
Operating system	Windows 8.1 single language, 64-bit	

Table 3 CPU (italics) and GPU (bold) execution time (in seconds) for various phases of the MoM, tested with system 1

N	4	20	448	1832	7408	8880
Basis	*0.0228*	*0.0239*	*0.2900*	*3.9152*	*61.610*	*85.144*
Impedance	*0.0078*	*0.0071*	*0.2501*	*4.5864*	*77.648*	*137.38*
	0.0005	**0.0004**	**0.0168**	**0.2698**	**3.8332**	**5.4303**
Excitation	*0.0001*	*0.0005*	*0.0056*	*0.0227*	*0.0932*	*0.1192*
Solution	*0.0002*	*0.0004*	*0.0173*	*0.5436*	*27.409*	*39.032*
Current	*0.2987*	*0.3865*	*0.3988*	*0.7529*	*5.9805*	*5.6473*

Table 4 Other performance parameters of strip dipole antenna, [13, 14]

Electric and magnetic field at point (0, 0, 5)	E-Field (V/m)	H-Field (A/m)
	$E_x = -0.0000 - 0.0000i$ $E_y = -0.0043 - 0.0052i$ $E_z = 0.0000 + 0.0000i$	$H_x = 0.1137 + 0.1379i$ $H_y = -0.0001 - 0.0000i$ $H_z = -0.0000 - 0.0000i$
Radiation intensity	$U = 6.0188e{-}04$ W/unit solid angle	
Radar cross-section	RCS = 5.6993 m^2	
Gain	$G = 2.1593$ dB = 1.6441	
Radiation resistance	$R_r = 92.0355\ \Omega$	

personal computers as a commodity graphics cards, the GPUs accelerated implementation presented is a cost effective solution, which makes high performance computing accessible for slower/older computers with a low addition of cost.

References

1. D. De Donno, A. Esposito, G. Monti, and L. Tarricone, "Parallel Efficient Method of Moments Exploiting Graphics Processing Units," Microwave and Optical Technology Letters, vol. 52, no. 11, pp. 2568–2572, Nov. (2010).
2. D. B. Davidson, "Computational Electromagnetics for RF and Microwave Engineering", ISBN 9780521518918, Cambridge University Press, (2005).
3. J. R. Mautz and R. F. Harrington, "Radiation and scattering from bodies of revolution," Applied Science Res., vol. 20, no. 1, pp. 405–435, (1969).
4. R. F. Harrington, "Field Computation by Moment Methods", ISBN: 978-0-7803-1014-8, Wiley-IEEE Press, (1993).
5. David Luebke, John Owens, "Intro to Parallel Programming Using CUDA to Harness the Power of GPUs", [Online]. Available: https://www.udacity.com/course/intro-to-parallel-programming–cs344.
6. John D. Owens, Mike Houston, David Luebke, Simon Green, John E. Stone, and James C. Phillips, "GPU Computing", Proceedings of the IEEE, Vol. 96, No. 5, May (2008).
7. E. Lezar and U. Jakobus, "GPU-acceleration of the FEKO electromagnetic solution kernel," International Conference on Electromagnetics in Advanced Applications (ICEAA), pp. 814–817, September (2013).

8. W. C. Gibson, "The Method of Moments in Electromagnetics", ISBN 13: 978 1 4200 6145 1, CRC Press, (2007).

9. S. M. Rao, D. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," IEEE Transaction Antennas Propagation, vol. 30, no. 3, pp. 409–418, May (1982).

10. "CUDA Zone," 25-Sep-2014, [Online]. Available: https://developer.nvidia.com/cuda-zone.

11. "CUDA C Programming Guide." [Online]. Available: http://docs.nvidia.com/cuda/cuda-c-programmingguide#axzz3b4d2t4b3.

12. EM Software & Systems-S.A. (Pty) Ltd: FEKO. 2010. Available at: www.feko.info.

13. Constantine A. Balanis, "Advanced Engineering Electromagnetics, 2nd Edition", ISBN 13: 978-0470589489, Wiley Press.

14. C.A. Balanis, Antenna Theory: Analysis and Design, 2nd edition, Wiley, New York, (1997).

Energy Saving Model Techniques in Wireless Ad Hoc Network

Ajaykumar Tarunkumar Shah and Shrikant H. Patel

Abstract A remote impromptu system is gathering of portable hubs that make a multi-jump self-governing framework with no altered foundation. The hubs utilization administration of different hubs in the system to transmit parcel to the destination hubs. Cell phone is battery worked and this is the constrained asset. So the vitality preservation is the basic issue in the system. There are numerous methodologies recommended for vitality protection. We have proposed two vitality effective procedures to decrease vitality utilization at convention level. In first strategies vitality preservation done by decreasing number of course demand message while in the second methods vitality protection done by force control systems.

1 Introduction

A remote ad hoc system is a gathering of portable hubs that form an element autonomous system. Remote ad hoc system does not faith on earlier period plan, for example, base station [1]. In remote uncommonly delegated framework every hub goes about as a switch and also source hub for sending information. Portable system can be characterized into foundation and Mobile Ad Hoc Network (MANET) as per their reliance on distorted structure. MANET can take after the self-motivated topology where hub can join or leave the system whenever. Portable Ad Hoc Networks (MANET) comprises of hubs that change position most of the time. The applications are fabric in devastation administration, salvage operation,

A.T. Shah (✉) · S.H. Patel
Computer Engineering Department, Alpha College of Engineering & Technology,
Khatraj, Gandhinagar 382721, Gujarat, India
e-mail: ajayshahnirma@gmail.com

S.H. Patel
e-mail: Shrikant.patel@alpha-cet.in

vehicular structure, agro detect, infectivity inspection, and some more. Moveable gadgets are battery lived up to expectations, so it is basic to decrease their vitality use. A large portion of exploration in vitality preservation procedures has focused on remote system that are organized around base station and concentrated servers, which do not have the limit connected with little, compact gadgets. In uniqueness, impromptu system is group of remote adaptable hub, which supportively routines a system independently of any altered association. We accept that power touchy plan and evaluation of system convention for that impromptu systems administration air requires functional information of the vitality utilization execution of genuine remote gadgets. Distinctive study recommends diverse procedures vitality in diverse path. This paper has two methods to diminish vitality utilization at convention Steps. The principal systems decrease the vitality utilization by coherently distributing the system by diminishing the quantity of course demand message. While in the second procedure we apply energy-aware method at center level to decrease the transmitting and receiving force of hub. The remaining of the formation as follows. Section 2 explains the related research on energy efficiency, Sect. 3 contains proposed energy techniques, Sect. 4 describes their simulation result as well as Sect. 5 describes the conclusion.

2   Related Work

2.1   *Energy Management-Based Protocols*

Wireless [2] standardization conventions have two sorts of force administrations. Initially, sort is known force spare type for foundation-oriented remote system and second sort is known as IBSS Power Saving mode for architectureless network systems. In the previous system hubs in PowerSaving mode expend less energy contrast with dynamic mode operation. Entrance point cushioned MAC administration date unit and transmits them at assigned time by the assistance of movement evidence map and deferred activity sign guide (DTIM). This sort of force sparing component is not suite for specially appointed system area as there is not focal provider like access point. Then again IBSS PowerSaving mode is relevant completely associated single bounce system where all the hubs are in the signal reach to one another. Facilitated sign interim is established by the hub that is starts the IBSS and it is kept up in a scattered manner. All the hubs wake toward the start of the reference point interim and wake till activity window is not complete. Hubs contributing in the development statement endure cognizant till the process end of reference point interim and the nonparticipator endeavors to rest to save power toward the end of the movement interface. Measure of vitality save by a hub relies on the time used in the rest part of process which can be influenced by the signal move from rest to dynamic mode of data transmission operation. Vitality sparing execution likewise relies on the system measure and additionally on the length of the ATIM and guide interim span. Time administration is must require when we

transmit data from sender node to receiver node through substitute node [3] in multinode system.

Element force sparing component [4] is a change of Wireless PS, as it works the idea of ATIM frame and guide interim. Examined before amid ATIM frame each and every hubs conscious and hubs not movement to get or transmission are swing to remain state mode after the procedure end of ATIM frame. Freeny [5] recommends if ATIM frame is altered then vitality sparing may be influenced. DPSM enhances execution by utilizing the schema ATIM frame. They allow transmitter and beneficiary hub for variety their ATIM frame outline powerfully. The ATIM frame size expanded when a few bundles remain after present frame is lapsed. The information bundles convey the present length of the ATIM frame and the hubs catch this change their ATIM frame size. DPSM permits the sender and collector hub transfer their radio signal promptly after their transmit be complete. This power economy implementation of DPSM is healthier as contend to Wireless DCF in definition of vitality sparing anyway its more computationally muddled.

In PAMAS [6] vitality proficiency objective is accomplished by utilizing two diverse areas, one area for organize then further information. RequestToSend/ClearToSend signs are sent over the control network where information are send over information network. Hubs with bundle to transmit sends a RequestToSend over the control channel, and sits tight for ClearToSend, if no ClearToSend then they gets inside of a particular period then hub enter to a back off part of network. Be that as it may, if CTS is gotten, then the hub transmits the information parcel over the information channel. The accepting hub transmits an occupied tone over the control channel for its neighbors showing that its information channel is occupied. The utilization of control channel permits hubs to focus when and to what extent to energy off. The duration of strength off time is dictate by diverse condition. Subsequent to awakening, hub gets to the channel over the information channel and discovered numerous transmissions going on. The hub utilizes a test convention as a part of this case to discover the amount of time it determination off. Recreation results demonstrate that great scope of force sparing is accomplished in PAMAS.

2.2 Energy Controlling-Based Protocol

Power Control MAC [1] accomplishes vitality sparing with creating debasement by executing distinctive kind of transmitting force. Information and ACKnowledge bundles are sent utilizing least energy while RquestToSend/ClearToSend parcels are sent utilizing greatest force. Beneficiary computes the base force needed by the sender to send information, contingent on the encompassing commotion and obstruction. At the point when the transmission happens the neighboring hub concedes their transmission. Amid information transmission same techniques are utilized for discovering least obliged force level that ought to be sufficient when sending of DATA and ACKnowledge. PowerControlMode requires a precise estimated signal quality based whereupon its energy control meets expectations.

Likewise elements like multipath engendering, blurring, and shadowing impacts may corrupt its execution.

An appropriated sending force save convention to remote system for accomplish vitality protection in the hub level. The reunion uses appropriate computation to fabricate the force sparing tree topologies without taking the neighborhood data of the hubs and give a basic approach to keep up network by converting the broadcast energy. Learn on outcome of energy Saving implementation on IEEE 802.11 remote systems [7] depicts about improvement of phantom again use in substantial scale systems. This convention displays that system with vitality control, getting away emitted hubs can reach higher entire system measure as contrasted and least transmit power technique. The proposed conveyed calculation tries to accomplish high otherworldly reuse by diminishing uncover hub while totally staying away from shrouded hub [8]. The advantages and disadvantages of basic reach and variable-range transmission force control on the physical and system layer integration are pleasantly portrayed by Gomic et al. [9]. The reproduction result demonstrated that variable-range transmission force control performs basic reach transmitter force control in type of vitality sparing and system limit.

2.3   Network Topology Energy Saving Based Protocol

SPAN [6] an appropriated force sparing convention adaptively chooses facilitator from all hubs in the system. Organizer hubs stay conscious constantly and perform multi-bounce parcel directing. Different hubs stay in force spare mode to monitor vitality. Compass accomplishes four objectives, for example, it chooses enough organizer hubs, turns the facilitator hubs to adjust available vitality, endeavors to reduce the quantity of facilitator and chooses the facilitator utilizing nearby data as a part of a decentralize way. Compass gives certification of system network by guaranteeing that each hub has no less than one dynamic hub in its radio extent. Reasonableness among the hubs is in view of the measure of lingering vitality and the extra neighbor combines that a hub can interface. It adjusts both decency and system integration. The whole dynamic hub in SPAN shape a joined spine, every hub intermittently show hi message which incorporates diverse data. From this, hub develops a table containing data like present condition of the hub, present condition of the neighbor, leftover vitality of the neighbor, and so forth.

3   Proposed Methods for Energy Efficient Techniques

In Network area we are propose three method vitality effective procedures in specially appointed system. The primary system reduces course demand message. Then Second strategy advances the transmission energy to every hub and in last methods builds system limit by topological control instrument.

3.1 Techniques for Reduce Route Request

Give a chance to include a multi-bounce homogeneous remote system in hubs are haphazardly conveyed to positive land region. In this manageable specially selected system administration environment every versatile hub can get to the web applications by means of one or more quantities of doors. The portable hubs correspond with the portals straightforwardly (single jump) or through multi-bounce contingent on the transmission scope of hub as indicated at Fig. 1.

Other than web utilize hubs can likewise transfer information among themself. Its expect that portals are stationary. That land region secured for passage divided in diverse sensible gathering and extraordinary gathering number allocated to portal. Gathering be cover and there are a few hubs introduce in the covering region. The apportioned depends on the quantity of hubs present. Hubs are classified in dynamic hub (Active Node) and common hub (Common Node). Hubs exhibit in the covering are gathering is known normal hubs when hubs fitting in with any specific gathering called dynamic hub of this gathering. Its expected to covering zone of distinctive gathering contain require number of regular hubs (Common Node) as between bunch correspondence will happen through this data. At end while active hub required to send course ask for (RouteREQuest) message add and gathering number in parcel exhibit idea. Data sent by the further node on the off chance that they having a place with the same gathering generally message will be drop. At the point where a CommonNode ready the RouteREQuest messages include to gathering number of gathering its having a place contingent on the offer file (SI) figure. SI is ascertained by the CN for the all gatherings it has a place with and adds the gathering number to the RREQ in light of augmentation view of ShareIndex. The ShareIndex is calculated as below.

$$\text{ShareIndex} = \frac{\left[\sum_{i=1}^{G} g_i\right]^2}{G\left[\sum_{i=1}^{G} g_i\right]^2}$$
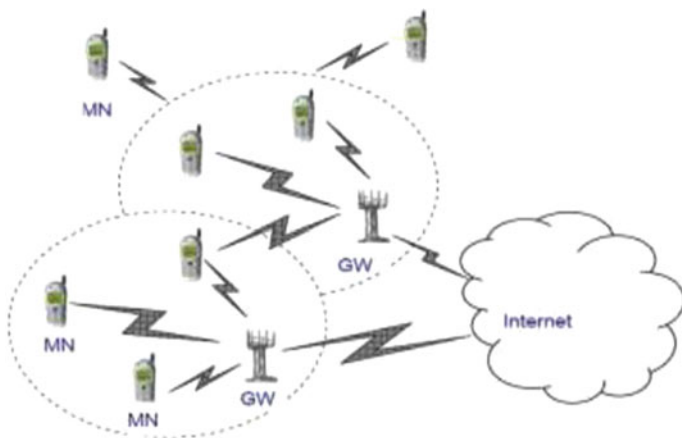
Fig. 1 One-hop and multi-hop transmission using gateways

where ShareIndex $= 1$: for eachone group contain same no. of sender hub

$$< 1 : \text{ in anyother situation}$$

G is quantity for gathering to regular hub and g_i is quantities for hub present ith bunch. ShareIndex counts is defeated burden adjusting reason. At the point when ShareIndex worth is demonstrates gatherings contains equivalent number of hubs. On the off chance that it is under one, then gatherings do not contain equivalent number of hubs. All things considered basic hub will connect for gathering it will boost the SI. The goal for proposed strategy is lessen quantity of demand for putting confinement at internal group correspondence. The node gathering would not forward RouteREQuest messages to gathering node. Just normal hub will bolster between gathering correspondence to diminish the quantity of RouteREQuest [10]. The calculations for ShareIndex computations and transmission method are explained and given below.

Algorithm for ROUTE REQUEST

Step 1: Calculation ShareIndex

$$\text{ShareIndex} = \frac{\left[\sum_{i=1}^{G} g_i\right]^2}{G\left[\sum_{i=1}^{G} g_i\right]^2}$$

Step 2: Procedure for TRANSMITING (node)

1. if (node == CommonNode)
2. Calculate ShareIndex
3. RREQ = RREQ ‖ gn

/* append the Group number (gn) depend upon the maximum value of ShareIndex */

4. Broadcast (RouteREQuest)
5. else if (node == ActiveNode)

   6. RouteREQuest = RouteREQuest ‖ GroupNumber

   /* RouteREQuest containing group number of ActiveNode */

7. Broadcast(RouteREQuest)
8. end of if statement
9. end of if statement

3.2 Power Control Techniques

System picks a domain where hubs conveyed haphazardly in a two-dimensional territory. Every hub has no less than two vitality stages, for example, Pmax and Pmin. Previous is the force needed to reach the most distant hub while the last is the

force needed to achieve the closest hub. The goal of vitality administration is diminishing force use a hub. Its accept that may exists some halfway energy level in the middle of Powermax and Powermin. Let Power(uv) be the force expected to bolster correspondence for hub u to v, then call symmetric if Power(uv) = Power (vu). Power necessity is known Euclidean in the event that it relies on Euclidean separation $\|uv\|$. They can be figured as Power(uv) = $c + \|uv\| \beta$ where c is a Saving steady genuine number, and $\beta \in$ [3, 6] relies transmitting situations. Powermax (u) and Powermin (u) are most extreme at least power of hub u. At the point when Power (uv) ≥ Powermin (u) and Power(uv) ≤ Powermax(u) where correspondence in the middle of u and v is conceivable else it is impractical. At the point when some middle vitality proficient way exists between hub u and v by means of halfway hub w then hub u will transmit to Power(uw) is opposed to Power(uv). In Fig. 2, if Power(uv) ≤ [Power(uw) + Power(wv)] and [Power(ux) + Power(xv)] is correspondence from u to v is occur through Power(uv) generally correspondence through halfway hub w or x by assistance of Power(uw) or Power(ux) separ.

Algorithm for ENERGY CONTROL

Input:
(1) Multihop wireless Ad-hoc network N
Output: Energy levels p for each node to communicate to other node
began
1. for every(u, v) do
    2. calculate Powermin(u) and Powermax(u)
/* for node u */
3. calculate Power(uv) by Euclidean distance
    4. if Power(uw) + Power(wv) ≤ Power(uv)
/* u finds any power efficient path to v via w */
5. then transmit with Power(uw) for v
6. else transmit with Power(uv) for v
7. end if statement
End of Statement

Fig. 2 Energy control methods through intermediate nodes
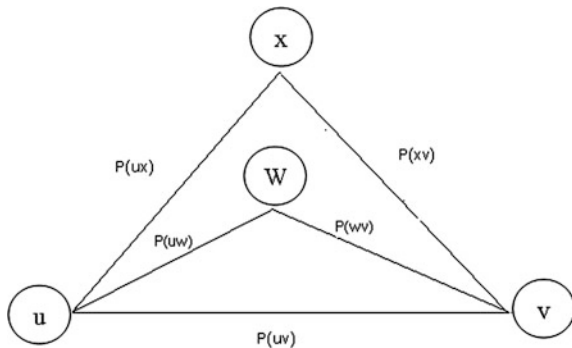
3.3 Topology Control Technique to Increase Network Capacity

Topology control techniques can characterize procedures by organize gadget take their own particular choice with respect to their transmission range, so as to fulfill some system imperatives. A system topology can be symbolized as a diagram. Chart is meant as Graph = (Vertex, Edge), where *V* signifies arrangement of hubs and Edges indicates the arrangement of *E*. Spread (vertexi, vertexj) implies hub vertexj is inside of transmitting scope of vertexi. At all hubs are transmitting to most extreme force Powermax [11]. Contingent on the Pmax esteem hub *u*, (*u* ∈ Vertex) has an immediate correspondence set known as DCS(*u*). P(*uv*) signifies the base force needed for hub u to convey specifically to hub *v*. By applying topology control we need to get subgraph *G′* = (*V*, *E′*) of *G*, in *G′* the hub has shorter and less quantities of edges as contrast with *G*. The goal of topology control here is to uproot the vitality wasteful connection from the system. At first all the hub in the system will figure their DCS relying on their greatest transmission power. Every hub will keep up a hub data table containing data like neighbor_id (NI), direct_communication_cost (DCC), as well as energy_efficient_cost (EEC) [12]. All hub has a novel NI. DCC speaks to the base transmission force needed for a particular neighbor hub. DCC of hub u to neighbor hub v is spoken to as DCC(*uv*) which is same as P(*uv*) as depicted prior. EEC is at first same as DCC, however, when any vitality effective way is gotten through interchange way ECC is redesigned. Every hub occasionally overhauls their hub data table and telecast the data to other hub. After a particular time every hub will figure their DCS by uprooting vitality wasteful connection (if any).

Algorithm for Network LINK REMOVAL

Input: (1) Multi hop remote system Graph=(*Vectore, Edges*),
          (2) Maximum transmission and receiving power
Output: Graph'= (*Vector, Edges'*) , Graph' has shorter and less
quantities of edges as contrast with Graph start

1. Every hub show a 一HI ‖ message with its hub data table.
2. On the off chance that a hub u gets the 一hi‖ message from its neighbor
3. On the off chance that u has a force proficient way to hub v through k
/ * k is a hub in interchange way */
4. upgrade (v, Power(uv), Power(uk) + Power(kv) into u's hub data   table.
/* energy_cost =(Power(uk) + Power(kv) ≤ Power(uv) */
5. Else statement for Insert (v, Power(uv), Power(uv) into u's hub data table.
/* as vitality cost= direct_communication_cost*/
6. End of if statement.
7. End of if statement.
8. End of main statement.

4   Simulation Results

Demonstrate the beneath figure, which figure of the Energy Saving Versus No of Nodes which show that when the number of hub build the vitality sparing may increment (Fig. 3 and Table 1).

The version of NS-2 that we use is NS-2.35. We set the experiment space to be a 670 m^2 × 670 m^2 and place nodes at random in the space. In these simulations we consider mobility. An instance of the NS-2.35 simulator is created, network topology is explained using the provided file and also trace files (Fig. 4).

In this figure we draw a graph between Energy Consumption (*Y*-axis) versus Node (*X*-axis). In this graph we show for 670 × 670 routing range energy consumption for different number of nodes for MANET.

In this figure we draw a graph between Energy Consumption (*Y*-axis) versus Node (*X*-axis). In this graph we show comparison for different scenarios of routing range energy consumption for different number of nodes for MANET (Fig. 5).

Fig. 3 Energy saving versus number of nodes
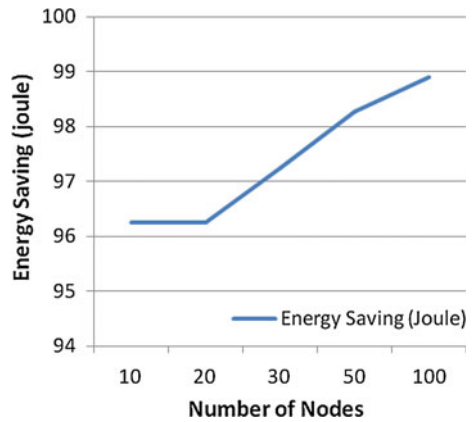
Table 1 Simulation parameter

Type	Value
Transmitting power	0.60 W
Receiving power	0.30 W
Traffic model	CBR
Packet size	512 Bytes
Maximum packet	10,000
Initial energy	100, 1000
Simulation time	200 s
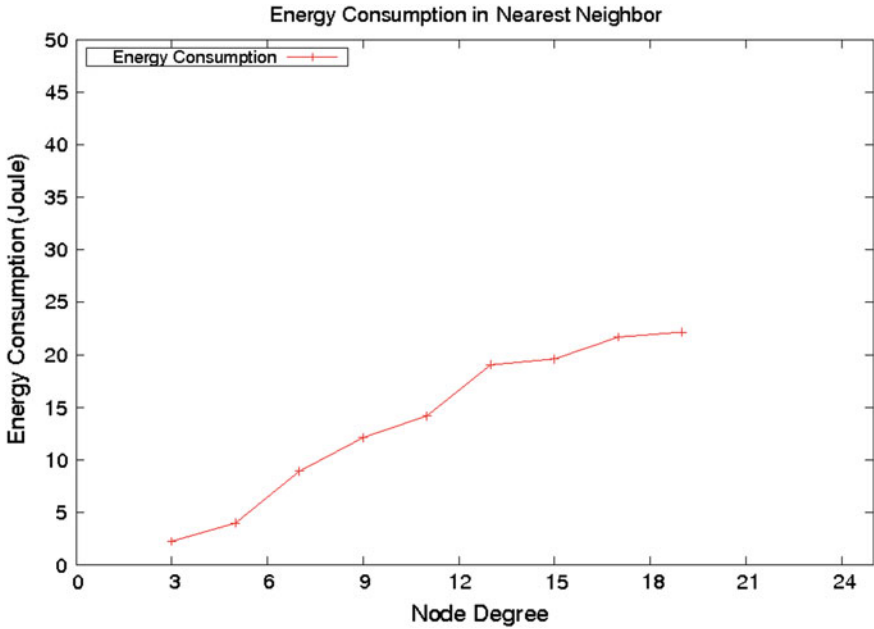
Figure. 4 Energy consumption in nearest neighbor vs. node degree
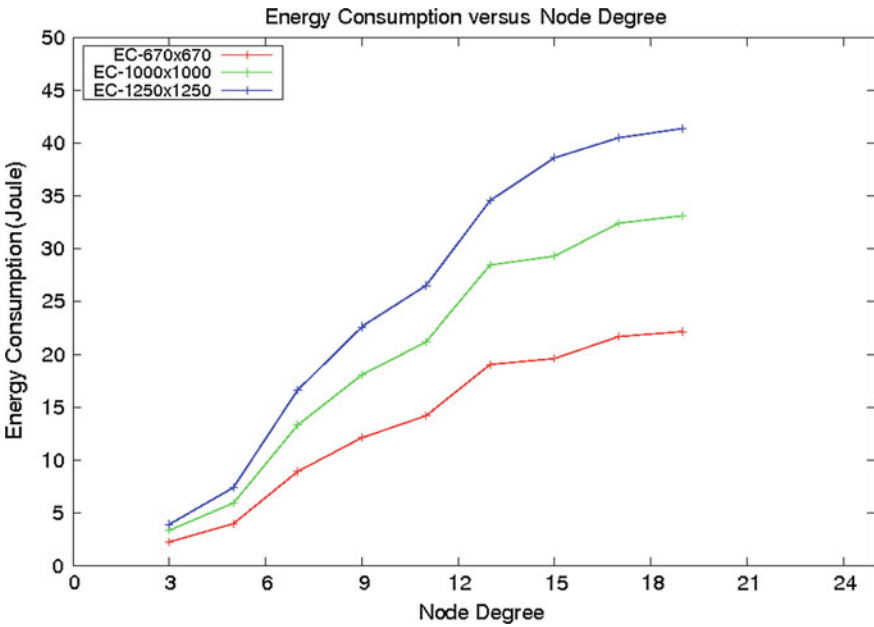
Fig. 5 Comparison of energy consumption versus node degree in different scenarios

5 Conclusions

We assessed a portion of late process complete in portable specially appointed system considering vitality as the main issue. This is find out that main prominent of this research talks vitality problems at information connection and system layer. We examine diverse vitality proficient convention taking into account power administration, force control on topology control process. We propose three vitality productive systems for MANET perspective. Course ask for minimization method should be possible by actualizing sensible gathering; force control strategies diminish the sender force of a hub while Network topology system control methodology expands the system life span by fulfilling system limitations. The reproduction results introduced in segment IV, recommends that multi-bounce is perfect for vitality perspective, however, the confinement is the build possibility of connection disappointment.

References

1. Ajay. Singh and H.S. Raghavendra, PAMAS power mindful multi- access convention with motioning for specially appointed systems, ACM Power Control Communication Review, volume 28(3), pp. 5–26, July 2001.
2. Ethernet Working Areas, Wireless LAN Medium Access Control (MAC) and Physical Layer stylisme, 2012.
3. D. Zhouo and P.H. Laib, A versatile and versatile clock synchronization convention of Ethernet-based multihop specially appointed systems, IEEE International Conference on Mobile Ad hoc and Sensor Systems Conference, Nov 2005, 2005.
4. A.S. Jung lasand N. H. Vaidya, A vitality productive MAC convention for remote LANs, IEEE INFOCOM, June 2002.
5. C.E. Perkins, AdHoc Network Eiddison and Aesley, 2013.
6. L.M. Abraham, –Energy produce procedure in specially appointed systems, Mobile AdHoc Network, Wiley-IEEE xplorer. 301–328, 2004.
7. S.K. Sahu, A.P. Sheoo and S.Y. Hseeeh, –Power control based topology development for the disseminated wireless remote sensor systems, Wireless Communications, vol. pp. 2774–2785, June 2007.
8. S.W. Holand and A.C. Liewer, Impact of Energy Control on execution of Ethernet Wireless system, IEEE Exchange on Mobile Computing, vol. 6(11), pp. 1245–1258, November 2007.
9. Jeng-Pang Shiiw et with out al.,–area topology control convention n remote impromptu systems, Elsevier, computer correspondences journal, vol 31(14), pp. 3410–3419, September 2008.
10. R. Ramana, R. Rosaaless-Haeen, Topological method of multi hop remote system utilizing transmit energy modification, in: Process of the 19th INFOCOM, Tel Aviv, Israel, pp. 403–411, March 2000.
11. A. Chenn, P. Jameson, H. Balakrushna with R. Morrees, Spaan: vitality proficient coordination calculation for topology support in specially appointed remote systems, ACM Wireless Networks Journal, vol. 8(5), pp. 481–494, September 2002.
12. E.S. Jung and N.H. Vaidya, A Power Control MAC Protocol for Specially appointed Networks, ACM Intl. Conf. Portable Computing and Organizing (MOBICOM), September 2002.

Weighted-PCA Based Multimodal Medical Image Fusion in Contourlet Domain

Aisha Moin, Vikrant Bhateja and Anuja Srivastava

Abstract Multimodal medical image fusion is used to fuse the complementary features from diverse modalities and abandon the superfluous information. The fusion of structural medical images-computed tomography (CT) and magnetic resonance imaging (MRI) scans provides to deliver an extensive fused image consisting of obligatory anatomical minutiae to improve medical diagnosis. This paper presents a weighted principal component analysis (PCA) based approach for multimodal fusion in Contourlet domain. The sole aim of using Contourlet transform is because of its adeptness to capture visual geometrical structures and anisotropy. Further, weighted PCA assists in reducing the dimensionality of the source images as well as helps in better selection of principal components. Maximum and minimum fusion rules are then applied to fuse the decomposed coefficients. Image quality assessment (IQA) is carried out using standard fusion metrics quantitatively to assess the fused image both in terms of information content as well as quality of reconstruction. Simulation results with the proposed fusion method depict an effective fusion response in comparison to other state-of-art approaches.

Keywords Multimodal · Weighted PCA · Contourlet transform · CT/MRI fusion

Aisha Moin (✉) · Vikrant Bhateja · Anuja Srivastava
Department of Electronics and Communication Engineering,
Shri Ramswaroop Memorial Group of Professional Colleges (SRMGPC),
Lucknow 227105, UP, India
e-mail: aishamn04@gmail.com

Vikrant Bhateja
e-mail: bhateja.vikrant@gmail.com

Anuja Srivastava
e-mail: anuja.srivastava009@gmail.com

1   Introduction

1.1   *Multimodal Medical Image Fusion*

The extensive advancement in technology has led to the easy availability of variants of imaging sensors in military and civilian applications such as battlefield surveillance, healthcare purposes, traffic control, and security monitoring [1]. But the obtained sensor response is often complementary and superfluous in nature. Image fusion thus aims to integrate this complementary and disparate data in order to improvise upon the information apparent in the individual source images, as well as to increase the reliability of interpretation [2]. During the past two decades medical image processing has enticed researchers from all over the world as it aims at improving the precision and performance of computer-aided diagnosis. Using medical images from sophisticated modalities like CT scan, MRI, positron emission tomography (PET), single photon emission tomography (SPECT), etc., has already enhanced the accuracy in biomedical analysis for the purpose of clinical diagnosis. But medical images are often superimposed by noises during acquisition or transmission. Thus, image pre-filtering is necessary to suppress the erroneous intensity fluctuations caused due to imperfection of imaging devices or transmission channels [3–5]. Medical image fusion facilitates physicians to extract diagnostic features from assorted modalities that may not be thoroughly visible in any of the individual source images. For example, the CT scan shows the dense structures like bones and implants with less distortion, but it cannot perceive details regarding soft tissues. On the other hand, MRI provides the pathological soft tissue information, but it restrains to support the bone information. This emphasizes upon the need of multimodal fusion of CT and MRI scans to a single composite image to provide the radiologists with all the necessary diagnostic information at a single glance. Here, image fusion algorithms can be cataloged into three levels: pixel, feature, and decision level. Among them, pixel-level fusion framework is the most extensively used method due to an advantage of easy implementation and better computational efficiency. Moreover, the latter two methods are based on a compromise between spectral consistencies and desired spatial enhancements.

1.2   *Related Works*

In the past literature, assortments of algorithms have been developed for effective fusion of multimodal images [6, 7]. Owing to the simplicity of implementation, multiscale transform-based image fusion techniques are amongst one of the most extensively used transformation methods. Among these the prominent ones are discrete wavelet transform (DWT) [8–11], Ridgelet transform [12], Curvelet transform [13, 14], and Contourlet transform [15]. DWT is one of the most conventional and widely used transformation methods as it is competent in dealing with

1-D point wise smooth signals and point singularities. But it is optimally inefficient in capturing sharp transitions such as edges and textured regions. Hence, the output image contains high spatial distortion leading to unsatisfactory fusion results [11]. In order to rectify these limitations of DWT, ridgelet transform was thus introduced. It poses to be a powerful instrument in capturing the mono-dimensional singularities [16], but it is optimally inefficient in representing curve singularities in the fused image [12]. Curvelet transform allows better optimal sparse representation of objects with curve singularities and can efficiently represent curved objects but does not provide multiresolution representation of the geometry. Contourlet transform is a true 2-D sparse representation of signals and it can efficiently capture 2D geometrical structures in visual information. Contourlet has been deployed for fusion as improvement over wavelets; as it offers better visual geometrical structures and anisotropy [17]. In this paper, an improved multimodal pixel-level fusion methodology based on a variant of PCA is proposed in Contourlet domain. The source images are decomposed using the proposed methodology employing weighted PCA on the coefficients obtained. Inverse Contourlet transform is then being applied to get the final fused image. Both qualitative and quantitative analyses have been carried out to validate the performance of the proposed fusion methodology. Simulations have been carried out on different sets of multimodal medical images, namely CT and MR-T1. The rest of the paper is organized as follows. The proposed fusion methodology is detailed in Sect. 2. Further, the simulation results and discussions are presented in Sect. 3, whereas the conclusions are drawn in Sect. 4.

2 Proposed Fusion Methodology in Contourlet Domain

The processing of image fusion has been carried out in this work in Contourlet domain owing to the aforesaid benefits of this over other transforms. But Contourlet requires directional filtering as it captures limited directional information. To improvise upon this constraint, Contourlet transform has been hybridized with the modified version of Principle Component Analysis. PCA not only counters the limited directionality limitation of Contourlet transform but also enhances the fusion of medical images [1, 2, 18, 19]. The source images from the diverse modalities are preprocessed first. It is assumed that the source images are free from any noise and are correctly registered. The medical images are then decomposed into approximation and detailed coefficients using Contourlet transform.

2.1 Decomposition into the Subbands Coefficients

The Contourlet transform consists of two steps which are the subband decomposition and the directional transform. The subband decomposition is carried out

using Laplacian pyramid (LP), which generates a low-pass version of the original image, resembling a pyramidal structure. Alternatively, directional filter bank (DFB) is applied to capture the high frequency of the source image which represents directionality. As, DFB alone cannot provide sparse representation of the images; it is combined with LP (a multiscale decomposition). The end result of this combination is pyramidal directional filter bank (PDFB), which decomposes images into multiscale directional subbands. In Contourlet transform, first multiscale decomposition is achieved by the Laplacian pyramid, and then a directional filter bank is applied to each band pass channel [1, 20].

2.2 Dimensionality Reduction Using Weighted PCA

Medical images generally contain high dimensionality or redundancy. Hence, PCA [11], being an orthogonal transform helps to reduce the redundancy present in CT and MRI source images. The subspace modeled by PCA captures the maximum variability in the data, and can be viewed as modeling the covariance structure of the data. Traditionally, PCA algorithm may select all the principal components from the same region of the image. This drawback is countered in the modified PCA methodology wherein the images are first divided into two static window blocks; this serves to improve upon the fusion methodology. The approximation coefficients are divided into two blocks each for the application of weighted PCA. The steps involved are shown in Algorithm 1. The weights deployed in the weighted PCA are adaptive in nature and are determined using Eqs. (1) and (2) respectively [19, 21, 22].

$$wt\_1 = \frac{entropy(d_i) \times std(image\_1)}{std(image\_1) \ + \ std(d)} \tag{1}$$

$$wt\_2 = \frac{entropy(d_i) \ \times \ std(image\_2)}{std(image\_2) \ + \ std(d)} \tag{2}$$

where, d_i denotes the ith block of approximation matrix and std() refer to standard deviation. Image_1 and image_2 denote the CT and MR images respectively. The feature vectors obtained as a result of weighted PCA are then fused using the maximum fusion rule. The detailed coefficients obtained are then fused using the minimum fusion rule [19]. Finally, reconstruction is carried out using the inverse Contourlet transform.

Algorithm 1 Procedural steps for weighted PCA algorithm

	BEGIN
Step 1:	*Input*: Approximation and Detail coefficients (of both modalities)
Step 2:	*Compute*: Block subdivisions of approximation coefficient matrices.
Step 3:	*Compute*: Mean adjusted Matrices.

(continued)

(continued)

Step 4:	*Process*: Diagonal elements of the covariance matrix.
Step 5:	*Compute*: Eigen vectors and Eigen Values of covariance matrix.
Step 6:	*Process*: Higher significance column vectors for each block of coefficient matrix.
Step 7:	*Compute*: Feature vectors.
Step 8:	*Compute*: Multiplication of feature vector with the weights in Eqs. (1) and (2).
Step 9:	*Process*: Fusion of the obtained feature matrices.
Step 10:	*Process*: Fusion of the detailed coefficients.
Step 11:	*Compute*: Coefficients in cell vector format.
	END

3 Results and Discussions

3.1 *Fusion Metrics for IQA*

The main requirement of image fusion is to preserve all the useful information and discard any distortion present in the fused image. Performance measures are essentially used to measure the fruitfulness of the fusion performed and compare the results from different algorithms [23, 24]. Relevant researches show that any single quality measure cannot be just sufficient to quantify the performance of approaches in consistency with human visual perception. Therefore, IQA metrics to evaluate the effectiveness of the fusion algorithms have been deployed in this work. These are Entropy (E), Edge Strength ($Q^{AB/F}$) and SSIM [14, 25]. Higher values of each of these indices demonstrate the effectives of the fusion.

3.2 *Simulation Results*

The test images in the present work include three sets of CT scan and MR images (namely Test Image 1, Test Image 2, and Test Image 3), downloaded from the Brain Atlas [26]. The foremost stage of the decomposition requires preliminary analysis of decomposition levels. In order to minimize complexity, the decomposition level of Contourlet is set to one. This is followed by application of Weighted PCA as in Algorithm-1 using weights of Eqs. (1) and (2) respectively; followed by Min–Max fusion rule [19] and then reconstruction using inverse Contourlet transform. Figure 1 shows the fusion response on three test image sets. It can be depicted in all the three test images that both the bones (white portion of CT scan) as well as soft tissues (dendrites like structures in MR-T2) are visible in the composite fused images. It can be therefore visualized that the fusion response in Fig. 1 effectively demonstrates the features of both the modalities.
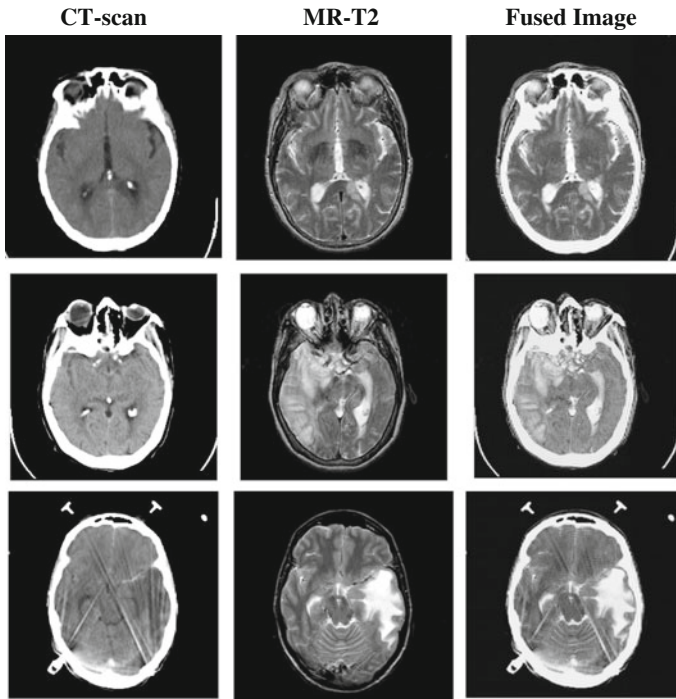
Fig. 1 Multimodal fusion response using proposed method on different combinations of CT scan and MR-T2 input images

Table 1 IQA of proposed fusion method using different metrics

Data set	E	SSIM	$Q^{AB/F}$
Image set 1	5.7389	0.9994	0.6276
Image set 2	5.8711	0.9991	0.6246
Image set 3	5.5885	0.9992	0.6337

The quality of the proposed fusion response is evident from the high values of the IQA metrics shown in Table 1. The high entropy values indicate that high amount of information content is present in the fused images. On the other hand, higher values of SSIM indicate the preservation of luminance, contrast and structural content. Above all, the amount of edge preservation in the fused images determines the major quality aspect of image fusion. Thus, higher value of $Q^{AB/F}$ indicates higher degree of edge preservation.

For the purpose of comparison and validation of fusion response; obtained results are also compared with those of state-of-art approaches which include DWT [27], Ridgelet [21], Curvelet [27], and Contourlet [27]. The results of comparison are shown in Fig. 2 and the computed values of IQA measure are enlisted under Table 2. The visual comparison as well as the high values of fusion metrics portrays
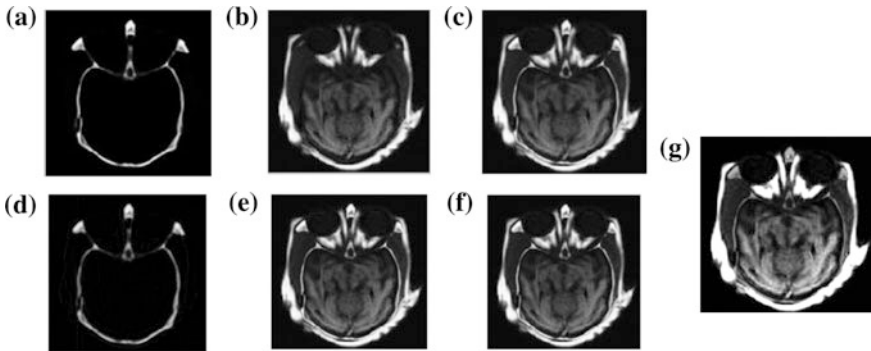
Fig. 2 A comparative analysis of fusion response of the proposed method along with other state-of-art approaches. **a** Input CT scan image. **b** Input MR-T2 image. **c** DWT. **d** Ridgelet. **e** Curvelet. **f** Contourlet. **g** Proposed method

Table 2 IQA of proposed fusion method compared with other state-of-art approaches

Approaches	E	SSIM	$Q^{AB/F}$
Wavelet [27]	6.7816	0.5901	0.5943
Ridgelet [21]	3.6803	0.5915	0.3970
Curvelet [27]	6.7938	0.5895	0.8060
Contourlet [27]	6.8079	0.9968	0.4861
Proposed method	6.3364	0.9957	0.6511

the efficiency of proposed methodology in capturing both soft and hard tissue information during fusion.

4 Conclusion

A weighted PCA-based image fusion method in the Contourlet domain has been presented in this paper. The results of proposed fusion method are superior to other methods as it counters the limited directionality constraint of Contourlet transform as well as assists in dimensionality reduction. The fused images thus obtained are more informative than individual source images as it depicts both soft tissues and the hard tissues to the finest. The aforesaid quality attributes have been validated by high values of fusion metrics; the proposed fusion method is therefore better than the other state-of-the-art fusion approaches. The fused images are anatomically consistent and have better spatial resolution; and are therefore beneficial for clinical applications. The future works would include further contrast improvement of the region of interest in the fused image by implementing enhancement approaches based on nonlinear [28–34] and Unsharp Mask (UM) filters [35].

References

1. Li, S., Yang, B., Hu, J.: Performance Comparison of Different Multi-resolution Transforms for Image Fusion. Information Fusion. 12(2), 74–84 (2011).
2. Singh, S., Gupta, D., Anand, R.S., Kumar, V.: Nonsubsampled Shearlet based CT and MR Medical Image Fusion using Biologically Inspired Spiking Neural Network. Biomedical Signal Processing and Control. 18, 91–101 (2015).
3. Bhateja, V., Tiwari, H., Srivastava, A.: A Non-Local Means Filtering Algorithm for Restoration of Rician Distributed MRI. Annual Convention of the Computer Society of India. 2, 1–8 (2014).
4. Srivastava, A., Bhateja, V., Tiwari, H., Satapathy, S. C.: Restoration Algorithm for Gaussian Corrupted MRI using Non-Local Average Filtering. In: 2nd International Conference on Information Systems Design and Intelligent Applications. 2(2), 831–840 (2015).
5. Tiwari, H., Bhateja, V., Srivastava, A.: Estimation Based Non-Local Approach for Pre-Processing of MRI. In: 2nd IEEE International Conference on Computing for Sustainable Global Development. 1622–1626 (2015).
6. Xu, Z.: Medical Image Fusion Using Multi-level Local Extrema. Information Fusion. 19, 38–48 (2014).
7. Bhatnagar, G., Wu, Q. M. J., Liu, Z.: A New Contrast based Multimodal Medical Image Fusion Framework. Neurocomputing. 157, 143–152 (2015).
8. Liu, Z., Yin, H., Chai, Y., Yang, S. X.: A Novel Approach for Multimodal Medical Image Fusion. Expert Systems with Applications. 41(16), 7425–7435 (2014).
9. Shrivastava, A., Alankrita, Raj, A., Bhateja, V.: Combination of Wavelet Transform and Morphological Filtering for Enhancement of Magnetic Resonance Images. In: International Conference on Digital Information Processing and Communications, Part-I. 460–474 (2011).
10. Bhateja, V.,Verma, R.,Mehrotra, R., Urooj, S., Lay-Ekuakille, A., Verma, V., D.: A Composite Wavelets and Morphology Approach for ECG Noise Filtering. In: 5th International Conference on Pattern Recognition and Machine Intelligence. 8251, 361–366. Springer (2013).
11. Krishn, A., Bhateja, V., Himanshi, Sahu, A.: Medical Image Fusion using Combination of PCA and Wavelet Analysis. In: 3rd IEEE International Conference on Advances in Computing, Communications and Informatics. 986–991 (2014).
12. Do, M. N., Vetterli, M.: The Finite Ridgelet Transform for Image Representation. IEEE Transactions on Image Processing. 12, 16–28 (2003).
13. Amini, N., Fatemizadeh, E., Behnam, H.: MRI and PET Image Fusion by Curvelet Transform. J. Advances in Computer Reasearch. 5, 23–30 (2014).
14. Himanshi, Bhateja, V., Krishn, Sahu, A.: Medical Image fusion in Curvelet Domain Employing PCA and Maximum Selection Rule. In: 2nd International Conference on Computers and Communication Technologies. Springer (2015).
15. Qiguang, M., Baoshu, W.: A Novel Image Fusion Method Using Contourlet Transform. In: International Conference on Circuits and System Proceedings in Communications. 1, 548–552 (2006).
16. Bhateja, V., Krishn, A., Himanshi, Sahu, A.: Medical Image Fusion in Wavelet and Ridgelet Domain: A Comparative Evaluation. International Journals of Rough Sets and Data Analysis. 2, 78–91 (2015).
17. Po, DY, D., Do, M. N.: Directional Multiscale Modeling of Images using the Contourlet Transform. IEEE Transactions on Image Processing. 15, 1610–1620 (2006).
18. Bhatnagar, G., Wu, Q. M. J., Liu, Z.: Directive Contrast based Multimodal Medical Image Fusion in NSCT Domain. IEEE Transactions on Multimedia. 15, 1014–1024 (2013).
19. Bhateja, V., Patel, H., Krishn, A., Sahu, A., Lay-Ekuakille, A.: Multimodal Medical Image Sensor Fusion Framework using Cascade of Wavelet and Contourlet Transform Domains. IEEE Sensors Journal. 1–8 (2015).

20. Sahu, A., Bhateja, V., Krishn, A., Himanshi.: Medical Image Fusion with Laplacian Pyramids. An Improved Medical Image Fusion Approach Using PCA and Complex Wavelets. In: IEEE International Conference on Medical Imaging, m-Health & Emerging Communication Systems. 448–453 (2014).
21. Krishn, A., Bhateja, V., Himanshi, Sahu, A.: PCA based Medical Image Fusion in Ridgelet Domain. In: 3rd International Conference on Frontiers in Intelligent Computing Theory and Applications. Springer 328, 475–482 (2014).
22. Himanshi, Bhateja, V., Krishn, A., Sahu, A.: An improved Medical Image Fusion Approach Using PCA and Complex Wavelets. In: IEEE International Conference on Medical Imaging, m-Health & Emerging Communication Systems. 442–447 (2014).
23. Gupta, P., Tripathi, N., Bhateja, V.: Multiple Distortion Pooling Image Quality Assessment. International Journal on Convergence Computing. 1(1), 60–72 (2013).
24. Srivastava, H., Mishra, A., Bhateja, V.: Non-Linear Quality Evaluation Index for Mammograms. In: 3rd Students Conference on Engineering and Systems. 269–273 (2013).
25. Gupta, P., Srivastava, P., Bharadwaj, S., Bhateja, V.: A Modified PSNR Metric based on HVS for Quality Assessment of Color Images. In: IEEE International Conference on Communication and Industrial Application. (23), 96–99 (2011).
26. Brain Atlas available online at http://www.med.harvard.edu/AANLIB/home.html, last accessed on 15th July, 2015.
27. Bindu, C. H., Prasad, K. S.: Performance Analysis of Multi Source Fused Medical Images Using Multiresolution Transforms. International Journal of Advanced Computer Science and Applications. 3, 54–62 (2012).
28. Bhateja, V., Urooj, S., Misra, M.: Technical Advancements to Mobile Mammography using Non-Linear Polynomial Filters and IEEE 21451-1 Information Model. IEEE Sensors Journal. 15(5) Spl. Iss. on- Advancing Standards for Smart Transducer Interfaces, 2559–566 (2015).
29. Bhateja, V., Misra, M., Urooj, S., Lay-Ekuakille, A.: A Robust Polynomial Filtering Framework for Mammographic Image Enhancement from Biomedical Sensors. IEEE Sensors Journal. 13(11), 4147–4156 (2013).
30. Jain, A., Singh, S., Bhateja, V.: A Robust Approach for Denoising and Enhancement of Mammographic Breast Masses. International Journal on Convergence Computing, Inderscience Publishers. 1(1) 38–49 (2013).
31. Pandey, A., Yadav, A., Bhateja, V.: Contrast Improvement of Mammographic Masses Using Adaptive Volterra Filter. In: 4th International Conference on Signal and Image Processing. 2, 583–593 Springer (2012).
32. Raj, A., Alankrita, Shrivastava, A., Bhateja, V.: Computer Aided Detection of Brain Tumor in MR Images. International Journal on Engineering and Technology. 3, 523–532 (2011).
33. Pandey, A., Yadav, A., Bhateja, V.: Design of New Volterra Filter for Mammogram Enhancement. In: International Conference on Frontiers in Intelligent Computing Theory and Applications. 199, 143–151 Springer (2012).
34. Bhateja, V., Devi, S.: An Improved Non-Linear Transformation Function for Enhancement of Mammographic Breast Masses. In: 3rd IEEE International Conference on Electronics & Computer Technology. 5, 341–346 (2011).
35. Do, M. N., Vetterli, M.: The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. IEEE Transactions on Image Processing. 14, 2091–2106 (2005).

Design Analysis of an *n*-Bit LFSR-Based Generic Stream Cipher and Its Implementation Discussion on Hardware and Software Platforms

Trishla Shah and Darshana Upadhyay

Abstract Pseudorandom numbers are at the core of any network security application. Also, security of satellite phones and cellular phones depends heavily on the pseudorandom numbers generated. In the network security domain, its use is particularly in key generation, re-keying, authentication, smart-phone security, etc. Also, current research shows that satellite-based telephony system, having GMR-1 and GMR-2 algorithms for secret key generation is prone to attacks. The algorithm A5/1 used in GSM technology is also cryptographically poor. Hence generation of strong sets of pseudorandom number is needed. These random numbers are produced through a pseudorandom number generator (PRNG). This generator in general terms is called a Cipher. Hence, if there is a flaw or the PRNG produces predictable sets of random numbers, then the entire application would be prone to attacks. Therefore, development of a generic framework for generating strong sets of pseudorandom numbers is proposed. The proposal aims to build an in-general framework and a unified model for enhanced security specifically for LFSR-based stream ciphers. The proposed generic model uses results from the above case study. For the hardware deployment, Spartan-6 FPGA toolkit is used and for the software part a parallel computing platform namely CUDA is used. The model is aimed at development of a framework which generates strong sets of pseudorandom numbers for its use in various network security, satellite and cellular applications.

Trishla Shah (✉)
Department of Computer Science and Engineering, B.H. Gardi College
of Engineering and Technology, Rajkot 361162, India
e-mail: tpshah@gardividyapith.ac.in

Darshana Upadhyay
Institute of Technology, Nirma University, Ahmedabad 382481, India
e-mail: darshana.upadhyay@nirmauni.ac.in

607

1 Introduction

1.1 Basic Concept

In current trends of security spectrum, LFSR-based stream ciphers form the backbone of critical security applications like military cryptography, encoding and higher order encryption mechanism. Recent research has open folds of attacks where these are most occurring like eavesdropping, snooping, masquerading, impersonation, and in the specific wireless network and telecommunications domain poor security mechanisms are explored. [1] Also, very critical applications like encryption scheme in military using GMR-1 and GMR-2 standards are prone to attacks. These LFSR-based stream ciphers currently are implemented on both hardware (A5/1, A5/2, KASUMA, E0, MICKEY, GRAIN, SNOW, FISH) and software (HC-256, Rabbit, Salsa20, SOSEMANUK) [2] platforms. These ciphers have been detected to be prone to various network attacks like dynamic cube attack, basic correlation attack, refinement attack, guess-and-determine attack, linear approximation attack, algebraic attack, Berlekamp–Massey attack, fast time memory trade-off attack which requires some pre-computation [3]. This survey demands a strong need for a co-simulation of hardware and software to resist these most common LFSR-based attacks.

1.2 Challenges

Hardware implementation for GSM stream cipher has already been implemented, under a particular segment of mobile communication. This project has been carried out under an Indo-Canadian collaboration—"Shastri Project and Research Grant," with University of Dalhousie, Canada. Progress was made in all four folds, namely —vulnerability assessment, protocol design, implementation on both software and hardware, and evaluation. The project has explored many research problems in this area, which opened up excellent research opportunities. Few of the challenges are as below.

Many stream ciphers have been designed [4] for the generation of a strong set of pseudorandom numbers, but certain limitations have been examined like: (i) In designing of hardware ciphers, the computational complexity over software performance decreases (ii) Very few ciphers have been designed, working for network security applications in both hardware and software domains. (iii) The software implementation is mostly done on sequential basis, i.e., CPU, thus increasing complexities of overhead and time. (iv) Ciphers compatible for generating good pseudo random series on a generic platform for diverse applications has not yet been designed.

This paper thus focuses on design implementation of an n-bit LFSR-based stream cipher and its implementation on hardware and software platforms.

2 Design Analysis

2.1 n-*Bit LFSR-Based Cipher Design*

The entire cipher is based on a parallel approach using n-LFSR's and different feedback polynomials.

3 LFSR Counter

The LFSR counter has an initial seed value of 7-bit. Here the LFSR counter is kept static, because there are total 119 LFSRs in the LFSR pool with largest LFSR of length 128 bits. Therefore, a total of 7-bit LFSR producing a combination of 27, i.e., 128 bits are used. The feedback polynomials to be used in 7-bit LFSR counter are as shown in following Table 1 (Fig. 1):

In one single clock cycle, it will generate one binary combination out of $2^7 - 1$ possible combinations. The decimal of the corresponding binary will be number of LFSRs to be chosen. For example, in first clock pulse it generates 1,000,000, i.e., 64 in decimal. Hence, it communicates with LFSR generator that a selection of 64 LFSRs have to be done. This makes the LFSR generator a 6-bit LFSR. After the counter has generated 27 possible combinations, its periodicity state will be gained. Hereafter, the feedback polynomial needs to be changed. Once all feedback polynomials have been exhausted, the seed value will change. The process will be repeated again.

Table 1 7-bit LFSR-based feedback polynomial

Primitive polynomial	Period
$x^7 + x^6 + 1$	127
$x^7 + x^6 + 1$	127
$x^7 + x^1 + 1$	127
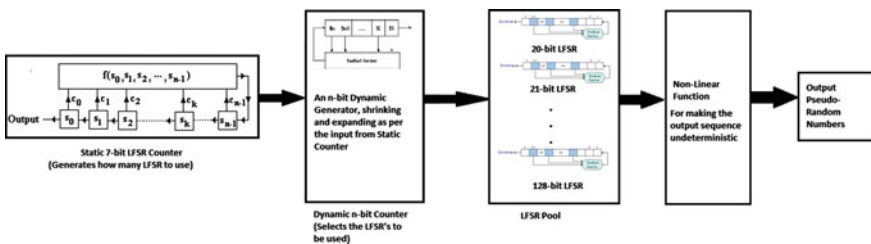
Fig. 1 *n*-bit LFSR-based cipher design 2

4   LFSR Generator

It generates the serial number of LFSR to be used from LFSR pool. The binary value generated from counter will be an input seed value to generator. Hence the generator will expand and contract based on the output generated from counter. To each n input, 2^n combinations will be generated for selection from LFSR pool. Since LFSRs has only 119 LFSRs (20 bit to 128 bit) in it, generation of decimal number 120 or more will make a circular shift to 0. If the pointed LFSR is already selected it will be compared by a comparator and incremented by 1. Hence, the counter will select number of LFSRs from LFSR pool.

5   LFSR Pool

It contains 119 LFSR ranging from 20 bit-128 bit. The seed value to every LFSR will be a secret key. Every n-bit LFSR stops generating the bit sequence after it has generated n random numbers, i.e., a 20-bit LFSR must stop generating numbers after it has generated 20 bits. After all LFSRs have generated bits, the initial counter will be clocked again. If a particular LFSR is chosen again during second iteration, it must be used with a different feedback polynomial. This design of LFSR pool makes it resistant toward most common LFSR attack, i.e., Berlekamp–Massey attack.

5.1   *Mathematical Model*

After understanding the basic design and security provided by the cipher, it is important to understand the internal mechanism and operation of the cipher.

A 7-bit counter produces 128 bits which are converted into a decimal number. Here an initial key value is decided up on a session key. The bits are produced until periodicity, once a periodicity is reached the feedback polynomials are changed. This reduces the possibility of repetition of decimal values. These decimal values suggest the number of LFSR's to be selected for the output. The feedback polynomials used for the counter is shown below. In second module, a mathematical function as per 5-bit, 6-bit, and 7-bit LFSR has been designed. A combination of 5-bit, 6-bit, and 7-bit have been designed. On the basis of the output produced, values are selected from LFSR pool. For example, for every decimal value produced, a particular combination of 5-bit, 6-bit, and 7-bit is chosen. A 5 * 5 matrix is formed, with decimal values in it. A median of it is chosen, as first decimal output. Similarly, from every row, a median is selected. And outputs are produced. A 6 * 6

Table 2 Primitive polynomials used in LFSR pool

No. of bits	Primitive polynomial
5-bit	(1) 5,3,2,1,0
	(2) 5,3,0
	(3) 5,4,3,1,0
	(4) 5,4,3,2,0
	(5) 5,4,2,1,0
6-bit	(1) 6,5,0
	(2) 6,1,0
	(3) 6,4,3,1,0
	(4) 6,5,3,2,0
	(5) 6,5,2,1,0
7-bit	(1) 7,3,0
	(2) 7,6,0
	(3) 7,6,5,4,0
	(4) 7,5,2,1,0
	(5) 7,5,4,3,0

and 7 * 7 matrix is formed similarly. Here median is chosen, because it is more unpredictable rather than mean. The feedback polynomials for it are as follows: [5]

5-bit: ((x-3)/25) mod 5 + 1
6-bit: ((x-3)/5) mod 5 + 1
7-bit: (x-3) mod 5 + 1

The feedback polynomials used in LFSR pool are as shown in following table:
In the third module, all LFSR's ranging from 20-bit to 128-bit are stored. Depending on the output from, second module, i.e., the decimals produced, LFSR's are selected and output bits are generated. It is as shown in figure. The feedbacks for these are as shown in Table 2. After the decimal values have been generated, the LFSR's to be selected are chosen from third module and binary bits are generated. The binary bits generated from every LFSR is in order of (length of LFSR * 2) − 3. This formula is applied for protection against Berlekamp–Massey Attack. All output bits are XORed, and a final output of 256 bits is produced (Table 3).

6 Experimental Basis

After the entire cipher was designed and mathematically verified, it is important to design an implementation flow for it.

The entire cipher is to be designed in mainly two levels:

Software Implementation—The main idea is to make a parallel implementation of the cipher on Nvidia's Parallel Computing Platform CUDA. The entire implementation idea is shown ahead in following subsection.

Table 3 Feedback polynomials for LFSR pool

N-bit LFSR	Feedback 1 and 3		Feedback 2 and 4
34-bit	1)34,4,3,1,0		2)34,27,2,1,0
	3)34,31,30,26,0		4)34,27,0
35-bit	1)35,2,0		2)35,34,28,27,0
	3)35,33,28,25,23,22,21,14,13,11,10,9,7,5,0		4)35,33,0
36-bit	1)36,5,4,2,0		2)36,35,29,28,0
	3)36,25,0		4)36,9,0
37-bit	1)37,6,4,1,0		2)37,36,33,31,0
	3)37,5,4,3,2,1,0		4)37,30,26,25,23,19,16,13,11,9,7,3,0
38-bit	1)38,6,5,1,0		2)38,37,33,32,0
	3)38,36,28,25,24,22,21,19,16,13,8,6,5,3,2,1,0		4)38,37,36,35,34,29,26,20,19,17,13,12,3,0
39-bit	1)39,4,0		2)39,38,35,32,0
	3)39,8,0		4)39,35,0
40-bit	1)40,5,4,3,0		2)40,38,21,19,0
	3)40,37,36,35,0		4)40,35,31,26,24,22,19,16,12,11,10,9,7,2,0
41-bit	1)41,3,0		2)41,40,39,38,0

N-bit LFSR	Feedback 1	Feedback 2	Feedback 3
45-bit	45,4,3,1,0	45,44,42,41,0	45,42,40,35,33,30,26,21,18,16,12,10,4,3,0
46-bit	46,1,0	46,45,26,25,0	46,40,39,38,0
47-bit	47,5,0	47,42,0	47,46,43,42,0
48-bit	48,5,3,2,0	48,47,21,20,0	48,44,41,39,0
49-bit	49,6,5,4,0	49,40,0	49,45,44,43,0
50-bit	50,4,3,2,0	50,49,24,23,0	50,48,47,46,0
51-bit	51,6,3,1,0	51,50,48,45,0	51,50,36,35,0
52-bit	52,3,0	52,49,0	52,51,49,46,0
53-bit	53,6,2,1,0	53,52,51,47,0	53,52,38,37,0
54-bit	54,8,6,3,0	54,51,48,46,0	54,53,18,17,0
55-bit	55,6,2,1,0	55,31,0	55,54,53,49,0
56-bit	56,7,4,2,0	56,54,52,49,0	56,55,35,34,0
57-bit	57,4,0	57,50,0	57,55,54,52,0
58-bit	58,6,5,1,0	58,39,0	58,57,53,52,0

N-bit LFSR	Feedback 1	Feedback 2	Feedback 3
99-bit		99,6,3,1,0	99,97,54,52,0
100-bit		100,6,5,2,0	100,63,0
101-bit		101,7,6,1,0	101,100,95,94,0
102-bit		102,6,5,3,0	102,101,36,35,0
103-bit		103,9,0	103,94,0
104-bit		104,4,3,1,0	104,103,94,93,0
105-bit		105,4,0	105,89,0
106-bit		106,6,5,1,0	106,91,0
107-bit		107,9,7,4,0	107,105,44,42,0
108-bit		108,6,4,1,0	108,77,0
109-bit		109,5,4,2,0	109,108,103,102,0
110-bit		110,6,4,1,0	110,77,0
111-bit		111,7,4,2,0	111,101,0

Table 3 (continued)

81-bit	81,4,0	81,77,0
82-bit	82,9,6,4,0	82,79,47,44,0
83-bit	83,7,4,2,0	83,82,38,37,0
84-bit	84,5,0	84,71,0
85-bit	85,8,2,1,0	85,84,58,57,0
86-bit	86,6,5,2,0	86,85,74,73,0
87-bit	87,7,5,1,0	87,74,0
88-bit	88,11,9,8,0	88,87,17,16,0
89-bit	89,6,5,3,0	89,51,0
90-bit	90,5,3,2,0	90,89,72,71,0
91-bit	91,8,5,1,0	91,90,8,7,0
92-bit	92,6,5,2,0	92,91,80,79,0
93-bit	93,2,0	93,91,0
94-bit	94,6,5,1,0	94,73,0
95-bit	95,11,0	95,84,0
96-bit	96,10,9,6,0	96,94,49,47,0
97-bit	97,6,0	97,91,0
98-bit	98,7,4,3,0	98,87,0
99-bit	99,6,3,1,0	99,97,54,52,0

N-bit LFSR	Feedback 1	Feedback 2
66-bit	66,3,0	66,65,57,56,0
67-bit	67,5,2,1,0	67,66,58,57,0
68-bit	68,7,5,1,0	68,59,0
69-bit	69,6,5,2,0	69,67,42,40,0
70-bit	70,5,3,1,0	70,69,55,54,0
71-bit	71,5,3,1,0	71,65,0
72-bit	72,10,9,3,0	72,66,25,19,0
73-bit	73,4,3,2,0	73,48,0
74-bit	74,6,2,1,0	74,73,59,58,0
75-bit	75,6,3,1,0	75,74,65,64,0
76-bit	76,5,4,2,0	76,75,41,40,0
77-bit	77,6,5,2,0	77,76,47,46,0
78-bit	78,7,2,1,0	78,77,59,58,0
79-bit	79,4,3,2,0	79,70,0
80-bit	80,9,4,2,0	80,79,43,42,0

126-bit	1)126,7,4,2,0	1)126,124,122,119,0
	2)126,125,90,89,0	2)126,21,0
127-bit	1)127,1,0	1)127,126,0
	2)127,126,124,120,0	2)127,120,0
128-bit	1)128,7,2,1,0	1)128,127,126,121,0
	2)128,126,101,99,0	2)128,127,124,120,118,113,112,111,109,105,103,101,100,98,95,94,92,91,86,83,77,76,74,72,70,64,62,61,58,57,56,52,45,44,43,37,35,33,32,29,25,24,19,18,17,9,8,7,4,3,0
120-bit	1)120,4,3,1,0	1)120,118,114,111,0
	2)120,113,9,2,0	2)120,116,113,110,107,104,103,102,99,98,97,96,95,92,86,82,81,78,75,72,60,57,56,52,50,48,47,46,42,41,40,39,36,32,31,28,22,20,14,13,11,10,8,7,6,5,2,1,0
121-bit	1)121,8,5,1,0	1)121,120,116,113,0
	2)121,103,0	2)121,91,0
122-bit	1)122,6,2,1,0	1)122,121,120,116,0
	2)122,121,63,62,0	2)122,121,120,119,117,113,112,111,108,107,102,101,98,96,93,91,84,82,70,65,63,62,61,58,57,53,52,51,50,48,44,40,38,35,34,29,28,26,25,23,21,20,17,16,13,12,11,9,8,7,5,1,0
123-bit	1)123,2,0	1)123,122,119,115,0
	2)123,121,0	2)123,122,121,120,119,115,113,109,107,105,100,97,95,94,93,92,89,88,85,84,80,73,72,67,66,65,62,59,58,51,49,46,45,43,38,35,29,28,24,18,17,16,11,10,8,5,0
124-bit	1)124,37,0	1)124,119,118,117,0
	2)124,87,0	2)124,105,0
125-bit	1)125,7,6,5,0	1)125,120,119,118,0
	2)125,124,18,17,0	2)125,123,122,121,120,118,113,112,110,102,99,98,97,96,94,92,89,88,86,85,79,76,75,74,73,72,71,69,68,67,65,60,58,56,47,43,42,37,33,31,30,24,19,18,15,13,9,8,7,3,2,1,0

128-bit to 112-bit

N-bit LFSR	Count:0 Feedbacks	Count:1 Feedbacks
112-bit	1)112,5,4,3,0	1)112,108,106,101,0
	2)112,110,69,67,0	2)112,108,105,104,103,101,94,93,91,90,88,87,85,83,82,81,74,73,71,70,69,67,64,61,59,58,54,53,49,48,42,41,40,37,33,28,26,22,21,13,10,9,3,2,0
113-bit	1)113,5,3,2,0	1)113,111,110,108,0
	2)113,104,0	2)113,98,0
114-bit	1)114,5,3,2,0	1)114,113,112,103,0
	2)114,113,33,32,0	2)114,112,108,107,104,103,99,98,94,92,87,86,78,77,75,74,71,67,64,61,60,59,58,57,55,48,44,41,39,38,37,34,32,29,25,23,18,15,13,12,11,9,7,6,5,4,3,2,0

Hardware Implementation—The entire cipher was designed at hardware level using VHDL programming language. The hardware implementation is as shown ahead.

6.1 Software Implementation

6.1.1 Implementing Sequentially Using C++

For checking the effectiveness of the parallel cipher against sequential, it was necessary first to build a sequential prototype of the cipher. Hence, first a sequential cipher was constructed in C++.

Following design method was adopted:

- Platform used for C++ is Visual Studio 2010 Express Version and Codeblocks.
- The entire code was divided into four modules and were integrated at last.
- The code had 6000+ LOC.

6.1.2 Code Profiling

As per the research going on in most accurate C++ profiling, some brief results were searched and found. There are two types of profiling:

Invasive Profiling—Modifies the program (instrumentation of code) and inserts calls to functions that record data.
Non-Invasive Profiling—Statistic sampling of the program and uses a fixed interval that records instruction pointer at each sampling event.

Various profiling tools used: [6]
gprof   Mixture of Invasive and Statistical Profiling
gcov   Analyses coverage of program code
valgrind JIT-compiler/translator
oprofile Noninvasive and kernel module

6.1.3 Implementing in Parallel Using CUDA

CUDA (Compute unified device architecture) is a parallel computing platform and programming model created by NVIDIA and implemented by the graphics processing units (GPUs) that they produce [7]. It is the process of using graphics card

Table 4 Time and memory consumption in sequential programming

Iteration	Time	Memory	Improvements
1st	4.533	Single thread	Elimination of getch
2nd	3.685	Single thread	Function rather than for loop
3rd	0.236	Single thread	Time and space optimization
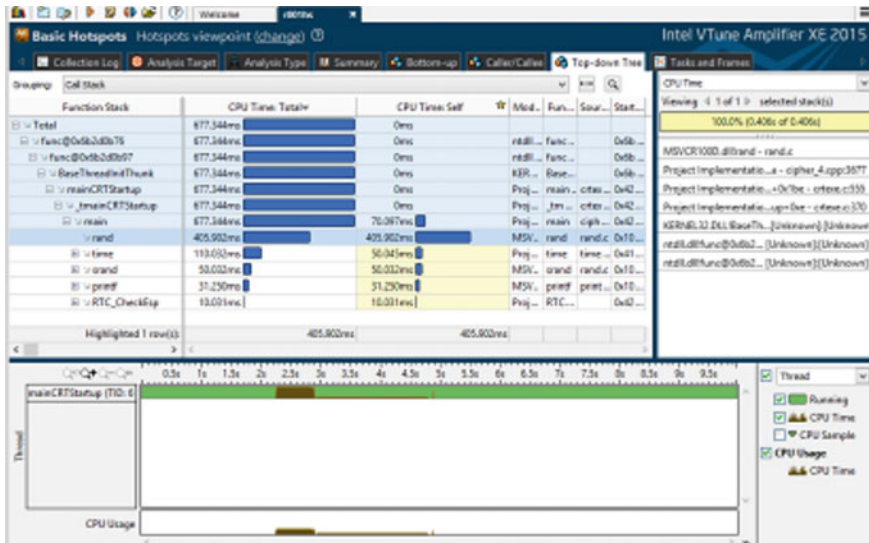
Fig. 2 Time and memory profiling in CUDA

for non-graphic processing unit. GPU's have evolved as a highly scalable processor for implementing parallel programming of large blocks of data [8].

The time and memory profiling for the code is as follows (Table 4, Fig. 2),

6.2 Hardware Implementation

RTL View is a register transfer level graphical representation of the design. It is generated at earlier stages of the synthesis process when technology mapping is still pending. The goal of this view is to be as close as possible to the original HDL code. The RTL design is independent of the targeted device and contains generic symbols, such as 58 adders, multipliers, counters, and basic gates (Figs. 3 and 4).
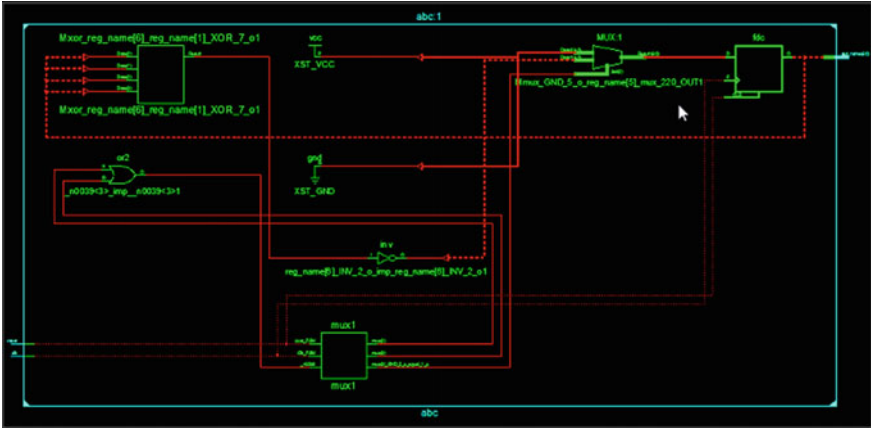
Fig. 3 RTL view
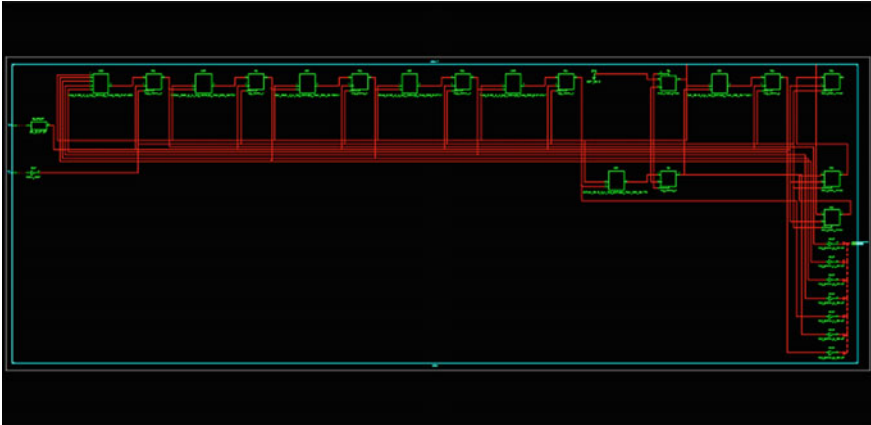
Fig. 4 Technology schematic

6.2.1 Implementation and Technology Used

Technology schematic for the post-synthesis netlist is generated after the optimization and technology targeting phase of the synthesis process [9]. Figure 6 shows a representation of the design in terms of logic elements optimized to the target Xilinx device in terms of LUTs, I/O buffers, carry logics, etc. The target Xilinx device used here is Spartan 6 XC6SLX45T in our case. Viewing this schematic allows us to see a technology-level representation of the HDL optimized for a specified Xilinx architecture [10].

On double clicking an LUT in the technology schematic gives an LUT dialog where the Schematic tab (Fig. 5) displays the under the hood logical schematic of
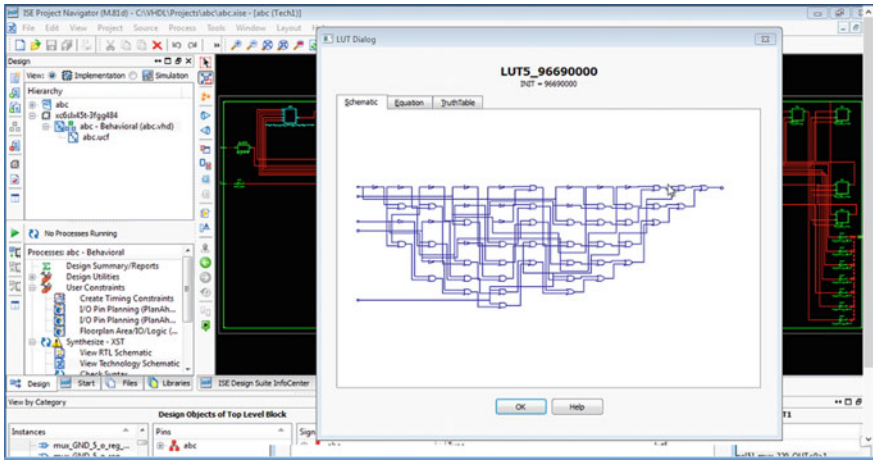
Fig. 5 Technology schematic
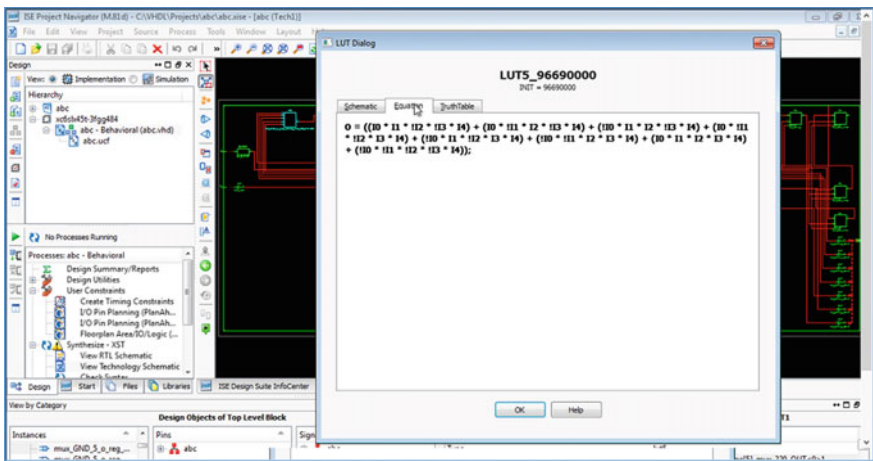
Fig. 6 Technology schematic

the LUT. It shows the LUT design using the AND gate, OR gate, and NOT gate. The Equation Tab (Fig. 6) displays the equation corresponding to the LUT in the sum-of-products form. The Truth Table Tab (Fig. 7) shows the corresponding truth table for the given LUT 59 Fig. 8: Technology schematic whereas the Karnaugh Map Tab (Fig. 5) displays the KMap for the LUT provided the number of inputs to the LUT is less and it is easily possible to generate a KMap. The above figure shows the snapshot of an LSIM simulator [11]. The objects selection screen on the top left allows us to select the objects to be viewed during the simulation. The console at the bottom uses the commands to perform the tasks. The required tasks can be done using the GUI or by providing appropriate commands to the console.
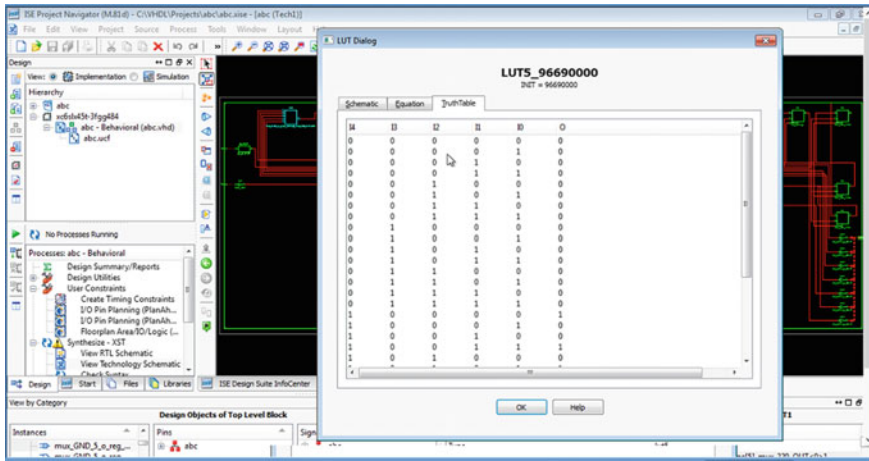
Fig. 7 Technology schematic
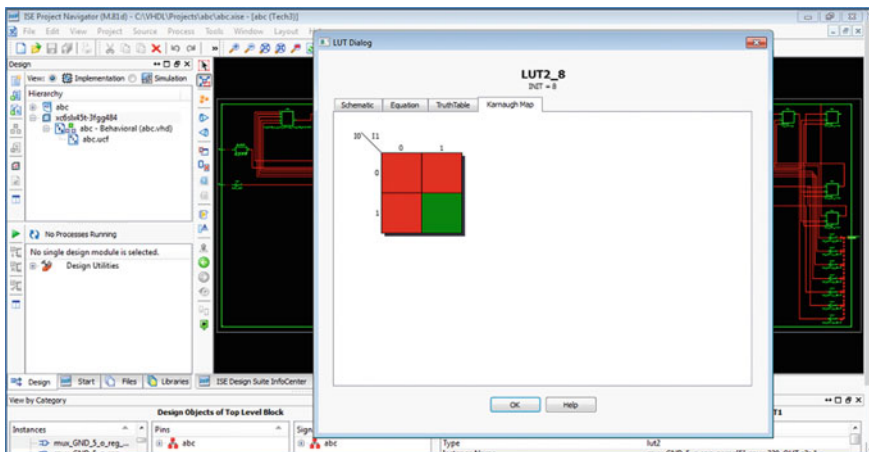
Fig. 8 Technology schematic

The screen on the topright shows the simulation of the selected objects. The Xilinx Design Suite in Fig. 6 also helps by creating a number of reports for the created project. It creates the report for Synthesis Phase, Translation Phase, timing statistics, power statistics, device utilization, and many more. All these reports help in understanding the whereabouts of the project, if actually dumped onto the hardware. It also helps to reduce any of the criteria according to the requirements. The synthesis tool also does the required optimization based on the user's requirement of power minimization or time minimization. Below mentioned are the reports of the partial project done using VHDL on the Xilinx Design Suite [12]. The above figure shows a part of the synthesis report where the general options for 60 Fig. 6:
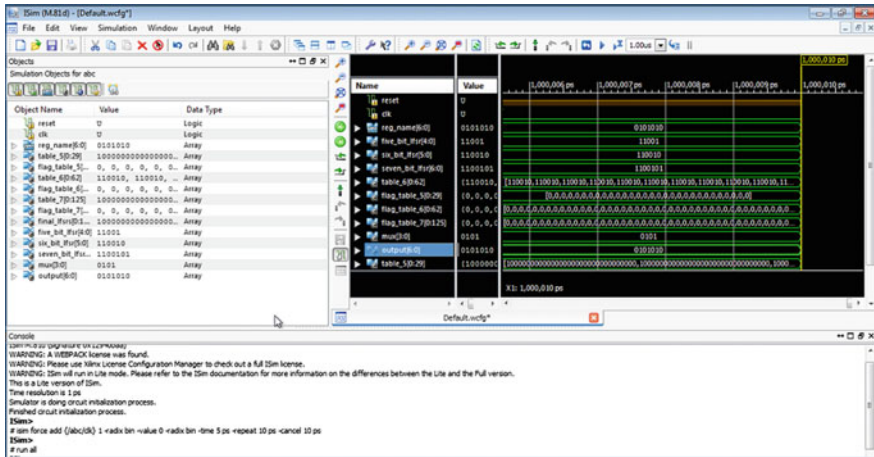
Fig. 9 Isim simulation

In technology schematic the given project are mentioned. This particular project aims at a higher speed and thus the optimization goal is speed. As the RTL output field is set to YES, the corresponding RTL view is created. The above figure is also a part of the synthesis report and it shows the VHDL synthesis which states the number of states, transitions, inputs, and outputs in the code and infers the corresponding required number of devices for carrying out the process. Here the macro-level for the device utilization is specified. The advanced HDL synthesis is also a part of the synthesis report. It shows the number of ip-ops, multiplexers, finite state machines, and XORs used. The device utilization summary of the synthesis report, as the name suggests, shows the utilization of the devices, the number of registers, LUTs, inputs, outputs, and buffers. It also mentions the total number of available devices and the percentage of devices used. The IOB properties is a part of the Map Report. The synthesis phase is followed by the implementation phase where four major tasks are done. Translate which merges the incoming source files and the constraints into a Xilinx design file; Map which fits the design file into the available resources on the target device; Place and Route which 61 Fig. 7: Technology schematic places and routes the design according to the timing constraints and programming file generation which generates a bit stream file which can be downloaded on the device. The slew rate shown in the figure is the maximum rate of change of output voltage per unit time. The place and route phase executes multiple phases of router. The router performs a process to find a solution that routes the design to completion and meets the timing constraints. The above figure shows the phases of the router wherein the design is routed in the first two phases but it still undergoes all the phases to improve the performance. The above figure shows the power consumption for the design. The supply voltage required, energy utilization by each device, total current, and all the other factors are displayed (Fig. 9).

7 Conclusion and Future Work

Toward the end of 1-year project titled "Development of an LFSR-based Framework for Cryptographically Secure PRNG using Parallel Computing to Enhance the Security of Network Applications" certain conclusions have been derived. Sequential Implementation of the cipher is working fine with average run time of 1.086 s. This time is very much fine for real-time network applications requiring at max 256 bits in a single transmission. Parallel implementation of the cipher on NVIDIA's Parallel Programming language, CUDA—Compute Unified Device Architecture is a very efficient way to divide the load and parallelize the given architecture. But it is a bit challenging process to do its parallelization, as the language is new and many platform architecture problems are there. Hence, a very precise optimization needs to be done. The entire cipher has been implemented on hardware using VHDL Spartan 6. A memory level optimization needs to be done.

References

1. M. Garland, S. Le Grand, J. Nickolls, J. Anderson, J. Hardwick, S. Morton, E. Phillips, Y. Zhang, and V. Volkov, "Parallel Computing Experiences with CUDA," IEEE Micro, vol. 28, pp. 13–27, July 2008.
2. Wasim A Al-Hamdani and Ivory J Griskell. A proposed curriculum of cryptography courses. In Proceedings of the 2nd annual conference on Information security curriculum development, pages 4–11. ACM, 2005.
3. Bart Preneel, Christof Paar, and Jan Pelzl. Understanding cryptography: a textbook for students and practitioners. Springer, 2009.
4. Tara Chand Singhal. Systems and methods for complex encryption keys, January 29 2013. US Patent 8, 363, 834.
5. Jianbin Fang, Ana Lucia Varbanescu, and Henk Sips. A comprehensive performance comparison of cuda and opencl. In Parallel Processing (ICPP), 2011 International Conference on, pages 216–225. IEEE, 2011.
6. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications.
7. Cuda toolkit documentation- Developer Zone.
8. Jonathan Passerat-Palmbach, Claude Mazel, and David RC Hill. Pseudo-random number generation on gp-gpu. In Principles of Advanced and Distributed Simulation (PADS), 2011 IEEE Workshop on, pages 1–8. IEEE, 2011.
9. Martin Feldhofer, Sandra Dominikus, and Johannes Wolkerstorfer. Strong authentication for rfid systems using the aes algorithm. In Cryptographic Hardware and Embedded Systems-CHES 2004, pages 357–370. Springer, 2004.
10. Darshana Upadhyay, Ankit Shah and Priyanka Sharma. "Design, Implementation and Analysis of GSM stream cipher: Software simulators v/s real test bed-FPGA", IEEE Conference: 2014 Sixth International Conference on Computational Intelligence and Communication Networks.

11. D. P. Upadhyay, P. Sharma and S. Valiveti, "Randomness analysis of A5/1 Stream Cipher for secure mobile communication," International Journal Of Computer Science & Communication, vol. 3, pp. 95–100, 2014.
12. Upadhyay, D.; Shah, T.; Sharma, P., "Cryptanalysis of hardware based stream ciphers and implementation of GSM stream cipher to propose a novel approach for designing n-bit LFSR stream cipher," in *VLSI Design and Test (VDAT), 2015 19th International Symposium on*, vol., no., pp. 1–6, 26–29 June 2015.

An Effective Strategy for Fingerprint Recognition Based on pRAM's Neural Nature with Data Input Mappings

Saleh A. Alghamdi

Abstract At present, information systems are highly susceptible to unauthorized human access. Various techniques have been introduced to secure these information systems. One of the most popular techniques for verifying the user for gaining access to information systems is fingerprint recognition. The popularity of fingerprint biometric is due to its invariant behavior over age and time. A novel and accurate fingerprint identification technique is presented in this paper based on neural nature of a pRAM. The proposed method uses a methodology incorporating the use of data mappings and reinforcement learning in order to maximize the efficiency and accuracy in identifying the scanned user prints. Since, the world is moving in the era of "Internet of things (IoT)," these biometric techniques are integral to the future information securing framework. pRAM-based network is a recently introduced technique employed in pattern recognition and is different from other classical neural network models reason being that a pRAM networks gets trained in relatively less time and can be implemented in minimal hardware setup. Here the application of the permuted mapping is derived using the proposed data-based input mapping with a bit plane-encoding scheme to cover multi-gray level images. Furthermore, binarization is also done using eight binary planes and a high-resolution image is processed by dividing it into sub-images so that it can be handled by several networks in parallel. The current recognition procedure has been applied on realistic fingerprint scans/images and the results drawn have shown significant improvements. The present results drawn here prove that a pRAM structure can provide highly reliable results by introducing the permuted mapping scheme for efficient identification.

Keywords Boolean neural network · pRAM · Pattern recognition · Input mapping · Reinforcement learning

S.A. Alghamdi (✉)
Department of Computer Engineering, Al-Baha University,
Al-Baha, Saudi Arabia
e-mail: salehatiah@bu.edu.sa

1   Introduction

Biometrics is used in both identity access management and access control. The design of a biometric recognition system requires a full-fledged study of all the related issues including explanation of pattern classes, pattern representation, feature extraction and selection, cluster analysis, classification design and learning. The use of neural networks further complicates the design by involving the challenges of network topology selection, training and re-training sets, and deriving test samples. This paper deals with the implementation of pattern matching using the pRAM paired with data input mappings to match and authenticate fingerprints. Also the paper uses neural network concepts to perform-effective pattern matching to identify fingerprints. In 1959, Bledsoe and Browning proposed the exhilarating N-Tuple method of pattern recognition, which acted as bedrock for more advanced schemes developed in later half of sixteenth century [1]. RAM technology was used later to develop hardware realizations of this method. The use of RAM technology permitted learning and classification in real time. The WISARD system which is a successful implementation for conventional pattern recognition used RAMs [2]. Kan and Aleksander proposed the concept of probabilistic logic neurons (PLN) and used them in a linked arrangement (logic-based network) [3]. These PLN's offered potential learning which was efficient than earlier used method. Gorse and Taylor proposed probabilistic RAM-based model [4]. The probabilistic RAM (pRAM) is a neural device, which is easily hardware realizable. It is stochastic in operation and is highly nonlinear. pRAM is seen as an extended version of probabilistic logic node (PLN), providing high levels of functionality even with small networks [5–7]. The probabilistic RAM (pRAM) model is a neural model that uses probabilities as weights called as synaptic weights. The memory in pRAM is assumed to be many bits where the values are treated as real numbers (Gorse and Taylor suggest a minimum of 16 bits) [8]. These synaptic weights are quantified precisely storing these values in a 16-bit register in the memory and get treated as an array in the range (0, 1). The synaptic weights are usually gathered by n-joint firing probabilities in the neural net of pRAM. The pRAM generates an output in the form of a spike train and works on the underlying principle of probabilistic automata which recognizes the class of synaptic weighted regular languages [9]. A Chomsky hierarchy general PNN (the probabilistic neural network) consists of d input units comprising the input layer, while each input unit is connected to the n pattern units forming a distinctive pattern. Each pattern unit in turn is connected to one of the c category units. The pRAM and the PNN are used with specialized system as in order to work out the pattern recognition. This specialized system starts with the data that is provided to the PNN's of pRAM in the form of mapping. The reinforcement learning is done internally with distinctive network formation as in order to identify the pattern in an effective manner. Use of pRAM with appropriate sampling scheme and the derivatives of sampling scheme help for the better and efficient fingerprint recognition. Fingerprint identification is known as dactyloscopy as discussed in previous studies [10–12]. Moreover pRAM is discussed in detail in the following section.

2   Probabilistic Neural Networks

As discussed in the previous section, the probabilistic RAM (pRAM) is considered as a Boolean neuron. [13, 14]. Famous neurophysiologist Warren McCulloch and Logician Walter Pitts invented artificial neurons in year 1943 [14]. These neurons are used generally as in form of computational units. The neurons actually behave like the biological neurons and are computationally very efficient. Use of these neurons is done as a VLSI processor in which large numbers of neurons are connected together for accomplishing a task. Figure 1 below shows a basic pRAM neural network.

   The basic form of pRAM comprises of following entities

1. N number of memory locations
2. A comparator
3. A noise generator

   The binary inputs to the memory module are used to select one of the locations. The weights (data stored in memory) in the location determines the probability of pRAM firing which is achieved by adding a RND number from the random number generator to the selected memory contents. If the result exceeds certain selected bits, pRAM fires, resulting in memory to shape up in 0, 1 pattern. It was the third generation of PRAMs where 16-bit memory module was used as suggested by (Gorse and Taylor) [8]. It offers "off chip" learning in which, once the synaptic weights have been calculated; only significant bit is copied into hardware for the next iteration of the training process. pRAM is a conceived model where n-tuples are used and operates on the generalization of synaptic lookup tables. pRAM is trained on the probability of firing which internally is analogous to the release of neuron-transmitter of a biological neuron. The weight gets modifications as learning is applied on the network of neurons using different algorithms for learning. The output of the pRAM is a spike train on which no arithmetic is performed. Multiple probabilities are fired at different joints in pRAM network. It is the mean of the firing probabilities at different inputs, which in return indicates the locations that are to be accessed.

Fig. 1 pRAM neural network
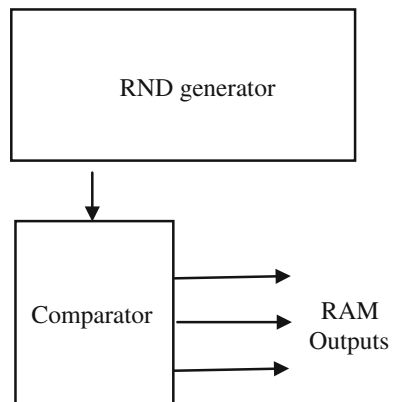
pRAM nodes in the network are used in different layers and these nodes are well connected. Nonlinear binary functions can be learnt by pRAM network, upon which the training is generalized. pRAM can store highly nonlinear set of responses mainly from its binary inputs. pRAM is used in recognition as well where nonlinear sets of data are to be stored or learned like in several temporal patterns, sequential pattern verification of signatures, etc., which is handled by the topology know as single layer sequential RAM-based network [15–18]. This pRAM network is the implementation of the *n*-tuple technique as already stated above, which has mostly been used for pattern recognition [1]. The present results drawn here prove that a pRAM structure can provide highly reliable results by introducing the permuted mapping scheme for efficient identification. The connections from the input to pattern unit's category unit represent modified weights, which are trained during the training iterations. Intrinsically connections as stated above uses pyramidal topology, which takes its shape when connecting input layer to the output layer. Each individual pyramid is employed to handle a p-pattern classification problem at any node [19]. Thus the use of pyramidal network sets the stage for the learning to occur as the input is mapped and is pre-analyzed that makes the learning efficient. Learning scheme based on reinforcement as discussed in next section seems to be promising, potentially ideal and effective. The use of reinforcement learning also suits the type of topology of the network used herein. Our previous work has been in same regard with an advancement provided in this work [20].

3 pRAM Learning Fundamentals

Learning in terms of a machine is a way of providing the sense to machines and the field which deals with this is artificial intelligence (AI). The methods which make a machine understand human actions and behavior is termed as learning. Explicitly, it is a learning that trains a machine to do a specific task in a specific stated manner. The learning schemes are used to train systems, search through data to look for patterns and compare the patterns and increase the knowledge base, i.e., the learning schemes use that data to improve the program's own understanding [12]. These learning programs detect patterns in data and adjust programs and actions accordingly. Learning is done using various types of learning schemes, but in this paper primary focus is on reinforcement learning among the schemes available, which differs from supervised and unsupervised learning. The learning here is also based on the neural nature of pRAM which is done in intervals of 0, 1, wherein these values are stored in the memory locations interchangeably known as weights, and the weights are modified owing to the firing in pRAM. It is the combination of stochastic and nonlinear behavior of pRAM which gives rise to learning ultimately. Technically, it has multiple address inputs, if the number of inputs is n, then memory locations in pRAM will be 2^n. Multiple layers are required in learning so as to overcome problems of parity and other logical problems. Thus multiple-layered models are used to cater all the problems in learning scheme. By having layers,

generalization problem is also solved as sometimes the weights are not large enough to make pRAM fire; there the encrusted scheme comes into the scene by generalizing the weights in the memory area. As already stated reinforcement learning algorithm used here, dates back to 1960, originated from dynamic programming, Monte Carlo method and Markov decision processes (MDP) [21]. The most common approach for training is the generalized delta rule. Basically, the learning method is an adapted method proposed by Barto and Jordan and later on Gurney elaborated the method with more facts, where convergence proof is also given. The model consists of a set of environment states, S; a discrete set of actions and A; a set of scalar reinforcement signals; typically 0 or 1, or the real numbers. In this paper, the proposed learning scheme method employs reward/punishment mechanism of reinforcement algorithm. This algorithm is the adaptation of method propounded by Barto and Jordan and later on it was described fully by Gurney [22]. This scheme works toward beneficial response behavior and the learning is done stochastically. The learning phase in the adapted learning scheme is based on the adjustment of the content (weights) of the pRAM memory i.e., "α's" at each time step. Gurney made the observation that for huge number of inputs, training results get sparsely populated, thus he suggested dividing the inputs into number of sub locations each with a separate address [23]. According to the rule given in Eq. 3 [1].

$$\Delta\alpha u = \alpha u(t+1) - \alpha u(t) = \Delta\alpha u r + \Delta\alpha u p = \delta(r(a - \alpha u) + \lambda(\bar{a} - \alpha u)p) \quad (1)$$

where a is the pRAM's output, r being the reward, and p being penalty signals.

$0 < \delta < 1$ and $0 < \lambda < 1$ are the learning constants. The reward r and the penalty p increases or decreases the memory contents in αu. Substitution occurs in a way to reinforce the probability of making similar moves in the next time step as explained below in the particular

case of $p = 1 - r$.
if $a = 1$ and $r = 1(p = 0)$:
$\Delta\alpha u = \delta(1 - \alpha u) > 0 \qquad \ldots \quad \alpha u$ increases
if $a = 0$ and $r = 1(p = 0)$:
$\Delta\alpha u = \delta(-\alpha u) < 0 \qquad \ldots \quad \alpha u$ decreases
if $a = 1$ and $r = 0(p = 1)$:
$\Delta\alpha u = \delta\lambda(-\alpha u) < 0 \qquad \ldots \quad \alpha u$ decreases
if $a = 0$ and $r = 0(p = 1)$:
$\Delta\alpha u = \delta\lambda(1 - \alpha u) > 0 \qquad \ldots \quad \alpha u$ increases

The two constants λ and δ for learning are selected by applying trial and error method. The mean value for R time steps is derived from updating the rule from Eq. (1) and is given in Eq. (2),

$$\Delta\alpha\,u = \sum t\,\Delta\alpha\,u\,r(t)/Rr + p + \sum t\,\Delta\alpha\,u\,p\,(t)/Rr + p \qquad (2)$$

where $Rr + p$ is the number locations accessed during R time steps. In the case of global reinforcement learning, the propagation of the reward or penalty signals from the output layer to hidden layers and back to the input layer at the last hidden layer is performed accordingly. After the previous step, the necessary training is done iteratively till low error rates are achieved. After each training iteration, the results are verified. The derived results showing highest value for convergence indicates the best match. The training results are updated in memory based on average contributions from pattern classes. The performance is measured based on the root mean square error given by Eq. (3).

$$\mathrm{RMS} = \sqrt{\sum p \sum j (\mathrm{tpj} - \mathrm{opj})2/\mathrm{npnj}} \qquad (3)$$

where np defines the number of pattern classes, nj is the different exemplars of each pattern, tpj is the desired output and, opj is the actual output mean firing rate. The training set is scanned and then batch processed to compute the error. However, the memory contents are updated on a pattern by pattern basis during R time steps as is illustrated in algorithm (1).

Algorithm(1)

Start
Initialize 1$^{\mathrm{st}}$ layer

```
 Do {
   Set exemplar number (Size of training set) NT
   Set pattern type (Number of different classes) PAT
   Connect input to 1ˢᵗ layer (input mapping selected)
   Set pulse train (Length of the spike train) R
         Record accessed addresses  u's
         Compute outputs
         Connect layers
  Derive reward/penalty (output layer) r,p
  Update/ change of memory contents(accumulate) Δαᵤ ᵣ ,Δαᵤ ₚ
  Test for the expiration of the spike train R
      Compute the output-mean value
      Derive the error for each pattern and accumulate
        Test for the pattern type PAT
        Update the memory contents in all layers αᵤ = αᵤ + Δαᵤ
    Test for the training set size NT
           Compute final error
          } while (error big)
 Go to   Do
    Save memory contents
  End
```

The learning scheme operates on the input data, which in this paper is the image which are pre-analyzed and pre-processed for better efficiency in recognition of patterns which in our case is fingerprint scans. The next section discusses about the input image which are pre-processed before being served as an input to the pRAM.

4 Advance Image Pre-processing

The 2D image scans are used to implement recognition by processing the input images in pRAM network. The focus is on the spatial representation of the image. Figure 2 shows the overview of the pre-processing done on the input image, owing to the variety of pictures or scans.

The images which are taken for processing vary in prints, so a strategy is adapted in which image scans and their noisy versions are processed in first phase using appropriate schemes and various normalizing algorithms available with the help of various tools and certain programming paradigms to generate the representations which are most suitable for the pattern recognition. The noise which may be part of a scanned fingerprint image makes fingerprint extraction quite intricate. The irregular distribution of the pixel intensity levels in a scanned fingerprint image is a more complicated for ridge orientation identification as can be seen in Fig. 3 below.
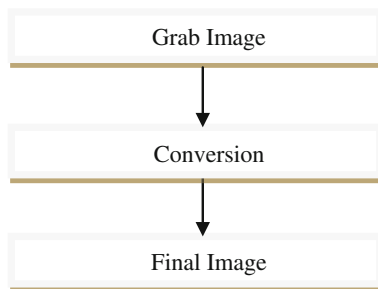
Fig. 2 Image conversion process overview
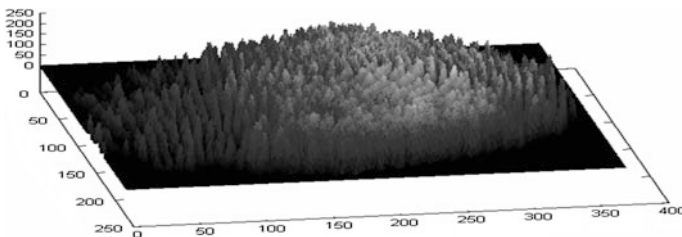
Fig. 3 Fingerprint with noise (uneven ridges)

Fig. 4 Original scan on left
and normalized scan on right

Many of these algorithms are used in the process of pattern recognition; it is worth mentioning that adaptation of this systematic approach in image conversion also helps us in quantifying inadequacies as seen in Fig. 4. While observing the whole process, emphasis is given on performance by fine-tuning the deficits that are encountered while quantifying the procedure for inadequacies.

These types of image scans are pre-processed using the standard conversion process. Then depending on the image scans, these varied image scans are treated based on the nature of the image scan, i.e., the dependencies, which are in form of gray levels and the resolution of the input scans.

As the part of the input images, encountering a varied gray leveled scans using pRAM net with binary images is forthright and fairly easy, since the binary image present pixels of values 1 (white) or 0 (black). Thus, the input to the pRAM network is the raw contents of the image itself. Bit plane decomposition method used is done in two phases.

Phase 1. Decoding the gray level of each pixel using digits,
Phase 2. Forming N binary planes for spatial resolution as given in the original gray-scale image.

This method is used is known as the bit plane decomposition technique there in exploiting spatial correlation and decomposing the image into set of binary layers.

5 Advanced Sampling Scheme and Fingerprint Recognition

The subprocess here initiates with the input sampling where n-pixels are to be extracted. Generally, n-tuple state is also referred to as an address into a single pRAM [24]. For decoding the corresponding n-tuple state, a novel input mapping based on data analysis is presented below.

1. Each image (image size in pixels) is represented in terms of number of pattern.
2. A training set for each pixel is defined.
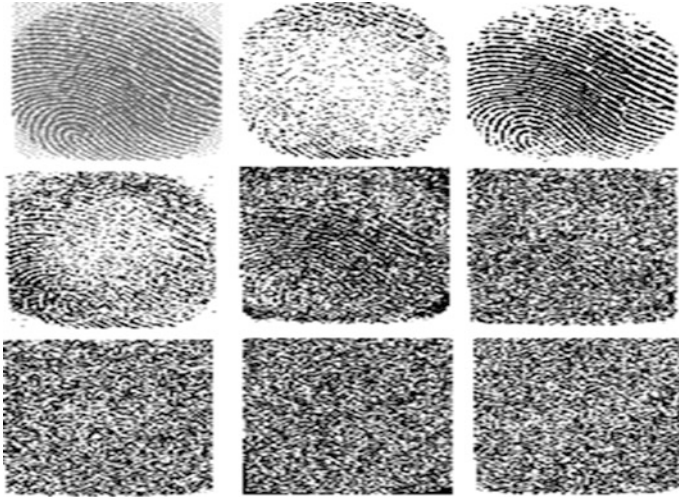
Fig. 5 Multi gray level image and its conversion to eight binary planes 4

3. The resultant probability of a given n-tuple is derived from the individual probability densities of the participating pixels. This sampling scheme which is based on the selective input mapping scheme is discussed in the next section.

As the above section defines a new sampling scheme, which will provide the enhancement in terms of recognition rates and the failure rates of the scans.

The fingerprint scan used 192×128 multi-gray level images, so we convert it into binary images using the previously stated method and we divide the resultant images into small chunks each of which has a size of 24×16, shown in Fig. 5 above.

6 Tests and Results

The above-proposed mapping was tested by computer-based simulations for recognition of 24×16 binary images. As many as three input mappings were tested, namely

1. Structured mapping
2. Data-based mapping
3. Permuted

Also Fig. 6 shows recognition error function of the number of pRAM network iterations for three different mappings. This figure also summarizes the proportion of each n-tuple type obtained from a corresponding n-tuple input mapping.
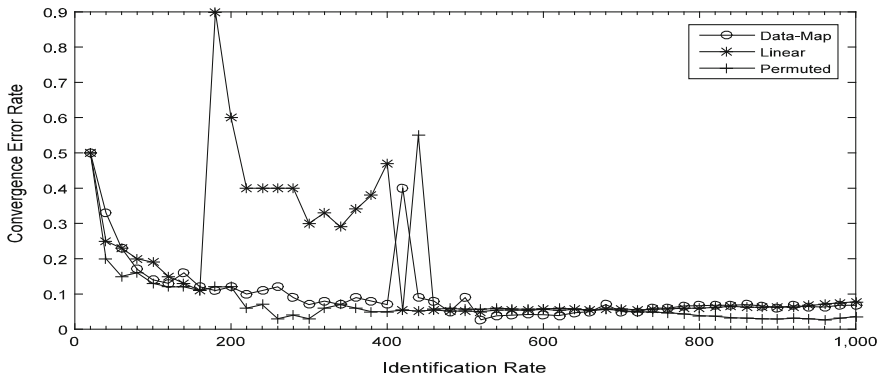
Fig. 6 Convergence network iterations error rate of number of neural

Figure 6 shows that the network's convergence error rate is less in permuted and data-based mappings when compared to the linear mappings plotted along y-axis and also the later mapping technique showed greater accuracy in identification with fewer errors.

In Fig. 6, the permuted mapping managed low convergence error rate of 2 % compared to 13 % for the structured mapping and 6 % in the case of the data based. The fluctuations were seen thus confirming the effect of changes in the pRAM memory contents whose values vary in the range [0, 1]. With these different mappings, the identification rate increases from the linear mapping to derived permuted mapping from the proposed data-based mapping. The Fig. 7 shows the identification rates in the given mapping techniques. The network using the permuted mapping on average achieves the highest identification rate of 93 % per hundred scans, whereas the non-permuted mapping only achieves 80 % identification rate as shown in the above Fig. 7. The Fig. 8 clearly shows that the n-tuples distribution rate is high in case of linear but systematizing the n-tuples in order to get better of recognition rates forms the basis of permuted mapping. The data-based
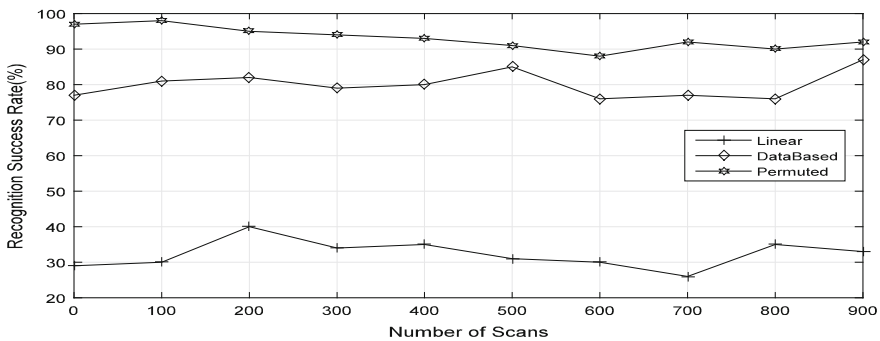
Fig. 7 Recognition rates of differentiated mappings
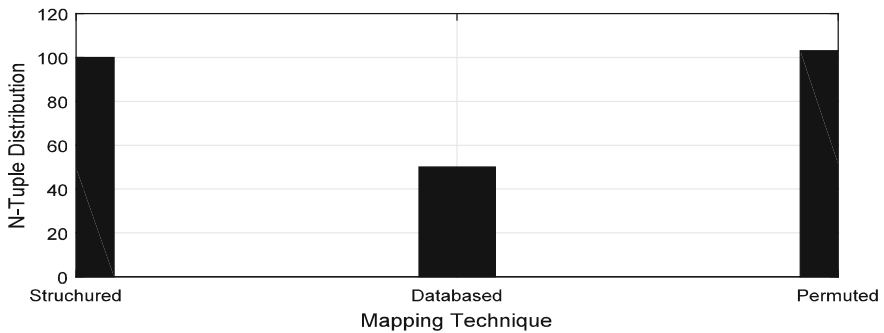
Fig. 8 Uniform distribution of *n*-Tuples among mappings

mapping shows relatively less *n*-tuple distribution as compared to other mapping techniques. The derived mapping has better efficiency which is clearly reflected in the results.

7 Conclusion and Further Work

The introduction of neural networks radically affects the results of fingerprint identification in a positive manner. The use of pRAM enhances the reliability of performance of fingerprint biometric recognition. The proposed advancements in this paper showed better results and further strengthened the use of pRAM in pattern recognition in future for other biometrics. The current results obtained are promising and the use of such schemes guarantees better results and increased efficiency in the field of user identification and authentication.

References

1. W. W. Bledsoe I. Browning, Pattern Recognition and Reading by Machine.
2. Eduardo Sanchez, Marco Tomassini, Towards Evolvable Hardware: The Evolutionary Engineering Approach p. 135.
3. I. Aleksander, M. De Gregorio, F. M. G. França, P. M. V. Lima, H. Morton brief introduction to Weightless Neural Systems.
4. D. Gorse and J. G Taylor. A continuous input RAM based Stochastic Neural Model.
5. W. K Kan and I. Aleksander. A Probabilistic logic neuron network for associative learning.
6. Clarkson T. G, Guan Y, Taylor J. G, Gorse D. Generalization in probabilistic RAM nets.
7. Clarkson TG1, Ng CK, Guan Y, The pRAM: an adaptive VLSI chip.
8. Gorse D, Taylor J. G, Training Strategies for Probabilistic RAMs, In Parallel Processing in Neural systems and computers, Eds. Eckmiller R, Hartmann G, and Hauske G, Elsevier, 1990.
9. Trevor G. Clarkson, pRAM: The Probabilistic RAM Neural Processor.

10. Ashbaugh, David R. "Ridgeology" (PDF). Royal Canadian Mounted Police. Retrieved 2013-10-26.
11. M Triplett, L Coonez, The etymology of ACE-V and its proper use: An exploration of the relationship between ACE-V and the scientific method of hypothesis testing.
12. Sergiy Stepanyuk, Neural network information technologies of pattern recognition, Perspective Technologies and Methods in MEMS Design (MEMSTECH), 2010 Proceedings of VIth International Conference on 20–23 April 2010.
13. T. G. Clarkson, The pRAM as a hardware realisable neuron, Proceedings Neural Networks, 1992 pp. 140–146.
14. M. Ouslim, Analysis of the n-tuple technique based upon two perspectives to determine n-tuple groupings, PhD thesis, University of Nottingham, 1997.
15. P. J Adeoadadato and J. G Taylor. Recurrent neural networks with pRAM S, pp. 607–612 Springer-Verla, 1995.
16. D. Gorse and J. G Taylor, ON the equivalence properties of noisy neural and probabilistic RAMnets.
17. T. B Ludermir. A Feedback network for temporal pattern recognition. In Parallel Proc., in Neural Systems and Computers pp. 395–398, Amsterdam, 1990.
18. J. E Hopcroft and J. D Ullman. Introduction to Automata Theory, Languages and computation. Addison Wiley Publishing Company, 1979.
19. M. Ouslim and S. A. Alghamdi, Enhancing the main properties of the pRAM neural network. 7th annual IEEE technical exchange meeting, April 2000, Saudi Arabia.
20. Saleh Alghamdi, "Fingerprint Recognition Upon a Data Analysis Input Mapping For a pRAM Neural Network", Minufiya University, Electronic Engineering Bulletin, Vol. 31, No. 3, July 2008.
21. Chelsea C. White III, Douglas J. White, Markov decision processes.
22. BARTO, AG Jordan I, Gradient following without back propagation in layered networks.
23. Gurney K, Training Hardware realisable sigma's Pi Units, Neural Networks, vol 5, no, 2, 1992.
24. M. Ouslim and K. M. Curtis, P pattern recognition based on a probabilistic RAM net using n-tuple input mapping, IEE Proc. Vis. Image, Signal Process., Vol. 145 No. 6, Dec. 1998, pp. 415–420.

Various e-Governance Applications, Computing Architecture and Implementation Barriers

Anand More and Priyesh Kanungo

Abstract To speed up the digital India mission, e-Governance applications and policies must be effectively designed and implemented. In this paper, various applications have been proposed that can be implemented using e-Governance to provide efficient online services to the citizens. We have also proposed the infrastructural requirement to implement these applications. Various barriers to the implementations have also been discussed.

Keywords e-Governance · Index for human capital · Web measure · Digital India · e-Governance preparedness · G2G

1 Introduction

e-Governance is the delivery of government services using Internet to the citizens, businesses or governmental departments. These are web-related services over the Internet by government agencies and is a gateway to major government services to support official processes to provide services to the citizens of the country. The interaction may be to collect the information or payment proceedings and similar activities using the World Wide Web [7].

In the developing countries e-Government defined by the Working Group as "using the information technologies to provide improved government, by means of more accessible services, allow information access to public, and more accountability of government to its citizens. e-Government involves services using the Internet, and mobile, wireless devices or other communications systems" [5]. Thus, the e-Governance services are classified as follows.

Anand More (✉) · Priyesh Kanungo
School of Computer Science, Devi Ahilya Vishvidyalaya, Indore, India
e-mail: sugamelec@yahoo.com

Priyesh Kanungo
e-mail: priyeshkanungo@hotmail.com

1.1 Government-to-Citizen (G2C)

G2C are the services in which the government delivers online to citizens. The G2C applications enable citizens to raise any questions related to their process and working model about government agencies and get satisfactory answers; pay income taxes, other taxes, traffic tickets, renew driving licenses, change address; appointment for driving tests, and make vehicle pollution report after inspections [2].

1.2 Government-to-Business (G2B)

G2B, refers to the government managed businesses using soft technology mode like the Internet and other ICTs tools. This includes two-way communications and dealing between government-to-business as well as business-to-government (B2G). B2G deals with business-to-government, i.e. businesses related to products or services or both to government. There are two main G2B areas, viz., e-Procurement and auction of stocks.

1.3 Government-to-Government (G2G)

This deals with processes and actions that take place among various government departments. Several such acts are designed to improve the productivity of operations.

To measure status of e-Government, several indices have been proposed that are display gauges for the implementation progress of e-Government services as developed by the UN countries. These indices are described in the following paragraphs:

- *Web Measure Index* Web Measure Index specifies various phases of government websites which indicate several levels of web presence, e.g. official web site, information (e.g. reports, policies and regulations, databases, newsletter, etc.), Interactive Presence, Transactional Presence, Networked Presence and so on.
- *Telecommunications Infrastructure Index* This index measures the utilization of a country's ICTs (Internet servers per ten thousand of people, percentage of nation's population that is online, and availability of Computers, landline phones and mobiles per thousand people);
- *Human Capital Index* It can be measured by indicators like the Information Access Index, Human Development Index, and ratio of urban to rural population (Fig. 1).

e-Government readiness was undertaken by the UN included 191 countries, is a overall measurement method of the capacity and inclination of countries for using e-Government with ICT-led development. It is dependent of the combined level of
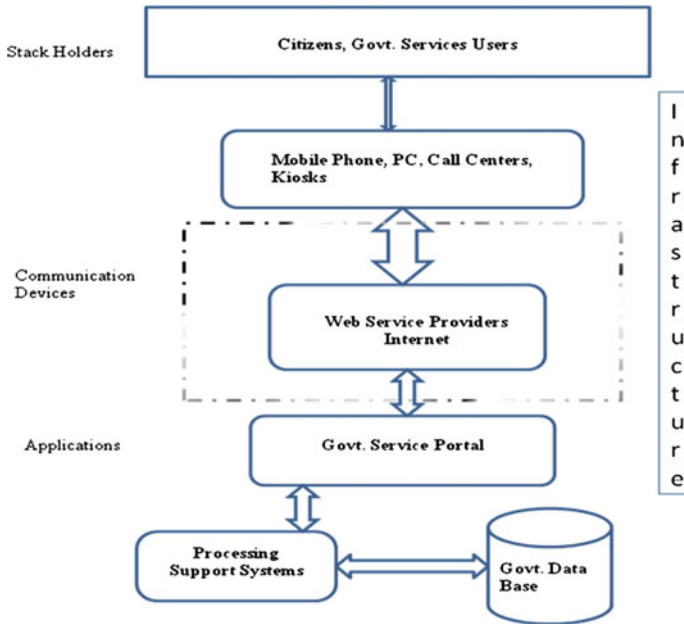
Fig. 1 e-Governance architecture

a country's readiness, level of technology, economy and HR development. e-Governance index includes these following indexes Human Capital Index, Telecommunication Infrastructure Index and Web Measure Index as described above [2].
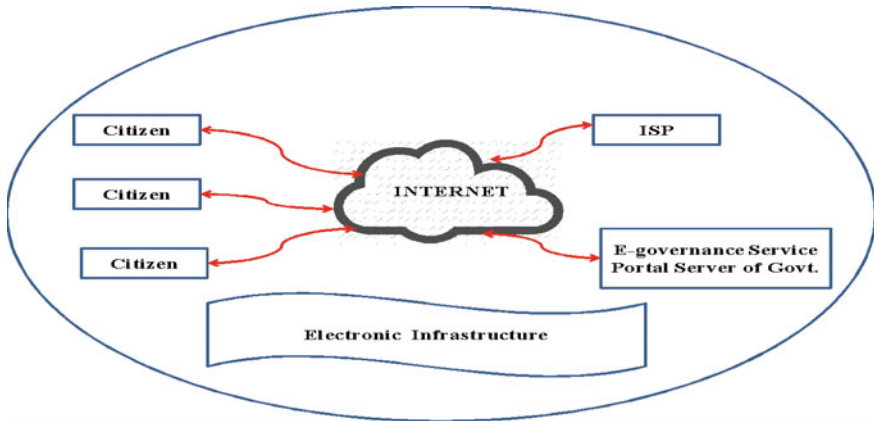
2   e-Governance Services

For exercising e-Governance system Government introduced various services for the benefit of citizens. Some of these services are discussed below [9].

2.1   Agriculture Services

Government needs to take several initiatives to support farmers to facilitate their agriculture production. As a geographical difference farmers are not aware of various important facts like:

- To use seeds, fertilizers, pesticides to increase productivity.
- Government's Schemes for the farmers.

Fig. 2 Connection between Government and Citizen

- Recommendations and standards related to Soil Health.
- Sale and Purchase of agricultural products.
- Weather forecasts.

Through e-Governance services for the farmers, they are well informed with the agricultural community information in their regional language about weather forecasts, insurance of crop, market prices, latest farming techniques and scientific farm practices, risk management. Prices for various commodities are available online so that farmer can directly negotiate the prices and can earn more profit. e-Choupal consists of the good number of farmers connected through digital infrastructure. Governmentt of India's various services are operational by Department of Agriculture and Cooperation (DAC), such as—Agriculture, Agmarknet, Portal for Farmers, Plant Quarantine System, Board for Central Insecticides and Registration Committee, Soil and Land Use Survey of India, Extension Reforms, Hornet, Dacnet, Farm Mechanization, KKMS, Nav Krishi, RKVY, Seednet [10] (Fig. 2).

2.2 Land Records

The Computerization of Land Records and land information accessibility is one of the significant initiatives of e-Governance. The main focus of the procedure has been to make use of state of art of IT to convert the existing land record system. Main objectives of this service are [8]:

- Ensuring that owners of land get computerized copies of ownership, updated copies of Records of Rights (RoRs) and crop and tenancy on demand.
- Easy to reproduce low-cost land record data through reliable and long-term old records maintenance and reservation.

- Ensuring speedy resolution, transparency and accuracy of any dispute related to land.
- Smoothing the process of fast and efficient recovery of information for decision-making.
- The Land Records which are implemented by Rural Development Ministry. Computerization of Registrar Office includes the following key functions:
- Online property registration on sale and purchase deeds.
- Online issue of copies of a previously registered deed.
- Details of Soil type, Irrigation facilities and Crop production.
- Pursuing the progress status of cases under mutation.

2.3 Rural Services

In the rural development IT initiatives had been undertaken. Panchayats are facing problems of inadequate financial and physical resources, technical skills due to which extremely limited computerization done at rural level. Inputs from stakeholders and experts in the various public and private sector are taken for the development process. The Ministry of Panchayati Raj, Government of India has therefore planned to initiate the computerization of Panchayat. The e-Governance project cover all aspects of Panchayats' functioning including Planning, Monitoring, Accounting, Budgeting, Social Audit and delivery of services to citizen like issue of certificates, licenses, etc. On the basis of collected information from various Panchayat the e-Panchayat was initiated. Computerization of Panchayat is developing and empowering the rural communities at the Panchayat levels through an interactive and collaborative web-based services. These portals are run, managed and sustained by the rural community through regular information and content management. The main objectives of e-Panchayat are:

- Data centre and gateway of e-Democracy.
- Create a web site and populate the web site with all relevant information of campaign.
- Create Internet and ICT awareness at Panchayat level
- System for citizens to lodge complaint directly and publicly.
- Provide e-Gov platform with G2C, B2C services, etc.

2.4 Municipal Services

e-Governance in Municipal services is aimed to improve efficiencies of operation within Urban Local Bodies (ULBs). With the acceptance of ICT in Municipal Corporation; they are excelling through e-Governance in the field of delivering

services to the citizens. To implement it, following aspects are required to be improved:

- Computerization across ULBs.
- Increasing staff with IT skill.
- Standardization of processes.
- The processes operated in a manual mode are to be transferred in digital mode.

e-Governance in municipal services focuses on developing citizen services through defining outcomes and service levels; easier interaction between local body, citizens and other stakeholders; improving transparency and accountability; developing citizen interface and enhancing service delivery to citizens. The main services in Urban Local Bodies (ULBs) have been identified to deliver quality services to the citizens are Birth and Death Certificate with Guidelines, New water Connection, Sanitary connections, Register Complaint, Property Tax, etc., Tender Notices, Trade License Builder License and Permit.

2.5   e-Choupal

e-Choupal is an initiative to support farmers to facilitate their agriculture business. This project benefited the farmers in all direction. In this project a person trained among the group of farmer and appointed to run the internet kiosk. Through e-Choupal farmers are well informed with the agricultural community information in their regional language about weather forecasts, insurance of crop, market prices, latest farming techniques and scientific farm practices, risk management. The service is mainly focused to facilitate the sale and purchase of farm products. Prices for various commodities are available online and hence farmer can directly negotiate the prices and can earn more profit.

3   e-Governance Infrastructure

The backbone infrastructure is needed for connecting blocks with districts and districts with capitals of state and also these capitals of state with country's Capital. The infrastructure will facilitate transmission, storage, processing of voice, data, and images.

- The infrastructure consists of LAN, servers, etc. It will include the communication equipments, networking, servers, etc. It also constitute with the satellite links, leased lines, copper links, etc. for connectivity [3].
- It comprises of a vast range of equipment like keyboards, computers, cameras, telephones, scanners, fax machines, compact disks, cable, wire, video–audio tape, televisions, satellites, transmission lines of OFC, microwave, switches, nets, etc. [3].

- The information can be sound recordings, scientific database, video programming, images, library archives, businesses data and other media.
- Software applications which not only allow end users to accomplish and access, but also to establish, make changes and accept the increasing information.
- The transmission data and networking standards facilitate to interconnect and interoperate between any networks that guarantee the persons' privacy and security of the information shared. This also increases the trust and security of the networks [6].
- Several people working in private industry like vendors, operators and service providers create information to develop applications and services, make the services available and provide training and e-Learning facility to increase the potential [4].

Digital India can be a successful programmed if following major aspects are taken into consideration:

- Internet should be economic and fast.
- Proper training to the users.
- Better connectivity.
- Continue electricity supply.
- Government services easily available at mobile.
- Online forms available.
- Bill and fee submission should be online.
- Proper SMS alert and notification.

4 Barriers to Implementation of e-Governance and Digital India

India's population is approx. 12.5 billion, out of which 10 billion population is out of the reach of Internet, still electricity has not reached to 0.40 billion population. Following are the major hurdles in the implementation of e-Governance in India [1]:

- *Lack of Awareness Among Public* It is not sufficient to build the infrastructure for e-Governance. It is need of an hour to create alertness among the public for the successful implementation of e-Governance which lies in growing and collectively use of electronic interactions between people and the government.
- *Resistance to Change* After following old skills and habits, governance system lean towards to develop disinterest and confrontation to change. If e-Governance has to accept, these procedures will be required and substituted with new skills and procedures. This is well known that in every kind of organization, there are a few types of people who are not easily convinced or ready about the benefits of e-Governance or people who would deceive the change in contrast with their personal interests. The resistance can be overcome

by showing that e-Governance is capable of strengthening the organization, easy is operation, increasing citizens' satisfaction level and creating the goodwill in the society.

- *Lack of top Management Support* e-Governance requires complex process reengineering task at various administrative levels. It also needs financial support and infrastructure development. Tall these need Top management or highest political level vision as well as commitment.
- *Lack of Training Facilities* Government officials and employees may not be interested in learning new techniques or change in way of working due to lack of awareness of Information Technology and operation of state-of-art devices. To make e-Governance as an achievable target, training facilities need to be developed so that the persons responsible for implementing it are not apprehensive of new technology.
- *Incentives* Preventing government organizations from using and implementing manual process to the adoption of e-Governance tools of application technology will require incentivizing and promotions of e-Governance at different levels of different units and individuals. Such incentives need to be separately allocated in budget. Government also wanted to zero the import for electronic items under "Make in India" program, and it itself is a big challenge for Digital India program, as our electronic item import is 65 %.
- *Nonavailability of Infrastructure* India is struggling to manage bare minimum infrastructure and finance. India's dream to become Digital India requires solving these problems.

(i) *Spectrum* Spectrum is the backbone to Telecommunication industry. Recently 1 billion Spectrum were distributed to different companies through auctions. Even than these numbers are not enough to fulfil the 3G demands. Singapore, Shanghai and Delhi have equal number of 3G users but in comparison to these cities Delhi's share of spectrum is 10 %.

(ii) *Broadband Connectivity* Lack of infrastructure and economic resources are the biggest challenge to develop required broadband connectivity. Till April 2015, broadband connectivity reaches to only 10 million people but if we compare other country like S. Korea broadband connectivity is available to 97 % of the population.

(iii) *Electricity* Officially 280 million people still facing electricity problem and there is a regular electricity disruption in small cities and towns of India. In such a case, high capacity spectrum or Speed Broadband connectivity have no use until and unless power generation is not increased.

(iv) *Mobile Network* 42,000 Indian villages are not under coverage of Mobile network. Around 950 million people are Mobile users out of which 210 million people use Internet on their mobile. Telecommunication infrastructure, lack of awareness, high cost would be additional reasons into the progress of Digital India Dream.

5   Conclusion

With the new dimensions in technology, India is looking forward to become Digitally Progressive Power. A strategy for e-Governance implementation is need of the hour for digital India. Implementation of strategies with global vision is required for the long-term success. But foremost and most important thing is the willingness of citizens as well as the employees to accept the innovation. The applications and infrastructural requirements mentioned in this paper can act as useful information in design of new e-Governance applications.

References

1. Gupta, P., Bagga, R.K.: Implementing e-Governance Reforms. 5 Inaugural address at IIT Delhi during International Conference on e-Governance, (reproduced in 'Compendium of e-Governance Initiatives in India', ed.; https://www.csi-sigegov.org/publications.htm) (Dec., 2003).
2. Jain, S. C., Palvia and. Sharma, S.S.: e-Government and e-Governance: Definitions/Domain Framework and Status around the World. Foundations of e-Governance.
3. Kanungo, P. and More, A.: "Design of Networking Infrastructure for Academic Institutions with Special Reference to Wireless Networking," International Journal of Science Engineering and Management," vol. I, pp. 82–86 (ISSN: 2250-0596) (January 2012).
4. Mehta, H., Kanungo, P. and Chandwani, M.: "Towards Development of a Distributed e-Learning EcoSystem," 2nd International Conference on Technology for Education (TforE-2010), IIT Mumbai (July 1–3, 2010).
5. National Informatics Centre (NIC): e-Governance Initiatives in India. Promoting e-Governance – The SMART Way Forward. Source: http://it-taskforce.nic.in/prem.htm, 26–58.
6. Sachdeva, S: e-Governance Strategy in India White Paper on e-Governance Strategy in India (December, 2002).
7. Sharma, R.K., Kanungo, P. and Chandwani M.: "An Intelligent Cloud Computing Architecture Supporting e-Governance," International Conference on Automation and Computing, University of Huddersfield, Huddersfield, United Kingdom (Sept 2011).
8. Sharma, S. K.: An e-Government Services Framework, Encyclopedia of Commerce, e-Government and Mobile Commerce, Mehdi Khosrow-Pour, Information Resources Management Association, Idea Group Reference, USA, pp. 373–378 (2006).
9. Various e-Governance Services: http://india.gov.in/e-governance/mission-mode-projects/state-mmps.
10. Various e-Governance Services for farmers: http://agricoop.nic.in/egov.html.

Fuzzy Logic-Based Expert System for Assessment of Bank Loan Applications in Namibia

Dharm Singh Jat and Axel Jerome Xoagub

Abstract In the present scenario, the majority of people in Namibia do not have enough capital to build houses, start-up businesses so they turn to banks to get a loan. In today's era, banks loans are considered to be one of the highest market risk phenomenon. Unsecured bank loans that are given without any collateral are a major stumbling in today's world, which can be seen from the economic recession that occurred across the world. The aim of this review paper is to present the exhaustive review and introduce a fuzzy logic-based automated artificial system, replacing the existing manual system, which will help in decision-making for disbursement of the bank loan and give the reason for the inferred decision.

1 Introduction

Nowadays one of the major component of banks and their main operations are the assessment of bank loan applications. This assessment is very important as it involves risk for both individuals, corporations, and the bank. In the current financial crises, banks have suffered losses from a continuous increase of customers that get default loans [1]. Therefore, banks should start to find a way to avoid approving loans, where the customers will not be able to comply with the terms. It is very important to approve loans which will satisfy all the important requirements and that will be able to manage to pay for corresponding demands. Various parameters such as age, the number of dependents, employment history, and repayment history all this which is related to applicants needs should be considered. So there must be a system that will help to cater for more robust and better banking

D.S. Jat (✉) · A.J. Xoagub
Namibia University of Science and Technology, Windhoek, Namibia
e-mail: dsingh@polytechnic.edu.na

A.J. Xoagub
e-mail: jeomex@gmail.com

analysis because banks have to write off bad debts between 20 and 40 % yearly due to not taking proper measure when the people apply for bank loans [2].

Most decisions in financial institutions are based on two factors, viz., past experience and a set of rules, which is usually done manually by analysts. Therefore, once the work becomes routine, enormous problems will arise based on human error or fatigue, etc. To overcome such problems, a computer-based system is needed to be developed [3]. The purpose of the paper is to try and reduce this risk factor by introducing a new system (expert system) which will be based on knowledge support system, which has the capacity to learn using the inference engines [4]. Currently, banks in Namibia take a long time to disburse the loan as a lot of information needs to be studied and analyzed by a lot of human resources and this normally causes delays in informing the customer whether their loan is accepted or declined.

This paper lays emphasis on the design, implementation, and evaluation of an expert system for Namibian bank.

2 Expert System

An expert system is a computer application that uses computer equipment, software, and special information given by a human expert to think like a human. As a subset of artificial intelligence, expert systems gives advice of a particular field and also states reasons on why it gives those particular advises. Typically, expert systems work best with certain activities or problems and with a discrete database which contains facts, rules, cases, and models.

Expert systems are currently being used in different fields such as in commercial and industrial settings which include medicine, finance, manufacturing, and sales. A knowledge-based expert system is a system where the application contains knowledge and it uses the knowledge it has to reason and come up with a conclusion. The knowledge is entered into the system with the help of a human expert.

An inference engine is a tool that is part of artificial intelligence. Fuzzy logic, as a part of inference mechanism, is a computational paradigm that is based on how humans think. The way fuzzy systems differ from classical logic is that statements are no longer just black or white, true or false. In the traditional logic, a variable can take on a value of either one or zero. In fuzzy logic, an object can assume any value between one and zero, representing the degree to which an element belongs to a given set.

In this paper, it has been tried to find how can we reduce the time it takes in assessing a bank loan application and reduce the number of default bank loan applications.

3 Literature Review

In a study in Greece, the authors have proposed the use of neural networks and fuzzy systems techniques [1]. A decision support system was proposed which is based on the credit scoring system [3]. This scoring evaluation system uses the numbers from one to ten. The study used the 4 Cs of Credit which are Capital, Character, Capacity, and Condition that are used to help in making the assessment of credit risk systematic. Each of the 4 C's is then dissected more so that Capacity contains income, the number of dependents, debt, and character contains age, repayment history, employment record, Capital contains debt ratio, conditions contain bank relationship, and job nature. Based on this criterion, a user just enters the value between 1 and 10, 1 being the least and 10 being the best. From all these data entered, the system will calculate the total the person has achieved and display it on the screen. The minimum total a user can get is between 60 and 70 to get the loan, but the problem with this system is that it does not take into consideration the applicant's qualifications and whether there is a possibility that in the near future the applicant might meet some of the criteria which they are not meeting now for example if this is a final year accounting student it means that in the next 2 years this particular applicant will get a job and the problem with the system is that it cannot infer when it comes to such situations [3].

MARBLE (Managing and Recommending Business Loan Evaluation) is a decision support system developed in California. It also utilizes a rule-based environment with an inductive inference engine [2]. It is more focused on the risk factor that comes with giving commercial businesses loans and thus does not cater for personal loans. MARBLE is a knowledge-based DSS that uses 80 decision rules for evaluating commercial loans. The MARBLE system was designed to use the lending judgment of experienced loan officers. The system was created in collaboration with the commercial bank of Chicago as the rules of how bank loan applications work can only be given by an expert from the bank.

Marble also has the capability to learn as this is an important aspect for artificial systems. In the study [2] there are two aspects in decision support tasks where learning comes into play: learning decision rules for the knowledge base, i.e., the knowledge-acquisition process and refining existing rules by observing prior problem-solving experience, i.e., the knowledge refinement process. To achieve these learning functions, MARBLE must be equipped with an inductive inference engine that is complementary to the deductive problem solver. This is an important design issue concerning the inductive inference technique formula learning and knowledge acquisition. It is a production rule system and is an inductive inference engine and with the capability to learn. The evaluation of loan application is by looking at the financial statement with qualitative information of the manager and the ability to repay the loan and collateral. In this system, the qualitative information weighs higher than the financial system [5].

A decision support system was developed in Nigeria for bank loan assessment. This system uses tree neuro-based model and eye disease diagnosis for the

development of the system, in other words, a new hybrid system was created using two models. This model allows for the development of a different type of rules for different types of borrowers [6]. In the first stage of the system, the bank customers or people applying for loans are segmented into clusters which have similar features. For the second stage, each group a decision trees is built to get the rules that are fed into the neural net for showing clients which are expected not to repay the loan or presence or absence of an eye disease. The system was created using C programming and an embedded MATLAB engine [5]. When the system was busy being developed there were two major parts. The first part is the decision tree that handles the basic decisions which are based on the fundamental decision rules of the system. Once the decision has been made the decision tree output becomes an input for the neural net which uses the result as a pad for the further refining which the system believes results in a much higher accuracy of decision-making [5].

The way the system functions is that users will select the action functions from the decision tree based on the lines given by the systems actions in complex organizational conditions. This condition present is built into a particular algorithm that decreases the complexity and computational effort required by the system on arriving at a result. The processing which is done by the neural net further refines the data based on the level of processing done by the decision tree and in the final stage the processing is taken further by the neural net which complete the final decision-making from the point where the complex decision stops [5].

A web-based intelligent system which focuses on Banking Applications developed in India. This system tries to be online so that its users can access it from anywhere via a thin client browser and over the Internet/Intranet [6]. Instead of a traditional decision support system, this system has an environment that supports centralized decision-making environment. The system using one particular technique like expert systems and neural networks because this technique can be integrated to reduce their weaknesses and to increase their strength and solve very complex tasks. The integration of analytical tools and intelligent systems make this system a powerful tool.

This system uses the Open XML technology and the power of decision support systems [6]. In this system, the web pages that use the open XML technology is called smart web pages. They deliver the information through dynamic web pages instead of static web pages. The system can control the flow of questions intelligently, by judging what user wants, based on user answers and the users must avoid asking irrelevant questions. An integrated shell environment was developed to deploy and run the server-based internet/intranet web applications which are rule-based, supports thin web clients, uses open standard technologies like XML (eXtensible Markup Language), XSL (eXtensible Stylesheet Language), JavaScript extensively, automatically generates feature-rich DHTML pages and has in-built fully configurable data access layer for data extraction, integration, transformation, and manipulation.

The system was developed by taking inputs from people from various banks as well as examining and collating input parameters from various applications forms used by the banks. The bank manager can select a specific group of input variable

which will be collected from the customer from their repository for a certain bank loan. From the variable and parameters selected a template form is generated where the customer details are entered and managed this is all done online and the customer can visit the bank to hand in supporting documents [6].

Backward chaining starts by the user having a list of goals and it works backward from the consequent to the antecedent to see if there is currently data available that will support the consequent. How the backward chaining method works is by searching the rules in the inference engine until it finds one which has a consequent that will match the desired goals. If the clause of that rule is not identified to be true, then it is added to the list of goals. Therefore, this will be implemented with an inference engine which would rather suit banks as this would mean that all the rules the bank uses to choose their successful candidates will be inserted into the systems and then the inference system can use those rules for complex analyzes [7]. Then deduce the successful candidates and also give reasons why the system things this candidate will not be able to pay back the loan and will not become a liability. All the rules which will be used will be gathered from a bank but most importantly from the person doing an assessment for bank loan applications as they will be a human expert in the particular field.

In this study, a knowledge-based support system is presented for a personal bank loan problem.

4 System Development

For the development of the system, MATLAB is used as a development tool. It is inference engine information will need to go through following steps.

Step 1: Determine the rules for the fuzzy system.
Step 2: Fuzzifying the inputs inserted by using membership function.
Step 3: Combine the inputs which were fuzzified to get the rule strength.
Step 4: Find the result of the rule by combining the rule strength with the output membership.
Step 5: Combine the results to get the output distribution.
Step 6: Defuzzify the output distribution.

The conceptional system for Namibia bank decides about the decision of loan based on three criteria which are; age, the number of dependents and the amount of capital available. The system uses the Mamdani Inference engine of MATLAB [8]. Membership functions expect the output to be fuzzy sets for this system and after the aggregation process, the fuzzy set for each output variable that needs to be de-fuzzified. It has been chosen as it is more suited for human inputs and easier to understand for users who do not have computer skills.

Table 1 Preliminary survey

Q. No.	Question	Relevance	Response
1	How much time is taken to handle a single bank loan application?	To know, either present system is time-consuming	Few months
2	How often a single aspect of the loan application is neglected?	To know whether default bank loans still occur	Frequently
3	Have they considered decision support system?	Ability to adopt the fuzzy-based expert system	Agreed

As shown in Table 1 the conceptional system was proposed, based on the three questions which were needed to be asked at the bank to find out whether an expert system for bank loan analysis is plausible and will be of advantage to the bank before the further study.

From preliminary investigation for Namibian bank, it was realized that an expert system for a bank loan analysis was needed to speed up the time it takes to handle an application since it is time-consuming and sometimes the human expert is not available and the banks are always looking at to new innovative techniques that can help the bank with their daily operations as banks are always looking at ways in which it can improve.

5 Results

The prototype is developed by following the methodology of fuzzy logic based expert system approach. The knowledge of the system will be represented by making rules as well as frames. In this preliminary study, sixteen set of rules has been framed to evaluate a loan application and inserted into the inference engine. Initially each rule assessment by some weight which the system is using when deciding the final outcome of the application. The rules do not have the same weights as some rules are more important when deciding on the final outcome of the application.

Figure 1 shows adjustments of the fuzzy regulator developed for the bank loan system in the editor of fuzzy inference systems, so quick decision can be taken using this fuzzy logic based expert system.

6 Conclusion and Future Work

Fuzzy logic-based expert intelligent system definitely has played the role in functional areas like in banking and finance. Literature survey covers some of well-known expert system applications in bank loan evaluation. There are many applications in banking where expert systems have been successfully implemented
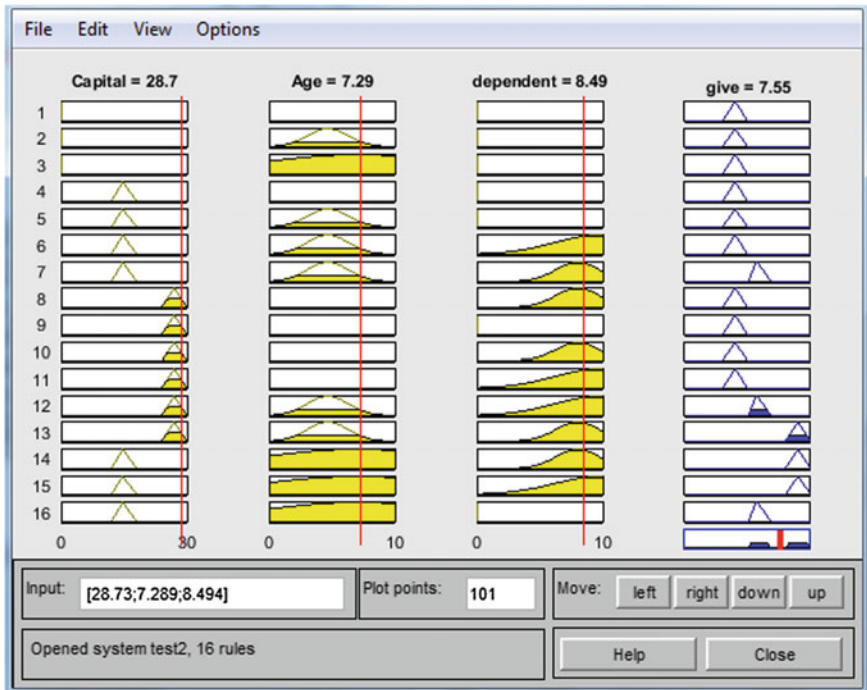
Fig. 1 Result of fuzzy inputs in the system

for the automating decision of various tasks. Even though the literature presents that researches are made in the field of banking and there is an expert system for banking sector but not for the loan appraisal for the individual. For the risk management, there is no perfect efficient expert system available for the personal loan appraisal in Namibia. Conceptual-based proposed system creates an inference engine which uses new fact in the knowledge base to trigger additional rules as this would make the system more robust and more efficient. In future, work review-based system will be extended to fully implement an expert system for the loan appraisal for individual in Namibian Banks.

References

1. L. Iliadis et al. (Eds.)., EANN/AIAI 2011, Part II, IFIP AICT 364, pp. 82–91 (2011).
2. Michael J. & Shaw James. (1987). A. Gentry Marble: A Decision-Support System for Business Loan Evaluation (1987). https://www.ideals.illinois.edu/bitstream/handle/2142/28792/marbledecisionsu1353shaw.pdf?sequence=1. (Accessed on 01.08.2015).
3. Wing, S.C., Knowledge-based support system for personal bank loan analysis, 346–351 (2015). http://staffweb.hkbu.edu.hk/vwschow/conf23.pdf. (Accessed on 7.08.2015).

4. Nils, J. Nilsson. Artificial Intelligence: A new synthesis. Morgan Kaufmann Publishers, INC. (1998).
5. Kabari, L.G, & Nwachukwu, E.O. (2013). Decision Support System Using Decision Tree and Neural Network (2013). http://www.researchgate.net/profile/Enoch_Nwachukwu/publication/258243184_Decision_Support_System_Using_Decision_Tree_and_Neural_Networks/links/02e7e52790790ed182000000.pdf. (Accessed on 07.7.2015).
6. Rajendra M Sonar, On Applications of Web-based Intelligent Systems: Focus: Banking (2005). http://www.som.iitb.ac.in/~rm_sonar/On%20Applications%20of%20Web-based%20Intelligent%20Systems-Focus%20Banking%20Applications.pdf.
7. Hatzilygeroudis, I., Prentzas, J.: New rules: Improving the Performance of Symbolic Rules. International Journal on AI Tools 9, 113–130 (2000).
8. Fuzzy Inference System Modeling, Build Mamdani, and Sugeno fuzzy inference systems http://www.mathworks.com/help/fuzzy/mamdani-fuzzy-inference-systems.html.

Author Index