N.R. Shetty · N.H. Prasad
N. Nalini *Editors*

Emerging Research in Computing, Information, Communication and Applications

ERCICA 2015, Volume 2

Springer

Emerging Research in Computing, Information, Communication and Applications

N.R. Shetty · N.H. Prasad · N. Nalini
Editors

Emerging Research in Computing, Information, Communication and Applications

ERCICA 2015, Volume 2

Springer

Editors
N.R. Shetty
Nitte Meenakshi Institute of Technology
Bangalore, Karnataka
India

N. Nalini
Nitte Meenakshi Institute of Technology
Bangalore, Karnataka
India

N.H. Prasad
Department of Master of Computer Application
Nitte Meenakshi Institute of Technology
Bangalore, Karnataka
India

ISBN 978-81-322-2552-2 ISBN 978-81-322-2553-9 (eBook)
DOI 10.1007/978-81-322-2553-9

Library of Congress Control Number: 2015943793

Springer New Delhi Heidelberg New York Dordrecht London

Printed on acid-free paper

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

Preface

The third International Conference on "Emerging Research in Computing, Information, Communication and Applications", ERCICA 2015 is an annual event organized at Nitte Meenakshi Institute of Technology (NMIT), Yelahanka, Bangalore, India.

ERCICA aims to provide an interdisciplinary forum for engineers and scientists to discuss and promote research and technology in the emerging areas of the core topics of Computing, Information, Communication, and their Applications. The conference brings researchers, educators, professional engineers and technologists under a single forum to discuss developments in research.

For ERCICA 2015, the organizers received more than 650 papers. All papers were reviewed thoroughly by experts of the Review Committee-ERCICA and the selected papers were chosen based on the quality of research and outcomes.

August 2015

N.H. Prasad
N. Nalini

Organizing Committee

ERCICA-2015

The Third International Conference on 'Emerging Research in Computing, Information, Communication and Applications' (ERCICA-2015) was held during 31 July–01 August, August 2015 at NMIT, Bangalore and organized by Department of CSE and MCA, NMIT.

Chief Patrons

Dr. N.V. Hegde, President, Nitte Education Trust
Dr. N.R. Shetty, Advisor, Nitte Education Trust

Conference Chair

Dr. H.C. Nagaraj, Principal, NMIT, Bangalore

Program Chairs

Dr. N.H. Prasad, HOD, MCA, NMIT, Bangalore
Dr. N. Nalini, Professor, CSE, NMIT, Bangalore

Publication Chair

Springer

Editorial Board

Prof. N.R. Shetty
(Former Vice-Chancellor Bangalore University and
Former President—ISTE, New Delhi and
Chairman—Karnataka State Universities Review Commission,
Bangalore, Karnataka)

Dr. N.H. Prasad
(Former Asst. Director-STPI, Ministry of C&IT, Government of India and
Former HOD–EdCIL (India) Ltd., Ministry of HRD, Government of India)
Professor and Head of the Department
Department of Master of Computer Applications and
Advisor—Foreign Students
Nitte Meenakshi Institute of Technology,
Yelahanka, Bangalore—560064

Dr. N. Nalini
Professor
CSE, NMIT
Bangalore

Dr. Kalidas Shetty
Associate Vice President
International Partnerships & Collaborations and
Professor, Department of Plant Science
North Dakota State University
214 Quentin Burdick Building
1320 Albrecht Blvd.
Fargo, ND 58108-6050 USA

Dr. Rajesh Kavasseri
Associate Professor/Graduate Coordinator
Department of Electrical and Computer Engineering
Room 101 H, Electrical Engineering
Dept. 2480, PO Box 6050
North Dakota State University, Fargo, ND 58108

Advisory Chairs

Dr. Jharna Majumdar, Dean (R&D), NMIT, Bangalore
Prof. Ranganatha Setty, Dean (Academic), NMIT, Bangalore

Advisory Committee

Dr. K.D. Nayak, Chief Controller, R&D, DRDO, India
Dr. Sathish Udupa, Dean and Professor, Michigan State University, Michigan
Dr. B.S. Sonde, Former Vice Chancellor, Goa University, Goa, India
Dr. D.K. Subramanian, Dean and Professor (Rtd.), Indian Institute of Science, Bangalore, India
Dr. L.M. Patnaik, Former Vice Chancellor, DIAT, Pune, India
Dr. K.R. Venugopal, Principal, UVCE, Bangalore, India
Dr. K. Chidananda Gowda, Former Vice Chancellor, Kuvempu University, Shimoga, India

Program Committee

Dr. K.G. Srinivas, Professor and Head—CSE, MSRIT, Bangalore, India
Dr. Jeevan Kumar, Professor-Physics, NMIT, Bangalore, India
Dr. Vittal S. Rao, National Science Foundation, USA
Dr. Rajkumar, Buyya-University of Melbourne, Australia.
Dr. Rajib, Mall-Indian Institute of Technology, Kharagpur, India.
Dr. P. Raveendran, University of Malaysia, Malaysia
Dr. K. Chandrasekaran, National Institute of Technology, Karnataka, India

Organizing Committee

Ms. Prathibha Ballal, Asst. Professor, CSE
Ms. Vijaya Shetty, Asst. Professor, CSE
Ms. Kavitha Sooda, Asst. Professor, CSE
Ms. Sujatha Joshi, Asst. Professor, CSE
Mr. Srinivas, Asst. Professor, CSE
Mr. Mohan, Asst. Professor, CSE
Ms. Geetha Priyadarshini, Asst. Professor, MCA
Ms. B.V. Shruthi, Asst. Professor, MCA
Ms. Deepthi J. Shetty, Asst. Professor, MCA

Ms. Ramya Shree, Asst. Professor, MCA
Ms. R. Uma, Asst. Professor, MCA
Mr. Mariyan Richard, Asst. Professor, MCA
Mr. Laxmi Narayana, Asst. Professor, MCA
Ms. Smriti Rai, Asst. Professor, MCA
Ms. S. Sushitha, Asst. Professor, MCA
Ms. K. VijayaLakshi, Asst. Professor, MCA
Mr. Navneeth, Asst. Professor, MCA
Ms. Sowmya Rai, Asst. Professor, MCA

About the Conference

ERCICA 2015

The Third International Conference on "Emerging Research in Computing, Information, Communication and Applications", ERCICA-2015 an annual event jointly organized by the Departments of CSE & MCA during 31 July–01 August, 2015 at Nitte Meenakshi Institute of Technology (NMIT), Yelahanka, Bangalore, India. The conference is supported by Technical Education Quality Improvement Program (TEQIP-II).

ERCICA-2015 is organized under the patronage of Prof. N.R. Shetty, Advisor, Nitte Education Trust. Dr. H.C. Nagaraj, Principal, NMIT served as the Conference Chair and the Program Chairs were represented by Dr. Prasad N. Hamsavath, Professor and Head, MCA and Dr. N. Nalini, Professor, CSE, NMIT, Bangalore.

ERCICA aims to provide an interdisciplinary forum for Researchers, Engineers, and Scientists to discuss and promote research and technology in the thrust areas of Computing, Information, Communication and Applications. The conference will bring researchers, educators, professional engineers and technologists into a single forum in order to discuss and debate on the emerging research in the above emerging areas.

For ERCICA 2016, authors are invited to submit the manuscripts of their original and unpublished research contributions to ercica.chair@gmail.com (ERCICA Website: http://nmit.ac.in/ercica.html). All the submitted papers will go through peer review process and the corresponding authors will be notified about the outcome of the review process. Authors of the selected papers can present their papers during the conference.

Acknowledgments

First of all, we would like to thank Prof. N.R. Shetty who has always been the guiding force behind this event's success. It was his dream that we have strived to make a reality. Our special thanks to Springer and especially the editorial staff who were patient, meticulous, and friendly with their constructive criticism on the quality of papers and outright rejection at times without compromising on the quality of the papers as they are always known for publishing the best international papers.

Finally, we would like to express our heartfelt gratitude and warmest thanks to the ERCICA 2015 Organizing Committee for all of the hard work and outstanding contributions as a member of the organizing committee. We know how much time and energy this assignment demanded, and we deeply appreciate all of your efforts to make it a grand success.

Our special thanks to all the authors who have contributed to publish their research work here in this conference and participated to make this conference a success and thanks to all of you for your exceptional work.

Regards,
Editorial Board
ERCICA 2015

Contents

Analysis and Visualization of Professional's LinkedIn Data

Puneet Garg, Rinkle Rani and Sumit Miglani

Abstract Social media has become very popular communication tool among internet users in the recent years. A large unstructured data is available for analysis on the social web. The data available on these sites have redundancies as users are free to enter the data according to their knowledge and interest. This data needs to be normalized before doing any analysis due to the presence of various redundancies in it. In this paper, LinkedIn data is extracted by using LinkedIn API and normalized by removing redundancies. Further, data is also normalized according to locations of LinkedIn connections using geo coordinates provided by Microsoft Bing. Then, clustering of this normalized data set is done according to job title, company names and geographic locations using Greedy, Hierarchical and K-Means clustering algorithms and clusters are visualized to have a better insight into them.

Keywords Data mining · Normalization · Greedy · Hierarchical · K-means clustering

1 Introduction

Social network is a web based service which allows users to create a public/semi-public or professional profile within a domain such that they can connect with other individuals within the network [1]. Now-a-days, social networking sites have become not only very popular but also a need for the new generations. These sites have eliminated the barriers of time, country, language, gender and money. Facebook, Twitter, LinkedIn, Google+ etc. are very popular social websites. Facebook has more than 1000 million active users. These users are increasing about 85 % annually since 2008 [2]. These days, the analytics are focusing to observe the

P. Garg (✉) · R. Rani · S. Miglani
Computer Science and Engineering Department, Thapar University,
Patiala 147004, India
e-mail: sonugargmmu@gmail.com

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_1

association and structure of various social networking sites. The main research area is to study and analysis of compactness, centrality and grouping with the promise of security and privacy of the users data in these networks. LinkedIn is a predominantly appealing and exciting objective, specifying the proficient environment of its subscribers. By investigating LinkedIn network information and data, one can understand the behaviour and other properties of individuals and group [3]. By collecting all the LinkedIn data i.e. occupation, employer, designation, locality etc., personality of the users can be exactly visualized. At first glance, LinkedIn appears like other social website, but, due to the responsive and professional characteristics, LinkedIn's data and its API are quite different from data available at other social network sites and have its own senses [11].

2 Clustering Techniques

Most of the time, an industry maintains its database to collect various types of information, but these datasets are not valid for solution of each and every problems. So, this data can't be assumed as universal datasets for all the problems e.g. there can be lacking in the design of the application's user interface, blank or over filled columns, abbreviations, misspelled etc. LinkedIn provides the text free service to the users to enter their information. To analyze the data sets in an effective manner, the data have to be normalized first, after that efficient clustering technique should be used [5]. Clustering is an unsupervised machine-learning method which is used to fastener in every data mining toolkits [12]. In this, all the information is first collected and then partitioned into a number of small clusters based on similarities in their properties. Various methods for clustering of the data sets can be legitimate as data mining tool kit [4]. The outcomes of the clustering depend upon the choice of clustering technique and type which is used. There are a number of clustering techniques available, but that technique should be adapted which gives the best solution for our problem [6]. There are some of the general similarity metrics which are very useful to compare the company or designation names such as Levenshtein Distance, N-gram similarity, jaccard distance etc. Now few real worlds LinkedIn data is accumulated and analyzed to obtain a deeper insight into the professionalism of this social network using various clustering algorithms such as greedy, hierarchical and k-means clustering.

2.1 Greedy Clustering

In Greedy clustering, a nested loop iterates over whole dataset and groups them together according to a threshold value defined by the similarity metric such as Jaccard Distance. If the Jaccard distance between two job titles is "close enough", they are greedily grouped together [3].

2.2 Hierarchical Clustering

It is a deterministic clustering technique in which it computes distances between all the items, stores them in a matrix, then traverses the whole matrix and clusters items that has a minimum threshold distance defined by the parameter such as Jaccard Distance. It is considered as hierarchical because it traverses over the whole matrix and clusters items together which creates a tree based on the relative distances between different items [7]. This is also called agglomerative clustering because it creates a tree by arranging each data element into several clusters, which hierarchically conflate into other top level clusters till the whole set of data clusters at the root of the tree. The leaves of the tree represent the data elements which are being clustered, whereas intermediate nodes hierarchically agglomerate these data items into clusters.

2.3 k-Means Clustering

Hierarchical clustering is an expensive technique having time complexity of the order $O(n^3)$, whereas k-means clustering normally executes with a time complexity of the order O(k*n). Even if the value of k is very small, the savings are very substantial. The savings in performance comes at the expense of results that are approximate, but they still have the potential to be quite efficient. The idea is that a multidimensional space containing n points is clustered into k clusters. Although k-means clustering can be run in two dimensional or two thousand dimensional space, the frequently used range is somewhere between tens of dimensions [3, 8]. When the dimension of the space is relatively small, this clustering technique can be effective because it runs fairly quickly and produces reasonable results. However we have to choose an appropriate value of k, but this is not always possible [9]. In k-means clustering, k points are picked randomly in the whole data space as initial seeds which are used to compute k clusters: K_1, K_2, K_3, …, K_k. Then each of n points present in data set is assigned to a cluster by searching for the nearest K_n. Then centroid of each cluster is calculated and K_i value of that cluster is reassigned to its centroid value. This procedure is repeated until the saturation comes and members of clusters become sTable between iterations. The name "K-means" is derived from the fact that K means are calculated in each iteration.

3 Dataset and Its Features

3.1 Source

The dataset in JSON and CSV format is obtained from LinkedIn API by exploring the API to get application's API key, Secret Key, OAuth User Token, and OAuth

Table 1 Description of Features

Features	Information
First name	First name of individual
Headline	Specifying position in particular company
Last name	Last name of the individual
Industry	Industry detail of the individual
Location name	Location details of work area
Country	Country code
Picture URL	URL of the profile picture
ID	Unique id provided by LinkedIn

User Secret identifications [11]. Table 1 shows various features present in the dataset obtained from LinkedIn API and corresponding description.

3.2 Data Normalization

LinkedIn provides the text free service to its users to enter their information. So, the users can fill the same information in thousands of ways. For example, if we want to collect all the members of "Thapar University" as their company or educational Institute, then our search should not only limited to "Thapar University", but also include "T.U." or "Thapar University, Patiala" or "Thapar Institute of Engineering & Technology" or "T.I.E.T." and many more in our search domain, because users are free to enter the text for their education institute name. So, to analyze the data set have to be normalized with considering various mind sets. The data is normalized by removing the redundancies present in job title name and company names. All the collective titles are split using a slash like convener/organizer and replaced with all the well known short forms. The data set is also normalized by removing redundancies present due to location of the users by getting geographic coordinates of various locations from Microsoft Bing. The final results have been obtained after process these data sets through frequency analysis as shown in Tables 2, 3 and 4.

Table 2 Normalizing company names and frequency count

Company	Frequency
Infosys	4
Amdocs	2
GENPACT	2
Tech Ahead	2
Self-employed	2
Smart data enterprises	2

Table 3 Normalizing job titles and frequency count

Job Title	Frequency
System engineer	5
Chief executive officer	2
Software engineer	2
Software developer	2
Subject matter expect	2
Business development	2

Table 4 Frequency count of tokens present in job titles

Token	Frequency
Engineer	21
Software	14
Developer	9
Senior	7
System	7
Executive	6
Manager	6
Associate	5
Analyst	4
Project	4
Assistant	4
Development	4
Business	3

4 Methodology

The methodology is described in Fig. 1. In the first step, the data set in JSON format is obtained from LinkedIn REST API [11]. To make data useful for further analysis, Normalization of data set is done in the second step. In third Step, similarity measurement and frequency analysis of job titles and company names are carried out. Clustering algorithms such as Greedy Clustering, Hierarchical Clustering and K-Means Clustering for clustering the data set according to different similarity criteria such as job title, company names and geographic coordinates are programmed in step 4. Finally, clusters are obtained and visualization of clusters for analysis is done. Geographic clusters obtained are visualized on Google Map [10].

5 Results and Discussion

Now LinkedIn data is clustered according to job titles by finding similarity between them using similarity metric Jaccard distance. The similar titles are clustered and displayed accordingly using greedy clustering technique as shown in Fig. 2.

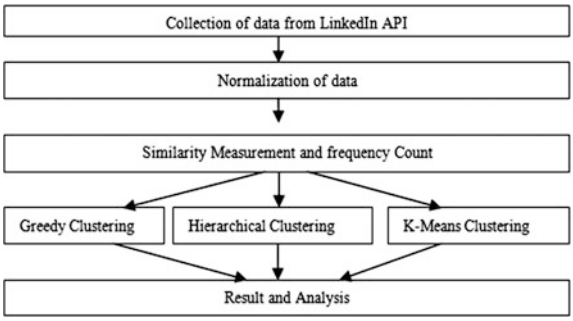


Fig. 1 Flowchart of methodology

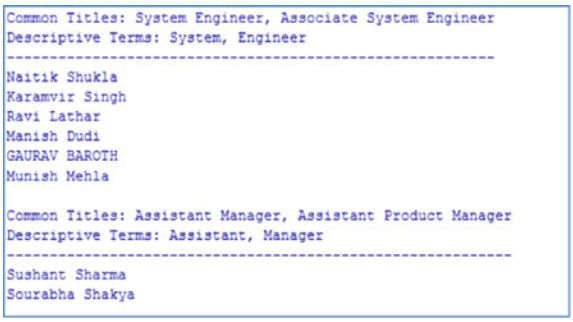


Fig. 2 Clustering connections according to job titles using greedy clustering

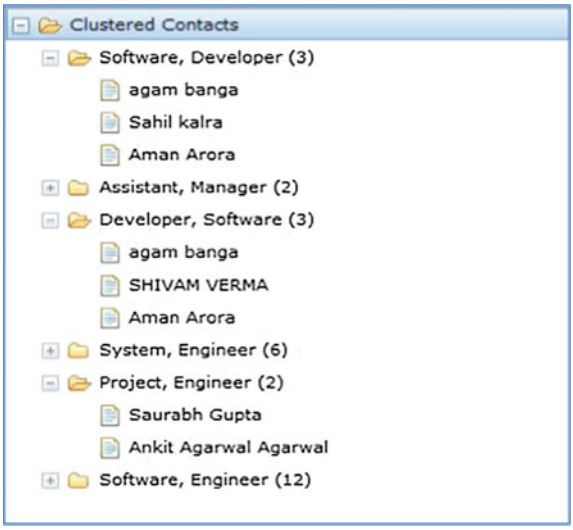


Fig. 3 Tree widget showing clusterd contacts by job title

Straightforward, efficient and easy clustering techniques and methods enhance user's experiences by influencing outcomes as smooth as the company name and designation name and title. A clustered content of data is represented in the Fig. 3, which shows a controlling option or point of the data sets using a simple and

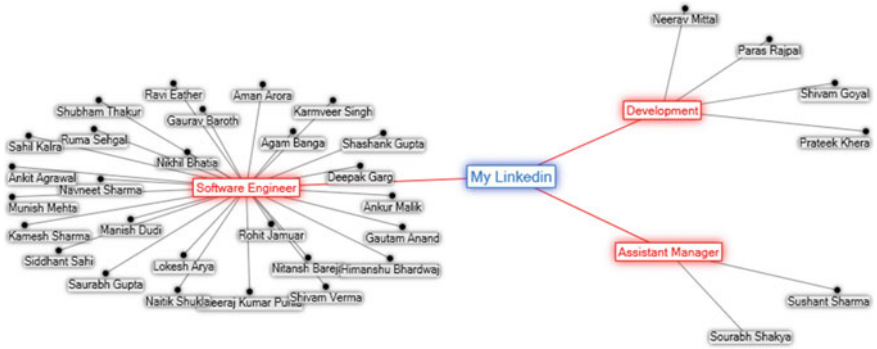


Fig. 4 A Network layout of contacts clustered by job title

efficient tree widget approach. This approach is useful to manage and filter data contents into sub parts as per various analyzing criterion.

A graph is generated by using Hierarchical clustering as shown in Fig. 4, the individual peoples are on the leaves of the tree, while the nodes such as "Assistant Manager" agglomerate all the leaf nodes into a cluster (small clusters having only one contacts are ignored in this graph).

Then data contents of LinkedIn connection are analyzed according to their geographical area. For this location of users is represented with geographical coordinates. An API key from Bing web service is requested to find out the geo-coordinates of every location. K-means clustering algorithm is used to cluster and visualize our professional's LinkedIn data and results are plotted on Goggle Map as shown in Fig. 5. After grouping our LinkedIn contacts according to geo-location, centroid of each cluster is also computed and plotted on the map as shown in Fig. 6.

The analysis of various clustering algorithms used for different criteria are shown below in the Table 5.

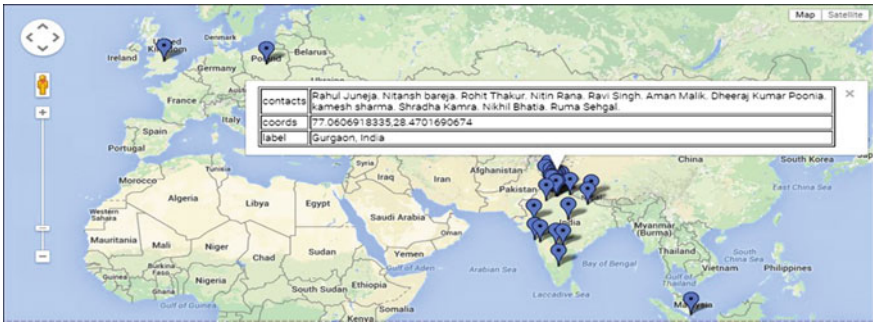


Fig. 5 Geographic visualization of clusters of LinkedIn connection

Fig. 6 Centroids of clusters computed by k-means

Table 5 Various clustering algorithms used for different criteria

Clustering techniques	Time complexity	Similarity criteria	Similarity metric	Clustered instances
Greedy	$O(n^2)$	Job title	Jaccard distance	Instances 255
				Clusters 9
Hierarchical	$O(n^3)$	Job title	Jaccard distance	Instances 255
				Clusters 7
K-Means	O(kn)	Geographic coordinates	Jaccard distance	Instances 255
				Clusters 3

6 Conclusion

We analyze LinkedIn data with different clustering techniques and address some common problems like normalization of messy LinkedIn data, and similarity computation. We also normalize geographic data present in our dataset using geo-coordinates and then proceed to geo clustering of LinkedIn connections and their visualization on Google Map. Geographic Cluster visualization can be utilized in some other problems such as solving the travelling salesman problem by different product based companies to reduce expenses occurred in supplying products to their customers. We can also explore LinkedIn extended profile using LinkedIn r_fullprofile API and can build a correlation between some aspects such as where people work, where they go to school and their relocation information. By mining LinkedIn data, we can analyze and visualize a lot of useful information about our professional contacts, which cannot be discovered manually. We can connect LinkedIn API with Twitter API to explore and measure the professionalism of our LinkedIn contacts by analyzing tweets made by them on twitter. We can also fetch information about various trendy topics in different professional communities by connecting LinkedIn API with Twitter API. By analyzing data from LinkedIn, we can also design a predictive system which recommends the preferred job for any person on the basis of his skill set and past experiences.

References

1. Olowe, M.A., Gaber, M.M., Stahl, F.: A survey of data mining techniques for social media analysis. J. Data Mining Digital Humanities (2014)
2. Ferrara, E., Meo, D.P., Fiumara, G., Baumagartner, R.: Web data extraction, applications and techniques: a survey. Knowled. Based Syst. **70**, 301–323 (2014)
3. Russell, M.A.: Mining the Social Web. O'Reilly Media
4. Revathi, S., Nalini, T.: Performance comparison of various clustering algorithms. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(2), 67–72 (2013)
5. Nagpal, A., Jatain, A., Gaur, D.: Review based on data clustering algorithm. In: Proceedings of IEEE conference on Information and communication Technologies (ICT), vol. 13, pp. 298–303 (2013)
6. Arora, S., Chana, I.: A survey of clustering techniques for big data analysis. In: Proceedings of IEEE 5th International Conference on Confluence the Next Generation Information Technology Summit (Confluence), vol. 14, pp. 59–65 (2014)
7. Srinivas, M., Mohan, C.K.: Efficient clustering approach using incremental and hierarchical clustering methods. In: Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN), 18–23 July 2010
8. Boutsidis, C., Zouzias, A., Mahoney, M.W., Drineas, P.: Randomized dimensionality reduction for k-means clustering. IEEE Trans. Inf. Theory **61**(2), 1045–1062 (2015)
9. Wilkin, G.A., Huang, X.: K-means clustering algorithms: implementation and comparison. Second Int. Multi Symp. Comp.Comput. Sci. IEEE **07**, 133–136 (2007)
10. Lu, C.A., Chen, C.H., Cheng, P.J.: Clustering and visualizing geographic data using geo-tree. In: Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, vol. 11, pp. 479–482 (2011)
11. https://developer.linkedin.com/
12. http://www.nltk.org/book/

Optimizing Websites for Online Customers

T.V. Vijay Kumar, Kumar Dilip and Santosh Kumar

Abstract The fast growth, along with the all encompassing presence, of the World Wide Web has given an unprecedented opportunity for organizations to maintain a strong online presence through a website catering to the requirements of varied users in an effective and efficient manner. In order to arrive at an optimal web site, relevant criteria need to be considered for selecting a set of web objects, from amongst a large number of web objects, which should be displayed on a web site. This being a combinatorial optimization problem would require simultaneous optimization of multiple relevant objectives based on relevant and key criteria for a given web site. In this paper, the multi-criteria web site optimization (*MCWSO*) problem, comprising of three criteria namely, download time, visualization score and product association level of web objects, has been addressed as a tri-objective optimization problem and solved using the vector evaluated genetic algorithm (*VEGA*). Experimental results show that the *VEGA* based *MCWSO* algorithm, in comparison to the *GA* based *MCWSO* algorithm, is able to select comparatively better web object sequences for a web site.

Keywords Usage mining · Web site adaptation · Genetic algorithm · Multi-objective genetic algorithm · VEGA

1 Introduction

The almost ubiquitous availability of the World Wide Web has given an unprecedented opportunity to many organizations to maintain an effective online presence in order to create a large user base, which would enable them to accomplish their business goals. This would necessitate the designing of an effective and efficient web site, which is capable of serving large number of users in the best possible

T.V. Vijay Kumar (✉) · K. Dilip · S. Kumar
School of Compter and Systems Sciences, Jawaharlal Nehru University,
110067 New Delhi, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_2

manner. Such web sites should have the potential to convert visitors into users even while retaining its existing users. This would enable them to sustain themselves and be competitive in the online environment. This task can be achieved by providing an effective and efficient browsing experience to its visitors. Designing such a web site poses many challenges due to changing navigational behavior of users, since every user visits a site for different purposes and the same user may visit the web site with different goals at different times [16]. No matter how carefully or judiciously a web site is designed, it may still not adequately cater to the needs and requirements of users due to highly unpredictable nature of user's future navigational behavior. In [16], adaptive web site design is suggested as an alternative to address these issues. An adaptive web site is one that is capable of automatically improving its organization and presentation based on access patterns exhibited by the user's navigational behavior. There are two approaches to web site adaptation namely customization based web site adaptation and transformation based web site adaptation [13]. The former approach mainly targets individual users, using their past navigation browsing behavior, based on which it speculates their possible future moves or preferences, and accordingly customizes the web page. In the latter approach, partial or, in some cases, almost total modification of the web site structure is done in order to cater to the browsing behavior of large number of users. This paper focuses on the transformation based web site adaptation, where there is a need to extract information from the web log data, comprising user-web interaction data, to understand and analyze the user navigation behavior. This extracted information can be used to improve the web site structure, which in turn would result in the improvement of the overall performance of the web site. Web site structure organization, which primarily aims to provide its users with easier navigation, has been an active area of research [1, 3, 7, 12, 13, 24]. In order to make the navigation easier and simpler for a user, several aspects or criteria need to be taken into consideration while optimizing the structure of a web site. These may include navigation, response time, credibility, content, usability, organization, interactivity, download time, presentation, visualization, product association level, number of clicks, search depth, information overload, web site links, access patterns, cost etc. [1–3, 7, 13–15, 17, 20, 22, 28]. The relevance and importance of these criteria may vary for different web sites. The relevant criteria need to be considered for selecting a set of web objects, from amongst a large number of web objects, which should be displayed on the web site. This being a combinatorial optimization problem would require simultaneous optimization of multiple objectives based on the relevant and key criteria for a given web site. Several approaches exist in literature [1, 3, 7, 13, 22, 24, 28] that address this problem as a multi-criteria web site optimization problem. In this paper, multi-criteria web site optimization problem, given in [1], has been addressed as a tri-objective optimization problem with the three objectives being minimization of download time, maximization of web site visualization and maximization of product association level. This tri-objective optimization problem is addressed using the vector evaluated genetic algorithm (*VEGA*) [19], which is a multi-objective genetic algorithm.

This paper is organized as follows: Sect. 2 briefly discusses web site optimization followed by the multi-criteria web site optimization (*MCWSO*) problem in Sect. 3. *MCWSO* using *VEGA* is discussed in Sect. 4 followed by an example based on it in Sect. 5. Section 6 discusses experimental results. The conclusion is given in Sect. 7.

2 Website Optimization

The galloping growth of the World Wide Web has eased our daily life by offering almost all sorts of services; while on the other hand, it has posed major challenges to the web designers. The biggest challenge for them is to provide a website that caters to a large number of users having varied interests and having dynamic navigation behavior. The goal is to provide users with a smooth and satisfying browsing experience on the web. Many attempts have been made in this direction by using the knowledge stored in the weblog data, and by considering the various criteria and constraints [1, 7, 12–15, 23, 25, 28], with the aim of improving the overall website structure and simplifying access to the web. In [7], user access pattern based website reorganization, that evolves the site into an adaptive web site, to cater to the user's requirements efficiently despite their changing navigation behaviors, has been proposed. This approach comprises three main steps namely, pre-processing, page classification and site reorganization. In [28], an approach for website link structure improvement based on the user visiting pattern, that optimizes and re-evaluates the link structure, in order to improve the average connectivity has been discussed. While improving the website structure several criteria are considered like the download time, product association level and visualization score [1]. In [27], design features for website optimization, based on empirical studies of the users perception, has been discussed. These design features include easier navigation, up-to-date information, visual design, multimedia, site responsiveness, search tool, comprehensiveness of information, security of information, accuracy of data etc. Among these features, the easier navigation is considered as one of the most important feature for an effective website design. Low download time and high website visualization helps in attracting users' attention. The up-to-date information makes the content more interesting and reliable. More comprehensive information makes the user stay longer and ease the navigation. Availability of the search tool facility provides for quick search of the desired items on a complex website. A quick responsive site can easily engage a visitor and can make the navigation smooth. Comprehensibility of information, and accuracy of data, will be helpful in increasing the interest and trust in the website and may result in the visitor revisiting the site. While optimizing the website design, there are many constraints that need to be considered in practice. These may include maximum out-degree of a page, maximum depth of the web site, overall site connectivity, basic links maintenance, page classification, compatible links and the security [7, 13, 22, 25, 26]. Fixing the maximum number of out-degree or links

from a web page would help in reducing the information overload. Keeping the website depth minimum would help in finding the desired information, for a user, in lesser number of clicks resulting in user satisfaction. Overall site connectivity constraints entail the provisioning of a path from the home page to the rest of the pages of the website. The basic links structure should remain intact subsequent to the web site structure optimization.

Approaches applied to achieve an optimal website are mostly based on, either the mathematical programming models [3, 12, 22, 24] or the metaheuristic based approaches [1, 10, 13, 18, 23, 26]. In [24], web site optimization has been considered as a structural optimization problem and a framework, which maps the website design problem into a graph model, has been proposed to which mathematical optimization has been applied for arriving at an optimal web design. In [22], a mathematical programming model has been proposed for improved user navigation efficiency, catering to both experienced and new users. With the intent of assisting existing users and site designers, structure stability constraint has been discussed in this approach. In [3], a mathematical programming model has been proposed for website structure enhancement, while minimizing the cost of user disorientation due to complete reorganization of the website. This approach advocates structure re-organization by making minimal changes to its existing structure. In [12], a 0-1 programming model has been proposed for website reorganization based on information overload and search depth. In order to reduce the computation time, a two stage heuristic has been discussed in this paper. This approach faces scalability issues for websites having large number of links.

The ability of metaheuristic based methods to provide good solutions in reasonable amounts of time makes them suitable for addressing website structure optimization problems. In [10], a simulated annealing based approach for website structure optimization, considering webpage relinking, has been proposed. This approach utilizes the combined user preference data for website structure improvement for easier navigation. In [13], an ant colony optimization based website structure reorganization using two criteria namely, information overload and search depth, has been proposed for effective and easier navigation. This approach has been shown to take less computation time for relatively small synthetic and real websites. In [23], a mathematical model with the objective being the minimization of the weighted distance between web pages for e-commerce website, has been proposed. A statistical Hopfield neural-network, and the strategic oscillation-based tabu-search algorithms, were applied to optimize the link structure of the web pages. In [18], a quadratic assignment problem based method has been applied to improve the link structure and navigation of the website. The degree of connectivity and distance between the web pages has been used in this approach, and an ant colony metaheuristic based technique is applied to solve this optimization problem. In [26], an enhanced tabu search based technique for website structure optimization, using web usage data, has been proposed. This approach allows a cyclic path and derives the bi-directional association between two pages. It also discussed the notion of compatible link constraints, in order to maintain the business logic by preventing the links between designated pages. In [1], a

multi-criteria website optimization (MCWSO) approach comprising three criteria namely, download time, visualization score and product association level has been proposed. This paper addresses this MCWSO problem, which is discussed next.

3 The *MCWSO* Problem

In order to achieve an effective and efficient web site that has the potential to convert visitors to users, there is a need to consider multiple relevant and key criteria to select an optimum sequence of distinct web objects for a web site with the aim of improving the browsing experience of a user. In [1], three such criteria namely product association level, download time and visualization have been considered for web site optimization. For an effective and efficient web site, the primary aims would be to minimize the download time (efficient access to web items), maximize the web site visualization (improve the look and feel of web site) and maximize the product association levels (increase the overall sale of items). This problem, considering these three criteria, has been formulated as a single objective optimization problem as given below:

$$F = DT * w1 + (1 - VS) * w2 + (1 - PS) * w3$$

Where

$$DT = \frac{\sum_{k=1}^{M} D(k)}{(Max_{k=1}^{N} D(k)) * M},\ VS = \frac{\sum_{k=1}^{M} V(k)}{(Max_{k=1}^{N} V(k)) * M} \text{ and } PS = \frac{\sum_{k=1}^{M} P(k)}{(Max_{k=1}^{N} P(k)) * M}$$

where DT is the download time, VS is the visualization score, PS is the probability of selling an object after accessing the previous object, $D(k)$ is the download time of the kth web-object in the sequence, $V(k)$ is visualization score of the kth web-object, $P(k)$ is the probability that the product or service represented by the kth web-object will be sold if followed by the $(k-1)$th web-object in the sequence, N is the total number of web-objects and M is the number of web-objects appearing in the sequence under consideration i.e. a sequence contains any M web-objects out of the given N objects. The objective is to minimize F.

The above-mentioned single objective optimization problem has been addressed using the genetic algorithm in [1]. The choice of weights, which requires profound knowledge of the domain and user preferences, is the major limitation of the above aggregate weighted sum approach [21]. Also, addressing a tri-objective optimization problem, as a single objective optimization problem, does not appropriately reflect the real life web site optimization problem. In this paper, an attempt has been made to address this problem as a tri-objective optimization problem with the three objectives being Minimize Download Time, Maximize Web Site Visualization and Maximize Product Association level i.e.

$$Minimize\ DT = \frac{\sum_{k=1}^{M} D(k)}{(Max_{k=1}^{N} D(k)) * M},\ Maximize\ VS = \frac{\sum_{k=1}^{M} V(k)}{(Max_{k=1}^{N} V(k)) * M} \text{ and } Maximize\ PS = \frac{\sum_{k=1}^{M} P(k)}{(Max_{k=1}^{N} P(k)) * M}$$

The objective would be to select a sequence of distinct web objects, from amongst all possible web objects, based on the simultaneous optimization of the above three objectives. The inherent nature of this problem makes it a combinatorial optimization problem and metaheuristic techniques have been widely used for solving such problems. In this paper, a multi-objective genetic algorithm *VEGA* has been used to solve this multi-criteria web site optimization problem. The use of *VEGA* addresses the major limitation, i.e. the choice of weights, associated with the aggregate weighted sum approach used in [1], where the weights are not associated with the objectives. *MCWSO* using *VEGA* is discussed next.

4 *MCWSO* Using *Vega*

The *MCWSO* problem [1] and briefly discussed above, requires the simultaneous optimization of three objectives namely Minimize Download Time, Maximize Web Site Visualization and Maximize Product Association Levels for selecting web object sequences (*wos*), comprising of web objects, that would result in the sale of product. For this multi-objective genetic algorithm *VEGA* is used, which is discussed next.

4.1 VEGA

VEGA is considered to be the first multi-objective genetic algorithm [4, 6]. *VEGA* evaluates an objective vector, instead of the scalar objective function used in the aggregated weighted sum approach, against each element of the vector representing each objective function [4, 6]. *VEGA*, which is a simple extension of GA, modifies the selection operator of GA to solve the multi-objective optimization problem.

In *VEGA* [4, 6, 19], for a problem comprising *M* objectives, the population of N individuals ($I_1, I_2, ..., I_N$) is divided into *M* sub-populations ($SP_1, SP_2, ..., SP_M$), each of size *N/M*. Each sub-population considers only one of the M-objectives for computing the fitness of the individuals in it. Thereafter, the proportionate selection operator is applied to select individuals in each sub-population to generate a corresponding selected sub-population ($SSP_1, SSP_2, ..., SSP_M$) having size N/M. These *M* selected sub-populations, comprising individuals for mating, are shuffled together to obtain a population of mating individuals ($MI_1, MI_2, ..., MI_N$) having size *N*. This is followed by applying the crossover and mutation operators, as carried out

in traditional *GA*, on the population of mating individuals to produce a population of N individuals ($I_1, I_2, ..., I_N$) for the next generation. *VEGA* has been used to solve the multi-criteria web site optimization problem discussed above. The proposed *VEGA* based *MCWSO* algorithm ($MCWSO_{VEGA}$) is discussed next.

4.2 $MCWSO_{VEGA}$

In order to address the *MCWSO* problem discussed above, there is a need to select optimal sequences of distinct web objects with respect to criteria download time, web site visualization and product association levels. The proposed *VEGA* based *MCWSO* algorithm ($MCWSO_{VEGA}$) attempts to generate such web object sequences. Algorithm $MCWSO_{VEGA}$ is given in Fig. 1.

As per algorithm $MCWSO_{VEGA}$, initially, a population of web object sequences P_{wos} is randomly generated. Each *wos* is comprised of distinct objects (M), each randomly chosen from amongst all the web objects (N). Since in the web site optimization problem under consideration, three criteria's are used, i.e. $|C| = 3$, the population P_{wos} is divided into $|C|$ sub-populations SP_{wos}, each of size $P_{wos}/|C|$. Each sub-population SP_{wos} uses only one of the criteria or objectives for evaluating the wos in it. This is followed by applying proportionate selection [9], based on the computed objective values of wos in each sub-population, to generate the corresponding mating sub-population of *wos* MSP_{wos}. Thereafter, crossover and mutation is applied to produce a new population P_{wos} of *wos* for the next generation. One of the major limitations of VEGA is the issue of speciation [8], i.e. it emphasizes on individuals that are good for individual objective functions. As a result, it is unlikely to find trade-off solutions that take into consideration all the objectives.

INPUT:
N: Total Number of web objects
M: Number of web objects appearing in a web object sequence (*wos*)
P_c: Probability of Crossover
P_m: Probability of mutation
$C=\{C_1, C_2, C_3\}$ are the criteria's considered for web site optimization where

$$C_1 = DT = \frac{\sum_{k=1}^{M} D(k)}{(Max_{k=1}^{N} D(k)) * M} \ , \ C_2 = VS = \frac{\sum_{k=1}^{M} V(k)}{(Max_{k=1}^{N} V(k)) * M} \ , C_3 = PS = \frac{\sum_{k=1}^{M} P(k)}{(Max_{k=1}^{N} P(k)) * M}$$

where *D(k):* download time of the k^{th} web-object in the sequence, *V(k):* visualization score of the k^{th} web-object, *P(k):* probability that the product or service represented by the k^{th} web-object will be sold if followed by the $(k-1)^{th}$ web-object in the sequence

OUTPUT:

$WOS_{Top\text{-}K}$: Top-K web object sequences

METHOD:

Step 1: Randomly generate an initial population of web object sequences P_{wos} of size $PSize_{wos}$
REPEAT
Step 2: Divide the population of web object sequences P_{wos} into $|C|$ sub-populations SP_{wos}, each of size $PSize_{wos} / |C|$
Step 3: For each sub-population $SP_{wos}(i)$ in SP_{wos}
- Compute C_i of each *wos* in $SP_{wos}(i)$
- Perform proportionate selection on *wos* in $SP_{wos}(i)$ to create a mating sub-population $MSP_{wos}(i)$

Step 4: Perform crossover with probability P_c and mutation with probability P_m to produce a new population of web object sequences P_{wos}
- *single point order crossover* is performed where one individual in each mating sub-population is considered and single point order crossover is performed between them to produce offspring individual.
- *swap mutation* is performed where two web objects in a *wos* are randomly chosen and thereafter swapped

UNTIL (stopping condition is satisfied)
RETURN $WOS_{Top\text{-}K}$

Fig. 1 Algorithm $MCWSO_{VEGA}$

In $MCWSO_{VEGA}$, an attempt has been made to address this problem by considering one *wos* from each sub-population and perform crossover between them to produce an off-spring *wos*. Since there is a need to generate a sequence of distinct web objects, single point order crossover [5] is used. Next, the offspring *wos* produced after the crossover undergoes mutation. Again, since the web objects in *wos* are required to be distinct, swap mutation [11] is used.

After crossover and mutation, the new population of *wos* P_{wos} is produced for the next generation. The algorithm runs till the stopping condition is met. The stopping condition can be the pre-specified number of generations for which the algorithms run or the stagnation condition is reached, i.e. there is not much improvement in the solution for a pre-specified number of generations. Thereafter, the algorithm $MCWSO_{VEGA}$ produces the *Top-K* web object sequences $WOS_{Top\text{-}K}$ as output.

An example illustrating the use of $MCWSO_{VEGA}$ to generate web object sequences, which is based on minimization of download time, maximization of web site visualization and maximization of product association levels is discussed next.

5 An Example

Suppose there are ten items to be sold through the website. Each of these items is represented by a web-object (*A, B, C, D, E, F, G, H, I, J*). The download time (*D*), visualization score (*V*) and probability P_{ij} (of buying the *j*th item just after visiting the *i*th item) for each of these web-objects are given in Fig. 2a. First, the population of web object sequences P_{wos} are randomly generated and are shown in Fig. 2b. Next, the population P_{wos} is divided into three sub-populations SP_{wos}*-1*, SP_{wos}*-2* and SP_{wos}*-3*, each containing four web object sequences. These are given in Fig. 2c.

Next, the fitness values of *wos* in each of the sub-populations are computed. For wos in SP_{wos}*-1*, SP_{wos}*-2* and SP_{wos}*-3*, the fitness is computed using *DT*, *VS* and *PS* respectively and are shown in Fig. 3. Next, proportionate Roulette Wheel selection is applied to each of the sub-population of *wos* SP_{wos}*-1*, SP_{wos}*-2* and SP_{wos}*-3* to select the corresponding mating sub-population of *wos* MSP_{wos}*-1*, MSP_{wos}*-2* and MSP_{wos}*-3*, given in Fig. 4a. Single point order crossover, as discussed before, with crossover probability $P_c = 0.75$ is performed between wos in each of the mating sub-population MSP_{wos}*-1*, MSP_{wos}*-2* and MSP_{wos}*-3* to generate offspring wos. This is followed by performing swap mutation with mutation probability $P_m = 0.05$. The new population of *wos* P_{wos} produced after crossover and mutation are given in Fig. 4b. The above process is repeated for a pre-specified number of generations whereafter the *Top-K wos* is produced as output.

In order to compare the quality of *wos* produced by $MCWSO_{VEGA}$ and the *GA* based *MCWSO* algorithm $MCWSO_{GA}$ [1], experiments were performed. The results and observations based on these experiments are discussed next.

(a)

Web Object		A	B	C	D	E	F	G	H	I	J
Download Time (*D*)		3	12	2	9	14	8	12	17	19	5
Visualization Score (*V*)		3	3	1	5	4	2	3	4	4	5
Probability (P_{ij})	A	0.01	0.17	0.82	0.20	0.23	0.84	0.97	0.74	0.02	0.61
	B	0.50	0.68	0.67	0.65	0.41	0.51	0.54	0.08	0.13	0.46
	C	0.61	0.81	0.45	0.26	0.98	0.84	0.87	0.75	0.12	0.75
	D	0.19	0.74	0.85	0.34	0.71	0.29	0.13	0.22	0.35	0.15
	E	0.84	0.83	0.84	0.19	0.75	0.46	0.13	0.29	0.11	0.60
	F	0.20	0.97	0.81	0.81	0.61	0.05	0.40	0.43	0.27	0.76
	G	0.67	0.96	0.49	0.17	0.32	0.76	0.88	0.71	0.33	0.26
	H	0.08	0.13	0.89	0.74	0.93	0.84	0.49	0.40	0.30	0.32
	I	0.21	0.73	0.23	0.67	0.95	0.09	0.59	0.96	0.03	0.49
	J	0.47	0.19	0.05	0.52	0.57	0.96	0.57	0.00	0.16	0.89

(b)

C A G E I	B A G F E
B A G E D	C A G I H
A G F D H	C B G E H
B F E D H	G F E D I
C A F D I	C B G E D
G F D I H	A G E I H

(c)

SP_{wos} - 1	SP_{wos} - 2	SP_{wos} - 3
C A G I H	B A G F E	C B G E H
C B G E D	G F E D I	A G E I H
C A G E I	A G F D H	B A G E D
B F E D H	G F D I H	C A F D I

Fig. 2 **a** D, V and P_{ij} of web objects. **b** Sub-population of wos $SP_{wos\text{-}1}$, $SP_{wos\text{-}2}$ and $SP_{wos\text{-}3}$

SP_{wos} - 1	$DT = \frac{\sum_{k=1}^{M} D(k)}{(Max_{k=1}^{N} D(k)) * M}$	SP_{wos} - 2	$VS = \frac{\sum_{k=1}^{M} V(k)}{(Max_{k=1}^{N} V(k)) * M}$	SP_{wos} - 3	$PS = \frac{\sum_{k=1}^{M} P(k)}{(Max_{k=1}^{N} P(k)) * M}$
C A G I H	(2+3+12+19+17)/(19×5) = 0.56	B A G F E	(3+3+3+2+4)/(5×5) = 0.60	C B G E H	(0.45+0.81+0.54+0.32+0.29)/(0.98×5) = 0.49
C B G E D	(2+12+12+14+9)/(19×5) = 0.52	G F E D I	(3+2+4+5+4)/(5×5) = 0.72	A G E I H	(0.01+0.97+0.32+0.11+0.96)/(0.98×5) = 0.48
C A G E I	(2+3+12+14+19)/(19×5) = 0.53	A G F D H	(3+3+2+5+4)/(5×5) = 0.68	B A G E D	(0.68+0.50+0.97+0.32+0.19)/(0.98×5) = 0.54
B F E D H	(12+8+14+9+17)/(19×5) = 0.63	G F D I H	(3+2+5+4+4)/(5×5) = 0.72	C A F D I	(0.45+0.61+0.84+0.81+0.35)/(0.98×5) = 0.62

Fig. 3 *DT*, *VS* and *PS* of *wos* in SP_{wos}-*1*, SP_{wos}-*2* and SP_{wos}-*3* respectively

(a)

MSP_{wos} - 1	MSP_{wos} - 2	MSP_{wos} - 3
C B G E D	G F E D I	C A F D I
C A G I H	B A G F E	B A G E D
C A G I H	G F D I H	C A F D I
C B G E D	G F E D I	C B G E H

(b)

C A G E D	C B E I H
G F D I H	B A G F E
B F E I H	C A E I H
C A E I H	C A G E I
C A G E D	G F E D I
B A G E D	C A G E I

Fig. 4 **a** Mating sub-population of *wos* $MSP_{wos\text{-}1}$, $MSP_{wos\text{-}2}$ and $MSP_{wos\text{-}3}$. **b** Population of *wos* P_{wos} for the next generation

6 Experimental Results

Algorithms $MCWSO_{VEGA}$ and $MCWSO_{GA}$ were implemented using JDK 1.6 in a Windows-7 environment. These algorithms were compared by conducting experiments on an Intel based 2.13 GHz PC having 4 GB RAM. First, in order to ascertain the performance of $MCWSO_{VEGA}$ and $MCWSO_{GA}$ for different crossover rate (*CR*) and mutation rate (*MR*), a graph showing the average value of fitness F, for weights $w1 = 1$, $w2 = 1$ and $w3 = 1$, against the number of generations for displaying 5, 10, 15 and 20 web objects were plotted. These graphs for $MCWSO_{VEGA}$ ($VEGA(CR, MR)$) and $MCWSO_{GA}$ ($GA(CR, MR)$), where $CR = \{0.6, 0.8\}$ and $MR = \{0.05, 0.1\}$, for displaying 5, 10, 15 and 20 web objects are shown in Fig. 5. It can be inferred from these graphs that for every *CR* and *MR*,

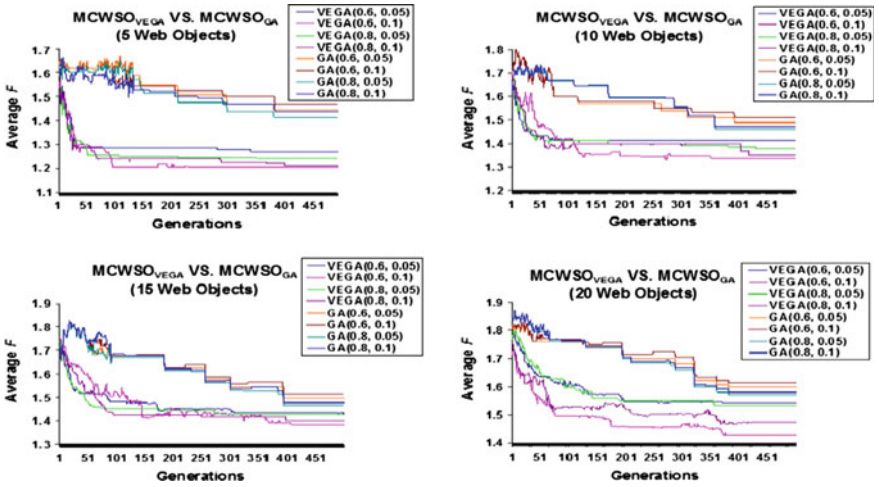


Fig. 5 MCWSO$_{VEGA}$ versus MCWSO$_{GA}$ (Average *F* Vs. Generations of *Top-10 wos*—5, 10, 15 and 20 Web Objects)

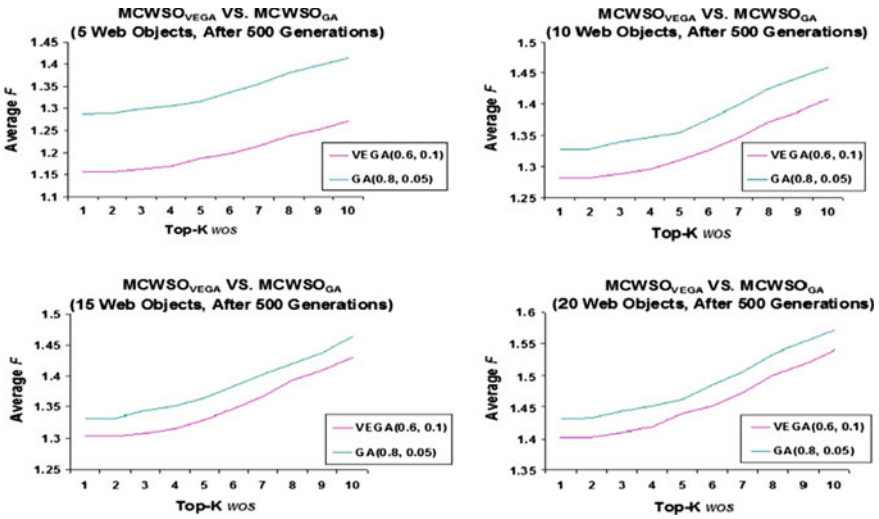


Fig. 6 MCWSO$_{VEGA}$ versus MCWSO$_{GA}$ (Average *F* Vs. *Top-K wos* after 500 Generations—5, 10, 15 and 20 Web Objects)

MCWSO$_{VEGA}$, in comparison to *MCWSO$_{GA}$*, converges to a comparatively lower average fitness value *F* of *Top-10* wos generated by them. Further, *MCWSO$_{VEGA}$* performs best for *CR* = 0.6 and *MR* = 0.1 and *MCWSO*$_{GA}$ performs best for *CR* = 0.8 and *MR* = 0.05.

Next, graphs to compare the average fitness *F* of *Top-K wos* generated by *MCWSO$_{VEGA}$* and *MCWSO$_{GA}$* for the observed crossover rates and mutation rates

$\{CR = 0.6, MR = 0.1\}$ and $\{CR = 0.8, MR = 0.05\}$ respectively for 5, 10, 15 and 20 web objects were plotted and are shown in Fig. 6. It can be observed from each of these graphs that the *Top-K wos* generated by $MCWSO_{VEGA}$ have a lower average fitness F than those generated by $MCWSO_{GA}$. This may be attributed to the fact that $MCWSO_{VEGA}$ simultaneously optimizes the three objective and produces offsprings resulting from crossover performed between *wos*, one each from each of the sub-population concerned with the three objectives.

7 Conclusion

A *VEGA* based *MCWSO* algorithm $MCWSO_{VEGA}$ has been proposed in this paper that generates optimal web object sequences for a given web site based on three criteria i.e. the download time, their visualization and the product association level of web objects. Unlike $MCWSO_{GA}$, $MCWSO_{VEGA}$ does not require assignment of weights to these three criteria. $MCWSO_{VEGA}$ selects web object sequences, from amongst all possible web object sequences, based on simultaneous optimization of the three objectives, namely the minimization of download time, the maximization of web site visualization and the maximization of product association levels. Further, the crossover operation has been adapted to ensure that the population produced after every generation comprises of web object sequences that reflects the three objectives. As a result, $MCWSO_{VEGA}$ is able to generate web object sequences that have an acceptable trade-off between the three objectives. Experimental results show that the $MCWSO_{VEGA}$ performs better than the $MCWSO_{GA}$ with regard to generating comparatively better web object sequences for a given web site. This would improve the browsing experience of the user and motivate the user to buy products online.

References

1. Asllani, A., Lari, A.: Using genetic algorithm for dynamic and multiple criteria web-site optimizations. Eur. J. Oper. Res. **176**(3), 1767–1777 (2007)
2. Boling, E.: Usability testing for web sites learning for the global community. In: Seventh Annual Hypermedia '95 Conference, 1995
3. Chen, M., Ryu, Y.U.: Facilitating effective user navigation through website structure improvement. IEEE Trans. Knowl. Data Eng. **25**(3), 571–588 (2013)
4. Coello-Coello, C.A., Lamont, G.B., van Veldhuizen, D.A.: Evolutionary Algorithm for Solving Multi-objective Problems. Genetic and Evolutionary Computation Series, 2nd edn. Springer, Berlin (2007)
5. Davis, L.: Applying adaptive algorithms to epistatic domains. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 162–164, 1985
6. Deb, K.: Multi-objective optimization using Evolutionary algorithms. Wiley, India (2010)
7. Fu, Y., Shih, M.Y., Creado, M., Ju, C.: Reorganizing web sites based on user access patterns. Intelligent systems in accounting. Finance Manage. **11**(1), 39–53 (2002)

8. Ghosh, A., Dehuri, S.: Evolutionary algorithms for multi-criterion optimization: a survey. Int. J. Comput. Inf. Sci. **2**(1), 38–57 (2004)
9. Goldberg, D.E.: Genetic Algorithm in Search, Optimization and Machine Learning. Pearson Education, India (1989)
10. Gupta, R., Bagchi, A., Sarkar, S.: Improving linkage of Web pages. INFORMS J. Comput. **19** (1), 127–136 (2007)
11. Kazarlis, S.A., Bakirtzis, A.G., Petridis, V.: A genetic algorithm solution to the unit commitment problem. IEEE Trans. Power Syst. **11**(1), 82–92 (1996)
12. Lin, C.C.: Optimal Web site reorganization considering information overload and search depth. Eur. J. Oper. Res. **173**(3), 839–848 (2006)
13. Lin, C.C., Tseng, L.C.: Website reorganization using an ant colony system. Expert Syst. Appl. **37**(12), 7598–7605 (2010)
14. Nielsen, J.: Designing Web Usability: The Practice of Simplicity. New Riders Publishing (2000)
15. Palmer, J.W.: Web site usability, design, and performance metrics. Inf. Syst. Res. **13**(2), 151–167 (2002)
16. Perkowitz, M., Etzioni, O.: Towards adaptive web sites: conceptual framework and case study. Artif. Intell. **118**(1), 245–275 (2000)
17. Pitkow, J.E., Kehoe, C.M.: Emerging trends in the WWW user population. Commun. ACM **39** (6), 106–108 (1996)
18. Saremi, Q.H., Abedin, B., Kermani, A.M.: Website structure improvement: quadratic assignment problem approach and ant colony meta-heuristic technique. Appl. Math. Comput. **195**(1), 285–298 (2008)
19. Schaffer, J.D.: Multiple objective optimization with vector evaluated genetic algorithms. In: Proceedings of the 1st International Conference on Genetic Algorithms, pp. 93–100, 1985
20. Shneiderman, B.: Designing the user interface: strategies for effective human-computer interaction, vol. 2. Addison-Wesley, Reading (1992)
21. Srinivas, N., Deb, K.: Multi-objective optimization using non-dominated sorting in genetic algorithms. Evol. Comput. **2**(3), 221–248 (1994)
22. Wang, M., Yen, B.: Web structure reorganization to improve web navigation efficiency. In: Proceedings of 11th Pacific Asia Conference on Information Systems. pp. 411–422, 2007
23. Wang, Y., Wang, Q., Ip, W.H.: Optimal design of link structure for e-supermarket website. In: IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 36, No. 2, pp. 338–355, 2006
24. Yen, B., Hu, P.J.H., Wang, M.: Toward an analytical approach for effective Web site design: a framework for modeling, evaluation and enhancement. Electron. Commer. Res. Appl. **6**(2), 159–170 (2007)
25. Yen, B., Hu, P., Wang, M.: Towards effective web site designs: a framework for modeling, design evaluation and enhancement. In: Proceedings of IEEE International Conference on e-Technology, e-Commerce, and e-Service, pp. 1–6, 2005
26. Yin, P.Y., Guo, Y.M.: Optimization of multi-criteria website structure based on enhanced tabu search and web usage mining. Appl. Math. Comput. **219**(24), 11082–11095 (2013)
27. Zhang, P., Von Dran, G.M., Blake, P., Pipithsuksunt, V.: Important design features in different web site domains: An empirical study of user perceptions. Am. Conf. Inf. Syst. (AMCIS) **1**(1), 77–91 (2001)
28. Zhou, B., Chen, J., Shi, J., Zhang, H., Wu, Q.: Website link structure evaluation and improvement based on user visiting patterns. In: The 12th ACM Conference on Hypertext and Hypermedia, pp. 241–244

Prevention Guidelines of SQL Injection Database Attacks: An Experimental Analysis

Vijaylaxmi Bittal and Soumi Banerjee

Abstract In today's global world every organization or enterprise, storing their essential data in terms of Databases and shares their data to authenticated users on the web through some security mechanisms. But now-a-days security is a big issue because of various types of database attacks. SQL injection is one type of such an attacks that inoculate a destructive SQL code to misuse database of an application. In this paper, we did experimental analysis on web-database applications as well as purely database applications and presented prevention guidelines of SQL injection attacks.

Keywords Database · Attack · SQL injection · Queries

1 Introduction

Today's databases are suffering from various database attacks. SQL injection is very dangerous attack among all data base attacks. The acronyms are SQLIA or SQLI [1]. The main reason for this kind of attack is improper design and development of database module, lack of validation controls on the user provided data during on the fly. Attackers easily run dangerous SQL statements on the server and change the database. So we considered SQLIA is a major attack and presented guidelines for preventing the same.

2 SQLI Types

SQLI can be classified into following major types [2].

V. Bittal (✉) · S. Banerjee
Department of IT, Ramrao Adik Institute of Technology,
Navi Mumbai 400706, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_3

2.1 Bypass Authentication

In this attack the attacker logs on to an application without valid authentication. This happens when an application does not have a filter on their authentication inputs e.g. *Username and Password*. There are many ways for Bypass Authentication in SQL injection. Here we are jotting down some of the ways

Or 1 = 1
0r 1 = 1—
Or 1 = 1#
Or 1 = 1/*
Admin' –
Admin' #
Admin' /*
Admin' or '1 '='1
Admin' or '1 '='1'—
Admin' or '1'='1'#
Admin' or '1'='1'/*
Admin' or 1 = 1 or ''='
Admin' or 1 = 1
Admin' or 1 = 1—
Admin' or 1 = 1#
Admin' or 1 = 1/*
Admin') or ('1'='1
Admin') or ('1'='1'–
Admin') or ('1'='1'#

2.2 Data Base Finger Printing

2.2.1 Using DBMS Error Messages

In this attacker inject some invalid queries and analyze the error messages. Those error messages [3] contains database information like server name, table name etc. Using these information attacker can attack on Databases

For example without entering the following valid input http://www.abc.com/Deals1.aspx?Campid=39934, an attacker have entered http://www.abc.com/Deals1.aspx?Campid=', and if it generates an error mes sage from that attacker may know the Database name or Table name and may attack on it.

2.2.2 Using String Concatenation

Here different strings concatenations the attacker can attack on Databases. There are different types of string concatenation operators for different databases

- For MS SQL- 'string1'+'string2'
- For MY SQL- CONCAT('string1','string2')
- For Oracle- 'string1'||'string2' or CONCAT('string1','string2')

For example, an attacker have entered the following input http://www.abc.com/Deals1.aspx?Campid=39934 AND 'str1str2'='str1'+'str2' then he may access the other values in Databases.

2.2.3 Using SQL Dialects

In this attack using some abnormal codes, statements and operators attackers can attack the database

For example during SELECT 1/0 if it returns NULL then Database is MY SQL Or if it returns Error: Division by zero then Database is Postgres Or if it returns ORA-00923: FROM not found where expected then Database is Oracle Or if it returns Server: Msg8134, Level 16, State 1, Line 1, Divide by zero encountered then Database is SQL SERVER.

2.3 SQL Injection with Union Query

Here an attacker joins two SQL statements using union query so that he may can access more than one tables of a Databases. Using these techniques an attacker can avail all types of data in Databases. Some of the examples are:

- Union All—It combines more than one SQL queries to avail all data in more than one tables http://www.abc.com/Deals1.aspx?Campid=39934 union select username, password from user
- Order By—using this an attacker can determine number of columns in a query. select name, password, address, birth_date from 'user' where 'id' = $id order by 4

3 Experimental Analysis

3.1 Bypass Authentication

As explained above bypass authentication is a well known database attack and with the above mentioned inputs we can bypass the authentication if the inputs are not validated.

In this analysis we created one test database with some relations using mysql and tested our inputs then observed vulnerability.

We start experimenting these inputs directly on test database, in that user table contains two columns username and password. Usually while writing SQL queries in where clause is for condition means left hand side equal to right hand side if it is not true the query will not return required fields.

Consider the above snapshot (Fig. 1) in that where sentence is having 1 = 1 condition is producing the sql query output. So wherever username and password inputs are there if the fields taking these inputs are not validated if you try with this option easily you are authenticated.

Similarly Consider next snapshot (Fig. 2)

Similarly next snapshot with double quotes in (Fig. 3)

With single quotes (Fig. 4)

Try some other options like as in (Fig. 5)

Suppose you are one of the user, you can access the other user information i.e. shown in below snapshot (Fig. 6).

Fig. 1

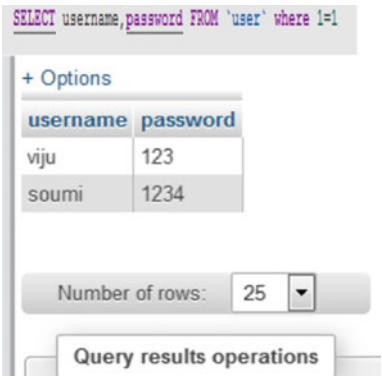


Fig. 2

Fig. 3

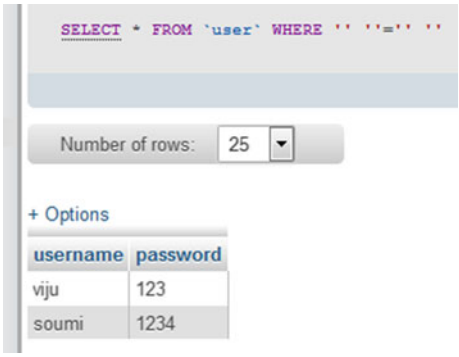


Fig. 4

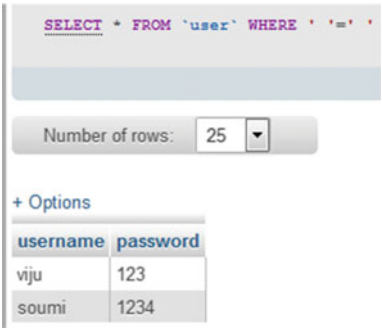
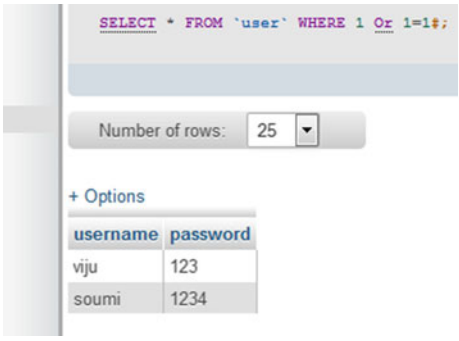
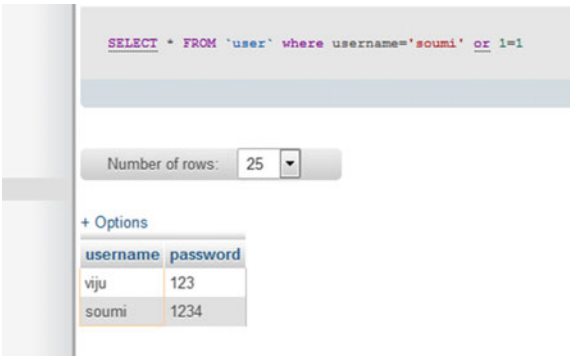


Fig. 5



Like this you can try the other options mentioned in (Sect. 2.1). From these snapshots we have analyzed that what is the effect of bypass authentication on the database and need to prevent this attack.

Fig. 6



3.2 Data Base Finger Printing

3.2.1 Using DBMS Error Messages

In this attacker purposely will try some invalid inputs and analyzes error messages.

See the above snapshots (Figs. 7 and 8), from these error messages attacker come to know that the application is using MySQL server, using these information an attacker can easily attack the database.

3.2.2 Using String Concatenation

In both the snapshot (Figs. 9 and 10) we got other rows just by adding concat function with already running query. Suppose you are the authenticated user so using concat you can get other users information that is not yours.

3.2.3 Using SQL Dialects

In this using some abnormal codes and some values attacker can attack database as shown in (Fig. 11).

Fig. 7

Fig. 8

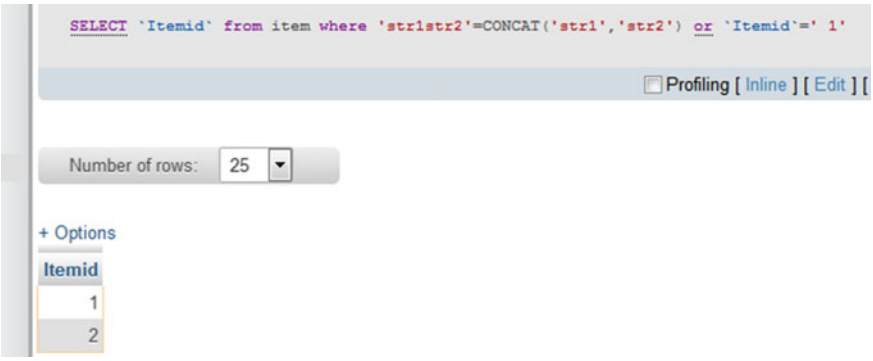


Fig. 9

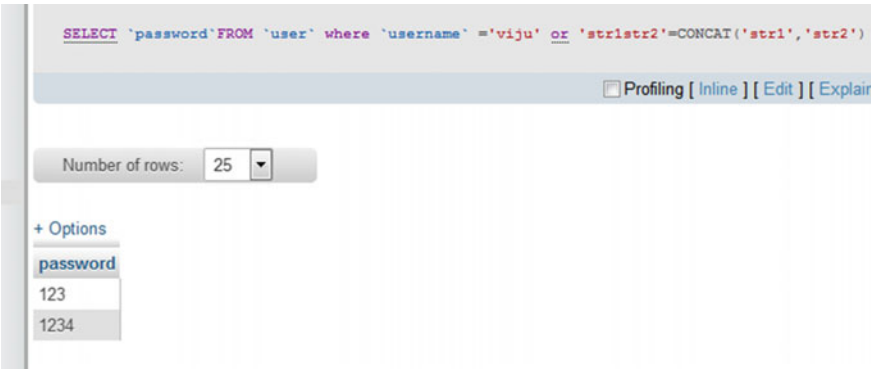
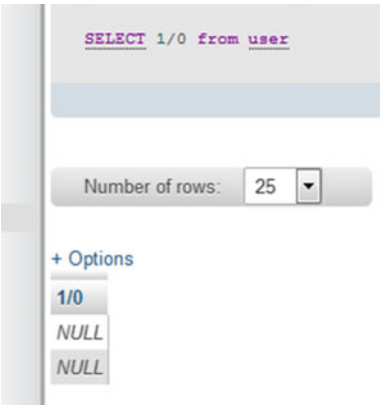


Fig. 10

Fig. 11



3.3 *Using Union Query*

- Union All

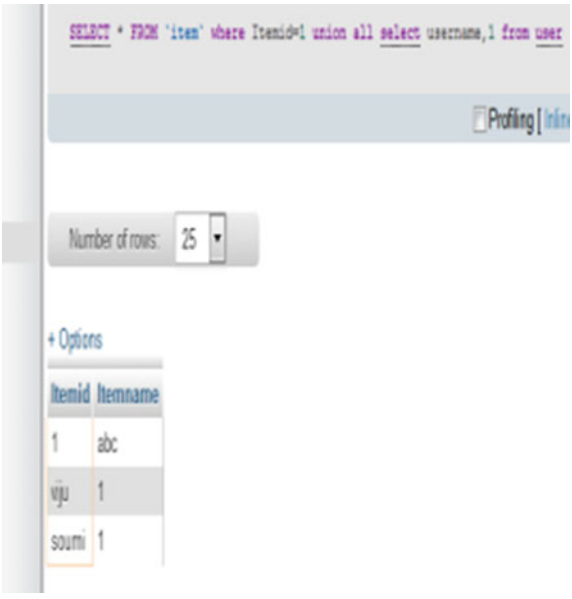
Using this we can join two tables.

In the above snapshot (Fig. 12) an additional query is combined along with running query. So result contains combination of two tables.

Similarly in this snapshot (Fig. 13) two tables combined using union all.

Fig. 12

Fig. 13



- Order By

The below snapshot (Fig. 14) showing the number of columns in table.

The Table 1 shows mapping between database attacks and their recommended and strongly recommended prevention guidelines.

Fig. 14

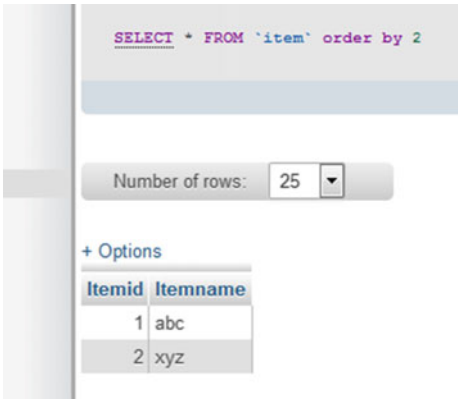


Table 1 Database attacks versus prevention guidelines mapping table

Prevention guidelines / Database attacks		A	B	C	D	E	F	G	H	I	J	K	L
Bypass authentication		☑	☑	☑	✓	✓	✓				✓	✓	
Data base finger printing	Using DBMS error messages	✓	✓	✓	☑				☑			☑	✓
	Using string concatenation		✓					☑	✓				
	Using SQL dialects		✓										☑
SQL injection with union query	Union all									☑		✓	
	Order by	☑	✓			✓							

☑ Signifies strongly recommended guideline
✓ Signifies recommended guideline
A—Validate the all possible inputs. Try almost all the alternative samples which we listed
B—Check all the web pages those who have link with the databases are vulnerable or not by putting single quote ‘at the end of url if it retuns ‘/’ server error means page is vulnerable
C—Inject some invalid inputs and check the effect
D—Error messages should not be generated from servers, it should be from application defined error messages
E—Create the cookies for each session
F—Carefully design ‘Back Buttons’
G—Check that any page is giving information about other users by using CONCAT
H—Inject some additional SQL code with running code and see the effect
I—Check UNION ALL with already running query
J—Don’t design the web pages with name of fields which are present in database
K—Try to avoid embedding SQL code with programming languages
L—Check some default values and peculiarities

If we follow these guidelines we can prevent SQL Injection attack to the some extent and can secure database [4].

4 Conclusion

From this experimental analysis we concluded that SQL injection is a very dangerous attack and has to be prevented. Following the above listed guidelines we can prevent the attack to the better extent and can secure your database.

References

1. Sharma, C., Jain, S.C.: SQL injection attacks on web applications. IJARCSSE (2014)
2. Sajjadi, S.M.S., Pour, B.T.: Study of SQL injection attacks and counter measures. IJCCE (2013)
3. Kulkarni, S., Uralogin, S.: Review of attacks on database s and database security techniques. IJETAE (2012)
4. Shar, L.K., Tan, H.B.K.: Defeating SQL injection. IEEE (2013)

Automated Data Analytics in Measuring Brand Sentiment Using Ml Technique from Online Conversations—Twitter

Karuna C. Gull and Akshata B. Angadi

Abstract With the ever growing amounts of data there is good motive to believe that smart data analysis will become more pervasive and an obligatory component in the progress of industry. Social Media act as a source between the company and the customers. People being at their work place can get to know the current status, discussions done by people on any trends and even he/she can express their opinions and views as per their interests through blogs, forums, emails etc. Twitter being a popular micro blog allows people to chat about their words with public in form of short texts (140 characters). Researchers of NLP and DM are attracted towards the sentiment analysis from past few years due to its many tricky research problems and purpose. In this paper, a novel Machine Learning Approach is used to classify the twitter dataset. The working of the Algorithm is explained with Sample set taken from twitter. If we have very little information i.e. the training sample, Algorithm defined may not give accurate estimation/probability of any object belonging to particular class as we have no other information to obtain a better estimation. This estimate will be reasonable if a training sample is very large and properly chosen. Thus in this paper Classification helps us to find the belongingness of an instance to a class to find the collective opinions of the users of twitter.

Keywords Classification · Data mining (DM) · Machine learning (ML) · Naïve bayes (NB) · Natural language processing (NLP) · Feature selection

K.C. Gull (✉) · A.B. Angadi
Department of CSE, K.L.E. Institute of Technology, Hubballi 580030, India
e-mail: karuna7674@gmail.com

A.B. Angadi
e-mail: akshata_angadi@yahoo.co.in

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_4

1 Introduction

Now a days "WORLD is GETTING MORE AND MORE SOCIAL ONLINE". Social Media provides way to the brand owners to connect to the targeted users increasing trust and loyalty. Reviews given by customer help the marketers to know how exactly their product is. Opinions like whether the product deployed attracting customers or lacking something!! This in turn helps marketers in raising their standards. Social helps to keep the customer relation strong with the brand. Social Networking sites being part of Media is attracting all youths, adults (of all age) by giving more facilities to interact with the people in form of short texts or posts. Examples of Social Media includes: Facebook, Twitter, Snapchat, Instagram. A study says 4.2 billion people uses mobile devices to access Social Media.

A study reveals that the use of Facebook, Twitter, YouTube and Instagram has become almost mandatory for brands, with more than nine in ten companies around the world using these platforms. Twitter being a popular micro blogging site acts as wonderful platform for research. As per the survey Twitter has 255 million active users who collectively send 500 million tweets each and every day. 50 % of technology companies have acquired a customer through twitter. Users in Mobile are more for twitter compared to using desktop. A statistics says there are 1 Billion+ users of Twitter. Figure 1 shows social media statistics for Twitter (18+ ages) [Source: comScore].

Sentiment Analysis refers to "[1] study of opinions, sentiments and emotions expressed in message. It can also be defined as an application of natural language processing, computational linguistics, and text analytics to identify and extract subjective information in source materials". Analysis helps to track the feelings, views, people ideas on web. This helps in knowing the products, brands and people.

[1] An opinion is a "QUINTUPLE", an object made up of 5 different things:

$$\left(O_j, f_{jk}, SO_{ijkl}, h_i, t_l\right)$$

where, O_j—The thing in question (i.e. product), f_{jk}—a feature of O_j, SO_{ijkl}—sentiment value of opinion of the opinion holder h_i in the feature f_{jk} of object O_j at time t_l.

Sentiment Analysis is not restricted to limited Application. Areas like automobiles, Books, Finance, Food and Health, Politics and Malls can use this analytics to

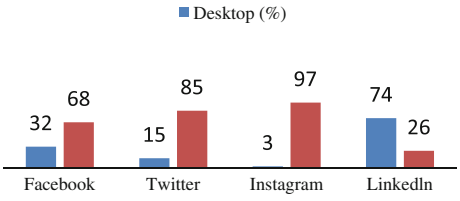


Fig. 1

find the sentiment value of customers. Sentiment Analysis requires to check many factors like: Statement of Negation, Conversion of words into Lower case Letters, Short forms, Acronyms, Normalization of Repeated Letters, Slang Letters to normal conversion, Spelling Corrections and many more.

If using a Machine Learning technique, system needs to be trained. It's a system learning technique. Human should train the system of how to act before testing the systems performance. Training should be done with more effort to reach accuracy level.

2 Discussions

Large volume of data is processed to construct a simple model with valuable use. If the system can learn and adapt to such changes, the system designer need not foresee and provide solutions for all possible situations. Thus building an automated system that makes the work simpler for the customer is a main motto. Two Approaches can be used as per the requirement. Dictionary based or Machine Learning Approach. Here we discuss about ML Techniques. ML help us to develop an automated system that analyze the dataset and help in predicting the data using supervised or unsupervised machine learning techniques.

Supervised Technique (Discriminant Analysis):

- Discover patterns in the data that relate data attributes with a target (class) attribute.
  - These patterns are then utilized to predict the values of the target attribute in future data instances.
- Highly dependent on the information given by the pre-determined classifications i.e. it Uses training data to infer model
- Model can be tested

Unsupervised Technique (Cluster Analysis):

- The data have no target attribute.
  - Explore the data to find some intrinsic structures in them.
- There is no training data. Thus the Model cannot be tested

Semi Supervised Learning:

- can use readily available unlabeled data to improve supervised learning tasks when the labeled data are scarce or expensive [2]
- Consists of both labelled and unlabelled data

Clustering is often called an unsupervised learning task as no class values denoting an a priori grouping of the data instances are given, which is the case in supervised learning. Classification fall under Supervised Learning Technique where

training the data set plays an important role. Training is an iterative process. Observation by giving different test cases, at this level is important.

Both Supervised and Unsupervised techniques can be valuable [3]. Which one you choose depends on the circumstances-what kind of problem is being solved, how much time is allotted to solving it (supervised learning or clustering is often faster than reinforcement learning techniques), and whether learning is even possible or not is to be taken into account [4].

Machine Learning Techniques get more accurate results when compared to the assumptions framed by human.

Machine Learning Model

$$B = M(A/P) \tag{1}$$

where M is a model developed using ML Technique, A is an input (Dataset taken), P is the parameters of M and B is the particular class. Equation 1 is a discriminator function that determines the belongingness of a tweet to a particular class. With machine learning algorithms, a common problem is over-fitting the data and learning the training set rather than a general classification technique.

For Machine Learning Techniques, applying first a method of automatic feature construction yields improved performance and a minimum set of features. Feature Selection helps to solve this problem by selecting the features. Feature Selection Technique needs to be used to evoke relevant features in learning process. If Feature selection is not considered before classification then two major problems are dealt (a) Greater computational cost (b) Over fitting. To add on, problems like lesser performance, more time and high storage requirements are observed.

Feature selection helps to ignore those input features that have little effect on the output, so as to keep the size of the Framework small. A set of features F and a target variable T is given as input, Finding minimum set F that achieves maximum classification performance of T for given classifier is an output. Thus it helps to improve the performance of classification algorithm and also reduce the cost. Numerous feature-selection algorithms have been proposed during the last few decades. Feature selection has been an active and fruitful field of research area in pattern recognition, machine learning, statistics and data mining communities [5, 6].

Minimizing the features of the original set S (x_1, x_2, x_3, x_4, x_5 …..x_m) of classifiers results in new subset S` (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5} …..x_{in}) with limited number of features say n where n < m (m is the initial features). Feature constraint reduces the feature set.

$$S \subset S' = \{x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5} \ldots x_{in}\} \subset \{x_1, x_2, x_3, x_4, x_5 \ldots x_m\}.$$

With feature constraint the final classifier must operate on n-dimensional subspace of the original m-dimensional space. This size reduction of the function space from which classifier is chosen reduces the likelihood of over fitting.

In this paper Classification Algorithm and the feature selection algorithm is used to classify the data input. The belongingness of a tweet is achieved better by iteratively performing the training using ML Techniques.

Given a set of tweets, each tweet is taken for analysis. If a particular tweet includes the attributes/features specified/predefined in trained set then the tweet get placed into any of the three classes as per the rules, '+' (positive) class,'−' (negative) class or neutral class.

3 Methodology

Designing of system begins with authentication process and extraction of tweets.

In authentication process, application need to be registered with Twitter, through which it issues a consumer-key and secret-key [7]. It uses these keys to generate a unique Twitter link for user authentication. The user needs to authenticate to authorize the application. Now Twitter verifies the user's identity and gets a PIN. User provides that PIN to application, so that it uses that PIN to get the session keys like Access-Token and Access-Secret. Using Access-Token and Access-Secret, the application authenticates the user to issue different API calls to extract the required data from Twitter.

In Extraction process, Once authentication process is completed, we can extract the data depending on our requirements of an application. Twitter is a structured model. Based on the kind of data present in social media, extraction method is applied. The data present in the webpage of Twitter come under structured whereas related data of the same, given by user come under unstructured one. We write the query to get the contents of web page of Twitter which is in the JSON format. We need to transform data from JSON to XML and there are a rich set of APIs and Tools available to do these transformations. Write the query with field name to extract required data from XML and dump them into either database or file for further processing.

In Preprocessing of training data or test data, we remove all URLs (e.g. www.example.com), hash tags (e.g. #topic), targets (@username), special Twitter words ("e.g. RT"), punctuations and exclamation marks. Replace all the words in the CAPs by lowercase. Do the spell corrections for slang words. Replace all the emoticons with their sentiment words. Remove all stop words. Now do the tokenization of the tweets collected. In sentiment analysis use of n-grams framework is preferable; if the number of n selected is 2 or 3 g will yield good results. But in our case we have taken 1-g framework. In this analysis the number of occurrences of the words/token in the tweet will not make a difference, as Binarized version of the algorithms perform better on single occurrences then multiple.

The main part of system is designed in two phases as showed in Fig. 2.

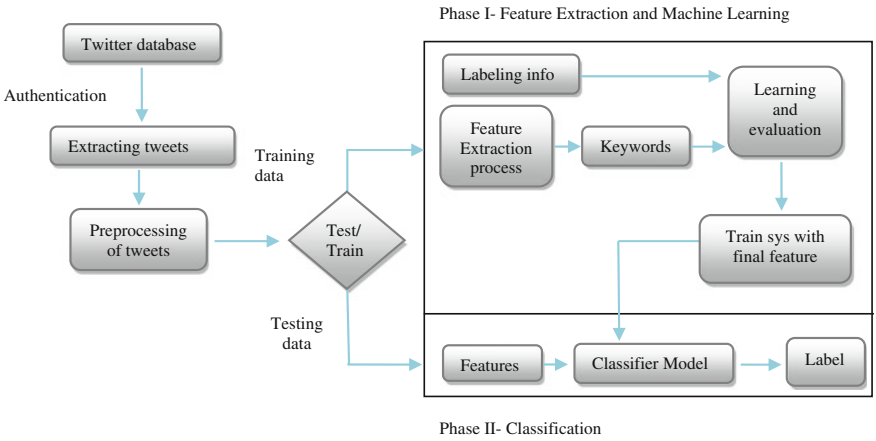


Fig. 2

Phase I: Feature Extraction—The steps involved in this phase are

a. Labelling information
b. Feature extraction to create feature subset
c. The learning algorithm

Phase II: Classification

a. Feature extraction
b. Classification model
c. Labeling

3.1 Feature Extraction to Create Feature Subset

In learning based techniques, before training or given as input to the classifier, we must select the better features. To do that task one can use different feature selection methods or combination of the same. As we can't use all the words that preprocessing or the tokenization has given us because there are several impertinent words within them. Depending on the domain on which we are working and text classifiers we are using the feature selection or combination of feature selection methods are going to change. For example, to select the feature set for restaurant reviews one may use combination Max Entropy and Chi-square feature selection methods, and for twitter Mutual Information feature selection is better [8]. There are two ideas to do the feature selection. One, the Hypothesis testing statistics: in which if there exists an association between a value of variable belonging to one category with the value of another, then use Chi-square method [9]. Another, the Information theory: in which if there exists any dependencies or indecencies of a value of variable belonging to

one category with the value of another then prefer Mutual information (MI). Thus one can commonly use these two feature selection algorithms in Text Classification [10]. Each algorithm evaluates the keywords in a different way and thus leads to different feature set. In our case we have used Mutual information to select feature set, followed by trial and error method to configure the system to work better. Thus in training process every token of each tweet with labelling information is undergone with mutual information feature selection method, which gives us the probabilities of each token with respect to labels. The tokens with highest probabilities and their labels are added in final token/training set [11]. Hence the fundamental idea of feature selection algorithm is that, it searches through all possible combinations of attributes/keywords in the tweets to find the subset of features works best for prophecy. So, the attribute vectors can be reduced in number where the meaningful ones are kept and the unrelated ones are removed. Here all the tweets in the training data are classified into three different classes such as positive, negative and neutral.

3.1.1 Labeling [12]

In training phase, once features are extracted from training data/tweets, the learning algorithm will use the label information and data or the feature itself to learn a map function. By this process features are assigned with proper label. In the classification phase, the features extracted from tweet under test are compared and analyzed with the feature set extracted along with their labels learned from mapping function in the training process, by the classifier. As we are using binarized Naïve Bayes algorithm, in which the multiple occurrence of a token in a train tweet is considered as single occurrence, which can be used in finding the conditional probabilities of each token of tweet under test. The output of this process will predict the label of the tweet under test.

3.2 Classification

Classification is to examine input data and build up the accurate portrayal model for each class using features present in data. Further that model is used to classify the test data for which class depict are not known.

In a text classification [13] problem, we will use the words/tokens of the tweets to find the tweets' appropriate class. Instead of assigning weights to the tweets, if we know some pertinent and some impertinent tweets, then it becomes easy for us to build a classifier that depends on probabilistic. One such model is the Naive Bayes model i.e.

$$\mathrm{P}(t_k|\mathrm{R}) = |\mathbf{D}_{rk}|/|\mathbf{D}_r| \text{and } \mathrm{P}(t_k|\mathrm{NR}) = |\mathbf{D}_{nrk}|/|\mathbf{D}_{nr}|$$

where t_k is a term; $\mathbf{D}_r$ is pertinent tweets set; $\mathbf{D}_{rk}$ is the subset which contains term t_k; $\mathbf{D}_{nr}$ is impertinent tweets set; $\mathbf{D}_{nrk}$ is the subset which contains term t_k.

We begin with Probability basics [Source: Wiki]: Study of Probability theory is one best approach for machine learning and classification. Bayes theorem is one, which works suitably on probability theory. Thus Bayes' theorem becomes building block of naive Bayes classification.

$$Posterior\ probability = \frac{conditional\ probabilty * prior\ probabilty}{evidence}$$

where, the posterior probability (w.r.t. classification) is a probability of belongingness of a particular object to particular class, given its trained feature values" [14].

3.2.1 Naïve Bayes Algorithm

Given: A description of words/tokens/features, $x \in X$, where X is the T*weet* and A fixed set of classes: $C = \{c_1, c_2,\ldots, c_J\}$.

Determine: The Class of a new instance/tweet X based on the feature values $X = \{x_1, x_2,\ldots, x_n\}$ into one of the classes $c_j \in C$ [15]

Table 1 shows the Algorithm of Naive Bayes.

Table 1

Naïve bayes algorithm:

Learning Process

- From training corpus, extract *Vocabulary*
- Calculate required $P(C_j)$ (Prior Probabilities of each class) and $P(x_k \mid C_j)$ (Conditional Probabilities of each (keyword/class))
- For each C_j in C do
  - Begin
    - $N(C_j) \leftarrow$ Number of tweets belonging to class C_j in corpus
    - $N \leftarrow$ Total number of tweets in corpus
    - *To find Prior Probability*

$$P(C_j) \leftarrow \frac{N(C_j)}{N}$$

    - *To find Conditional probabilities*
    - *Testdata* $\leftarrow$ Single tweet under test
    - for each word/token x_k in *Testdata*
      - *Begin*
        - $count(x_k, C_j) \leftarrow$ Total number of occurrences of keyword x_k picked from *Testdata* in the class C_j
        - $count(C_j) \leftarrow$ Total number of words/keywords in class C_j *(Neglect the multiple occurrence of the same word in a tweet)*
        - *Vocabulary* $\leftarrow$ Total number of terms in all classes (without duplications)

$$P(x_k \mid C_j) \leftarrow \frac{count\ (x_k, C_j) + 1}{count\ (C_j) + |Vocabulary|}$$

        No training cases for a particular attribute of test data, results to zero probability. Addition of 1 is necessary to avoid over-fitting .
      - End
  - End

Classifying Process:

- positions $\leftarrow$ All word positions of current *Testdata*
- Calculate C_{NB}, for all the classes using

$$C_{NBj} = \arg\max_{C_j \in C} P(C_j) * \prod_{k \in positions} P(x_k \mid C_j)$$

- Multiplication of probabilities (0 to 1) results in floating point underflow, so use summing log for better performance

$$C_{NBj} = \arg\max_{C_j \in C} \log P(C_j) + \sum_{k \in positions} \log P(x_k \mid C_j)$$

- To find the class of Test data using Max a Posterior(MAP) of C_{NBj}

$$C_{NB} = \text{Max}(C_{NBj})$$

4 Analysis of Tweets Using Naïve Bayes Algorithm

Table 2 are the sample tweets collected from Twitter account.

In the preprocessing we remove all stop-words, replace jargons/short forms of words and replace words in the CAPs by lowercase and emoticons with actual words etc. Then tokenize them into tokens which are acting as input to feature extraction process. This process uses mutual information technique to select the features. Learning process uses both features and labeling information given by user to evaluate the feature quality so that they can be placed into training set along with their labels. Thus the final sample training set is as follows:

Table 3 shows final sample training set collected after learning and evaluation process [16]:

Table 2

Sample tweets collected from Twitter account for Sony brand	Labeling information
1. Sony Xperia is beautiful 5-in. full HD display ☺	C_1
2. Sony mobile sleek, strong, waterproof design	C_1
3. Sony Xperia has pleasing OS and UI combination	C_1
4. Sony Xperia Z3 water proof sensational smart phone	C_1
5. Sony Xperia Z3 yes it is stunning when viewed straight	C_1
6. It's basically faster than you'll ever need	C_1
7. Sony Xpeia the small apps feature is a wonderful addition	C_1
8. Sony latest model fantastic Phone!	C_1
9. Sony simply wow!	C_1
10. Sony the best phone I've had with an exemplary and stylish model	C_1
11. It has Powerful quad core processor ☺	C_1
12. Superb camera quality	C_1
13. Sony mobile Becomes excessive hot during use	C_2
14. Sony handset quickly becomes smeared mess	C_2
15. Sony mobile disappointing battery life	C_2
16. Sony Xperia Z3 stunning but with minor flaws	C_3
17. Sony's glass panel at the back scratches easily	C_2
18. Sony's flaps covering headphone and usb ports are annoying	C_2
19. The battery of Sony got so hot whilst watching a film, that the screen cracked	C_2
Test Case 1—build quality of Sony Xperia Z3 is stunning, its waterproof and the attention to detail is exemplary	?
Test Case 2—Sony Xperia is average in lookwise, battery becomes hot soon and found flaws when I used it	?

Table 3

Tweet No.		Class
Training data		
1	Beautiful 5-in. smile	C_1
2	Sleek strong waterproof	C_1
3	Pleasing	C_1
4	Waterproof sensational	C_1
5	Stunning	C_1
6	Faster	C_1
7	Wonderful	C_1
8	Fantastic	C_1
9	Wow	C_1
10	Best exemplary stylish model	C_1
11	Powerful core smile	C_1
12	Superb camera quality	C_1
13	Problem excessive hot	C_2
14	Smeared mess	C_2
15	Disappointing	C_2
16	Stunning flaws	C_2
17	Scratches problem	C_2
18	Annoying	C_2
19	Hot whilst cracked	C_2
Test data		
TC-1	Build quality stunning waterproof exemplary best	?
TC-2	Average hot flaws	?

Now the steps followed by us using Naïve Bayes text classification to find the class or label of the tweet under test:

Step 1: Find Prior Probability of all different classes

$P(c_j) \leftarrow \frac{\lvert N_{Cj}\rvert}{\lvert N\rvert}$	where N_{Cj} = Number of tweets in class c_j	$P(C_1)$ = 12/19 = 0.6316
		$P(C_2)$ = 6/19 = 0.3158
	N = Total number of tweets	$P(C_3)$ = 1/19 = 0.0526

Step 2: Find Conditional probabilities of all the tokens, by picking each from Test data

$P(x_k\|c_j) \leftarrow \frac{count(x_k,c_j)+1}{count(c_j)+\|Vocabulary\|}$	*count*(x_k,c_j) = Total number of occurrence of the token picked in the class c_j	*count*(c_j) = Total number of keywords in the class c_j (instead of considering multiple occurrence of a token in a tweet, consider it once for a tweet (duplicates allowed)
	Vocabulary = Total number of keywords in all the classes (without duplications)	
For C_1 class:	For C_2 class:	For C_3 class:
P(build|C_1) = (0 + 1)/(24 + 35) = 0.016949	P(build|C_2) = (0 + 1)/(12 + 35) = 0.021277	P(build|C_3) = (0 + 1)/(2 + 35) = 0.027027
P(quality|C_1) = (1 + 1)/(24 + 35) = 0.033898	P(quality|C_2) = (0 + 1)/(12 + 35) = 0.021277	P(quality|C_3) = (0 + 1)/(2 + 35) = 0.027027
P(stunning|C_1) = (1 + 1)/(24 + 35) = 0.033898	P(stunning|C_2) = (0 + 1)/(12 + 35) = 0.021277	P(stunning|C_3) = (1 + 1)/(2 + 35) = 0.054054
P(waterproof|C_1) = (2 + 1)/(24 + 35) = 0.050847	P(waterproof|C_2) = (0 + 1)/(12 + 35) = 0.021277	P(waterproof|C_3) = (0 + 1)/(2 + 35) = 0.027027
P(exemplary|C_1) = (1 + 1)/(24 + 35) = 0.033898	P(exemplary|C_2) = 0 + 1)/(12 + 35) = 0.021277	P(exemplary|C_3) = (0 + 1)/(2 + 35) = 0.033333
P(best|C_1) = (1 + 1)/(24 + 35) = 0.033898	P(best|C_2) = (0 + 1)/(12 + 35) = 0.021277	P(best|C_3) = (0 + 1)/(2 + 35) = 0.027027

The resultant matrix is

Test Case 1:						
Data 1	Build	Quality	Stunning	waterproof	Exemplary	Best
$P(x_k\|c1)$	0.016949	0.033898	0.033898	0.050847	0.033898	0.033898
$P(x_k\|c2)$	0.021277	0.021277	0.021277	0.021277	0.021277	0.021277
$P(x_k\|c3)$	0.027027	0.027027	0.054054	0.027027	0.027027	0.027027

Test Case 2:			
Data 2	Average	Hot	Flaws
$P(x_k\|c1)$	0.016667	0.016667	0.016667
$P(x_k\|c2)$	0.021277	0.063830	0.042553
$P(x_k\|c3)$	0.027027	0.027027	0.054054

Step 3: Maximum a posterior-MAP class (C_{NB}) is used to find the best class of test data in NB classification. If the C_{NB} is underflow then Log is applied and instead of multiplying, adding Log of Profanities yields to probable result.

Thus Final Results that reveals the belongingness for two test cases as shown in Table 4.

Table 4

	Class	$c_{NB} = Max(\arg\limits_{c_j \in C} P(c_j) * \prod\limits_{k \in positions} P(x_k \mid c_j))$	$C_{NBj} = Max(\arg\limits_{C_j \in C} \log P(C_j) + \sum\limits_{k \in positions} \log P(x_k \mid C_j))$
Test case 1	C_1 (Positive)	7.18700E-10	−9.14345
	C_2 (Negative)	5.85923E-11	−10.5332
	C_3 (Neutral)	4.10266E-11	−10.3869
Test case 2	C_1 (Positive)	2.92398E-06	−5.5340
	C_2 (Negative)	1.82497E-05	−4.7387
	C_3 (Neutral)	2.07812E-06	−5.6823

Final Result shows the belongingness of sample tweets. First Test case belongs to the Positive class where as second Test case belongs to the Negative class. The result is completely dependent on the training set. If for instance we consider the Neutral Test case we will not get the correct classified result as in trained set we did not consider neither predefined neutral keywords nor tweets of that category. Note that the Training set should be huge. Here we have taken few dataset. Fact is Data Mining Approaches work better if the data is huge. Hence we conclude saying "Training and the Classifier used should be best; if not, better for proper classification".

5 Conclusion

On Web it is difficult for a person to find pertinent basis and get related opinions. Even though an individual finds data, organizing all the data into useful form is difficult task. Research on finding opinion sources and monitoring them is increasing due to the remarkable growth of huge data on diverse sources. Thus, innovations in field of automated opinion discovery and summarization systems are essential. Main tasks included are Mining objects and their features that have been commented or expressed by people through tweets and determining whether the comment/opinion on each entity feature is positive, negative or neutral and lastly summarizing the results.

In this paper Twitter is taken as case study. Opinions in Twitter are in form of short texts. The Experimental Results shows the mined tweets with their sentiment value done using NB classifier. As NB is an ML Technique, accuracy can be reached if the trained set is at the high level. The Analysis is done using two test cases, Final result shows in which class instance, the tweets (test cases) belong to. Thus the probability of correctly classifying the randomly selected data is dependent on the Classification Method and randomly selected training set. Hence we conclude that estimating the accuracy of a supervised (ML Technique) classification method is, its ability to determine the class of unseen data instance.

References

1. Liu, B.: Sentiment Analysis and Subjectivity. In: Indurkhya, N., Damerau, F.J. (eds.) Handbook of Natural Language Processing, 2nd edn (2010)
2. Zhu, X., Goldberg, A.B.: Introduction to Semi-Supervised Learning, p. 130 (doi:10.2200/S00196ED1V01Y200906AIM006) (2009)
3. Padhye, A.: Comparing Supervised and Unsupervised Classification of Mmessages in the Enron Email Corpus. University of Minnesota, Duluth, August 2006
4. Supervised and Unsupervised Learning AI Horizon
5. Mitra, P., Murthy, C.A., Pal, S.K.: Unsupervised feature selection using feature similarity. IEEE Trans. Pattern Anal. Mach. Intell. **24**(3), 301–312 (2002)

6. Miller, Subset Selection in Regression, Chapman & Hall/CRC, 2nd edn (2002)
7. http://twittercommunity.com/t/twitters-search-api-and-user-application-authentication-can-anyone-clarify-my-concerns/1188
8. Al-Ani, A., Deriche, M.: Feature Selection Using a Mutual Information Based Measure, pp. 1051–4651/02. IEEE (2002)
9. https://www.statpac.com/statistics-calculator/correlation-regression.htm
10. Tsimpiris, A., Vlachos, I., Kugiumtzis, D.: Nearest neighbor estimate of conditional mutual information in feature selection, Elsevier, doi:10.1016/j.eswa.2012.05.014 (2012)
11. Ting, S.L., Ip, W.H., Tsang, A.H.C.: Is naïve bayes a good classifier for document classification? Int. J. Softw. Eng. Appl. **5**(3), July 2011
12. K-nearest_neighbors_algorithm Wikipedia Source
13. Korde, V., Namrata Mahender, C.: Text classification and classifiers: a survey. Int. J. Artif. Intell. Appl. (IJAIA) **3**(2), March 2012 DOI : 10.5121/ijaia.2012.3208.pp.85–99
14. Raschka, S.: Naive Bayes and Text Classification I-Introduction and Theory, Article, 04 October 2014
15. http://didawiki.cli.di.unipi.it/lib/exe/fetch.php/magistraleinformatica/eln/4-textclassification.ppt
16. Gupta, G.K.: Introduction to Data Mining with Case Studies, 2nd edn. PHI Publication (2013)

Design of Five Port Structure and Evaluation of Its Performance at 5.8 Ghz for SDR Application

K.V. Karthikeyan

Abstract In this paper, the design of five port structures using power dividers, hybrid coupler and phase shifter was done using Advance Design Systems. The design termination considered as 50 ohm and the microstip based methodology is followed. The performance analysis of the design is carried out with s-parameter simulation and the corresponding return loss and mismatch loss are obtained. The results obtained are in greater coherence with the theoretical observations.

Keywords Hybrid coupler · Microstrip · S-parameter · Phase shifter and VSWR

1 Introduction

Software Defined Radio mainly consists of two sections: Software and Hardware. In the course of these developments, it has become predominant that although software section of these radios is flexible, the hardware section also needed to be upgraded or made flexible to incorporate the growing needs of the customer [1, 3]. In this work, the hardware aspects of the SDR are addressed, which is the recent area of interest for researchers. The software section consists of DSP Processors or FPGA. The hardware part consists of mainly BPF, LNA and Mixer. Researchers have been working on exploring alternative solutions for the hardware. Port structures are useful alternatives in the RF front end sub system of radio receivers. There has been considerable research in this area since last two decades and the exploration is wide open because of changing demands of communication technology [2, 3]. These port structures replace mixer stage in direct demodulation receivers and their applications lies in software defined radio. In this research, software defined radio application is considered and the design of port structures suitable to the changing needs is addressed. Port Structures prominently are of two types, Six Port and Five Port Structures. Port structures consist of hybrid coupler

K.V. Karthikeyan (✉)
Department of ECE, Sathyabama University, 600119 Chennai, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_5

and power dividers along with the phase shifters. These structures are not useful for homodyne receiver operation. This is because of the direct path existence between LO and RF ports of these structures. These port structures replace mixer stage of a radio receiver and low noise amplifier design [6], which facilitates direct demodulation of received signal. In this work an attempt is made successfully to design five port structures for a particular range of specifications, where a great degree of performance is obtained when compared with their counterparts. In this the methodology adopted contains micro strip lines which were designed and used to make these structures as competent enough with the changing market needs.

2 Port Structures in Software Defined Radio

Port structures find their applications in Direct Conversion Receivers. Basic difference in the receivers exists in the number of stages sufficient enough to convert RF signal to baseband signal. Direct conversion consists of down conversion; whereas super heterodyne receivers have two or more [10]. The complexity of the receiver increases along with the number of stages in down conversion.

The direct down conversion technique, as shown in Fig. 1 [11] does down converting the signal (i.e. from RF to base band) where there is no mixer thereby no IF frequency. The receiver unit consists of a down-conversion of the transmit RF signal received from the antenna to base band unit. The receiver unit performs first a filtering and amplification before the signal is fed to our main direct down converter circuit after a second, amplification [4, 7]. The signal is then down converted from RF to base band using the Five-Port direct down conversion techniques to base-band. The outputs obtained at the three output ports were amplified before A/D conversion. In the next stage IQ values are calculated. In this work, micro strip [6] based design method is used and these micro strips are very useful particularly at RF frequency. Micro strip structure consists of a metal, which is a shaped layer on one face of dielectric material, whereas it is covered with metal on the opposite side.

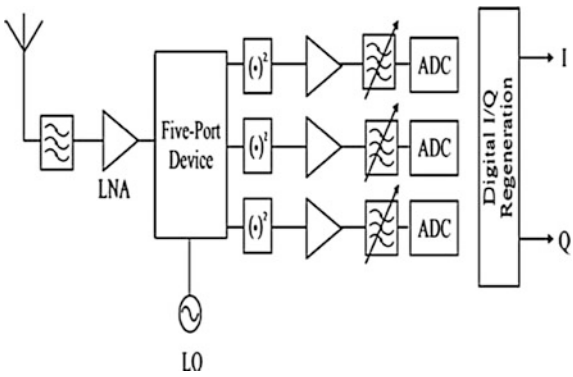


Fig. 1 Base band five port architecture (11)

Circuit shape is printed on one side of the layer i.e., front side and the other side which is grounded or consists of information pertained to the shape of circuit in accordance with the application. These two conducting layers are in general made of copper, which are very thin around micrometres and these are separated by a dielectric layer (substrate). Advantages of using a micro strip line are numerous. These structures are famous for their low loss and flexibility in implementation. These are high frequency components and therefore surface mount capacitors, transistors can be easily fabricated on to the printed circuit board micro strip metallization layer. The basic design philosophy of the five port structure is elaborated in the Sect. 3.

3 Design Philosophy of Five Port Structure

Port Structures are designed using power dividers, hybrid couplers and phase shifters

3.1 Power Divider

Power divider is used in power splitting, fault testing in transmission line and in other applications related to processing of signal. Power divider receives an input signal and gives multiple output signals. This is a passive device, which divides RF signal into different signals [5]. Power dividers properties include, equal amplitude, Zero degrees phase among the output ports, high isolation between output signals, impedance matching at ports. Typical power divider is shown in the Fig. 2 [9].

3.2 Wilkinson Power Divider

Wilkinson power dividers shown in Fig. 3 [8] are often used because of the fact that there is a necessity to replace the feed structures, which are lossy into the linear

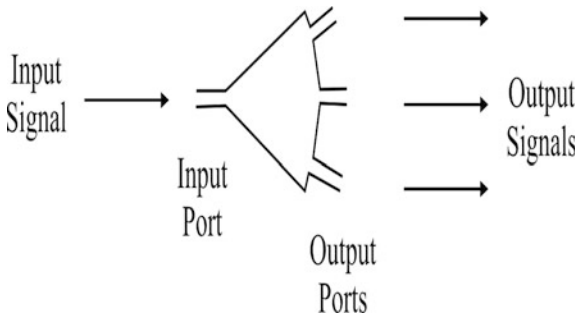


Fig. 2 Power divider (9)

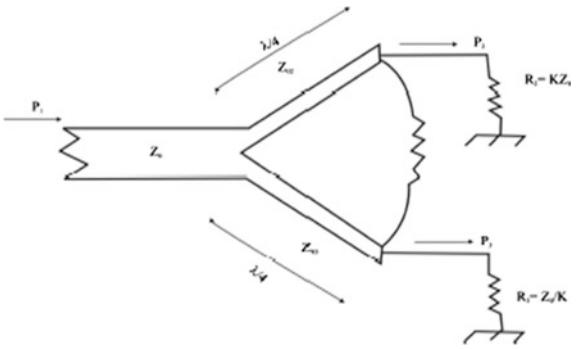


Fig. 3 Wilkinson power divider for unequal power splits (8)

array of antenna elements. Low to Medium power transmission is possible with these components and it requires quarter wavelength multiple lines, these structures have got topological advantage also because of near linear alignment of output ports. Voltage Standing Wave Ratio of Wilkinson's power divider is a low value at all the ports and there exists high isolation between the output ports. The i/p and o/p impedances at each port are designed to be equal to the characteristic impedance (Z_o) of microwave system [8]. The Wilkinson power divider has advantages as Lossless; Output ports are isolated and arbitrary power division.

3.3 Hybrid Coupler

In the field of radio technology these power dividers and directional couplers are used for power division or combining. A Hybrid coupler is a four port component, shown in Fig. 4 [7], which can be used to either split an input signal equally with 90° phase shift between output signals. Quadrature hybrid couplers play an important role in various RF systems like receivers, transmitters, signal analysis/processing circuits etc., for phase shifting, power combining, and power applications. These are often connected with the other components those are used in

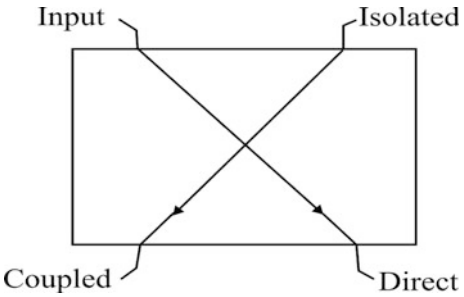


Fig. 4 Quadrature hybrids (7)

mixers, amplifiers and other RF circuits. The overall performance of the design also depends upon the effectiveness of the hybrid coupler design.

The extent of coupling is determined by the coupling coefficient and the coupling factor is calculated using $k = (\rho - 1)/(\rho + 1)$, where $\rho = (Z_{oe})/(Z_{oo})$. Z_{oe}, Z_o and Z_o are related by, $Z_{oe} * Z_{oo} = Z_{o2}$, where Z_o is the characteristic impedance. All 90° Power Dividers and Combiners are reciprocal four port networks. These are also known as also known as quadrature hybrids or Quad hybrids [5].

3.4 Phase Shifter

It is often a requirement to change the phase of signals in many electronic applications. Microwave Phase Shifters have numerous applications in various equipment's like phase discriminators, phase array antennas and linearization of power amplifiers. It can be used to shift a signal to required phase location. Circuit parameters selection is important to achieve maximum phase shift. Properties of Phase Shifters are Frequency Range, Bandwidth, Phase Variance, Insertion Loss, Power Handling, VSWR, Return Loss and Switching Speed. Analog Phase Shifters based on varactor diodes can achieve a good amount of phase shift, high speed and require lesser number of diodes than in digital phase shifters [12].

The design shown in the Fig. 5 is designed with the philosophy comprising of all the components in the Sect. 3. Advantages include area optimization, easy to fabricate and higher return loss, which is evident from the results. Hybrid couplers in this design contain Tees, which has a W of 3.009020 mm and this value is prone

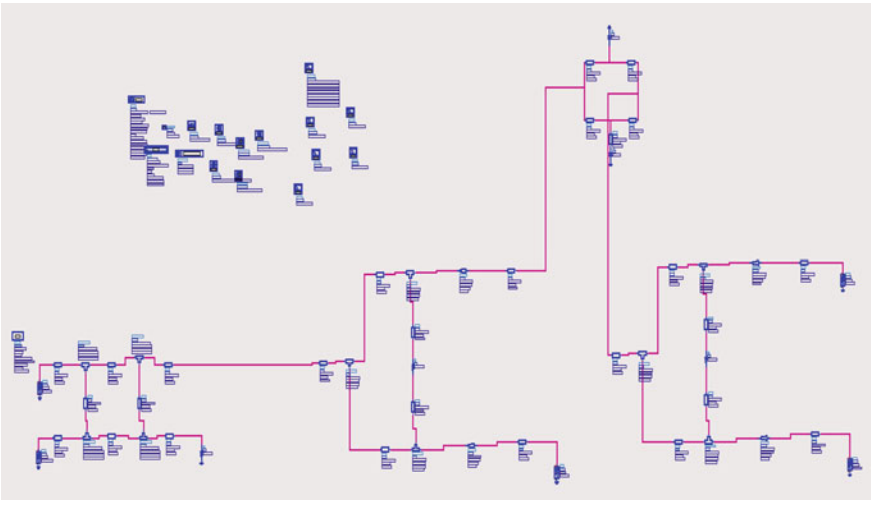


Fig. 5 Design of five port structure using micro strip design methodology at 5.8 GHz

to alteration in accordance with the centre frequency as 5.8 GHz. Scattering parameter simulations SP1 has been incorporated for efficient analysis of the proposed design. Dielectric constant of substrate here is 4.6 for all the lines that are used in this design. Msub specifications are in accordance with micro strip technology.

4 Results and Discussion

The simulated results obtained are as shown in the Figs. 6, 7, 8 and 10. Figure 11 represents the Layout extracted for the design shown in the Fig. 5. At 5.8 GHz the variation of VSWR is obtained at three output ports and is represented in the Figs. 6, 7 and 8.

The below set of conversion formulae are used in calculating the Mismatch loss and Return Loss of the design.

$$\Gamma = \frac{VSWR - 1}{VSWR + 1} \tag{1}$$

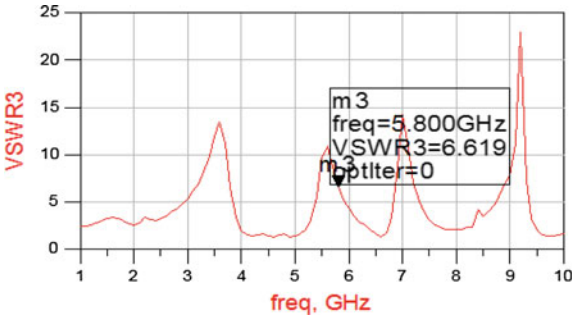


Fig. 6 VSWR at the output port 3

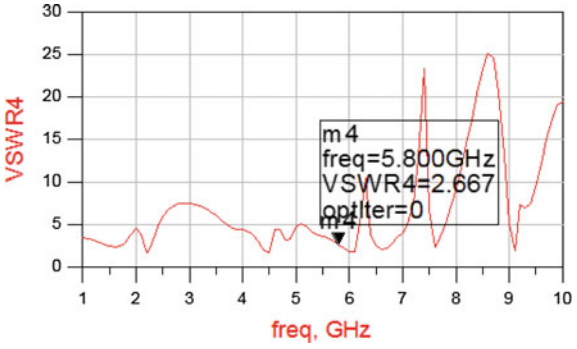


Fig. 7 VSWR at the output port 4

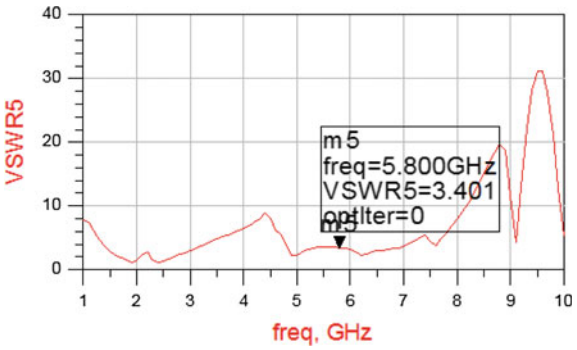


Fig. 8 VSWR at the output port 5

Table 1 Performance analysis of five port structure design

PORT number	VSWR	Reflection coefficient	RL (return loss in dB)	ML (mismatch loss in dB)
3	6.619	0.74	2.64	3.41
4	2.667	0.45	6.85	1.01
5	3.401	0.55	5.26	1.53

$$\mathrm{RL} = -20\log\left[\frac{VSWR-1}{VSWR+1}\right] \tag{2}$$

$$\mathrm{ML} = -10\log\left\{1\left[\frac{VSWR-1}{VSWR+1}\right]^2\right\} \tag{3}$$

.

Table 1 represents the performance analysis of the five port structure shown in the Fig. 5 and it tabulates the calculated values of ML and RL in dB.

Figure 9 represents the variation of the losses at the ports which in turn depends on the VSWR values obtained at the respective ports.

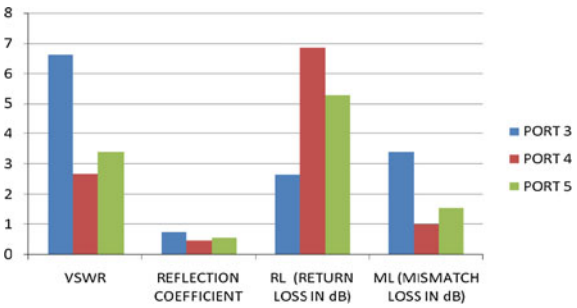


Fig. 9 The variation of RL and ML w.r.t to the ports

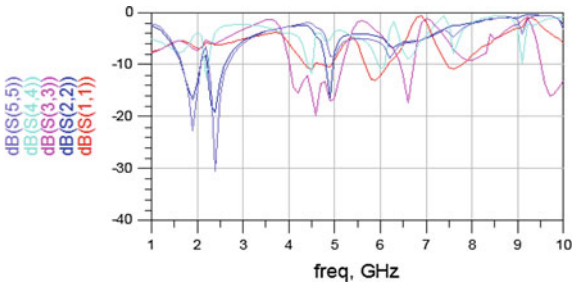


Fig. 10 S-parameter variation w.r.t Frequency

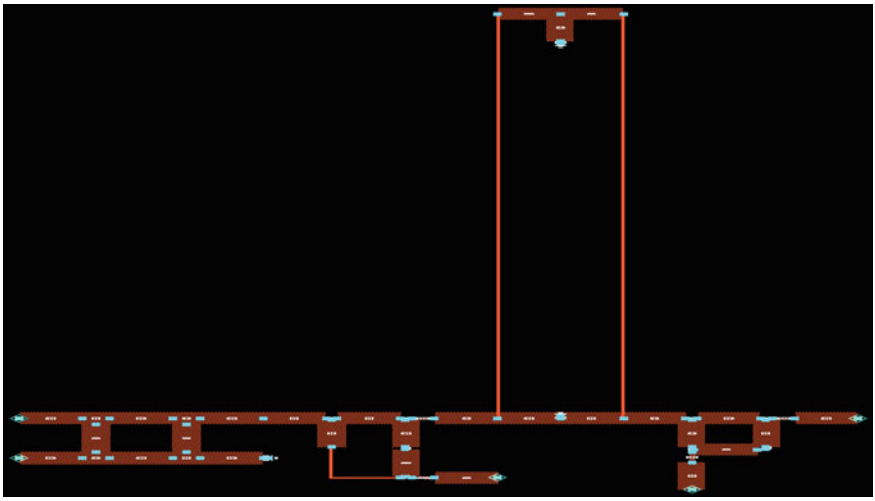


Fig. 11 Layout of the five port structure design

Figure 10 represents the variation of various scattering parameters at 5.8 GHz as center frequency and within a bandwidth of 10 GHz. This variation is well in line with the previous investigations.

Figure 11 represents the layout extracted for the design shown in the Fig. 5 and the hybrid coupler, power divider and phase shifter are very much evident from the layout.

5 Conclusion

These results are very encouraging so that this novel design of five port structure will serve as the efficient alternative for the mixer stage in the RF front end receiver architectures. The variation of the VSWR at the input and output ports were obtained and the return loss as well as mismatch loss was calculated at the output ports. In most of the cases, it is observed that the return loss is more than mismatch loss expressed in dB. This also leads in future works towards reduction of the MMIC area, there by the cost of the production.

References

1. de la Morena-Alvarez-Palencia, C., Burgos-Garcia, M., Rodriguez-Aparicio, D.: Three octave six-port network for a broadband software radio receiver, Proceedings of the Microwave Conference (EuMC), pp.1110–1113 (2010)
2. Wallace, E.K., Ping, Y.S.: Planar micro strip microwave power divider with co-linear output ports. Eng. Res. **13**(1), 10 (1998)
3. Gatti, F., Bozzi, M., Perregrini, L., Wu, K., Bosisio, R.G.: A new wide-band six-port junction based on Substrate Integrated Coaxial Line (SICL) technology. In: Proceedings of the IEEE Mediterranean Electro technical Conference, pp. 367–370 (2006)
4. Karthikeyan, K.V., Vasudevan, N.: Analysis of performance of sigma delta architecture in Software Defined Radio. In: Proceedings of the International Conference on Recent Advances in Space Technology Services and Climate Change—2010, RSTS and CC—2010, Article number 5712826, pp. 164–166
5. Karthikeyan, K.V., Vasudevan, N.: Performance analysis of power dividers of five port structures for SDR application. Nat. J. Electron. Sci. Syst. **2**(2), 7–13 (2011)
6. KishorSonti, V.J.K., Kannan, V.: An analysis of the design and performance of RT duroid substrate-based MODFET low noise amplifier. IUP J. Electr. Electron. Eng. **VI**(1), 49–56 (2013)
7. Mathan, N.: CNTFET based Highly Durable Radix-4 Multiplier using an Efficient Hybrid Adder. Biosci. Biotechnol. Res. Asia **11**(3), 1855–1860 (2014)
8. Nassiri.: Argonne National Laboratory, Lecture on Power Dividers and Couplers 10, Microwave Physics and Techniques UCSB (2003)
9. Pozar, D.M.: Microwave Engineering, Wiley, 3rd edn (2005)
10. Tuttlebee, W., Beach, M., Warr., P., MacLeod, J.: Software Defined Radio: Enabling Technologies, Wiley, pp. 32 (2002)
11. Xu, X., Bosisio, R.G., Wu, K.: Analysis and implementation of software defined radio receiver platform. In: Asia Conference Proceedings on Microwave, vol. 5 (2005)
12. IulianRosu—Phase Shifters www.qsl.net

Optimization of Code Blocking in W-CDMA Based on Modified Solid State Genetic Algorithm (MSSGA)

B. Suresh and K. Murugan

Abstract Orthogonal variable spreading factor (OVSF) codes are with various spreading factors that are used to allocate the variable data rates to users with different bandwidth requirements in 3G wideband code division multiple access (W-CDMA) networks. This paper investigates genetic algorithm based approach for optimization of code blocking and dynamic assignment of OVSF code. The ability of the genetic algorithm used in MSSGA structure build over the improvised crossover, mutation operators and the selection operator includes an age factor criteria that overrides the dominance and diploidy structure for early convergence of the result. The simulation results show that the MSSGA has improved spectral efficiency and reduced code blocking probability to result in high efficiency of code usage and to increase the capacity of the system.

Keywords OVSF · WCDMA · Multi-rate service · Blocking probability · Code assignment · Code blocking · Code allocation · Genetic algorithm

1 Introduction

In 3G Wideband-CDMA (W-CDMA) defined by the International Telecommunication Union (ITU) for UMTS networks, have high throughput, wideband services and flexibility. W-CDMA technology specifications are ready to satisfy the basis for UMTS/IMT-2000 [2]. Technical specifications and Technical Reports for a W-CDMA based 3G mobile systems provided by 3GPP [3]. The W-CDMA protocol structure has 3 layers of the subset of OSI layers namely Physical, Data Link and Network. The layers visualization are termed as physical, transport and logical channels of the W-CDMA system. The physical layer uses the FDD (Frequency Division Duplex) mode which includes the major RF functions like frequency synchronization, time synchronization, spreading and modulation [1].

B. Suresh (✉) · K. Murugan
Ramanujan Computing Centre, Anna University, Chennai, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_6

The data link layer includes the radio link control (RLC) and MAC functionality which are responsible for user date transfer and ciphering and multiplexing with dedicated error correction/flow control and recovery of logic channels onto a common transport channel. Logical channels are one on which the information content will be transmitted. Logical channels are categorized as Broadcast Control Channel (BCCH), Paging Control Channel (PCCH), Common Control Channel (CCCH), and the Dedicated Control and Traffic Channels (DCCH, DTCH).

The interaction between the various layers with the associated channels is shown in Fig. 1. The transport channels support sharing physical resources between multiple services such as data, voice, or signalling with various data rates. W-CDMA is principally based on the spread spectrum the bandwidth that is shared among multiple users concurrently. The user's signals are differentiated through Direct Sequence Spread Spectrum (DS-SS) which uses a discrete code per user to differentiate the user's channel. The original signal is multiplied with a high bandwidth signal known as a channelization code (OVSF) with various wideband services like data, voice, and Multimedia services with variable data rates. Then the signals are scrambled using scrambling code and transmitted through the physical channel. The high bandwidth signal is referred to as a spreading or channelization 'code' symbol called 'chips' [3]. The spread rate is 3.84 Mega-chips-per-second (Mcps) for W-CDMA and a new approach makes use of a multicarrier (MC) CDMA waveform using the OVSF spreading factors. The W-CDMA defines OVSF code tree of 8 layers in the forward link (SFs) with the data rates handled are between R and 256R and in reverse link 512R where R is the basic data rate of 7.5 kbps.

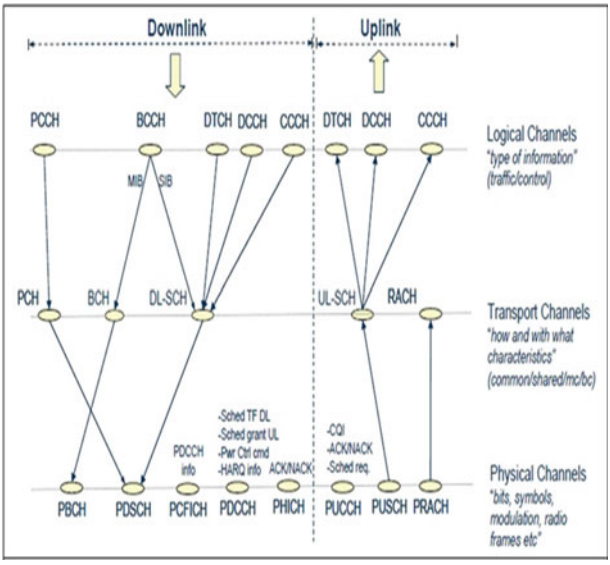


Fig. 1 Interactions between channels

The OVSF codes are orthogonal in nature, two codes are said to be orthogonal if they have zero cross-correlation. The codes are generated using technique of Hadamard transform. Walsh codes are generated by applying Hadamard transform upon 1 repeatedly. OVSF code tree SF is shown in Fig. 2. The OVSF code allotted to the newly requested user, if all the code from root to leaf are free. The code having smaller SF is used for user requesting the higher data rate, so that the overall bandwidth (data rate x SF) of the system is same. The busy codes are the one allocated to the user, the blocked codes are the one orthogonal to the busy codes and cannot be used by the system. The vacant codes are the free codes of the tree remaining can be allocated to the requested user by the system. The existent of the code in the requested data rate will impact the deficiency of the system and introduces the term of call blocking. The efficient reassignment of users in the system of W-CDMA deployment, the higher data rate users such as High Speed Downlink Packet Access (HSDPA) in downlink and High Speed Uplink Packet Access (HSUPA) in uplink are allotted either with the smaller SF or combined higher SFs to enable the data rate requested. Normally in overall capacity of the system only 15–30 percent is used for higher data access users. The dynamic code tree allotment of one sector of NodeB (3G-BTS of one of the ISP (BSNL)) shown in Fig. 3.

The incoming call to the system is not allowed even when the system is having enough capacity but not having the code to support the requested data rate of the incoming user. Due to the nonlinear constraints of the code blocking the soft computing methods have the edge over the analytical methods. For the purpose of overcoming the code blocking complications, small number of studies on exploitation of Genetic Algorithm (GA) for OVSF code assignment has been investigated where efficient reallocation of OVSF code tree with an arbitrary initial population is addressed. Though, GA has been observed to produce good optimization results for the nonlinear systems, problems of convergence and prematurity occurred in GA. The local optima of the solution of GA can be converted to global optima if the

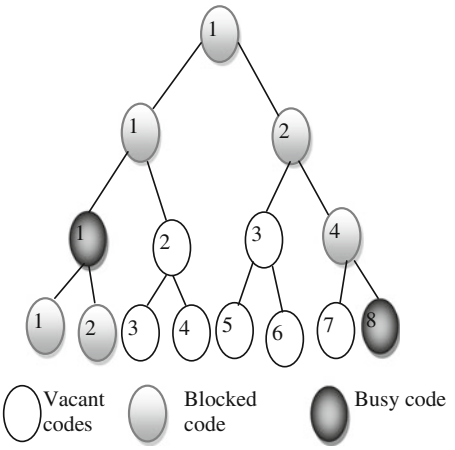


Fig. 2 Code tree spreading factor

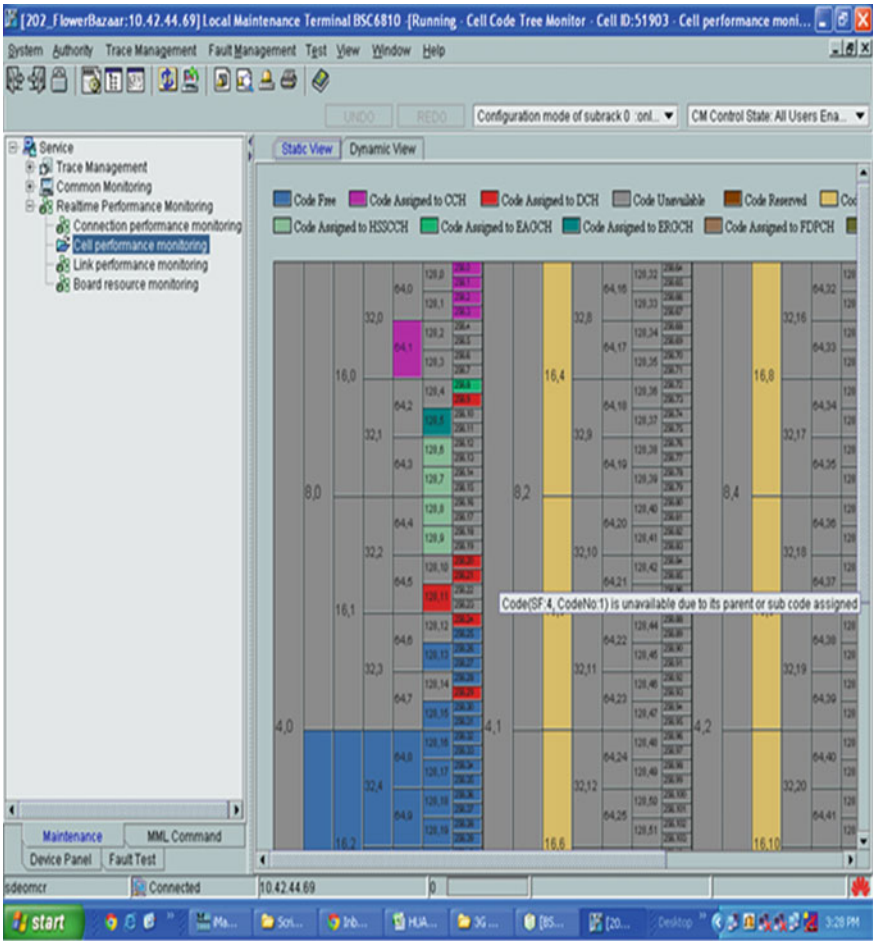


Fig. 3 Code tree of a nodeB of an ISP (BSNL)

selection of GA operators is assigned correctly. This paper introduces the new techniques to overcome the overhead problem discussed in D&D-GA [6]. This paper compares the results in terms of spectral efficiency, blocking probability and convergence delay times of the algorithms of Simple Genetic Algorithm (SGA), Dominance and Diploidy Genetic Algorithm (D&D-GA) and Modified Solid State Genetic Algorithm (MSSGA).

The next section contains the related works and back ground knowledge and Sect. 3 prescribes the SGA, D&D-GA and MSSGA schemes and the simulation results are presented in Sect. 4. The conclusion and the future work are detailed out in Sect. 5.

2 Related Works

The main objective behind the OVSF codes schemes presented earlier is that the reassignment can be possible by rearranging the current user's codes into the code tree such that the fragmented codes are combined to free the code for new users. Tseng and Chao, introduced single code in Ref. [4] and developed the multi-code placement in Ref. [5] which produces the solutions to the code blocking problem, the multi OVSF code replacement scheme [5] produces good results [8]. Minn et al. [6] introduced a dynamic code assignment—DCA which provided the statistical multiplexing and thus improved the spectral efficiency. Balyan and Saini [7] proposed an OVSF code scheme which simplifies the complexity involved in the crowded first assignment. Kavipriya et al. [11] presented the various schemes and techniques; the majority of the papers under comparative study employ the analytical methods to solve the reassignment problem in OVSF code blocking.

Regarding GA adaptation in code blocking optimization Karakoc et al. [8–10] presented the different methods, code tree reassignment with adaptive population size in Ref. [8]. The results show that the D&D-GA provided the better results compared to the conventional (CCA) and dynamic schemes (DCA) [11]. The methods adopted SGA and D&D-GA in Ref. [8] is compared with the proposed system in this paper. The Basic concepts and specifications of GA are detailed in Ref. [12], which illustrates the basic GA operators and their attributes and selection of fitness function and their use to derive the optimization to the given problem. The improvements in GA using adaptive techniques are elaborated in Ref. [13] using the population diversity in crossover and mutation operators and redefined the fitness value. The MAGA modified algorithm for OVSF reassignment is presented in Ref. [14] by Kavipriya et al. The modified enhanced steady state genetic algorithm is explained for scheduling and optimization techniques is adopted to assign externals for viva-vose or examination in other colleges under its jurisdiction by the university [15].

3 MSSGA Based Dynamic OVSF Code Allocation

Genetic algorithm is a heuristic approach that relates the evolution of biological genetic processes, and uses the stochastic optimization methods. Genetic algorithms have many characteristics such as global search, internal parallelism, simple and universal, robust strongly and fast convergence etc. [13]. GAs iteratively updates populations of individuals; the individuals are evaluated using a fitness function. The new population is obtained by probabilistically selecting fitter individuals from the current generation and is subjected to genetic operators such as crossover and mutation to create new offspring. A small number of studies on exploitation of Genetic Algorithm (GA) for OVSF code assignment have been investigated where efficient reallocation of OVSF code tree with an arbitrary initial population is

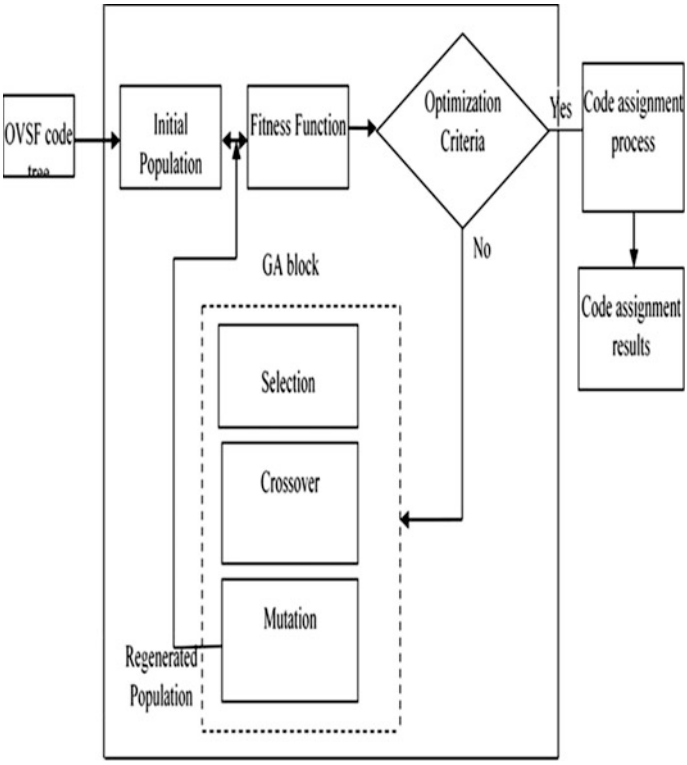


Fig. 4 Block diagram of the system

addressed. The design of GA operators, with the criteria of the OVSF tree, convergences into the global optima of the reassignment with optimal cost. Figure 4 shows the block diagram of the GA based OVSF code reassignment, the execution is not essential for resource assignment in idle state, the calls are allocated as the requested data rate by the resource manager to assign resources for a traffic channel, the call can be supported, if the code available as requested [8]. Due to unavailability of the free code with the data rate in the code tree by the requested user or all orthogonal to the assigned user codes in the system, GA block is invoked for reassignment process of OVSF code.

The index number of the code belonging to active users at that moment in the given OVSF code tree which is input to the SGA block is named as initial chromosome (Chini). Each index number of the active users is called as a gene in respective chromosome. The initial chromosome (Chini) and the age and the data rate are also memorized for further process. The size of initial population is defined according to Eq. (1), with Ref. [8], where population is 2 times the n, 'u'—no of users at the moment,

$$2 \times n = SF - \sum_{i=1}^{U} S(i) \tag{1}$$

Each chromosome in the population is generated using the random permutation with the data rate it originally belongs. The comparison of operator's evaluation of three systems are shown in the Fig. 5. Random permutation is used to generate the population OVSF code tree from 1st gene to last gene, with the consideration the orthogonal principle. Then, the fitness value of chromosomes is evaluated according to fitness function value of the jth chromosome in f(j) and the replaced chromosome of the population is P(j) as defined by Eq. 2.

$$f(j) = \frac{1}{\sum_{i=1}^{V} (Ch_{ini}(j) - P(j,i) \times S(i))} \tag{2}$$

According to the fitness values, the roulette wheel function, the selection operator generates a new population of n chromosomes and they should be hybridized through the crossover operation to attain better attributes. Pairs of chromosomes are chosen from the population with respect to crossover rate (pc). But here we take the advantage of individuals age where old individuals have high Pc and Low Pm while young individuals have low Pc and high Pm. We can improve the crossover and mutation, and we can make their probabilities adapt according to the age of individuals which will participate to gain successful schedule. Depending on the Fitness and Age the value of Pc and Pm will be determined according to the rules in Table 1 and 2 respectively. It was observed that during the initial generations there was good amount of variance of fitness in the population. But as the generations increase the variance reduces resulting in two or more individuals having same fitness levels. To overcome this, pure insertion as per

Fig. 5 Block diagram of systems GA operators

Table 1 Age dependent crossover probability

Crossover (Pc)	Age	Individual I Young	Individual I Old
	Individual II young	Low	Low
	Individual II old	High	High

Table 2 Age dependent mutation probability

Mutation (Pm)	Individual young	Individual old
	High	Low

Table 3 Age Dependent Insertion Probability

Insertion (Pi)	Individual young	Individual old
	High	Low

Fig. 6 Pseudo code for MSSGA

```
Proposed Modified Solid state Genetic Algorithm ()
Initialize 2 times the population;
Evaluate fitness function of the population;
Select the elite individuals;
Crossover operation
{
Select the random Pc;
Assign the age factor;
Adapt crossover probability rule;
Evaluate the crossover;
}
Mutation operation
{
Select the random Pm;
Assign the age factor;
Adapt Mutation probability rule;
Evaluate the Mutation;
}
If fitness is same
{
Adapt insertion probability rule;
}
While convergence not achieved
{
Scale population fitness;
Select solutions for next population;
Perform improved crossover and mutation probability;
Evaluate population;
}
}
```

Table 3 will simply replace the any one of the parent from the population having the maximum fitness with the new child which is having lower fitness. If an OVSF code tree denoted by best chromosome, can allocate the requested data bit rate to appropriate user. Once the optimization criterion is verified and requested data rate is allocated to users. If not, other chromosomes in the population are checked.

As stated, the D&D-GA has considerable overhead and convergence time as two times the population of SGA is involved. In MSSGA the elitism population is generation and selection process are same as in the D&D-GA. The major difference in MSSGA is “age” for every individuals in the population. This value is initially

zero for a new individual but is gradually incremented depending on its survival in each generation. Old age of an individual means that individual has an effective impact on successful schedule which avoids hard constraints. The pseudo code for the proposed MSSGA is shown in Fig. 6. In contrast, young age individuals do not effectively participate in schedule process. Now two individuals may have same fitness but different age. Individuals having same fitness but high age should be replaced last during insertion as they have survived more number of generation and have a higher potential of providing better solutions in the next generations. Thus individuals having same fitness the probability of insertion is kept high for low age and low probability of insertion for individuals with high age.

4 Performance Evaluation

Main objective of the optimizations is to increase free codes by reassigning the fragmented codes. In order to assess the reassignment process of the GA based code reassignment types of SGA and D&D-GA and MSSGA by simulations.

In the simulations parameters are used as Poisson process is used for to determine the call arrival while the mean arrival rate l calculated from 4 to 64 calls/unit. For Call duration parameter exponential distribution of 0.25 units of time mean value is used. 256 is taken as Maximum spreading factor. Possible OVSF code rates are generated using Uniform distribution between R and SF × R.

4.1 Blocking Probability

Call blocking probability which is the ratio between the numbers of blocked calls to the incoming call requests. It is shown in the Eq. (3), where N_B is the number of

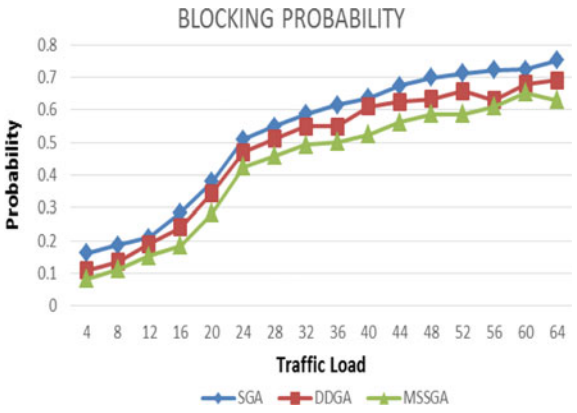


Fig. 7 Blocking probability

blocked calls and N_T is the total number of all incoming calls, which is given by Eq. 3. The results show that the blocking probability of MSSGA is improved to 22 and 12 % SGA and D&D-GA respectively as shown in Fig. 7.

$$Pr(blocking) = \frac{N_g}{N_T} \tag{3}$$

4.2 Spectral Efficiency

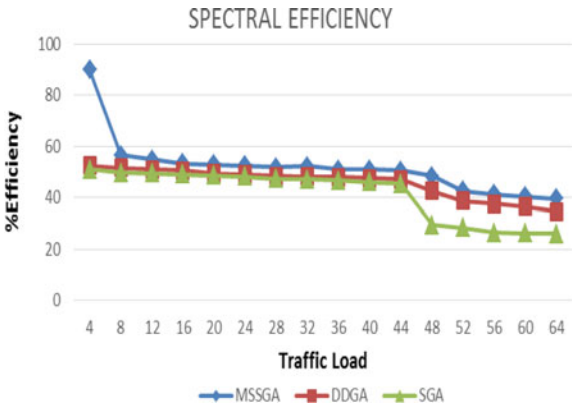
Spectral Efficiency is the ratio between the Rassigned – (assigned data rate) and the total Rrequested data rates– (all incoming calls data rate), which is given by Eq. 4. The results show that the spectral efficiency of MSSGA is improved to 25 % and 11 % SGA and D&D-GA respectively as shown in Fig. 8

$$n(\%) = \frac{R_{assigned}}{R_{requested}} \times 100 \tag{4}$$

4.3 Delay of Optimization

The time delay between the new user and the code blocking to the reassigned code for the new user. The time measurement of code allocation to the new user, who is blocked earlier. This factor is a crucial on in selecting the algorithm because it adds the overhead of the system. The result shows that the MSSGA delivers 11 % reduction in time delay compared to D&D-GA as shown in Fig. 9.

Fig. 8 Spectral efficiency



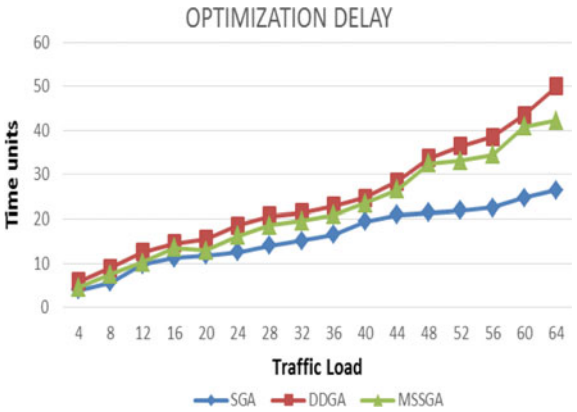


Fig. 9 Optimization delay

Even the cost involved in the reassignment also may be considered and the MSSGA has edge over the others with the effect that it improves the spectral efficiency that is the no of users allowed is higher and converges the optimum reassignment time.

5 Conclusion and Future Work

The research work focuses on code blocking avoidance through Genetic algorithm optimization approaches. The orthogonal property of OVSF codes makes more appropriate for W-CDMA. OVSF codes assignment have high influence on the code utilization and system performance. This research work utilizes three efficient Genetic algorithms namely MSSGA, D&D-GA and SGA for dynamic OVSF code assignment in W-CDMA systems to reduce the call blocking. The spectral efficiency of the system also increased with increased user code allotment. The simulation results exhibits that the MSSGA approach provides the lesser blocking probability and increased spectral efficiency of the system compared with SGA and D&D-GA approaches. The future work of this approach would be to use hybrid heuristic optimization algorithm to seek better results in terms of call blocking and spectral efficiency. The main limitation of the research work is the utilization of optimization algorithms. Hence, the future enhancement of this research work would be to use hybrid optimization approaches. In future work optimized assignment of code is performed based on the time slots for different order of priority within priority users like HSUPA and HSDPA, which will measure the probability density function and cumulative distributive function.

References

1. Spreading and modulation (FDD), Technical specification TS.25.213 V4.2.0, http:// www.3gpp.org1999
2. Dhalman, E., Gudmunson, B., Nilsson, M., Skold, J.: UMTS/IMT-2000 based on wideband CDMA. IEEE Commun. Mag. **36**, 70–80 (1998)
3. Holma, H., Toksala, A.: W-CDMA for UMTS. Wiley, Nokia, Finland (2000)
4. Tseng, Y.C., Chao, C.M.: Code placement and replacement strategies for wideband CDMA OVSF code tree management. IEEE Trans. Mob. Comput. **1**(4), 293–302 (2002)
5. Chao, C.M., Tseng, Y.C., Wang, L.C.: Reducing internal and external fragmentations of OVSF codes in W-CDMA systems with multiple codes. IEEE Trans. Wirel. Commun. networking **1**(2003), 693–698 (2005)
6. Minn, T., Kai, Y.S.: Dynamic assignment of orthogonal variable spreading factor codes in w-CDMA. IEEE J. Sel. Areas Commun. **18**(8), 1429–1440 (2000)
7. Balyan, V., Saini, D.S.: Flexible assignment of OVSF Codes for data calls in CDMA wireless networks. International Conference on Communication Systems and Network Technologies (CSNT), pp. 5–9 (2011)
8. Karakoc, M., Kavak, A.: Genetic approach for dynamic OVSF code allocation in 3G wireless networks. Elsevier on applied soft computing **9**, 348–361 (2009)
9. Karakoc, M., Kavak, A.: Using diploidy genetic algorithm for dynamic OVSF Code allocation in W-CDMA networks. Radio and Wireless Symposium, pp. 15–18 (2007)
10. Karakoc, M., Kavak, A.: Genetic algorithm and simulated annealing for dynamic OVSF Code allocation in W-CDMA Networks. IEEE 15th Signal Processing and Communications Applications, pp. 1–4 (2007)
11. Kavipriya, P., Gomathy, C.: A comparative study of OVSF Code allocation techniwues in 3G wireless networks. Indian J. Appl. Eng. Res. **1**(1), 127–136 (2013)
12. Elisaveta, G., Natasha, G.: BASIC—a gentic algorithm for engineering problems solution. Elsevier comput. chem. Eng. **30**(8), 1293–1309 (2006)
13. Youchan, Z., Feng, S.: An improvement adaptive genetic algorithm. IPCSIT **43**, 113–117 (2012)
14. Kavipriya, P., Gomathy, C.: An efficient dynamic orthogonal variable spreading factor code allocation approach in WCDMA throguh modified adaptive genetic algorithm. Maxwell J. **7** (12), 2545–2553 (2013)
15. Shiburaj, P., Talele, K.T., Mehul, K.V.: Modified enhanced steady state genetic algorithm for scheduling & optimization problems, IEEE Annual Conference, pp. 1–5 (2013)

Augmented Reality Simulation to Visualize Global Warming and Its Consequences

K.R. Parvathy, M.L. McLain, Kamal Bijlani, R. Jayakrishnan and Rao R. Bhavani

Abstract Augmented Reality (AR) technology is considered to be an important emerging technology used in education today. One potentially key use of AR in education is to teach socio-scientific issues (SSI), topics that inure students towards social conscience and critical thinking. This work uses multiple markers and virtual buttons that interact with each other, creating a life-like visual spectacle. Learning about issues such as global warming by using AR technology, students will have an increased sense of experiencing immersion, immediacy, and presence, thereby enhancing their learning as well as likely improving their ability to make better informed decisions about considerations of such issues. Another advantage of AR is that it is a low cost technology, making it advantageous for educators to adapt to their classrooms. Also in this work we compare the effectiveness of AR versus ordinary video by polling a group of students to assess the content understandability, effectiveness and interaction of both the delivery methods.

Keywords Augmented reality · Global warming · Multiple markers · Virtual buttons

1 Introduction

According to Johnson et al., among the different types of technologies that can be applied to education, AR is among the leading emergent technologies that will be utilized in the next five years [1]. The work conducted here demonstrates that AR can be a successful tool used for educational purposes. Socio-scientific issues

K.R. Parvathy (✉) · K. Bijlani · R. Jayakrishnan
Amrita E-Learning Research Lab, Amrita Vishwa Vidyapeetham University, Amritapuri Campus, Kollam, Kerala, India

M.L. McLain · R.R. Bhavani
Ammachi Labs, Amrita Vishwa Vidyapeetham University, Amritapuri Campus, Kollam, Kerala, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_7

(SSI) are important in education as they foster social conscience as well as encourage critical thinking. SSIs involve the purposeful use of scientific topics that entail students to hold in dialogue, discussion, and debate. They are usually divisive in nature but have the added constituent of requiring a degree of ethical reasoning or the assessment of moral concerns in the development of arriving at decisions concerning possible motion of those issues [2].

Global Warming is a crucial environmental threat and demands special attention. Students of today need to be well-informed because the proportion of GHGs (Greenhouse Gases) in the atmosphere increases significantly each year, thus the effects of Global warming will continue to increase, with potential to be seriously deleterious in the not-too-far future. Teaching by AR will allow the learners to have an almost direct virtual experience by seeing the process and the effects in a 3 dimensional lifelike view. Research shows that this type of experience increases learning by virtue of the immersion and immediacy of this method. Until now relatively little has been done in this field, which makes doing this type of work of particular importance [3].

Learning about issues of socio-science such as global warming by using AR technology, students will potentially improve their ability to make better informed decisions about considerations of such important current issues [4].

At present the concept of global warming is typically conveyed with the help of 2 dimensional medias such as animated videos and images. But often these media do not give the full impact of the adverse effect of global warming because of the limitations of 2 dimensional media. By teaching these concepts using the visual appeal of AR, the learners tend to be more attentive, thereby increasing their potential to retain the presented information.

Augmented reality (AR) combines both real world and digital data. At present, most AR research uses live video images, which the system processes digitally to add computer-generated graphics. In other words, the system *augments* the image with digital data [5]. Today Augmented Reality is used in many areas. Processor, display, sensors and input devices are the hardware components used for Augmented Reality. The objective of the project is to study the use of Augmented Reality in teaching and learning methodology. AR offers opportunities for more reliable learning and appeals to many learning styles, providing students an extra personalized and explorative learning experience [6].

The work includes the technology of Augmented Reality in explaining a global phenomenon. Here, using the technology of Augmented Reality, the satellite views of earth and sun are projected, rays emitted from the sun, detailed specification of the earth's technology of augmented reality in explaining a global phenomenon. Using Virtual Buttons the consequences are shown, so that learners could have the feel of actually experiencing the effect of global warming. Virtual buttons allows the user to interact directly.

Augmented Reality can be used with different game engines with the help of SDK. Unity 3D is the game engine used in the proposed system. We can do augmented reality applications using unity3D. Different SDKs used in Augmented Reality are Catchoom, Metaio, Obvious Engine, PointCloud, Sphero, SSTT,

Vuforia, Xloudia, NyARToolkit. The system needs external SDK to use Augmented Reality in Unity 3D. The proposed system used Vuforia SDK. The Vuforia SDK allows building vision-based augmented reality applications for the Android platform. The basics used in augmented reality are Marker Detection, Transform Mapping, and Rendering.

Marker based tracking is used in the proposed system. Knowledge and information on hard-copy book can be readily and interactively manipulated through images and 3D animation in the form of markers [7]. The marker is located mainly by using the device webcam. The content should be hidden behind the marker, unless the marker is shown to the web cam. In order to find a predetermined marker the webcam monitor the physical reality near the user. The positions, rotation, scale of the marker are interpreted by the webcam. Then the 3D model is being updated and changed to the screen. In Augmented Reality instead of using main camera, AR camera is used. Target image is tracked using the AR camera. Target image is nothing but the marker.

The use of an AR system provides several benefits over traditional teaching methods. One of the most important benefits is the stimulation of several sensory modalities: touch, sight and hearing. As a consequence, it makes students actively involved in the learning process. With traditional teaching techniques, students only receive information by one sense at a time [8].

The above Fig. 1 shows an example of a marker technology [9], which is used to load the real object. Marker technology is used in the system. The input given is the marker, which is cost effective and the expected output received is the real visualization of an object. Marker based AR system allows user to input predesigned markers, and create new scenes blending together the real world and computer generated 3D images [9].

Fig. 1 Example of marker

1.1 Augmented Reality Toolkit

Augmented Reality Toolkit is a well-liked planar marker structure for Augmented Reality and Human Computer Interaction (HCI) systems due to its accessible source code. The first stage of the recognition procedure is finding the markers black borders that are finding connected groups of pixels below a certain gray value threshold [10].

2 Related Work

Zhu et al. [6] describes the various functionalities and use of augmented reality (AR) in the field of healthcare education is explored in this paper. Students and medical professionals require more situational experiences in medical care, particularly for the sake of patient protection there is an obvious need to additional learning the employ of AR in healthcare knowledge. It improves the learning effect by the acquisition of skills and knowledge, understanding of medical concepts, enhancing learning retention, given material in a suitable and appropriate manner. It seems that AR is very useful in the field of health care education.

Kaufmann [8] describes the impending and challenges of using collaborative Augmented Reality (AR) in learning contained by the better circumstance of immersive virtual learning environments. The paper takes an instance of a collaborative AR claim purposely intended for mathematics and geometry learning called Construct3D.The refuse in hardware costs, the use of mobile immersive virtual or augmented reality become practicable for educational institutions within this decade. The augmented reality or virtual reality features need to be carefully reflected in translating into educational efficiency.

Ivanova and Ivanov [11] describe the benefits of Augmented Reality (AR) for the educational sector and in particular for a computer graphics course using low-cost interactive marker augmented reality technology. The innovation of the solution is that it can propose students high interactive human-computer interface for model manipulation and thus observing the details in 3D space. Several of the students like AR technology, because there presentations are portable and easy to make, allowing them to observe the structures in more details and also to get a clear image. The AR technology can be applied in self-paced learning, where individual learners are able to manage their directions of exploration as well as in group-based learning where communication, ideas sharing and interaction among participants are the main methods for learning.

Cai et al. [12] discusses about the importance of AR tool in improving middle school students cognitive test performance on corresponding content, and has relatively larger influence of low-achieving students.AR tool is an innovative interactive technology used especially in subjects like chemistry. AR tool can act as remedial learning tool to express the contents in a middle school chemistry course

so that the students can memorize chemical structures and concepts. Students could control, combine and interact with the 3D model of micro particles by means of markers and perform a sequence of inquest based experiments.

Chang et al. [13] explored whether an online SSI unit improved by AR can get better students' accepting of the science content concerned. The reason of the study was to reveal a work in development of how to get benefit of mobile AR affordances to plan curricular actions that speak to significant science education goals. Results showed that learning was significantly improved using AR technology.

3 Solution Approach

This study demonstrates the effect of global warming with the help of Augmented Reality. Unity 3D is a novel portion of technology that strives to create life superior and easier for game developers. Unity is a game engine or a game authoring instrument that enables imaginative people to build games [14]. The system needs external SDK to use Augmented Reality in Unity 3D. Vuforia SDK is used in this study.

The proposed system gives awareness to people about the impact of global warming. Being a different technology most of the people will have keen interest to know about it. There are many different ways to classify the different augmented reality; they can be separated by function, hardware and software requirements [15].

Vuforia is the software stage that enables the most excellent and most imaginative recognized augmented reality (AR) app experiences across the majority authentic world environments [16]. Vuforia uses image identification to sketch and record image marker [17]. The study shows the use of Augmented Reality in teaching and learning methodology. The proposed system shows the impacts of global warming. Marker technology is used to detect different objects. Unique markers are used to show every different object. For example, to show the picture of earth, a particular marker is used and to show an image of the sun, another marker is used. For each object each marker is used. For example, one marker is used to represent the earth model and another is used to represent the sun model.

In Augmented Reality, contents are hidden behind the marker image. The webcam connected to computer captures the "marker" reads the information/pattern encoded in it and sends this information to the computer. The information is being recognized and the marker is replaced with an image encoded in it. The movement of the marker is being tracked by the computer and adjusts the size of the image according to the movement.

When using this technology, it is easy for the people to know the causes of global warming, the reasons behind why the earth is getting heated up, the effects of global warming. This helps to bring realization in people that these are caused by their human activities and they will try to reduce these harmful activities which cause global warming.

The above Fig. 2 illustrates the augmented reality system. The system captures and detects the target image (marker) and renders the real object (3D objects, graphics) on the top of the target image and displays it. When the target image is exposed to the camera, the camera captures and detects the image. And then the image is being tracked and compared with the image stored in the database (Vuforia target manager tool). When the tracked image is detected in the database, it renders the real object corresponding to the target image.

The flow diagram shows in Fig. 3 demonstrate the cause of global warming with the help of Augmented Reality. Marker technology is used to detect different objects. The markers detect the sun and the earth model simultaneously. When the marker is shown to the webcam the real objects will be loaded.

The flow diagram shows in Fig. 4 demonstrate detailed explanation of the effect of global warming. Virtual Buttons are used to show the consequences of global warming. Using the marker containing two virtual buttons it shows the consequences of rays on the layers of the Earth.

When the marker is detected by the webcam, the particular area is detected. After detecting the area the real object is loaded. The same process is repeated for the other markers. So by explaining the consequences of global warming using Augmented Reality, its impact could be more pronounced. AR enables learners to

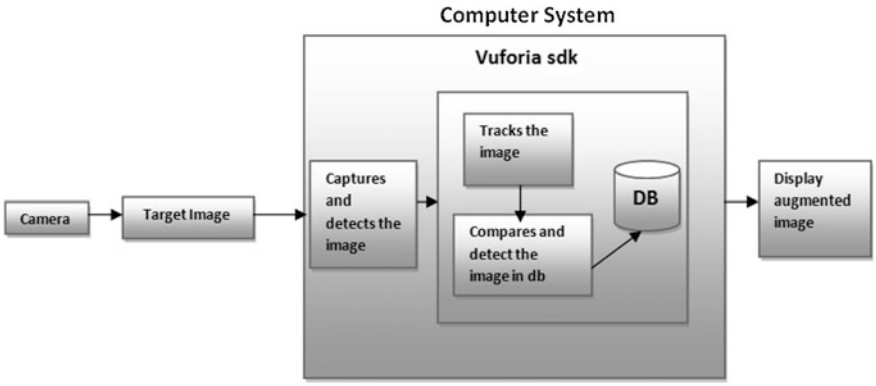


Fig. 2 Flow diagram for an augmented reality system

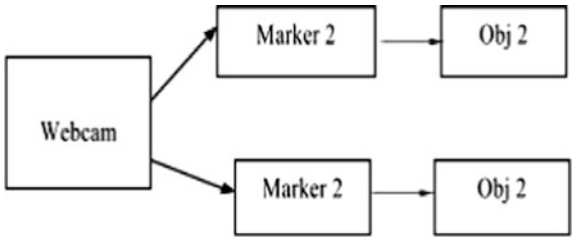


Fig. 3 Flow diagram to load the sun and earth model

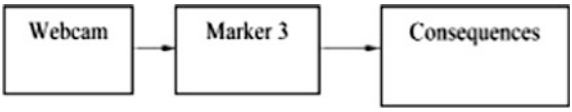


Fig. 4 Flow diagram to load the consequences of global warming

extend the attention span and make unseen more tangible. Thus the demonstration of global warming is made more effective and thought provoking.

The first phase shows the earth model with single marker technology. The earth and the sun models were also loaded with multiple marker technology. The Target manager tool is used to create the target image. Target manager tool extracts the features of the target image and stored in the database. Target image is tracked using the webcam and then the virtual content or the object is loaded. Within seconds the sun emits the rays that results the color change of the earth. The consequences happened due to the result of human activities are also shown using the virtual buttons.

The above Fig. 5 shows the screenshot of earth's heat up. Due to the rays from the sun the earth gets heated up. That results in Global Warming. Multiple marker technology is used to demonstrate the above process.

The above Fig. 6 shows that the human activities are the major cause of global warming. Here carbon dioxide is released to the atmosphere when humans chop down the trees.

4 Performance Analysis

The above Fig. 7 illustrates the statistical analyses of a comparative study of the AR system with ordinary video. The analyses redone on the basis of a survey conducted in a batch of students. The students were divided into two groups and were

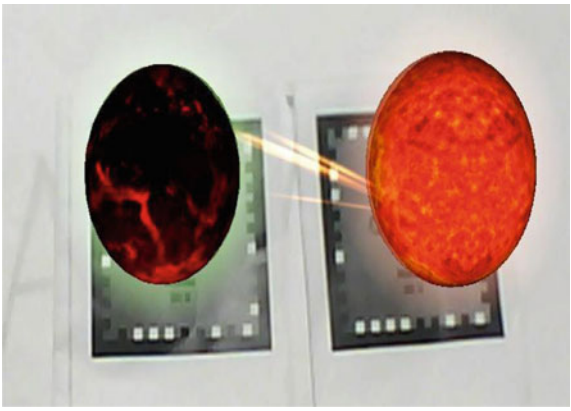


Fig. 5 Screenshot showing the effect of global warming

Fig. 6 Screenshot showing carbon dioxide emission from trees

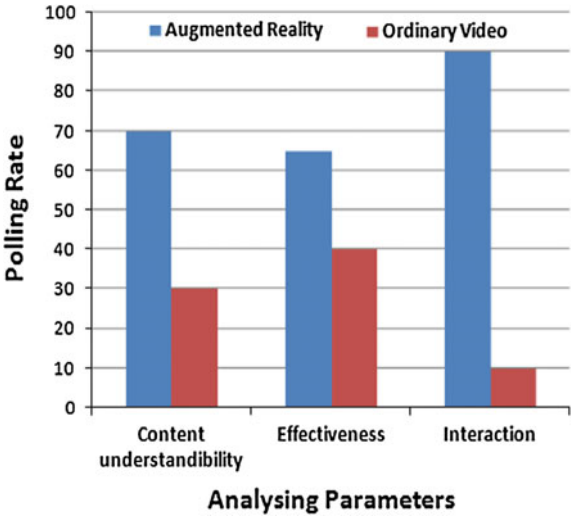


Fig. 7 Comparison of augmented reality and ordinary video

familiarized with both ordinary videos as well as the Augmented Reality. Polling was conducted to analysis the content understandability, effectiveness and interaction of both the delivery methods. Most of the student's gave their approval for Augmented Reality. The study demonstrates that the proposed work gives an impact upon the user as how much effectively the technology of Augmented Reality can be used in the educational system. The proposed system is beneficial for educational purpose.

5 Conclusion

This study demonstrates the use of multiple marker Augmented Reality as an educational tool to teach an important socio-scientific issue: Global Warming. This emergent technology incorporates the use of multiple markers and virtual buttons that interact with each other giving a life-like sense of direct experience to the learner. The visual impact of the display creates an immersive experience that helps learners understand the deeper implications of SSIs which leads to more improved critical thinking and ability to discuss and debate suchlike issues.

We use AR technology to demonstrate the cause and adverse effect of global warming by showing how the radiation from the sun affects the earth. Multiple markers and virtual buttons are used in the work so that interactions and visual impact are more effective.

The study reveals that AR-based modeling is a more successful teaching tool than video based modeling by analyzing a survey conducted which showed increased interest and comprehension of content in AR technology over ordinary video.

Acknowledgments We are grateful for the support and facilities provided by the Amrita E-Learning Research Lab and Amrita Vishwa Vidyapeetham University.

References

1. Johnson, L.F., Levine, A., Smith, R.S., Haywood, K.: Key emerging technologies for elementary and secondary education. Educ. Digest **76**(1), 36–40 (2010)
2. Zeidlera, D.L., Nichols, B.H.: Socio scientific issues: theory and practice. J. Elementary Sci. Educ. **21**(2) (2009)
3. Bronack, S.C.: The role of immersive media in online education. J. Continuing High. Educ. **59** (2), 113–117 (2011)
4. Klopfer, E.: Augmented Learning: research and design of mobile educational games. MIT Press, MA (2008)
5. Hirzer, M.: Marker Detection for Augmented Reality Applications, ICG-TR-08/05Graz, October 27 (2008)
6. Zhu, E., Hadadgar, A., Zary, N.: Augmented Reality in Healthcare Education: an integrative review, PeerJ on 8 July (2014)
7. Chang, R.-C., Chen, S.-N., Lin, H.-J., Yu, H.-M.: DUIRA: an interactive learning platform for mixed reality. International Conference on Multimedia and Expo (ICME) (2010)
8. Kaufmann, H.: Collaborative augmented reality in education, Inst. Softw. Technol. Interact. Syst. (2014)
9. Chang, R.-C., Chung, L.-Y.: A study on augmented reality applications in situational simulation learning. International Conference on Ubi-Media Computing and workshops (2014)
10. Perez Lopez, D., Contero, M.: Delivering educational multimedia contents through an augmented reality: a case study on its impact on knowledge acquisition and retention. J. Educ. Technol. October (2013)
11. Ivanova, M., Ivanov, G.: Enhancement of learning and teaching in computer graphics through marker augmented reality technology. Int. J. New Comput. Architectures Appl. (2011)

12. Cai, S., Wang, X., Chiang, F.-K.: A case study of augmented reality simulation system application in a chemistry course. Comput. Human Behav.(2014)
13. Chang, H.-Y., Wu, H.-K., Hsu, Y.-S.: Integrating a mobile augmented reality activity to contextualize student learning of a socio scientific issue. British J. Educ. Technol. (2013)
14. Creighton, R.H.: Unity 3D Game Development by Example, pack publication, ISBN 13:978-93-5023-278-1 (2010)
15. Hawkinson, E.: Informal learning environments in tourism: augmented reality, International Academic Conference (2014)
16. Augmented Reality (Vuforia), https://developer.qualcomm.com/mobile-development/add-advanced-features/augmented-reality-vuforia
17. Xiao, C., Lifeng, Z.: Implementation of mobile augmented reality based on Vuforia and Rawajali. International Conference on Software engineering and service science (2014)

Design and Modeling of Microfluidic Channel in a Dielectric Planar Waveguide Using Cosmol Multiphysics

K. Asha, N.K. Suryanarayana, K. Narayan and P.K. Pattnaik

Abstract Integration of Micro Electro Mechanical System (MEMS) with integrated optics is known as Micro-Opto-Electro-Mechanical-System (MEOMS). There is a wide development happening in the Lab-On-a-Chip fabrication industry for bio-medical application. In this paper we demonstrate simulation and modeling microfluidic channel embedded in a dielectric planar waveguide. As a fluid constituent the core of the waveguide is made up of glass and cladding consists of air. The model designed can be used as Refractive Index based sensor. In this simulation velocity is kept constant throughout the microfluidic channel. The pressure at the inlet is higher compared to the outlet.

Keywords Laminar flow · Modal analysis · Electromagnetic waves · Lab-On-a-Chip · Micro-Opto-Electro-Mechanical-System · Dielectric planar waveguide

K. Asha (✉) · N.K. Suryanarayana · P.K. Pattnaik
Department of ECE, Sai Vidya Institute of Technology, VTU,
Bengaluru 560064, India
e-mail: asha.k@saividya.ac.in

N.K. Suryanarayana
e-mail: Suryanarayana.nk@saividya.ac.in

P.K. Pattnaik
e-mail: pkpattnaik@hyderabad.bits-pilani.ac.in

K. Narayan
Department of EEE, BITS Pilani, Hyderabad Campus, Secunderabad,
Telangana State 500078, India
e-mail: deanrnd@saividya.ac.in

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_8

1 Introduction

The Micro Opto Electro Mechanical System (MOEMS) is a wide area having a lot of research potential. MOEMS devices can be used as a Lab-On-a-Chip (LOC) for bio-medical application. The Dielectric slab waveguide is one of the simplest optical constructions used in the integrated optics based devices. This planar waveguide consists of the three layers namely middle layers is having core with refractive index of 1.47 (glass) and the surrounding two layers are called cladding with refractive index 1 (air medium). The planar waveguide simulation is explained that can be used as Refractive Index sensors.

The photonic applications mainly use Slab waveguide [1]. A single cell of bacteria or a particle can be noticed and studied on the optofluidic platform as in [2]. The waveguide fabricated using the Micro Opto Electro Mechanical System technology (MOEMS). The MOEMS based devices are used in the industry for the various applications such as bio-medical, structural, fluid flow and Plasmon resonance sensors. A single particle can be identified using the planar waveguide [3]. Sensitivity analysis of the Evanescent field based sensors is explained [4]. These optical waveguide arrangements can be used as a Lab-On-Chip based bio-sensors.

Waveguide component with microfluidic channel can be used to detect Deoxyribo Nucleic Acid [3]. In the microfluidic channel, the flow rate is very significant measuring constraint depends on the density of the liquid used. Waveguide with a microfluidic channel can be used as sensor for bio medical applications [3, 4]. The reactive chemicals in the environment can be monitored by colorimetric microfluidic devices as designed [5]. Planar slab waveguide are suitable for use in the mili- meter wave region for both communication and bio-medical application.

The Eigen value problems for planar waveguide demonstrate the modal analysis of planar optical waveguide based on the wave equation method. Modal analysis method is used to solve the mode propagation constant by solving second order differential equations [6, 7]. The solutions are calculated by using suitable boundary conditions. Finite element method is used to find numerical solutions for practical situations like optical loss or optical gain involved.

A General Approach towards derivation of Eigen equation for planar slab optical waveguides is explained for the single clad/core devices and also to the multiple clad/core devices with proper assumptions and derivations [8]. Transmission line methods can be used to solve the planar waveguide. Micro-Optofluidic system is a new branch of system where interaction of fluid and light wave is involved.

2 Waveguide Modal Analysis

Planar waveguides has homogeneous wave equation. A profile with uniform Refractive index in core is chosen. The Electromagnetic wave fields are confined in core region and decays exponentially in the cladding region, the electric field

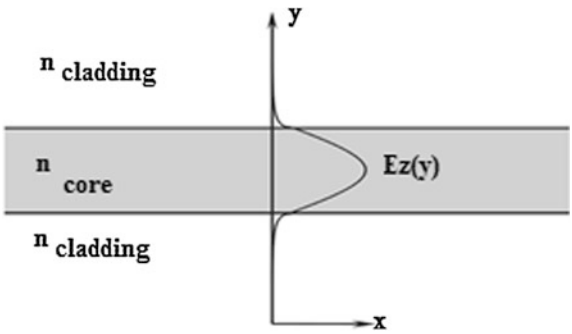


Fig. 1 Evanescent wave pattern in the waveguide

pattern is expressed [9, 10]. The modes are transverse Electric (TE) and Transverse Magnetic (TM). For linear waveguides in the XY coordinates use the Gauss' law [11, 12] where E and H given by equation (Fig. 1)

$$\nabla(\epsilon \mathrm{E}) = 0 \tag{1}$$

$$\nabla \cdot \mathrm{E} = -(\nabla\epsilon/\epsilon) \cdot \mathrm{E} \tag{2}$$

$$\nabla \cdot \mathrm{H} = 0 \tag{3}$$

Here modal analysis of electromagnetic waves in frequency domain is done. There are two ports one is input port where wave excitation is ON. And the other port is output port where the wave excitation is OFF.

$$\mathrm{E_z} = \mathrm{E(y)exp}(-i k_x \mathrm{x}) \tag{4}$$

$$\mathrm{E(y)} = \begin{cases} C_1 \cos(\mathrm{k_y y}) & \text{for core region} \\ \mathrm{C_0\, exp}(-\alpha\left(|\mathrm{y}| - \left(\frac{\mathrm{h}_{slab}}{2}\right)\right) & \text{for cladding region} \end{cases} \tag{5}$$

Here the electric and magnetic fields must be continuous at the interface, for that reason the guidance condition is

$$\alpha = \mathrm{k_y} \tan\left(\mathrm{k_y} \frac{\mathrm{tslab}}{2}\right) \tag{6}$$

where $\mathrm{k_y}$ and α satisfy

$$\mathrm{k_y^2} = \mathrm{k_{core}^2} - \mathrm{k_{cladding}^2} = -\alpha^2 \tag{7}$$

With

$$k_{core} = 2 * \pi * n_{core}/\lambda \tag{8}$$

$$k_{cladding} = 2 * \pi * n_{cladding}/\lambda \tag{9}$$

The Eqs. (6) and (7) can be solved using the Newton-Raphson method. Generally this method is used for non-linear system.

2.1 Laminar Flow

If the flow of fluid is uniform throughout the entire core region of the waveguide then it is called as laminar flow. The Velocity of the fluid is uniform in the microfluidic channel. For Laminar flow, to study the mathematical analysis we use the Navier-Stokes equation in the time domain. The computational fluid dynamics problems are solved by using the Navier–Stokes equations. The flow of water in the waveguide is defined by these equations. Navier-Stokes equations are used to solve Many engineering problems. The application of these equations includes modelling of the water flow in a medium, air flow around the wing and blood flow in a static magnetic field [13].

$$\rho \frac{\partial u}{\partial t} \nabla \cdot \mu\left(\nabla u + (\nabla u)^T\right) + \rho u \cdot \nabla u + \nabla p = 0 \tag{10}$$

$$\nabla \cdot u = 0 \tag{11}$$

In Eqs. (10) and (11) are the Navier-Stokes equation.
where ρ is density (kg/m3), u is velocity (m/s), μ denotes dynamic viscosity (Pa·s), and p is pressure (Pa).

3 Design of Planar Waveguide

The simulation of planar waveguide using COMSOL Multiphysics is explained in this paper. The parameter used for the simulation is shown in Table 1. planar waveguide is having many application in the Lab on a Chip fabrication industry. The planar waveguide geometry is shown in Fig. 2; it contains the core and cladding regions, core being the Glass material, cladding is the air. A slab waveguide of thickness 1 μm is used. The light wave of wavelength (λ) 1550 nm is used here. The fluid in this case is water (other fluid blood can be used), with the corresponding density and viscosity values as shown in Table 2. The boundary conditions assumed for the inlets and the outlet are set pressures 1 kPa;

Table 1 Parameters used for planar waveguide

Name	Global parameters	
	Expression	Value
lambda0	1.550 (μm)	1.550E-6 m
n_core	1.47	1.4700
n_cladding	1	1.0000
h_core	1 (μm)	1.0000E-6 m
h_cladding	5 (μm)	5.0000E-6 m
w_slab	12 (μm)	12.0000E-6 m
k_core	2*pi[rad]*n_core/lambda0	6.0805E6 rad/m
k_cladding	2*pi[rad]*n_cladding/lambda0	4.0537E6 rad/m
f0	C_const/lambda0	1.9341E14 l/s
D	1*exp(−10)	1.0000E-10

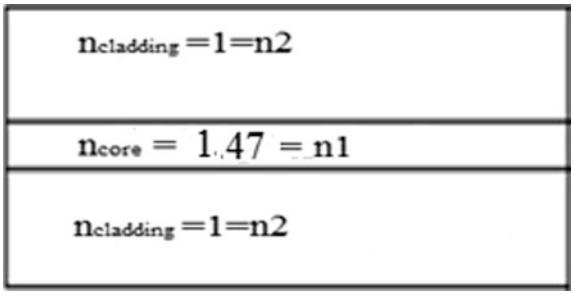


Fig. 2 Geometry of the planar waveguide

Table 2 Material contents for liquid

Name	Parameter	
	Symbol	Value
Density (pressure)	rho (pa)	1000 kg/m³
Dynamic viscosity	mu	1e-3 Pa·s

3.1 Finite Element Method

Finite Element Method (FEM) can be used solve by the boundary condition problems for a given partial differential equations. FEM is a numerical technique, where complete problem is subdivided into Finite Elements; the different types of mesh used for three-dimensional and two-dimensional geometry. The triangular mesh is used for two dimensional geometric modelling whereas tetrahedron mesh used for the three dimensional problems. The element size parameter for the cladding region is given by expression lambda0/n_cladding/8. Lambda0/n_core/8 is the element parameter size expression for the core region. The Fig. 3 shows the triangular mesh applied to the dielectric planar waveguide.

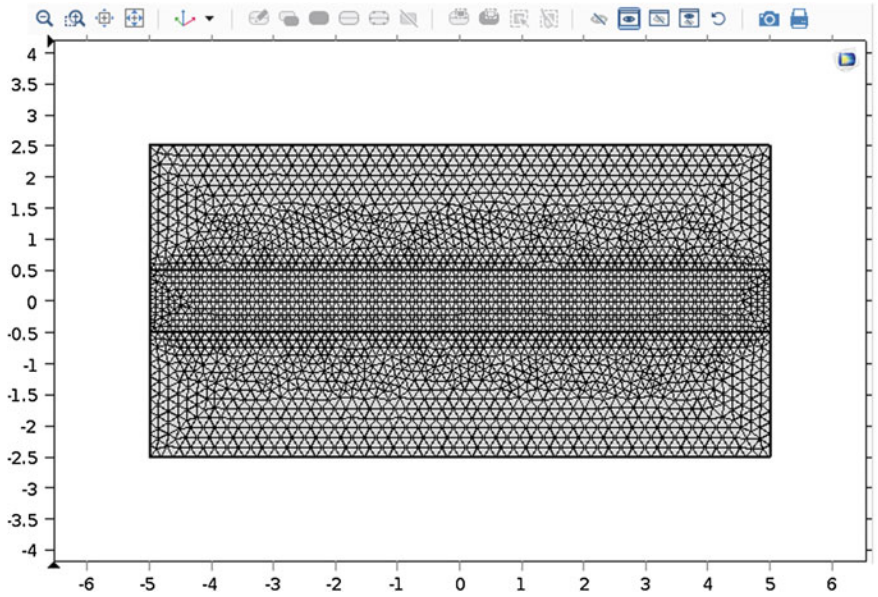


Fig. 3 The triangular mesh

4 Results and Discussion

The Schematic diagram of the planar waveguide is shown in Fig. 4. The electromagnetic wave propagating along the waveguide when the core RI in the evanescent field is 1.5 and cladding medium is air. The boundaries1, 3 and 5 are the input port they are used to transmit the light wave here the wave excitation is ON. The boundaries 8 and 10 are the output port they are used to receive the light wave here the wave excitation is OFF. The different boundaries for the planar waveguide are shown in Fig. 5.

The Electromagnetic wave passes through the microfluidic channel as below. The simulated result of electromagnetic wave propagating along the fluidic channel

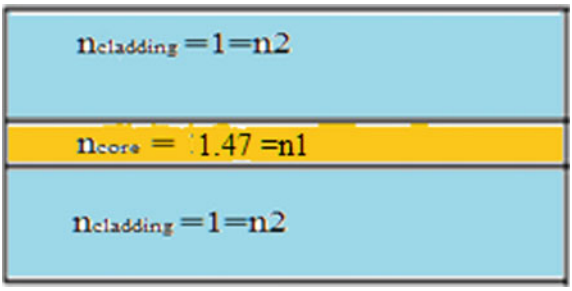


Fig. 4 Schematic of the planar waveguide

Fig. 5 Boundaries in the waveguide

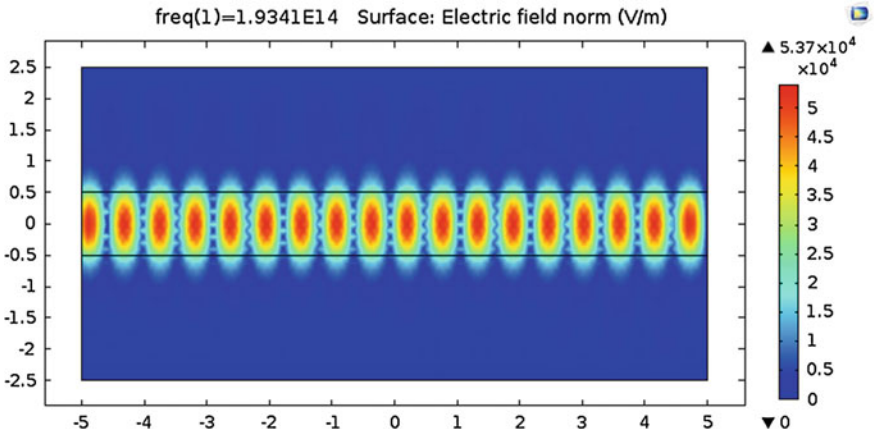


Fig. 6 The EM wave propagating in a fluidic channel

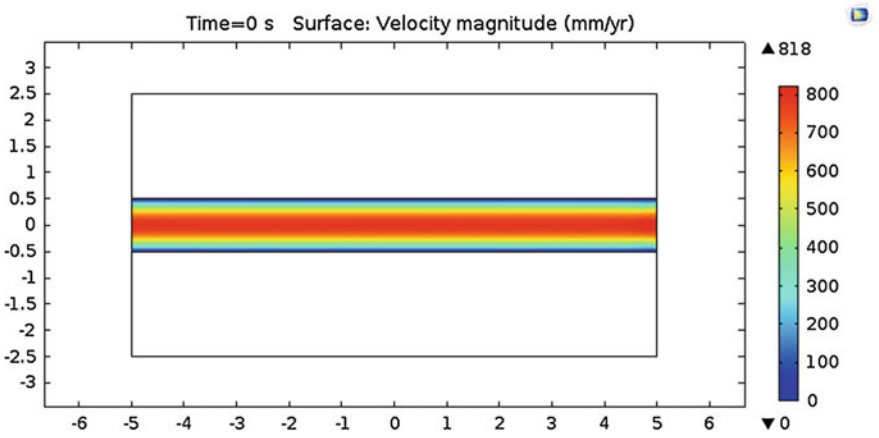


Fig. 7 The velocity of fluid flow in microfluidic channel

in the waveguide is shown in Fig. 6. In figure electromagnetic wave is not Prominent due to the presence of fluid which causes dispersion of light in the form of EM wave in the fluidic medium. The fluid flow is laminar which results in uniform fluid flow in the microfluidic channel.

Figure 7 shows the velocity magnitude of the fluid (water) contained in the microfluidic channel. As evident from the figure we can see that the flow is almost laminar. The Fig. 8 shows the pressure of the fluid in the waveguide. As seen from Fig. 8 we can see that pressure is higher at inlet compared to outlet.

The variation of reciprocal of step size versus time steps of the fluid flow is shown in Fig. 9.

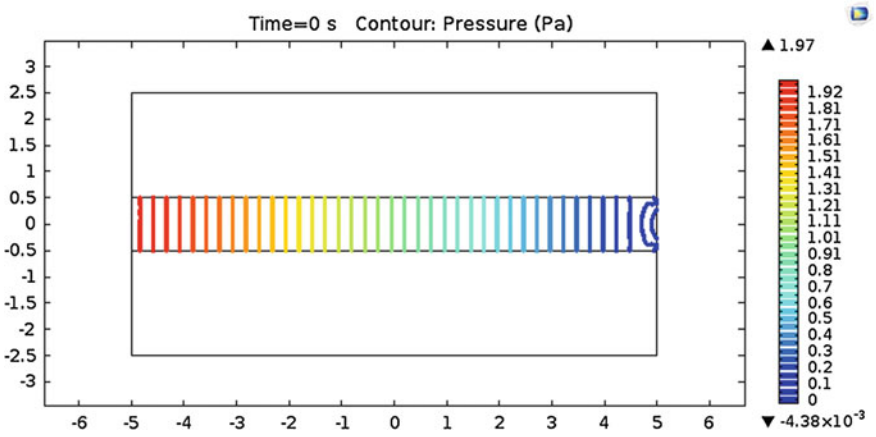


Fig. 8 Contour pressure in the waveguide

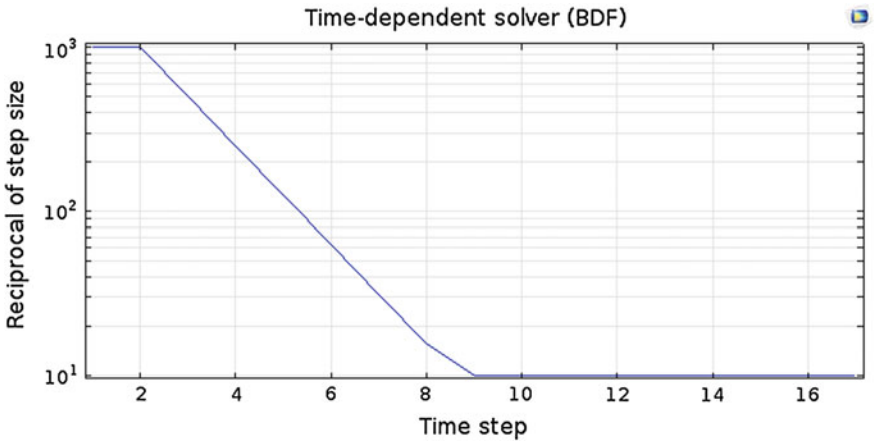


Fig. 9 variation of reciprocal of step size versus time step

5 Conclusion

In this paper we have analyzed and simulated a planar waveguide along with a microfluidic channel in which fluid flows. We have demonstrated how absorbance of the electromagnetic wave happens with in the microfluidic channel. By incorporating the characteristics of the wave optics and microfluidic modules of COMSOL Multiphysics it is shown here that planar waveguide can be used as the Evanescent field sensor. Velocity of the fluid flow is uniform in the planar waveguide. The pressure is very high at the inlet side when compared to the outlet, and as the fluid moves in the guided area the pressure decreases. In this work we have used a fluid flow which is uniform in the planar waveguide. This work can be expanded for bio-medical application.

Acknowledgments We would like to thank VTU Belagavi for promoting research culture and thanks to Dr T. Srinivas, Associate Professor, Department of ECE, IISc Bengaluru for constant encouragement. The special thanks to SVIT management for their inspiration. We would also like to thank Dr Ramesh Babu, Principle SVIT for motivating us to go in towards research. Last but not least we would like to thank Professor Y. Jayasimha for constant guidance and support.

References

1. Jozef, C., Koza, E et al.: Design and investigation of SiO2 slab waveguide grating sensing application, IEEE, ISBN 978-1-4799-3715-8/14 (2014)
2. Shao, H., Kumar, D., Lear, K.L.: Single cell detection capability of an optofluidic spectroscopic biosesor, IEEE, pp. 183–186, ISBN- 0-7803-9056-3/05 (2005)
3. Yin, D., et al.: Single planar optofluidic chip for single particle detection manipulation, and analysis. Lab Chip **7**, 1171–1175 (2007)
4. Veldhuis, G.J., Parriaux, O., Hoekstra, H.J.W.M., Lambeck, P.V.: Sensitivity enhancement in evanescent optical waveguide sensors. J. Lightwave Technol. **18**(5), 677–682 (2000)
5. Helmers, H., Bettech, P., Rimet, R.: Integrated optical components employing slab waveguides for sensor applications. IEEE Photon. Technol. Lett. **8**(1), pp. 81–83 January (1996)
6. Wang, R et al.: A microfluidic-colorimetric sensor for continuous monitoring of reactive environmental chemicals. IEEE Sens. J. **12**(5) May (2012)
7. Marcuse, D.: Theory of Dielectric Waveguides. Academic, New York (1974)
8. Kogelnik, H.: Theory of dielectric waveguides in Integrated Optics, Springer-Verlag (1975)
9. Raghuwanshi, S.K., Kumar, V., Pandey, R.R.: Derivation of eigen value equation by using equivalent transmission line method for the case of symmetric/asymmetric planar slab waveguide structure. J. Int. Acad. Phys. Sci. **15**, 1–14 (2011)
10. Raghuwanshi, S.K., Kumar, S.: Derivation of dispersion equation by using the equivalent transmission line method for the case of planar slab optical waveguide structure. In: International Conference on Optical Engineering (ICOE), IEEE 978-1-4673-2463-2/12/ 2012 (2012)
11. Introduction to Fiber Optics, Ghatak, A., Thyagarajan, K. (eds.)
12. Optical Fiber Communications, John M. Senior (eds.) 2nd edn.
13. Atalla, A., Nagarkar, K., Ashe, J.: Modeling the Magnetic Disturbance of Pulsatile Blood Flow in a Static Magnetic Field, IEEE, pp. 1406–1409, 978-1-4244-7929-0/14/ (2014)

Game Based Learning to Reduce Carbon Foot Print

U.V. Archana, Abhishek Rajeshwaran, M.L. McLain, Kamal Bijlani, Bhavani Rao and R. Jayakrishnan

Abstract Technology in the classroom is changing the way educators teach. When a serious game is combined with a pedagogical learning strategy the probability of learning and synthesizing new concepts can become much greater. In this work we take the concept of learning about one's carbon footprint and combine it with the step by step process of Bloom's Taxonomy, a proven and well-established learning method, and incorporate it into a serious game. The objective is to demonstrate that learning in this way will result in higher learning motivation and a better understanding of the subject matter. In the game the player needs to reduce their carbon footprint in an office building. Unlike in a traditional classroom where the student is only a receiver of knowledge, in this serious game the student has a more direct experience of the subject, eliciting good decision making and improving subject comprehension.

Keywords Serious game · Blooms taxonomy · Carbon footprint · Green house gases · Sustainable practices

1 Introduction

Bloom's taxonomy is largely applied in course design. The Bloom's taxonomy criteria are knowledge, comprehension, application, analysis, synthesis and evaluation. Here every task of this game is implemented by considering the five levels of

U.V. Archana (✉)
Amrita E-Learning Research Lab, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, India

A. Rajeshwaran · K. Bijlani · R. Jayakrishnan
Amrita E-Learning Research Lab, Amritapuri, Kollam, India

M.L. McLain · B. Rao
AMMACHI Labs, Amritapuri, Kollam, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_9

Bloom's taxonomy. For implementing each criteria different scenario is considered within the game.

Due to globalization and rapid industrialization, commercial establishment is utilizing a lot of energy whose source is mostly fossil fuels. As a result large amount of greenhouse gases is emitted into the atmosphere. The emission of these gases leads to global warming and overall environmental degradation. This game focuses on the reduction of usage of electricity and greenhouse gas emission within a building and its surrounding. The game play teaches the use of energy efficient technologies and sustainable practices will reduce carbon footprint. The technologies such as photovoltaic cell, biogas plant, various kinds of sensors, renewable energy sources, furniture made up with bamboo, etc. encompassing the overall concept of "Reduce, Recycle and Re-use."

Bloom's taxonomy's levels are incorporated here because they can provide a step by step manner of content delivery. Effective learning is always happening in a step by step manner. It starts from the basic level and reaches the higher level. Every task is linked with another task. This pedagogical aspect plays a major role in the dissemination of the concept of sustainability in this game. The task of the game starts with the knowledge level of the Bloom's taxonomy.

Knowledge is the lowest level of Bloom's taxonomy. This level is incorporated in this game by providing a chance—a pretest to the player to memorize what he has already learned. *Comprehension*—the translation of knowledge from one format to another is built-into the game with the replacement of existing technology with the new one. *Application* is the response of the learner to a particular problem, which is included as choosing an option of new technology. *Analysis* is the level in which analysis of each action taken place happens. The player analyses the technology, chosen one by one, on the basis of carbon footprint and energy consumption. As *Synthesis* is building an application on the basis of what he learned, here the player is given a situation to assemble the technologies on the basis of carbon footprint and energy consumption. Since the player can judge himself by the post-test conducted after the game play, the *Evaluation* level of Bloom's taxonomy is directly not included in this game.

Serious game is a tool which provides a virtual environment for learning complicated real world problems through gaming. Here students learn serious game concepts like sustainable building while playing. Through this game students will learn about several environmentally sustainable and energy efficient technologies.

According to [1] Game design, the game should provide an outcome along with the process. In this game the player will have the options to choose various technologies which can reduce the carbon footprint and bring overall efficiency in the entire system and implementing sustainable energy practices like renewable energy sources inside and in the vicinity of the establishment. The efficiency of various technologies and methods can be measured in terms of carbon footprint. Carbon footprint is the total amount of greenhouse gases produced to directly and indirectly support human activities, usually expressed in equivalent tons of carbon dioxide (CO_2) [2].

The various technologies are assigned to various scores, in which the technology has the least carbon footprint and less energy consumption has the highest score. On choosing the technology, there is a pop up box which shows the corresponding technology's energy consumption and carbon footprint.

2 Related Work

Serious game has a great scope in the field of sustainable building. Korina katsaliaki and Navonil Mustafee conducted a survey [3] to show that serious game is a superior tool to teach sustainability compared to traditional teaching methods. This paper focuses on learning and teaching sustainable development in an educational context. Serious games can be considered as one of the training methods to increase awareness. Here the learners can experience situations that are impossible in the real world for safety, cost, time. Based on the survey of Sustainable game evaluations they founded that the objective is easily achieved for various age groups and also serious games improve the leaner's analytical skill, spatial skill, strategic skill, insight, psychomotor skill, decision making, self monitoring, problem recognition. Thus the serious game can be a best teaching tool for sustainability among students.

The rate of carbon in the atmosphere is increasing. There are companies that are making efforts to reduce their carbon footprint; one example is the telecom industry [4]. With the help of the Alliance of Telecom Industry Solutions (ATIS) they bring together the subject matter experts from various service providers and equipment manufacturers and put forward some energy efficiency standards where they conducted a study on the basis of the carbon footprint from various industries in various countries. They are using a Greenhouse Gas Emissions Inventory to determine the carbon footprint of various companies. They are suggesting methods such as waste recycling, using energy star equipments, and renewable energy source used to reduce the carbon footprint. These methods are used in my game to reduce the carbon footprint of plastic furniture and electronic equipments. Research shows that gamification provides better learning than traditional learning [5].

The discrepancy in carbon storage and carbon boundary shows a significant difference in global warming on considering the comparative study of different methods EPD, PAS 2050, and ISO LCA [6]. The usage of wooden products does not contribute much to the carbon footprint. In the proposed game wooden furniture is used as one of the methods to reduce the carbon footprint of furniture.

Bloom's taxonomy criteria are used as a course designing tool. For making learning simple the entire course topics are divided into six on the basis of the levels in Bloom's taxonomy [7]. For each level they made the learning of each game simple. The knowledge level game is provided for the beginners, as the level of player goes on then the game level provided to the player is also increases. This clearly shows the learning is happening in step by step manner.

3 Solution Approach

On the basis of the criteria in Bloom's Taxonomy the game is designed with the Unity game engine. The player gets motivated with scores. It can be intrinsic or extrinsic motivation. Various criteria of Bloom's taxonomy is included into this game. Each movement in this game is designed according to the levels of blooms taxonomy.

3.1 Game Flow

- At first the player is provided with instructions, play and quit buttons. He can choose instructions or he can play the game or he can quit. On clicking the play button there is a pretest incorporated in the form of a quiz section. The pretest, which is comprised of three questions, checks the prior knowledge of the player. Here he recalls his memory and selects an option for every question.
- Before the game play starts, the player gets to know about the pre-test score. Every correct answer is awarded with a score of fifty. It should be noted that, this score is not taken into account with the total score of the player.
- The game has three levels. Interior of school and office building is the environment taken for first level. Interior of cafeteria is the environment for second level. Vehicles and office premises are the environment for third level. These environments are chosen for different levels because these are the areas that contribute carbon foot print into the atmosphere from our day to day life.
- On choosing level 1, the interior of the building with plastic furniture will open with the incandescent lamps burning and centralized air conditioners running. Desktop computers switched on with and without users are also brought into picture.
- To make the learner aware of carbon footprint and energy consumption the game play starts with the element Lamps and proceeds further with Centralized AC, Desktop computers and Plastic furniture respectively. On moving the mouse pointer to each of these, its carbon footprint and energy consumption will appear as a pop up.
- The game play is replacement of existing elements with the newer one. To support the player, newer technologies and methods having less carbon footprint are listed. The player gets maximum score by choosing the technology or method with least carbon footprint. The other listed technologies or methods are also awarded with score corresponding to its carbon footprint.
- As the game play starts with the incandescent lamps burning, the listed technologies and methods for replacing it are natural lighting, light emitting diode, chloro fluoro light and halogen bulbs. On clicking any of them the carbon footprint and energy consumption of the particular technology or method will

appear as a pop up. After replacing, the score of that technology chosen will appear in the next screen.

- Figure 1 represents the replacement of incandescent lamps with new technologies. Figure 2 represents the score of the player attained after replacing incandescent lamp with LED lamp. Here the score for natural lighting, light emitting diode (LED), chloro fluoro lamps (CFL), and halogen bulbs are 50, 40, 30, 20 respectively. The game play of other elements also proceeds in a similar fashion.
- The technologies and methods to replace the centralized air conditioning are fans, natural ventilation, air conditioners and usage of sensors. Figure 3 represents the replacement of centralized air conditioners. Figure 4 shows the score after replacing the centralized air conditioner with fan.
- Likewise, to replace the system related gadgets the technologies listed are use of laptop computers; switch off when not in use and reduction in the number of

Fig. 1 Replacement of incandescent lamp with led

Fig. 2 Score of the player after replacing the first element

Fig. 3 Replacement of centralized airconditioner with decentralized airconditioner

Fig. 4 Score of the player after replacing the second element

gadgets. Figure 5 represents the replacement of desktop computers. Figure 6 shows the score after replacing desktop computers with laptops.

- To replace the final element, Plastic furniture, the technologies and methods listed are recycling plastic, wooden furniture and bamboo furniture. Figure 7 represents the replacement of plastic furniture. Figure 8 shows the score after replacing plastic furniture with recycling plastic.
- After the replacement of four elements, total score of the player will be displayed. Figure 9 shows the total score after the game play.
- A minimum threshold is defined for the player to move to *Synthesis* phase of the game play, failing to which the player needs to repeat the process of replacement
- On attaining a total score beyond the minimum threshold, the player needs to reflect the knowledge acquired (For example, suggest the best carbon free technologies for the interior design of an office).
- The player advances to post-test after the *Synthesis phase*. The post-test comprising of objective questions, provides the player with self assessment. The first

Fig. 5 Replacement of desktop computers with laptops

Fig. 6 Score of the player after replacing the third element

Fig. 7 Replacement of plastic furniture with recycling of plastic

level ends with the post-test. The game play progress with level 2 and 3 by choosing the replacement strategy for cafeteria and vehicles and office premises respectively. The flow of the game can be easily understood from the Fig. 10.

Fig. 8 Score of the player after replacing the fourth element

Fig. 9 Total score

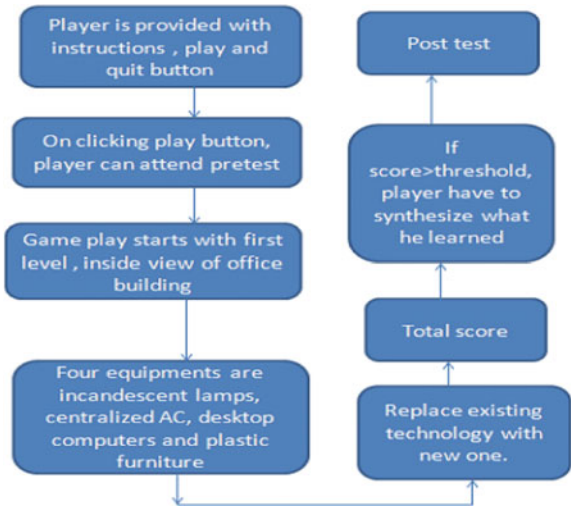


Fig. 10 Flow diagram

As mentioned earlier, the five of the six levels of Bloom's Taxonomy is used to design the game play as follows. The pre-test, game play, process of replacement strategy, comparing the technologies after the process of replacement and suggesting the learning outcome to a given situation corresponds to the Knowledge, Comprehension, Application, Analysis and Synthesis levels of Bloom's taxonomy respectively.

4 Experimental Results

This section includes the experimental results of comparing traditional and game based leaner. The proposed solution to teach the concepts effectively using GBL, the topic carbon foot print reducing technologies and methods have been chosen. The game provides information about various technologies which reduce carbon footprint and energy consumption. It can improve learner motivation in various approaches. Rewards and levels are to motivate the learners. Usage of appropriate challenge can avoid boredom. Motivational factors in a single player game and a multiplayer game are different. The elements in the game are interconnected. The scores are the rewards, in the proposed game. Whenever the player chose a technology or method he gets score and scores maximum when the chosen technology is the most appropriate method to reduce carbon footprint.

For the experimental setup, high school students of rural area were grouped into two groups of eight students each. One group was subjected to traditional teaching sessions and the other to GBL. On completion of these sessions, the two groups were given an objective test and was evaluated. From the analysis of the scores obtained, it has been found that scores of the students who learned through the game is higher than the scores of students who learned in the traditional classroom. Figure 11 shows the analysis of scores.

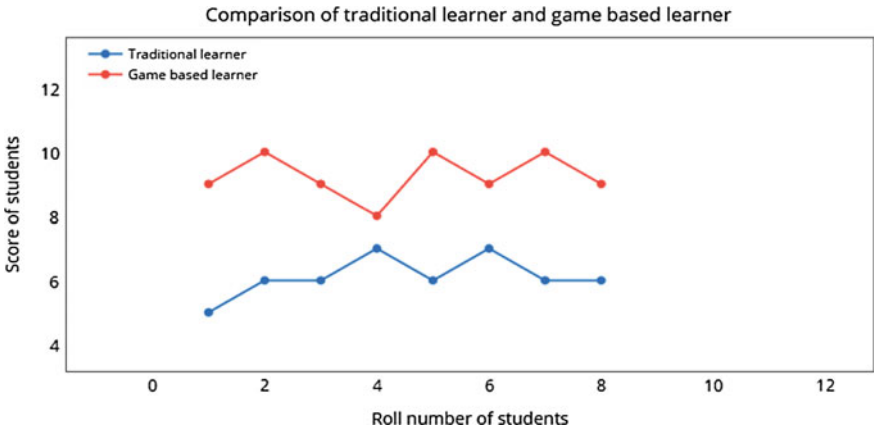


Fig. 11 Analysis of the student's score

Bloom's taxonomy's various levels affect the cognitive domain of the player through recalling in the quiz section of the game, psychomotor domain through applying the appropriate technology from the list of options provided and affective domain through scores awarded for each technology. It can guide the player from the very basic learning skill, memorizing to synthesize the best technology to the situation provided.

5 Conclusion and Future Extension

This pedagogical serious game is a great technology for learning the methods and technologies which reduce the carbon footprint in classrooms and office rooms. Reduction of carbon footprint with the technologies and methods can be demonstrated effectively and efficiently in this game using a learning method such as Bloom's Taxonomy. Knowledge, Comprehension, Application, Analysis levels provide the awareness of carbon footprint of various electrical and electronic devices and the Synthesis level nominates the player to synthesize in a given situation. Serious games help the learner to understand the underlying concept in a profound way. Serious games are an emerging tool of education because the concept is comprehended by the player at a deep level without the risk of boredom. The decision making ability of the player is improved as they replace of each devices which have a large carbon footprint with a newer one that saves more energy. When serious games are blended with pedagogy the motivation and decision making ability of the player improves.

Future work will include further implementation of methods and technologies in level 2 and level 3 which reduce individual carbon footprint through their lifestyle and food style.

Acknowledgments We are grateful for the support and facilities provided by the Amrita Vishwa Vidyapeetham University and Amrita E learning Research lab.

References

1. Honeycutt, B.: Applying bloom's taxonomy to game design. In: Complimentary Article
2. http://timeforchange.org/what-is-a-carbon-footprint-definition
3. Katsaliaki, K, Mustafee, N.: A survey of serious game on sustainable development. In: Winter Simulation Conference (2012)
4. Elkenaney, M., Kujan, B.: Where we are with energy efficiency in telecom equipment in north America INTELEC. Telecommunications Energy Conference (2008)
5. Bohnke, P., Wierzbicki, R.J.: Cognitively engaging learning through gameful experience. International Workshop on Online Teaching Initiatives (2012)
6. Garcia, R., Freire, F.: Comparative assessment of environmental life-cycle-based tools: an application to particleboard. In: Sustainable Systems and Technology (ISSST), IEEE International Symposium (2011)
7. Buchanan, L., Wolanczyk, F., Zinghini, F.: Blending bloom's taxonomy and serious game design

Analysis of Black Hole Attack in Ad Hoc Network Using AODV and AOMDV Protocols

S.H. Jayachandra, R. Manjunatha, Nayab Hussain, B.U. Sujan, H.L. Gururaj and B. Ramesh

Abstract Security is an essential factor in wireless ad hoc network to have safety in transmitting data packets between two wireless sensor nodes. The nodes posses a unique characteristics and it leads to consequential challenges to security design. Comparing to other wireless networks WSN has more security problem; this may be due to its nature of broadcasting messages, resources and their environment. One of the traditional and main attacks of WSNs is Black Hole Attack.

Keywords Black hole · AODV · AOMDV

1 Introduction

Black Hole Attack: Black Hole is an active and routing attack method where attacker node promotes itself as a best node path to reach the destination and all other nodes. In this attack, the attacker node waits until neighboring nodes initiate the RREQ packet. When the attacker node gets the request it sends a fake reply packet RREP with a new sequence number. The source thinks it is the active and best node to reach the destination. So; it ignores the other nodes and sends all its data packets through attacker node. The malicious node accepts all the incoming data packets and drops it. It does not forward it to other nodes. As all the data packets are concentrated at a single node it is called as 'Black hole' and the region is called as 'Black region' [1–3].

S.H. Jayachandra (✉) · R. Manjunatha · N. Hussain · B.U. Sujan · H.L.Gururaj · B. Ramesh
Department of Computer Science and Engineering, Malnad College of Engineering, Hassan, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_10

2 Ad Hoc Routing Protocols

Ad hoc network consists of dynamic nodes with router functions. The important sector of ad hoc network is the routing protocols, because network topologies keep on changing according to the movement of active nodes.

We have two types of routing protocols in ad hoc network, they are Table-driven type (proactive) and On demand type (reactive). The two main routing protocols we are focusing in this paper are AODV and AOMDV which comes under the on demand type of routing. On demand protocol is a routing protocol which creates a path between nodes only during requirements. It minimizes the overhead on ad hoc network.

2.1 AODV

AODV is one of the on demand and typical routing protocol, to avoid the breakings in the routing protocol we go for this algorithm. It has symmetrical path between the nodes. AODV has routing table and sequence number for the nodes. The sequence numbers are assigned by the destination node to obtain the freshness of routing information [1, 2].

AODV uses Client-server method that is Request-reply method for finding a valid path between sources to destination. The source node broadcast RREQ (route request), the neighboring nodes receive a route request and routing table gets updated. Finally the RREQ reaches destination node, after receiving the RREQ destination node creates a RREP (route reply) and unicasts a route reply. The communication starts between the nodes when the RREP reaches the source. If the links get fail the respective nodes create a message called RERR (route error).

2.2 AOMDV

Ad hoc On-demand Multi path Distance Vector Routing (AOMDV) protocol is considered as the continued part of the AODV protocol. AOMDV is used to find number of loop free paths and also link disjoint paths in any network. AOMDV protocol generates a large number of packets than the AODV protocol. Upon comparing the AOMDV with AODV in terms of number of packets dropped, AOMDV shows a large number of packets that are dropped, since it does not have the larger receiving capacity as in AODV protocol. The main benefit of using AOMDV protocol is that it supports intermediate nodes to respond to RREQ's [4, 5].

3 Network Simulator 2

NS2 is a event simulator mainly intended to network research. NS is an interpreter of TCL scripts of the users; they work along with C++ codes. It uses Object Oriented Tool Command Language to evaluate user simulated scripts. TCL and C++ are totally compatible with each other.

4 Performance Metric

4.1 Packet-to-Delivery Ratio (PDR)

The ratio of the packets delivered to destination when compared to that of the packets sent from the source.

The below equation gives the formula to evaluate PDR:

$$\text{PDR} = (\text{Number of packets receive}/\text{Number of packets send})$$

4.2 Average Throughput

Throughput is the total number of packets sent successfully from sender to receiver in a specified time.

$$\text{Throughput} = (\text{PacketSize} * \text{ReceivedPackets} * 8)/100$$

5 Simulation and Performance Analysis

5.1 Black Hole Attack Under AODV Protocol

The Throughput and the packet to delivery ratio of the AODV protocol using black hole attack can be analyzed by introducing an attacker on a particular node.

Whenever an attacker claims for a specific node, there is a possibility, that several parameters like throughput, performance, packet to delivery ratio etc. can vary accordingly. The change in throughput can be shown in the below Fig. 1.

As we can see in the above graph, the average throughput in the absence of the attacker is very high, such that the majority of the packets sent from the source will reach the intended destination without any packet loss. Hence high is the throughput when there is no attacker in a networking area during the transmission of data.

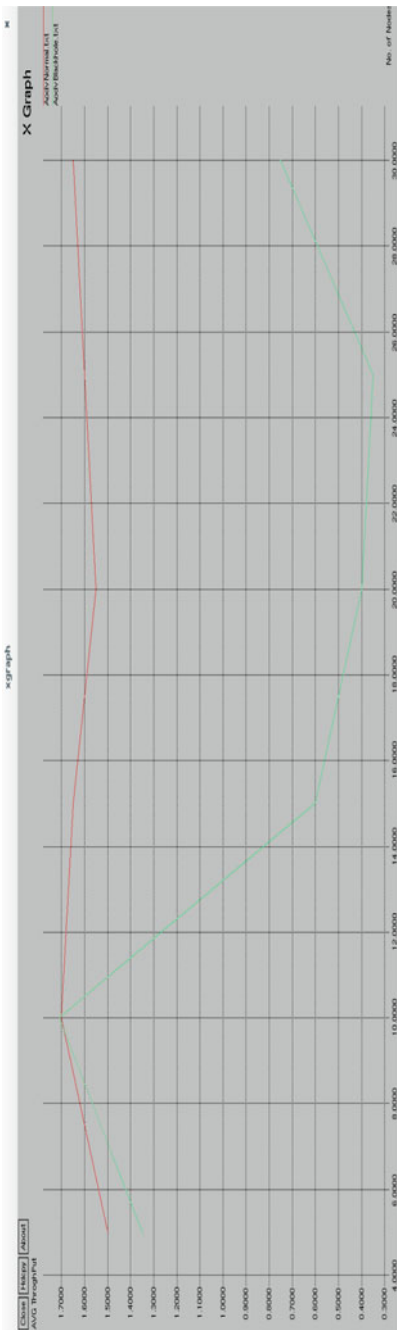


Fig. 1 Analysis of throughput under AODV

Similarly, the green line which indicates the throughput in the presence of attacker, shows that when a node is attacked the throughput gradually decreases. Hence we can see the difference in the throughput because of attacker. In the sense the packet sent by the sender will not be successful in reaching the destination. Hence there is decrease in the throughput [6, 7].

The Packet to Delivery ratio can also be analyzed using the similar method by introducing an attacker to any of the nodes.

Here in case of normal transmission of packets in the absence of attacker, the packet to delivery ratio is very high and remains constant throughout the transmission. Such that the ratio of number of packets sent to number of packets received will be very high. All the packets are assumed to be received at the destination.

While in the presence of an attacker, we can clearly notice that the packet to delivery ratio has come down to very low values. The attacker captures the sent packets from the source leading to low delivery of packets to the destination. Hence the Packet to Delivery ratio is very low if the attacker is present in the network. We can observe the variation of PDR in the above Fig. 2.

5.2 Black Hole Attack Under AOMDV Protocol

To evaluate Throughput we have taken some nodes, it is evaluated by plotting the graphs, and numbers of malicious nodes are taken in the x-axis and throughput in the y-axis.

We can observe from the Fig. 3 that as the number of malicious nodes increases there is a drastic low in the throughput value. In the normal transmission of data the data transmitted will be more compare to value with the attacker. Because in the presence of attacker the intended data may not reach the destination. The attacker consume the bandwidth and gives the wrong routing information and sender sends the data packets to wrong destinations which leads to lower in the throughput value. To evaluate the Packet to delivery ratio we have taken Number of malicious nodes in the x-axis and PDR in the Y-axis.

Higher the PDR value, the performance is best. Here in our x-graph Fig. 4, in the presence of Black hole attacker PDR value is very low, sender is failing to send the complete data to destination. In the absence of attacker the data packets are sent without any miss guide in routing information to the destination, as a result we have higher PDR value. We can conclude that in the presence of attacker (Black hole) the Throughput and PDR value are very low, and higher value in the absence of attacker [8, 5, 9].

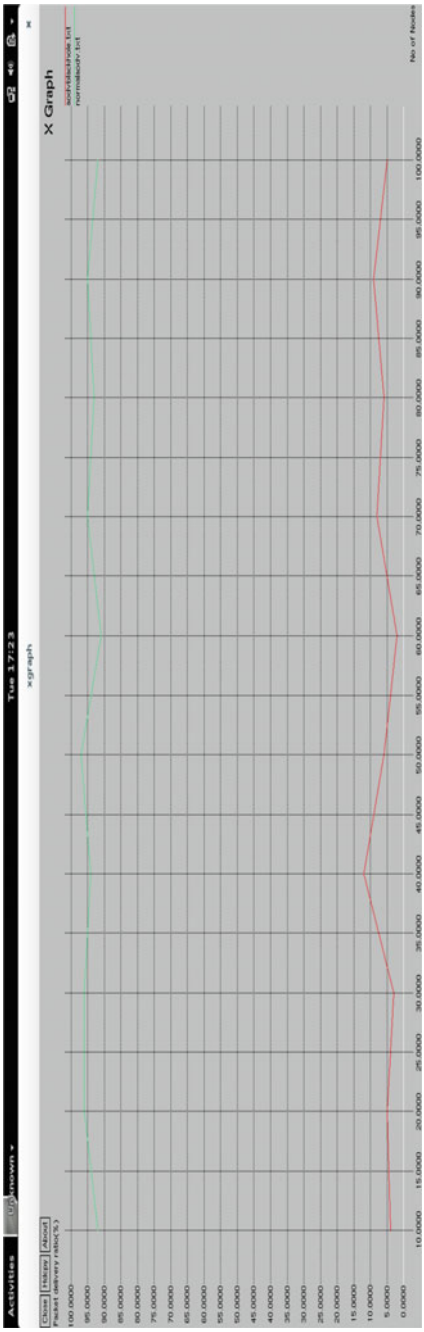


Fig. 2 Analysis of PDR under AODV protocol

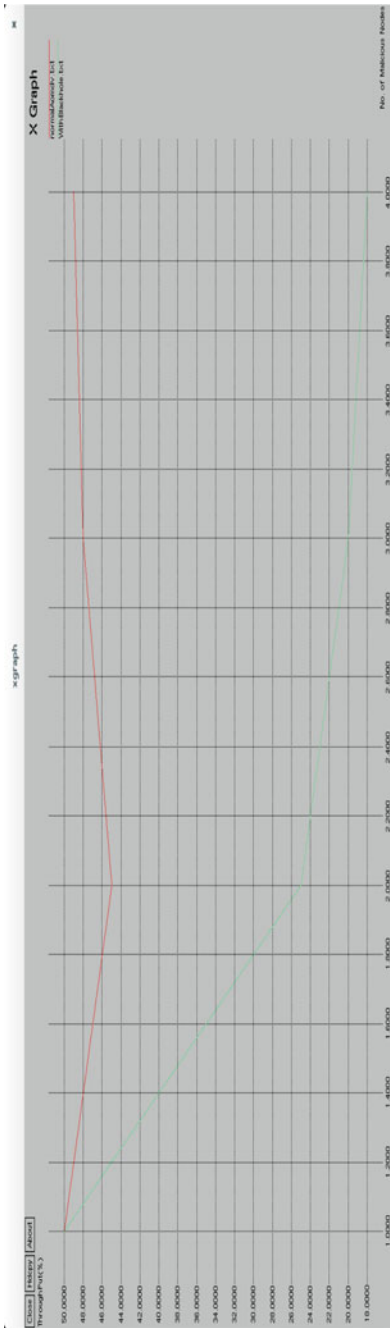


Fig. 3 Analysis of throughput under AOMDV

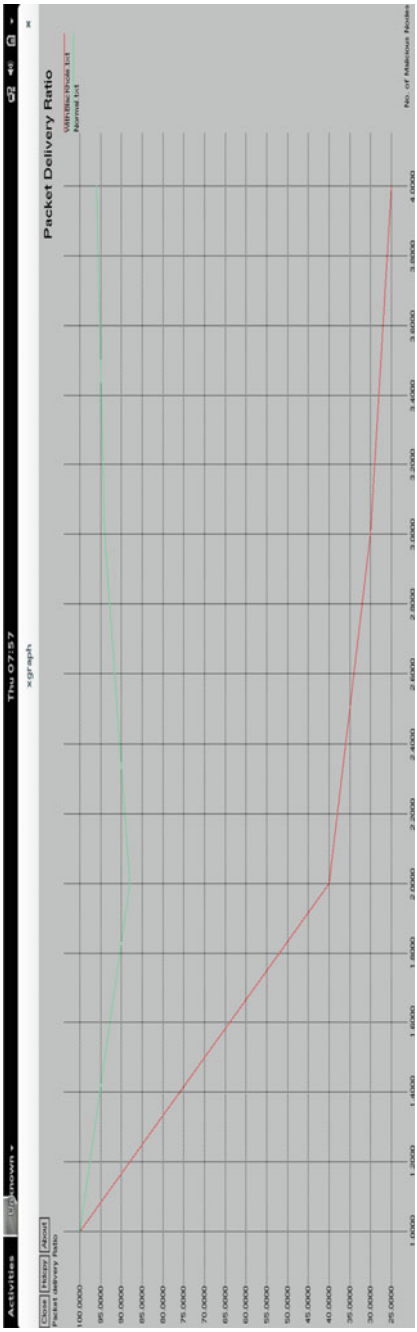


Fig. 4 Analysis of PDR under AOMDV protocol

5.3 Comparative Analysis of PDR in AODV and AOMDV

In AODV protocol, almost all the packets will reach the destination when there is no attack. When we consider the attack it almost none of the packets will reach destination. When a link fails, a routing error is passed back to a transmitting node, and the process repeats.

Where as in AOMDV multiple paths are available between the source and destination. When we consider the attack, AOMDV protocol has multiple paths, if one node fails, the packets will try to reach destination from other paths, we can see the decreasing line this shows it will try to find another path to reach the destination from all the possible paths. Due to attack it will not reach the destination.

5.4 Comparative Analysis of Throughput in AODV and AOMDV

We consider higher the throughput higher the performance by using AODV protocol it will have better throughput means packets will be sent fast and overhead will be avoided. When we consider attack the average throughput will decrease thus the performance also decreases. The No. of packets travelling from one node to another decreases this will affect the entire performance.

When we consider AOMDV throughput is better than AODV protocol because however RREQ is transmitted through all the nodes to find destination. Due to AOMDV Protocol it makes entry in routing protocol where as in AODV protocol it is discarded means duplicates packets are discarded. Hence we have better throughput in AOMDV than in AODV protocol [10, 11].

6 Conclusion

It is quite clear from the above graphs and propositions when in case of Black Hole Attack that AOMDV protocol is better than AODV. But at the same time, overhead is quite less in AODV when compared to that of AOMDV. It is because, in case of RREQ when a destination is set in AODV it takes only that destination and discards all other nodes. But in AOMDV, if RREQ fails for a certain path, then the inputs in the route request table is maintained throughout. Thus in case of Black Hole Attack, if we want to avoid higher overhead, we intend to choose AODV protocol. And if we require higher throughput and PDR, we intend to choose AOMDV protocol.

References

1. Bhosle, A.A., Thosar, T.P., Mehatre, S.: Black-hole and wormhole attack in routing protocol AODV in MANET. IJCSEA **2**(1), (February 2012)
2. Khin, E.E., Phyu, T.: Impact of black hole attack on AODV routing protocol. IJITMC **2**(2), (May 2014)
3. Virmani, D., Soni, A., Batra, N.: Reliability analysis to overcome black hole attack in wireless sensor network. In: ICCIN-2K14| January 03–04, 2014
4. Naseera, K.M., Chandrasekar, C.: Prevention of black hole attack using AOMDV. IJERA **3** (6), (Nov–Dec 2013)
5. Bhardwaj, N., Singh, R.: Detection and avoidance of blackhole attack in AOMDV protocol in MANETs. IJAIEM **3**(5), (May 2014)
6. Pooja, Jaglan, S., Gupta, R.: Int. J. Adv. Res. Comput. Sci. Softw. Eng. **4**(8), (August 2014). ISSN:2277 128X
7. Manikandan, T., Shitharth, S., Senthilkumar, C., Sebastinalbina, C., Kamaraj, N.: Removal of selective black hole attack in MANET by AODV protocol. IJIRSET **3**(3), (March 2014)
8. Vyas, T.G., Rana, D.J.: Survey on black hole detection and prevention in MANET. IJARCSSE **4**(9), (September 2014)
9. Jain, A.: Trust based routing mechanism against black hole attack using AOMDV-IDS system in MANET format. IJETAE **2**(4), (April 2013)
10. Mohebi, A., Kamal, E., Scott, S.: Simulation and analysis of AODV and DSR routing protocol under black hole attack. IJMECS (November 2013)
11. Verma, Monika, Barwar, N.C.: A comparative analysis of DSR and AODV protocols under black hole and gray hole attacks. IJCSIT **5**(6), 7228–7231 (2014)

An Approach to Overcome Startup Issues Using Serious Game

Aiswarya Gopal, Kamal Bijlani, Vinoth Rengaraj and R. Jayakrishnan

Abstract In the business world, the term "startup" is frequently ringing the bell with the high frequency of young ventures. The main dilemma of startups is the unsuccessful management of the unique risks that have to be confronted in the present world of competition and technology. This research work tried to bring out a game based methodology to enhance enough real-world experience among entrepreneurs as well as management students to handle risks and challenges in the field. The game will provide experience to the player to overcome challenges like market problems, running out of cash, poor management, and product problems which can be resolved by a proper strategic approach in the entrepreneurship world. The proposed serious game works on the life cycle of a new software enterprise where the entrepreneur moves from the planning stage to secured financial stage, laying down the basic business structure, and initiates the operations ensuring the increment in confidence level of the player.

Keywords Startup · Simulator · Business model · Serious game · Poor management · Risk management · Game based learning

1 Introduction

Based on the study conducted by Ministry of Labour and Employment, Government of India [1], the reason for poverty in our country is due to the increased number of unemployment. Even though there are a lot of talented and qualified young people in our country. Sadly, no entrepreneurs are trying to start a business on their interested area. One of the main problems [2] is the lack of experience in running a business and people are fearful of taking up the risks. These factors demotivate people from initiating new ventures. The real world experience

A. Gopal (✉) · K. Bijlani · V. Rengaraj · R. Jayakrishnan
Amrita E-Learning Research Lab, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, Kerala, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_11

and practical knowledge are two important factors for bringing up a successful business. Lack of experience of a person will affect procurement and utilization of funds, hiring and management of employees, competing effectively in the market, reaching the consumers properly, etc. So the entire business will be affected by the lack of experience of the entrepreneur.

Usually, to get an experience in a business, one should work in the industry for 4–5 years minimum. This will normally provide ample experience for a person to run a business on his own. If a person is not getting an experience like this, the possibility of losing his business is high. The lack of experience is one of the reasons that most people holding back to start a business on their own. But joining an organization and working for more than 5 years may not always provide an initiative to start a business; however this is a long period. For resolving this issue, this paper is introducing a method to teach how to start and manage a business successfully using the media of serious game.

The main objective of the game is to train an aspiring entrepreneur. The game helps to get an idea about how to start and manage a business successfully. This is helpful for persons with no real life business experience to manage a business on his own. It can also be used as a tool for teaching management students. The teacher can use the game as a tool for explaining in a traditional class and this can also be used in a flipped class. In this method students are given game and they are asked to play the game at their home up to a particular portion (Here up to the 8th scenario). After this, the teacher should clarify their doubts and engage them to play the last scenarios. These scenarios are like an experiment to test the students/players understanding level. The last scenario will give the reports based on the game play of the students or players with the help of this the teacher can evaluate the students' understanding level. After completing each scenario the student must bring up some questions based on the previous scenario. If the answer is correct, the player can go to the next scenario otherwise the player must complete the same scenario again. The last two scenarios are like an experiment.

The learning styles of people are not the same. Some people like visual learning, some others prefer auditory learning, some others favor kinesthetic learning style or tactile way of learning [3, 4]. This game provides visual effects like instructions in the popup box, picture of different elements in the game, the result of following each instruction, etc. Apart from these, there will be a verbal guidance and sound effects for each instruction and results with a musical background to give an audible effect for the game [5]. In this game we have to drag and drop according to the instructions, which make the game more interactive and interesting to the player. Throughout this game the player has to engage with hands on activities. Especially in the last scenario, the player has to play the game on his own with the experience he got from the previous scenarios.

The game based learning have many advantages compared to other learning methods. Game based learning [6] helps the learner to understand the concepts and remember the contents easily. Serious Games [6] provide fun, entertainment and also can be used for educational purposes, training, advertising, etc. In addition, Serious Games have more learning objectives. The business field is generally full of

tension and stress, business people have to face a lot of problems daily. The game can also be used as a self training tool. If a business person is getting an experience by playing Serious Game with real life situations of business, he will be able to overcome the problems by recollecting the solutions he learned from the game. The next thing is this game gives the visualization effects and experience on different problems and solutions. So it affects the cognitive mind [6] of the player and this helps him while he is facing similar situations in real business. Due to these factors the player can easily recollect similar situations in the game and he can correlate this in the real world. So the player can easily achieve the solutions of the problems. Through the game the player will first learn all the steps of starting a business, solutions for solving some problems etc. with the help of instructions. The player will engage with the game and learn things easily. After the ninth scenario the instructions will be removed level by level, following the scaffolding methods [7].

The game provides different scenarios, situations, problems and their solutions in a business environment. After completing the first eight scenarios the player will get an idea about the common challenges and their solutions in the business field. In the last scenario the player can start his own software startup. The challenges and problems will overcome with the help of previous scenarios. And this will give a clear idea about starting and running a business and the player will get a real life experience in the field of business. This game will also enhance the business ideas of a person. In this way this will invoke the business dreams of the player and will improve the number of self employed people in our country. So this will be helpful to the development of a person and the overall growth of the economy of the country.

In developing this game, the main focus is on some common challenges in the business world that are [8]; Market problems, Business Model failure [9], Poor Management, Running out of cash, Product problems, etc. Most of the Startups fail due to the important factors like poor time management, over usage of the resource, extravagance over the quoted budget etc. The source of these problems' is the lack of real life experience.

The proposed system is targeted to solve the above problems by creating a real life business situation through Serious Game. By playing this game the aspirant will get an experience and an idea about running a business. This game includes most of the scenarios of a business organization with its problems. As a result of playing this game the confidence level of the player will be increased.

2 Related Work

Mario Allegra, Valentina DalGrande and two others discussed [10] about the tourists company management game though a research work. The learning by doing and learning by failing methodologies for learning complex situations are mention in their work. The game follows the experience based teaching method where the student can learn the management of a tourist resort. This serious game is integrated

into a blended educational path [10] that combines both face to face and distant learning, but the importance of the evaluation part is less.

Stefan Greuter, Susanne Tepe and some other colleagues [11] created a game related to the construction workers and different problem they encounter on sites. Due to the increased number of injuries and accidents in construction sites, as a safety measure in this industry, this game helps. The Australian Building and construction commissioner made it compulsory for everybody to complete an occupational Health and safety, construction induction process to make themselves eligible to work [11] on a construction site. The main problems or challenges in this induction training are to identify hazards, the management of hazards through implementing controls to avoid accidents and to engage learners [11, 12] in the training programs. Through this work Stefan Greuter and Susanne Tepe, presents a serious game as a tool to engage learners, motivate them and help users to retain their knowledge of obstacles at the workplace and the way to manage them and to use their knowledge in the real world situations.

Study through the game [11] provided lots of motivational aspects like relaxation, fun, passing the time, challenge, social aspects, excitement and fantasy. Because of these aspects players become very immersed and engaged in active learning through games. Advantage of active learning [11] helps to achieve deep learning rather than surface learning.

More than entertainments serious games also have learning effects, which is proved [13] by Matt Critelli, David I. Schwartz and Steven Gold. The serious games in business which introduces three games, namely Enterprise, American's Army, and Grip is described that. Serious games are important, because of the improved and effective way of learning and training where the players will learn and work [13] at the same time. Thus, it will contribute to their profession.

The first game is "Enterprise-The Game" [13] it is an entrepreneurial game to learn more about the successful operation of a business. This game will develop the entrepreneurial competency of the players. The second one is "American's Army" [13] this game is considered as the most popular first person shooter game genre in the world and it is designed, developed, and distributed by US Army. The objective of this game is to recruit and train young adults to the army. By this game participants will be recruited and at the same time they will get trained also. The last one is [13] "GRIP Self management for children between the age of 7 and 11 years with diabetes and their parents" the objective of this game is to teach the patients and their parents how to handle and treat diabetes. By providing the information about the patient and their behaviour GRIP helps patients and their parents to find out the illness and take action to prevent the effects of the behaviours.

Lynda Donovan is a pedagogical lead who discussed about the serious game. From the paper [7] learner can understand each and every attribute and elements of the game and also describe the relation between the blooms taxonomy and the serious game. It describes all the basic factors in the game and also describes the learning theories. Also portraits some games related to training, recruitment, marketing and sales. It furnish with the best idea about business related serious games. The usage of serious game for training helps to increase employee engagement. The

paper [7] also describes the games for financial training, hospitality training, skill training etc. From the analysis of the results of these three games we can conclude that serious games are effective and give more results comparing to any other mode of learning/training. Through game we can improve the learning skill.

3 Solution Approach

Serious game is a game designed with a learning objective. Even though it provides a certain kind of entertainment [6], the main purpose is learning and not entertainment. This research work using this platform as pedagogy by including real life situations. The player will be taught by the instructions provided with the game and evaluation will be done at the end using score, manner of utilization of fund and completion time of the game, final report etc. At first, the player will be given complete instruction for achieving the objectives. In the last scenario, there will be less instructions and player has to complete the game on his own. At this stage the player has to overcome the obstacles he faces while playing the game with the knowledge he received at the early stages.

The Fig. 1 shows the learning methodology, that the first instruction to play the game. From the instructions, the player will learn the objectives. Then play each scenario based on the instructions and achieve the objectives. If the player achieves the objective, he/she can go to the next scenario otherwise the player can go back and again read the instructions. Finally, evaluate the player with the help of last scenarios. The last scenario is the actual game play section.

The last two scenarios mainly used for evaluation. This will give the report. This report will show the correct way and the path played by the player. In comparison of these two the player or teacher can evaluate the knowledge level. In the last scenario the player can play in his own way. At this stage the player need to face the challenges and the surprise elements handling etc. The previous scenarios will help the learner/player to solve the challenges and problems.

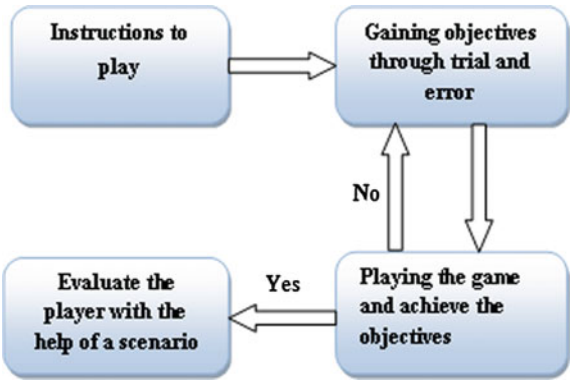


Fig. 1 Learning methodology

3.1 Game Description

The game includes a set of scenarios with certain procedures that player has to follow. Scenarios are mostly related to the common problems and its solutions in the business field those are arranged in a sequence. Each scenario will have one or more objectives and the player has to play according to the instruction to achieve the objectives. Timely instructions will be given as a popup. This will also help the user to get familiar with the game. The user has to follow these instructions to complete the game and achieve the objectives successfully. Checkpoints are provided in the form of score, fund or bank balance and time. Evaluation is done by verifying the scores, effective utilization of fund, final report and timely completion of the game. The success rate is decided on the basis of the speed of the player in learning the game. At the end, the player has to complete the basic scenarios. And the last scenario will be managing the whole stuff from pitching the idea to become a better entrepreneur.

The Fig. 2 show the game overview. This game has ten scenarios. The first eight scenarios are in the learning stage and the last two scenarios will come under the evaluation stage. The first three scenarios will help to understand the game objects and the game play. Next five scenarios will describe basic problems and constraints faced in the startups. In the next scenario, the player will be learning the game play in managerial or best entrepreneur mode. The last main scenario helps to become a successful entrepreneur.

The game gives the problem statement which explains different problems that will be faced in the real business world. This helps the player to understand the problems that he may face in the future in his own business. This is helpful to avoid the problems in advance. This will help the player to prepare mentally to face the problem and this will avoid the effects of unexpected shock. So this will reduce the risk of running a business. And the solutions to the problem help the player to face

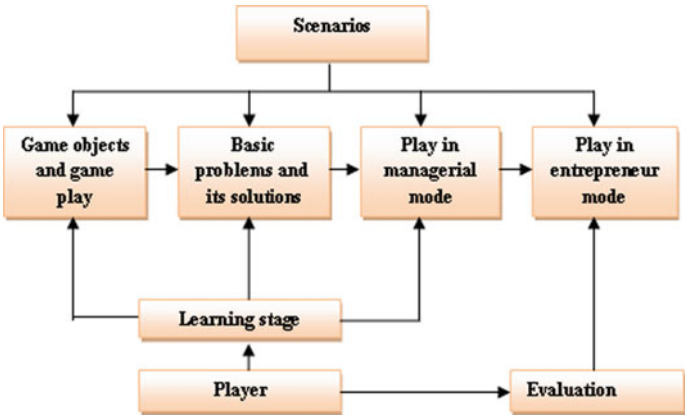


Fig. 2 Flow of game

unavoidable problems without fear. And this will improve his confidence to run a business. This mentality improves the positive energy in the player and in turn this will improve the overall performance of the business. The known solutions will help the player to play on a safer side. The given path will help the player to approach the solution. In the students perspective it will improve the theoretical knowledge and practical knowledge.

3.2 Game Design

The main game play will include in the last scenarios. With the help of learning stage scenarios the player can play the last scenarios easily.

The first three scenarios will give the idea about game objects and game play. That is, these scenarios show the player from where to select the objects to continue with the game. The player will select objects on his own and he will learn to proceed by trial and error mode, like every human being's natural way of learning things without specific teaching.

Next five learning scenarios explain basic problems and constraints faced in startups. With the help of these scenarios the player can play the last scenario in an effective way. An entrepreneur has to face many problems related to the startup. We are summarizing these problems and its solutions in five scenarios, the five scenarios are:

- Basic requirements, hiring, upgrading technology, bottleneck.
- Affiliation and tax problems.
- Customer satisfaction or decision management or dead line problems.
- Market problem or product problems.
- Cash flow.

From first scenario, the player will get good knowledge and experience in the following fields.

From basic requirements the player will learn what the relevant points in the business proposal are and how to make an acceptable business proposal etc. The player has to give his business ideas and budgeting here and if it is acceptable to the venture capitalist the player can collect the seed fund. If it is rejected, the player has to revise his business proposal. Through this way player will learn how to prepare a best business proposal and what the different methods for collecting fund etc. For example for preparing the business plan, the player must select seven or eight relevant points from a set of points. That set of points contains many relevant and irrelevant points. After playing this section, the player achieves a good idea about the relevant points in the business plan.

In hiring, the player has to give his priorities for selecting his team of employees. That means number of technical staff, number of non technical staff, skilled employees, semiskilled employees, unskilled employees, year of experience etc. For example at the time of team building the player will select more number of unskilled employees

and only less or zero number of supervisors. At that time the popup will come and give some instructions like the proportionality missing or your team is not good. Then give a chance for improving the selection. The Fig. 3 related to the proportionality of employees. For example sometimes the designer is more than as compared to the developer. In the case of Software Company we need more developer as compared to the designer; otherwise we cannot implement our project at correct time.

From updating technology, the player has to proceed with latest technology and facilities. For this he has to update his departments, buildings and proper training should be given to the staffs regarding latest technologies. The Fig. 4 related to upgrading the employees. If the employee is low rated we can give the training for them with the help of this training they can improve their work. The game will be provided with choice of technology to be selected. If the player selects the appropriate technology, he can proceed with the game.

Bottleneck is a problem of lack of sufficient resources to produce the required number of output. In an organization certain resources will be enough and some other resources will not be enough to fulfil the requirement. This will affect the smooth production and marketing adversely. So the player has to arrange his available resources in a proportion required for fulfilling the market requirements. Example for bottle neck is lack of employees with required skill, non availability of raw materials improper allocation of funds etc.

From the second scenario, the player will get the idea about Affiliation and Tax related problems. After completing this scenario, the player will solve such

Fig. 3 Maintain the proportionality of employees

Fig. 4 Upgrade the employees

problems in the business. Through the third scenario, the player can learn about customer satisfaction or deadline problems. From here the player learns to select successful plans to handle customers, take decisions for producing quality product, fulfil the customer requirements etc. The fourth scenario will give the idea about the problems and solutions related to market and the product. The marketing of a product involves various stages starting from understanding customer requirements, educating the customer about the product, forecast future requirements and fixing competing prices in the market etc. The fifth scenario will give the idea about the problems and solutions related to cash flow. Sufficient funding is necessary to invest in building construction; machineries etc. and adequate flow of working capital should be there for meeting day to day requirements. Cash flow includes finding out the sources for initial investments, working capital, credit policy, payment policies, etc. The player should select the cheapest and safest mode for ensuring smooth cash flow in the firm.

The last two scenarios will help to evaluate the player, i.e. in the second last scenario is a consolidation of all the above said scenarios. The player has to face all the problems one after another and the player will handle these problems with the help of the experience he got from the previous scenarios. If the player will solve this problem and challenges correctly, he/she will get high score otherwise player will get a chance to correct it and try again. Through trial and error method the player will learn the correct path for solving the problems. This will make the player able to take decisions in all the critical stages that can arise while running a

business. The player will become an experienced managerial person who can take up all the challenges in a business.

The main aim of last scenario is to evaluate the player, i.e. how much the player learned from the scenarios, how they face the challenges and how they overcome the problems etc. In this scenario the player has to play the game himself. He will play the game on his own from starting up of a business and this go through the growth phase also. He has to face all the risks outlined in the above scenarios. Apart from this he has to solve some specific risks also.

The last scenario is the actual game, i.e. the player has to develop an idea of particular software development. He has to present his idea before seed fund providers at various technical festivals. For presenting idea he/she need to prepare the business plan. Through this game the player will get an idea about to create a good business plan. The Fig. 5 related to the business plan preparation. That is for creating business plan the player need to select the seven relevant points from a set of relevant and irrelevant points. These seven relevant points will set with the help of some best business plans in the real world. If the player has any difficulty for creating the business plan, at that time he must recollect the knowledge from first scenario. That is he already learned the first scenario that gave the idea for preparing the business plan. After getting the approval of any of the seed fund provider he can start the business. At this stage he should be able to manage all the expenditure of the business. In such a way that it should not go beyond the fund provided by the seed fund provider or venture capitalist. Through this game the player can understand the procedures for starting and implementing of the software projects. The player has to select the building according to his requirements within the budget.

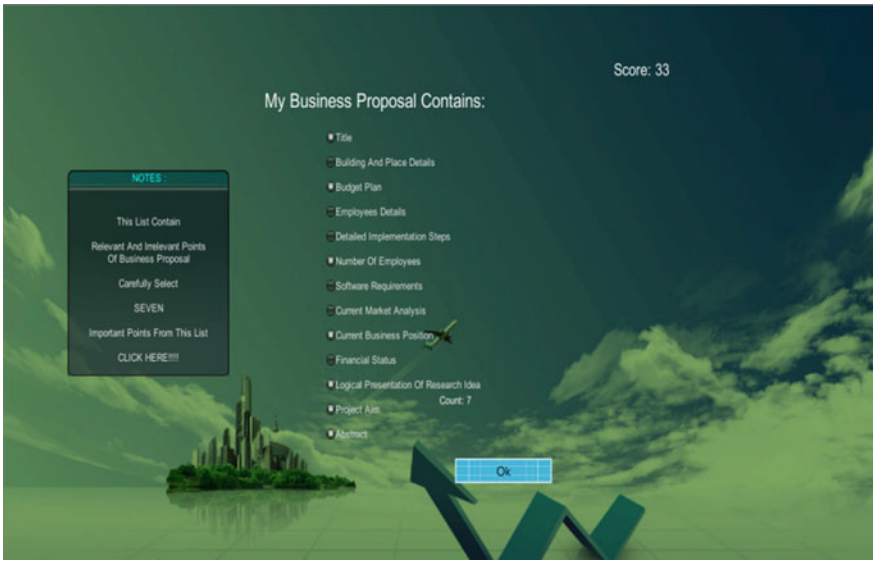


Fig. 5 Business plan preparation

And he should appoint employees to fulfil our requirements. Appoint of the labor should be in such a manner that would give maximum utility of the skill of the employees. After completing the project successfully, the player has to think the next stage of expansion of his business. One of the main challenges in this game is to spend the seed fund in an effective and profitable manner. The fund utilization is another challenge. Sometimes the fund is not enough for managing all operations. At that time will give another options like bank loan, accept sponsor ship or some other methods. With the help of the learning scenarios the player can easily solve these problems. The player will get the score based on his decisions. The higher the score the better will be the chance of success. After the last scenario we will provide the report of the player's game play. This will help the teacher to evaluate the student's performance and also help to self evaluation of the player.

4 Implementation Details

Unity 2D is the tool used for implementation of game. Unity is a flexible [14] and user friendly language. The study over real life situations to start a new business, different steps, challenges and its remedies are the initial step. Preparation of the scripts and storyboard of the game is followed behind. In implementation phase, build up a game based on the storyboard, while including many challenges and steps for a start up Software Company.

5 Performance Analysis

A comparative study has been conducted with a batch of management students, to analysis the effects of different sorts of learning methods like text based learning, learning using videos or presentation, and game based learning. The various factors like understandability of the content, interactivity, involvement, motivation and role of teachers have been scrutinized through this study. Figure 6 shows the variation in the proportionality for each methodology.

The understandability level for serious game is relatively high compared to other methodologies like video/presentation based leaning or text based learning; which indicates the knowledge accruing on the content will be more effective through game based learning. The growth in interaction level and engagement/motivation level indicates that the serious game is capability for involving students into the subject; whereas text based learning gives less involvement compared to others. The responsibility of a tutor is limited in the serious game as the students are learning on their own by playing the game. The tutors are required only for giving instructions for playing the game. In other methodologies, the involvement of tutor is inevitable. Text based learning is more bounded with a tutor, where he should be impart knowledge by giving more expansions and interpretations.

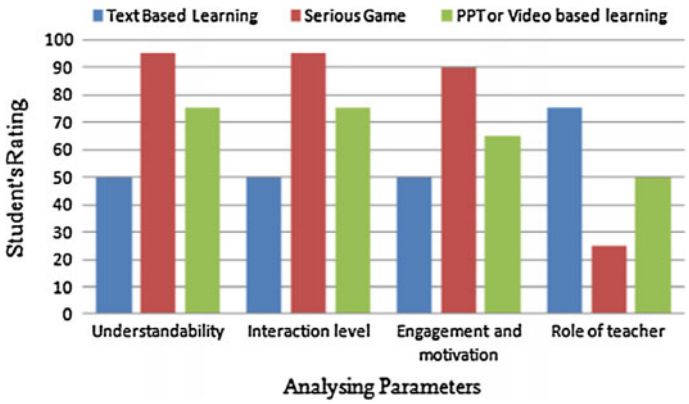


Fig. 6 comparative study of different methodology

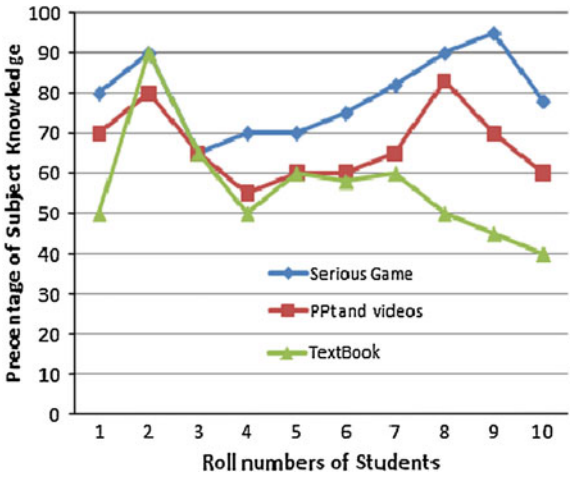


Fig. 7 Study on subject Knowledge

Figure 7 indicates the graphical representation of a study on subject knowledge, which clearly indicates the advantage of game based learning. The knowledge acquisition on the subject shows an improvement in serious game while comparing the other two methodologies.

6 Conclusion

It is clearly understood from performance analysis that, use of the serious game in learning renders a higher level of understandability, interaction and engagement of the students. Ninety percentages of students supported the game based learning

method. This research work ensures to accomplish the enhancement of business interest in people's mind. Educating upcoming entrepreneurs or managers through this gaming environment will save time for training and produce efficient employees. Though the game is business oriented, it not only aimed for entrepreneurs in the business sector, but business schools can also be benefited by utilizing it for the management students and teachers. The report will point out the pros and cons to the player related to a startup and running of a business in accordance with the way he played the game. The game ensured that the player could learn and evaluate their knowledge, over common problems and challenges related to a startup as well as running a business. Comparative study indicates the effectiveness of game based learning with other methodologies for the knowledge acquiring as well as critical thinking for facing the challenges in the business. The next phase of the game will include more challenges and steps for implementing new startups.

Acknowledgments We thank Amrita E-Learning Research Lab and Amrita Vishwa Vidyapeetham University for all the support and facilities provided for completing this paper.

References

1. Ministry of Labour and Employment, Government of India: In: Report on Employment and Unemployment survey, October 2010
2. Major Reason for Small Business Failure: https://www.sba.gov/content/what-are-major-reasons-small-business-failure
3. Different Learning Styles of People: http://blc.uc.iupui.edu/academic-enrichment/study-skills/learning-styles/3-learning-styles
4. de Freitas, S., Jarvis, S.: A Frame Work for Developing Serious Games to Meet Learner Needs (2006)
5. Management Games: www.learn4good.com/game/tycoonbusiness
6. Huynh-Kim-Bang, B., Wisdom, J., Labat, J.-M.: Design Patterns in Serious Games: A Blue Print for Combining Fun and Learning (2011)
7. Donovan, L., Pedagogical Lead: The use of serious game in the corporate sector. In: Astate of the Art Report, December 2012
8. For Enterpranurs Expertise, Wisdom and Resources for Startup and Enterpranurs: http://www.forentrepreneurs.com/business-models/why-startups-fail/
9. Wiegers, K.: 21 Project Management Success Tips (2013)
10. Allegra, M., Grande, V.D., La Guardia, D., Gentile, M.: A serious game to promote and facilitate entrepreneurship education for young students. In: Proceedings of the 2013 International Conference on Education and Educational Technologies
11. Greuter, S., Tepe, S., Peterson, J.F., Boukamp, F., d'Amazing, K., Quigley, K., van der Waerden, R., Harris, T., Goschnick, T., Wakefield, R.: Designing a Game for Occupational Health and Safety in the Construction Industry
12. Blazic, A.J., Arh, T.: Immersive Business Simulation Games: An Innovative Pedagogical Approach to e-Learning and Education (2013)
13. Critelli, M., Schwartz, D.I., Gold, S.: Serious Social Games: Designing a BusinessSimulation Game (2012)
14. Doran, J.P.: Unity Game Developments Blue Prints (2014)

Integrating Different Machine Learning Techniques for Assessment and Forecasting of Data

P. Vidyullatha, D. Rajeswara Rao, Y. Prasanth, Ravindra Changala and Lakshmi Narayana

Abstract Machine learning techniques are useful for solving different problems in many applications. Different machine learning techniques are available for assessment and forecasting of data. For illustration, this paper is focused alight on four machine learning techniques that are Weka, Tanagra, R software and MATLAB for showing different views to analyze and forecast the data. Weka is the most effective machine learning technique for regression and classification problems. Tanagra, the data mining tool, which is a supervised learning technique and also suitable for statistical analysis, classification and clustering problems. R software is a flexible programming accent for statistical computing and graphical settings. Finally, MATLAB is exclusive for technical computing representing the data in 2D, 3D and it is very effective tool for predictive analysis.

Keywords Machine learning · Meteorological parameters · nntool · Tanagra R software · WEKA

1 Introduction

Machine learning could be a set of mechanisms that permit to instruct data processing machines to execute tasks by giving datasets of what way they ought to be carried out. For the Artificial Intelligence View, Learning is fundamental to human knowledge and intelligence likewise, and vital for building intelligent machines. Machine learning [1] methods are mangled into two phases such as "Training" in which a model is learned from a group of training data and "Application" where the model is worn to build results around any recent test data. Some of the primary types of machine learning are as 'Supervised Learning', that is marked as label and the most common types are classification and regression. Where as in 'Unsupervised learning', the data is unlabeled that to analyze and discover patterns within. The two

P. Vidyullatha (✉) · D.R. Rao · Y. Prasanth · R. Changala · L. Narayana
Department of CSE, KL University, Guntur, Andhra Pradesh, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_12

most imperative cases are measurement diminishment and clustering. Machine learning [2] structure is in a general sense a cycle of activities that need to be performed. You can procure information from numerous sources; it may be information that is held by your association or open information from the Internet. There may be one dataset, or there could be ten or more. You must come to acknowledge that information will need to be cleaned and checked for quality before any transforming can occur. These processes happen at the prepare phase. The processing phase is the place where the work accomplishes. The machine learning schedules that you have made perform this phase. At last, the outcomes are exhibited. Reporting can happen in a mixed bag of routes, for instance, reinvesting the data again into a data store or reporting the results as a spreadsheet or report (Fig. 1).

For this, Frank et al. [3] Witten a Machine Learning framework on WEKA for Data Mining. Ramamohan et al. [4] presents an overview of the data mining tools like Weka, Tanagra, Rapid Miner and Orange. Smith et al. [5] presents the design of Tanagra and discussed the editing operations and evaluation of the expressivity of its generator. Devika et al. [6] studied the Neural Network Tool in MATLAB for Classification of Glass using. Demuth and Beale [7] gave a guidance on Neural Network Toolbox For Use with MATLAB. Kaundal et al. [8] forecast the blast disease in rice by comparing the Multiregression, backpropagation neural network and generalized neural network and he proved the backpropagation using matlab machine learning is efficient to predict the results. In this paper, the machine learning techniques on weka, Tanagra, R software and matlab.

2 Methods

2.1 Data Collection and Data Preparation

Information is gathered an hourly premise of both meteorological and surrounding information from thermal industry of 6 months of winter season from the months August to January of 2014. The meteorological parameters and surrounding information were ascertained for relationship studies, which were executed in

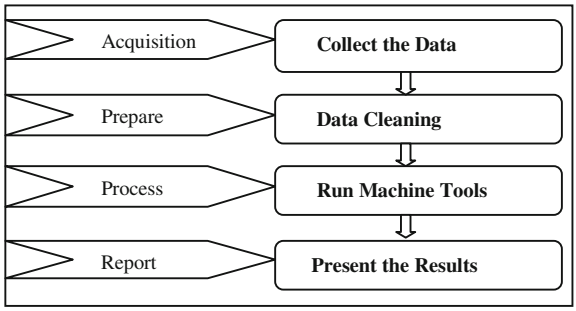


Fig. 1 The machine learning process

Tanagra apparatus in figure. After accumulation of information, the information was arranged by investigating the noise and missing information [9]. The some information is missed because of instrument calibrations or breakdowns. This missing information is filled by the method using interpolation.

2.2 Machine Learning Tools

Waikato Environment for Knowledge Analysis, Weka [10] which is an accumulation of visualization tools for information mining assignments like information investigation and prescient examination utilized as a part of a wide range of utilization regions. It is delivered at the University of Waikato, New Zealand and it is completely upheld in the Java programming dialect, free accessibility and effortlessness of utilization because of its graphical user interfaces. Information is preprocessed for importing the data as ARFF (property social level record) position. The Classifier empowers the user to apply regression process, Bayesian network, multilayer perceptron, decision tress are utilized to gauge the precision of the predictive model. Tanagra is a free DATA MINING programming for scholastic and exploration ideas. It presents a few information mining routines from initial information investigation, statistical learning, machine learning and databases zone. Tanagra is the replacement of SIPINA which executes different regulated training calculations, particularly an intelligent and visual development of decision trees. TANAGRA [11] is more effective, it contains some supervised learning techniques but also other models such as clustering, factorial analysis, parametric and non-parametric statistics, association rule, feature selection and construction algorithms. The user can plan outwardly an information mining process in a chart. Every node is a measurable or machine learning method, the association between two nodes speaks to the information exchange. The outcomes are shown in a HTML format. So it is anything but difficult to fare the yields keeping in mind the end goal to envision the outcomes in a program. It is likewise conceivable to duplicate the outcome tables to a spreadsheet. Tanagra [12] makes a decent compromise between the measurable methodologies; the multivariate examination systems for instance factor analysis, correspondence analysis, cluster analysis, regression and the machine learning techniques such as neural network, support vector machine, decision trees and random forest. The R Project for Statistical Computing (or simply R) is a capable information analysis tool. It is both a programming dialect and a computing and graphical environment. R [13] is free, open source programming built accessible under the GNU General Public License. It executes on Mac, Windows, and UNIX working frameworks. There are a few courses in which R will advantage you, be it as a data innovation proficient, business expert. R gives you chances to outline and execute perplexing, altered investigations that other programming bundles don't. In the meantime, R stays available and significant to a substantial group of onlookers of potential clients. R will aid you in settling on the best conceivable choices and permit you to rise to greatness as a premier strategist.

MATLAB [14] is an elite dialect for technical computing. It incorporates processing, visualizing and programming in an easy to utilize environment where controversies and solutions are related in natural scientific documentation. MATLAB is generally utilized as a part of all regions of connected arithmetic, in instruction and examination at engineering, and in the business. MATLAB remains for MATrix LABoratory and the product is developed around vectors and networks. This makes the product especially helpful for linear algebra however MATLAB is likewise an incredible instrument for comprehending logarithmic and differential comparisons and for numerical mix. MATLAB has intense realistic devices and can create decent pictures in both 2D and 3D. It is likewise a programming dialect, and is one of the most effortless programming dialects for composing scientific projects. MATLAB likewise has some tool kits valuable for sign handling, picture preparing, advancement, and so forth. In this paper, the proposed back propagation has been simulated using neural net toolbox, MATLAB. The model takes four inputs and one output which are shown in the figure. For training a neural network, levenberg marquardt (LM) [15] is considered which the combination of gauss Newton is and gradient descent method for reducing the errors.

3 Results and Discussions

In this paper, the results are analyzed using four powerful machine learning techniques which are shown in the figures. In Weka, Fig. 2, the classifier multilayer perceptron is used for forecasting ambient data at the learning rate 0.3 and momentum 0.2 of epoch 500. The multilayer perceptron is having 4-3-1

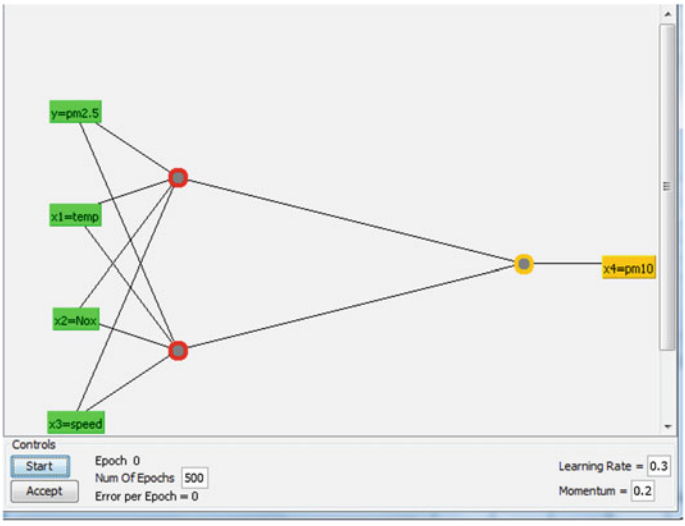


Fig. 2 Multilayer perceptron using machine learning technique WEKA

architecture, where the data input variables are four and one hidden layer with three neurons and one output layer. In Tanagra, Fig. 3, the relationship between variables is shown and It is clearly evident that independent variables x_1, x_3 are negatively

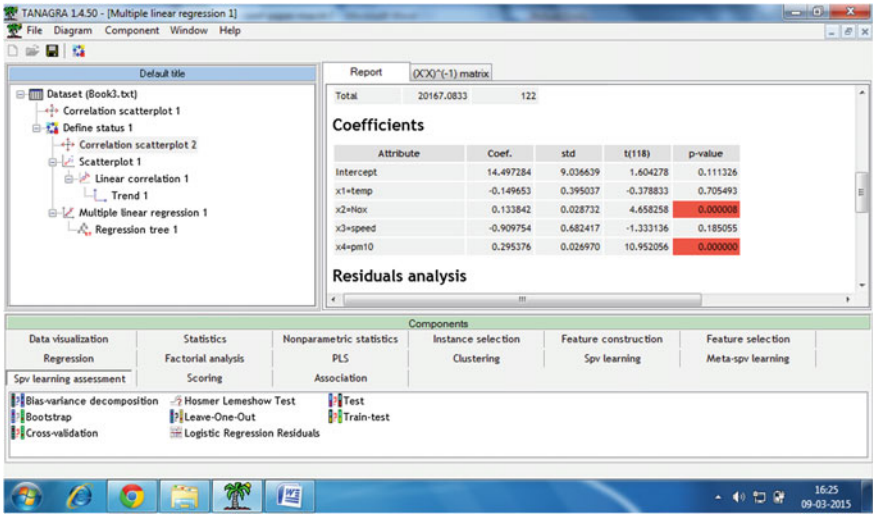


Fig. 3 Multilinear regression method using machine learning technique TANAGRA

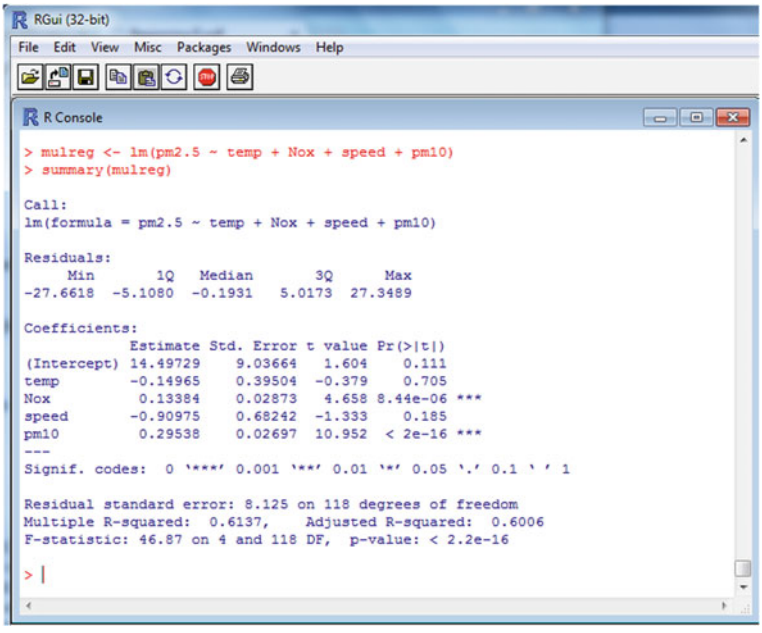


Fig. 4 R console showing correlation coefficients between the variables

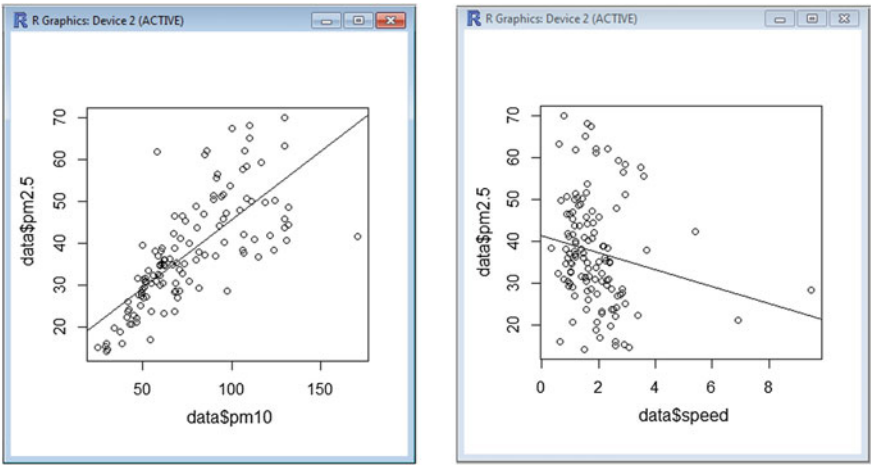


Fig. 5 Data plotting using machine learning technique R software

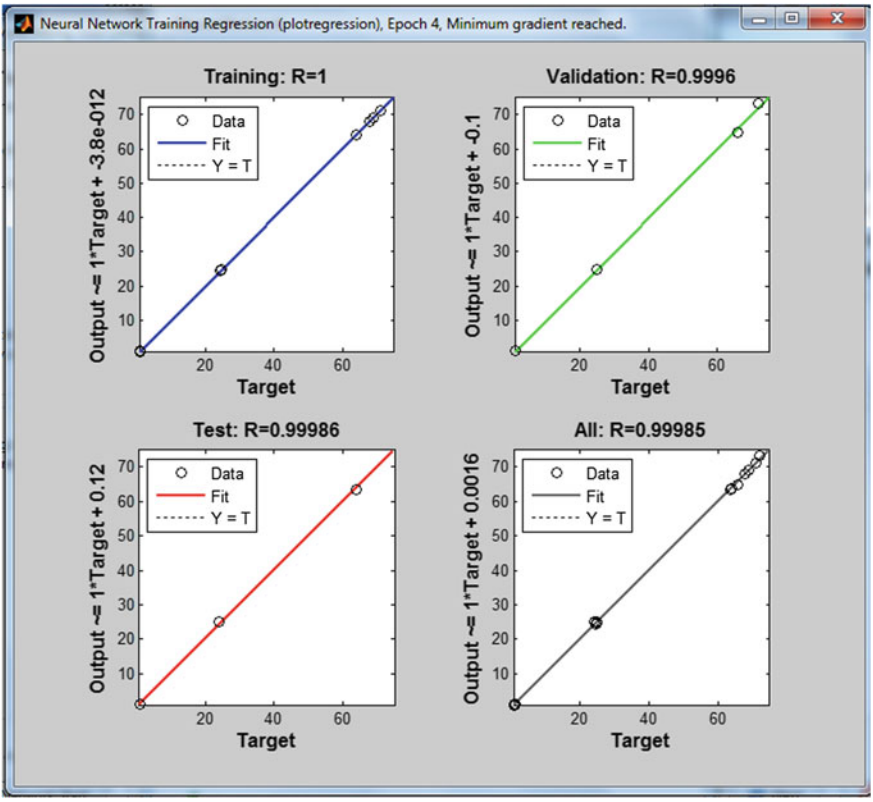


Fig. 6 Neural network training regression method using MATLAB

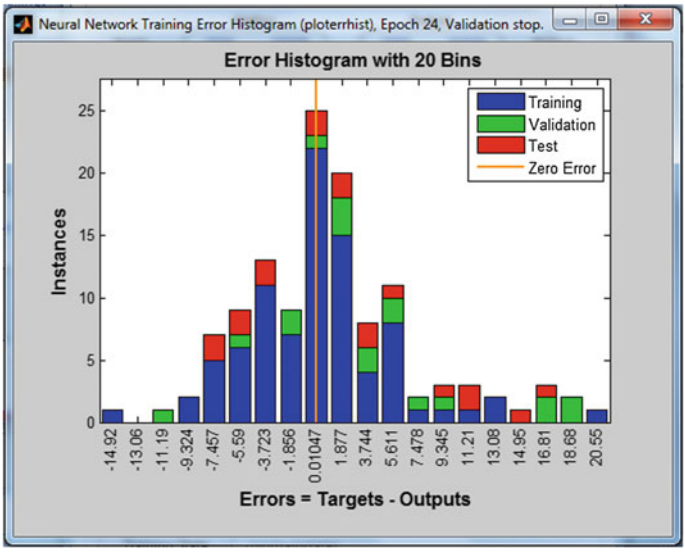


Fig. 7 Histogram for training data using MATLAB

correlated and x_2, x_4 are positively correlated with dependant variable 'y'. As the probability of F-test is less than 0.05 for variables x_2 and x_4 it can be concluded that the relation is not linear between x_2, x_4 and y. For further analysis, second order polynomial equation is formulated to calculate probability. If it is again less than 0.05, then go to third order polynomial equation and so on. Thus, Tanagra is a powerful statistical analysis tool. By using R software the graphical representation between variables are shown in Figs. 4 and 5, the flexible language R has statistical and capability for representing correlation between the variables by using simple functions. It is easy to download from the R archive network (CRAN) website. The nntool in mat lab is using to forecast the ambient data where the 80 % data handled as training data and 20 % as testing data and the network automatically stops when minimum gradient is reached at epoch 4 and R^2 is reached at one which is shown in Fig. 6. The error 0.0047 is shown in histogram representation and the validation stops automatically at epoch 24 which is in Fig. 7.

4 Conclusion

The main contribution of this review is to discuss the various Machine-Learning Techniques for data assessment and forecasting. In this paper, the ambient data is collected and predicted by using different machine learning techniques and also gives a relative comparison of all the techniques based on their advantages and limitations. Weka is the most effective machine learning technique for classification

and regression problems with graphical representation. The TANAGRA is an open source and supervised learning technique for statistical analysis, classification and clustering problems but not represented the data graphically. R software is a command-line interpreted language used for both statistical and graphical capabilities and it provides commercial statistical packages such as SPSS, SAS, and Minitab by analyzing the data. Among all Machine-Learning Techniques MATLAB is an elite dialect for technical computing and is efficiently used to forecast the data.

References

1. Dutton, D., Conroy, G.: A review of machine learning, knowledge engineering. Knowl. Eng. Rev. **12**(14), 341–367 (December 1997)
2. De Mantaras, Armengol, E.: Machine learning from examples: inductive and Lazy methods. Data Knowl. Eng. 99–123 (1998)
3. Frank, E., Hall, M., Holmes, G., Kirkby, R., Pfahringer, B., Ian, H.: WEKA A Machine Learning Workbench for Data Mining
4. Ramamohan, Y., Vasantharao, K., Kalyana Chakravarti, C., Ratnam, A.S.K.: A study of data mining tools in knowledge discovery process. Int. J. Soft Comput. Eng. (IJSCE) **2**(3), (July 2012). ISSN:2231-2307
5. Smith, G., Whitehead, J., Mateas, M.: Tanagra: A Mixed-Initiative Level Design Tool
6. Devika, A.S., Gupta, M.: Case study on classification of glass using neural network tool in MATLAB. In: International Conference on Advances in Computer Engineering & Applications (ICACEA-2014) at IMSEC, GZB 11
7. Demuth, H., Beale, M.: Neural Network Toolbox for Use with MATLAB
8. Kaundal, R., Kapoor, A.S., Raghava, G.P.S.: Machine learning techniques in disease forecasting: a case study on rice blast prediction. In: BMC Bioinformatics Article, 1–16, Nov 3 2006
9. Kotsiantis, S., et al.: Data preprocessing for supervised leaning. Int. J. Comput. Sci. **1**(2), 111–117 (2006)
10. www.cs.waikato.ac.nz/ml/weka/
11. http://tanagra.software.informer.com/
12. http://www.statisticalconsultants.co.nz/blog/tanagra-a-free-data-mining-program.html
13. http://cran.r-project.org/doc/FAQ/R-FAQ
14. www.mathworks.com
15. Camargo, L.S., Yoneyama, T.: Specification of training sets and the number of hidden neurons for multilayer perceptrons. Neural Comput. 2673–2680 (2001)

Simulative Analysis of DWDM Systems Using Loop Control in Inter Satellite Optical Wireless Communication Channel for 10,000 km Long Distance Transmission

Prabhdeep Kaur, Amit Gupta and Jaskaran Kaur

Abstract In this paper, the correlation among the various performance parameters of Inter satellite optical wireless communication (ISOWC) systems has been scrutinized for achieving a highly proficient system for satellite communication covering a distance of 10,000 km. This work further investigates the performance of IS-OWC link using a combination of loop control and Erbium doped fibre amplifiers. Besides this, an improved system for high data rate and long distance transmission using the Optisystem Software has also been modelled and analyzed.

Keywords DWDM · Inter satellite optical wireless communication · Eye diagram · Q-factor · Bit error rate · Loop control · Long distance transmission

1 Introduction

Communication systems have been evolving since the time communication has been started. This evolution has been taking place either in terms of technological advancements or in terms of cost effective communication systems. Earlier communication systems employed the use of bulky copper cables which were applicable to only for a limited distance communication. Then with the passage of time the focus shifted towards the optical fibers which in comparison to the copper cables were more reliable, secure and could cover a longer distance. But still the optical communication systems suffered from problems like higher cost and lengthy installation time. This led to the development of Optical Wireless Communication systems which have proved beneficial for communication in terms of providing a huge bandwidth, modest power and diminutive cost alternate to the presently used microwave communication systems [1].

P. Kaur (✉) · A. Gupta
Department of ECE, Chandigarh University, Gharuan 140413, India

J. Kaur
Department of ECE, Guru Nanak Dev University, Amritsar 143001, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_13

The Optical wireless communication systems are based on the basic principle of data transmission through air and using light as the carrier. The signal carrying information is being modulated on a laser which acts as a light source and is transmitted to another satellite in the free space. On the receiver side, this light signal is detected using a photo detector and then converted back into electrical signal. Depending on the type of communication required, the number of satellites used can be increased or decreased and the satellites' positions in their orbits can be adjusted accordingly to form a communication network [2]. Satellites are the objects which revolve around any other object in the space. Depending upon the distance of the satellite from the earth's surface, satellites can generally be classified as LEO, MEO, HEO and GEO i.e. Lower Earth Orbit, Medium Earth Orbit, Highly Elliptical Orbit and Geosynchronous earth Orbit respectively. An inter satellite link is a communications link that connects two separate satellites directly. One satellite could have several links to numerous other satellites. Inter satellite links are very important for communication of two satellites in same orbit or two different orbits like communication between a LEO and GEO satellite. Depending on the type of application a LEO or a GEO satellite can be used for the purpose of communication [3].

The number of satellites is increasing every year to fulfill the increasing demand for more applications. Optical wireless communications are thus providing an alternate for the bandwidth hungry communications [4]. Thus the deployment of DWDM in hybrid optical and wireless communication systems has increased in the past few decades. The inherent properties like huge bandwidth, no requirement of licensing and easy deployment make these systems to be used efficiently for broadband services. Yet the systems suffer from many setbacks like bandwidth inefficiency, noises and errors in transmission, thus leading to a degraded or less competent communication system [5, 6].

The previously reported works have undergone transmissions up to 5000 km and that too at data rates below 10 Gbps [7]. In this work, an enhanced communication system has been proposed covering a distance of 10,000 km using a combination of loop control and erbium doped fiber optical amplifiers for a 32 channel DWDM Inter satellite optical wireless communication system.

The paper is organized as follows: Sect. 2 describes the IS-OWC system block diagram, Sect. 3 explains the system specifications used for simulation, Sect. 4 highlights the results obtained and finally this paper is concluded in Sect. 5.

2 IS-OWC System Block Diagram

2.1 DWDM Transmitter

A transmitter can be elucidated as a component in DWDM system which provides the source signals to be multiplexed further. While designing of DWDM systems, it becomes imperative to carefully analyze the characteristics of optical transmitters in

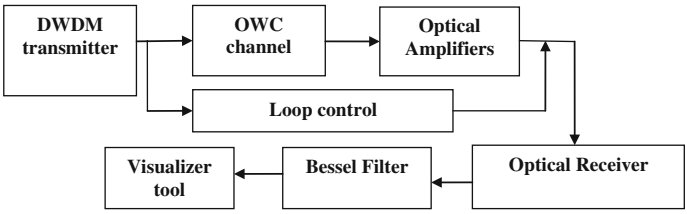


Fig. 1 The proposed DWDM system for ISOWC communication

terms of their performance characteristics. Depending on the system requirements, single or multiple transmitters can be utilized as light sources in the system. The incoming data bits in the form of a binary '1' or '0' can be modulated accordingly. A binary '1' indicates the presence of light and a binary '0' represents absence of light [8] (Fig. 1).

Generally Lasers are used for creation of light pulses. A light pulse usually constitutes a light carrier having a unique wavelength generally articulated in nanometres and symbolized as lambda. For transmission of data, a stream of digital data is transmitted over a physical layer device, which is being acknowledged by a light laser that may be an LED or a Laser that interfaces a fiber optic cable. The transmitter performs the electro-optical conversion i.e., it converts the electrical signal into optical signal (electrons to photons). Combinations of ones and zeros are responsible for triggering of the light source. These light pulses are passed through the optical channel and then received at the receiver by a photo detector like photodiode and the optical signal is again recovered back as an electrical signal at the receiver [9].

2.2 OWC Channel

The term optical wireless refers to transmission of optical data through wireless medium but the transmission takes place through unguided media. If the transmission of data is undertaking between two satellites, the same communication can be called as Inter satellite Optical Wireless Communication and the OWC channel being used here can be called as IS-OWC channel. This implementation can be performed using Optisystem software by using the OWC channel between an optical transmitter and optical receiver having an optical antenna with 15 cm aperture at each end. The channel is modelled at a wavelength of 1550 nm and two optical channels are modelled each covering a distance of 50 km. For the sake of simplicity, the transmitter and receiver are assumed to be ideal having 100 % optical efficiency and gains are also kept to be 0 dB. Errors due to mispointing and scintillation losses are also assumed to be zero. As the satellites are at a very high altitude from the earth's surface, the attenuation resulted because of atmospheric effects of the various layers of atmosphere is neglected [10].

2.3 Loop Control

Loop control methodology can be opted for increasing the range of optical communication systems. The number of loops decides the distance travelled by the optical signal from transmitter to receiver. As the number of loops increases, the distance travelled multiplies the OWC link distance and hence results in long distance transmission system for communication. The system is modelled using 20 loops configuration, hence covering a maximum distance of 10,000 km which is very large as compared to the previously reported works [11].

2.4 Optical Amplifiers

Optical amplifiers are the devices mainly used for the purpose of optical signal amplification (Fig. 2).

Two main classes of optical amplifiers generally exist:

2.4.1 Semiconductor Optical Amplifiers (SOAs)

- **Fabry Perot Amplifiers (FPA)**: These amplifiers are based on the principle of reflection of light. As the light enters into the amplifier, amplification of light takes place. After amplification, this light gets reflected back and forth in between the mirrors unless emitted at a higher intensity. These amplifiers are sensitive to temperature and input optical frequency.
- **Travelling–Wave Amplifiers**: TWAs are approximately same as the FPA except that the end facets are either anti reflection coated or cleaved at an inclination such that no total internal reflection could take place and the input signal could be amplified only single time during one pass through the instrument. A variety of advantages make TWAs more popular such as low polarization diversity and high bandwidth [12].

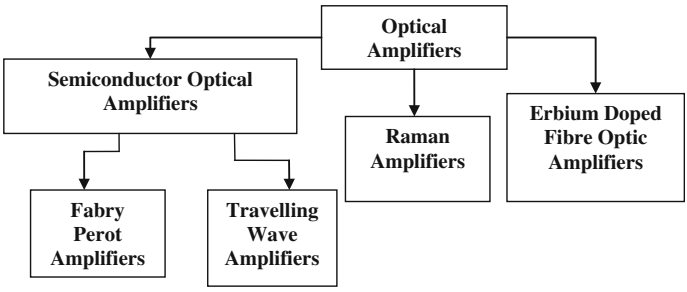


Fig. 2 Different types of optical amplifiers

2.4.2 Raman Amplifiers

In Raman Amplifiers, unlike EDFA, the amplification is realized through non linear interaction between the signal and a pump laser inside an optical amplifier. A Raman amplifier utilizes the basic principle of Raman Effect in which the power is transferred from lower wavelengths to higher wavelengths [13].

2.4.3 Erbium Doped Fibre Amplifiers

The most widely used fibre amplifiers in optical communication are the Erbium doped fiber amplifiers for the purpose of long distance communication. EDFAs are usually preferred for the long range communication because they provide in-line amplification of the signal without the requirement of electrical conversion. Moreover, EDFAs provide high power efficiency from pump to the signal power and possess relatively flat gain characteristics which make them suitable for long distance communication. EDFAs are generally operated near 1550 nm window for achieving higher efficiency. The EDFAs are formed by doping the rare earth ions like the Erbium ions (Er3+) which when excited through pumping of lasers to higher level results in a high signal gain. In the proposed model, an enhanced version of EDFA i.e. the EDFA black box has been utilized for system modelling and simulation [13].

2.5 Optical Receiver

While designing an efficient optical communication system, the receiver modelling is as crucial as the transmitter [14]. A number of parameters need to be accessed for designing a competent receiver such as the receiver sensitivity which is mathematically equal to the minimum average for a specified bit error rate and besides the link budget parameters including the affect due to various losses needs to be taken into consideration while modelling of a receiver. Thus all the important parameters should be analyzed carefully prior to designing of optical receiver for maximum efficiency and better output. In the proposed model, an Avalanche photodiode is being utilized at the receiver. An APD is suitable for applications requiring high sensitivity and long distance communications [14, 15].

2.6 Bessel Filter

Bessel filter is the most commonly used filter due to linear phase response characteristics. Also this filter is free from ripples in pass band and possesses monotonic

decay in stop band. This type of filter is generally preferable only for signals like square and triangular wave having a broad spectrum suitable for the signal such as square wave and triangle wave which have wide spectrum [16].

2.7 Visualizer Tool

The visualizer tools are available in the form of optical and electrical analyzers for visualizing the results at the output. The different tools analyze the various parameters based on different simulation efforts and helps to establish the relationship among them using input output graphs. For instance eye diagram analyzer, BER analyzer, Optical Spectrum analyzer and many more [17].

3 Simulation Specifications

The modelled system consists of a WDM transmitter operating at a wavelength of 1550 nm, having an input power of 10 dBm and NRZ modulation format. The optical wireless link is simulated using Optical wireless communication channel is modelled at 10 Gbps for a distance of 10,000 km using loop control at a wavelength of 900 nm. As the number of loops is equal to 20, and two optical wireless channels are employed at a distance of 250 km, the total distance thus travelled equals to 10,000 km. EDFA black box has been used for the amplification purpose and at the receiver side, Avalanche photo diode has been used as a detector. The output ports 1, 8, 16, 24, 32 are connected to the receiver and the outputs are analyzed using various analyzers such as the BER analyzer, optical spectrum analyzer, RF spectrum analyzer and the optical time domain analyzer, at the output side [18, 19].

The proposed IS-OWC system has been premeditated keeping in mind the various system performance parameters in order to achieve a high throughput. The use of loop control in combination with a set of optical amplifiers has been employed to accomplish a high Quality factor and a low bit error rate. The adjustments in operating modes and parametres of the different components has thus resulted in an optimized system for inter satellite optical wireless communication systems. The simulation specifications for the proposed work have been summarized in Table 1.

4 Results and Discussions

The results reveal that the quality factor decreases with increase in distance and correspondingly BER increases. At 10 Gbps, a high Quality factor of 37.5 and a BER of 2.65×10^{-312} is achieved using loop control and EDFA amplifiers

Table 1 Simulation specifications for the configuration set-up

Parameter	Value/type
Input power (dBm)	10
Transmitter wavelength (nm)	900
Data rate (Gbps)	10
Distance (km)	10,000
Channel wavelength (nm)	1550
Modulation scheme	NRZ
Amplifier used	EDFA black box
DWDM transmitter	32 channel

for a distance of 10,000 km. The eye diagrams have been verified through simulations. The proposed work is better than the previously reported works in terms of high Q-factor, low BER and a long distance transmission leading to a highly proficient system for communication.

Figure 3 shows the BER pattern observed at a distance of 10,000 km. It depicts that a clear and open eye diagram is obtained having an eye height of 49.89, threshold value of 8.14 and a decision inst. of 0.5303 is attained.

Figure 4 shows the Q-factor obtained at a distance of 10,000 km. The maximum achievable Q-factor is observed to be 37.74 which is acceptable for such a long distance transmission. The value of Quality factor plays a crucial role in deciding the system performance. Greater the value of Q-factor better will be the system performance and lower the value of Q-factor, lesser will be the efficiency of the system. But this value needs to be parameterized accordingly with the increase in

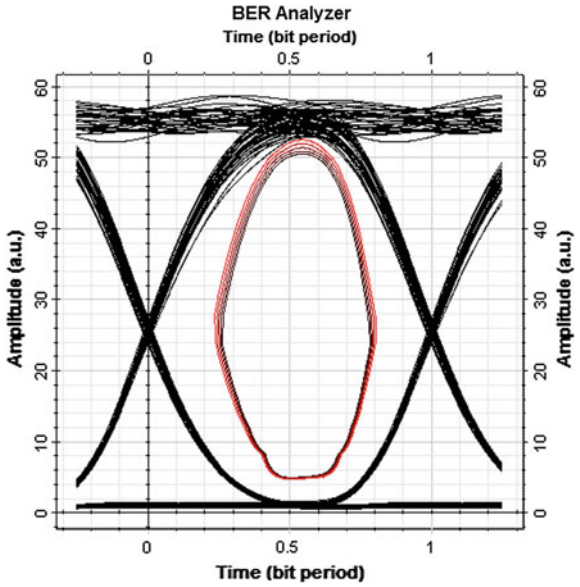


Fig. 3 The BER pattern obtained at a distance of kilometres at 10,000 km

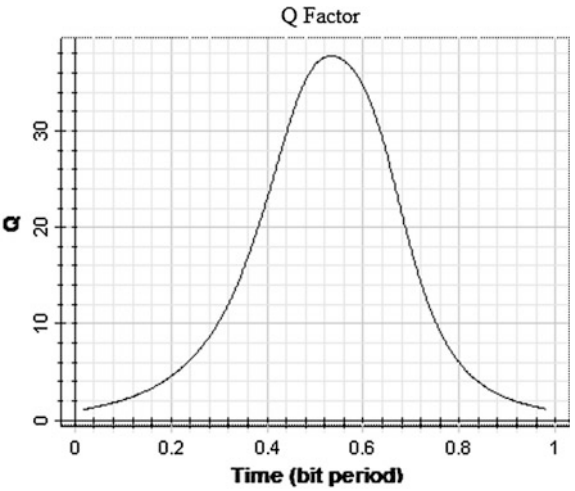


Fig. 4 The Quality-factor curve obtained through simulation at 10000 a distance of 10000 kilometres

distance. Thus a trade of needs to be maintained between the distance travelled and the quality of the communication [20].

Figure 5 implies that the min. BER corresponding to a Quality factor value of 37.74 is observed to be 2.65×10^{-312}. The indications thus disclose that both the parameters satisfy the requirements of an effective communication system leading to an improved 32 channel DWDM system for IS-OWC channel implementation at a data rate of 10 Gbps [21].

Figure 6 indicates the optical spectrum analysis for the DWDM signal transmitted through the IS-OWC link. The green colored lines indicate the noise power

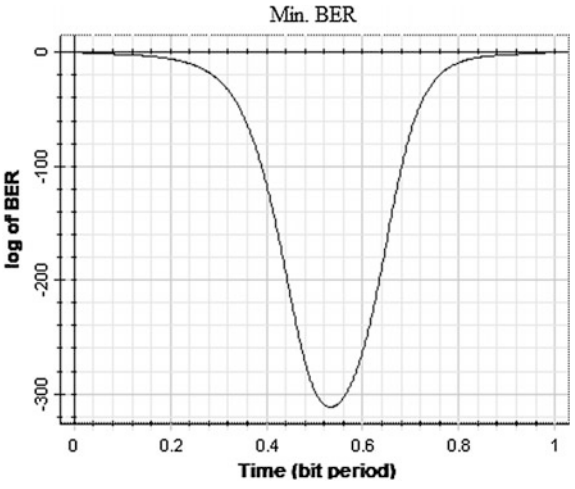


Fig. 5 The BER curve at a distance of 10,000 km

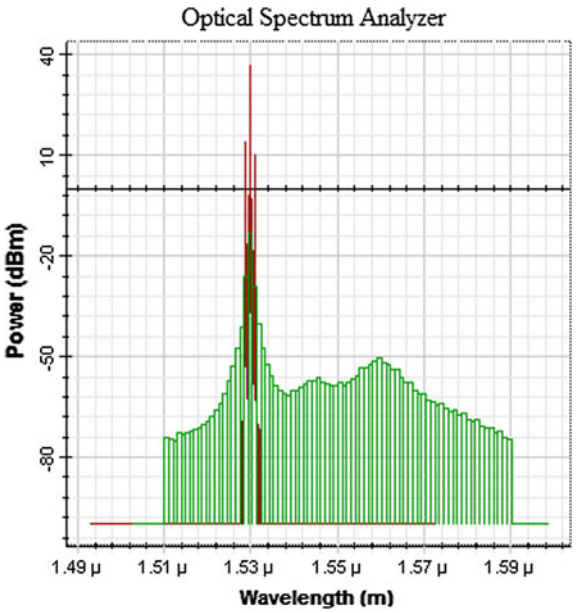


Fig. 6 Optical spectrum analysis for an input power of 10 dBm and a distance of 10,000 km

in dBm and the red portion indicates the sampled power of the DWDM signal. The results reveal that the maximum sampled power of 43.6 dBm and a minimum power of −106.8 dBm have been realized for the proposed system scenario in the case of sampled power. In the case of noise power, a maximum of −8.59 dBm and a minimum of −104.35 dBm has been reported for the above mentioned communication system.

Figure 7 depicts the RF spectrum for the DWDM signal transmitted at a distance of 10,000 km. The graph indicates that a maximum of 29.55 dBm and a minimum of −93.03 dBm of signal power have been reported for the given specifications. However, for the RF spectrum the noise power is negligible as compared to the optical spectrum.

Figure 8 shows the time domain analysis of the optical signal. The graph is plotted against signal power along Y axis in watts and time along the X-axis in seconds. A maximum of 24.57 W and a minimum of −1.17 of signal power are achieved when simulated in time domain.

The noise power obtained in time domain of optical noise signal is indicated in Fig. 9. A noise power of 0.606 W has been observed as the maximum noise power level and a noise power of 0.029 has been observed as the minimum noise power level for the system simulated at 10,000 km long distance and a data rate of 10 Gbps [22].

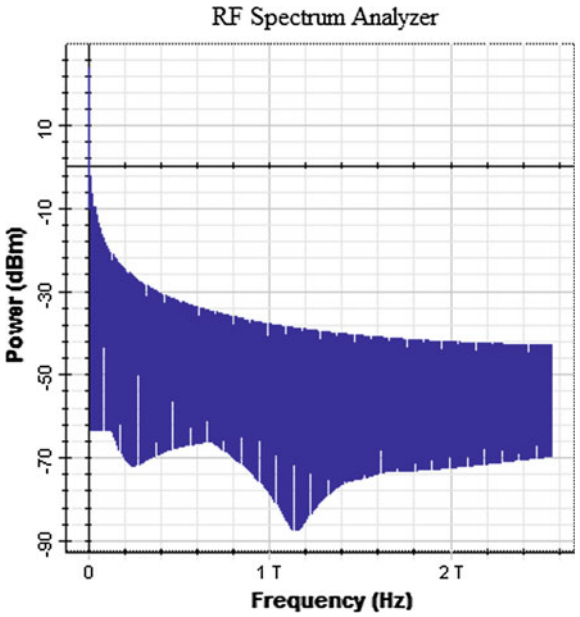


Fig. 7 RF spectrum analysis of the signal

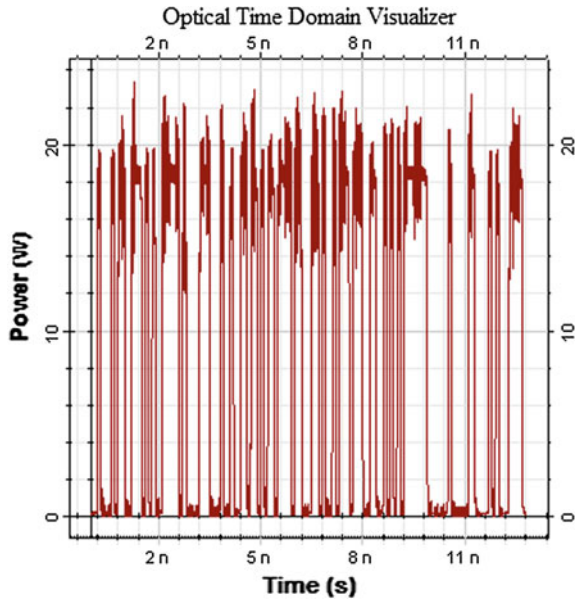


Fig. 8 Signal power in optical time domain

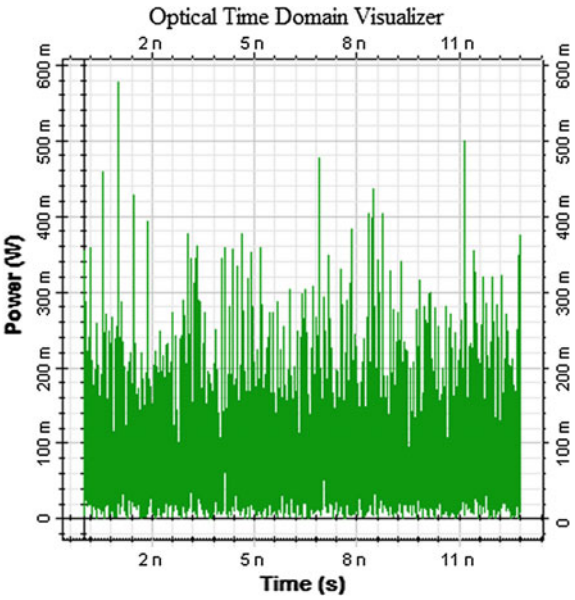


Fig. 9 Time domain analysis of optical noise signal

5 Conclusion

This paper analyzed a long distance transmission up to a distance of 10,000 km at 10 Gbps using combination of loop control and EDFA amplifiers for a 32 channel DWDM system. It has been observed that a high Quality factor of 37.5 and a low BER of 2.65×10^{-312} has been realized for DWDM communication systems using ISOWC link modelling at such a long distance. The Q factor and BER values offer a superior communication system than the previous systems that could cover below 10,000 km and the data rates were also below 5 Gbps. The proposed work is thus better in terms of distance coverage and data rates accomplishments. However a trade-off has to be maintained among the values of distance travelled, the quality factor and the bit error rate obtained for realizing a well-organized system for communication.

References

1. Kaplan, L.: Optimization of satellite laser communication subject to log-square-hoyt fading. IEEE Trans. Aerosp. Electron. Syst. **47**, 4 (2011)
2. Naji, A.W., Khateeb, W.A.: Performance analysis of a free space optical link with multiple transmitters/receivers. In: IEEE International Conference on Space Science and Communication. Penang, Malaysia (2012)

3. Tan, L.: Wave-front distortion and its effect on inter-satellite laser communication systems. Harbin Institute of Technology National Key Laboratory of Tuneable Laser Technology Harbin, China (2010)
4. Nadeem, F., Kyicera, V., Awan, M.S., Leitgab, E., Muhammad, ss, Kandus, G., Muhammad, S.S., Kandus, G.: Weather effects on hybrid FSO/RF communication link. IEEE J. Sel. Areas Commun. **27**(9), 1687–1697 (2009)
5. Ijaz, M., Ghassemlooy, Z., Ansari, S., Adebanjo, O., Minh, H., Rajbhandari, S., Gholsmi, A.: Experimental investigation of the performance of different modulation techniques under controlled fso turbulence channel. In: Telecommunications (IST), 5th International Symposium, pp. 59–64 (2010)
6. Hameed, W., Muhammad, S.S., Sheikh,N.M.: Integration scenarios for free space optics in next generation (4G) wireless networks. In: Communication systems networks and digital signal processing (CSNDSP), 7th International Symposium, pp. 571–575, Newcastle upon Tyne (2010)
7. Singh, K., Kaur, K.: Performance analysis of subcarrier multiplexing technique on intersatellite optical wireless communication and its comparison with wavelength division multiplexing. Int. J. Res. Comput. Eng. Electron. **3**(3) (2014)
8. Sharma, C., Singh, S., Sharma, B.: Investigations on bit error rate performance of DWDM free space optics system using semiconductor optical amplifier in inter satellite communication. Int. J. Eng. Res. Technol. (IJERT), **2**(8) (2013)
9. Antil, R.P., Beniwal, S.: An overview of DWDM technology & network. Int. J. Sci. Technol. Res. **1**(11) (2012)
10. Chaudhary, A., Singh, S., Minocha, G., Rana,H.: Optimization of performance of inter-satellite optical link with effect of bit rate and aperture. Int. J. Sci. Res. Eng. Technol. (IJSRET), **3**(2) (2014)
11. Optical systems-WDM designs. www.optiwave.com. Accessed 2 March 2015
12. Keiser, G.: A Book on Optical Fiber Communications, 3rd edn. McGraw-Hill
13. Singh, M., Raghuwanshi,S.K.: Analysis of the crosstalk in optical amplifiers. ICTACT J. Commun. Technol. **4**(2) (2013)
14. Winzer, P.J., Pfennigbauer, M., Strasser, M.M., Leeb, W.R.: Optimum filter bandwidth for optically pre amplified NRZ receivers. J. Lightwave Technol. **19**(9), 1263–1273 (2001)
15. Lima, I.T., Lima, A.O., Zweck, J., Menyuk, C.R.: Performance characterization of chirped return-to-zero modulation format using an accurate receiver model. IEEE Photon Technol. Lett. **15**(4), 608–610 (2003)
16. Jingbo, X., Xiaohong, X., Pengyu, Y., Haijun1, L., Xin, G.: The design of anti-aliasing analog filter for data acquisition in the surface measurement. Int. J. Signal Process. Image Process. Pattern Recog. **7**(5), 169–176 (2014)
17. Rawi, N.I.M., Jamaludin, M.Z., Abdullah, F.: Optisystem: an alternative to optoelectronics and fiber optics teaching e-laboratory. Int. J. Asian Soc. Sci. Special Issue: International Conference on Teaching and Learning in Education (2013)
18. Sarup, V., Gupta, A.: Performance Analysis of an Ultra High Capacity 1 Tbps DWDM-RoF System for Very Narrow Channel Spacing. WECN, IEEE (2014)
19. Aggarwal, I., Chawla, P., Gupta, R.: Performance evaluation of inter satellite free space optical communication system with varied parameters and transceiver diversity. Adv. Electron. Electr. Eng. **3**(7) (2013)
20. Mili, M.R., Hamdi, K.A.: Minimum BER analysis in interference channels. IEEE Trans. Wirel. Commun. **12**(7) (2013)
21. Saini, S., Gupta, A.: Modelling and performance analysis of DWDM based 100 Gbps low power inter-satellite optical wireless communication (LP-IsOWC) system. SOP Trans. Signal Process., **2**(1) (2015). ISSN-2377-0546
22. Kalbani, A.I., Yuce, M.R., Redouté, J.M.: Design methodology for maximum power transmission, optimal BER-SNR and data rate in biomedical implants. IEEE Commun. Lett. **17**(10) (2013)

A Novel Rate Based Overload Control Method for SIP Servers

Abdullah Akbar, S. Mahaboob Basha and Syed Abdul Sattar

Abstract Voice Over IP applications are becoming popular due to proliferation of smart phones. Many people are using Internet on these devices. To support millions of VOIP users and increase QoS (Quality of Service) VOIP installations employs large number of SIP Proxy servers to route calls. To achieve high throughput and minimum response times usually a load balancer is used to dispatch calls to SIP proxy servers. Overload control in SIP is a well known problem in SIP proxy servers. Local overload control, hop-by-hop and end-to-end overload control methods are used in SIP networks. Stand-alone overload control method monitors the processor and memory resources of the proxy server where overload control is implemented. It does not consider call-rates to avoid the overload in SIP server. In this paper we developed a novel rate-based overload control method to detect and control overload in SIP servers. We have implemented the algorithm by modifying leading open source Kamailio SIP proxy server. We have evaluated our algorithm by experimental test setup and found results are outperforming the standard standalone overload control algorithms in terms of throughput and response time.

Keywords Voice over—IP · Session initiation protocol · Overload control · Kamailio

A. Akbar (✉)
Jawaharlal Nehru Technological University, Hyderabad, Telangana, India
e-mail: akbar.jntuphd@gmail.com

S. Mahaboob Basha
Al Habeeb College of Engineering & Technology, Hyderabad, Telangana, India
e-mail: smbasha.phd@gmail.com

S.A. Sattar
Royal Institute of Technology & Science, Hyderabad, Telangana, India
e-mail: syedabdulsattar1965@gmail.com

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_14

1 Introduction

Session Initiation Protocol [1] is a signalling protocol to maintain multimedia sessions in VOIP networks defined in RFC 3261. It is also adapted by 3GPP [2] to maintain multimedia sessions in Internet Protocol Multimedia System (IMS). The RFC defines SIP as an application-layer control protocol that can establish, modify and terminate multimedia sessions or calls. SIP is often seen as a good replacement of H.323 and today most of VOIP manufacturers have adopted SIP in their product portfolios. To support millions of VOIP users and thousands of calls large scale ISP use cluster of servers. There should be a dispatcher which distributes call requests to SIP Proxy servers. Load balancing dispatchers use load balancing algorithms to take intelligent decisions to distribute load. Overload detection and control is essential to avoid sudden collapse of SIP service due to flash crowds and Denial of service attacks. SIP protocol define addresses to identify Callers and callee's. Basic SIP operation is depicted in Fig. 1. To make a SIP call, a caller uses location service to extract the SIP address of the callee and make a call by sending INVITE request to the proxy server. Users of SIP service can register their location with SIP servers. The proxy server resolves the callee address and forward the

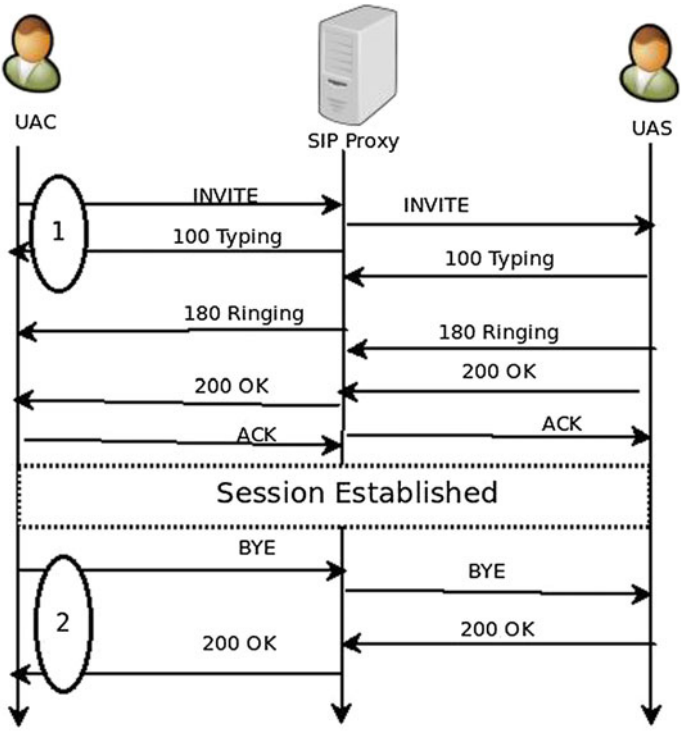


Fig. 1 Basic SIP transaction flow

INVITE request to the callee and reply with 100 status message (trying) to the caller.

Once the caller phone rings it send 180 status message to proxy. Proxy server forwards 180 status to caller. If caller and callee agree to communicate they respond with 200 OK. Then onwards call will established media is transmitted independent of proxy server. In SIP protocol call is distinguished from transaction. SIP protocol mandates all responses should contain the same values in the Call-ID, CSeq, To, and From fields (with the possible addition of a tag in the To field) for corresponding SIP requests. This enable responses to be identified with requests. SIP is independent of transport protocol. SIP can use TCP (Transmission Control Protocol), UDP (User Datagram Protocol) or SCTP (Stream Control Transmission Protocol). A successful SIP invitation consists of two requests, INVITE followed by ACK. The ACK request is not part of the SIP transaction which is followed by INVITE request. If the user want to disconnect the call he has to send BYE request. BYE request may come from either caller or callee. 200 OK response should commence after BYE request in order to complete SIP transaction. A Redirect server will forward the INVITE if one of the Proxy server not reachable or failed.

This paper is based on client-to-server overload control in SIP network. It proposes a rate-based overload control method to protect a SIP proxy server from sudden collapse and performance degradation under overloaded condition. INVITE rate is measured in SIP proxy server. Overload control is implemented by cooperatively controlling the incoming INVITE rate by receiving feedback from UAC through SIP message parameters.

The remaining sections of the paper are divided as follows. Section 2 gives related work done in SIP overload control. Section 3 describes overload control problem and its control methods. Section 4 explains the proposed rate-based SIP overload control method and its design aspects. Section 5 provides experimental test bed details. Section 6 presents our experimental results and Sect. 7 concludes the paper.

2 Related Work

The Load balancing algorithms and fail over schemes for web servers has been evolved in literature. Session Initiation Protocol is different from HTTP, there is less significance for sessions in HTTP where as SIP dispatcher should maintain sessions in-order to route requests of same session to the same proxy server. For providing high availability in web servers usually IP address takeover, MAC address takeover schemes are used. Y. Hong, C. Huang etc. identified two broad SIP overload control mechanisms: Load reducing and Load balancing approach. Load reducing approach reduces the traffic by rejecting incoming call requests to avoid overload where as Load balancing equally distribute traffic among local SIP servers to avoid overload in SIP network [3]. Load balancing algorithms for SIP server clusters are classified as static and dynamic. Cisco's architecture uses static

load balancing based on DNS SRV and Naming Authority Pointer records. Load dispatcher selects the server based on priority and weight fields of these records. Kundan Singh and Henning Schulzrinne applied web server redundancy techniques to SIP telephony. They have described various load sharing methods based on DNS, SIP identifier, redundant servers with unique IP address and Network Address Translation on a two-stage SIP server architecture. They have also considered fail over techniques based on DNS, clients, database replication and IP address takeover [4]. Wright etc. studied SIP performance on multi core systems [5]. Hongbo Jiang and ArunIyengar etc. introduced three novel dynamic load balancing algorithms for SIP Server clusters. These algorithms are based on Session Aware Request Assignment technique. Each algorithm works on features combining knowledge of SIP protocol, distinguish between calls and transactions and dynamic estimates of server load. Transaction Least Work Left outperforms all remaining algorithms in terms throughput and response time [6]. Luca De Cicco and Giuseppe Cofano etc. proposed a novel local overload controller using control theoretic technique called extremum seeking. They had practically implemented in Kamailio server but does not control the overload in entire SIP network [7]. Jianxin Liao and Jinzhu Wang etc. describe a distributed end-to-end SIP overload control method on edge servers of SIP network. They have given theoretical analysis and extensive simulations are carried out on NS-2 simulator to prove the control system [8]. Ahmadreza Montazerolghaem and Mohammad Hossien Yaghmaee Moghaddam implemented window based adaptive overload control in Asterisk Proxy servers [9].

3 SIP Overload Control Problem

When a SIP proxy server cannot handle call requests at a rate greater than or equal to rate at which they are supposed to handle for an extended period of time, then the server is said to be overloaded. Overload control is a method which detect the overload and degrade the performance of the server gracefully. In SIP networks overload occurs due to lack of resources, flash crowds, denial of service attacks, in emergency situations, unintended traffic due to software and hardware errors in UAC's. If overload is not properly controlled it not only degrade the performance of the SIP server but it halt the entire VOIP service. Overload causes inconvenience to customers and effect the trust worthiness of VOIP Service Network Providers. If overload has not been taken care by VOIP service providers they may not compete traditional PSTN service providers. IETF does not provide proper standards and procedure to deal overload in SIP network. It defines only 503 response mean Service Unavailable returned by SIP servers. SIP specifications does not provide sophisticated methods to deal overload. Three straight forward approaches to deal overload problem is dropping excess traffic, rejecting with 503 status code indicating service unavailability and rejecting with 500 indicating retransmitting after retry period.

3.1 SIP Overload Control Algorithms

Many congestion control algorithms are developed for IP based systems which assume all incoming events consume same resources [10–12]. We cannot employ these algorithms for SIP Servers since SIP requests and responses does not consume same resources. INVITE request consume more resources compared to BYE. REGISTER request involves data base look up. So there is a need to develop new overload control algorithms for SIP based systems. Overload control algorithms can be classified as local, hop-by-hop and end-to-end depending on the decision where the excess traffic should be regulated. Local overload control algorithm monitors the SIP server CPU and memory resources and detects the overload and take decision according to the algorithm implemented in the server. Hop-by-hop algorithms take feedback from downstream server and regulate the excess traffic. End-to- end algorithms take consideration of entire SIP network and control the overload. Implementation of algorithm is easy in case of Local SIP overload control algorithms rather than end-to-end overload control.

3.2 Rate Base Overload Control

Overload control mechanism is implemented at the SIP server to protect from collapse due to excess SIP traffic. Server monitors its processor and memory resources to detect the overload. This type of control is called stand-alone overload control. Most of the current implementations use stand-alone overload control because it is easier to implement and can independently deploy on the server without relying on other SIP components. Disadvantage of this approach is it does not reduce overall overload in SIP network even though it reduces the excess traffic coming to itself. Sometimes a situation occurs overload prevails in overall SIP network to lead to halt of entire VOIP service. Rate-based overload control is a method which takes decision on information received from user clients. Client and Proxy server should exchange information regarding how much amount of traffic it receiving by a Via header of SIP message. In this paper we have implemented a novel rate-based overload control algorithm on a Kamailio server.

4 The Proposed Rate-Based Overload Control Method

In this section we describe the proposed cooperative overload control method, it is easy to implement without modifying the existing SIP protocol.

4.1 Design of Overload Control Algorithm

Our overload control method depends on the system model specified in RFC 6357 [13]. In Fig. 2 S is the load samples measured by the Monitor and reports to Control Function. T is the throttle to be set by Control Function to adjust the load sent to the SIP processor. F is the feedback sent by the receiving entity to sending entity [14]. Our algorithm works on measuring fraction of INVITE messages to be rejected on both upstream and downstream SIP servers described in [7]. In our model we have assumed that cost of forwarding one INVITE message is unitary whereas rejecting one INVITE message is $1/\beta(\beta > 1)$. Let us consider $\lambda(t)$ is incoming INVITE load and $\alpha(t)$ is the fraction of INVITE messages to be rejected. O(t) is the instantaneous processor occupancy given by

$$O(t) = (1 - \alpha(t))\,\lambda(t) + (1/\beta)\,\alpha(t)\lambda(t) + d(t) \tag{1}$$

The first additive term caters for the cost of accepting INVITE calls and second one covers the cost for rejecting calls and the last additive term covers the cost due to disturbance like OS system calls overhead [7].

5 Experimental Test Bed

To evaluate the performance of our cooperative overload control algorithm, we have used experimental test bed consisting of three PC's. Hardware used in the test bed is shown in the Fig. 3. Individual components of the test bed are:

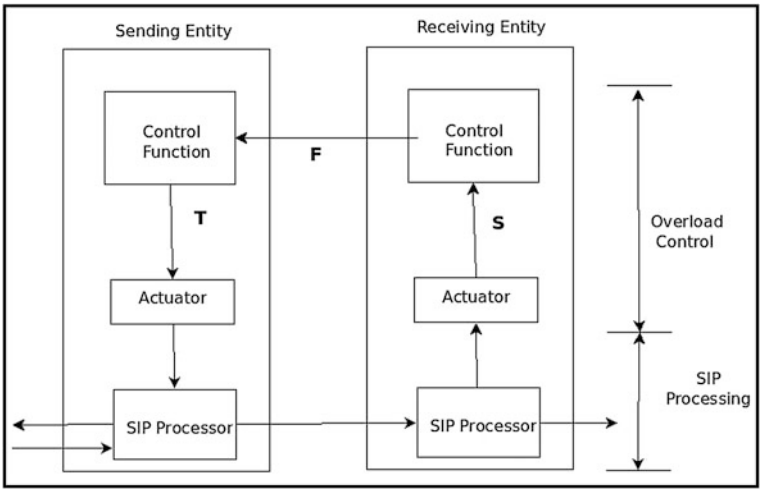


Fig. 2 Rate based overload control model

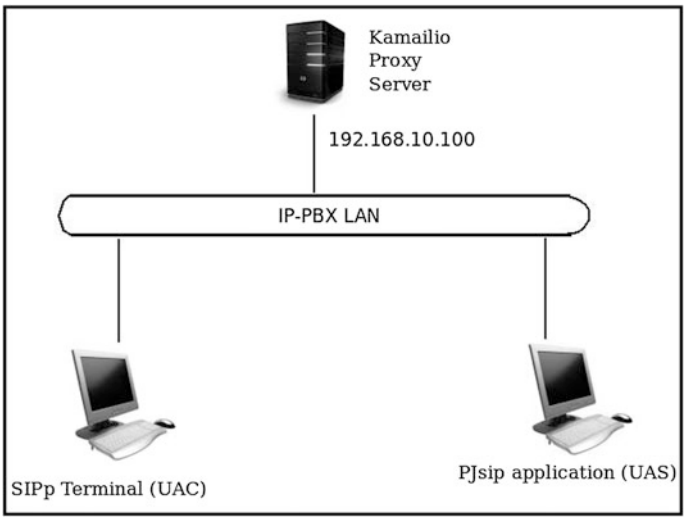


Fig. 3 SIP experimental test bed

5.1 Load Generator

To generate load we have used one of the SIP performance test tool, SIPp which is an open source platform. We have used subversion 3.3 of SIPp [15]. Using SIPp software we can emulate UAC or USA depending on the XML framework. We emulate UAC defining our own custom XML scenario to define timers for enumerating statistics to evaluate response times. We started SIPp with increasing the call rate 1 cps for every 5 s and stopped the SIPp once the maximum limit of 20,000 calls reached.

5.2 Kamailio Proxy Server

Kamailio is a leading open source SIP server used in many VOIP installations. It is evolved from Open SIP Express Router Project [14]. It can be configured as Proxy, Registrar, Presence and Application Server. Kamailio have a basic core which implements SIP protocol and does memory management. Kamailio configuration file provides a way to handle different SIP messages. The default configuration file provided by Kamailio can be used for Proxy Server. We have modified rate control module of Kamailio to implement our overload control method. Rate control module have a feature to define queue to hold incoming SIP requests. We have used two queues one for incoming INVITE messages and another queue for holding other SIP requests.

5.3 User Application Server

To terminate the calls we emulate UAS developed by using PJSIP stack library. UAS continuously send 200 OK requests in-order to close the loop. INVITE and BYE transactions are successful once they receive 200 OK messages from the UAS.

6 Results

In order to measure Invite response time and bye response time we have used a customized XML scenario file in SIPp. Whenever INVITE request comes timer 1 started and it is stopped when corresponding 200 OK was received similarly timer2 measure the BYE response time. We have gathered the statistics by enabling—trace_stat and—trace_ rtt options of SIPp. We had generated the SIP calls by gradually increasing the call rate until 400 cps reached for SIP proxy alone and with rate-based overload control. Invite response time and bye response time graphs shown in Figs. 4 and 5 clearly indicate that rate-based overload control method have 30 % improvement in response time. In Tables 1 and 2 summarized the critical parameters call set-up delay, throughput, INVITE retransmission rate, BYE retransmission rate for SIP proxy alone and rate-based overload control. For each run, SIPp instance was executed with fixed call rate and Packets are captured using Wireshark. From the table we can conclude that the cooperated overload control method definitely increased the throughput and reduce the call setup delay. Retransmission rates of INVITE and BYE are less in rate-based overload control.

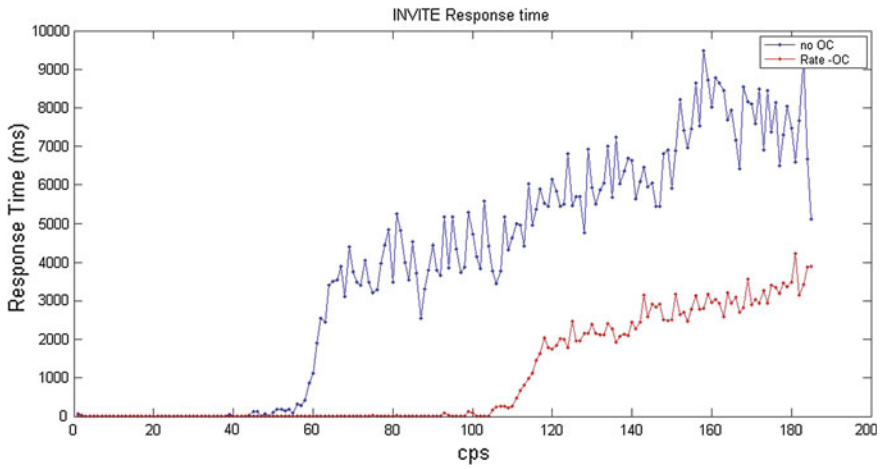


Fig. 4 NVITE response time

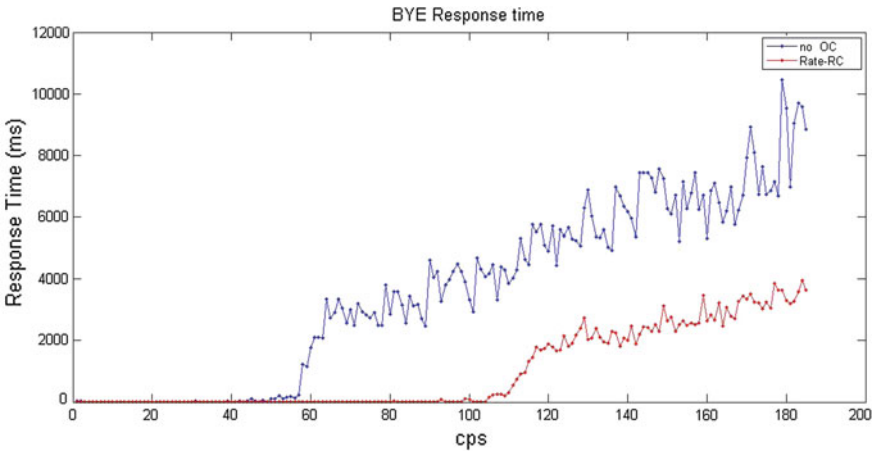


Fig. 5 BYE response time

Table 1 Results without overload control

Parameter	100 cps	300 cps	500 cps	700 cps	900 cps	1100 cps
Call setup delay (msec)	533	3481	5690	5822	6235	6492
Throughput (cps)	100	290	445	631	729	885
Retransmission rate (rps)-INVITE Msg	0	184	452	852	1306	2110
Retransmission rate (rps)-BYE msg	24	290	592	760	1056	1857

Table 2 Results rate-based overload control

Parameter	100 cps	300 cps	500 cps	700 cps	900 cps	1100 cps
Call setup delay (msec)	136	1288	1935	3168	4470	5099
Throughput (cps)	100	300	500	700	899	1053
Retransmission rate (rps)-INVITE Msg	0	0	18	87	383	556
Retransmission rate (rps)-BYE msg	12	158	460	798	1080	1782

7 Conclusion

We have proposed a rate-based overload control algorithm for SIP proxy server. The algorithm is a simple rate-based congestion control method implemented in Kamailio, a leading open source SIP Proxy server. The results have shown that the proposed overload control algorithm dramatically increase the throughput and reduce the call set-up delay and retransmission rate.

References

1. Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., Schooler, E.: SIP: Session initiation protocol, Internet Engineering Task Force, RFC 3261 (2002)
2. 3rd Generation Partnership Project. http://www.3gpp.org
3. Hong, Y., Huang, C., Yan, J.: A comparative study of SIP overload control algorithms. Network and Traffic Engineering in Emerging Distributed Computing Applications, IGI Global, pp. 1–20 (2012)
4. Singh, K., Schulzrinne, H.: Failover and load sharing in SIP telephony. J. Comput. Commun. **35**(5), 927–942 (2007)
5. Wright, C., Nahum, E.M., Wood, D., Tracey, J., Hu, E.: SIP server performance on multicore systems. IBM J. Res. Dev. **54**(1) (2010)
6. Jiang, H., Iyengar, A., Nahum, E., Segmuller, W., Tantawi, A.N., Wright, C.P.: Design, implementation, and performance of a load balancer for SIP server clusters. IEEE/ACM Trans. Netw. **20**(4), 1190–1202 (2012)
7. De Cicco, L., Cofano, G., Mascolo, S.:Local SIP overload control: controller design and optimization by extremum seeking. IEEE 52nd Annual Conference on Decision and Control (CDC), pp. 3281–3286, 10–13 Dec 2013 (2013)
8. Liao, J., Wang, J., Lib, T., Wang, J., Wang, J., Zhu, X.: A distributed end-to-end overload control mechanism for networks of SIP servers. Sci. Direct, Comput. Netw. **56**(12), 2847–2868 (2012)
9. Montazerolghaem, A., Yaghmaee, M.H.: SIP overload control testbed design, building and evaluation. Int. J. Ambient Syst. Appl. (IJASA) **1**(2), 17–26 (2013)
10. De Cicco, L., Carlucci, G., Mascolo, S.:Experimental Investigation of the Google Congesion Control for Real-Time Flows, FhMN '13 Aug 16 (2013)
11. De Cicco, L.: Skype video congestion control: an experimental investigation. Comput. Netw. **55**, 558–571 (2011)
12. Zhu, X., Pan, R.: NADA: A unified congestion control scheme for real-time media. IETF Draft (2013)
13. Hilt, V. (ed.): Design considerations for session initiated protocol (SIP) overload control. IETF Draft (2009)
14. Rosenberg, J.: Requirements for management of overload in the session initiation protocol. IETF RFC 5390 (2008)
15. Gayraud, R., Jacques, O.: SIPp. http://sipp.sourceforge.net

Discovery of Multi-frequent Patterns Using Directed Graph

Kuldeep Singh, Harish Kumar Shakya and Bhaskar Biswas

Abstract In this paper, an algorithm is proposed for mining frequent maximal itemsets. Discovering frequent itemsets is the key process in association rule mining. One of the major drawbacks of traditional algorithms is that they work only for items with single frequency per transaction. Proposed algorithm works with multiple frequency of an item per transaction. Proposed algorithm scans base database only once. The proposed algorithm took lesser time to find candidate itemsets. It uses directed graph for candidate itemsets generation. We conducted experiments on two datasets, Mushroom and Chess. Experimental results showed that proposed algorithm can quickly discover maximal frequent itemsets and effectively mine potential association rules.

Keywords Association rule · Frequent itemsets · Support count · Directed graph · Maximal itemsets

1 Introduction

The Association rule mining has been a major and basic task in the data mining. Association rule mining finds interesting association or correlation relationship among a large set of items. It was first presented by Agrawal [1]. The mining of association rules is typically done in two steps: first, find all the frequent itemsets and second step is generating association rule base on frequent itemsets [2].

Many algorithms have been developed for searching association rules. The main challenge in mining association rules is developing fast and efficient algorithm that can handle large volume of data, minimum time scans database and find associated rule very quickly. Most of the proposed apriori-like algorithms for mining association rules are wasting lots of time to generate candidate itemsets. FP Growth algorithm is also very useful for finding association rule. It is faster than apriori

K. Singh (✉) · H.K. Shakya · B. Biswas
Department of CSE, IIT-BHU, Varanasi 221005, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_15

algorithm and consumes lesser memory. The FP Growth algorithm does not generate candidate itemsets so take less time to find frequent itemsets [3]. But it also have limitation in respect of space and time. The proposed algorithms overcome these limitations. The proposed algorithm mining association rule with minimum candidate itemsets and items with multiple frequency of an item per transaction. Frequent itemsets can be of two types closed and maximal. A frequent itemset is called maximal if it has no superset that is frequent [4]. An itemset is closed if it is frequent but none of its superset has the same support count [5].

Directed graph is used in the proposed algorithm for candidate itemsets generation. It saves time and memory consumption because it generates minimum number of candidate itemsets, those are likely to be frequent. Frequent 2-itemsets required for the directed graph generation. First item of a frequent 2-itemset works as origin and second item like destination for an edge in the directed graph [6]. Directed graph gives us minimum candidate itemsets that are likely to be frequent itemsets.

Presently association rule mining algorithms is enhanced using bottom-up techniques and feature discovery based techniques. PIETM (Principle of Inclusion–Exclusion and Transaction Mapping) discovers frequent itemsets by using the bottom-up technique. It reduces database scanning only two times [7]. PIETM does not scan the database to count the support of itemsets. Instead, it presents a new method that calculates the support of itemsets using a well-known set theory property. It employs the Principle of Inclusion–Exclusion to calculate the support of candidate itemsets. Improved Apriori Algorithm Based on Features discovers more reasonable association rules [8]. The algorithm shows that different items should be given different importance. In this every transaction item has its own feature(s) to carry more information. With adding feature(s) to these items, when mining the association rules, just those transaction data with same feature(s) will be scanned and computed. The algorithm in which features are added to the transaction items to make the mining association rules more reasonable and more efficient. But in many conditions the algorithm need to repeat the calculation.

2 An Algorithm Based on Directed Graph

2.1 Definitions

Definition 1 Let $I = \{I_1, I_2, \ldots, I_n\}$ be a set of items. Let D be a database of transactions. And $D = \{T_1, T_2, \ldots, T_m\}$, Each transaction T is a set of items

Definition 2 Let m be the number of transaction in D. And n be the number of distinct items in the D

Definition 3 Item has their own item_freq. item_freq is the occurrence of an itemset in a transaction

Definition 4 A matrix (m x n) with elements value corresponding to item_freq of the items

Definition 5 Let X be an itemset, itemset X contains k items. The set of all frequent k-itemset is denoted by L_k (Fig. 1)

Input: Matrix (m x n), min_sup;
Output: Set of Maximal frequent itemset FS.
1. Find total_item_freq of all the items by adding frequency of an item in all transactions.
2. If total_item_freq <=min_sup than
3. remove those items.
4. Otherwise add 1-itemsets into L_1
5. sort the items in ascending order according to their total_item_freq.
6. Generate candidate 2-itemsets using lexicographical order with L_1.
7. Calculate total_item_freq of candidate 2-itemsets. Using math function min(item_freq(I_i), item_freq (I_j))
8. If total_item_freq < min_sup than
9. remove those itemsets.
10. Otherwise add 2-itemsets into L_2.
11. Generate directed graph using L_2.
12. If (I_i,I_j) in L_2 than draw an directed edge from I_i to I_j in directed graph.
13. Traverse the directed graph and generate candidate k-itemsets. Traverse I_i to all the directed paths.
14. calculate total_item_freq of k-itemsets using min(item_freq(I_i), item_freq (I_j),... item_freq (I_k))
15. If total_item_freq <=min_sup than
16. remove those itemsets
17. otherwise add k-itemsets into L_k.
18. Find maximal frequent itemset form L_k.
19. return Maximal frequent itemsets.

Fig. 1 The proposed algorithm

2.2 Algorithm Details

The proposed algorithm can be divide into fives main steps. These steps explained below:

2.2.1 Transform the Transactional Database D into a Matrix

The transactional database transformed into an asymmetric $matrix_{m*n}$. Total number of distinct items are the total number of columns (n) in the matrix. Total number of transaction T are the rows (m) in the matrix as shown in the Fig. 2. Item_freq of the items are value of corresponding column of that particular transaction. If the transaction does not have an item, then 0 (zero) inserted into corresponding column location. After the transformation step a transactional database became an asymmetric matrix [9].

We denote $Matrix_{m*n}$:

m = Total number of transactions &
n = Total number of distinct items in transaction database D.

Fig. 2 The frequency matrix$_{7*6}$

	I_1	I_2	I_3	I_4	I_5	I_6
T_1	3	2	0	0	4	2
T_2	1	0	0	1	3	3
T_3	0	0	5	2	4	0
T_4	0	1	4	0	1	4
T_5	0	0	3	0	0	3
T_6	4	3	0	1	2	2
T_7	2	2	0	0	3	1

2.2.2 Generate the Set of Frequent 1-Itemsets L_1

We scan the Matrix and calculate total_item_freq of items by counting item_freq of items in a column. If total_item_freq of I_j item lesser than min_sup, itemset I_j in not a frequent 1-itemset, so remove jth column. Otherwise itemset I_j is the frequent 1-itemset and is added to the set of frequent 1-itemset L_1. Sorting has been done of all frequent 1-itemsets by total_item_freq in ascending order. Sorted elements help us to generate minimum candidate itemsets as in [10], these candidate itemsets are likely to be frequent.

2.2.3 Generate the Set of Frequent 2-Itemsets

In this step, we use sorted 1-itemsets L_1 and generated candidate 2-itemsets in lexicographical order. Count the total_ item_freq of the candidate 2-itemset $\{I_i,I_j\}$. Minimum item_freq of I_i and I_j are the total_item_freq of the candidate 2-itemset. If the total_item_freq of candidate 2-itemset $\{I_i,I_j\}$ greater than min_sup, then itemset $\{I_i,I_j\}$ frequent 2-itemset and added into L_2, otherwise removed.

2.2.4 Construct the Directed Graph

We construct the directed graph by using frequent 2-itemsets L_2. If itemset $\{I_i,I_j\}$ frequent 2-itemset then draw a directed edge from I_i to I_j in directed graph. First item of 2-itemset are origin node and second item are destination node. We draw all directed edge by using itemset of L_2. After drawn all the edges, we present a directed graph [11, 12].

2.2.5 Generate All the Candidate k-Itemsets Using Directed Graph

The proposed algorithm traverses once to the directed graph and generate candidate itemsets by using the directed neighbour nodes. Start from I_i and traverse all the possible reachable nodes. After completion of traversing, it leads to formation

simple paths; these simple paths are our candidate itemsets. Nodes of possible simple paths are element of these itemsets.

2.2.6 Find Frequent Itemsets

In final step, we collect only frequent itemsets. For finding frequent k-itemsets, we traversed the matrix and count the total_item_freq of the candidate k-itemsets $\{I_i,I_j, \ldots I_k\}$. Minimum item_freq of I_i, I_j up to I_k are the total_item_freq of the candidate k-itemsets. If the total_item_freq of candidate k-itemsets $\{I_i,I_j,\ldots I_k\}$ greater than min_sup, than $\{I_i,I_j,\ldots I_k\}$ itemsets are frequent k-itemset and added into L_k. Otherwise remove infrequent itemsets. Finally we got the final set of frequent k-itemsets L_k. Now we verify that frequent k-itemsets L_k are maximal or not. If those itemsets who do not full fill the property of maximal itemsets are removed, otherwise add into final resulted sets.

2.3 Example

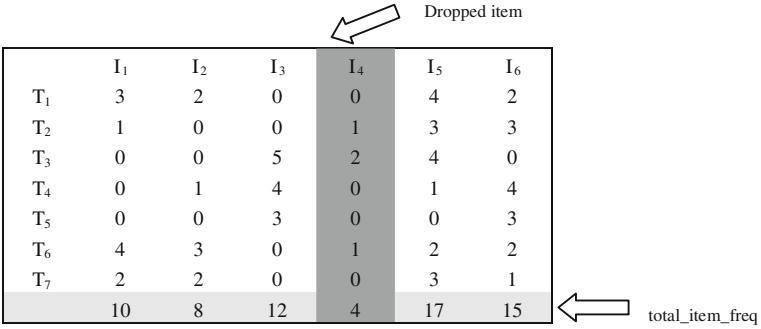
The transactional database showed in Table 1. The number of items is 6 and the number of transactions is 7. The minimum support count (min_sup) is 5.

Table 1 Transactional database

T_{ID}	List of items with their counts
T_1	I_1-3, I_2-2, I_5-4, I_6-2,
T_2	I_1-1, I_4-1, I_5-3, I_6-3
T_3	I_3-5, I_4-2, I_5-4
T_4	I_2-1, I_3-4, I_5-1, I_6-4
T_5	I_3-3, I_6-3
T_6	I_1-4, I_2-3, I_4-1, I_5-2, I_6-2
T_7	I_1-2, I_2-2, I_5-3, I_6-1

2.3.1 Transform the Transactional Database D into a Matrix

The transactional database transformed into an asymmetric matrix$_{7*6}$. The items are the columns of the matrix and transactions are the rows of the matrix as shown in the Fig. 2. Item_freq of the items are values of corresponding column in particular transaction. In T_1 transaction I_1,I_2,I_5 and I_6 having item_freq 3,2,4 and 2 respectively inserted in first row of the matrix as shown in Fig. 2. All other rows inserted in the same manner as T_1 row was inserted.



	I_1	I_2	I_3	I_4	I_5	I_6
T_1	3	2	0	0	4	2
T_2	1	0	0	1	3	3
T_3	0	0	5	2	4	0
T_4	0	1	4	0	1	4
T_5	0	0	3	0	0	3
T_6	4	3	0	1	2	2
T_7	2	2	0	0	3	1
	10	8	12	4	17	15

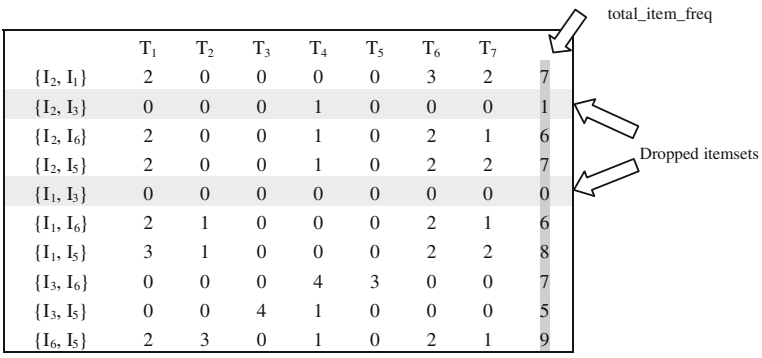
Fig. 3 The frequency matrix with total_item_count

2.3.2 Generate the Set of Frequent 1-Itemsets L_1

In this step, we scan the frequency Matrix and calculate total_item_freq of items by counting values of the columns. The total_item_freq of items are $I_1 = 10$, $I_2 = 8$, $I_3 = 12$, $I_4 = 4$, $I_5 = 17$ and $I_6 = 15$. Total_item_freq of I_4 smaller than min_sup, so removed as shown in Fig. 3. I_1, I_2, I_3, I_5 and I_6 itemsets are frequent. Sort the frequent 1-itemsets according to their total_item_freq and add these into L_1. Now $L_1 = \{\{I_2\}, \{I_1\}, \{I_3\}, \{I_6\}, \{I_5\}\}$

2.3.3 Generate Candidate 2-Itemsets

We generated candidate 2-itemsets using sorted L_1. Count the total_ item_freq of the candidate 2-itemset $\{I_i,I_j\}$. Minimum item_freq of I_i and I_j are the total_item_freq of the candidate 2-itemset. The total_item_freq of itemsets are $\{I_2,I_1\} = 7$, $\{I_2,I_3\} = 1$, $\{I_2,I_6\} = 6$, $\{I_2,I_5\} = 7$, $\{I_1,I_3\} = 0$, $\{I_1,I_6\} = 6$, $\{I_1,I_5\} = 8$, $\{I_3,I_6\} = 7$, $\{I_3,I_5\} = 5$ and $\{I_6,I_5\} = 9$ as shown in Fig. 4. In itemset $\{I_2,I_1\}$, item_freq are 2, 3 and 2 in



	T_1	T_2	T_3	T_4	T_5	T_6	T_7	
$\{I_2, I_1\}$	2	0	0	0	0	3	2	7
$\{I_2, I_3\}$	0	0	0	1	0	0	0	1
$\{I_2, I_6\}$	2	0	0	1	0	2	1	6
$\{I_2, I_5\}$	2	0	0	1	0	2	2	7
$\{I_1, I_3\}$	0	0	0	0	0	0	0	0
$\{I_1, I_6\}$	2	1	0	0	0	2	1	6
$\{I_1, I_5\}$	3	1	0	0	0	2	2	8
$\{I_3, I_6\}$	0	0	0	4	3	0	0	7
$\{I_3, I_5\}$	0	0	4	1	0	0	0	5
$\{I_6, I_5\}$	2	3	0	1	0	2	1	9

Fig. 4 The candidate 2-itemsets matrix$_{10*7}$

transactions T_1, T_6 and T_7 respectively as shown in the Fig. 4, and total_item_freq is 7. Itemset $\{I_2,I_3\}$ and $\{I_1,I_3\}$ are infrequent, so removed. Now the frequent 2-itemsets are $L_2 = \{\{I_2,I_1\},\{I_2,I_6\},\{I_2,I_5\}, \{I_1,I_6\}, \{I_1,I_5\}, \{I_3,I_6\}, \{I_3,I_5\}, \{I_6,I_5\}\}$.

2.3.4 Generate Directed Graph Using Frequent 2-Itemsets

Construct the directed graph using L_2. If $\{I_i,I_j\}$ in L_2, then draw a directed edge from I_i to I_j in directed graph. A graph consists of two things node (V) and directed edged (E). V is the set of nodes and E is the set of edges of a graph. V = $\{I_1, I_2, I_3, I_5, I_6\}$ or $\{L_1\}$ and E = $\{I_2,I_1\},\{I_2,I_6\},\{I_2,I_5\}, \{I_1,I_6\}, \{I_1,I_5\}, \{I_3,I_6\}, \{I_3,I_5\}, \{I_6, I_5\}\}$ or $\{L_2\}$. For itemset $\{I_2,I_1\}$, I_2 is the origin point and I_1 is the destination point of an edge. A graph having the number of edges is equal to the number of itemsets in L_2. Completed directed graph is shown in Fig. 5.

2.3.5 Generate Candidate k-Itemsets Using Directed Graph

Candidate k-itemsets generated by traversal of directed graph. This process is very easy and simple as shown in Fig. 6. Start from each source node and traverse to all reachable nodes. This process gave us all candidate itemsets. Our candidate k-itemsets are $\{I_2,I_1,I_6\}$, $\{I_2,I_1,I_6,I_5\}$, $\{I_2,I_1,I_5\}$, $\{I_2,I_6,I_5\}$, $\{I_1,I_6,I_5\}$, $\{I_3,I_6,I_5\}$.

2.3.6 Find Frequent Itemsets

In last, we check whether each candidate k-itemsets is frequent or infrequent itemsets. Calculate total_item_freq of all the itemset as shown in the Tables 2 and 3. Suppose we calculate total_item_freq of itemset $\{I_2,I_1,I_6\}$, now we calculate minimum item_freq value among I_2,I_1 and I_6 for all the transaction, and add all these minimum item_freq as shown in Table 2. Minimum item_freq for T_1, T_2, T_3, T_4, T_5, T_6 and T_7 are 2, 0, 0, 0, 0, 2, and 1 respectively, and total_item_freq is 5 as shown in Table 2. Now the total_item_freq of candidate k-itemsets are $\{I_2,I_1,I_6\} = 5$, $\{I_2,I_1,I_6, I_5\} = 5$,$\{I_2,I_1,I_5\} = 6$, $\{I_2,I_6,I_5\} = 6$, $\{I_1,I_6,I_5\} = 6$, $\{I_3,I_6,I_5\} = 1$. The total_item_freq of $\{I_3,I_6,I_5\}$ itemset is less than min_sup, So we remove this itemset as shown in

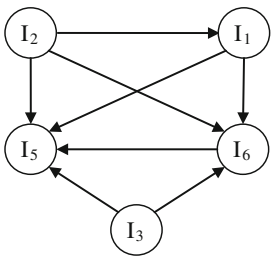


Fig. 5 Directed graph

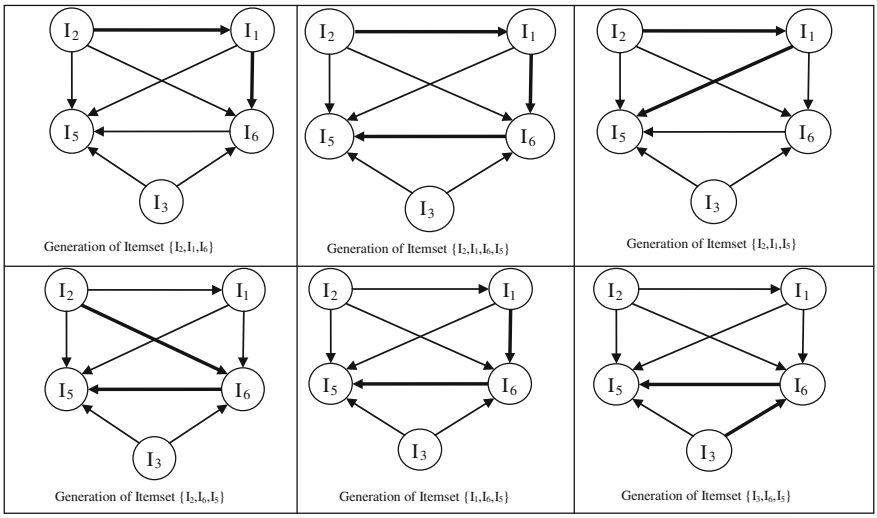


Fig. 6 Generation of candidate k-itemsets

Table 2 Calculation of item_freq using mathematical function minimum

	$\{I_2,I_1,I_6\}$	$\{I_2,I_1,I_6,I_5\}$	$\{I_2,I_1,I_5\}$	$\{I_2,I_6,I_5\}$	$\{I_1,I_6,I_5\}$	$\{I_3,I_6,I_5\}$
T_1	min(2,3,2)	min(2,3,2,4)	min(2,3,4)	min(2,2,4)	min(3,2,4)	min(0,2,4)
T_2	min(0,1,3)	min(0,1,3,3)	min(0,1,3)	min(0,3,3)	min(1,3,3)	min(0,3,3)
T_3	min(0,0,0)	min(0,0,0,4)	min(0,0,4)	min(0,0,4)	min(0,0,4)	min(5,0,4)
T_4	min(1,0,4)	min(1,0,4,1)	min(1,0,1)	min(1,4,1)	min(0,4,1)	min(4,4,1)
T_5	min(0,0,3)	min(0,0,3,0)	min(0,0,0)	min(0,3,0)	min(0,3,0)	min(3,3,0)
T_6	min(3,4,2)	min(3,4,2,2)	min(3,4,2)	min(3,2,2)	min(4,2,2)	min(0,2,2)
T_7	min(2,2,1)	min(2,2,1,3)	min(2,2,3)	min(2,1,3)	min(2,1,3)	min(0,1,3)

Table 3. All other itemsets is having greater total_item_freq than min_sup, so these are added into frequent k-itemset L_k as shown in the Table 4.

In the end, we find maximal frequent itemsets from L_k. $\{I_2,I_1,I_6\}$,$\{I_2,I_1,I_5\}$ $\{I_2,I_6,I_5\}$, $\{I_1,I_6,I_5\}$ are the subsets of the itemset $\{I_2,I_1,I_6,I_5\}$. In maximal itemset we did

Table 3 Calculte total_item_freq

Dropped itemset

	$\{I_2,I_1,I_6\}$	$\{I_2,I_1,I_6,I_5\}$	$\{I_2,I_1,I_5\}$	$\{I_2,I_6,I_5\}$	$\{I_1,I_6,I_5\}$	$\{I_3,I_6,I_5\}$
T_1	2	2	2	2	2	0
T_2	0	0	0	0	1	0
T_3	0	0	0	0	0	0
T_4	0	0	0	1	0	1
T_5	0	0	0	0	0	0
T_6	2	2	2	2	2	0
T_7	1	1	2	1	1	0
	5	5	6	6	6	1

total_item_freq

Table 4 Frequent k-itemsets

Candidate sets	Support count
$\{I_2,I_1,I_6\}$	5
$\{I_2,I_1,I_6,I_5\}$	5
$\{I_2,I_1,I_5\}$	6
$\{I_2,I_6,I_5\}$	6
$\{I_1,I_6,I_5\}$	6

not include frequent subsets. So remove the subsets of the maximal itemsets. For the example our final maximal frequent itemset is $\{I_2,I_1,I_6,I_5\}$.

3 Experimental Results

Experiment results showed that proposed algorithm is simple and efficient. To find experimental results, we have used mushroom and chess dataset obtained from UCI [13] and synthetic dataset which are created by us. The algorithms were implemented in Java and tested on a windows platform. Mushroom dataset have total Number of Instances: 8124, Number of Attributes: 23, Number of items: 119. Another dataset used for experiment result is chess [14]. Chess dataset have total Number of Instances: 3196, Number of Attributes: 37, Number of items: 75. We have divided the mushroom and chess datasets into groups of 50 instances. Hence, each instance contains their corresponding count. We analysis that as support count increases, the execution time goes down as shown in the Fig. 7.

Proposed algorithm took less time when minimum support count of threshold was higher. The frequent pattern generated by traditional algorithms and proposed algorithm are almost same. But proposed algorithm gives more information about item frequency. This item frequency can used to find, how many items sold together with some item frequency. This additional information can help us to determine the number of items sold with associated items. Proposed algorithm took less time as compare traditional algorithms to find frequent patterns.

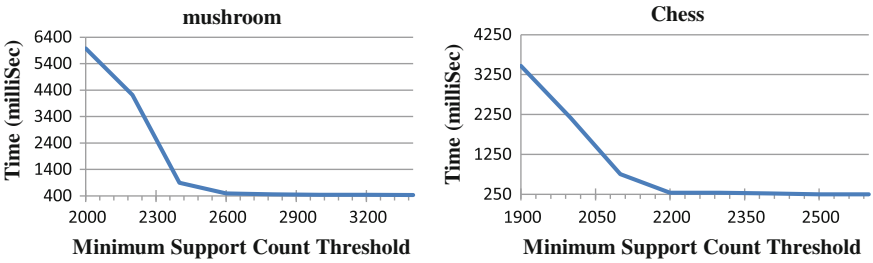


Fig. 7 Execution time with different minimum support count thresholds for mushroom and chess datasets

4 Conclusion

The proposed algorithm use directed graph for generate candidate itemsets. The candidate itemsets are likely to be frequent, so provide efficiency and accuracy in results. So the proposed algorithm spends less time to find candidate and maximal frequent itemsets. Directed graph generates candidate itemsets by graph traversing instead of join and prune steps. The proposed algorithm scan base database only once. So the proposed algorithm suitable for find frequent itemsets when items have multiple counts or frequency. Not much difference in patterns was with proposed and traditional algorithms but the execution time of proposed algorithm is lesser than traditional algorithms.

References

1. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: Proceedings of the 20th International Conference on Very Large Databases, pp. 487–499. Santiago, Chile (1994)
2. Han, J., Kamber, M.: Data Mining Concepts and Techniques, 3rd edn. Morgan Kaufmann Publishers, San Francisco (2012)
3. Han, J., Pei, J., Yin, Y.: Mining Frequent Patterns without Candidate Generation. In: Proceedings of the ACM-SIGMOD conference management of data, pp. 1–12 (2000)
4. Amir, A., Anumann, Y., Feldman, R., Fresko, M.: Maximal association rules: a tool for mining associations in text. J. Intell. Inf. Sys. **25**(3), 333–345 (2005)
5. Zaki, M.K.: Closed itemset mining and non-redundant association rule mining. In: Encyclopedia of Database Systems. Springer-Verlag (2009)
6. Rosen, K. H.: Discrete mathematics and its application. In: 7th Edition, McGraw-Hill publication (2012)
7. Lin, K.C., Liao, I.E., Chang, T.P., Lin, S.F.: A frequent itemset mining algorithm based on the principle of inclusion-exclusion and transaction mapping. Inf. Sci. **276**, 278–289 (2014)
8. Yang. J., Li.Z., Xiang.W., Xiao.L.: An improved apriori algorithm based on features. In: Ninth International Conference on Computational Intelligence and Security, pp. 125–128 (2013)
9. Jain, G., Bingwu, L., Mingru, Z., Pingping, D.: A new method of computing the reachability matrix of directed graph. In: Second International Workshop on Education Technology and Computer Science, pp. 774–776 (2010)
10. Liu, H., Wang, B.: An Association Rule Mining Algorithm Based on a Boolean Matrix. Data Sci. J. **6**, S559–S565 (2007)
11. Liu, N., Ma, L.: Discovering frequent itemsets an improved algorithm of directed graph and array. In: 4th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 1017–1020 (2013)
12. Fang, L.: Frequent itemset mining algorithm based on graph and two-directional search. Comput. Eng. **38**(1), 59–61 (2012)
13. Schlimmer J., https://archive.ics.uci.edu/ml/machine-learning-databases/mushroom/agaricus-lepiota.data
14. Shapiro, A., https://archive.ics.uci.edu/ml/datasets/Chess (KRKPA7)

Efficient Identification of Users and User Sessions from Web Log Repository Using Dimensionality Reduction Techniques and Combined Methodologies

G. Shivaprasad, N.V. Subba Reddy, U. Dinesh Acharya and Prakash K. Aithal

Abstract Web Based Applications are data intensive. In addition to web content and structure, they collect huge amount of data in the form of User interactions with the web, forming Web Log Repository. Application of data mining techniques over the Web Log Repository to extract useful knowledge is referred to as Web Usage Mining. Web Usage Mining consists of three phases—Web Log Preprocessing, Knowledge Discovery and Pattern Analysis. In this paper, an efficient implementation for Web Log Pre-processing using Dimensionality Reduction Techniques and Combined Methodologies is presented.

Keywords Web usage mining · Web log repository · Web log Pre-processing

1 Introduction

Web Usage Mining is the process of extracting useful information from Web Log Repository by the application of Data Mining technique. Extracted patterns represent user browsing behaviors. Accurate analysis of these patterns leads to understanding of users visiting the web site thereby improved user satisfaction. Improved customer satisfaction is the key to success of business. Thus, Web Based Applications can improve their business by the application of Web Usage Mining to Web Log Repository. Targeted Marketing, Location Based Marketing, Web Personalization, Fraud Detection and Improved Web Administration are some of the application areas of Web Usage Mining.

G. Shivaprasad (✉) · N.V. Subba Reddy · U. Dinesh Acharya · P.K. Aithal
Department of CSE, Manipal Institute of Technology, Manipal University, Manipal 576104, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_16

Web Usage mining consists of three main steps: Web Log Preprocessing, Knowledge Discovery and Pattern Analysis. Among these tasks, Web Log Preprocessing is most complex and critical for the efficient extraction of useful patterns. Especially, the Web Log Cleaning is more demanding in order to eliminate noisy and irrelevant data and to make the Log Data suitable for Knowledge Discovery. Also the Web Log is memory intensive and pruning irrelevant data reduces the input load of the Knowledge Discovery phase. This paper presents dimensionality reduction techniques to eliminate the noisy data and combined methodologies to efficiently identify users and user sessions from Web Log Repository.

The paper is organized as follows: Sect. 2 presents a brief literature review, Sect. 3 presents the structure of Web Log, and Sect. 4 presents the Web Log Preprocessor, Sect. 5 presents the Experimental Setup and Result analysis and finally, Sect. 6 presents the conclusion.

2 Literature Review

The method of extracting useful information from server log files and different application areas of Web Usage Mining is presented in [1]. A framework for Web Usage Mining consisting of Preprocessing, Pattern Discovery and Users classification, is proposed in [2]. This framework classifies the users based on country, site entry and access time. Information extraction from user navigation history using Web Usage Mining is explored and discussed in [3, 4]. A detailed survey on data collection and pre-processing stage of web usage mining is discussed in [5]. Several data preparation techniques of access stream to identify the unique sessions and unique users are presented in [6]. Educational data mining techniques to analyze learners' behavior, to help in learning evaluation and to enhance the structure of a given course is implemented in [7]. A new algorithm for preprocessing and clustering of web log is proposed in [8]. A specific methodology to extract useful information from an e-commerce website is proposed in [9]. A critical analysis and comparison of the common web robot detection approaches is presented in [10].

3 The Web Log Repository

Web Log Repository is a pool of user activities on a web site. When activated by the web site administrator, it automatically collects the user navigation activities on the web, the moment he enters the web site till the moment he leaves the web site. In Extended Common Log Format (ECLF), a web log usually contains entries with regard to Host IP Address, User Authentication, Date and Time of visit, HTTP Request, Referrer Field and User Agent Field. Details with reference to each field are given below:

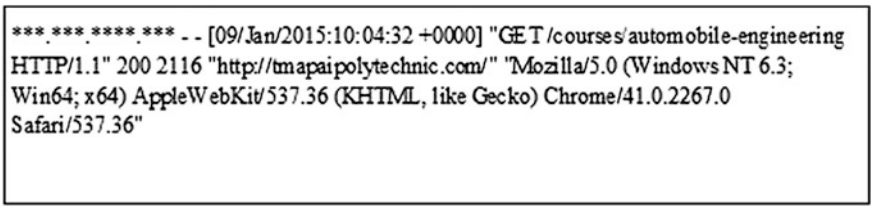


Fig. 1 Extract of experimental data

- *Host IP address*—Used to identify the user visiting the web site.
- *User Authentication*—Contains the Username and password of the user visiting web site, usually empty due freeform of websites.
- *Date and time of the visit*—Tells when the user has visited the web site
- *HTTP Request*—Represents collective information—the Request Method (GET, POST, HEAD, etc.), the Requested Resource (a HTML page, an Image, a CGI program, or a script, etc.) and the Protocol Version (HTTP protocol being used along with version number).
- *Request Status*—Status of the request (200 Series—Successful Transmission, 400 Series- Client Error, etc.).
- *Page size*—Size of the document downloaded in Bytes.
- *Referring Agent (RA)*—Gives the details of the web site from which the user has traversed to the web site. If the user has directly enters this website by typing the web site URL, this field will be "-".
- *User Agent (UA)*—Gives the details with regard to the browser and operating system of the client.

The web access log was collected from the web server of Dr. T.M.A. Pai Polytechnic, Manipal web site [11]. The web site hosts information about courses offered, admission details, facilities available and the placement details, etc. A sample of web log record is given in Fig. 1.

The above log entry indicates that user with IP Address ***.***.***.*** (masked here) requested the link automobile-engineering under courses on 9th Jan 2015 at 10:04:32 AM and he traversed from the link http://tmapaipolytechnic.com. The request was successful and a total of 2116 bytes have been downloaded. Also, it indicates that Mozilla (compatible) 5.0 was the browser and Windows NT 6.3 was the operating system used.

4 Web Log Preprocessor

Web Log Preprocessing plays an important role in Web Usage Mining. The data collected in Web Log Repository is not suitable for Data Mining algorithms. The Log Data needs to be cleansed and converted into structured format before being

processed by Knowledge Discovery Phase. Web Log Preprocessor takes Web Log Repository as input and identifies the Users and User Sessions. We begin Pre-processing phase by Feature extraction and Time Stamp Creation. It is then followed by Data Cleaning employing Dimensionality Reduction Techniques. The original Web Log Repository is blended with relevant and irrelevant information leading to huge log size. Direct processing of this raw data puts unnecessary burden on the Knowledge Discovery Phase. Hence, efficient cleaning of Web Log Repository is necessary to extract useful patterns from Web Log. Once the log is cleansed effectively, Users visiting the web site are identified. Then, User activities in the Web are grouped into meaningful sessions before being processed by Knowledge Discovery Phase. Thus, Web Log Preprocessor mainly contains—Feature Extraction and Time Stamp Computation, Data Cleaning, User Identification, and User Session Identification.

4.1 Feature Extraction and Timestamp Computation

In Feature Extraction step, features are extracted from fields representing collective information so that preprocessing algorithms can be applied. Also, from the date and time entries of web log, time stamp is computed so as to estimate the duration of the user's visit to the Web site and to maintain the sequence of web requests across days. The steps for creation of time stamp are as follows:

(i) Compute the number of days between the web log entry date and a reference date.
(ii) Multiply this number of days by 86,400.
(iii) Find the time in seconds since midnight that is represented by the time in the web log entry.
(iv) Add (ii) and (iii).

4.2 Web Log Cleaning

In Data Cleaning, all irrelevant entries from the log record are eliminated to minimize the burden on the processor. A web log usually contains all the requests to the web server. This includes actual user requests and automated requests. The automated request represents the requests from automated programs like web bots, spiders and crawlers. Similarly, when a user requests a page from the web server, along with the page, any images associated in the requested page is also downloaded and a record for each such image downloads is created in the log. As Web Usage Mining intends to model the user browsing patterns, all such requests need to be eliminated. Similarly, unsuccessful requests are also eliminated. Also, only the request for getting a resource from the web server is retained. Thus, Web Log cleaning consists of the following sub steps:

Fig. 2 IsRobotRequest algorithm

```
Input: Web Log-Record from Web Log Repository
Output: TRUE/FALSE
Algorithm IsRobotRequest
Begin
  For the current Log-Record
      Check whether the Requested Resource Field = 'robot.txt'
      If YES, return TRUE.
      Otherwise, check whether the User Agent Field contains
        words like 'crawler', 'spider', 'bot' or 'robot'.
      If YES, return TRUE. Otherwise, return FALSE.
End
```

Fig. 3 IsImageRequest algorithm

```
Input: Web Log-Record from Web Log Repository
Output: TRUE/FALSE
Algorithm IsImageRequest
Begin
  For the Log-Record
      Check whether the URI field contains file extensions
        like .JPG or .CSS or .png or .js
      If YES, return TRUE.
      Otherwise return FALSE.
 End
```

Robot Request Eliminator: Robot requests can be identified using 2 methods:

1. Robot identification based on Requested Page Field
2. Robot identification based on User Agent field.

For efficient identification of web robots, combined methods were employed. The algorithm IsRobotRequest takes as input each log record and returns TRUE/FALSE is given in Fig. 2.

Image Request Filter: The files with the extensions like GIF, JPEG, CSS are also downloaded along with requested page. They are not actually the user interested web page; rather it is just the documents embedded in the web page. So, it is not necessary to include in identifying the user interested web pages. So, the cleaning process eliminates these unnecessary entries from web logs by scanning the Uniform Resource Identifier (URI) field of every record. This step drastically reduces the size of web log. The algorithm for filtering out the Image Requests is given in Fig. 3. The algorithm checks each web log record and returns TRUE/FALSE.

Unsuccessful Request Remover: Successful web requests represent the user actual request to the web server using which the user profile can be modeled. Hence, log records with status codes other than 200 (successful request) are removed. This cleaning process will further reduce the evaluation time for

Fig. 4 IsSuccessfulRequest algorithm

```
Input: Web Log-Record
Output: TRUE/FALSE
Algorithm IsSuccessfulRequest
Begin
    For the Log-Record
            Check whether the Status Code = 200.
            If YES, return TRUE.
            Otherwise, return FALSE.
 End
```

Fig. 5 IsGetMethod algorithm

```
Input: Web Log-Record
Output: TRUE/FALSE
Algorithm IsGetMethod
Begin
    For the Log-Record
          Check whether the HTTP Request Method = GET.
          If YES, return TRUE.
          Otherwise, return FALSE.
 End
```

determining the user interested patterns. The algorithm for Removal of Unsuccessful HTTP requests is given in Fig. 4. The algorithm checks each record of web log and returns TRUE/FALSE.

nonGET Request Remover: A GET method in the HTTP Request Field indicates that the user has requested a resource from the web server. Hence, log records having the value of GET in the Method field of HTTP Request are retained, while all other records are eliminated. This step again reduces the volume of the data to be processed further. The algorithm for elimination of non GET methods from web log is given in Fig. 5. The algorithm takes each record and returns TRUE/FALSE.

The modified Data Cleaning algorithm using above algorithms is given below. The algorithm scans each log record and either retains or discards the record by calling the above algorithms.

```
Input: Web Log Repository obtained after Time Stamp Computation Method
Output: Cleaned Log File
WebLogCleaner Algorithm
Begin
While not eof (WebLogRepository) Do
    Read the next record of WebLogRepository into Log-Record
        If (NOT( IsRobot(Log-Record)) AND NOT(IsImageRequest(Log-Record)) AND IsSuccRequest(Log-Record)
        AND IsGetMethod(Log-Record))
                Write Log-Record to New-LogFile.
        End If
End While
End
```

4.3 Modified User Identification

Identification of each distinct user visiting the website is important and complex task in Web Usage Mining. Apart from the user-id field, the IP address, UA and RA fields can be employed for user identification. In this paper, User Identification based on combined methods using all the three fields has been implemented to uniquely identify the users. In UA field, both the browser and operating system are considered for distinguishing between two users. The Modified User Identification algorithm is given below:

```
Input: N records of cleaned web log file (New-LogFile)
Output: User set U
ModifiedUserIdentification Algorithm
Begin
      User-Count = 0
      While not eof (New-LogFile) Do
         Read the current record and next record of New-LogFile into Cur-Record and Next-Record.
         Let IP_cur, IP_next, UA_cur, UA_next, RU_cur and RU_next be the IP Address, UA and Referrer URL in Cur-Record and Next-Record.
         If IP_cur <> IP_next Identify both entries as belonging to different user, Increment the User-Count by 1.
         Else If UA_cur <> UA_next [Both browser and OS are unique]
                Identify both entries as belonging to different user; Increment the User-Count by 1.
                Else If RU_cur<>"-"and RU_next = "-"
                          Identify both entries as belonging to different user; Increment the User-Count by 1.
                    Else  Assume as same user.
                    End If
                End If
            End If
      End While
End
```

4.4 Modified User Session Identification

User Session identification is the process of segmenting the access log of each user into individual access sessions. For Session Identification heuristics based on Time and Navigation are employed. Time based methods are not reliable because users may involve in some other activities after opening the web page. Hence in this paper, a combined technique based on both the heuristics is employed for Session Identification. This method uses web topology and page stay time. The Session Identification algorithm is given below.

```
Input: User sets with N records
Output: Constructed Sessions SessionSet
ModifiedSessionIdentification Algorithm
Begin
    Let PageStayTime ←10 minutes
    SessionSet = { }
    K←J←0
    Let Lj, URIj, tj , RUj and Uj denote log entry, URI, time stamp, Referrer URL and user respectively.
    For each unique user Uj do
        For each Lj do
            If RUj = "-" and (tj- tj-1)>PageStayTime
                K←K+1; Sk ← URIj; SessionSet = SessionSet U Sk
            Else If RUj is Present in any Si of Uj, where i=1,2,3,…sessions of User Uj,  Si ←URIj
                Else  K←K+1;  Sk ← URIj; SessionSet = SessionSet U Sk
                End If
            End If
        End For
    End For
End
```

5 Experimental Setup and Results

The web access log was collected from the web server of Dr. T.M.A. Pai Polytechnic web site from 31st Dec 2014 12:09:56 through 11:18:07 15th Jan 2015, a total of 15 days. A total of 5817 requests were recorded during this period. The algorithms were implemented in MATLAB.

5.1 Web Log Cleaning

The Web Log Cleaning Algorithm was applied to the log data after feature extraction and time stamp computation. The algorithm eliminated a total of 4648 records containing multimedia objects, robot requests and failed requests with a total of 1169 clean log records ready for further processing. This means that the size of the log file was reduced to 20 % of the original log size. The Tables 1 and 2 shows the statistics about individual request category and aggregated results of Data Cleaning Step. Figure 6. depicts the distribution of irrelevant Data in Web Log. It is observed that a major portion of Web Log usually consists of irrelevant and redundant data which has to be eliminated to speed up the upcoming mining process.

Table 1 Statistics of individual request category

Request Category	Number of records	Percentage
PNG	408	7
JPEG	2501	43
CSS	291	5
JS	582	10
NOT GET	58	<1
NOT 200	1164	20
Robot	931	16

Table 2 Aggregate results of data cleaning

Statistics	Number of records
Original size	5817
Failed requests (other than 200)	1164
Multimedia objects	3782
Robots	931
Cleaned log size	1169
Percentage in reduction	80

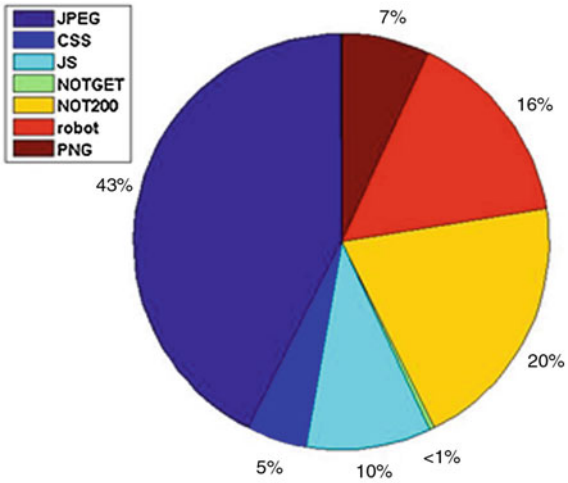


Fig. 6 Distribution of irrelevant data in web log

5.2 *User Identification and User Session Identification*

The User Identification Algorithm uniquely identifies the users of Web Site. A total of 235 users were identified in the given log. The session identification splits all the pages accessed by each user into individual access sessions using the combined technique based on time oriented and navigation oriented heuristics. It was

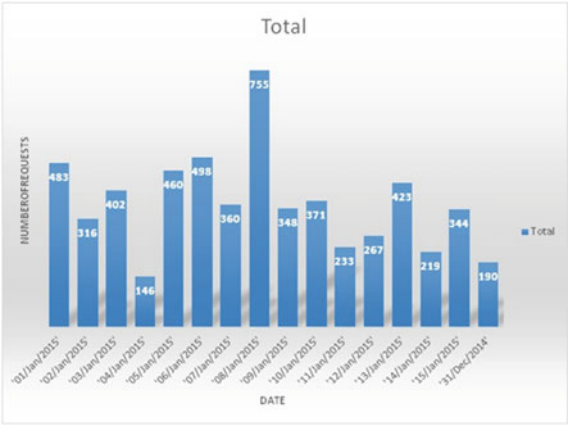


Fig. 7 Days versus no. of requests

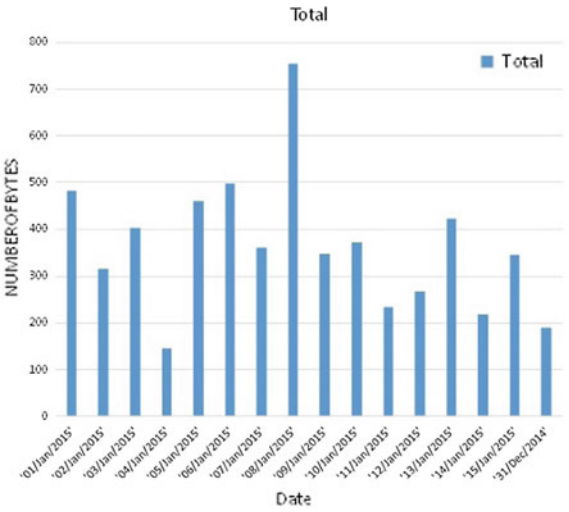


Fig. 8 Days versus no. of byes downloaded

observed that each user having one session with maximum number pages in a session = 26.

5.3 *Analysis of Web Log Pre-processing Results*

The simple analysis of Web Log Repository, after Web Log Preprocessing, could be useful for the web site administrator. The chart of Requests across days and chart of downloads across days is given in Figs. 7 and 8.

6 Conclusion

Web Log Preprocessing is one of the complex tasks of Web Usage Mining. Modified Web Log Preprocessing eliminates the noisy data and drastically reduce the input log thereby lessen the burden on the further tasks. In this paper, the Web Log Preprocessing algorithms based on Dimensionality Reduction Techniques and Combined Methodologies on Web Log Repository from a real time web server is been implemented. The Web Log Preprocessing Algorithm has identified around 16 % robot requests in the log. Results of preprocessing have shown that the input web log size is reduced by 80 %. The results show that the Web Log Preprocessing techniques based on various dimensionality reduction techniques and combination of methods improve the performance of Web Log Preprocessor.

References

1. Neelima, G., et al.: An overview on web usage mining. In: Emerging ICT for Bridging the Future—Proceedings of the 49th Annual Convention of the Computer Society of India, vol. 2, Advances in Intelligent Systems and Computing, vol. 338, pp. 647–655. Springer International Publishing (2015)
2. Bhargav, A., et al.: Pattern discovery and users classification through web usage mining. In: Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on, IEEE, pp. 632–635 (2014)
3. Eltahir, M.A., et al.: Extracting knowledge from web server logs using web usage mining. In: Computing, Electrical and Electronics Engineering (ICCEEE), 2013 International Conference on, IEEE, pp. 413–417 (2013)
4. Malik, S.K., et al.: Information extraction using web usage mining, web scrapping and semantic annotation. In: Computational Intelligence and Communication Systems, 2011 International Conference on, IEEE, pp. 465–469 (2011)
5. Varnagar, C.R., et al.: Web usage mining: a review on process, methods and techniques. In: Information Communication and Embedded Systems (ICICES), 2013 International Conference on, IEEE, pp. 40–46 (2013)
6. Sudheer Reddy K., et al.: An effective data pre-processing method for web usage mining. In: Information Communication and Embedded Systems (ICICES), 2013 International Conference on, IEEE, pp. 7–10 (2013)
7. Sael, N., et al.: Web usage mining data pre-processing and multi level analysis on moodle. In: Computer Systems and Applications (AICCSA), 2013 ACS International Conference on, IEEE, pp. 1–7 (2013)
8. Maheswari, B.U., et al.: A new clustering and pre-processing for web log mining. In: Computing and Communication Technologies (WCCCT), 2014 World Congress on, IEEE, pp. 25–29 (2014)
9. Carmona, C.J., et al.: Web usage mining to improve the design of an e-commerce website: OrOliveSur.com. Expert Syst. Appl. **39**(12), 11243–11249 (2012). Elsevier
10. Doran, D., et al.: Web robot detection techniques: overview and limitations. In: Data Mining and Knowledge Discovery, pp. 1–28. Springer, US (2010)
11. T.M.A. Pai Polytechnic Web Site: http://www.tmapaipolytechnic.com/

Co-operative Co-evolution Based Hybridization of Differential Evolution and Particle Swarm Optimization Algorithms in Distributed Environment

Suma Nambiar and G. Jeyakumar

Abstract Evolutionary computing algorithms play a great role in solving real time optimization problems. One of the evolutionary computing algorithm is Particle Swarm Optimization algorithm (*PSO*). The aim of this paper is to propose a model to improve the performance of *PSO* algorithm. Hybrid models of Particle Swarm Optimization (*PSO*) algorithm and Differential Evolution (*DE*) has already proved to be one of the better approaches for solving real world complex, dynamic and multimodal optimization problems. But these models hybridize *PSO* and *DE* to form a new serial algorithm. In these serial hybridization models, we are losing the originality of both *DE* and *PSO* algorithms since the structure of both the algorithms is being modified to get the hybridized *PSO* and *DE* algorithm. In this paper, we develop a model for *PSO* in distributed environment with improved performance in terms of speed and accuracy. The proposed model is a hybridized distributed mixing of *DE* and *PSO* (*dm-DEPSO*) which improves the performance of *PSO* algorithm. In this model, algorithms are implemented in a cluster environment to perform co-operative co-evolution. Better solutions are migrated from one node to another in the cluster environment. Co-operative co-evolving model shows better performance in terms of speed and accuracy. The algorithm is applied to a set of eight benchmarking functions and their performance are compared by mean of objective function values, standard deviation of objective function values, success rate, probability of convergence and quality measure.

Keywords Co-operative co-evolution · Distributed *DE* · Distributed *PSO* · Distributed mixing of *DE* and *PSO*

S. Nambiar (✉) · G. Jeyakumar
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Amrita Nagar PO, Coimbatore 6411 112, India
e-mail: suma.c.nambiar@gmail.com

G. Jeyakumar
e-mail: g_jeyakumar@cb.amrita.edu

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_17

1 Introduction

Evolutionary computing algorithms are computationally good and can be used in case of complex multimodal and noisy optimization problem domains. Each evolutionary algorithm in evolutionary computing has its own strength and weakness. The Particle swarm optimization (*PSO*) algorithm is one of the evolutionary algorithms, which is easy to understand and implement. Differential evolution (*DE*) is a new addition to evolutionary algorithms group, which is most suitable for real valued optimization problems. The weakness of both *PSO* and DE lies in the premature convergence, in case of large and complex problems. Our motivation for this paper is to design a structure where *PSO* can work in an environment to improve its performance. Our experimental results prove that *PSO* works better in a distributed environment. This distributed model will work similar to a ring topology and migrates better result from one node to another. This paper proposes a distributed mixing of *DE* and *PSO* (*dm-DEPSO*) which shows a better performance when compared to the performance of *PSO*. Xingquan Zuo and Li Xiao have proposed an algorithm to hybridize *DE* and *PSO* for dynamic optimization problems [1]. In this paper, they have proposed a serial hybridization of *DE* and *PSO* algorithms where both the algorithms are losing their originality. To retain the originality of the *DE* and *PSO* algorithms, we are developing an environment in which *DE* and *PSO* can be run separately and parallel in a cluster. There are different variants of *DE* algorithm based on their crossover and mutation operations. In our experimental setup, we are considering *DE/rand/*1*/bin*, which is a most popular *DE* variant, for its simplicity and efficiency. It says that a single pair of parents is selected randomly and their difference is added to another random parent to generate a mutant vector. The *PSO* also has different variants based on its constraints. This paper uses inertial weight of 0.729 for *PSO* algorithm.

The objective of this paper is to develop a hybridized distributed environment in which *PSO* and *DE* can run parallel in a cluster to improve the performance of *PSO* in terms of speed and accuracy. Our experimental results prove that *dm-DEPSO* outperforms *PSO* and distributed *PSO* (*d-PSO*).

This paper is organized into six sections. Section 2 contains literature survey of existing techniques. The proposed methodology is explained in Sect. 3. Design of experiment is explained in Sect. 4. Results and analysis are discussed in Sect. 5. Section 6 describes conclusion and scope for future work.

2 Literature Survey

Evolutionary computing tools build intelligent systems that can solve a complex problem. Traditional optimization algorithms use some deterministic rule to move from one from one point to another in a hyperspace. But there is a great likelihood of getting stuck at local optimum. After each generation, evolutionary computing

algorithms generate same number of members as that of initial population so that many maxima and minima can be obtained simultaneously.

2.1 Particle Swarm Optimization Algorithm

With the Inspiration of bird flocking and social behavior, Eberhert and Kennedy developed *PSO* where algorithm keeps track of the fitness of its particles and their respective positions in the hyperspace where best solutions are obtained [2]. *PSO* uses only few parameters and its concept is simple and easy to implement. The main problem faced by *PSO* is that it results in premature convergence in case of complex problems. *PSO* algorithm also initializes a random population of solutions. These solutions are called particles and are considered to be flying in the search space in the direction of current optimum solution. Each particle is associated with their fitness values and velocities which direct them each time to the current optimum solution. Particle gets updated by two values: *pbest* and *gbest*. *pbest* stores the fitness value or the best solution achieved by a particle so far and *gbest* stores the global best solution achieved so far by any particle. *vid* represents the velocity of the particle, *xid* represents their position and c_1 and c_2 are acceleration constants. *pid* and *gid* represents *pbest* and *gbest* values, respectively.

$$vid = vid + c_1 rand()(pid - xid) + c_2 rand()(gid - xid) \tag{1}$$

$$xid = xid + vid \tag{2}$$

2.2 Differential Evolution Algorithm

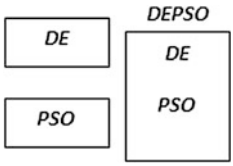
DE algorithm generates an initial population of solutions randomly. Fitness or objective value of each member in the population is calculated. *DE* algorithm chooses 3 solutions from the population set randomly (*DE/rand/*1*/bin*) for each parent to create an offspring using *DE* operators (mutation and crossover). This is repeated a number of times equal to the population size. If the fitness of offspring is more, then replace parent with the offspring. There are different variants of *DE* based on different mutation and crossover logics. *DE/rand/*1*/bin* is proved to be the best and popular among all the variants of *DE*. This variant uses the binomial crossover and uses the following mutation strategy.

$$V_{i,G} = X_{r_1^i,G} + F\left(X_{r_2^i,G} - X_{r_3^i,G}\right) \tag{3}$$

Table 1 *DE* and *PSO* implementation results

Functions	DE		PSO	
	Mean	*StdOV*	Mean	*StdOV*
f_1	2.92E − 12	1.25E − 12	2.46E + 09	5.38E + 09
f_2	3.08E + 00	7.02E + 00	2.63E + 04	3.57E + 03
f_3	4.90E − 03	1.63E − 03	4.90E + 01	2.21E + 00
f_4	2.64E + 01	1.65E + 01	9.81E + 07	1.63E + 07
f_5	1.30E − 12	2.42E − 12	3.70E + 02	2.27E + 01
f_6	5.13E − 09	3.60E − 08	6.31E + 00	7.65E + 00
f_7	4.23E − 20	2.27E − 20	1.29E + 02	2.27E + 01
f_8	1.32E − 23	1.04E − 23	1.31E + 08	1.01E + 08

Fig. 1 Serial hybridization



In [3], Storn and Price proved that *DE* converges faster when compared to their counterpart algorithms. The other advantage of using *DE* is that it is robust, easy to implement and can be used for parallel computation. The results after implementing *DE* and *PSO* are recorded in Table 1.

2.3 Serial Hybridization

In [4], Xin, Chen and Peng proposed *DEPSO*, a serial hybrid optimizer by combining the strength of both *DE* and *PSO*. A variant of *PSO*, *PSO* with von Neumann topology and variant of *DE*, *DE/rand/1/bin* was hybridized in their model. For each solution, either PSO or DE is chosen as evolution method based on their success ratio. Their experimental results showed that *DEPSO* gives good performance in solving multimodal functions. For global optimization of complicated multimodal functions, this algorithm gives a better solution. The main drawback is that originality of both *DE* and *PSO* is lost when this model is used. Figure 1 shows an operational scheme of serial hybridization of *DE* and *PSO* (*DEPSO*) where both *DE* and *PSO* logics are used in the same algorithm.

3 Proposed Method and Design

For improving the performance of *PSO*, a distributed cluster environment (*d-PSO*) was created and results were recorded. A hybridized distributed model is expected to show even more better performance than *PSO* algorithm and hence one of the efficient evolutionary computation algorithm *DE*, was selected for creating such a model. *DE* was also implemented in a distributed environment (*d-DE*). A new model was proposed in [5] called *dmvDE* which mix effective *DE* variants in a distributed environment. Also, the experimental results showed that distributed *d-DE* is much suitable for optimization problems. Hence, the hybridized distributed mixing of *DE* and *PSO* (*dm-DEPSO*) is proposed to show a better performance than *PSO* algorithm. This section describes the proposed model and algorithm for distributed *DE*, distributed *PSO* and distributed mixing of *DE* and *PSO* (*dm-DEPSO*).

3.1 Distributed dE *(*d-dE*) and Distributed* PSO *(*d-PSO*)*

A distributed environment was set up to implement *DE* and *PSO* to ensure that both algorithm works better in such an environment. In our model for *d-DE* and *d-PSO*, we created a distributed environment (cluster) of four nodes. Divide the population set equally into all nodes and implement *DE/PSO* algorithm separately among these nodes. These nodes are running in parallel. The sub populations are co-evolved and better results are migrated from one node to another. This approach is called co-operative co-evolving approach. The operational scheme of d-DE and d-PSO is shown in Fig. 2. The algorithmic description of *d-DE/d-PSO* is as follows

Algorithm

Step 1. Initialize population.
Step 2. Divide the population into 4 subpopulations and input each into 4 different nodes.
Step 3. Set up *DE/PSO* separately in all nodes.
Step 4. Each subpopulation is evolved by their respective algorithm in the node
Step 5. Migrate the better results from one node to another, to solve the given problem.

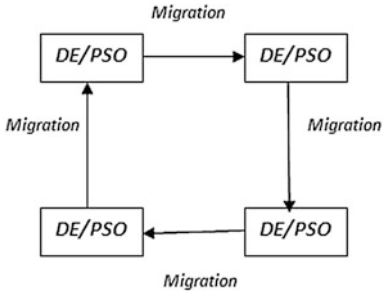


Fig. 2 Operational scheme of *d-DE*/*d-PSO*

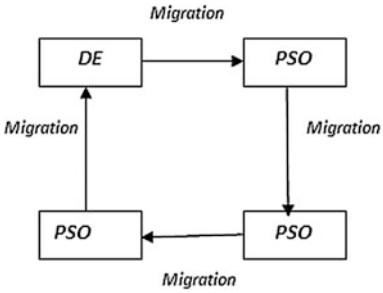


Fig. 3 Operational scheme of *dmDEPSO*1

3.2 Hybridization of DE *and PSO (*dm-DEPSO*)*

The implementation of algorithms in distributed environment was successful and then a hybridized mixing of both *DE* and *PSO* was proposed. The hybridization model gains the strength of one algorithm to overcome the weakness of other. The operational scheme is similar to *d-DE* or *d-PSO*. However, instead of implementing the same algorithm in all the four nodes of the cluster we make use of following cases:

DE in 1 node, *PSO* in 3 nodes: *dm-DEPSO*1
DE in 2 nodes, *PSO* in 2 nodes: *dm-DEPSO*2
DE in 3 nodes, *PSO* in 1 node: *dm-DEPSO*3

Algorithm

Step 1. Initialize population.
Step 2. Divide the population into 4 subpopulations and input each into 4 different nodes.
Step 3. Set up *dm-DEPSO* (3 cases) in all nodes.
Step 4. Each subpopulation is evolved by their respective algorithm in the node.
Step 5. Migrate the better results from one node to another, to solve the given problem.

The operational scheme of *dm-DEPSO*1 is shown in Fig. 3. This model also makes use of co-operative co-evolving approach.

4 Design of Experiments

Our aim is to compare the performance of *PSO* and *d-PSO* with our proposed *dm-DEPSO* based on certain performance metrics chosen from evolutionary computing literature. The performance metrics used for comparing the performance of each model are:

Metrics for measuring accuracy

- Mean Objective Function Value (*MOV*)
- Standard Deviation of Objective Function Values (*StdOV*)

Metrics for success performance

- Success Rate (*SR*)
- Probability of Convergence (P_c *(%)*)

Metrics for measuring speed

- Quality Measure (Q_m)

Mean Objective Function Value (*MOV*): The Mean Objective Function Value is measured as the average of all the objective function values provided by an algorithm for any benchmarking function in all the independent runs.

Standard Deviation of Objective Function Values (*StdOV*): It is measured as the standard deviation of all the objective function values provided by an algorithm for any benchmarking function in all the independent runs.

Success Rate (*SR*): It is the percentage between the number of successful runs made by an algorithm out of total number of runs it performed for a function. A run is regarded as successful run if it reaches the tolerance error or the solution before the stopping criteria. The *SR* value for an algorithm is calculated as follows

$$SR = \left({}^{nsr}/_{ntr}\right)\% \tag{4}$$

where *nsr* is the number of successful runs and *ntr* is the total number of runs.

Probability of Convergence (P_c*(%)*): The probability of convergence [6] is a measure similar to *SR*. But it calculates the percentage between the total number of successful runs and total number of all the runs made by an algorithm for all the benchmarking functions considered in the experimental set up. This metric is used to identify the algorithm with highest probability of convergence, by considering all the different benchmarking functions. It is calculated as follows

$$P_c = \left({}^{nsr\_a}/_{ntr\_a}\right)\% \tag{5}$$

where *nsr_a* is the number of successful runs for all the functions and *ntr_a* is the total number of runs for all the functions.

Quality Measure (Q_m): Quality measure or simply Q- measure [6] is a metric which combines the convergence speed and probability of convergence of an algorithm as a single value. It is calculated as follows

$$Q_m = \left({}^{C_m}/_{P_c}\right) \tag{6}$$

In Eq. (6) the P_c value is calculated as described in Eq. (5), the C_m value is calculated as follows

$$C_m = \left(\sum_{i=1}^{nsr\_a} fesr_i / nsr\_a \right) \tag{7}$$

where nsr_a is the number of successful runs for all the functions and $fesr_i$ is the number of function evaluations for the successful run i

On comparison, the algorithm with lesser C_m and higher P_c are regarded as competitive algorithm. Such algorithm will result lower Q_m value, comparing to other algorithms.

The experimental setup includes parameter setup and fixing the benchmarking suite. For implementing *DE/rand/*1*/bin*, the values for the following parameters were decided based on [7].

1. Mutation Scaling Factor (F) = 0.3 to 0.9
2. Crossover Rate (CR) = 0.9, 0.5, 0.9, 0.1, 0.9, 0.5, 0.9 and 0.1 for f_1; f_2; f_3; f_4; f_5; f_6; f_7 and f_8 respectively.
3. Population Size (NP) = 100

For the implementation of *PSO*, the parameters used (based on the literature) are:

1. Inertia Weight (w) = 0.729
2. Cognitive Weight (c_1) = 1
3. Social Weight (c_2) = 1

Apart from these, the other parameters common for both *DE* and *PSO* are decided as follows:

1. Dimension (D) = 30
2. Stopping criteria = $1*10^{-12}$(tolerance error)
3. Maximum Generation ($MaxG$) = 2000
4. Total number of runs= 50

The experimental setup includes 8 benchmarking functions (chosen from [7]) with different categories: unimodal separable, unimodal non-separable, multimodal separable and multimodal non-separable.

f_1: Schewefel's function
f_2: Schewefels function 2.2
f_3: Schewefels function 1.2
f_4: Generalized Rosenbrocks function
f_5: Generalized Rashtrigins function
f_6: Ackley's function
f_7: Generalized Rashtrigin's function
f_8: Generalized Penalized function

For the distributed environment of *dm-DEPSO*, the parameters are decided as follows based on [8].

1. Number of nodes (ni) = 4
2. Number of migrants (nm) = 1

3. Migration frequency (*mf*) = 45
4. Selection Policy (*sp*) = Migrates the best individual (Individual with good fitness value)
5. Replacement Policy (*rp*) = The best individual that migrates to a target node replaces a random individual in target node

5 Results and Discussion

5.1 Results for dE *and* PSO

For each of the benchmarking function, we have evaluated the mean and standard deviation of their objective values. (Table 1). The number of successful runs and the success rate is calculated and tabulated for *DE* and *PSO* in Table 2. For *DE*, the success rate value is zero for functions f_1, f_3, f_4 and f_6 because the number of successful runs made by *DE/rand/*1*/bin* for these functions is zero. The values of Probability of Convergence (P_c) for *DE* and *PSO* are shown in Table 2. Quality measure is calculated and shown in Table 3.

In case of implementation of *PSO*, the success rate values are zero for all the functions because the number of successful runs made by *PSO* for all these functions are zero. So the Probability of Convergence (P_c) and Quality Measure (Q_m) are also zero. From Tables 1, 2 and 3, it is noticed that in all the performance metrics *DE* is outperforming *PSO*. This motivated us to choose *DE* to mix with *PSO*.

Table 2 Probability of convergence for *DE* and *PSO*

Functions	DE		PSO	
	No. of success runs	*SR*	No. of success runs	*SR*
f_1	0	0	0	0
f_2	50	1	0	0
f_3	0	0	0	0
f_4	0	0	0	0
f_5	35	0.7	0	0
f_6	0	0	0	0
f_7	50	1	0	0
f_8	50	1	0	0
	P_c =46.25%		**P_c=0**	

Table 3 Quality measure for *DE* and *PSO*

Functions	*DE*	*PSO*
	nFE	*nFE*
f_1	0	0
f_2	5,078,113	0
f_3	0	0
f_4	0	0
f_5	235,065	0
f_6	0	0
f_7	4,051,623	0
f_8	3,479,384	0
	Q_m = 2057.25	Q_m = 0

Table 4 Distributed-*DE* and Distributed-*PSO* Implementation Results

Functions	*d-DE*		*d-PSO*	
	Mean	*StdOV*	*Mean*	*StdOV*
f_1	9.57E − 13	3.80E − 14	4.71E + 01	4.10E + 00
f_2	1.72E − 15	1.15E − 14	2.06E + 04	1.09E + 03
f_3	9.66E − 17	4.48E − 17	3.94E + 01	4.48E + 00
f_4	2.56E + 00	3.34E + 00	1.47E + 07	7.95E + 06
f_5	9.29E − 13	6.48E − 14	2.97E + 02	1.94E + 01
f_6	7.06E − 17	1.01E − 16	1.62E + 00	8.09E + 00
f_7	9.26E − 13	5.75E − 14	7.51E + 01	1.33E + 01
f_8	9.31E − 13	5.45E − 14	5.76E + 06	3.84E + 06

5.2 *Results for* D-dE *and* D-PSO

The mean and standard deviation of the objective function values measured for *d-DE* and *d-PSO* is shown in Table 4. On comparing the results in Tables 1 and 4, it is observed that *d-DE* and *d-PSO* are showing better performance than their serial versions. The number of successful runs, the success rate and Probability of Convergence (P_c) measured for *d-DE* and *d-PSO* are shown in Table 5. The Quality measure is calculated and shown in Table 6.

The results in Tables 5 and 6 shows that the *d-DE* has shown increased performance by Probability of convergence and Quality measure than classical *DE*. However, the *d-PSO* does not show any performance difference by its Probability of convergence and quality measure than the classical *PSO*.

Table 5 Probability of Convergence for *d-DE* And *d-PSO*

Functions	*d-DE*		*d-PSO*	
	No. of success runs	*SR*	No. of success runs	*SR*
f_1	50	1	0	0
f_2	50	1	0	0
f_3	50	1	0	0
f_4	50	0	0	0
f_5	0	1	0	0
f_6	50	0	0	0
f_7	50	1	0	0
f_8	50	1	0	0
	P_c = 87.5 %		P_c = 0 %	

Table 6 Quality Measure for *d-DE* and *d-PSO*

Functions	*d-DE*	*d-PSO*
	nFE	*nFE*
f_1	4,657,680	0
f_2	6,000,000	0
f_3	6,000,000	0
f_4	0	0
f_5	5,025,420	0
f_6	6,000,000	0
f_7	3,882,120	0
f_8	3,293,400	0
	Q_m = 1138.24	Q_m = 0

5.3 Results for Dm-DEPSO

We implemented all the three cases of *dm-DEPSO* (*dm-DEPSO*1, *dm-DEPSO*2, *dm-DEPSO*3) and compared their performance based on the performance metrics with *DE*, *PSO*, *d-DE* and *d-PSO*. It is observed from the results (Tables 7, 8 and 9) that among the 3 cases of *dm-DEPSO*, the *dm-DEPSO*3 shows better performance in terms of accuracy of the solution.

On comparing *PSO*, *d-PSO* and *dm-DEPSO*, it is noticed that the *dm-DEPSOs* have provided more accurate solutions to the considered benchmarking problems. However, the *dm-DEPSOs* could not provide any successful run for any of the functions.

Interestingly, on comparing *DE*, *PSO*, *d-DE*, *d-PSO* and *dm-DEPSOs*, we found that the *d-DE* algorithm has shown better performance in terms of both accuracy and speed.

Table 7 *dm-DEPSO*1 Implementation Results

Functions	Mean	*StdOV*	*SR*
f_1	1.02E+00	7.36E−01	0
f_2	1.75E+03	9.63E+02	0
f_3	1.32E+01	2.67E+00	0
f_4	1.01E+05	1.76E+05	0
f_5	9.88E+01	1.55E+01	0
f_6	5.77E+00	1.05E+00	0
f_7	1.86E+00	9.82E−01	0
f_8	4.59E+00	3.75E+00	0

Table 8 *dm-DEPSO*2 Implementation Results

Functions	Mean	*StdOV*	SR
f_1	8.83E−01	1.02E+00	0
f_2	1.24E+03	6.01E+02	0
f_3	1.25E+01	2.50E+00	0
f_4	2.20E+04	3.13E+04	0
f_5	5.84E+01	1.24E+01	0
f_6	3.80E+00	6.67E−01	0
f_7	1.14E−01	1.15E−01	0
f_8	2.11E−02	9.69E−02	0

Table 9 *dm-DEPSO*3 Implementation Results

Functions	Mean	*StdOV*	SR
f_1	8.17E−05	8.33E−05	0
f_2	6.16E+02	2.91E+02	0
f_3	7.34E+00	2.92E+00	0
f_4	6.67E+03	1.52E+04	0
f_5	6.58E−03	4.48E−02	0
f_6	3.34E+00	1.00E−00	0
f_7	2.27E−03	3.53E−03	0
f_8	5.12E−06	1.50E−05	0

6 Conclusion and Future Scope

This paper proposed a distributed model to mix *DE* and *PSO* in a co-operative co-evolutionary framework (*dm-DEPSO*). The implementation of *dm-DEPSO* necessitated 3 different cases of *dm-DEPSO* algorithms. The experimental results on implementing *DE*, *PSO*, *d-DE*, *d-PSO* and *dm-DEPSO* have shown that the performance of *PSO* algorithm is improved to a higher extent by mixing it with *DE*. The identified performance difference between *PSO* and *dm-DEPSO* may be due to the co-operative co-evolutionary approach where the weakness of *PSO* is balanced by the strength of *DE*. The future work includes validating the performance of *dm-DEPSO* on some real time applications.

References

1. Zuo, X., Xiao, L.: A DE and PSO based hybrid algorithm for dynamic optimization problems. Springer, Berlin (2013)
2. Poli, R., Kennedy, J., Blackwell, T.: Particle swarm optimization: an overview. Springer Science + Business Media (2007)
3. Storn, R., Price, K.: Differential Evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces. J. Glob. Optim. (1996)
4. Bin, Xin, Jie, Chen, ZhiHong, Peng, Feng, Pan: An adaptive hybrid optimizer based on particle swarm and differential evolution for global optimization. Sci. China Inf. Sci. **53**(5), 980–989 (2010)
5. Jeyakumar, G., Shunmuga Velayutham C.: Distributed mixed variant differential evolution algorithms for unconstrained global optimization. Memet. Comput. **5**(4) 275–293, (2013) (Springer)
6. Feoktistov, V.: Differential evolution in search of solutions, Springer, New York (2006)
7. Jeyakumar, G., Shunmuga Velayutham C.: An empirical comparison of differential evolution variants on different classes of unconstrained global optimization problems. In: Proceedings of the international conference on computer information systems and industrial management application, pp. 866–871 (2009)
8. Jeyakumar, G., ShunmugaVelayutham C.: Empirical study on migration topologies and migration policies for island based distributed differential evolution variants. Lecture Notes in Computer Science, Springer, vol. 6466, pp. 95–102 (2010)

Application of Multiple Databases Based on Hand Features in Hand Biometry

V.R. Prakash, Kumaraguru, G. Vimala Rani and S.K. Indumathi

Abstract We propose the significance in use of multiple databases for the purpose of authentication using the palm-print of the hand for increased speed. The technique is used for authentication using the acquired image of the hand. The hand could be held in any pose hence providing pose independence. We use Appearance analysis for feature extraction. This technique also suggests hand independence for biometric authentication. It also provides a robust technique for detecting the extremities using a neighborhood scan method.

Keywords Neighborhood method · Palm-print · Valley-peak points · Pose correction

1 Introduction

The evolution of hand biometry can be classified into four generations. Constrained and contact based [1, 2]. In the beginning when the hand geometry was introduced, the sensors were contact based; it also used pegs and other constraints to held the hand in position. Unconstrained and contact based [3–6]. Here the hand geometry was developed to do away with pegs and other constraints. Unconstrained contact free and position dependent Contact free [7–10] technique was developed for hand biometry, but the disadvantage for the developed technique is that it assumes that the

V.R. Prakash (✉) · Kumaraguru · G. Vimala Rani · S.K. Indumathi
School of Electrical Science, Hindustan University, Chennai, India
e-mail: vrprakash@hindustanuniv.ac.in

Kumaraguru
e-mail: pkguru@hindustanuniv.ac.in

G. Vimala Rani
e-mail: gvrani@hindustanuniv.ac.in

S.K. Indumathi
e-mail: skindumathi@hindustanuniv.ac.in

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_18

person held his hand perfectly parallel to the camera or sensor. Using either geometric or palm print information For identifying hand biometry either geometric or palm print information used and no techniques specify the use of both for identification. In all of the previous work either left or right hand was used, not both.

For hygiene, increased security and ease of use people generally prefer contact free technique for biometric identification/Authentication. For hand biometry people may wish to use either right or left hand. By separating out the right or left hand and creating separate databases for each increases the speed of hand biometry

Recently more emphasis is given for biometric securities. One of the main challenges in using the hand biometry is the pose of the hand. The pose of the hand will be varying on each use. The solution for this is pose correction. Principle component analysis is a good technique that can be used for feature extraction. Euclidean distance and curvature properties can be used for extracting the valley and peak points of the hand efficiently.

2 Methodology

The Fig. 1 depicts the block diagram of proposed system used for hand biometry.

2.1 Hand Image Acquisition

For image acquisition any digital cameras can be employed. All the images are captured with homogeneously colored background, this helps in extraction of the hand much easier. The images are taken in a well-lit environment. The fingers must be kept separately for the acquisition so that the extremities can be easily found.

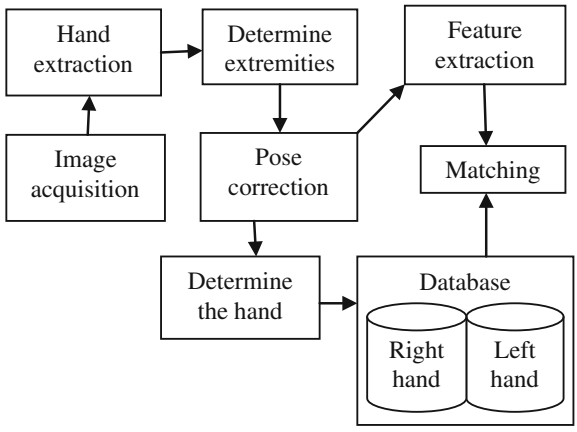


Fig. 1 Block diagram of the used scheme

2.2 Hand Extraction

The easiest approach in hand extraction is Otsu's method [11] which separates the background image from the foreground image.

2.3 Extracting Valley Points

The best technique for finding the extremities is by determining the curvature of the hand [12] as shown in the Fig. 2.

1. Detect edges using the Sobel edge detector to obtain a binary edge map.
2. Extract edge contours from the edge map. When the edge reaches an end point, fill the gap and continue the extraction if the end point is nearly connected to another end point, or mark this point as a T-junction corner if the end point is nearly connected to an edge contour, but not to another end point
3. After contour extraction, compute the curvature at a fixed low scale for each contour to retain the true corners, and regard the local maxima of absolute curvature as corner candidates.
4. Compute a threshold adaptively according to the mean curvature within a region of support. Round corners are removed by comparing the curvature of corner candidates with the adaptive threshold.
5. Based on a dynamically recalculated region of support, evaluate the angles of the remaining corner candidates to eliminate any false corners.

A more robust alternative technique was provided by Euclidean distance. Hamming distance across the rows is first calculated. The highest hamming distance is provided by a line following the centre of the hand. The image area below this line corresponding to the hamming distance is removed. Then filling is done. Filling

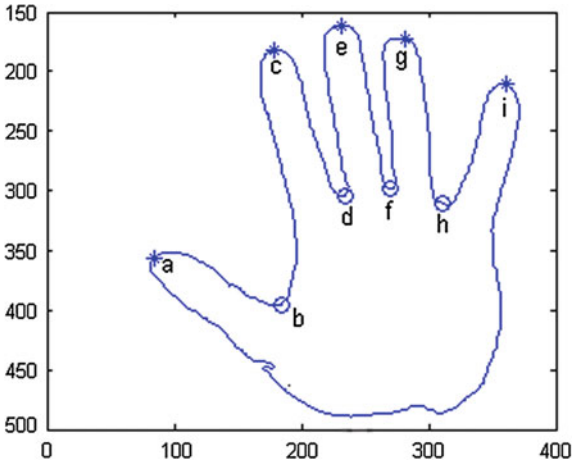


Fig. 2 Figure showing various valley and peak points

is nothing but making all the pixels below any white pixel in the image white. This helps in finding error free Euclidean distance.

Then edge is extracted using Sobel transform. Then Euclidean distance with respect to the base of the image is calculated. The resulting sequence of radial distances yields minima and maxima corresponding to the sought extreme points. The maxima and minima are extracted from the image by dividing the graph into various segments. The resulting maxima and minima are moderately accurate but not perfect. These points are marked in an image having intensity value zero throughout with one.

Detecting and localizing the hand extremities, that is, the fingertips and the valley between the fingers is the first step for hand geometry. Since both types of extremities are characterized by their high curvature, curve gram of the contour, that is, the plot of the curvature of the contour at various scales along the path length parameter was computed. The accuracy is high but with false points. Secondly points were extracted based on Euclidean distance. This have low accuracy but without any false points.

The combination technique combines both approaches to get the best suited output. The algorithm is a decision based algorithm, Suppose X is the image obtained after Euclidian approach & Y is the image obtained after Corner approach. First locate the first white pixel in X. Scan the neighbourhood of the index of white pixel found in X in Y. If any white pixel is found then it is taken. Else the point originally found in X is taken.

Third method implemented uses following algorithm

- Generate edge map as in curvature method
- For first point, find its location.
- Search location with higher value after current location −X it becomes next maxima and consider this as next first point.
- Search location with lower value after current location +X it becomes next minima and considers this as next first point.

This third method is found to be the best suited for hand biometry.

2.4 Pose Correction

A landmark is a point of correspondence on each object that matches between and within populations. An example of a hand annotated with 28 landmarks is illustrated in the Fig. 3.

For getting the landmark points-

1. Interpolate points threw the contour between the extremities
2. Draw line between these points
3. Find the distance of each line through the point

Fig. 3 Landmark Points

4. Using this distance information interpolate equally distance points
5. These points becomes the landmark points for comparison.

The database contains large number of hand. To obtain a true shape representation we have to filter out all the location, scale and rotational effects. Solution-establish a coordinate reference with respect to position, scale and rotation to which all the shapes will be aligned. This reference coordinate system is represented in Fig. 4. From the database we select one image as target and remaining images has to be aligned to it.

First the average of the image is found and this is added to the coordinates to relocate the image to centre. Average of the thickness is found and ratio of it with that of target is found using this we scale the image. The angle formed between different points are found. The average of difference of angle from the target is used to correct the rotation.

2.5 Palm-Print Extraction

Palm-Print is the pixel intensities across the object in question (if necessary after a suitable normalization). For this mapping is done from one arbitrary point set $\{X_1 \ldots ... X_n\}$ into another $\{X'_1 \ldots ... X'_n\}$.

To partition the image for palm-print analysis, connect three landmark points in such a way that a triangle will be created. Then every point inside the triangle can be expressed by their relative distances from each of the vertices. This method is called Delaunay triangulation. Delaunay triangulation connects an irregular point set by a mesh of triangle's each satisfying the Delaunay property. The Delaunay property is as follows:

No triangle must have any point inside its circumcircle, which is the unique circle that connects all three points (vertices) of a triangle.

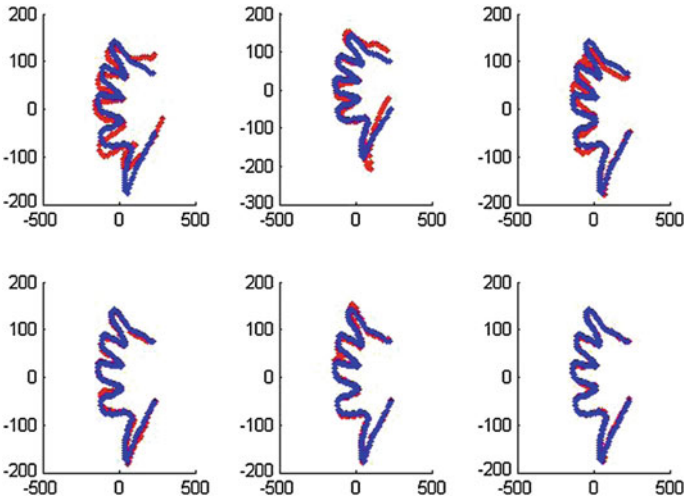


Fig. 4 Pose corrected hand

With implementation in mind, one should notice the common denominator. In pseudo code the piece-wise affine warp can be written as:

1. Determine the triangle, X belongs to
2. Find the relative position of X given by

$$\mathrm{X} = \alpha \times a + \beta \times b + \gamma \times c \tag{3}$$

$$\alpha = 1 - (\beta + \gamma) \tag{4}$$

$$\beta = \frac{yx_3 - x_1y - x_3y_1 - x_3y_1 - y_3x + x_1y_3 + xy_1}{-x_2y_3 + x_2y_1 + x_1y_3 + x_3y_2 - x_3y_1 - x_1y_2} \tag{5}$$

$$\gamma = \frac{yx_2 - xy_1 - x_1y_2 - x_2y + x_2y_1 + x_1y}{-x_2y_3 + x_2y_1 + x_1y_3 + x_3y_2 - x_3y_1 - x_1y_2} \tag{6}$$

   where $\alpha + \beta + \gamma = 1$
   And X is inside a triangle if $0 \leq \alpha, \beta, \gamma \leq 1$
3. To obtain the mapped position

$$X' = \mathrm{f(X)} = \alpha x_1' + \beta X_2' + \gamma X_3' \tag{7}$$

4. $$\mathrm{Set}\, I'(\mathrm{X}) = \mathrm{I}(\mathrm{f}(\mathrm{x})) \tag{8}$$
5. Interpolate the points between the obtained points.

2.6 Databases

Here two databases are used instead of one for speeding up the operations. The two databases are maintained and classified based on left or right hand. To determine the hand to be detected is left or right

1. During the pose correction the target image to which the images has to be aligned should be held vertically.
2. Determine the coordinates of the extremities.
3. Select the two extremities in the rightmost and leftmost of the hand.
4. If the coordinates in the leftmost side is more than in the rightmost side then it is a right hand.
5. If the coordinates in the rightmost side is more than in the leftmost side then it is a left hand.
6. Based on the determined hand select the appropriate database.

3 Results

When using neighborhood scan based approach it can be seen that accuracy has been improved to a large extend. And can be verified by the following Tables 1, 2 and 3. In which X and Y are the coordinates.

For matching small variations are allowed in z-axis. In most scenarios for an angle up to 15° recognition provided satisfying results. The pose correction technique prescribed here can correct rotational effect up to 40 , but the information becomes too unreliable in the case of hand features. The main problem of hand biometry was difficulty in obtaining reliable feature point. Even the techniques that

Table 1 Extremities determined using corner method

x	188	175	173	241	209
y	132	159	195	238	308

Table 2 Extremities determined using Euclidean method

x	111	188	71	175	55	173	74	241	197
y	108	132	144	159	188	195	217	238	259

Table 3 Extremities determined using third method

x	111	189	71	177	55	168	74	239	197
y	108	134	144	166	188	206	217	250	259

produced good information were iterative and require more than 2 h for extracting the features of twenty five hands. This problem has been addressed here but restricting rotational motion in x-y plane.

4 Conclusion

By using third method shown in the Table 3 we obtained better detection of extremities. The multiple database technique had also provided significant speedup. Further improvement can be achieved by modifying the pose correction to work on three dimensional images and implementation more database based on more hand features for increased speed.

References

1. Sanchez-Reillo, R., Sanchez-Avila, C., Gonzalez-Macros, A.: Biometric identification through hand geometry measurements. IEEE Trans. Pattern Anal. Mach. Intell. **22**(10), 168–1171 (2000)
2. A.K. Jain, A. Ross, S. Pankanti, A prototype hand geometrybased verification system. In: Proceedings AVBPA, pp. 166–171, Mar 1999
3. Xiong, W., Toh, K.A., Yau, W.Y., Jiang, X.: Model-guided deformable hand shape recognition without positioning aids. Pattern Recognit. **38**(10), 1651–1664 (2005)
4. Ribaric, S., Fratric, I.: A biometric identification system based on eigenpalm and eigenfinger features. IEEE Trans. Pattern Anal. Mach. Intell. **27**(11), 1698–1709 (2005)
5. Woodard, D.L., Flynn, P.J.: Finger surface as a biometric identifier. Comput. Vis. Image Understand. **100**(3), 357–384 (2005)
6. Methani, C., Namboodiri, A.M.: Pose invariant palmprint recognition. In Proceedings ICB, pp. 577–586 (2011)
7. Malassiotis, S., Aifanti, N., Strintzis, M.G.: Personal authentication using 3-D finger geometry. IEEE Trans. Inf. Forensics Security **1**(1), 12–21 (2006)
8. Kumar, A.: Incorporating cohort information for reliable palmprint authentication. In Proceedings ICVGIP, pp. 583–590, Dec 2008
9. Morales, A., Ferrer, M., Díaz, F., Alonso, J., Travieso, C.: Contact free hand biometric system for real environments. In: Proceedings 16th Europe Signal Process. Conference Laussane, Switzerland, Sept 2012
10. Kanhangad, V., Kumar, A., Zhang, D.: Combining 2-D and 3-D hand geometry features for biometric verification. In: Proceedings IEEE Workshop Biometrics, Miami, FL, pp. 39–44, June 2013
11. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybernet. **9**(1), 62–66 (1979)
12. He X.C.: Corner detector based on global and local curvature properties. Optical Engineering

Serious Game on Flood Risk Management

P. Meera, M.L. McLain, Kamal Bijlani, R. Jayakrishnan and Bhavani R. Rao

Abstract The applications of using serious games as a teaching tool are vast. One of the fields of knowledge that is now being implemented in serious games is Disaster Relief Management. Serious games use a variety of multimedia and strategies that allow the learner to participate in a graphic simulation of a disaster. This work gamifies the topic of flood risk management. Floods are a unique type of disaster in that they have an element of predictability. This predictability can be utilized as a gaming element involving time limits to reduce the amount of destruction and loss of life due to flooding. Of the many benefits of using games for educational purposes, it can help people who live in rural areas who may have limited education comprehend this complex material in a meaningful way. A learning approach called Kolb's Learning Model is used to convey the material. It familiarizes people with the different responses and terminologies of the hazard while achieving high concentration and interest from the learner. After developing this game, the effectiveness of this gaming method is analyzed by a comparative study of text book learning and this serious game.

Keywords Serious games · Flood · Risk management · Disaster management · Kolb's learning model

1 Introduction

Natural disasters are a massive problem all around the world. It can happen because of the natural processes on the earth. Floods, earthquakes, volcanic eruptions, and tsunamis are some of the natural disasters that can occur. Floods are the most

P. Meera (✉) · K. Bijlani · R. Jayakrishnan
Amrita E-Learning Research Lab, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, Kerala, India

M.L. McLain · B.R. Rao
Ammachi Labs, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, Kerala, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_19

common natural disasters worldwide. A flood is the overflow of water that submerges land [1]. Frequency of worldwide flood disasters is increasing year by year [2], so it is important to plan and construct a safer environment. Taking the correct actions for a disaster management is important. One way this can be achieved is to instruct people who live in flood zones and areas at risk of flooding by offering them specialized courses in flood management. This can help reduce the adverse effects of floods by presenting different disaster risks that require solutions. We can reduce the vulnerability of an at risk location by taking preventive measures such as raising the land, making firebreaks, etc. [3]. Nowadays human activities contribute to these disasters by effecting climate change, deforestation, and reduction of wetlands, which leads to flooding so it is important to have an effective disaster management system. The adverse effects of floods can be reduced through some structural and non-structural measures and should have an established method of disseminating information about flooding. It is essential to this information system is trusted [4]. Often floods can have enormous adverse effects while people are not prepared due to lack of early warning systems, preparedness and mitigation measures. Thus, it is important to educate people about the causes and risks of disasters.

Game based learning is one of the best methods to help people understand a phenomenon. Each participant can experience emotions in a very personal and enduring manner [5]. There are many benefits in game based learning method. It can aid learning to those who have limited education because of its pictorial simulation and display. Also, it can implicitly support the learner to improve their knowledge [6]. Serious games are the best opportunity to learn and experience different flood risk situations and also it provides an admirable motivation and engagement. With the help of serious games complex real life situations can be simulated in a simple manner in the game world. By taking the correct responses at the correct time, we can reduce the causalities of flood disaster. This is done by taking precautions at the right time with the help of various disaster management institutions and also making use of military resources. It will help reduce the potential losses from any harmful hazards and it is important for an effective management of an accident. The player identifies the roles and responsibilities for each member of the community and the preparedness measures. By describing different types of hazards and the preparedness measures can be of both personal benefit and benefit for the community for mitigating the worst impacts of flood disaster. Serious games are the best tool to get insights in decision making also raising awareness for people [7]. Therefore, it will be an excellent tool to look at different future implications and to facilitate the involvement of people in these strategies to initiate some movement during deadlock situations.

The game provides a teaching tool for learning the different phases of flood risk management including Mitigation, Preparedness, Response and Recovery. Then it demonstrates the responses related to the flood. Usually, children and adults are most affected during the flood because of lack of awareness and preparation. In these years floods are becoming a major hazard. The serious game is demonstrating all vulnerable situations and their mitigation methods. The vulnerable situations like soil erosion, evacuation, medical care, and sewage collection and risk awareness.

All citizen have the right to be informed of how to respond to flooding, also ensuring that all the emergency responders have adequate resources and skills. All these issues are demonstrated by this serious game. It will improve the player's ability to respond against any vulnerable situations during disaster.

The rest of the paper is organized as follows, Sect. 2 provides the background studies of flood disaster, Sect. 3, conveys how the system is organized, risk management, Sect. 4 includes the solution approach, Sect. 5 is Result analysis, and Sect. 6 shows the conclusion.

2 Related Works

Several works have been done on disaster risk management. Meesters et al. [8] have proposed a serious game for disaster information management as a team player; each team should manage different scenarios and can communicate each other. As per their research, all the first responders should evacuate people into their prospective shelters. If medical support is needed anywhere, then responders can make use of mobile application provided to the participants by means of proper communications. The study provides methods for evacuating people properly and how to effectively manage information about disasters. The responder should manage available data and experience with real-time feedback. In another study [9], Felicio and the team have introduced a serious game for different disasters. The player should need to select one disaster and then get into the play. But the drawback in this paper is that the graphics are poor and also the game is not user-centric. The Player is unable to Stop or Pause the game and there are no multiplayer and different levels. The Score is another concern. Based on another research work by Syukril et al. [10], the plot's challenges arise as different monsters. The player needs to solve each monster with different challenging environment like puzzles and acquire scores. Each player is acting as a magician. This is an example of active learning; the learner could get knowledge from their own experiences. We evaluated all the existing game based learning method and found out that there is the gap in between the game and the learner. The method of teaching used in our serious game is analyzing the learner capability and to motivate them to take suitable actions on the game through the psychological aspects.

3 System Design

Serious games realize the primary purpose of learning along with entertainment. Kolb's learning model is a four stage learning method. This is used in the game for facilitating the learner in learning. The stages are Experience, Reflection, Conceptualize and Test. Here the serious game is for developing awareness about different vulnerable situations in flood disaster and the mitigation methods. It

reduces the impact of the disaster by allowing people to take proper precautions. The main targeted audiences are students for learning about disaster management as part of their syllabus. In addition to that, it is also helpful for general people who lived in the flood prone area. The people will get awareness and motivate them to make further steps on real time flood disaster. During the crisis, each organizer should be aware about their tasks and responses. The serious game gives awareness for all such responders. An effective learning is achieved by reinforcing the learner for learning concepts and key steps regarding the disaster throughout the game. An exercising method with that is providing an initial experience for the player. In the game, the player would need to identify appropriate resources and to satisfy each and every need regarding a flood disaster.

Kolb's learning model is used to demonstrate the game. In the first stage the player should analyze and identify the risk situation. Then the player finds out the active and proper solution for the risk from the organizers. We provide an information button for the player to observe the nature of that particular area. From that information they will have a perception about the risk of that situation. It will help them to select the proper choice for risk management. In the last phase, we test the knowledge that the learner acquired from the game. A quiz will evaluate the players content acquisition level. Quiz questions are created on the basis of Bloom's Taxonomy for facilitating the cognitive thinking of the learner. The stages of cognitive domain, such as knowledge, comprehension, application and evaluation are used to frame the quiz questions.

4 Solution Approach

When flooding occurs, it involves economic, people and environmental consequences. We are considering all those consequences throughout the game play.

The economic consequences such as

- Repair of roads
- Reservoir destruction
- Electricity and communication
- Bridges
- Salvaging damage to agriculture
- Automobile destruction

Environmental consequences such as

- Water pollution
- Introduces insects
- Soil erosion

People consequences such as

- Flood walls
- Improved early warning

The flood management system has 4 steps such as mitigation, preparedness, response and recovery. The game contains three levels. First level includes mitigation and preparedness. In this phase the player obtains knowledge about the preparedness and mitigation measures of flood disaster. Mitigation is to prevent hazards from developing into disasters. Preparedness measure will reduce the consequences of disaster while they occur. In this, we can focus on long term risk reduction measures like different cropping patterns for reducing crop loss, preventing massive leakage from nuclear plant which causes high radiation in that area, as well as short term measures. The second level conveys the response measures on flood disaster. To take a direct action to save lives, protect property, care for victims, etc. and also mitigate the amount of damage. The third level is the recovery stage. In this, the player takes actions that involve rebuilding destroyed property, re-employment and repair. We are considering a flood prone environment with high population and risks. Figure 1 shows the game flow. At first, the collector in that area is conducting a conference with the main organizers in that area such as health officers, civil engineers, PWD officers, fire fighters, etc. He announces that there is a possibility of flood within 48 h. He then assigns tasks for every member for flood risk management. The fire safety officers should be concerned about fire operations and other accidents, and the medical department should assure the medical support. In the case of bridging of two coastal areas for easy evacuation, the PWD officers will take care about that. Like that each department should do their work properly. Maintenance should complete within the time limit. We are assigning scores for the number of lives saved. If the player loses any lives the score will be reduced. Then the player needs to play the game as each responder. If there is a need for military resources, the player can make use of it. This also increases the score. Then the high

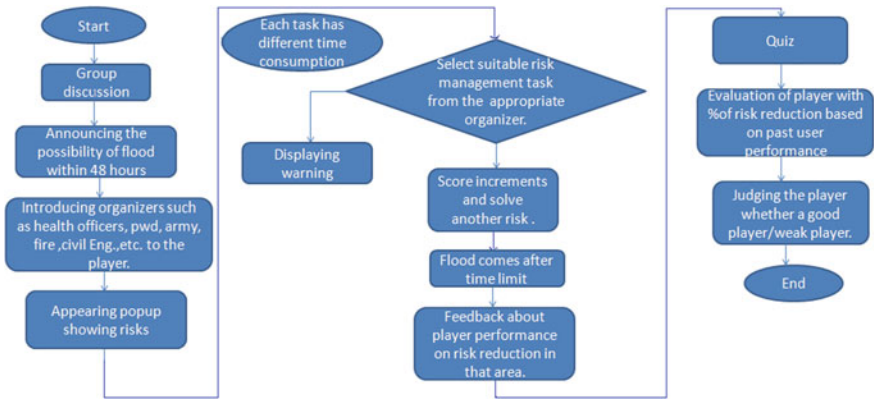


Fig. 1 Architecture of serious gaming for flood risk management

Fig. 2 Game screen shot of group discussion

flooded area could be considered first for getting a good score. The flood will come after the time limit. If the player has overcome the risks, then he is allowed to take a quiz, otherwise we will give them feedback and allow them to play again by resolving problems that he was not able to solve earlier. In the end, we are displaying Score, Feedback and Badges for motivation and reinforcement.

Figure 2 shows the conference with the important organizers as mentioned above. They will discuss the current challenging situations in the city and also discuss the measures that can be used to mitigate them. Risk assessment can be done in this phase as well as taking appropriate actions to reduce the vulnerabilities. First, shows an introduction about that area with population, economic status and percentage of vulnerability. Figure 4 shows the game play scenario. There is a group of organizers and each organizer is assigned to particular tasks. When a popup comes out, the player comprehends the risk situation from the pop-up and selects the appropriate action from the assigned tasks of organizers. If the selection of task is wrong, that will reduce the score and also gives feedback for that response. The popup arises along with a 5 s timer. When the timer completes, the popup will end up automatically. And then shows another risk situation. The player should complete all the risks within 10 min time limit (Fig. 3).

There is a risk bar at the top, which shows the remaining risks. Before 10 min completion, a warning for flood will appear. It also shows some clues for the remaining risks. Once the time completes, the flood will occur. Then, we give feedback about the performance of the player and score. As like this, the player needs to complete level 2 and 3. After that there is a quiz. The quiz is created on the basis of Bloom's Taxonomy, assesses the player content knowledge. Figure 4 shows the quiz after the game play. Through this we can evaluate the player as to whether he is a good player or a weak player. Throughout the game we are

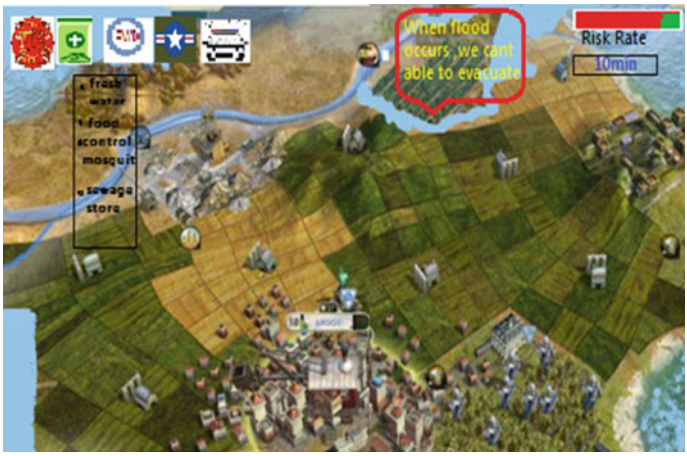


Fig. 3 Identifying risk and taking the appropriate task from the organizer

visualizing Kolb's learning model. Kolb's model includes the phases such as experience, reflection, conceptualize and test. In the serious game, first player experiences the flood risks, then responding to that risk. From this they will get to know about the risk situations and the different mitigation methods.

The evaluation part intends to check whether the player understands the content knowledge. The timing is another important concern. The player should complete almost all the risks within the time limit. So he needs to identify more time consuming and less time consuming work. The serious game has implemented with the help of Unity3D. Unity is a game development ecosystem. It has a powerful

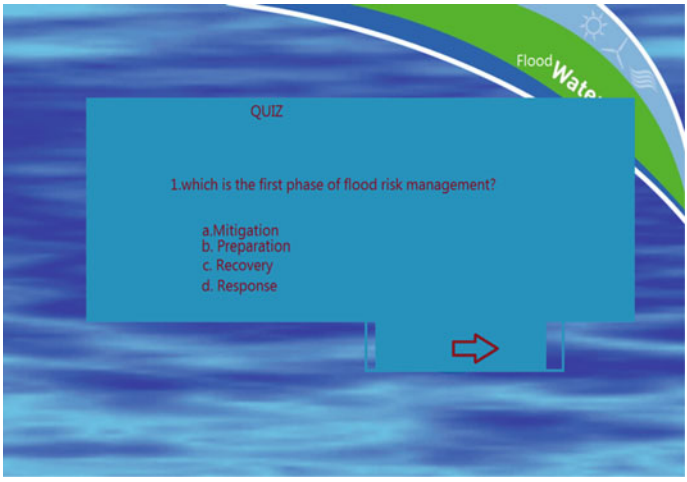


Fig. 4 Quiz after the quiz for checking the content knowledge

Table 1 Comparative study of serious game with text based learning

	Textbook-based learning	Game-based learning
Understandability of the content	Comparatively less	High
Interaction	Tutor-learner interaction	Direct involvement
Real-world experience	Nil	High
Self assessment	Less	High
Motivation	Less	High
Engagement	Comparatively less	High
Effectiveness	Less	High

rendering engine fully integrated with a complete set of intuitive tools and rapid workflows to create interactive 3D and 2D content [11]. It is also an easy multi-platform.

5 Result Analysis

The comparative study of textbook-based learning with game-based learning is shown below in the Table 1. Different factors such as effectiveness, interaction, self-assessment, motivation, experience, understanding and involvement have compared. It has drawn attention to the effectiveness of game-based learning.

A group of high school students was chosen for analysis of the effectiveness of the game. A quiz program was conducted for the evaluation of knowledge acquirement while providing two different resources to learn about flood management. Textbook material and developed game were the two different resources provided to them.

Scrutinizing the results from both learning shows the significance in game based learning. Students answered most of the questions correctly while playing the game as compared to text book learning method. Figure 5 conveys that the learner understands more about the flood management using serious games. Thus, game based learning is one of the best educational technologies for improving quality of education.

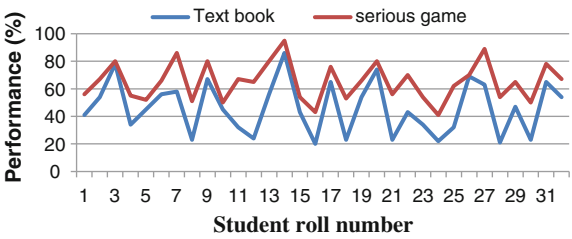


Fig. 5 Analysis of knowledge acquisition

6 Conclusion

The work in this paper shows that serious games can be a very effective teaching tool to convey the process of flood risk management. It can be especially useful to help people learn who have limited education in rural areas at risk. The design of this game explores a flood risk management practice through which it ensures an effective decision making at the time of disaster alert. Through this game, people can develop the confidence to respond to each flood risk. The psychology based learning methods used here: Kolb's model and Bloom's Taxonomy, assure motivation among learners. This also helps to improve the effectiveness of the serious game. The scope for serious game in learning process has been validated by the significant increment seen in the performance of students in the conducted testing phase. With a relatively good accuracy ratio, the game succeeded in providing better awareness and opportunity for understanding the content of the flood management. It helps to prepare, manage, reduce and cope with different risks at the time of flood occurrence.

Acknowledgments We are grateful for the support and facilities provided by the Amrita Vishwa Vidyapeetham University and Amrita E learning Research lab.

References

1. Uche, D.O.V.: The environmental effects of flood disaster in Anambra state. Pelagia Res. Libr. Adv. Appl. Sci. Res. **4**(1), 499–505 (2013)
2. EM-DAT: The OFDA/CRED international disaster database: www.em-dat.net, universitecatholique-Brussels-Belgium
3. Xuan, W., Zhao, G.: Early warning monitoring and management of disasters. In: Proceedings of the Workshop on Applications of Spatial Information Technology on Disaster Prevention and Reduction, 1-4244-1212-9/07/$25.00 ©2007 IEEE, 08 Oct 2005
4. Samaddar, S., Misra, B.A.: Flood risk awareness and preparedness: the role of trust in information sources. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics, COEX, Seoul, Korea 14–17 Oct 2012
5. Barthélémy-Ruiz, C., Carpier, B., Clément, N.: United nations high commissioner for refugees ©1995. http://www.unhcr.org/473dc1772.pdf
6. Soler, J.L., Mendoza, P., Pons-Vazquez, S.: Impact of video games-based learning on academic performance and motivation: a case study
7. Iacovides, I., Cox, A.L.: Moving Beyond Fun: Evaluating Serious Experience in Digital Games, CHI 2015, Seoul, Republic of Korea, Publication rights licensed to ACM, 18–23 April 2015
8. Comes, T., Fiedrich, F., Fortier, S., Geldermann, J., Müller, T., Meesters, K., van de Walle, B. (eds.): Disaster in my backyard: a serious game introduction to disaster information management. In: Proceedings of the 10th International ISCRAM Conference, Baden-Baden, Germany, May 2013
9. Felicio, S., Dargains, S.: Stop disasters game experiment with elementary school students in Rio de Janeiro: building safety culture. In: Proceedings of the 11th International ISCRAM Conference, USA, May 2014

10. Syukril, M., Pramono Hadi, M., Straatssma, M.W.: Differences in disaster response due to varying data availability a serious game for flooding disaster research in Surakarta, Indonesia. (ISSN 0024-9521, IJG vol. 43, no. 1, June 2011 (49–62)© 2011 Faculty of Geography GadjahMada University the Indonesian Geographers Association
11. Create the Game You Love with Unity: Unity game development engine. http://unity3d.com/unity

SS_PAR: Signal Strength Based Petal Ant Routing Algorithm for Mobile Ad Hoc Network

M. Manjunath and D.H. Manjaiah

Abstract Mobile ad hoc network (MANET) is a collection of mobile node which acts as both host and router at the same time during communication. Each node in MANET usually runs on battery and capable of communicating to other without fixed base station. Since node has limited transmission range, limited bandwidth and limited battery, obtaining optimal path with less overhead is a challenging task. In this paper, we proposed Signal Strength based Petal Ant on-demand routing (SS_PAR) algorithm which is based on swarm intelligence technique and the signal strength of mobile node. The proposed SS_PAR reduces overhead by analyzing signal strength inside the petal and optimize the node that participated during route discovery. We compared our routing algorithm with SARA and simulation results shows that SS_PAR performance is better than SARA.

Keywords SS_PAR · PAR · Petal ant routing · SARA · MANET

1 Introduction

Mobile ad hoc network (MANETs) [1] is an infrastructure-less, Multi-hop, self-configuring ad hoc network [2]. Each device in MANETs is mobile in nature and has limited bandwidth, transmission range which usually runs on battery. In Manet, when source node wishes to establish a communication channel to destination, source may don't have direct path to destination because of limited transmission range. Source may need to rely completely on intermediate node for obtaining path to destination. In Manet, all the control packets, data packets must traverse through the multiple nodes and multiple links to reach destination. So obtaining optimal path with accurate topological information during data communication is a challenging task [3]. The core problem in Manet is still, the finding the accurate path between source and destination during high mobility [4–6]. In many

M. Manjunath (✉) · D.H. Manjaiah
Department of Computer Science, Mangalore University, Mangalore, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_20

literature proposal such as AODV [7], DSDV [8], AOMDV [9], ARA [4], SARA [10], MDSR [11], DSR [12], HOPNET [13], ANTNet [14], etc., one can find different routing technique to handle this problem but none of them provide solution for all the scenarios [15].

The remaining part of this research article is organized as follows. Section 2 describes the literature survey. Section 3 presents the proposed work. Section 4 presents the Simulation results and comparison. Finally conclusion and acknowledgement is described at the end of the report.

2 Related Work

Manjunath and Manjaiah [16] proposed routing algorithm called Spatial DSDV (SDSDV) routing algorithm for mobile ad hoc network. The SDSDV consists of three phases that is identification of activity area, identification of representative node and the routing process. SDSDV algorithm eliminates redundant node during route discovery and provide better performance in terms of PDF with respect to number of packet generation across the network.

An intelligent based routing is proposed in [17], where author has considered the power saving management (energy) as a major Quality-of- Service (QoS) parameter for construction of routes in ad hoc wireless networks. The ABIRP improves the energy efficiency by analyzing the power status of node either as active or idle mode and forward the data packets using ACO routing technique by establishing path between end nodes using FANT and BANT. The author has compared the ABIRP with AODV and stated that ABIRP is better than AODV.

Spatial Ad hoc On demand Distance Vector (SAODV) [18] is a reactive and improved version of Ad hoc On Demand Distance Vector routing protocol. Spatial AODV (SAODV) routing protocol allows the source node to identify the route quickly to destination by eliminating redundant node that participated during the route discovery process. Spatial AODV routing protocol consist of 3 phases Identification of Activity area, Identification of Representative nodes and Routing process. SAODV uses of route request (RREQ), route reply (RREP) packets for identifying routes between end nodes and uses HELLO message to detect the broken link between the neighboring nodes. In SAODV, when source needs a communication channel for transmitting data or require a path to destination, it identifies the activity area and representative node and non representative node in the network and eliminates the redundant node during the routing process.

3 SS_PAR: Signal Strength Based Petal Ant Routing

This section details the problem definition by define a scenario and the proposed SS_PAR algorithm in details.

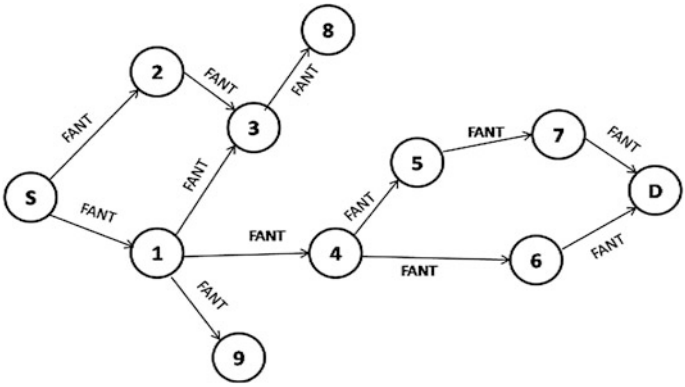


Fig. 1 FANT propagation of ACO

3.1 Problem Definition

In popular ACO algorithm, when source require a communication path to destination, Source flood the route request packet called Forward Ant [FANT] in the network. Figure 1 illustrates FANT propagation of ACO. The node which receives the FANT will replicate and rebroadcast again until it reaches destination. The node in the network may receive multiple FANT for the same route, which update the routing table and rebroadcast it again. Flooding mechanism of FANT and BANT in the network is disadvantage which increases the time required for searching route during route discovery, increases additional time to update routing table, increases overhead and decreases the network performance [4, 10]. Hence aim of our proposed work is to eliminate redundant FANT and BANT propagation during routing process by analyzing the signal strength with Ant colony technique.

3.2 SS_PAR

The SS_PAR is a routing algorithm based on swarm intelligence technique called Ant based routing for mobile ad hoc network, which is an extended variant of SARA [10] and extract few feature of petal routing [19]. The main objective of SS_PAR is to minimize the overhead by eliminating redundant Forward Ant [FANT], Backward Ant [BANT] and the node that processed during the route discovery phase. The proposed SS_PAR calculate the petal region between source and destination and analyze the signal strength inside the petal region during routing process. The algorithm is divided into two phases that is route discovery and route maintenance.

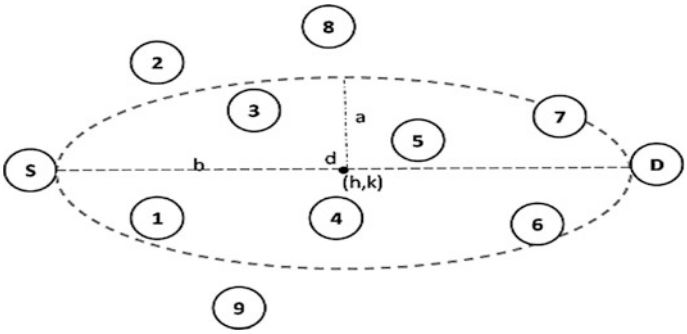


Fig. 2 SS_PAR petal calculation

3.2.1 Route Discovery

When source has data to transfer to destination, initially the source searches destination in its transmission range, as destination lies in the transmission range, the source starts transmitting packets directly without any routing technique otherwise SS_PAR constructs route as follows.

Source $S(x_s,y_s)$, Destination $D(x_d, y_d)$ and the Intermediate node $I(x_i, y_i)$ {i = 1, 2, 3, …, n}. SS_PAR merge the concepts of geographic routing and addresses the network node based on latitude and longitude along with node id. The (x, y) coordinate of a network node represents (longitude, latitude) respectively. When destination doesn't lies in its transmission range, source compute petal region/width of the petal $[P_w]$ using expression (1), (2), (3) and (4) as shown in Fig. 2.

$$d = \sqrt{(x_d - x_s)^2 + (y_d - y_s)^2} \tag{1}$$

$$h = (x_s + x_d)/2 \tag{2}$$

$$k = (y_s + y_d)/2 \tag{3}$$

$$P_w = \pi ab \tag{4}$$

Once the P_w is calculated, the source flood the special packet called Petal-Forward Ant [P_FANT] in the network. The P_FANT and P_BANT are small route discovery packets which consist of width of the petal $[P_w]$, sequence number, pheromone count and the threshold $[T_i]$ (signal strength). SS_PAR uses expression (5) to identify node inside or outside the P_w and uses expression (6) to set threshold value T_i.

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} \leq 1 \tag{5}$$

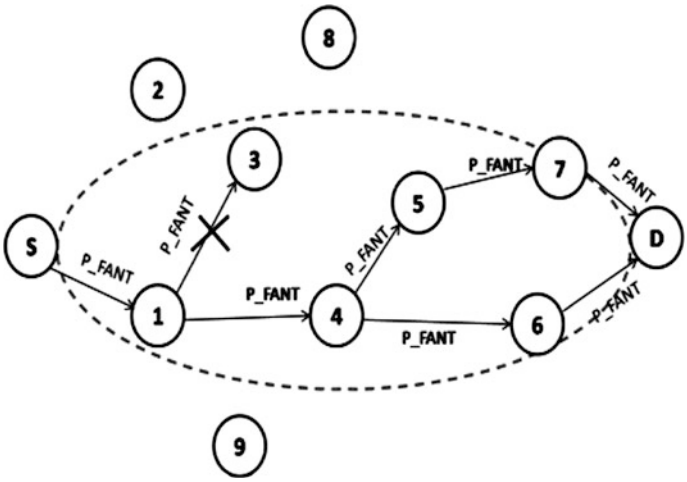


Fig. 3 SS_PAR P_FANT propagation

$$T_i(d) = \frac{P_t \times G_t \times G_r \times h_t^2 \times h_r^2}{d^4 \times L} \tag{6}$$

where P_t is the Transmitted Signal Power (0.28183815 W), G_t is the Transmission Gain of Antenna, G_r is the Receiver Gain of Antenna, h_t is the height of Transmitted Antenna, h_r is the height of Receiving Antenna, d is the distance and L is the System Loss. Depending upon the petal region P_w and threshold T_i, the network node accept the P_FANT packet, update the routing table, update the pheromones value and rebroadcast it again to its next neighboring node. For example, from the Fig. 3, the source node has two adjacent neighboring nodes that are node 1 and node 2. As node in the network receives the P_FANT packet for the first time, it verifies P_w in P_FANT using expression (5) and analyses the receiving signal strength (RSS) threshold T_i using expression (6). The node during route discovery if satisfy both the conditions will probably accept P_FANT and rebroadcast it again. The setting of value P_w and T_i makes P_FANT to propagate within the petal region rather than whole network. The network node which does not lie in P_w will drop the P_FANT packet and doesn't participate in routing process. This process continues until the P_FANT reaches destination. Once the P_FANT packet reaches destination, it replies with P_BANT packet back to the source nodes. Upon the reception of P_BANT packet, source starts transmitting data through shortest path updated by each intermediate node in the network. Figures 3 and 4 schematically depict the SS_PAR route discovery process respectively.

3.2.2 Route Maintenance

In mobile ad hoc networks, nodes are highly dynamic which moves away from one place to other. A route may be in valid due to dynamic nature of mobile or link being broken or by pheromone evaporation during data communication. SS_PAR need no special packet for maintaining active route, which work similar as SARA routing algorithm.

4 Simulation Experiment Setup Using NS2

The performance of SS_PAR is evaluated through five different simulation scenarios. The proposed routing is compared with Simple Ant Routing Algorithm (SARA) [10]. The SARA uses the code written by Fernando Correia, Teresa Vazao. SS_PAR algorithm is simulated using Network Simulator [20], version NS2.35 on Ubuntu operating system. The SS_PAR can be deployed at network layer as shown in Fig. 5.

The simulation is carried for 160 s and consists of 15, 20, 25, 30 and 35 network nodes which move randomly with minimum speed of 0 m/s to maximum speed of 10 m/s according to two ray propagation mobility model.

The Fig. 6 shows the figure of packet generated of SARA and SS_PAR. The proposed SS_PAR generates an average of 2.071 % packets more than SARA. The Fig. 7 shows the figure of comparison of two routing algorithm in terms of received packet across the network. The proposed SS_PAR receives an average of 2.108 % packets more than SARA. The Fig. 8 shows the performance of two routing

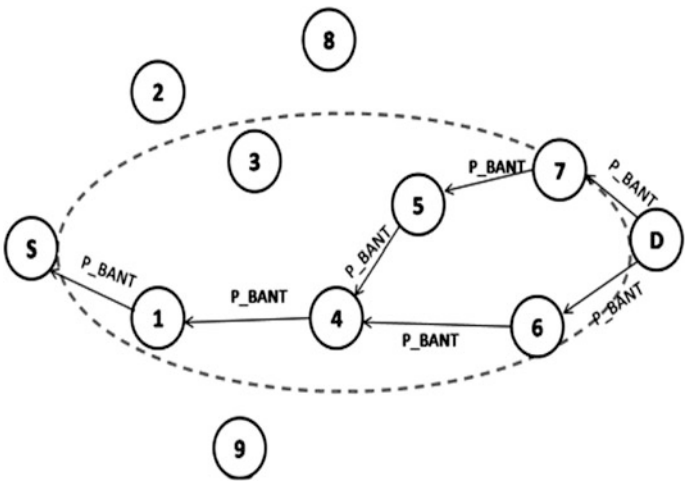


Fig. 4 SS_PAR P_BANT propagation

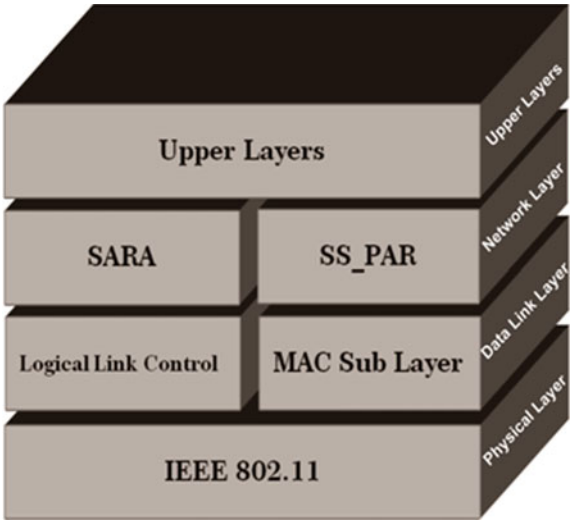


Fig. 5 Deployment of SS_PAR in the OSI protocol stack

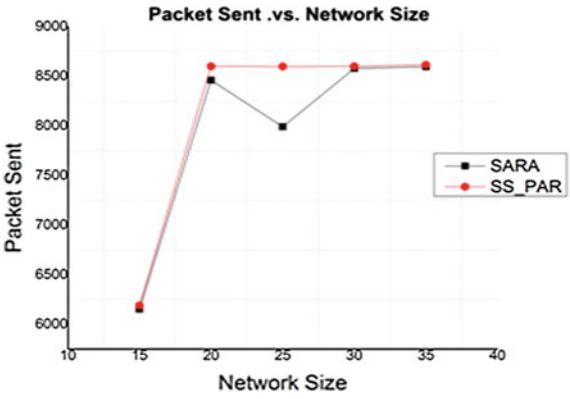


Fig. 6 Packet sent versus network size

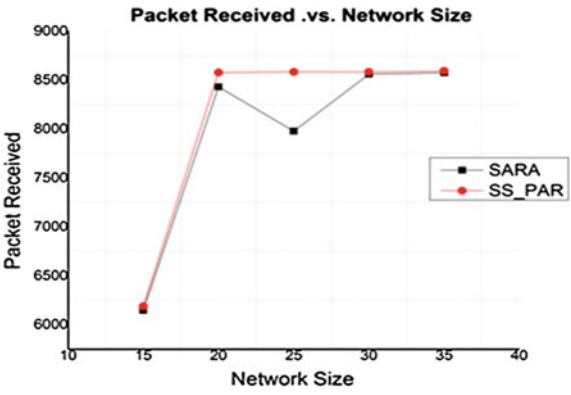


Fig. 7 Packet received versus network size

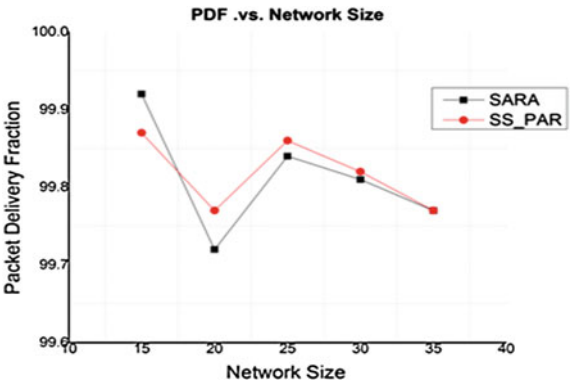


Fig. 8 PDF versus network size

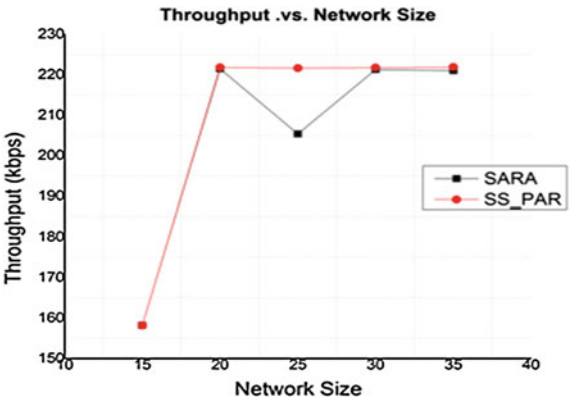


Fig. 9 Throughput versus network size

algorithm in terms of packet deliver fraction (PDF). The proposed SS_PAR shows better performance by increasing an average PDF to 0.0060 % more than SARA. The Fig. 9 shows the figure of comparison of throughput with respect to different

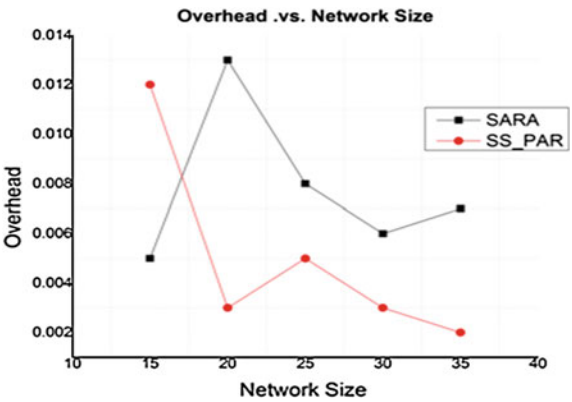


Fig. 10 Overhead versus network size

network size. The proposed SS_PAR provide an average of 1.745 % more throughput than SARA. The Fig. 10 shows the performance of two routing algorithm in terms of Overhead. The proposed SS_PAR reduces overhead of an average 35.897 % less than SARA.

5 Conclusion

The proposed SS_PAR is a routing technique for Manet, which analyses the signal strength inside the petal using ant colony technique and make P_FANT and P_BANT to propagate within the petal region which optimizes the network node during route discovery. The SS_PAR is simulated in NS2 and the simulation results of both routing show that the proposed routing algorithm has better performance by reducing the overhead of an average of 35.897 % less than SARA.

Acknowledgments The author's wishes thanks to UGC, Indian government, for providing funding support under Rajiv Gandhi National Fellowship (RGNF) scheme Ref. No. F1-17.1/2012-13/RGNF-2012-13-SC-KAR-17563/(SAIII/Website) dated February 28, 2013. Thanks also go to the dedicated research group in the area of Computer Networking at the Dept of Computer Science, Mangalore University, Mangalore, India. Lastly but not least the author would like to thank everyone, including the anonymous reviewers.

Appendix

The performances of routing protocols were studied with respect to different network size. The experimental result shows that our proposed model performs better than SARA routing algorithm. Tables 1 and 2 show the analysis results of SARA and SS_PAR routing algorithm respectively.

Table 1 Results of SARA routing algorithm

Network size	Packet sent	Packet received	PDF	Throughput	Overhead
15	6151	6146	99.92	158.19	0.005
20	8460	8432	99.72	221.46	0.013
25	7990	7978	99.84	205.37	0.008
30	8577	8561	99.81	221.28	0.006
35	8594	8575	99.77	221.02	0.007
Total	39,772	39,692	499.06	1027.32	0.039
AVG	7954.4	7938.4	99.812	205.464	0.0078

Table 2 Results of SS_PAR routing algorithm

Network size	Packet sent	Packet received	PDF	Throughput	Overhead
15	6189	6189	99.87	158.18	0.012
20	8599	8579	99.77	221.8	0.003
25	8596	8584	99.86	221.63	0.005
30	8598	8583	99.82	221.75	0.003
35	8614	8594	99.77	221.89	0.002
Total	40,596	40,529	499.09	1045.25	0.025
AVG	8119.2	8105.8	99.818	209.05	0.005

References

1. Giordano, S.: Mobile ad hoc networks. In: Handbook of Wireless Networks and Mobile Computing, pp. 325–346. Wiley, New York (2002)
2. Tarique, M., et al.: Survey of multipath routing protocols for mobile ad hoc networks. J. Netw. Comput. Appl. **32**(6), 1125–1143 (2009)
3. Manjunath, M., Manjaiah, D.H.: Spatial DSDV (S-DSDV) routing algorithm for mobile ad hoc network. In: International Conference on Contemporary Computing and Informatics, [IC3I—2014] IEEE, pp. 625–629. Sri S.J. College of Engineering, [SJCE], Mysore. ISBN: 978-4799-6629-5. 27–29 Nov 2014
4. Gunes, M., Sorges, U., Bouazizi, I.: ARA-the ant-colony based routing algorithm for MANETs. In: Proceedings of the International Conference on Parallel Processing Workshops, IEEE (2002)
5. Singh, G., Kumar, N., Verma, A.K.: Ant colony algorithms in MANETs: a review. J. Netw. Comput. Appl. **35**(6), 1964–1972 (2012)
6. Moussaoui, A., Boukeream, A.: A survey of routing protocols based on link-stability in mobile ad hoc networks. J. Netw. Comput. Appl. **47**, 1–10 (2015)
7. Perkins, C., Belding-Royer, E., Das, S.: Ad hoc on-demand distance vector (AODV) routing. Internet RFCs, vol. 285. pp. 1–38 (2003)
8. Perkins, C.E., Bhagwat, P.: Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. ACM SIGCOMM Comput. Commun. Rev. **24**(4), 234–244 (1994)
9. Marina, M.K., Das, S.R.: Ad hoc on-demand multipath distance vector routing. Wireless Commun. Mob. Comput. **6**(7), 969–988 (2006)
10. Correia, F., Vazão, T.: Simple ant routing algorithm strategies for a (Multipurpose) MANET model. Ad Hoc Netw. **8**(8), 810–823 (2010)
11. Mohanapriya, M., Krishnamurthi, I.: Modified DSR protocol for detection and removal of selective black hole attack in MANET. Comput. Electr. Eng. **40**(2), 530–538 (2014)
12. Johnson, D., Hu, Y., Maltz, D.: The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4, vol. 260. RFC 4728 (2007)
13. Wang, J., et al.: HOPNET: a hybrid ant colony optimization routing algorithm for mobile ad hoc network. Ad Hoc Netw. **7**(4), 690–705 (2009)
14. Di Caro, G., Dorigo, M.: AntNet: distributed stigmergetic control for communications networks. J. Artif. Intell. Res. (JAIR) **9**, 317–365 (1998)
15. Karia, D.C., Godbole, V.V.: New approach for routing in mobile ad-hoc networks based on ant colony optimisation with global positioning system. IET Netw. **2**(3), 171–180 (2013)
16. Manjunath, M., Manjaiah D.H.: Spatial DSDV (S-DSDV) routing algorithm for mobile Ad Hoc network. In: International Conference on Contemporary Computing and Informatics,

[IC3I—2014] IEEE, pp. 625–629. Sri S.J. College of Engineering, [SJCE], Mysore. ISBN: 978-4799-6629-5. 27–29 Nov 2014
17. Karthikeyan, D., Dharmalingam, M.: Ant based intelligent routing protocol for MANET. In: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), IEEE (2013)
18. Manjunath, M., Manjaiah D.H.: Spatial AODV (S—AODV) routing algorithm for mobile Ad Hoc network, pp. 799 –804. Elsevier Publications, ISBN: 978935 1072638. Aug 2014
19. Biswas, T., Dutta, R.: Spatially diffuse pathsets for robust routing in Ad Hoc networks. In: Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE (2011)
20. Simulator, Network: ns-2 (1989)

A Novel Framework for Data Processing and Computation of Wireless Sensor Networks on Cloud

Payal Pahwa, Deepali Virmani, Akshay Kumar and Sahil

Abstract This paper proposes a novel framework to combine two emerging technologies—wireless sensor networks and cloud computing. One of the major limitations of any wireless distributed system is energy conservation. The proposed framework tackles the energy problem in wireless sensor networks by combining it with cloud. The proposed framework aims to shift the processing and computational overhead of the wireless sensor network towards the cloud, by utilizing the cloud for routing and providing a platform to carry out complex processes such as clustering. The worst case of the proposed framework is compared with the best case of existing scenario. The results prove the furtherance of the proposed framework over the existing frameworks by reducing the energy consumption to 63.7 % and in-directly enhancing the network lifetime.

Keywords Wireless sensor networks · Cloud · Framework · Energy consumption · Cluster heads

1 Introduction

In the past decade Wireless Sensor Networks [1] have been showing a lot of promise and are under rampant development. Today these networks find application in environmental sensing, health care, area monitoring etc. But the development of wireless sensor networks face hurdles which are caused by their own unique properties. There is a limit to what can be achieved from these systems considering their characteristics such as low power, limited computation ability and limited

P. Pahwa (✉) · A. Kumar · Sahil
Department of CSE, Bhagwan Parshuram Institute of Technology,
New Delhi 110089, India

D. Virmani
Department of IT, Bhagwan Parshuram Institute of Technology,
New Delhi 110089, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_21

communication ability. Wireless sensor networks largely find application in temporally sensitive purposes such as intrusion detection in homeland security, climate monitoring, forecasting tsunamis or sensing earthquakes. This calls for the wireless sensor networks to sense and report data with real time reliability, which is a challenge for small enterprises which use their own wireless sensor network infrastructure. In this scenario cloud computing can emerge as the solution.

As Cloud computing may become ubiquitous in future [2, 3], researchers are considering how cloud computing can solve the issues faced by wireless sensor networks. The combination of the two can overcome the shortcomings of processing power, storage and connectivity, as cloud offers the weak characteristics of wireless sensor networks as its strengths.

In this paper, a novel framework is proposed for the combination of cloud with wireless sensor networks aimed at increasing the network lifetime of wireless sensor networks..

2 Related Work

In [2] paper, a pub/sub based framework of the combination paradigm has been presented and discussed about the advantages and disadvantages including the security challenges of the new paradigm. Solutions of the security problems faced in this paradigm are also discussed. In [3] paper, upcoming research topics relating to cloud computing and distributed systems are discussed. In [4] paper, the time synchronization problem and need for synchronization in sensor networks has been reviewed and the basic synchronization methods explicitly designed for sensor networks have been discussed in detail. In [5] paper, an energy efficient and reliable data acquisition approach for time-critical and real time traffic in wireless sensor networks has been discussed. In [6] paper, Cluster Head selection protocol using Fuzzy Logic (CHUFL) has been presented. In [7] paper, opportunities in cloud computing were presented. It also presents the comparison on why data should be stored on cloud. In [8] paper, comparison of simulation results of two different deployment models has been discussed using MICAz mote [9]. An Energy Consumption Model is proposed.

3 Proposed Framework

Taking care of the existing short comings of wireless sensor networks, this paper introduces a new framework that is reliable and energy efficient. This framework links the wireless sensor networks to the cloud to reduce the processing overhead

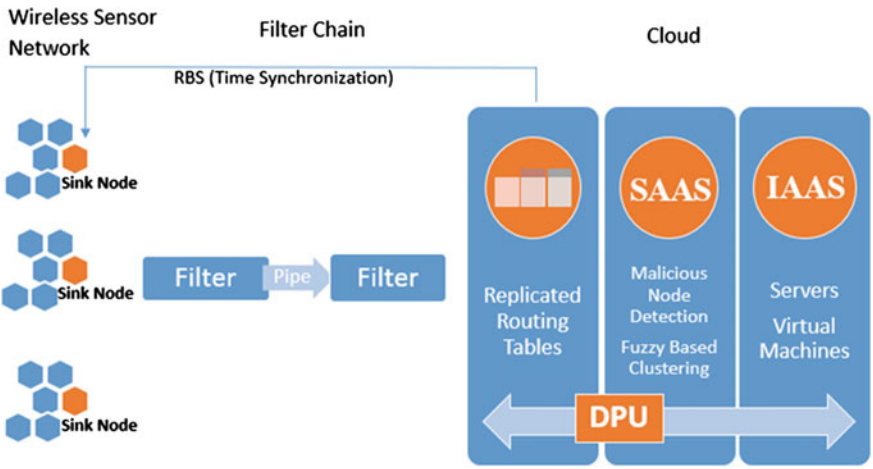


Fig. 1 Proposed framework

and provide hard real time computation. The proposed framework will not only be instantly deployable but can adapt to any hostile environment and can also be applicable to preconfigured networks.

As shown in the figure Fig. 1 our framework is broadly divided into 3 parts-wireless sensor network, filter chains [2] and cloud [10].

The wireless sensor network consists of clusters of nodes, which communicate through a cluster head (sink node).

The sink node selection is based on fuzzy logic [6]. After the sink nodes are determined, they send selective information through a chain of pipes and filters which further process the data which is to be sent to the Cloud through a gateway. A gateway aids in the transmission of data by compression at the source and decompression at the receiver.

The cloud consists of the following:

1. Replicated Routing Tables
2. SAAS: Consists of all the services which a wireless sensor network can utilize through cloud such as keeping track of malicious nodes [11, 12], selection of cluster heads, management of nodes etc.
3. Data Processing Unit (DPU): The DPU performs computations using the routing tables and algorithms in SAAS [12].

The cloud's processing power and memory enables us to use AI based algorithms to take care of essential processes in a wireless sensor network such as clustering. This will greatly reduce the processing overhead on sensors.

This framework utilizes Reference Broadcast Synchronization [4] for time synchronization in the wireless sensor network and cloud.

4 Comparison

In a wireless sensor network without cloud, for a node *x*1 to transmit data to another node y1 the following chain of events take place:

1. *x*1 (*node*), *X* (*cluster head*) *x*1 → *X*
2. *X forwards the data to Y* (*cluster head*) *by following a path determined by a routing algorithm, which involves multiple hops* via *other cluster heads.*
3. *Y* → *y*1(*destination node*)

Figure 2 shows the transmission of data in a wireless sensor network without cloud. In this method the multiple hops among cluster heads in step 2 consume substantial energy from the sensor network. [8] shows that when data packets have to be sent from a source node (N_n) to a destination node (N_1), data packets emerging from node N_n have to be retransmitted at every intermediate node N_{n-1}, N_{n-2}, …N_2 before reaching the destination node N_1. On each instance of packet transmission from one node to another, the total energy consumed by a single node is the sum of its transmission energy, reception energy and transition energy. Transmission energy is the energy consumed to transmit a packet, reception energy is the energy consumed while receiving a packet and transition energy is the energy consumed to change the node's energy state. A node can be in one of these energy states: shutdown, idle, reception or transmission. For each hop in the transmission of packets from cluster head X to cluster head Y, intermediate nodes have to change

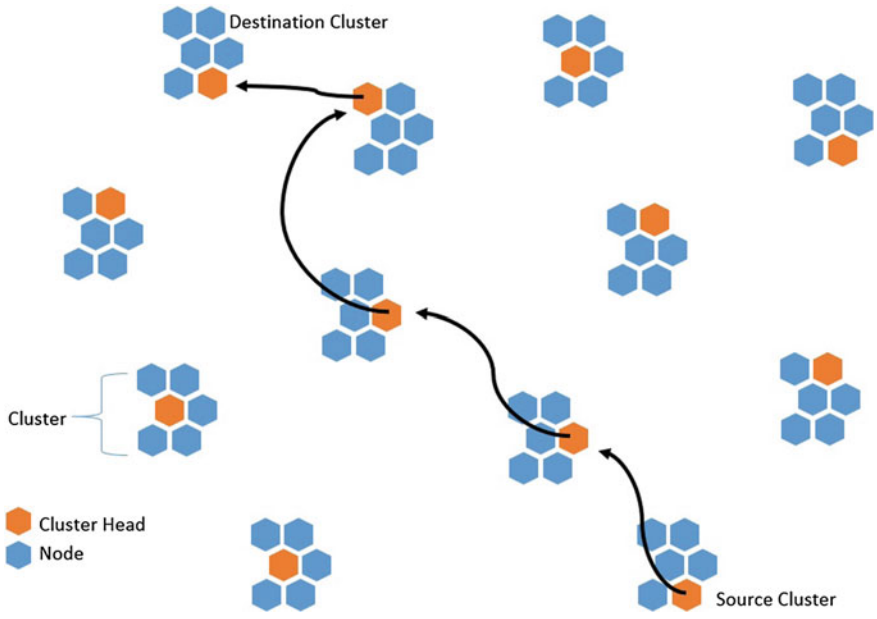


Fig. 2 Transmission of data in wireless sensor network

states at least once and forward the packets. Nodes which could be in idle or shutdown state have to participate in the transmission, thereby consuming energy. As intermediate nodes also have to spend energy, by taking part in transmissions meant for other nodes, the net total energy consumption of the sensor network increases.

At any instant in the wireless sensor network an intermediate node could be transmitting its own data packets, rendering it responsible for multiple data transmissions. This can introduce bulk traffic in the network which makes it vulnerable to collisions and packet loss. Retransmission of lost packets will add to energy demand in the network.

In our framework for cloud integrated wireless sensor network, node to node communication takes place through the cloud as shown below:

1. *X* (*cluster head*), *x*1 (*source node*) *x*1→ *X*
2. *X* → *CLOUD*
3. *CLOUD* → *Y* (*cluster head*)
4. *Y* → *y*1 (*destination node*)

Figure 3 shows the transmission of data in the proposed framework. In the proposed framework, Cloud acts as the lone intermediate node to carry out transmissions between two cluster heads. With cloud as the single intermediate node, hops are minimized to 2. Energy consumption at the cloud does not affect the energy of the sensor network, only the nodes belonging to source and destination clusters have to spend energy during packet transmissions. The cloud in this

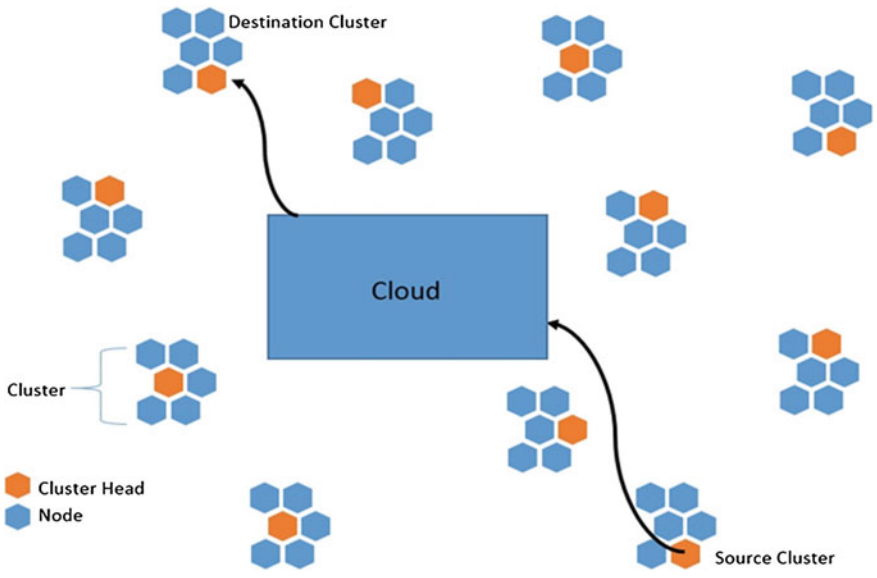


Fig. 3 Transmission of data in the proposed framework

framework handles the network traffic to minimize packet loss, which frees up the bandwidth and increases throughput for each cluster. The reduced energy consumption added with higher bandwidth and throughput of our framework ensures a network lifetime which is significantly larger than that of a wireless sensor network without cloud.

5 Results

Change of states in head nodes involved in data transmission in Fig. 2 is shown in Table 1. Table 3 shows the energy used in changing states. Table 4 depicts the energy used for reception. Taking the best case in which signal power level P_{out} at output transceiver is least required Table 2 energy consumption of Fig. 2 is calculated in Table 5.

5.1 Energy Consumption Model

Total Energy consumption of a node as given in [8] can be calculated as

$$TotalEnergy_{node\ i} = \sum_{State\ j} \left(P_{state\ j} \times t_{state\ j}\right) + \sum Energy_{transitions} \quad (1)$$

In which $State_j$ refers to the energy state-idle, shutdown, reception or transmission taken by the mote. $P_{state\ j}$ refers to the power consumption in $state_j$, and $t_{state\ j}$ refers to the time spent in the state. $Energy_{transitions}$ is the total energy spent during transition among states.

Table 1 State transitions in head nodes in Fig. 2

N_1	N_2	N_3	N_4	N_5
Idle	Tx	Idle	Rx	Idle
⬇	⬇	⬇	⬇	⬇
Tx	Rx	Rx	Tx	Rx
	⬇	⬇		
	Tx	Tx		

Table 2 Energy values for transmission [8]. Packet length = 320 (bits) = 1.28 (ms) (320 bits @ 250 Kbps)

Case	P_{out} (dBm)	I_{tx} (mA)	P_{tx} (mW)	$E_{tx\text{-}1packet}$ (μJ)	E_{tx}/bit (μJ)
Worst	0	17.04	30.67	39.26	0.122
Best	−25	8.42	15.15	19.39	0.060

Table 3 Transition energy from Idle state [8]

States	Transition energy from idle state (μJ)
Shutdown	0.000691
Receiving	6.7
Transmission	5.8

Table 4 Energy values for reception [8]

I_{rx} (mA)	P_{rx} (μJ)	$E_{rx\text{-}1packet}$ (μJ)	E_{rx}/bit (μJ)
19,6	35,28	45,15	0,14

Table 5 Results of existing frameworks (Fig. 2)

$Energy_{transitions}$ (μJ)	$Energy_{transmission}$ (μJ)	$Energy_{reception}$ (μJ)	Total energy consumption (μJ)
43.3	18,060	7756	25,859.3

Table 6 State transitions in head nodes in Fig. 3

N_1	Cloud	N_5
Idle		Idle
⬇		⬇
Tx		Rx

Table 7 Results of proposed framework Fig. 3

$Energy_{transitions}$ (μJ)	$Energy_{transmission}$ (μJ)	$Energy_{reception}$ (μJ)	Total energy consumption (μJ)
12.5	4515	3926	8453.5

Change of states in cluster heads involved in data transmission in Fig. 3 is shown in Table 6. Using the worst case in which signal power level at output P_{out} transceiver is maximum Table 3, energy consumption is calculated in Table 7.

Energy conserved = (25,859.3–8453.5) μJ = 17,405.8 μJ
Energy gained = 67.3 %

6 Conclusion

In this paper, a novel framework is presented which combines the two nascent technologies- wireless sensor networks and cloud computing. The proposed framework can significantly reduce energy consumption in wireless sensor networks. The processing and computation of wireless sensor networks is shifted to cloud using the proposed framework. The worst case of proposed framework is compared with the best case of existing scenario. The results prove the betterment of the proposal showing 67.3 % energy conservation and network lifetime enhancement.

References

1. Jena, S., Agarwal, P.: Thesis on energy aware multipath routing in wireless sensor networks national institute of technology rourkela (2013)
2. Xie, Y., Zhang, J., Fu, G., Wen, H., Han, Q., Zhu, X., Jiang, Y., Guo, X.: The security issue of WSNs based on cloud computing. Commun. Netw. Secur. IEEE (2013)
3. Birman, K., Chockler, G., Renesse, R.: Towards a cloud computing research agenda. ACM SIGCAT News **40**(2) (2009)
4. Sivrikaya, F., Yener, B.: Time synchronization in sensor networks: a survey. IEEE Netw. **18** (4), 45–50 (2004)
5. Soyturk, M., Altilar, D.T.: Reliable real-time data acquisition for rapidly deployable mission-critical wireless sensor networks (c), IEEE (2008)
6. Gajjar, S., Sarkar, M., Dasgupta, K.: Cluster head selection protocol using fuzzy logic for wireless sensor networks. Int. J. Comput. Appl. **97**(7), 0975–8887 (2014)
7. Armbrust, M., Fox, A., Griffith, R., Joseph, A., et al.: Above the clouds: a berkeley view of cloud computing. University of California, Berkeley (2009)
8. Barberis, A., Barboni, L., Valle, M.: Evaluating energy consumption in wireless sensor networks applications. In: 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools (DSD 2007) © 2007 IEEE (2007)
9. Crossbow: MICAz Datasheet: http://www.xbow.com
10. Demarest, G., Wang, R.: Oracle cloud computing. An oracle white paper, May 2010
11. Wazid, M., Katal, A., Sachan, R.S., Goudar, R.H., Singh, D.P.: Detection and prevention mechanism for Black hole attack in wireless sensor network. In: International Conference on Communication and Signal Processing, India (2013)
12. Virmani, D., Hemrajani, M., Chandel, S.: Exponential trust based mechanism to detect Black Hole attack in wireless sensor network. In: Proceedings of the International Conference on Computing, Informatics and Networks, pp. 59–63 (2014)

Online Signature Verification Using Single Spiking Neuron Model

K. Venkatanareshbabu and Pravati Swain

Abstract Integrate-And-Fire model is one of the latest neuron model. A biologically plausible Tonic Non-linear Integrate-And-Fire neuron model (TIFN) is proposed in this paper. The complete solution of the proposed model is derived and used for the construction of aggregation function in Multi-layer perceptron model, which is named as Tonic Single Neuron model (TSN). This modified neuron model is used for verification of online signature data. The comparison study of the Tonic Neuron model and the well-known Single Integrate-And-Fire neuron model (SIFN) are presented. It is observed that the proposed model perform well in terms of classification accuracy. Moreover, It is found that inclusion of biological phenomenon in an artificial neural network makes efficient for biometric authentication.

Keywords Biological neuron models · Back propagation · Single Integrate-And-Fire neuron model

1 Introduction

The human brain consists of millions of neurons which are connected and transfer information through neural networks. Many computational biologically realistic mathematical models have been proposed for the efficient simulation of a biological neuron as well as mathematical description for the generation of action potential. Integrate-And-Fire neuron model is designed using an electric circuit consists of a parallel capacitor and resistor [2]. A class of nonlinear ordinary differential equations was proposed which approximates the electrical characteristics of excitable

K. Venkatanareshbabu (✉) · P. Swain
Department of Computer Science and Engineering,
National Institute of Technology Goa, Goa 403401, India
e-mail: venkatanaresh@nitgoa.ac.in

P. Swain
e-mail: pravati@nitgoa.ac.in

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_22

cells such as neurons and explains the ionic mechanisms of action potentials. Leaky Integrate-And-Fire neuron model, quadratic model [4] and Exponential Integrate-And-Fire neuron model [3] were reproduced as special cases of generalized Non-Linear Integrate-And-Fire neuron model. For a spike with sharp intention, Exponential Integrate-And-Fire neuron model is equivalent to leaky Integrate-And-Fire neuron model.

A biologically realistic artificial neuron model has been proposed whose weighted aggregation function is calculated from the complete solution of a proposed Non-linear Integrate-And-Fire model. The sequence of timings of the spikes comprises neuronal information rather than the geometrical shape of the action potential. Hence the aggregation function is rendered from the map between injected current and Inter-spike interval of the proposed Non-linear Integrate-and-Fire neuron model. Multilayer perceptron [5] has been used as learning algorithm. Only a single neuron is sufficient for the network architecture. The effectiveness of the proposed neuron model has been shown by comparing with the performance of single Integrate-And-Fire neuron model (SIFN) [6] on online signature verification. It has been observed that the proposed model reduces the computational time and improves the classification accuracy. It is noticed that the inclusion of biological phenomenon in an artificial neural network can make it robust for classification.

Biometric security is aimed at method of verifying a person's identity based on his/her body and/or physical attributes. Various forms of biometric security exist including fingerprint, iris, speech, heart sound and keystroke based recognition. Owing to wide use of signatures, there is an inevitably bigger risk of people performing forgeries. The design of an automatic signature verification system has to be tested regarding its robustness to forgeries. Therefore, there is a clear emphasis in the field towards research on online methods of signature verification.

Organization of the paper is as follows. Section 2 presents an overview of neuron models discussed in the literature. Section 3 presents the mathematical derivation of proposed neuron model and its complete solution is used to prepare aggregation for proposed tonic single neuron model. Section 4 deals with application of proposed network onto online signature verification. In Sect. 5 simulation results obtained from proposed neuron model and single Integrate-And-Fire neuron model on online handwritten signature verification are discussed. Conclusions are given in Sect. 6.

2 Biological Neuron Models

Various mathematical models for biological neurons have been designed in the literature to represent biological activities of a neuron. This section deals generalized nonlinear IFN model.

2.1 Non-linear Integrate-and-Fire Neuron Model

A generalized Non-linear Integrate-And-Fire neuron model [1] is denoted by

$$\tau_m \frac{dv}{dt} = P(v) + Q(v)I^{ext} \tag{1}$$

The dynamics of membrane potential v are stopped if "v" reaches the threshold v_θ then reinitialized at $v = v_{rest}$.

A novel Tonic Non-linear Integrate-And-Fire (TIFN) neuron model has been proposed. Its solution has been used for online signature verification.

3 Proposed Tonic Non-linear IFN Model (TIFN)

The proposed tonic Non-linear Integrate-And-Fire neuron model is represented by

$$\tau \frac{dv}{dt} = a\frac{v}{1+v^2} + RI(t) \tag{2}$$

where "v" denotes the membrane potential, I denotes the external current, τ denotes membrane time constant, R denotes the input resistance

$\frac{v}{1+V^2}$ denotes a non-linear function of membrane potential. The dynamics of membrane potential v are interrupted if v reaches the threshold v then reinitialized at $v = v_{rest}$

Equation (2) is solved by using variable and separable method as given by

$$\frac{dv}{a\frac{v}{1+V^2} + RI(t)} = \frac{dt}{\tau} \tag{3}$$

The solution is given by

$$\frac{(a)^2 \tan^{-1}\left(\frac{a+2RIv}{\sqrt{4R^2I^2-a^2}}\right)}{R^2I^2\sqrt{4R^2I^2-a^2}} - \frac{a\log(av + RIv^2 + RI)}{2(RI)^2} + \frac{v}{RI} = \frac{t}{\tau} + Cons \tag{4}$$

Equation (4) represents the Interspike interval and the frequency f is given by

$$f = \frac{1}{\tau\left(\frac{(a)^2 \tan^{-1}\left(\frac{a+2RIv}{\sqrt{4R^2I^2-a^2}}\right)}{R^2I^2\sqrt{4R^2I^2-a^2}} - \frac{a\log(av+RIv^2+RI)}{2(RI)^2} + \frac{v}{RI}\right)} \tag{5}$$

Spikes can be observed in response to the proposed Non-linear Integrate-And-Fire neuron model.

3.1 Proposed Tonic Single Neuron Model

Biological neurons exchange neuronal information [7] based on the timing of the spikes generated by the action potential. Motivated from the non-linear map between injected current and Interspike interval of proposed Integrate-And-Fire neuron model as shown in Eq. (4), following aggregation function is assumed instead of the weighted sum of a traditional neuron

$$\text{net} = \prod_{i=1}^{n} \left(\left(p_i tan^{-1}(q_i + r_i x_i) - s_i \log\left(t_i x_i + u_i(x_i^2 + 1)\right) \right) + v_i x_i \right) \tag{6}$$

where x_i denotes ith input variable, n denotes number of input neurons and p_i, q_i, r_i, s_i, t_i and u_i denotes weight parameters.

In comparison with Eq. (4), $\frac{a^2}{R^2I^2\sqrt{4R^2I^2-a^2}}$ is denoted as p_i, $\frac{a}{\sqrt{4R^2I^2-a^2}}$ is denoted as q_i, $\frac{2RI}{\sqrt{4R^2I^2-a^2}}$ is denoted as r_i, $\frac{a}{2(RI)^2}$ is denoted as s_i, a is denoted as t_i, RI is denoted as u_i and $\frac{1}{RI}$ is denoted as v_i.

Proposed neuron model comprises of an input layer and a single neuron in the output layer without using any hidden layer. Weighted product of input neurons is used as aggregation to the single output neuron as shown in Eq. (6). To minimize the error function, Gradient descent rule is applied for the patterns in the training dataset. If t denotes the target and y denotes the actual output of the neuron then the error function is given by Eq. (7).

$$E = \frac{1}{2}(y - t)^2 \tag{7}$$

In gradient descent, weight parameters are modified in proportion to the negative of the error derivative with respect to each parameter p_i, q_i, r_i, s_i, t_i, u_i and v_i as shown below.

$$\Delta d = - \in \frac{\partial E}{\partial d} \tag{8}$$

where d varies from p_i, q_i, r_i, s_i, t_i, u_i and v_i, ϵ denotes the learning parameter. Partial derivatives of the Error function with respect to each parameter p_i, q_i, r_i, s_i, t_i, u_i and v_i is given by the following equations.

$$\text{Let } Z = \left(\left(p_i \tan^{-1}(q_i + r_i x_i) - s_i \log\left(t_i x_i + u_i(x_i^2 + 1)\right) \right) + v_i x_i \right)$$

$$\frac{\partial E}{\partial p_i} = (t - y)y(1 - y)(\text{net})\left(\frac{1}{Z}\right)\left(\tan^{-1}(q_i + r_i x_i)\right) \tag{9}$$

$$\frac{\partial \mathrm{E}}{\partial \mathrm{q_i}} = (\mathrm{t} - \mathrm{y})\mathrm{y}(1 - \mathrm{y})(\mathrm{net})\left(\frac{1}{\mathrm{Z}}\right)\left(\frac{1}{1 + (q_i + r_i x_i)^2}\right) \tag{10}$$

$$\frac{\partial \mathrm{E}}{\partial \mathrm{r_i}} = -(\mathrm{t} - \mathrm{y})\mathrm{y}(1 - \mathrm{y})(\mathrm{net})\left(\frac{1}{Z}\right)\left(\frac{r_i}{1 + (q_i + r_i x_i)^2}\right) \tag{11}$$

$$\frac{\partial \mathrm{E}}{\partial s_\mathrm{i}} = -(\mathrm{t} - \mathrm{y})\mathrm{y}(1 - \mathrm{y})(\mathrm{net})\left(\frac{1}{\mathrm{Z}}\right)(\log(\mathrm{t_i x_i} + u_i(x_i^2 + 1))) \tag{12}$$

$$\frac{\partial \mathrm{E}}{\partial t_\mathrm{i}} = -(\mathrm{t} - \mathrm{y})\mathrm{y}(1 - \mathrm{y})(\mathrm{net})\left(\frac{1}{Z}\right)\left(\frac{x_i}{(\mathrm{t_i x_i} + u_i(x_i^2 + 1))}\right) \tag{13}$$

$$\frac{\partial \mathrm{E}}{\partial u_\mathrm{i}} = -(\mathrm{t} - \mathrm{y})\mathrm{y}(1 - \mathrm{y})(\mathrm{net})\left(\frac{1}{Z}\right)\left(\frac{(x_i^2 + 1)}{(\mathrm{t_i x_i} + u_i(x_i^2 + 1))}\right) \tag{14}$$

$$\frac{\partial \mathrm{E}}{\partial v_\mathrm{i}} = -(\mathrm{t} - \mathrm{y})\mathrm{y}(1 - \mathrm{y})(\mathrm{net})\left(\frac{1}{Z}\right)(x_i) \tag{15}$$

Parameters are updated as follows

$$\mathrm{y_i}(\mathrm{k} + 1) = \mathrm{y_i}(\mathrm{k}) + \Delta \mathrm{y_i} \tag{16}$$

Here y is variable that is denoted as p, q, r, s, t, u and v.

4 Application to Online Signature Verification

Handwritten signature verification is a biometric authentication technique to verifying a signature by capturing information about shape of the signature (i.e., static features) and the characteristics of how the person signs his/her name in real-time (i.e., dynamic features). There are two broad approaches to signature verification: Offline and Online modes. In offline signature verification, the digital image of the signature is used as input to the algorithm. The main drawback of this method is lack of dynamic information which makes this approach very poor for fraud detection. In online signature verification, dynamic features of the process of drawing are also considered. A feature contains information about x-coordinate, y-coordinate, timestamp, adjusted pressure of pen tip and its angle with vertical etc.

4.1 Implementation

Data acquisition, pre-processing and feature extraction are significant steps for implementing the proposed method. The data acquired for the signature recognition is from an online source [7]. It contains signature data of 100 persons with 20 original and 20 forged signatures for each.

4.1.1 Feature Extraction

The following features are used for signature recognition. The features are token length, mean and standard deviation values of observation components and trend coefficients.

(1) ***Token length***; i.e., the number of token's points ***N***.
(2) ***Average values*** of observation components

$$\overline{s_j} = 1/N \sum_{k=1}^{N} s_j(k)\,;\; j = 1,\, 2,\, 3 \tag{27}$$

(3) ***Standard deviations*** of observation components

$$\sigma_j = \sqrt{\overline{s_j^2} - \overline{s_j}^2}\,; \quad \mathrm{j} = 1, 2, 3 \tag{28}$$

(4) ***Trend coefficients***; i.e., the slopes α_j of trend lines of each observation component. The trend lines are given by

$$s_j^0(k) = \overline{s_j} - \alpha_j k\,; \quad \mathrm{j} = 1, 2, 3 \tag{29}$$

where

$$\begin{gathered} \alpha_j = \rho_j \sigma_j / \sigma_k,\; \rho_j = (\overline{ks_j} - \bar{k}\overline{s_j})/\sigma_j \sigma_k,\; \overline{ks_j} = 1/N \sum_{k=1}^{N} k s_j(k) \\ \sigma_k^2 = \overline{k^2} - \bar{k}^2,\; \bar{k} = \frac{1}{2} N(N+1),\; \overline{k^2} = \frac{1}{6} N(N+1)(2N+1) \end{gathered} \tag{30}$$

The feature vector has 12 components. The first three of them are the token sizes of angle, x and y. It is the same for all the three, i.e. angle, x and y coordinates. The second three of them are the means of angle, x and y coordinates. The next three of them are the variances of angle, x and y. The last three of them are the trend coefficients of angle, x and y. The feature vector is:

$$\mathrm{Fv} = [\mathrm{size(x)},\, \mathrm{size(y)},\, \mathrm{size}(\theta),\, \mathrm{avg(x)},\, \mathrm{avg(y)},\, \mathrm{avg}(\theta),\, \mathrm{var(x)},\, \mathrm{var(y)},\, \mathrm{var}(\theta),\, \mathrm{tr(x)},\, \mathrm{tr(y)},\, \mathrm{tr}(\theta)].$$

5 Results

This section explains the performance evaluation of the proposed tonic single neuron model and compared with the single-integrate-neuron model [7] on online signature dataset [8]. p_i, q_i, r_i, s_i, t_i, u_i and v_i are random values drawn from uniform distribution on the open interval (0, 1). We now provide 5 genuine and 5 forged signatures per user for the training of tonic Single Integrate-And-Fire neuron model. We use five genuine signatures and five forgeries to train the neural network and fifteen genuine signatures and fifteen forgeries to test the neural network.

The forgeries are used in the training to make the neural network understand the difference. The same person's forgeries are also given. Figure 1 shows sample three out of five Genuine Signature for User 1 and Fig. 2: Sample three out of five Forgeries signature for User 1. Figure 3 show that 14 out of 15 genuine signatures were accepted and 1 one was rejected. Figure 4 indicates that for the user1 using TSN, all the forgeries are detected correctly and hence rejected. Figure 5 shows that for user1 using SIFN algorithm, the recognition rate is 100 %. But, one forgery as shown in Fig. 6 is also falsely accepted. For real systems, this is unacceptable as the false acceptance rate should be very low. Hence it has been concluded that TSN is far superior in comparision with SIFN.

It is observed that the number of training iterations needed to obtain MSE of order 0.0001 is lesser for the proposed neuron model when compared with Single Integrate-And-Fire (SIFN) model for signature verification dataset. Experiment results were carried out on 200 users. For 92 % users, proposed TSN is exactly

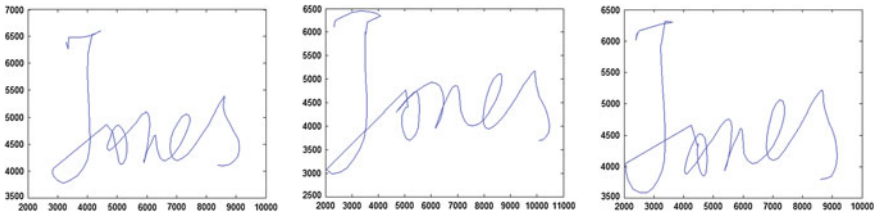


Fig. 1 Sample three out of five genuine signature for user 1

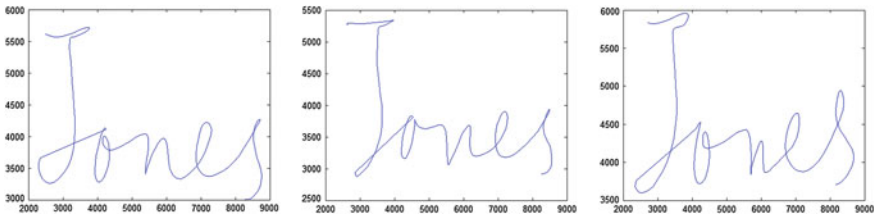


Fig. 2 Sample three out of five forgeries signature for user 1

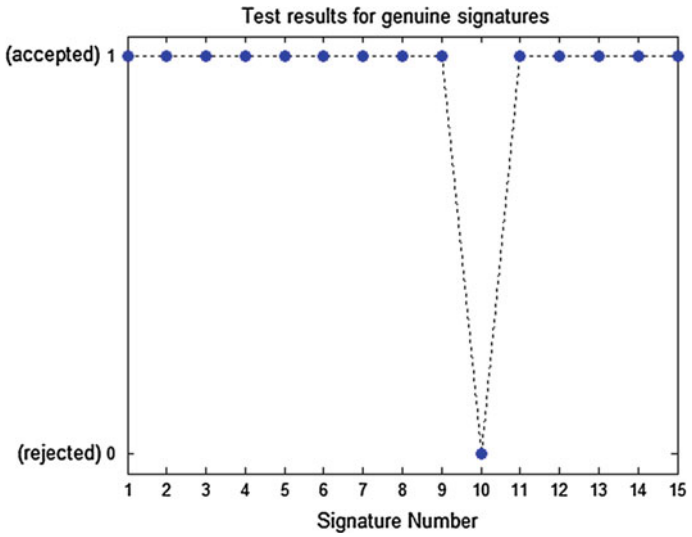


Fig. 3 Genuine signature results (TIFN) for user 1

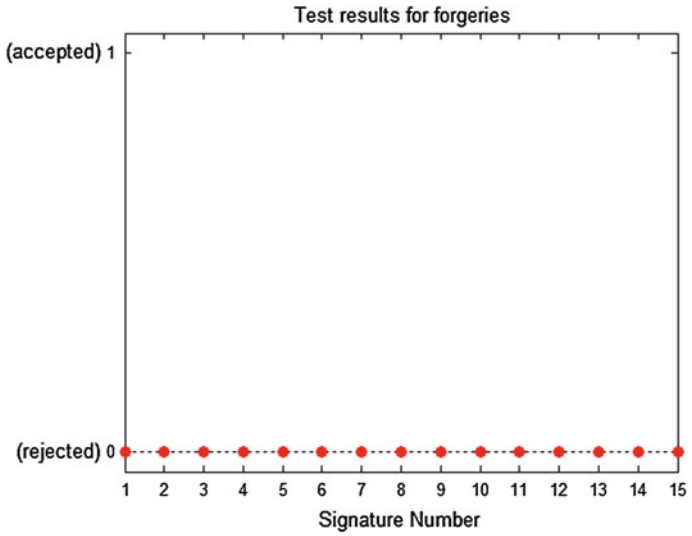


Fig. 4 Forgery results (TIFN) for user 1

detected the forgeries where as SIFN algorithm works well on only 76 % users. It is found that a single neuron in output layer for the proposed neuron model is sufficient for the applications that need a number of neurons in different hidden layers of a traditional neural network.

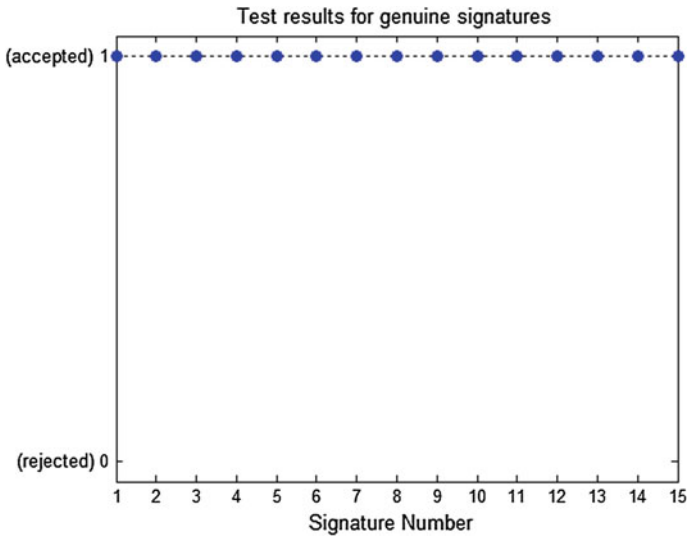


Fig. 5 Genuine signature results (SIFN) for user 1

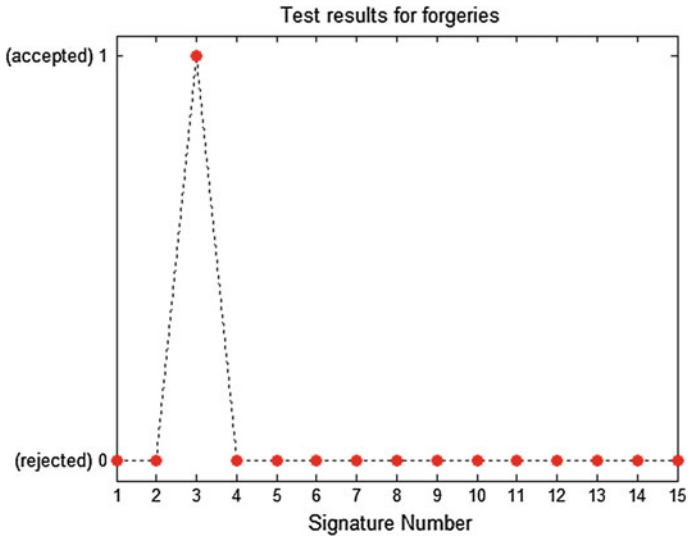


Fig. 6 Forgery results (SIFN) for user 1

6 Conclusions

This paper proposes a new tonic single neuron model (TSN) whose aggregation function is designed from the complete solution of a proposed single Integrate-And-Fire neuron model (SIFN). The concept of neuronal information

exchanged from one neuron to other depending on the timing of the spikes generated has been used to design the proposed neuron model. Training of Online signature verification dataset shows that a single neuron in the output layer is capable of performing classification task better in terms of testing accuracy as compared with single Integrate-And-Fire neuron model.

References

1. Abbott, L.F., van Vreeswijk, C.: Asynchronous states in a network of pulse-coupled oscillators. Phys. Rev. E **48**, 1483–1490 (1993)
2. Abbott, L.F.: Lapique's introduction of the integrate-and-fire model neuron (1907). Brain Res. Bull. **50**(5/6), 303–304 (1999)
3. Brette, R., Gerstner, W.: Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. J. Neurophysiol. **94**, 3637–3642 (2005)
4. Ermentrout, G.B., Kopell, N.: Parabolic bursting in an excitable system coupled with a slow oscillation. SIAM J. Appl. Math. **46**, 233–253 (1986)
5. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Nature **323**, 533–536 (1986)
6. Yadav, A., Mishra, D., Ray, S., Yadav, R.N., Kalra, P.K.: Learning with single integrate-and-fire neuron. In: Proceedings of IEEE International Joint Conference on Neural Networks, vol. 4, pp. 2156–2161 (2005)
7. www.cse.ust.hk/svc2004/index.html

Comparative Study of Classification Algorithms for Spam Email Detection

Aakanksha Sharaff, Naresh Kumar Nagwani and Abhishek Dhadse

Abstract Spam in emails has become a major issue. Spam messages consume space, network bandwidth and are of no use to the receiver. It is very difficult to filter spam as spammers try to tackle the processes carried out by the filtering mechanism. Various classification algorithms are used to classify a mail as spam or non-spam (ham). The present paper compares and discusses the effectiveness of four machine learning classification algorithms, belonging to different categories (Probabilistic, Decision Tree, Vector Machines and Lazy Algorithms) on the basis of various performance measures, using WEKA, a data mining tool to analyze different algorithms. Enron dataset is taken in a processed form from Athens University of Economics and Business and it is found that J48 and BayesNet algorithms perform better than SVM.

Keywords Email classification · Performance measures · Spam · Email spam filtering

1 Introduction

With the advent of email, communication has become a lot easier than before. Emails have become an important part of communication in any organization among the world. Originally, email was designed to create, send and receive messages from one another. However the increase in the popularity of emails has also resulted in spam e-mails over the course of time. Spam mails are the messages which are sent in bulk, i.e. sent to many users at the same time for commercial purposes. The sender of spam mails has no relation with the receivers. The spammers obtain a list of mails from various sources such as address books, so as to send mails. On clicking of spam mail, the spammers get benefit. These mails can also be infected, so upon clicking, these can be sent from your address to other

A. Sharaff (✉) · N.K. Nagwani · A. Dhadse
Department of CSE, National Institute of Technology, Raipur 492010, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_23

contacts in your list, or simply introduce viruses in your system as well. These can also contain phishing links [6] aimed to retrieve your passwords. Overall, this reduces a firm's productivity regarding communication, and is very risky regarding security [14]. Hence study of email and spam detection has become an important area in the field of computer science. Each message cannot be detected as spam or not by human alone because that would be very time consuming task. Hence classification algorithms are used for the purpose.

Many classification algorithms can be used for the task of text classification, such as probabilistic algorithms, decision tree algorithms, lazy algorithms etc. [8, 10]. Despite of so many algorithms, the task of complete accuracy is not yet fulfilled. So these algorithms need to be compared for performance with each other. In this paper, we have used four classification algorithms, namely BayesNet, SVM, J48 (C4.5) and Lazy IBK for comparison. Spam dataset is a processed dataset taken from the library of Natural Language Processing Group, Department of Informatics—Athens University of Economics and Business. These mails are analyzed using WEKA, a data mining tool from University of Waikato, New Zealand. In this work, the same dataset is used for training and testing in the form of folds. The contents of a mail are used to train the four above mentioned algorithms, which learn from the dataset [4, 5]. The results obtained are used to decide the performance of algorithms.

2 Related Work

Over the time, due to the risk factor involved in the reception of spam mails, many works have been carried out to counter these mails. Many spam classification systems are designed for this purpose. The goal of such systems is to check whether a mail is spam or not, and if it is spam, move it to another destination or folder. Many research works are carried out to search for the best classification algorithm. But it mainly depends on data. The four algorithms we have used are of four different categories, i.e. Probabilistic, Decision Tree, Lazy algorithms and Support Vector Machines. Many studies have been done to compare the performances of these algorithms, the classification algorithms are used to test whether a mail is spam or not and based on the results, the algorithms are ranked. The various performance measures included accuracy, TP rate, precision, recall and F-measure in in paper [1, 11].

Spam Classifiers are built and tested on publically available datasets for evaluation. For example, J48 is used for medical diagnosis, which decides whether a person has certain disease or not on the basis of his symptoms [2]. A study of Naive Bayes is done in [3] which consider six types of Naive Bayes on six datasets, derived from the original Enron corpus [12], and are made publically available, which contain proper ham to spam ratio and are more realistic than previous comparable benchmarks. In this paper, a comprehensive review of recent machine learning approaches to Spam filters was presented in which a quantitative analysis of the use of feature selection algorithms and datasets was conducted [16].

3 Dataset

Enron dataset [12] is taken in a processed form from Athens University of Economics and Business. The dataset used in our study contains ham and spam messages of particular users of Enron. The ratio between ham and spam is maintained. These messages contain body and different headers such as subject, sender etc. Out of which we have considered subject and body for our study.

4 Classification Algorithms

As mentioned before, we have used four algorithms, each belonging to different category.

1. **J48** is the open source java implementation of C4.5 algorithm which is an improvement over ID3 algorithm, and it works on data by splitting into different parts. Each node of the tree splits its instances into one part or other. It is based on an impurity measure, called entropy. The difference in entropy, called information gain is calculated and the node providing maximum information gain is used to split the data. It handles both continuous and discrete values. It handles missing values as well.
2. **SVM**, i.e. Support Vector Machines are the supervised machine learning models which work by classification and regression analysis. SVM considers data as points in space mapped in a way such that the difference between the closest data points is maximum [7]. New examples are then put into the graph depending on which side of the margin they fall on. A good separation has more distance between the closest data points, since larger margin indicates the lower generalization error.
3. **BayesNet**, which stands for Bayesian Network is a probabilistic classifier model which works on the data by creating directed acyclic graphs using their conditional dependencies. The conditional probabilities are associated with the words in the email. These are then used to determine a mail is spam or not. Because a Bayes net only relates nodes that are probabilistically related by some sort of causal dependency, it can save enormous amount of time.
4. **LazyIBK** belongs to the lazy class of algorithms [9]. It is also called the K Nearest Neighbour algorithm, which works by classifying an instance depending on the majority in the nearest neighbours [15]. It works directly on test-data. If the majority of nearest mails are spam, the new email is likely to be classified as spam. LazyIBK is the Weka [13] implementation of KNN algorithm. It may return more than k neighbours if there are ties in the distance. Lazy algorithms are suitable when instances are not available beforehand, but occur online one by one.

5 Performance Measures

The classification results into a confusion matrix, which consists of four parts, True Positive (TP), True Negative, (TN), False Positive (FP) and False Negative (FN) [11]. These values can be used to determine performance measures like Accuracy, FP rate, Precision, Recall and F-Measure. The following performance measures are used for comparison of algorithms-

1. Accuracy:
   Accuracy = (TP + TN)/(TP + TN + FP + FN)
   It tells how much classification is done correct. TP and TN together is the correct number of classifications done by the classifier. It does not consider positives and negatives separately and hence other measures are also required for the analysis other than accuracy.
2. FP Rate:
   FP Rate = FP/(FP + TN)
   It tells how the model has performed in detecting the negatives. A low FP rate is desirable as a model classifying positives as negatives is not desired.
3. Precision:
   Precision = TP/(TP + FP)
   It indicates how many instances, which are classified as positive, are actually relevant. A high precision is desirable because high relevancy in detecting positives is desired, i.e. less FP is desired.
4. Recall:
   Recall = TP/(TP + FN)
   It is also called as TP rate, and is an indication of how good a system can detect positives.
5. F-Measure:
   F-Measure = 2*(Precision*Recall)/(Precision + Recall)
   Since a high Precision and Recall is desired, hence high F-Measure is also desired.

6 Methodology

The methodology of our experimentation consists of four main parts, preparation of dataset, pre-processing, application of algorithms, and performance evaluation based on the above performance measures.

1. **Preparation of Dataset** The dataset needs to be in the proper format for applying machine learning algorithms. The dataset is converted into .arff file using weka library, with 4307 mails, out of which 635 are spam. The first attribute contains the filename, which is ignored later, second attribute consists

of the subject and body of the email, and the last attribute of the dataset is the nominal attribute, which consists of the value whether a mail is spam or not.

2. **Pre-Processing** The dataset is loaded into the WEKA tool and the first attribute is removed, as it does not contribute to spam detection. The second attribute, which is a string of words, has to be converted into a vector of words, which is done by string to word vector. Then the words are separate entities. These words are then passed through stopword removal, which consists of WEKA's list of stopwords, and then Snowball stemming algorithm is applied to the data.
3. **Application of Algorithms** The algorithms discussed above are applied with the use of filtered classifier. The split between training and test data set is done using 10-Fold Cross validation.
4. **Performance Evaluation** The results obtained are then compared on the basis of the performance measures discussed above. The results obtained are shown below.

A. **Accuracy**
In terms of accuracy, BayesNet and J48 performed better than SVM and Lazy IBK. This indicates a deviation in the existing trend that SVM performs better than most of the algorithms. Here, J48 has the highest accuracy, followed by BayesNet and lazyIBK while SVM has the lowest accuracy. Figure 1 shows the accuracy chart over different classifiers.

B. **Precision and Recall**
The same type of result is obtained when precision and recall are considered shown in Fig. 2. BayesNet performs the best, followed by J48 and Lazy IBK while SVM has the lowest precision and recall values. This is also against the existing norms. Usually SVM performs the best amongst all the above mentioned algorithms.

C. **TP Rate and FP Rate:**
In terms of TP Rate, J48 has the highest value, followed by BayesNet and LazyIBK and again SVM has the lowest value of all as shown in Fig. 3. It is

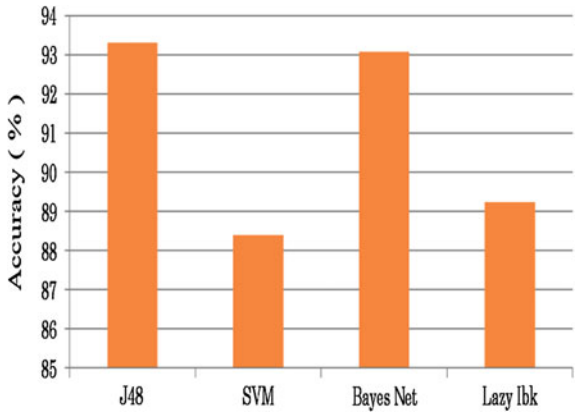


Fig. 1 Accuracy

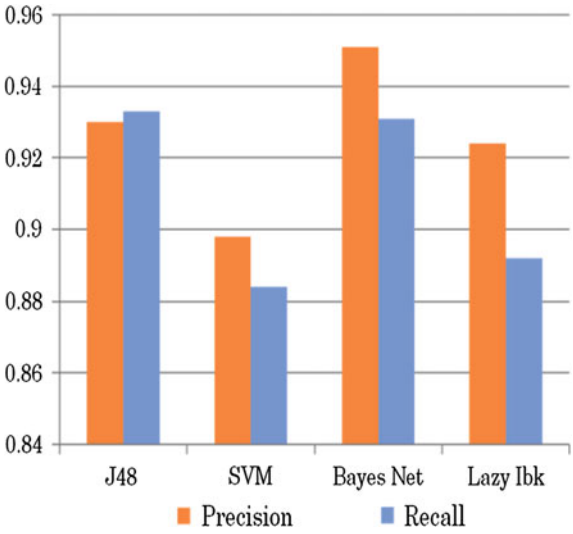


Fig. 2 Precision and recall

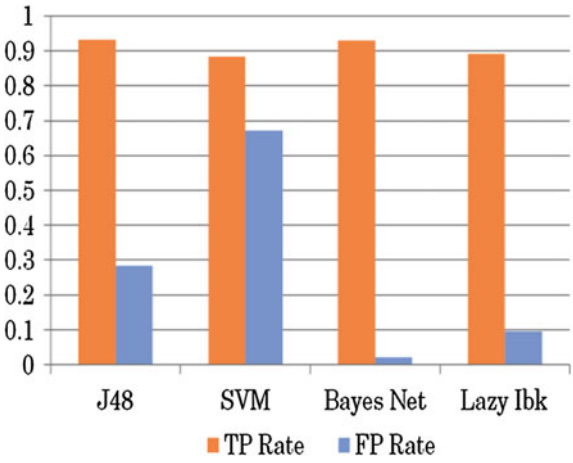


Fig. 3 TP rate and FP rate

equivalent to Recall. While in terms of FP Rate, Bayesnet again performs the best, followed by LazyIBK and J48, while SVM is found to have exceptionally high FP Rate.

D. **F-Measure**

F-Measure is dependent on Precision and Recall, hence the result is almost the same as above obtained results, i.e. BayesNet performs best followed by J48 and Lazy IBK and after that SVM. Figure 4 shows the F-measure performance of different classifiers

E. **Overall Results**

Table 1 describes the overall performance of different classifiers (J48, SVM, BayesNet, LazyIBK)

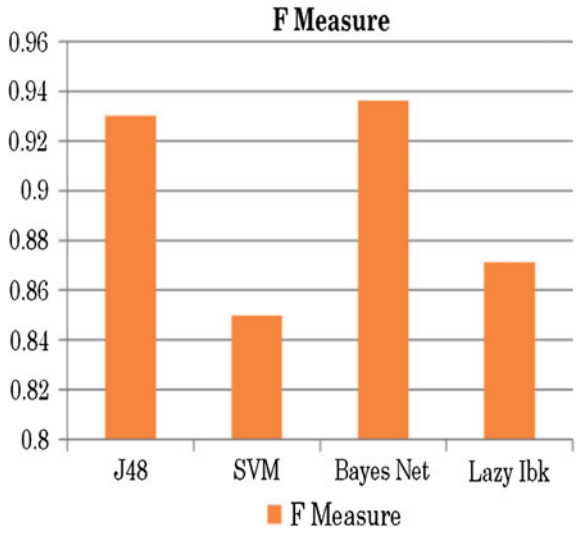


Fig. 4 F-measure

Table 1 Describes the overall performance of different classifiers (J48, SVM, BayesNet, LazyIBK)

Classification algorithms	Accuracy (in %)	Precision	Recall	FP Rate	F-measure
J48	93.3132	0.93	0.933	0.284	0.93
SVM	88.391	0.898	0.884	0.671	0.85
BayesNet	93.081	0.951	0.931	0.022	0.936
LazyIBK	89.236	0.924	0.892	0.095	0.901

7 Conclusion

In this paper, four classification algorithms, belonging to four different categories are compared against various performance measures. The results obtained indicate a deviation from the pre-established norms, i.e. SVM is one of the best classification algorithms, but here the results indicate otherwise. BayesNet and J48 perform the better than SVM. This deviation indicates that the performance of algorithm depends on data more than the algorithms. LazyIBK since belongs to Lazy algorithms, works at prediction time, i.e. on test data, hence it cannot be used as a basis to judge the performance of SVM. But J48 and BayesNet are having better results indicate that the algorithm performance depends on data. In future, we would like to optimize the classification time by the use of sub classifiers and then evaluate the performance of various classification algorithms. This can be done by using incremental sub-classifiers. Each sub-classifier represents a small part of a typical classifier and hence takes less time. This will improve the classification results which will be compared again with the existing results.

Acknowledgments Aakanksha Sharaff, Naresh Kumar Nagwani would like to thanks National Institute of Technology, Raipur, India for providing their support, imparting their knowledge, and making all the required facilities available for completing this research.

References

1. Panigrahi, P.K.: A comparative study of supervised machine learning techniques for spam email filtering. In: 2012 Fourth International Conference on Computational Intelligence and Communication Networks, 2012
2. Alvestad, S.: Early warnings of critical diagnoses, 2009
3. Metsis, V., Androutsopoulos, I., Paliouras, G.: Spam filtering with Naive Bayes—which Naive Bayes?. In: Proceedings of the 3rd Conference on Email and Anti-Spam (CEAS 2006), Mountain View, CA, USA, 2006
4. Ackoff, R.L.: From data to wisdom. J. Appl. Syst. Anal. **16**(1), 3–9 (1989)
5. Sharma, N.: The origin of the data information knowledge wisdom hierarchy. Data Inf. Knowl. Wisdom hierarchy (2008)
6. Chandrasekaran, M., Narayanan, K., Upadhyaya, S.: Phishing email detection based on structural properties. In: Proceedings of 9th Annual NYS Cyber Security Conference, June 2006
7. Ozarkar, P., Patwardhan, M.: Efficient spam classification by appropriate feature selection. Global J. Comput. Sci. Technol. Softw. Data Eng. **13**(5) (2013)
8. Lim, T.S., Loh, W.Y., Shih, Y.S.: A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. Mach. Learn. **40**(3), 203–228 (2000)
9. Hsiao, W.F., Chang, T.M.: An incremental cluster-based approach to spam filtering. Expert Syst. Appl. **34**(3), 1599–1608 (2008)
10. Awad, W.A., Elseuofi, S.M.: Machine learning methods for E-mail classification. Int. J. Comput. Appl. **16**(1), 39–45 (2011). doi:10.5120/1974-2646
11. El-Alfy, E.S.M., Abdel-Aal, R.E.: Using GMDH-based networks for improved spam detection and email feature analysis. Appl. Soft Comput. **11**(1), 477–488 (2011)
12. Ion Androutsopoulos, Enron Dataset http://nlp.cs.aueb.gr/software_and_datasets/Enron-Spam/index.html
13. Weka- Data Mining Tool tutorials, documentation http://www.cs.waikato.ac.nz/ml/weka/documentation.html
14. Ahmed, Kh: An overview of content-based spam filtering techniques. Informatica **31**(3), 269–277 (2007)
15. Geetha Ramani, R., Sivagami, G.: Parkinson disease classification using data mining algorithms. Int. J. Comput. **32**(9) (2011)
16. Guzella, T.S., Caminhas, W.M.: A review of machine learning approaches to Spam filtering. Expert Syst. Appl. **36**, 10206–10222 (2009)

CSCR-A Context Sensitive Concept Ranking Approach

Niyati Kumari Behera and G.S. Mahalakshmi

Abstract CSCR is a context sensitive ranking approach which combines both linguistic and statistical approach to exploit the contextual information around a set of concepts to rank them. It represents each concept as a very high dimensional vector of context terms gathered from its surrounding. This paper discusses a probabilistic approach to assign weight to those terms in the vector and there by rank a set concept. The objective of this work is to bring the semantically related concepts closer so that the result can be used for application where there is a need to consider semantic closeness between concepts e.g. incremental ontology learning. The effectiveness of CSCR is evaluated with a set of 300 concepts extracted using frequency based approach by C Finder algorithm.

Keywords Concept · Context · Context windowing · Concept ranking

1 Introduction

In this era of continuously evolving semantic knowledge where meaning plays crucial role rather than mere terms, this paper presents a novel approach to incorporate contextual information to figure out the concepts whose meanings are closely related in a particular domain. The main aspect of this paper is to exploit this information around a concept to estimate its closeness to other concepts under consideration. This semantic relatedness information among concepts can be used for applications where there is a need to consider semantic closeness between terms

N.K. Behera (✉)
Department of CSE, Research Scholar, CEG, Anna University, Chennai
Tamilnadu 600025, India

G.S. Mahalakshmi
Department of CSE, CEG, Anna University, Chennai
Tamilnadu 600025, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_24

like finding suitable place for newly discovered concept within an existing ontology in case of incremental ontology learning.

The research contribution of this paper is discussed elaborately in the following sections which include literature surveys related to the work prospective, detailed description about the proposed frame work, result evaluation. Finally it concludes with some open issues which can be further explored.

2 Related Work

The main objective of this proposed work is to extract and use the contextual information around a term/concept which generally conveys the sense in which it has been used. Moreover it aims to rank the concepts based on these information so that semantically related concepts can be brought closer. There have been considerable amounts of research done mainly for discovering domain specific terms based on their frequency of occurrence [1]. *C Finder* algorithm [2] uses a unique technique to rank the concepts (also multi term words) by considering the frequency of occurrence of included candidate terms. The authors in [3] have used context window to extract words from a domain corpus which are correlated to the concepts existing in ontology. Some other researchers as in [4] have used a Multivariate Probability based clustering approach to find the semantic relatedness between concepts. Authors in [3] have proposed a semantic ordering of terms based on term clustering and term togetherness to generate a linear order. Authors in C*/NC* value based approach [5] for automatic recognition of multiword terms have achieved considerably better result by using context information in their automatic term recognition system. The proposed approach considers the probabilistic occurrence of each and every context word appearing around the target concept within the window range. So that the concept doesn't miss out any major or minor information conveyed by the context word which can bring out its meaning or semantic relatedness with other concepts. In the following section the proposed methodology has been discussed in more detail.

3 CSCR-Context Sensitive Concept Ranking

CSCR is a novel concept ranking approach which incorporates context information to rank a set of concepts extracted from a domain specific text corpus. Section 3.1 discusses the work flow details of the proposed architecture for incorporating context information. The following subsection also discusses the preprocessing steps in involved to extract context information regarding concepts. The obtained results of the proposed algorithm has been analyzed and discussed in Sect. 4.

3.1 Context Formation for Concept

The context formation idea evolves from an assumption that the neighboring key words of a concept are very much informative to convey its meaning or the context in which it is used. This subsection gives a detailed insight of the proposed method which intend to incorporates context information for ranking the concepts. Figure 1 shows the overall flow diagram of CSCR.

The Context Formation problem is considered as follows:

Given an occurrence of a target concept, CSCR tries to obtain a set of words around it throughout the corpus which is supposed to convey its context of use. This collection of words for a particular concept is referred to as context vector throughout the discussion here after. For this purpose only the verbs, nouns and adjectives surrounding the target word are considered. To rank the concept, a probabilistic weight is assigned to it by computing the frequency of occurrence of words appearing in its context vector as well as all other context vectors under consideration.

The ranking of concepts using context information takes place as follows:

Windowing has been a very common practice in many NLP applications [6] to extract a set of words surrounding a target word. Those words are extracted by using a window of size 2n + 1 where n is the number of words on either side of the target word. In our case we have chosen a context window of size n = 5 on either side of the target concept. From the collected words, only the nouns, adjectives and verbs are filtered by using POS tagger. All these filtered words together form the context vector for a particular concept. For each concept, we obtain the context based new rank by adding up: a probabilistic weight for its Individual context words, multiplied by their frequency of appearing with this particular concept.

For example, suppose a concept C has the following context vector $\{W_1 \cdot C_1,\ W_2 \cdot C_2,\ W_3 \cdot C_3 \ldots W_n \cdot C_n\}$ where:

C_n = any context word in the context vector
W_n = frequency of occurrence of the context word C_n

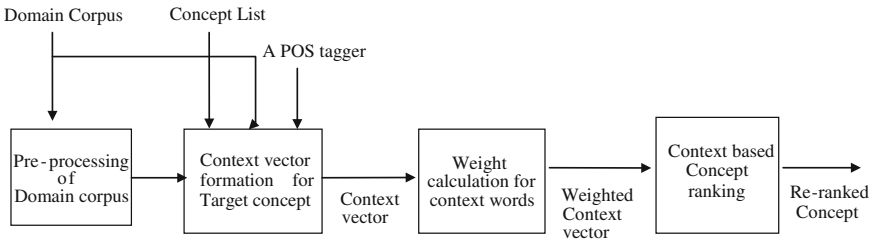


Fig. 1 Flow diagram of CSCR

Then the context based new weight for any concept is calculated as

$$CW(c) = \sum_{i=1}^{n} Wc(i) * Pc(i) \quad (1)$$

where:

C = a concept from the concept list
n = no of distinct context words in the context vector of c
Wc(i) = frequency of occurrence of each context word i
Pc(i) = probability of occurrence of i = T(i)/N
T(i) = no of concept terms the context word i appeared with
N = total no. of concepts in the concept list

3.2 Pseudo Code for Concept Ranking

The Pseudo code used to compute the rank of a concept using the contextual information around it is discussed below.

Input: a list of concepts, a domain specific corpus
Output: a ranked list of concepts based on context

Steps:

- Preprocess the corpus for stop word removal and POS tagging
- Use context windowing to extract context words around each concept in the list
- Create a context vector for each concept including all those context words in the whole corpus
- Compute the probabilistic weight for each context word in the context vector
- Combine frequency of occurrence of the context words along with its probabilistic weight to find the rank of each concept in the input list

4 Results and Discussion

In this paper, a context based ranking approach has been discussed that can be effectively used to bring the semantically related concepts closer. The data set used for this work is the "Compendium of Mass Gatherings "corpus for both statistical analysis as well as evaluation of the proposed approach. All sort of required preprocessing like stop word removal, tagging etc. have been done on the corpus before the context information extraction phase begins. The concepts that have been considering for ranking process are already extracted and ranked by the *C-Finder*

algorithm. This *C-Finder* algorithm uses a very unique method to extract the concepts from the above mentioned domain specific corpus by considering the internal structural pattern of the concept (mainly multi term concepts). It has extracted 300 concepts from the corpus 'Compendium of Mass gathering' corpus [2] which has been validated by domain experts [2]. It has used statistical approach based on frequency of occurrence of key concepts in the corpus and has ordered them accordingly. Table 1 list out few group of concepts which are semantically related.

In Table 1 few semantically related concept groups are listed. Many such groups are available in the original list.

For every group of such semantically related concepts a n × n matrix is created where n is the number of concepts in the group. Each entry in the matrix is the linear difference in rank between a pair of concepts. Then based on this matrix we introduce a performance measure "Average Ranking Distance (AvRD)" to compare the performance of the proposed system over the base system.

$$\text{AvRD} = -\sum_{n}^{1} \text{ranking distance among all concepts} \tag{2}$$

Since the main motive behind this context based ranking approach is to bring the semantically related concept closer, the system performs better if the AvRD value is less. Figure 2 shows a representation of how the proposed system performs. The graph clearly shows that the proposed context based ranking performs better compared to the base frequency based ranking method for most of the cases. The performance dip in case of GI and GVI can be due to the fact that the proposed method is purely dependent on effectiveness of context formation process. If a concept does not occur frequently in the corpus then it will affect the accumulation of context words for context formation.

Table 1 Few semantically related concepts

Group	Concepts
I	Event, public events, mass-gatherings events, event planning, event site
II	Emergency medical services, medical care, medical services, mass-gathering medical care, health care systems, medical care system, medical care services, medical responses, appropriate health care
III	Patient presentation rate, patient presentations, potential patient presentations, casualty rates, patient presentation patterns, casualty presentation rate
IV	Illness, injury, mass casualty incident, casualties, injury patterns, specific injury types, casualty presentations
V	Health services, health care, patient management
VI	Mass gathering, mass gathering phenomenon, large event, large public event, mass gathering guidelines, mass gathering situations

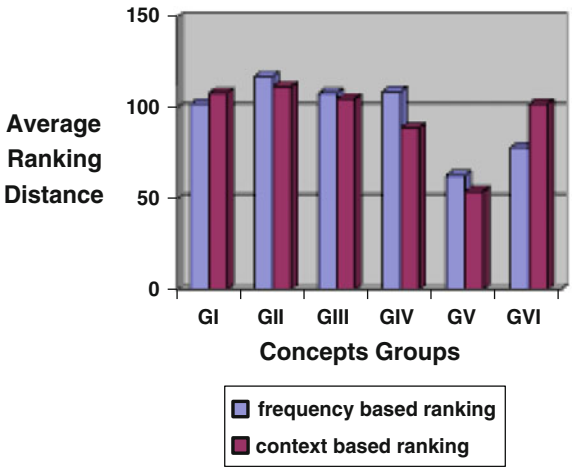


Fig. 2 A sample graph showing performance of proposed method

Table 2 lists out few concepts whose rank has improved after using the context information. It clearly shows that the proposed approach has brought concepts like Mass, Mass gathering, Gathering closer. This is due to the high probabilistic weight of the context words in their context vectors. Similarly concepts like mass gathering health and mass-gathering health though different from keyword point of view, their semantic relatedness i.e. the context of their use has brought them nearer. Again concepts like care, medical services, hospital are also ranked relatively closer compared to their position in frequency based approach. Although it shows that the proposed method is effective in bringing the semantically similar concepts closer,

Table 2 Few concepts whose ranking has improved after context based ranking

Sl. no	Concept name	Frequency based ranking	Context based ranking
1	Service	144	1
2	Mass	62	2
3	Mass gathering	1	3
4	Gatherings	289	7
5	Care	28	5
6	Medical services	79	6
7	Medical care	14	17
8	Hospital	22	11
9	Health	164	14
10	Physician	12	18
11	Planning	190	21
12	Analysis	267	44
13	Study	59	47
14	Illness	81	65
15	Health care	61	66

Table 3 Few concepts whose ranking has not improved after context based ranking

Sl. no	Concept name	Frequency based ranking	Context based ranking
1	Patient	3	8
2	Injury	10	30
3	Event	2	10
4	Event type	13	58
5	Public events	57	110
6	Medical care	14	17
7	Health care	61	66
8	Physician	12	18
9	Emergency medical services	4	20
10	Emergency care facilities	243	291
11	Concert	23	22
12	Rock concert	16	57
13	Crowd	9	24
14	Crowd size	65	81
15	Crowd behavior	27	97

there are certain cases where the algorithm sets back to make the claim. Table 3 lists out few such concepts.

For example concepts like (‘mass gathering’ and ‘large gathering’), (concert and Rock concert), (medical care, health care), (Patient presentations, patient presentation rate) though semantically related; they are placed apart in t[he new list. Some concepts like ‘table’, ‘types’,’ zeitz’, their semantically relatedness to the concepts in their close vicinity are not very convincing. This can be due to the reason that these concepts don’t appear frequently in the corpus. So their associated context vectors fail to gather sufficient contextual information about them.

5 Conclusion and Future Work

In this paper, a probabilistic technique has been discussed to rank a set of concepts using the linguistic information around them. It mainly aims at using the method for application where there is a need to bring the semantically related concepts closer. For example, in case of dynamic ontology learning, semantic relatedness between a newly discovered concept and existing ontological concepts is accessed to update the ontology. Using this approach a set of concepts can be suggested that are semantically close to the existing ontological concepts. Work is in progress to incorporate some other semantic similarity measures along with this context information to strengthen the claim. Since the context vector plays very important role in ranking the concepts, in future the context vector generation aspect can be

improvised by considering other syntactic and semantic features of surrounding words. Since this work has been tested in one corpus, as future work performance of CSCR can be analyzed with other related algorithms and corpuses.

References

1. Faatz, A., Steinmetz, R.: Ontology enrichment with Texts from the www. In: Proceedings of the 2nd Semantic Mining Workshop at ECMLI/PKDD, WS'02 (2006)
2. Ha, M.S.: Semantic information retrieval based on Wikipedia taxonomy. In: Int. J. Comput. Appl. Technol. Res. **2**(1), 77–80 (2013)
3. Wittek, P., Dáranyi, S., Tan, C.L.: An ordering of terms based on semantic relatedness. In: Proceedings of the 8th International Conference on Computational Semantics, pp. 235–247 (2009)
4. Ye, N., Gauch, A.P.S.: Mining hidden concepts for ontology extension using multivariate probabilistic modelling. In: Web Intelligence and Intelligent Agent Technologies, WI-IAT'09, IEEE/WIC/ACM International Joint conference (2009)
5. Frantziy, K., Ananiadouy, S., Mimaz, H.: Automatic recognition of multi-word terms: the C-value/NC-value method. In: Natural language processing for digital libraries, Springer, vol. 3, pp. 115–130 (2000)
6. Kang, Y.-B., Haghighi, P.D., Burstein, F.: CFinder : An intelligent key concept finder from text for ontology development. In: International Journal of Expert Systems with Applications, Elsevier (2014)
7. Harb, A., Hajlaoui, K.: Enhanced semantic automatic ontology enrichment. In: IEEE (2010)
8. Liur, H.Z., Bao, H., Wang, J. Xu, D.: A novel vector space model for tree based concept similarity measurement. In: IEEE (2010)
9. Liu, H.Z., Bao, H., Xu, D.: Concept vector based similarity measurement based on hierarchical domain structure. Comput. Inform. **30**, 881–900 (2011)
10. Altintas, Ergin, Karsligil, Elif, Coskun, Vedat: The Effect of Windowing in Word Sense Disambiguation. LNCS 3733, pp. 626–635. Springer, Berlin (2005)

Energy-Aware Application Scheduling and Consolidation in Mobile Cloud Computing with Load Balancing

L. Shakkeera and Latha Tamilselvan

Abstract Mobile Cloud Computing (MCC) extends cloud computing with the advantages of mobility and wireless networks to create a new infrastructure where cloud takes over mobile devices' responsibilities of executing tasks and storing enormous amounts of data. Through offloading, all the major data processing work takes place in the cloud instead of the mobile devices. The main aim of MCC is to achieve a rich user experience by enabling wide range of mobile devices to execute rich mobile applications. Scheduling of tasks require minimum completion time, better performance, effective utilization of resources and quick response time for which cloud uses virtualization concept. For task allocation, cloud provides virtual machines which are scalable but scheduling them while efficiently utilizing the idle service capacities of the mobile devices are still remains major problem. Likewise, there are other issues faced in MCC such as insufficient resource, low connectivity and limited energy due to which utilizing its full capability is a challenge. The existing application scheduling algorithms in MCC do not take each task's profit or the overall energy consumption of mobile devices into consideration. Also it cannot increase the profit of the system, which is an import target for scheduling the tasks in commercial mobile cloud environment. In this paper, E-MACS (Energy-aware Mobile Application Consolidation and Scheduling) algorithm is proposed to make the mobile devices contribute their computing and sensing capabilities to attain efficient scheduling of application in hybrid cloud model. The consolidation of application minimizes the overall energy consumption in cloudlet. The proposed system minimizes the response latency, cost of application migration and it improves quality of service like throughout and scalability among resources using load balancing techniques by mobile cloud computing.

Keywords Mobile cloud computing · Cloudlet · Application offloading · E-MACS · Application scheduling and consolidation · Load balancing

L. Shakkeera (✉) · L. Tamilselvan
Department of IT, B.S. Abdur Rahman University, Chennai 600048, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_25

1 Introduction

Mobile Cloud Computing (MCC) [1–3] is growing as one of the most significant branches of Cloud Computing as it offers greater flexibility and availability of resources. MCC can be defined as a mobile computing technology that provides mobile users with unlimited resources, functionality, data storage and mobility. Users can access the cloud when and wherever they need it using any mobile device via Internet or Ethernet based on pay-per-use principle. MCC helps the users to overlook the fact that mobile devices generally have limited capacity and energy. This is because all resource-intensive computational and processing work takes place within the cloudlet providing mobile users with more facilities. This greatly enriches the types of applications on mobile devices and enhances the quality of service of the applications. Mobile Cloud Computing is a rapidly rising trend in the mobile computing technology. Though its usage has been increasing in an alarming rate, it is difficult due to its inherent problems like resource scarcity, frequent disconnections and mobility. Section 3 explains the different challenges faced in mobile cloud environment.

Application scheduling is the most effective and important task for a mobile cloud to basically decide the order of execution of tasks when multiple tasks are in the queue. A good scheduling mechanism must keep the load of a system balanced. Load balancing focus on distributing the tasks to be executed equally among the resources of the system and avoid overload of one machine with many tasks. This will prevent any mobile device getting overloaded and leaving the system too early in a mobile cloud environment.

Application consolidation is also referred to as server consolidation where the key goal is to reduce energy consumption, infrastructure complexity and increase the system availability. This is done by grouping multiple applications into one server thereby reducing the total number of servers in a data center. Idle or least used applications in the cloud are identified and migrated to other servers. The idle server is then temporarily shut down in order to prevent wastage of energy. The concept of server virtualization can be used as an effective way of doing this. According to this concept each virtual machine is dedicated to a single application. These virtual machines are in turn mapped to their corresponding physical servers. Section 4 includes the proposed system for energy-aware application scheduling and load balancing.

2 Related Work

Wei et al. [4] provided suitable definitions of critical aspects and proposing efficient algorithms and approaches. Their simulation results revealed that when the load of the system is heavy, HACAS algorithm can select those applications with maximum profit and minimum energy consumption. With the parameters setting in the

simulation, the profit of HACAS algorithm is about 30 % higher than that of FCFS algorithm. Besides, when the load of the system is light, the provider selection scheme adopted in HACAS can effectively balance the load of the devices in the system. Concretely speaking, HACAS algorithm's load variation is about 60 % better than that of the random provider selection scheme.

Lin et al. [5] studied the MCC task scheduling problem and came up with the first task scheduling work that minimizes energy consumption under a hard completion time constraint for the task graph in the MCC environment, taking into account the joint task scheduling on the local cores in the mobile device, the wireless communication channels, and the cloud. They proposed a novel algorithm that starts from a minimal-delay scheduling and subsequently performs energy reduction by migrating tasks among the local cores and the cloud. A linear-time rescheduling algorithm was also proposed for the task migration such that the overall computation complexity is effectively reduced. Simulation results demonstrated significant energy reduction with the overall completion time constraint satisfied.

Wua et al. [6] proposed a scheduling algorithm that computes the priority of the tasks by using their attributes based on QoS-driven in Cloud computing. This algorithm sorts the tasks according to their priority and then schedules each task onto a service which has minimum completion time. The experimental results showed that the algorithm achieved good performance. Though the algorithm based on QoS-driven could attain load balancing, it failed to consider energy consumption.

Yamauchi et al. [7] considered mobile devices' battery power, mobility and imbalanced load as three important aspects and proposed a distributed parallel scheduling methodology for mobile cloud computing. As per this methodology a master device observes the above mentioned parameters of the system and the other devices are its slaves. If the parameters are not sufficient for parallel processing to take place, the master device will select other slave devices for processing those threads. The performance of mobile devices was improved and unlike other algorithms it considered network quality in addition to load balancing too. It reduced the battery consumption of mobile devices but not the total power of the overall system.

Mishra and Jaiswal [8] designed a heuristic load balancing algorithm based on Ant colony optimization concept considering various performance parameters such as both CPU and network load, and available memory. The pheromone update mechanism proved to be an efficient tool for achieving load balancing of cloud resources and minimizing makespan. It does not consider the fault tolerance issues in cloud environment.

Suryadevera et al. [9] proposed a load balancing algorithm for grid computing using Ant colony optimization. High pheromone value of a path means it is shorter. It considers various resource parameters (MIPS, communication bandwidth, Number of processors and memory) for calculating pheromone i.e. the resource capability for executing various jobs thus allocating the best resource for the job and also balance the load of all the resources. The algorithm is expected to achieve

better throughput and hence increase the overall performance in the grid environment. However, it does not consider QoS requirements.

Yang et al. [10] proposed and described a service for offloading heavy mobile applications from resource-constrained MHs to nearby surrogates with rich resources. These applications may be downloaded from the internet. The service was able to efficiently offload only those applications that were implemented in Java. The experimental and simulation results showed that the offloading service was very effective.

3 Problem Definition

In mobile cloud computing, offloading to remote cloud infrastructure will introduce long response latency. Hence, to reduce the response latency, the offloaded applications should be handled by local cloud infrastructure, which usually has fewer resources than those of the remote ones. Therefore, how to efficiently schedule the limited resources while fulfilling the requirements of a large number of offloaded applications is a critical problem for MCC to address. Different from the scheduling algorithm in cloud environment, the scheduling algorithms for MCC should take the energy consumption into consideration. To make the system last longer, the scheduling algorithms should balance the load [11] of the mobile devices to avoid some heavy-loaded nodes leaving the system too early. Therefore, these scheduling algorithms for cloud computing cannot be applied to the MCC environment directly.

4 Hybrid Cloud Model for Application Scheduling and Consolidation

As the number of users of cloud computing systems has increased, the tasks to be scheduled in cloud have also increased proportionally. Task Scheduling algorithms in cloud computing aim at minimizing the makespan of tasks with minimum resources efficiently. Load balancing and reducing energy consumption has to be done efficiently to achieve quality of service. The architecture model of our hybrid cloud system for energy-aware application scheduling is shown in Fig. 1. The working mechanism is as follows: (1) The client (mobile users) submits their application requests. (2) These requests are offloaded to the cloudlet with the help of internet. (3) The applications are executed in the cloudlet servers. (3.1) The idle applications and resources are found. (3.2) The idle applications are migrated to other servers within the cloudlet and the idle server is shut down thereby preventing wastage of energy. (4) The application scheduling algorithm (E-MACS) is executed. (5) The applications are offloaded to the local mobile cloud which comprises

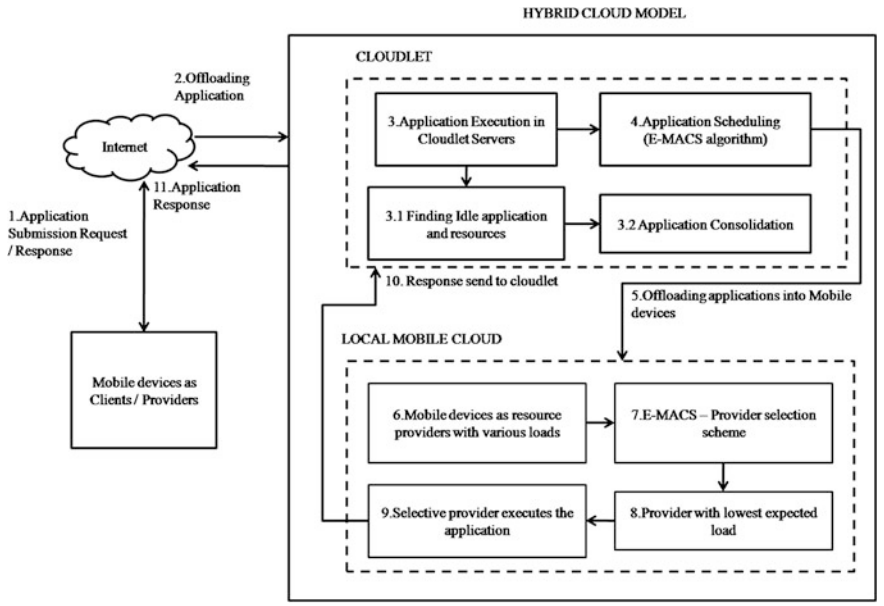


Fig. 1 Hybrid cloud model

of mobile devices. (6) Various mobile devices (resource providers) with various loads are available. (7) The provider selection scheme of the E-MACS algorithm gets executed. (8) The mobile device with the least load is selected as the resource provider. (9) The selected resource provider executes the application. (10) The execution result is sent to the cloudlet. (11) The cloudlet then sends the result to the mobile device (client).

The system is divided into three important phases:

- Application Scheduling and Consolidation
- Load Balancing
- Resource Provisioning

4.1 Application Scheduling and Consolidation

A cloud must be able to handle numerous tasks and provide its clients' desired services within a minimum time delay. Any task has to be scheduled in order for it to get sufficient resources for execution. In a Hybrid cloud model, the application requests received from the clients are scheduled among the cloudlet servers. Multiple requests from various clients are likely to keep the cloudlet resources busy. During such situations the arriving requests are offloaded to the local mobile cloud to prevent them from waiting in queue (thereby reducing response latency). For this

the E-MACS algorithm is implemented. This scheduling algorithm's provider selection scheme is responsible for selecting the resource provider for the application. Idle applications are believed to consume more energy than running applications. For better utilization of servers, application consolidation is done. Hence idle applications are migrated to other servers within the cloudlet that can accommodate them. For reducing the overall energy consumption, the idle servers are shut down. The E-MACS algorithm includes the consolidation process of applications.

The prime objective of the algorithm given below is to efficiently schedule applications, balance the load and achieve QoS in a mobile cloud environment. The algorithm reduces the response latency, improves resource utilization and reduces the overall energy consumption. The following performance metrics Eqs. (1), (2), and (3) are used for measuring the final results.

Makespan: It is the total time taken to complete the processing of the initial task and ending task. It can be calculated using (Eq. 1).

$$M = t(a_n)_f - t(a_1)_s \tag{1}$$

Here, $t(a_n)_f$ is the execution finish time of the last application and $t(a_1)_s$ is the execution start time of the first application.

Average Latency: It is the difference between the current time and the time at which the application entered the queue initially.

$$L = \Sigma(t_c - t_q)/n \tag{2}$$

Here, t_c is the current time while t_q is the time at which an application entered the queue. n is the total number of applications.

Energy Consumption: It is the amount of energy consumed by the mobile devices (clients) and cloudlet server resources.

4.2 Load Balancing

Load balancing, as the name implies, manages network load by distributing traffic among the various nodes in a network thereby increasing resource usage and throughput. By this way, more tasks get executed and all users get served faster. It is responsible for equally distributing load across the various available resources preventing any single node getting overloaded. Load balancing is critical for networks where the number of incoming requests is unpredictable. It prevents single point of failure and hence improves the overall responsiveness and availability of applications. In Local mobile cloud resources with various loads are present. When

applications have to be offloaded to the local mobile cloud, using E-MACS's Provider selection scheme, the mobile device with the least load is selected as the resource provider. Because offloading to a provider with high load will make it overloaded and leave the system too early.

4.3 Resource Provisioning

Resource provisioning is the process of making resources available for the clients' application requests. After execution, the mobile device (resource provider) sends the results to the cloudlet. The cloudlet in turn sends the results to the corresponding client.

Resource Utilization: It indicates how efficiently resources are utilized.

$$Ru = \Sigma r_x(tf_i - ts_i) \tag{3}$$

Here, r_x is the resource, tf_i is execution finish time, and ts_i is the execution start time.

Algorithm: Energy-aware Mobile Application Consolidation and Scheduling (E-MACS)
Input: Input size, Output size, Pe count and Number of applications (noapp), Queue of applications app[] to the cloudlet
Begin
(1) Initialize i, j, noapp, applilen[], arr1[], arr2[], app[], k, scheduled_applications, mobile_service_providers
(2) **for**(i=1;i<=noapp;i++)
(3) **for**(j=1;j<=noapp;j++)
(4) **if** (applilen[i]<applilen[j]) **then**
(5) Swap applilen[i] and applilen[j] to sort them in ascending order
(6) **else**
(7) Order of applilen[i] and applilen[j] is retained
(8) **end if**
(9) applilen[] is divided into arr1[] and arr2[]
(10) Applications with highest load executed in cloudlet
(11) **if** (app[i]_status = idle) **then**
(12) migrate app[i] to host k
(13) **end if**
(14) Calculate the expected loads of the feasible mobile service providers
(15) Sort the feasible providers in an increasing order of their expected loads
(16) Select the provider with lowest expected load
(17) Offload the applications with less load to local mobile cloud
(18) **end for**
(19) Send application execution results to client
(20) **end for**
(21) Send results from local mobile cloud to cloudlet
(22) Cloudlet sends result to client
(23) Print scheduled_applications
(24) Print mobile_service_providers
End

5 Operational Workflow

The Fig. 2 shows the entire operational workflow of our proposed system. For this paper we have created a mobile cloud environment in which a single mobile user is enabled to offload multiple applications to the cloudlet via internet. The applications are initially placed in the queue. It is checked if resources are available in the cloudlet. If yes, the applications are executed within the cloudlet. If no, they are offloaded to the local mobile cloud. The E-MACS algorithm is executed. Idle applications are identified and migrated to other servers and then the idle server is shut down. The provider selection scheme of the E-MACS algorithm checks for the provider with least load. If available, the applications are offloaded to that provider. The mobile service provider runs the application and sends the response to the cloudlet which in turn sends it to the mobile user. If a provider with least load is not available, the provider selection process gets executed again.

6 Experimental Results and Discussions

6.1 Experimental Setup

In order to simulate a cloud environment, Cloudsim 3.0 with NetBeans IDE is used. Using Cloudsim a cloud environment is set up by creating datacenter with hosts, a

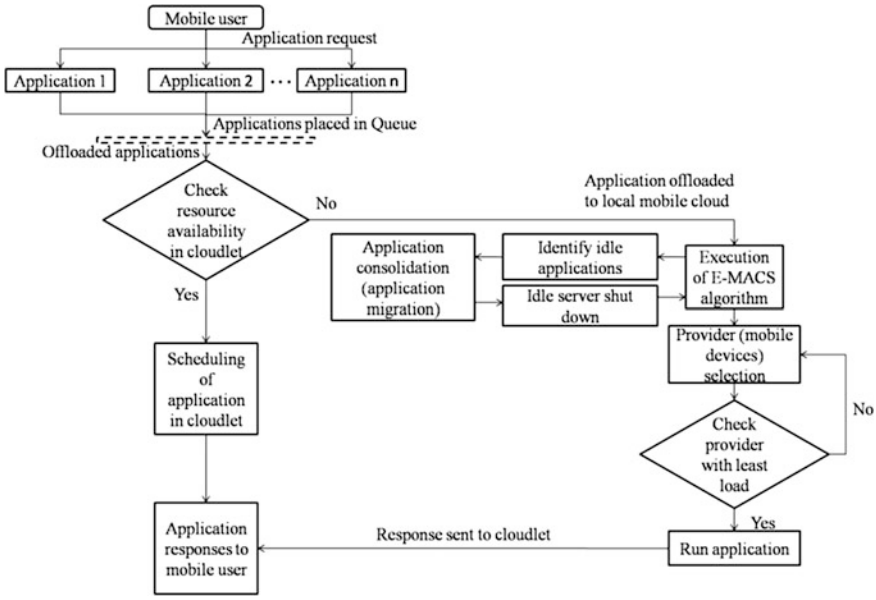


Fig. 2 Operational workflow

cloudlet, and mobile users (clients) and mobile service providers. VMs are created to provide cloud services to the clients. Scheduling of applications is done by the E-MACS algorithm. The proposed algorithm is compared with the HACAS and ACO algorithms. The performance such as makespan, average latency, energy consumption, resource utilization and load of the resources are monitored using performance analysis graph.

To analyze theses performance metrics for execution of offloaded applications in a mobile cloud environment, various experiments are conducted implementing the E-MACS, HACAS, and ACO algorithms. All these experiments are performed for a maximum of 50 offloaded applications from a single mobile client. Experiment 1 discusses about makespan. Experiment 2 analyses the average latency. Experiment 3 compares the energy consumption. Experiment 4 analyses load balancing and experiment 5 discusses about the resource utilization when each algorithm is used.

6.2 Results and Discussions

6.2.1 Makespan

The Fig. 3 shows the makespan of the ACO, HACAS and the E-MACS algorithm. The makespan is calculated using Eq. (1). It is evident from the comparison that E-MACS have a better makespan of about 70 % than ACO or HACAS. This is because applications are executed in both the cloudlet as well as the local mobile cloud. Hence the applications are prevented from waiting in the queue for a long time.

6.2.2 Average Latency

Figure 4 shows the average latency of E-MACS, HACAS and ACO algorithms. It is calculated as the difference between the current time and the time at which the application enters the queue using Eq. (2). It is observed from the graph that the E-MACS had about 60 and 40 % better average latency than the HACAS and ACO

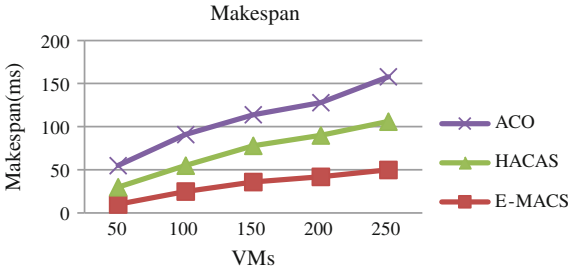


Fig. 3 Makespan

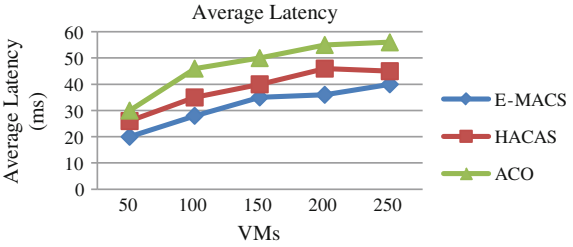


Fig. 4 Average latency

algorithm respectively. This is because accessing the remote cloud will introduce high response latency. Since we use cloudlet, the average latency is reduced.

6.2.3 Energy Consumption

Figure 5 shows the comparison of energy consumption of the ACO, HACAS and the E-MACS algorithms. It shows that our proposed E-MACS algorithm reduces the overall energy consumption by 80 % when compared to that of ACO. This is because application consolidation prevents wastage of energy in cloudlet servers. The response latency for applications are minimized which will in turn keep the energy of the mobile devices in stable manner.

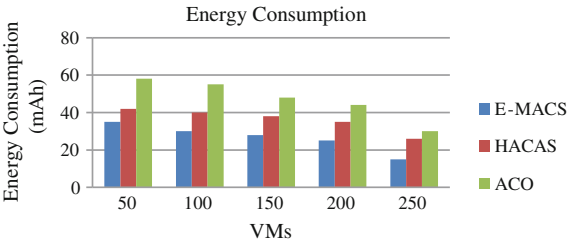


Fig. 5 Energy consumption

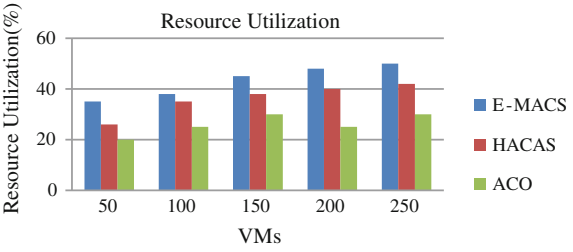


Fig. 6 Resource utilization

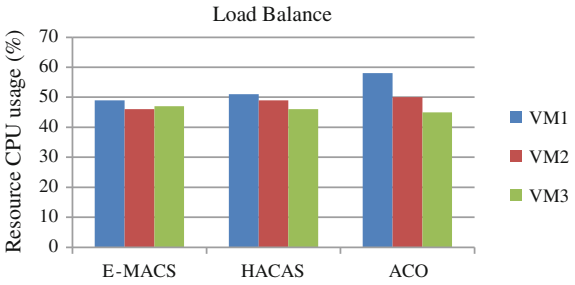


Fig. 7 Load balance

6.2.4 Resource Utilization

Figure 6 shows the resource utilization of ACO, HACAS and the E-MACS algorithms. The E-MACS algorithm seems to utilize the resources more than the other two algorithms. This is due to application consolidation. The consolidation process eliminates idle resources and makes use of other active servers. Thus, our proposed system effectively utilizes the resources.

6.2.5 Load Balance

Figure 7 shows how the load is balanced by the ACO, HACAS and E-MACS algorithms. Out of 50 VMs actually created, three VMs are considered for comparison. It is observed that when E-MACS algorithm is implemented the load among the three VMs are almost equally distributed. It is slightly better than the HACAS algorithm. This is because the E-MACS algorithm's provider selection scheme offloads the application only to a mobile service provider whose load is less. Thus overload of resources are prevented.

7 Conclusion

This paper discusses about energy aware application scheduling and consolidation in mobile cloud computing with load balancing. Most approaches concentrate on scheduling and load balancing but have not considered energy consumption. So we have proposed an energy aware application scheduling algorithm called E-MACS (Energy-aware Mobile Application Consolidation and Scheduling). This algorithm migrates idle applications to other active servers. The idle servers are shut down. Finally the makespan, average latency, energy consumption, resource utilization, and load balance are monitored. The experimental results show that the proposed system implementing the E-MACS algorithm gives a better performance than the existing system by reducing the overall energy consumption and response latency.

References

1. Wang, Y., Chen, I.-R, Wang, D.-C.: A survey of mobile cloud computing applications: perspectives and challenges. Wireless Pers. Commun. (2014)
2. Fernando, N., Loke, S.W., Rahayu, W.: Mobile cloud computing: a survey. Elsevier Trans. Future Gener. Comput. Syst. **29**(1), 84–106 (2013)
3. Dinh, H.T., Lee, C., Niyato, D., Wang, P.: A survey of mobile cloud computing: architecture, applications, and approaches. Wireless Commun. Mobile Comput. **13**(18), 1587–1611 (2013)
4. Wei, X., Fan, J., Lu, Z., Ding, K.: Application scheduling in mobile cloud computing with load balancing. Hindawi Publishing J. Appl. Math. (2013)
5. Lin, X., Wang, Y., Xie, Q., Pedram, M.: Energy and performance-aware task scheduling in a mobile cloud computing environment. In: IEEE International Conference on Cloud Computing (2014)
6. Wua, X., Denga, M., Zhanga, R., Zengb, B., Zhoua, S.: A task scheduling algorithm based on QoS-driven in cloud computing. Elsevier First Int. Conf. Inf. Technol. Quant. Manag. **17**, 1162–1169 (2013)
7. Yamauchi, H., Kurihara, K., Otomo, T., Teranishi, Y., Suzuki, T., Yamashita, K.: Effective distributed parallel scheduling methodology for mobile cloud computing. In: SASIMI 2012, the 17th Workshop on Synthesis and System Integration of Mixed Information Technologies, pp. 516–521 (2012)
8. Mishra, R., Jaiswal, A.: Ant colony optimization: a solution of load balancing in cloud. Int. J. Web Semant. Technol. **3**(2) (2012)
9. Suryadevera, S., Chourasia, J., Rathore, S., Jhummarwala, A.: Load balancing in computational grids using ant colony optimization algorithm. Int. J. Comput. Commun. Technol. **3**(3) (2012)
10. Yang, K., Ou, S., Chen, H.-H.: On effective offloading services for resource-constrained mobile devices running heavier mobile internet applications. Mobile Internet Technol. Appl. (2008)
11. Raghava, N.S., Singh, D.: Comparative study on load balancing techniques in cloud computing. Open J. Mobile Comput. Cloud Comput. **1**(1) (2014)

Classroom Collaboration Using an In-Home Gaming Device

R.K. Lekshmipriya, Ashwini Kumar, Divya Mahadevan and Kamal Bijlani

Abstract Few common devices currently used in classrooms for collaboration include computers, projectors, interactive whiteboards and multi-touch devices. Such devices are often expensive and not easily acquirable to all. We propose a low-cost multi-touch collaborative surface that can be created using a simple and inexpensive gaming device, Nintendo's Wiimote. Wiimote and hand-made IR-LED pen was combined to setup the multi-touch surface. IR sensor in Wiimote was programmed to accurately track the IR signal from the IR-LED pen. Our system can be used to create a multi-touch surface anytime-anywhere and thus overcoming the mobility constraints that exist with current systems. Proposed system was found to be effective for collaborative classroom activities through the closed-group experiments conducted at multiple workshops on certain digital-media based classroom activities. Through the proposed system we have created a pathway to bring in gaming devices to classroom to make learning fun and more engaging with less complex setup and at an optimized cost.

Keywords Collaborative learning · Future classroom technologies · Emerging education technology · Interactive whiteboard technologies. Wiimote multi-touch surface

1 Introduction

Learning can be made fun and more engaging. *Smart classrooms* [1] aim at creating an interactive and engaging learning environment for students. Here, we provide a model which enables anybody with no particular technical expertise to calibrate a gaming device that would help them to create a *multi-touch surface* for classroom

R.K. Lekshmipriya (✉) · A. Kumar · D. Mahadevan · K. Bijlani
Department of E-Learning, Amrita University, Karunagapally 690525, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_26

use. The paper focus on using a gaming device, the *Wiimote* that enables anyone to use it in classrooms and make classroom based learning more fun. We propose a cost-effective solution to create a multi-touch surface anywhere-anytime using an in-home gaming device, *Nintendo's* [2] *Wiimote*. Wiimote [3] was released by Nintendo on December 2006 as its fifth home videogame console. The device resembles a TV remote and contains an IR sensor, an accelerometer and a few controlling buttons. Even though Wiimote was released as a videogame console, its efficient IR sensing capability have enabled it to be used for several other purposes such as for creating multi-touch surfaces, for finger tracking, head tracking etc.

Closed-group experiments were conducted at multiple workshops to demonstrate how easily one can use a gaming device to make classroom learning fun, interactive and more engaging to students. The experiments also analyze the user satisfaction of the proposed system. The study mainly focuses on cost-effectiveness, accuracy of tracking, user friendliness, time taken to complete the tasks etc. *Cost optimization* of a smart classroom device is the main contribution of the proposed model. Thus the model could be recommended to be used at remote and rural areas where the schools and colleges have difficulty to acquire expensive commercial multi-touch devices that help in student's learning activities. *Collaborative activities* were given priority because collaborative learning is more engaging and exploratory when compared to the traditional way of long hours of lectures.

With technology becoming more and more integrated to everyone's daily lives, using it in classrooms can help students in their future [4]. Apart from this, there are specific benefits to using technology in classrooms [5]. Here we have focused on how technological advances can be used to improve collaborative learning in classrooms. Few popular and commonly used tools in classrooms are iPads, Smart Boards, e-Readers, PDAs, Netbooks etc. The main drawback of these devices is that they do not encourage serious learning or support serious collaborative activities in classroom [6, 7]. The proposed model overcomes the above drawbacks by making classroom learning fun using a gaming device to create a cost-effective multi-touch surface that can be used for collaborative activities that add fun to education.

The upcoming technological advances in the *IoT era* have revolutionized the way of living of people. In classrooms, an interactive multi-touch system can help teachers to foster collaborative learning easily and effectively. Few popular devices that support multi-touch interface are Mitsubishi's *DiamondTouch* [8], *Microsoft's PixelSense* [9], *Apple's iPhone* [10], *CUBIT* [11] etc. These commercially available devices allow many features including multiuser experience, object recognition, multi-touch contact and much more. But the main disadvantage of these devices is that they are expensive. Also, most of them have mobility constraints and are platform restricted. It is not easy for students at remote areas to acquire such devices and use it for learning. So creating a multi-touch surface using a cost-effective model anywhere-anytime is a possible solution.

2 Related Work

Many approaches have been developed for creating an interactive whiteboard using Wiimote. Nintendo's Wiimote is a handheld device resembling a television remote. Wiimote, primarily intended to be used as a gaming device, was reverse-engineered to be used for finger tracking, interactive whiteboards and tablet displays, gesture recognition, head tracking for VR displays, spatial augmented reality, 3D motion tracking etc. as explained by Lee [12]. Sreedharan et al. [13] discusses about how gestures in 3D can be performed with the Wiimote and used to interact with a 3D environment. Through user studies the paper examines which gestures are appropriate to reflect common interactions like waving, with the 3D virtual world, Second Life [14]. Three applications using Wiimote was used for interaction by Shirai et al. [15]. Only data from the accelerometer is used. This problem is also described in [16]. For the drawing application both accelerometer and IR camera is used. In general the previous work within the field has mainly been focusing on the accelerometer in the Wiimote [13, 15–18] while only few have used the IR camera and sensor bar [6, 19]. In [20], Lorento-Leal et al. discuss about how a gesture recognizer takes Wiimote as an input device to estimate the movements of the user to compare the detected trajectory with the previously learnt movements in order to carry out the associated actions. The paper [21] presents a novel, inexpensive, coarse tracking system that determines a person's approximate 2D location and 1D head orientation in an indoor environment. Dooly [22] talks about collaborative learning and its importance. The various methods of taking collaborative learning beyond the classroom walls are explained with various examples.

Very recently in 2013, Aleotti et al. [23] discusses about how a Wiimote can be effectively used to create a 3D virtual world at low cost and with full user satisfaction. In [24], Cheong et al. describes about a multi-touch teaching station that have both standard educational collaborative software such as Classroom Presenter and the customized multi-touch teaching module that supports fingers' gesture controls on teaching materials. Even though the paper describes about using Wiimote to create a multi-touch surface, no survey or experiments were performed to highlight its use for performing collaborative activities in classroom and how students are more benefited with such a setup. We try to overcome some of these shortcomings through the survey conducted at multiple workshops on students from different cultural and geographical backgrounds (Fig. 1).

3 Hardware Setup

Nintendo's Wiimote is the main hardware device used in the model. It is paired with the PC using *Bluetooth* wireless connectivity. The *Wiimote's IR sensor* can sense any source of IR light. So once the Wiimote multi-touch surface is set up, an *IR-LED pen* could be used to write on the projected surface. The IR sensor in the Wiimote will track the motion of the IR-LED pen and the manipulations made on

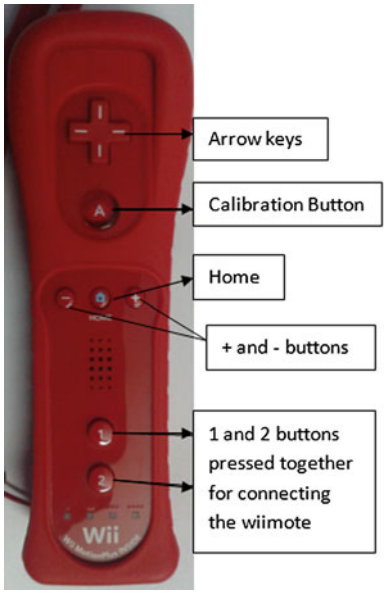


Fig. 1 Wiimote

the whiteboard are reflected on the system accordingly. An IR-LED array was also used for finger tracking using Wiimote for one of the experiments conducted. The IR- LED array and the reflector tape together can be used to track fingers.

We made the IR-LED pen, shown in Fig. 2, for writing on the multi-touch surface using hardware components such as an IR-LED, a momentary switch, a 2.2

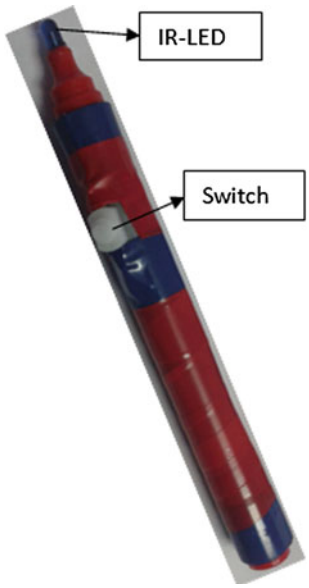


Fig. 2 Hand-made IR-LED pen

Ω resistor, two battery cells, few connection wires and the shell of a used marker pen. An IR-LED array was also made for finger tracking. The array strip board. The other basic hardware components used were windows 7 PC and a projector for projecting the computer screen onto a tabletop surface and also to wall according to the experiment being conducted. Nowadays it is common to have a PC or Laptop and a projector in smart classrooms. So the extra hardware cost to set up the model used for experimentation was Rs. 3,705.

3.1 Triangulation

In order to define the coverage area of Wiimote, it is necessary to calibrate and specify the area. Calibration process is repeated until the maximum area is found. Table 1 indicates some of the values obtained during the calibration of Wiimote producing maximum area of coverage. From Table 1, a generalized equation can be obtained which state that the distance of the Wiimote to the projected screen should be at least twice the projected height.

$$\text{i.e.,} d = 2 * d_h$$

Wiimote was projected from the ceiling in order to get maximum coverage area on the projected surface. So it is essential to make sure that the Wiimote and the IR-LED light source are placed as suitable angles so that the IR-LED light is not blocked from the Wiimote's camera which could affect tracking of the IR light source. We calculated the orientation of the Wiimote using a simple triangulation algorithm as explained below (Table 2).

Consider the coordinates as given in Fig. 3.

$$\alpha + \beta + \gamma = 180^\circ$$
$$\gamma = 180^\circ - \alpha - \beta$$
$$\theta = \alpha + \gamma/2 \quad (1)$$
$$\theta = \sin^{-1}(y/d) \quad (2)$$

Table 1 Values obtained from trial and error that gives maximum coverage area for the Wiimote during calibration

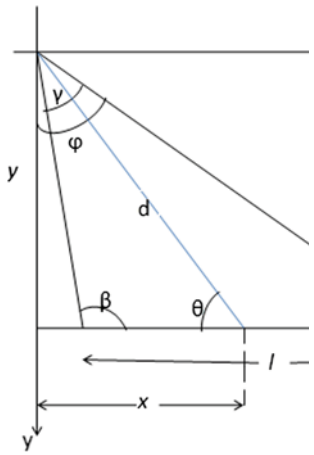
Width of screen (d_w)	Height of screen (d_h)	Wiimote distance (d)
68	55	118
57	53	101
49	38	70
30	23	50

Table 2 Hardware cost to setup the interactive multi-touch collaborative surface

Components used	Cost (in rupees)	Recurring cost[a]
Windows 7 PC/Laptop	Previously owned	Yes
Projector	6,000	Yes
Wiimote	2,600	No
Rechargeable batteries	400	Yes
Wiimote stand	300	Yes
Soldering kit	235	Yes
Bluetooth dongle	120	Yes
IR-LED pen	50	No
Total cost	9,705 INR	

[a]Recurring cost implies that it is one time purchase that add only to the initial setup cost. Cost may vary with location

Fig. 3 Triangulation coordinates



Equating Eqs. (1) and (2),

$$\alpha + \gamma/2 = \sin^{-1}(y/d)$$

$$\gamma = 2[\sin^{-1}(y/d) - \alpha] \quad (3)$$

The angle, γ, is the angle of coverage of Wiimote. γ needs to be optimized in order to achieve maximum coverage area on the projected screen. '*l*' is the length of the projected screen. α and β are the angles measured from the baseline, i.e., projected screen, to the Wiimote. Based on different values of 'd', we obtained few

possible Wiimote camera angles out of which one was most preferred for the experimentation. The most optimized coverage angle was 45° which gave the maximum coverage area of the projected surface. The rectangular coverage area of Wiimote increased with increase in distance. At 150 cm from the surface, area of coverage was 110 × 80. And at 240 cm, the coverage area was increased to 170 cm × 130 cm.

3.2 Data Transfer Speed

Bluetooth connectivity is used to connect Wiimote with the PC. An external Bluetooth dongle with Microsoft Bluetooth stack was used to identify Wiimote as a HID device in PC. Range up to which PC could detect Wiimote was within 7 m radius. The data transfer rate was around 10 Mbit/s; but the initial set-up time was 4–5 min on an average. Even after pairing the Wiimote once, it was necessary to pair Wiimote again with the same PC every time the whiteboard software was run. One main drawback of using Bluetooth connection was the quick draining of batteries in the Wiimote. The batteries had to be constantly replaced every few hours after continuous use (Fig. 4).

4 Wiimote Whiteboard Software

There are multiple open source Wiimote whiteboard softwares that could be used to setup the multi-touch surface. Few of them on which we worked on were as follows:

1. Johnny Lee's Wiimote whiteboard [12]
2. Uweschmidt Wiimote whiteboard [25]

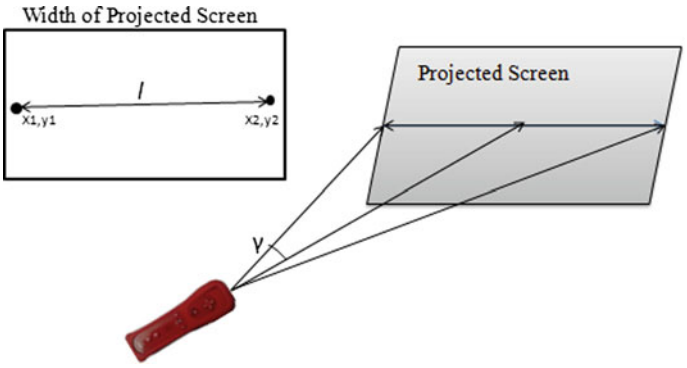


Fig. 4 Illustration of Wiimote's camera view and width of the projected surface. 2γ is the horizontal field of view

3. Linux Whiteboard [26]
4. iwiiboard [27]
5. Stéphane Duchesneau's Gtkwhiteboard [28]

All the Wiimote whiteboard softwares had drawbacks. Here, we modified the *Stéphane Duchesneau's Wiimote whiteboard software* written in Python and added the following properties:

- Tracked multiple IR-LED sources (4 sources was tracked instead of 2)
- Enabled scrolling
- Performed double click smoothly

LinuxWiimoteLib is the python Bluetooth library used to connect Wiimote to Windows 7 PC using the Microsoft Bluetooth stack. The modified software was then added to the open source software repository.

5 Experimental Design

We conducted closed group experiments at multiple workshops on volunteering participants within the age group of 15–20. The experiments were focused on school students. About 40 % of the students had previous experience of being in a virtual classroom while the rest have only heard of such online classroom setup. Most of them have at some point been part of a collaborative learning group and were familiar with few collaborative activities. Before the start of the experiments, all the students were given an initial introduction about how to use an IR-LED pen and also troubleshoot if any problem was to occur with the software or hardware. Figure 5 shows few participants during the experimentation. A timer was kept to track the time within which the tasks were completed. The different groups were:

- Two small groups containing 5 students each in a traditional classroom
- On a single traditional classroom with a strength of 35 students
- A networked classroom with a teacher present at a remote location and monitoring the activities of the remote classroom.

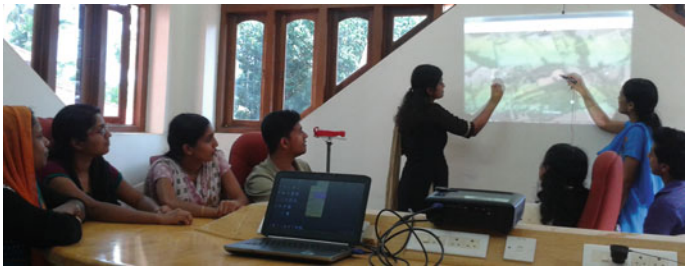


Fig. 5 Participants doing map navigation as part of the experiments conducted

Higher priority was given to those activities that were designed to be used with digital media. Classroom activities can be classified as productive and non-productive based on how well they contribute towards active learning. Here we have concentrated on improving the productive classroom activities to improve classroom learning. Collaborative activities were given priority in the experiments rather than individual activities.

5.1 Cognitive Reasoning for Choosing the Specified Collaborative Activities

- The activities were simple and of 5–10 min duration. So the students could complete the experiments faster.
- Activities such as Concept Mapping and Word-storming involved peer discussions that created an exploratory learning environment.
- Activities were chosen in such a way that they provide best way to measure the ease-of-use of the pen.
- It was also easy to measure the accuracy of tracking of the IR light source with these activities.

Three modes with varying number of participants were considered for the experiments to analyze the system's applicability with different groups. They were as follows:

Small Group Collaboration: Here, we chose two groups of 5 students each. An instructor was present live to monitor the activities of the groups. The Wiimote was setup over a round table. We used a projector to project the computer screen onto the table-top surface. The students were given 4 IR-LED pens to work on the activities.

Large Classroom Collaboration: Here, we conducted the experiment in a traditional classroom with 35 students. The computer screen was projected onto the classroom wall and the Wiimote was made to face the screen at the triangulated angle. Both the students and the teacher used the multi-touch surface for collaborating.

Networked Classroom Collaboration: Here, the teacher was present remotely and monitored the work carried out by students at a large classroom and multiple small individual groups. A-VIEW [7] software was used as the software collaborative tool. A-VIEW is an online collaborative E-learning tool which allows teachers to teach in a live interactive mode to various geographical locations across the world. The whiteboard application in A-VIEW was used by students to draw on the projected screen using the IR-LED pen (Fig. 6).

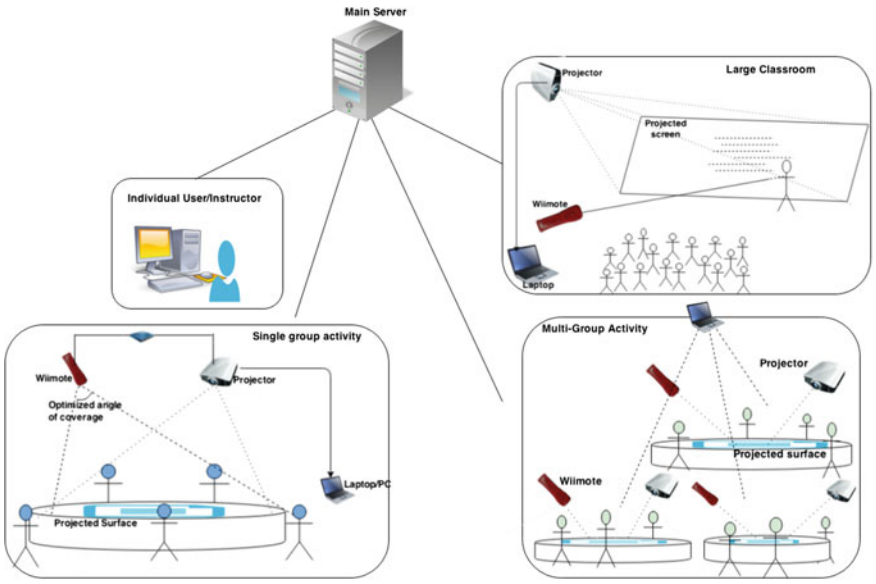


Fig. 6 Networked mode of setup including the different individual modes of setup

6 Optimizations Achieved

Using this approach we are offering certain optimizations as listed in Table 3.

7 Observations and Discussions

The main contributions offered (Table 4) through the proposed model and experiments were as follows:

- Few classroom activities designed mainly for digital media courses were chosen to show how they can be made more engaging and interesting using this setup; thus making learning more fun.
- Optimization of cost, software resources, time taken in performing such activities in classrooms, complexity of certain collaborative activities etc. was achieved.
- We also show how easily a person from a non-technical background can setup a multi-touch system effectively in a classroom that assist in active learning.

Even though there are many multi-touch displays currently available in the market, not all are suited well for collaborative activities. Specifications for a good collaborative multi-touch surface in a classroom are:

Table 3 Percentage of the optimization achieved for each parameter optimized using the proposed model

Parameters optimized	% Optimization[a]
Cost	79 %
Effort in setting up the multi-touch surface	73 %
Software	
1. Point tracking (calibration)	60 %
2. Accuracy of line drawing	63 %
3. Single and double clicks using the IR-LED pen	67 %
Time for completing classroom activities	71 %
Complexity of collaborative activities	68 %
Hardware	In progress
• Reducing the size of Wiimote	

[a]May vary due to uncontrollable factors

Table 4 Various classroom activities considered for the experiments and student response

Classroom activities	Designed for digital media?	Preferred by students?
Map navigation activity	Yes	Yes
Online collage making	Yes	Yes
Online polling	Yes	Yes
Highlighting notes	Yes	Yes
Motion sensing application	Yes	Yes
Activities done on A-VIEW whiteboard application	Yes	Yes
Virtual role-play	Yes	No
Finger tracking application	Yes	No
Problem solving—individual and collaborative	Not always	No
Wordstorming	Not always	Yes
Whiteboard square activity	Not always	No
Concept mapping	Not always	Yes

- Collaborative surface need to be at an optimal dimension of about 170 cm × 130 cm.
- At least 4–5 students should be able to collaborate on the surface freely.
- Devices used to setup the system should be simple and easy to use.
- It should be easily portable so that any surface can be used for collaboration at any time.
- The proposed system ensures that all the above specifications were met. The experimental results show that the system could be used to improve certain collaborative activities in classrooms. But not all the activities chosen for experimentation were preferred to be done using the proposed model (Fig. 7).

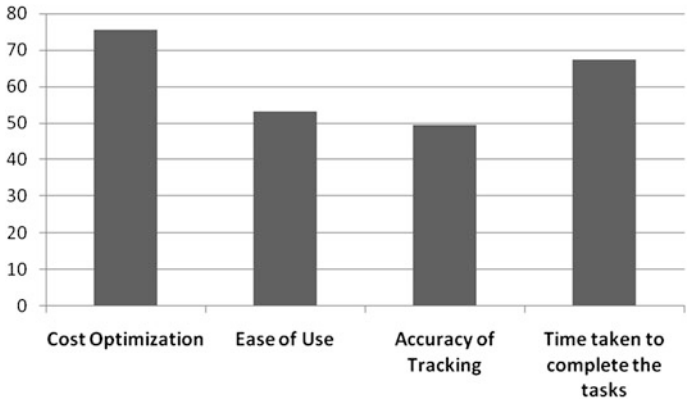


Fig. 7 Summarization of the experimental survey results

8 Conclusion and Future Works

Devices such as Wiimote, Kinect etc. though designed to be used in gaming, have multiple applications. In modern education where students are more tech-savvy it will be more effective to bring in such devices to classrooms that could aid in collaborative and exploratory learning experiences. The proposed system was found to be an effective tool for classroom collaboration based on the experimental results obtained. Here we have achieved cost optimization as well as optimized the effort for setting up a multi-touch surface in a classroom. The future work would include more optimization of hardware and software resources. Only the IR sensor in the wiimote is used to capture the IR light source from the pen. So the hardware can be modified to be used as a wearable device. Wiimote could be paired with the computing device through Wi-Fi so that Internet connectivity is also available to the user. Multiple Wiimotes can be purchased second hand. We are currently working on providing an optimized software solution to the hardware problems faced during calibration and setup.

Acknowledgments We are grateful for the support and facilities provided by Amrita Vishwa Vidyapeetham University and Amrita E learning Research lab in successfully completing the experimental work and in data analysis.

References

1. Shi, Y., Xie, W., Xu, G., Shi, R., Chen, E., Mao, Y., Liu, F.: The smart classroom: merging technologies for seamless tele-education. IEEE Pervasive Comput. **2**(2), 47–55 (2003)
2. Nintendo: www.nintendo.com
3. Nintendo: In depth regional Wii coverage. http://wii.nintendo.com/controller.jsp. Accessed 23 March 2007

4. Shih, M., Feng, J., Tsai, C.-C.: Research and trends in the field of E-learning from 2001 to 2005: a content analysis of cognitive studies in selected journals. Comput. Educ. **51**(2), 955–967 (2008)
5. Cope, C., Ward, P.: Integrating learning technology into classrooms: the importance of teachers' perceptions. Educ. Technol. Soc. **5**(1), 2002ISSN 1436–4522 (2002)
6. Lai, C.-C.: The advantages and disadvantages of computer technology in second language acquisition. Doctoral Forum Natl. J. Publishing Mentoring Doctoral Student Res. **3**(1) 2006
7. A-View: http://www.aview.in/
8. DiamondTouch: http://www.circletwelve.com/
9. Microsoft PixelSense: http://www.microsoft.com/en-us/pixelsense/default.aspx
10. Apple: https://www.apple.com/in/
11. Cubit: http://labs.nortd.com/cubit/
12. Lee, J.C.: Hacking the Nintendo Wii remote. IEEE Pervasive Comput. **7**(3), 39–45 (2008)
13. Sreedharan, S., Zurita, E.S., Plimmer, B.: 3D input for 3D worlds. In: OZCHI '07: Proceedings of the 2007 Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of Australia on Computer-Human Interaction: Design: Activities, Artifacts and Environments, pp. 227–230. ACM (2007)
14. Linden Lab, Second Life: www.secondlife.com
15. Shirai, A., Geslin, E., Richir, S.: Wiimedia: motion analysis methods and applications using a consumer video game controller. In: Sandbox '07: Proceedings of the 2007 ACM SIGGRAPH Symposium on Video Games, pp. 133–140. ACM (2007)
16. Pedersen, L.J.F., Tøgern, A.S.: Styring af objekter i 3D vha. Wii-controller. Under graduate report at DIKU (2007)
17. Schou, T., Gardner, H.J.: A wii remote, a game engine, five sensor bars and a virtual reality theatre. In: OZCHI '07: Proceedings of the 2007 Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of Australia on Computer-Human Interaction: Design: Activities, Artifacts and Environments, pp. 231–234. ACM (2007)
18. Looser, J., Billinghurst, M., Grasset, R., Cockburn, A.: An evaluation of virtual lenses for object selection in augmented reality. In: GRAPHITE '07: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia, pp. 203–210. ACM (2007)
19. Slay, H., Sieborger, I., Hodgkinson-Williams, C.: Interactive whiteboards: real beauty or just, lipstick? Comput. Educ. **51**(3), 1321–1341 (2008)
20. Lorento-Leal, A., Fernandez, J., Montero, J.M.: Development of a Wiimote-based gesture recognizer in a microprocessor laboratory course. Innovation Project, UPM
21. Hallaway, D., HoLLerer, T., Feiner, S.: Coarse, inexpensive, infrared tracing for wearable computing. In: 2013 IEEE International Symposium on Wearable Computers
22. Dooly, M.: Constructing knowledge together. In: Telecollaborative Language Learning: A Guidebook to Moderating Intercultural Collaboration Online (2008)
23. Aleotti, J., Caselli, S., Micelli, V.: Assessment of the Wiimote as a tangible user interface for interactive virtual environments. In: 2013 IEEE International Conference on Systems, Man and Cybernetics
24. Cheong, S.N., Yap, W.J., Chan, M.L.: Design and development of a cost effective Wiimote-based multi-touch teaching station. Inf. Technol. J. **9**(3), 589–594 (2010). ISSN 1812-5638
25. Uweschmidt.org, wiimote-whiteboard: http://www.uweschmidt.org/wiimote-whiteboard
26. Linux-Whiteboard: https://code.google.com/p/linux-whiteboard/
27. Iwiiboard: http://www.iwiiboard.com/index.html
28. GTK Wiimote Whiteboard: http://www.stepd.ca/gtkwhiteboard/

An Effective Implementation of Exudate Extraction from Fundus Images of the Eye for a Content Based Image Retrieval System Through Hardware Description Language

C. Gururaj, D. Jayadevappa and Satish Tunga

Abstract Data retrieval plays a critical role in the progress of the technology in the present day technological scenario with huge databases. Content Based Image Retrieval (CBIR) is one of the popular image retrieval techniques which find its application in varied fields including medical image analysis, historical research, military applications etc. Teleconferencing is gaining widespread acceptance in the field of medical analysis. One of the methods for testing through the above method is to use images of the fundus of the eye. Although there are CBIR algorithms that are accurate, their implementation on PC based systems suffer from long execution time. The paper proposes to accelerate the algorithm through its implementation in a mixed hardware/software platform. The first towards this process is implementation through a Verilog HDL code which can be used on a VLSI system. The extraction of features from these images may indicate the presence of infirmities namely exudates that are determined and its possible implementation through Verilog HDL is addressed in this paper. An improvement in terms of effective implementation was observed using the standard DRIVE database. The algorithm has been implemented and optimized on Xilinx ISE Design Suite version 14.2 and simulated on Modelsim simulator version 10.1d.

Keywords Exudate · Content based image retrieval · CBIR · Fundus · Verilog · Optical disc

C. Gururaj (✉)
Department of Telecommunication, BMS College of Engineering,
Jain University, Bengaluru, India
e-mail: gururaj.c@gmail.com

D. Jayadevappa
Department of Instrumentation Technology, JSSATE (VTU), Bengaluru, India
e-mail: devappa.22@gmail.com

S. Tunga
Department of Telecommunication, MSRIT, Jain University, Bengaluru, India
e-mail: satish.tunga@msrit.edu

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_27

1 Introduction

Over the last few decades, the most notable development in regard of technological progress has been with respect to improvements in the medical field. Following the invention of the computer, these developments have helped medical technology leapfrog into a modern era of medical equipment which renders 24 × 7 help to doctors on hand, bringing with them a consistency and accuracy at levels humans find extremely hard to achieve. These technologies have been a great boon to doctors especially, as it relieves some or most of their work in terms of human effort. There is, however, still room for improvement of the scope of the technologies in use today. For example, wide spread, cost-effective implementation of modern technology in the medical field is yet to be completely reached. One of the many such fields of medical aid is the detection of infirmities of the eye. This process is accomplished through a popular image retrieval technique called Content Based Image Retrieval (CBIR) technique.

CBIR is a technique used for retrieving images on the basis of automatically derived features such as color, texture and shape. Although content-based retrieval provides an intelligent and automatic solution for efficient searching of images, the majority of current techniques are based on low level features. One of the many applications of CBIR is in the area of medical image analysis of exudate extraction from the fundus images of the eye.

The fundus of the eye can be examined thoroughly through Ophthalmoscopy and fundus photography which can be used to detect cases of diabetic retinopathy. The above methods are used to find the presence of exudates in the fundus. An Ophthalmoscope is used in Ophthalmology to examine the eye, and it helps in on spot examination of the patient by the doctors; whereas in fundus photography, using a high power microscope and camera the image of the fundus of the eye can be captured. One of the major virtue of the fundus photography is that it allows patients data to be verified even though the patient themselves are in a different part of the world.

The authors Ibrahim et al. [1] discuss about one such area where work is being done is the detection of afflictions of the eye. There has been a lot of work done by research scholars from all over the world, to automate the process of taking the image of the fundus of the eye, detect the afflictions, if any, and provide feedback to the patient, and perhaps suggest a course of action as well. The authors Salazar-Gonzalez et al. [2] highlight that key part of course, is the detection of affliction.

2 Content Based Image Retrieval

Content based image retrieval (CBIR), as the name suggests, involves retrieving images based on the content of the images and not based on textual tags embedded within them. Usually CBIR is performed based on certain features such as color, texture and shape. The input to the CBIR may be a keywords or an image, and the

query run is capable of acquiring images similar to the supplied input. CBIR is attractive because of its proportionate effort and time taken in labeling large sets of images. A system that can acquire related images without using these labels poses a striking solution.

Using CBIR, the images of the patient can be used to find similar images of other patients and by studying the outcomes of the histories of the obtained through CBIR, a prediction of the current patient's condition can be computed.

Since a massive database can be made available from hospitals all over the world, a very detailed and accurate prediction can be made using the information at hand, which would otherwise be impossible for anyone to manually compare large numbers of similar cases. Furthermore, possibility of human error in diagnosis of any affliction is reduced, and this system can also be used to cross-verify any cases of positives or false positives.

A method for automatic diagnosis of retinopathy has been tried previously by Chaum, Edward et al. in their research [3]. They present a system to assist doctors to assess their patients and also their decisions regarding their medical state. They have used distinctive features for extraction of images, and have achieved an accuracy of 91 % for appropriate image retrieval.

In order to do apply CBIR on fundus images of the eye for detection of afflictions, an algorithm is required to first extract these defects from the fundus image of the eye. The requirement for this implementation is its efficiency. For CBIR, the database has to be queried very quickly, which is possible through a hardware implementation.

The implementation should be such that it cannot take a large time to compute or find the exudates for one image. There needs to be a design method that is accurate and fast enough to list and find the required images in the database. A novel algorithm that separates these exudates from other components in the image was developed by the authors of this paper [4] and the present paper deals with an attempt towards the effective implementation of this algorithm through the Verilog HDL on a Xilinx ISE design suite.

3 Exudate Extraction

One of the major tests performed by an ophthalmologist during an eye checkup is the examination of the internal region of the eye, also known as the Fundus. This region includes all major regions of the inner eye, found right opposite to the lens. This includes the retina, optic disc, macula, fovea and also the posterior pole [5]. Blood vessels within the eye are also a part of this region. This region is usually examined with the help of ophthalmoscope [6], by keeping the instrument as close as possible to the eye and pointing bright light at an angle to the eye. This test is usually performed on the eye, after it has been dilated. In order to document the results obtained, or in order to perform diagnosis at a later stage, a camera called a "Fundus camera" is usually used in order to perform ophthalmoscopy [7].

Most of these medical signs stem from various diseases such as diabetic retinopathy, Retinal Vasculitis, Obliterative Retinal Vasculopathy to mention a few. With the advancement in science and technology, various image processing techniques to automate the process of detection of these medical conditions have been researched upon. This would in turn assist the doctor in his/her diagnosis. Sometimes, there is also a need by the doctor to refer previous cases of patients to improvise his diagnosis. In such a case, a speedy access to the database with accurate information would be a must. Therefore, an efficient CBIR algorithm becomes the need for the day.

Out of all the diseases related to the eye, diabetic retinopathy is considered as the most sight-threatening disease existing [8]. One of the most important medical sign pointing towards the existence of this disease is the presence of precipitates of lipoproteins or other proteins leaked from the blood vessels into the eye. These proteins form what is known as hard exudates [9].

4 Methodology

Histograms are primarily used to find the optic disk in the image [10]. The logic followed is that in the fundus image, the values of pixel intensities that fall in the optic disk regions are similar, but a threshold cannot be used to find this region as this will result in the algorithm that varies with changes in the lighting and/or the camera that captures the fundus images. A more elegant solution is required—one that is accurate and time efficient. The following steps are used to detect the optic disk, and then mask it to obtain the exudates present in the picture (Fig. 1).

RGB Color space is generally used to represent the input image. This image is separated into its constituent channels and only the green channel is retained. This is an essential step in the process of detecting the exudates, and is popular in its use for the same. Green channel offers balance in terms of intensity of the pixels acquired. The RBCs present would give a red hue to the fundus of the eye due to which separation of optic disc and exudates become challenging. In case of the blue channels, the opposite happens i.e. the image appears dark in color thereby separating the features is a daunting task. The comparison of the same fundus image in the three different channels is seen in Fig. 2 which is the original RGB colour image, Fig. 3 is the image with red channel only, Fig. 4 is the image with green channel only and Fig. 5 is the image with blue channel only. Therefore the green channel of the RGB image is chosen for the further steps of the algorithm.

From the histogram of the green channel it can be noticed that maximum number of pixels inhabits values between 0 and 5 and these pixels are not required to detect the optic disc. On the other pointer, the optic disc has the highest intensity value. The difference in the number of pixels, between adjacent bins in the histogram, will provide a threshold to separate the optic disk from the rest of the information contained in the image. To eliminate exudate region, image erosion is performed with three different image resolutions i.e. 1500 × 1152 pixels, 375 × 288 pixels and

Fig. 1

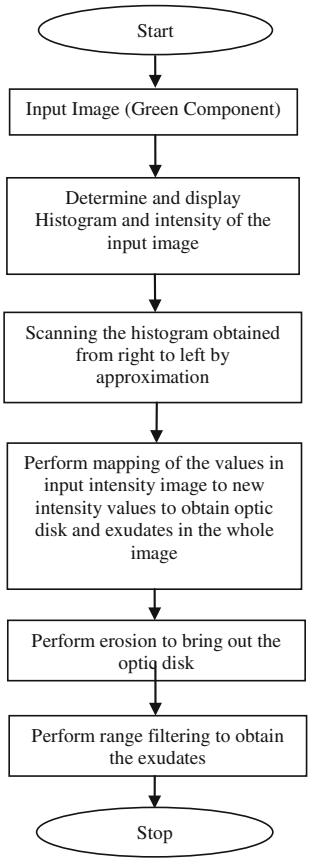


Fig. 2

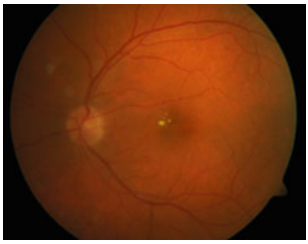


Fig. 3

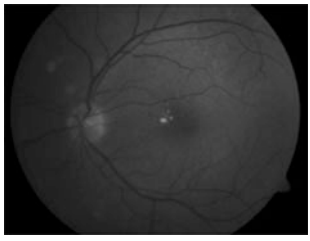


Fig. 4

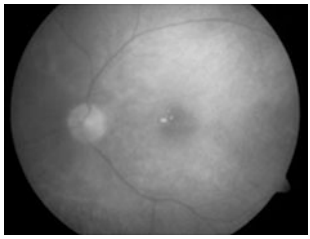
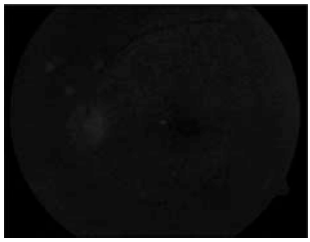


Fig. 5



on the DRIVE database images which have a resolution of 768 × 584 pixels. The eroded image may not contain the total area of the optic disk. In order to reconstruct the optic disk, a range filter is applied [4]. The range filter is a transform that is used to set all the pixels in a given window to the local range of the window. The application of the range filter is summarized in (1)–(3). The reason for the application of the range filter is that the pixel values, if spread non-uniformly, are brought together using the range filter. The portions of the optic disk that were not detected in the previous steps and the portions that were lost due to image erosion are retained in this step.

$$\mathrm{mx} = \max(\mathrm{V}(\mathrm{i}, \mathrm{j})) \tag{1}$$

$$\mathrm{mn} = \min(\mathrm{V}(\mathrm{i}, \mathrm{j})) \tag{2}$$

$$\mathrm{V}(((\mathrm{i}+\mathrm{j})/2), ((\mathrm{i}+\mathrm{j})/2)) = (\mathrm{mx}-\mathrm{mn}); \tag{3}$$

where V(x, y) = value of intensity of the pixel at (x, y) and ∀ i, j is the extremes of the local neighborhood. The formulae are repetitively applied for all the pixels in the image, which will result in an image where the optic disk appears brighter than the background. The above steps are illustrated from Figs. 6, 7, 8, 9, 10, 11 i.e. input image, contrast adjusted image, after image erosion, application of range filter, resultant bit mask, exudates detected image respectively when verified through the MATLAB code [4].

The above algorithm was also implemented on a VLSI platform to improvise the efficiency of the algorithm. The code was written through the Verilog Hardware Description Language (HDL) platform on a Xilinx ISE Design Suite version 14.2 and simulated on Modelsim simulator version 10.1d.

Fig. 6

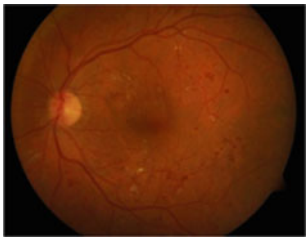


Fig. 7

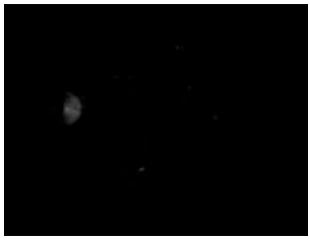


Fig. 8

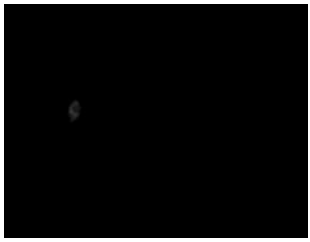


Fig. 9

Fig. 10

Fig. 11

5 Results

The algorithm was tried for images of three different resolution databases. A total of 40 images were considered in each of the database. The first database of 40 images had a resolution of 1500 × 1152 pixels with images having mild cases of retinopathy. The second database had images with a resolution of 375 × 288 pixels; also having diabetic retinopathy. The final database was from the popular DRIVE [11] database. The photographs for the DRIVE database were obtained from a diabetic retinopathy screening program in The Netherlands. The screening population consisted of 400 diabetic subjects between 25–90 years of age. Forty photographs have been randomly selected, 33 do not show any sign of diabetic retinopathy and 7 show signs of mild early diabetic retinopathy. Each image has been JPEG compressed. The MATLAB based results are as shown in Figs. 12, 13, 14, 15 [4]. The simulation results considered with a common clock frequency of 330 MHz is obtained through the Modelsim wave window and the screenshots of the same are shown in Fig. 16.

Fig. 12

Fig. 13

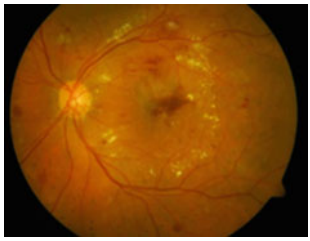


Fig. 14

Fig. 15

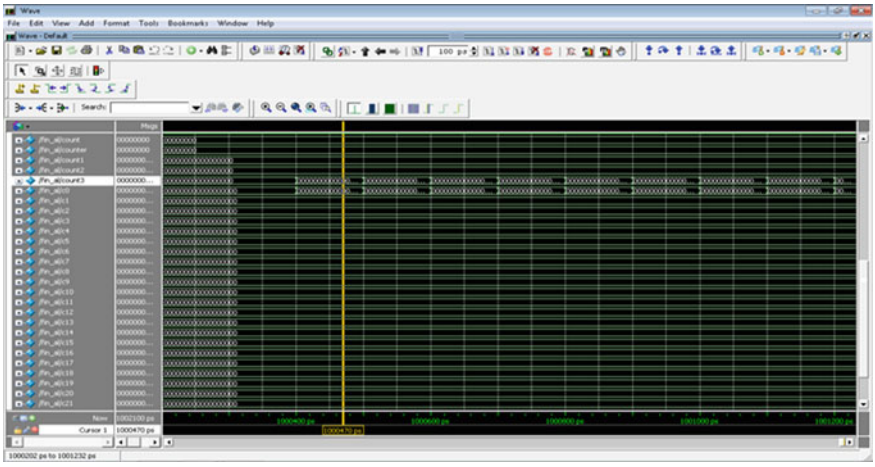


Fig. 16

The detailed analysis of the algorithm has yielded the graph given in Fig. 17. The graph consists of four important stages and their timing analysis for each stage of execution. The process of determining the threshold consumes very less time as compared to that of the contrast adjustment and exudate extraction. The erosion and dilation process which forms the bottleneck of the algorithm proceeds with the

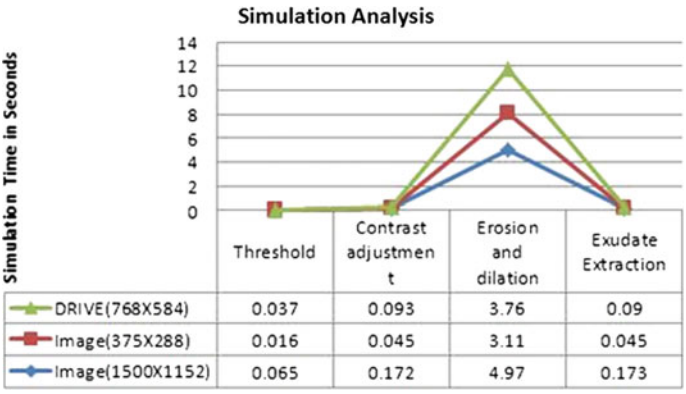


Fig. 17

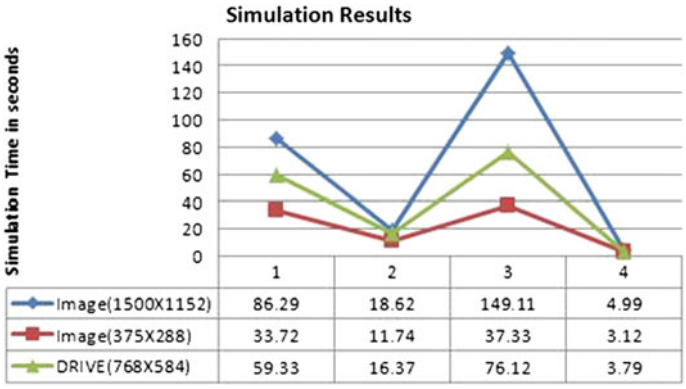


Fig. 18

maximum time occupancy towards the simulation. The graph is considered for all the three databases.

The timing analysis of the entire process with respect to different databases and algorithms are shown in Fig. 18. It has to be noted that x-axis indicates various algorithms that have been used to implement the exudate extraction process. The first algorithm considered is the Machine learning approach to automatic exudate detection in retinal images from diabetic patients [12] wherein only the optic disk extraction is being considered. However the paper also details feature extraction, which has not been considered for the purpose of analysis. The second algorithm considered is the automatic tracing of optic disc and exudates from colour fundus images using fixed and variable thresholds [13]. The third algorithm deals with fundus image features extraction [14]. Our algorithm is the fourth on the given graphs.

6 Conclusions

The algorithm described in this paper achieves better results in terms of both time complexity and efficiency as compared to the other algorithms in detecting the optic disk and the exudates from the fundus images of the eye, while using simple image processing techniques. But the algorithms efficiency is yet to be tested on larger databases. Also the implementation of the Verilog HDL on a hardware platform such as System-on-Chip or FPGA would encompass challenges. However it is a known fact that the implementation once completed on a hardware platform would be highly efficient in terms of the processing times when compared to a general PC implementation. This scenario would especially be useful in case of the medical image processing scenario wherein as the day progresses, several terra bytes of image data that are being produced in hospitals.

Acknowledgments The authors would like to acknowledge the management of BMS College of Engineering, Bangalore for sponsoring the publishing of this work through the TEQIP—II grants.

References

1. Ibrahim, F., Ali, J.B., Jaais, A.F., Taib, M.N.: Expert system for early diagnosis of eye diseases infecting the Malaysian population. In: TENCON 2001. Proceedings of IEEE Region 10 International Conference on Electrical and Electronic Technology, vol. 1, pp. 430, 432 (2001)
2. Salazar-Gonzalez, A., Kaba, D., Li, Y., Liu, X.: Segmentation of blood vessels and optic disc in retinal images. IEEE J. Biomed. Health Inf. **18**(99)
3. Chaum, E., Karnowski, T.P., Govindasamy, V.P., Abdelrahman, M., Tobin, K.W.: Automated diagnosis of retinopathy by content-based image retrieval. Retina **28**(10), 1463–1477 (2008)
4. Gururaj, C., Jayadevappa, D., Tunga, S.: Novel algorithm for exudate extraction from fundus images of the eye for a content based image retrieval system. In: 4th IEEE International Conference on Control System, Computing and Engineering (ICCSCE—2014), ISBN-978-1-4799-5685-2, pp 395–400, Penang, Malaysia, 28–30 Nov 2014
5. Asim, K.M., Basit, A., Jalil, A.: Detection and localization of fovea in human retinal fundus images. In: 2012 International Conference on Emerging Technologies (ICET), pp. 1, 5. 8–9 Oct 2012
6. Pereira, T., Barbeiro, P., Lemos, J., Morgado, M., Silva, E.: Digital image acquisition for ophthalmoscope. In: 2012 IEEE 2nd Portuguese Meeting in Bioengineering (ENBENG), pp. 1, 6. 23–25 Feb 2012
7. http://www.nlm.nih.gov/medlineplus/ency/article/003881.htm. Accessed 2 Feb 2014
8. Vidyasari, R., Sovani, I., Mengko, T.L.R., Zakaria, H.: Vessel enhancement algorithm in digital retinal fundus microaneurysms filter for nonproliferative diabetic retinopathy classification. In: 2011 2nd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), pp. 278, 281. 8–9 Nov 2011
9. http://www.aao.org/theeyeshaveit/optic-fundus/hard-exudates.cfm. Accessed 2 Feb 2014
10. Russ, J.C.: Processing binary images. In: The Image Processing Handbook, CRC Press, North Carolina, p. 468 (2011)
11. Gupta, A., Moezzi, S., Taylor, A., Chatterjee, S., Jain, R., Goldbaum, I., Burgess, S.: Content-based retrieval of ophthalmological images. In: Proceedings., International Conference on Image Processing, vol. 3, pp. 703–706. IEEE (1996)

12. Sopharak, A., Dailey, M.N., Uyyanonvara, B., Barman, S., Williamson, T., Nwe, K.T., Moe, Y.A.: Machine learning approach to automatic exudate detection in retinal images from diabetic patients. J. Mod. Opt. **57**(2), 124–135 (2010)
13. Reza, A.W., Eswaran, C., Hati, S.: Automatic tracing optic disc and exudates from color fundus images using fixed variable thresholds. J. Med. Syst. **33**(1), 73–80 (2009)
14. Li, H., Chutatape, O.: Fundus image features extraction. In: Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 4 (2000)

Energy Efficient Dynamic Reconfiguration of Routing Agents for WSN Data Aggregation

S. Prabhavathi, A. Subramanyam and A. Ananda Rao

Abstract Wireless Sensor Network (WSN) has encapsulated the major attention from the research group owing to its potential features as well as some of its unsolved issues. From the various issues, an energy issue is particularly found to be a root cause of majority of the problems. In the past, there are various research attempts to enhance the network lifetime of WSN, but very few of the studies has proved fruitful. Moreover performing optimization of energy from routing as well as load-balancing perspectives are not witnessed any standard research attempts. Hence, this paper discusses about a framework that uses standard radio and energy model and performs energy optimization by introducing a novel routing agent. The routing agent is incorporated within aggregator node and base station to perform dynamic reconfiguration in case of energy depletion. Compared with standard LEACH algorithm, the proposed technique has better energy efficiency within optimal data aggregation duration.

Keywords Energy · Load balancing · Routing · Wireless sensor network

1 Introduction

Wireless sensor network (WSN) is one of the most upcoming areas of commercial utilization for faster and effective communication without the presence of any infra-structures [1]. Generically, a wireless sensor network will consist of the huge quantity of electronic sensing motes with high radio frequency and base station,

S. Prabhavathi (✉)
Department of CSE, RYMEC Ballary, Ballari, India

A. Subramanyam
Deparmentt of CSE, AITS, Rajampet, India

A. Ananda Rao
JNTUA, Anantapur, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_28

which characteristically act as a relative link with many other processing networks facilitating potential processing of data, location of storage, and an access point to the sensor motes in its network. Sensor motes sense its environment, gather the data sensed and transmit it to the base station. Unfortunately, the area of wireless sensor is also accompanied by various issues e.g. restriction of power, computational capability, and memory. Wireless sensor network can be deployed in various application right from military, biomedical, and environmental utilization [2]. It is very difficult to detect the route and retain it as the confined quantity of power and frequent alteration in the locus of the position of the sensor nodes gives rise much changes which are very dynamic and tough to forecast. Although there has been an extensive research work towards energy [3–6] but still it has remain a major obstruction towards development of the effective routing algorithm in wireless sensor networks. In the cluster first approach, the network selects the cluster formation initially and then cluster leader is selected whereas in leader first approach, cluster leader is selected first and then formation of the cluster takes place. Initially the sensor nodes are randomly deployed. Usually, the different sensor nodes are used in real times and therefore heterogeneous network is formulated with varied scales of energy retention. Usually, the sensor nodes which have highest score of residual energy are selected as cluster leader where the entire cluster leader formulated their personal communication range and thereby formulates clusters. The sensor nodes residing in each clusters transmits their gathered information and forward it to the cluster leader which in turn forward it to the sink. In majority of the networking situation, cluster leader was assumed to transmit the data directly to the sink located at a specified distance from the cluster leader. The phenomenon of clustering basically assists in diminishing the memory utilization of the routing table updates, preserves the communication channel capacity, enhances the cumulative lifetime of the network, and minimizes the chances of collection of redundant data [7].

2 Related Work

This section discusses about the prior significant contribution of various researchers towards energy efficiency in wireless sensor network. Some of the prior study e.g. [8] shows that energy efficient route can be mechanized for enhancing the lifetime of network. This fact was preliminarily supported by the study conducted by Chang et al. [9]. The author exclusively highlighted that network lifetime tends to degrade if the routes are established among the nodes with less energy. Haggerty et al. [10] presented results from @scale; one of the largest deployment in terms of mote-years yet published. They were successful in achieving their science goals while testing several hypotheses about network dynamics, and reinforced emerging design practice on the construction of this type of network.

Gu et al. [11] have investigated various sorts of trade-offs existing in the energy-aware concepts of WSN and presented a solution for it using task planning.

The evaluation was performing using real-time sensors using their routing protocol. The outcome of the study shows prominent energy preservation over arbitrary synchronization managements.

Alwadi and Chetty [12] examined Isolet, ionosphere and forest cover type datasets from the UCI repository to emulate the wireless sensor network scenario. From their simulation results, they show that it is possible to achieve two important objectives using the proposed scheme: (1) Increase the lifetime of the wireless sensor network, by using optimal number of sensors, and (2) Manage sensor failures with optimal number of sensors without compromising the accuracy. Nikravan et al. [13] have presents a technique that indirectly attempts to reduce the energy depletion using fuzzy logic. The system selects both energy factor as well as transmission rate for selecting the next immediate node for performing data aggregation process. The outcome of the study shows that proposed system furnishes optimal enhancement of packet delivery ratio as well as restores energy conservation even in care of dynamic wireless environment.

Lakshmi [14] proposes novel energy efficient algorithm Fuzzy Dynamic Power Control Algorithm for Wireless Sensor Networks. Although it can be seen that there are huge blocks of research work being done in the field of routing issues related to power consumption in wireless sensor network using Fuzzy Logic. But every approach has yielded enhancement in the performance using Fuzzy only 50 %, whereas we have achieved around 80 % in performance efficiency using dual level of fuzzy inference system. Therefore, it can be seen that majority of the prior research work has been focused on distance between the cluster leader and the sensor motes which ignore majority of the other parameters which influences the power draining and maximizing the network lifetime. Understanding the research gap from the review of the prior research work, the current research work considers tri-factors e.g. magnitude of remoteness between the cluster leader and the sensor motes, Euclidean's distance between the cluster leader, and finally the power factor of the cluster leader.

3 Problem Identification

Organizing sensor networks into clustered architectures has been extensively explored over the previous couple of years, leading to the surfacing of an excellent variety of tasks-specific clustering protocols in WSN. Clustering is a cost-effective approach for information aggregation within the WSN. Every sensor node within the network sends information to the aggregator node, meaning the cluster leader, then the cluster leader performs the aggregation method on the received information and sends it to the bottom station. However, performing the aggregation operation over the cluster leader causes the vital energy drains. In case of a homogeneous type of WSN, the cluster head can soon die out, and once more re-clustering should be done, which once more causes energy consumption. In this study, the focus is placed on designing a protocol that can extract information for data dissemination within the range of clusters. The termination of the old clustering technique and the

design of a new clustering technique will be highly avoided to maintain algorithm efficiency. Along with energy, routing and load balancing is another parameters that also requires significant attention for designing optimization solution in wireless sensor network. From the literatures, the existing geographic routing is found with various challenges. Two factors of load balancing techniques with respect to optimization was ignored much viz. Data aggregation duration and energy consumption. The data aggregation duration is very often estimate from the time taken by the physically sensed aggregated data by the aggregator node to the base station. However, possibly, if a new model is design that aims to enhance the data aggregation by introducing a new module within aggregator node and base station need to be involved while estimating data aggregation time. This types of evaluation will show effectiveness of the proposed module with respect to energy optimization as well as load balancing, which is missing in the literatures. Hence, the proposed system aims to overcome such issues.

4 Proposed Model

The proposed model introduces a novel model that can perform dynamic reconfiguration of the new routing agents as the designated routing agents depletes their residual energy to a large extent in the process of data aggregation. However, the primary goal of the study will be to accomplish much better energy efficiency while disseminating the data from the sensors to the sink in wireless sensor network. The adoption of the routing agent in the current study is motivated by the work discussed by Shakshuki et al. [15]. According to Shakshuki et al. [15], the term 'routing agent' is a critical module that has the capability to perform communication with all the sensor nodes in the wireless environment. The routing agent is also responsible for assigning the cluster leader and evaluates their decision making strategies for energy efficient routing in wireless sensor network. In the current work, a novel technique will be introduced that will incorporate the capabilities of an agent to explore the energy efficiency. The block diagram of the study is as shown in Fig. 1.

The proposed study formulates a routing agent using graph theory that ensures energy efficient data aggregation technique within less duration and more computational capability. The proposed system discusses about the agent based routing structure where the nodes are managed into numerous unit stratum cluster of multiple size for the purpose that cluster can communicate with the routing agent. Consider a WSN topology, considering C as a cluster with N_{sen} as total number of sensor nodes.

$$\sum_{i=1}^{C-1} \eta_i < N_{sen} \leq \sum_{i=1}^{C} \eta_i \tag{1}$$

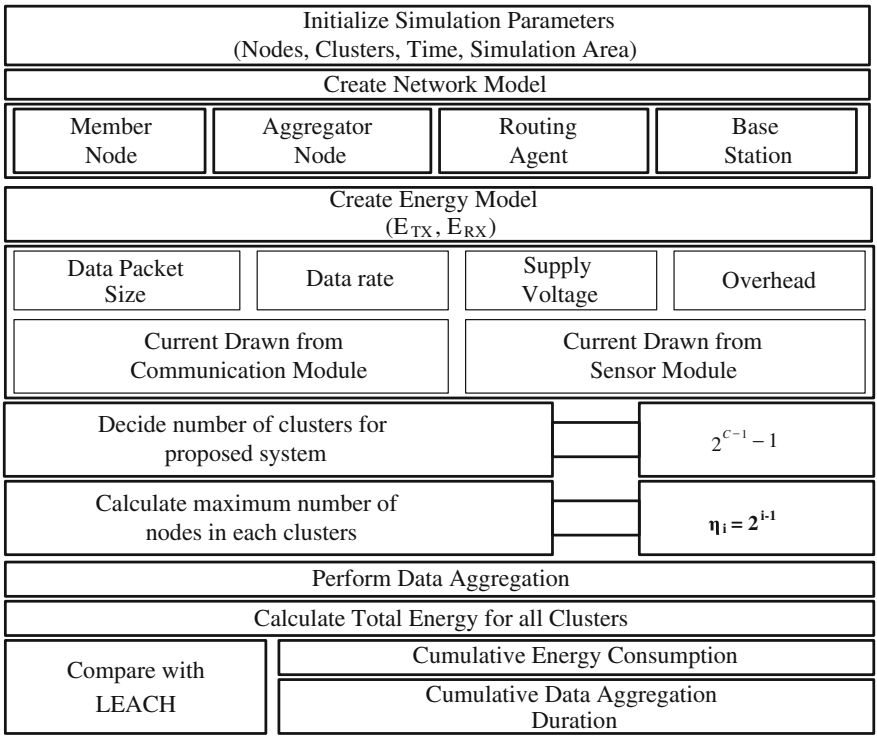


Fig. 1 Block diagram of the proposed study

In the above equation, η is acting as the highest number of the sensor node in the ith cluster that can be represented in the form of geometric progression as follows:

$$\eta_i = 2^{i-1} \tag{2}$$

Therefore, substituting the value of $i = 1, 2, 3, \ldots..$, it can be seen that the ηi will generate the values of 1, 2, 4, 8 etc. as a geometric progression. Therefore, Eq. (1) can be amended as below,

$$2^{C-1} - 1 < N_{sen} \leq 2^C - 1 \tag{3}$$

It is quite evident that a smaller dimension of the cluster size can reduce the process of data dissemination technique. Therefore, the proposed system considers the Cth cluster (last cluster) to be the ultimate member to perform communication with the routing agent. Hence, during the cluster formation phase, the members are allocated to (C−1)th clusters initially coming down to Cth cluster for member node allocations within that cluster. The system considers the time period of the cumulative data aggregation to be lower bound by the total number of sensor nodes.

Considering a topology with Eq. (3), it can be seen that the initial (C−1) clusters are completely filled with the member nodes and needs a following cumulative timeslot for forwarding all the aggregated data to routing agent as,

$$\sum_{i=1}^{C} \eta_i = 2^{C-1} - 1 \tag{4}$$

And the last cluster (Cth cluster) is allocated with $N_{sen} - 2^{C-1}$ member nodes. Which means that the aggregator node will consume time period of $(2^{C-1}-1)$ to aggregate the physical data and $(N_{sen} - 2^{C-1}+1)$ time period to forward the data to the routing agent. Therefore the complete time period (T_p) to perform data aggregation can be represented as

$$\begin{aligned} T_P(N_{sen}) &= \left(2^{C-1} - 1\right) + \left(N_{sen} - 2^{C-1} + 1\right) \\ &= N_{sen}(\max) \end{aligned} \tag{5}$$

5 Implementation & Results

The implementation of the proposed research work is done on Windows 32-bit OS with 1.84 GHz dual core processor considering Matlab as programming language. The proposed work is experiment on a simulation area of 100 × 100 m^2. We have considered a various scenarios with 100 nodes with arbitrary distribution. The preliminary power of sensor motes is initialized at 0.1 Joules with 100 rounds. The proposed system is compared to frequently used LEACH [15] protocol. The network simulation parameters considered are as shown in Table 1.

The current framework considers that all the cumulative energy depletion are happening overheads as well as due to spontaneous receiving and transmitting of physical data. Another fact is different hardware components of a sensor node consumes different rate of energy. As a property of sleep scheduling algorithm, whenever a sensor node is in inactive stage, it will preserve the energy consumption of the different hardware components in a sensor node.

The result evaluation of the study is classified into multiple unit stratum groups or clusters. It is already know that networks deployed using LEACH principles

Table 1 Simulation parameters

Total number of nodes	100
Size of network	100 × 100 m^2
Position of BS	(60,150) m
Preliminary energy initialization	0.1 J
Feasibility of CL election	0.5
Size of packet	6500 bits

have uniform cluster nodes to total number of nodes in many cases. Therefore, the simulation is being studied considering two parameters mainly i.e. energy consumption and total time of data aggregation. Both the above performance evaluation parameters are also evaluated with respect to three different values of compression ratio i.e. when compression ratio is 0.5, 0.75, and 1 respectively. Figure 2 represents the total energy consumption in joules when compression ratio is initialized to 0.5 and the simulation is done considering 100 nodes. The outcome highlights that energy consumption of proposed system is less as compared to LEACH protocol. In the next phase as exhibited in Figs. 3 and 4 when the compression ratio is increased to 0.75–1, s similar trend is observed which shows that proposed system has better energy conservation as compared to standard LEACH protocol with increased compression ratio. One of the basic reasons behind this is

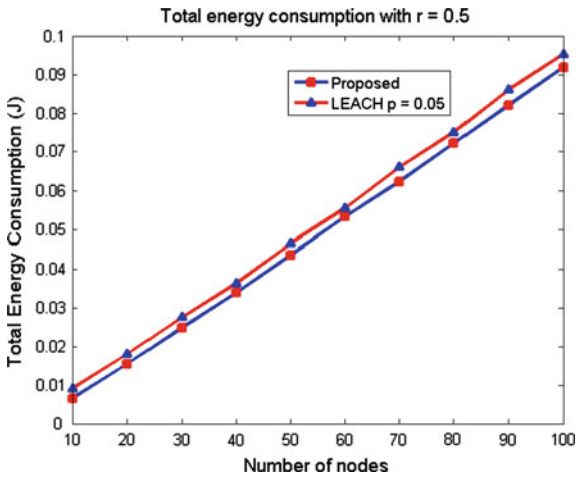


Fig. 2 Total energy consumption at CR = 0.5

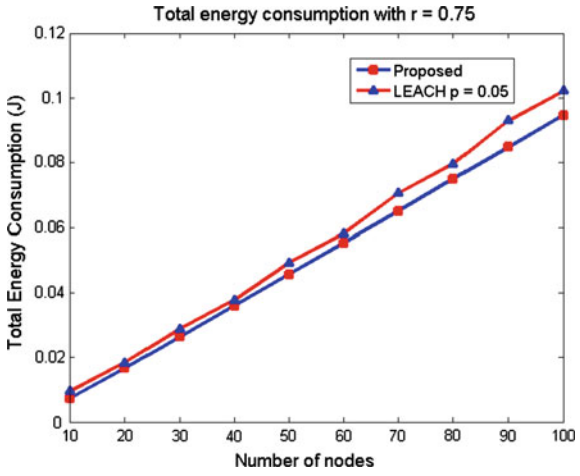


Fig. 3 Total energy consumption at CR = 0.75

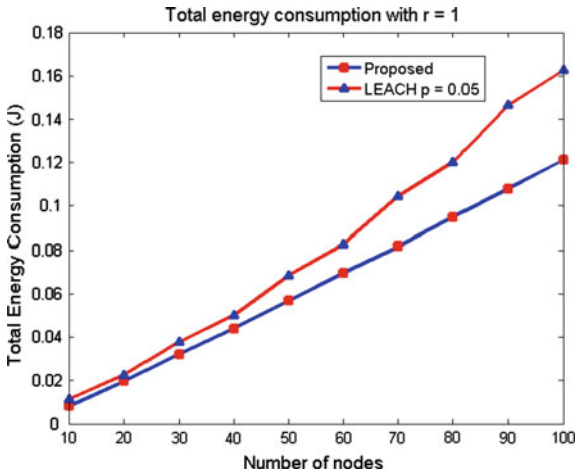


Fig. 4 Total energy consumption at CR = 1

the proposed tree based approach which don't allow the aggregator node to send the data directly to the base station, but creates a better hoping strategy thereby mitigating the enough traffic load using the new module of routing agent. Hence, if in the duration of data aggregation, if the aggregator node is about to die off, the routing agent performs selection of new aggregator node even before the old aggregator a node dies off. In this manner, the cumulative network's energy is highly preserved by dynamically reconfiguring the new routing agents when the previous routing agent is about to die off. The system therefore ensures better energy optimization along with better balancing traffic load using the novel schema of routing agents introduced in the current study.

The performance evaluation of the current study is cumulative data aggregation time. In this evaluation phase, each simulation is studied by varying the value of compression ratio to 0.5, 0.75, and 1. Figure 5 shows that cumulative data aggregation duration for proposed system is higher compared to LEACH at compression ratio 0.5. This increased dimension of the duration for the proposed system is basically due to the initial routing stage that attempts to perform preliminary communication within various clusters after performing data aggregation. The prime reason here is although the system might experience a bit high delay in the preliminary rounds, but the delays will be highly minimized in the consecutive cycle of data aggregation. However, when the compression ratio is increased to 0.75 from 0.5, the system is witnessed with increasing linearity in curves with increasing number of sensor nodes. Hence, the data aggregation time for proposed system is now found to be minimized as compared to preliminary outcome when compression ratio was 0.5.

A closer look into the curves in Fig. 6 and 7 shows that data aggregation duration for both LEACH as well as proposed system exponentially differs with respect to performance as the current study provides a better routing schema that can direct its packets, minimizing its redundancies, and routing them in proper

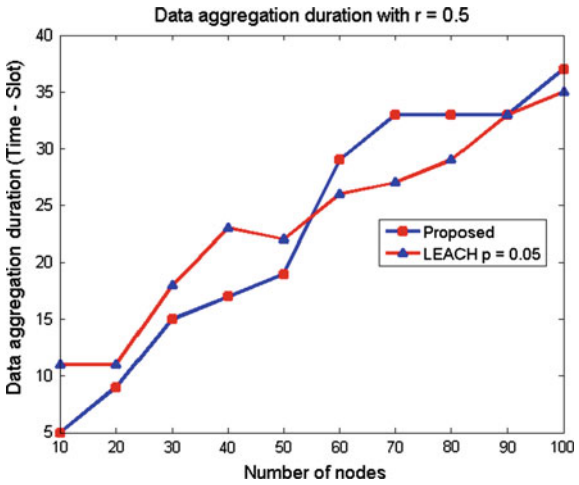


Fig. 5 Total data aggregation time at CR = 0.5

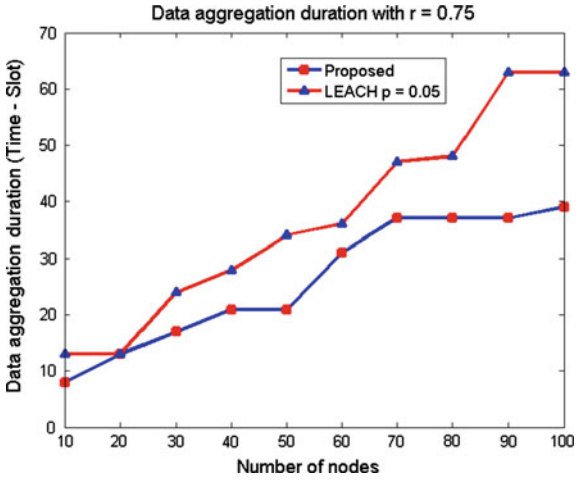


Fig. 6 Total data aggregation time at CR = 0.75

direction (either base station or next hop) using routing agents. Moreover, due to higher level of sustenance (owing to better energy conservation), the algorithm can perform thrice times better than LEACH in minimizing the duration of data aggregation. As the system is designed based on the radio and energy model using multiple layer grouping of nodes, the accomplished outcome is therefore easy to measure and hence scalable. With the increasing number of nodes, the number of cluster will possibly increases, however, it is still not found with any significant effect in the data aggregation duration. The outcome shows that data aggregation duration of the proposed system is exponentially minimized to larger extent as compared to the standard LEACH protocol.

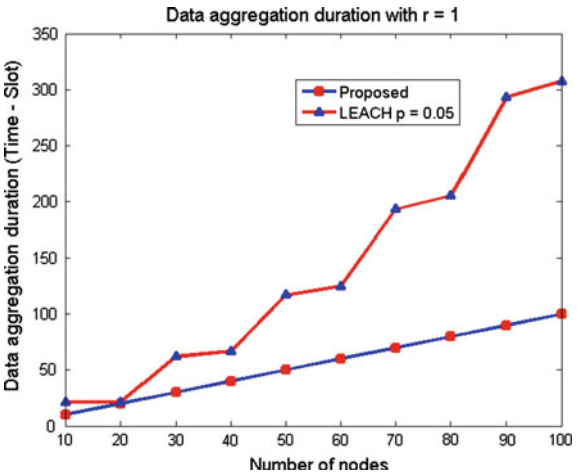


Fig. 7 Total data aggregation time at CR = 1

6 Conclusion

The current paper has discussed about a novel energy aware routing schema that introduces a module called as routing agent for the major purpose of accomplishing energy efficiency. However, along with energy optimization, the system is also found to posse's better load balancing scheme due to its unique routing tactics. The prime responsibility of the new module is to collect the aggregated data and perform energy efficiencies task by taking the decision of forwarding the aggregated data to the base station or to the new hop. The outcome accomplished from the study is evaluated with respect to standard LEACH protocol, where it was seen that proposed network has complete utilization of energy optimization leading to minimizing in data duration time.

References

1. Chong, C.Y., Kumar, S.P.: Sensor Networks: Evolution, opportunities, and Challenges. Proc. IEEE **91**(8), 1247–1256 (2003)
2. Li, Y., Thai, M.T., Wu, W.: Wireless Sensor Networks And Applications, Springer, Computers, p. 441 (2008)
3. Zou, Y., Chakrabarty, K.: Energy-aware target localization in wireless sensor networks, pervasive computing and communications. In: Proceedings of the 1st IEEE International Conference, pp. 60–67 (2003)
4. Sharma, A., Shinghal, K., Srivastava, N., Singh, R.: Energy management for wireless sensor network nodes. Int. J. Adv. Eng. Technol. **1** (2011)
5. Cardei, M., Thai, M.T., Yingshu, L., Weili, W.: Energy-efficient target coverage in wireless sensor networks, 24th annual joint conference of the ieee computer and communications societies. INFOCOM **3**, 1976–1984 (2005)

6. Jang, K.Y., Kim, K.T., Youn, H.Y., Suwon, S.U.: An Energy Efficient Routing Scheme for Wireless Sensor Networks, Comput. Sci. Appl. pp. 399–404 (2007)
7. Katiyar, V., Chand, N., Soni, S.: Clustering algorithms for heterogeneous wireless sensor network: a survey, Int. J. Appl. Eng. Res. **1**(2) (2010)
8. Akyildiz, F.: A survey on sensor networks. Comput. J. IEEE Commun. Mag. **40**(8), 102–114 (2002)
9. Chang, H., Leandros, S., Tassiulas, L.: Maximum lifetime routing in wireless sensor networks. Comput. J. IEEE/ACM Trans. Netw. **12**(4), 609–619 (2004)
10. Haggerty, S.D., Lanzisera, S., Taneja, J., Brown, R., Culler, D.: Scale: insights from a Large, Long-Lived Appliance Energy WSN. In: Proceedings of the ACM (2012)
11. Gu, Y., He, L., Zhu, T., He, T.: Achieving energy-synchronized communication in energy-harvesting wireless sensor networks. ACM Trans. Embed. Comput. Syst. (TECS), **13**(2) (2014)
12. Alwadi, M., Chetty, G.: Feature selection and energy management for wireless sensor networks. IJCSNS Int. J. Comput. 46 Sci. Netw. Sec. **12**(6) (2012)
13. Nikravan, M., Jameii, S.M., Kashani, M.H.: An Intelligent Energy Efficient QoS—Routing Scheme forWSN. (IJAEST) Int. J. Adv. Eng. Sci. Technol. **8**(1) (2011)
14. Lakshmi, A., Manisekaran, S.V.: Fuzzified dynamic power control algorithm for wireless sensor networks. Int. J. Eng. Sci. Technol. (IJEST) **3**(4), ISSN: 0975-5462 (2011)
15. Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: IEEE Proceedings of the Hawaii International Conference on System Sciences (2000)

Load Based Migration Based on Virtualization Using Genetic Algorithm

S. Sandhya, N. Usha and N.K. Cauvery

Abstract Load balancing is the process of distributing tasks among different nodes in a network. The nodes may be either on different machines or nodes (virtual machines) on the same machine. Based on the availability of the nodes, processes from fully loaded node can be migrated from one node to another having less load known as process migration. Process migration can also be done in the virtualization environment by using a hypervisor called Xen hypervisor. This hypervisor safely multiplexes the hardware resources of the physical machine leading the resource allocation in the Virtual Machine (VM) to improve the utilization and performance. To optimize the task of balancing load among nodes, Genetic algorithm (GA) may be used for selecting the destination. GA is a search algorithm, based on natural genetics and principle of evolution and is been widely used in optimization with binary and continuous variables. With the adaptive crossover operation of GA and the searching heuristic and fitness function, lots of possible solutions are searched and the best one is selected as the destination for migration. GA is been proved to be effective in discovering the global optimum even in a very complex searching space.

Keywords Process migration · Load balancing · Xen hypervisor · GA

1 Introduction

Process migration is an approach of transferring the load between nodes. Generally the process has an address space, the required data with the program code in addition to the state specific to the OS. The process state comprises of the infor-

S. Sandhya (✉)
Department of CSE, RVCE, Bangalore, India

N. Usha
Department of ISE, SJBIT, Bangalore, India

N.K. Cauvery
Department of ISE, RVCE, Bangalore, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_29

mation with respect to process execution, files, related task, signal handlers, accounting information and more [1]. The state of the process, address space and registers are sent to the destination machine either prior migration or on-demand by the destination as specified by the migration policy. Since resources on the source will be released immediately upon the transfer of the task, the load is brought down on the source. Load balancing is carried out in different methods by using several algorithms including static and dynamic approaches which collaborated with Genetic Algorithm performs efficiently [2, 3]. Xen [4, 5] is an open-source virtual machine which supports virtualization, creating virtual machines that can operate on different host OS (can be LINUX flavours). Xen is a virtual machine monitor that allocates resources manges efficient migration of process results in low-overhead [4], providing system isolation with efficient execution. Different applications like a text file, audio file, video file, and performance-critical HPC kernels [4] applications can be given to different VM's to perform load balancing.

2 Virtualization Using Xen Hypervisor by Considering HPC Applications

Xen [5] generally supports two types of domains Dom0 and DomU. The systematic management services are controlled by Dom0 for the guest domains. DomU has the access to the virtual devices but restricted to the actual physical devices. The services to the domains is done through the hypercalls. A resource management system model based on Xen virtual machine [5] will ensure domain's resource usage with high priority. The main goal is to maximize the performance of VM by allocating resources for VM based on priorities and resource management system which has low overhead. Resource management system model consists of two modules: (a) System monitoring module- which access the VM run time information's like resource utilization and so on. (b) Resource scheduling module—consideration of resource scheduling policy based on different priorities of the domain request. The priority of applications are considered in three different levels like core application level, service application level, common application level. For testing Domains with different priorities, a benchmark is used in both highest and lowest priority Domains and attempts to utilize guest domain resources. This also monitors guest domain and analyze runtime information for automating resource allocation. But still the performance using these two models is poor comparatively which is a challenging issue. Both communication, computation [6] are compared to evaluate the impact of Xen. Generally Xen supports virtualization techniques like Full-virtualization [7] which simulates the base architecture and Para-virtualization [7] which abstracts the base architecture. To compare the performance of paravirtualized (which has improved scalability and performance) kernel against three popular OS like Red Hat Enterprise Linux (RHEL4) and its versions (v2.6.9 and v2.6.12) a method called statistic sound methods [6] are used. The specific

hardware platforms is considered with Extended Memory 64 Technology machines, each node consists of four Intel Xeon 3.40 GHZ processors with 4 GB RAM. The experiments are performed by repeatedly executing wide range of standard benchmarks and performance data collection. Here micro-benchmark like serial Fortran implementation, which includes programs for evaluating computational performance and macro-benchmarks, real HPC applications are used to evaluate overall performance of the paravirtualized system. Therefore results indicates that there is no statistical significant overhead by Xen and also system maintenance, management and customization are flexible. And Xen is a practical and very efficient hypervisor for HPC systems. The specific mismatch comes generally at the cost, performance degradation, assumptions and algorithms provided by legacy OS. The analysis of Xen driver split mechanism [8] inspires the possibility of building a communication channel between guests and host OS. The very basic legacy design considered in this paper has disadvantages like read/write access to all, incomplete shared data, emergency termination of shared behavior unnecessarily. To improve this, a framework with different modules is designed and methods like Communication method [8] is used to build communication channel and data sharing control method is to signify virtual disk controller module on each operations for data sharing and access. The method Data sharing which is based on virtual disk between guest OS's of full-virtualization is implemented using cross-level data sharing method. Another method, Virtual disk access control includes access to the disks for data via full-virtualization. By series of experiments, it promises the efficiency beyond other sharing methods and makes data access more controllable and much safer. The results proved that data sharing & data transferring is safer than traditional data sharing with good speed of access. Yet the complete efficiency and performance is the major concern for the researchers.

Virtualization [6, 9] is a technology which provides abstraction between the hardware and the applications using VM. It is very attractive which runs hypervisor on the host OS to benefit the HPC applications [10] in the dimensions like OS variety productivity, performance, simplicity along with other parameters. An HPC application and its execution environment [4] can be embedded within a guest domain along with domains of conventional OS which is used for the purpose of administration only. Hypervisors or virtual machine monitor are much more stable compared to usual OS. The benefits of this hypervisors provide reliability and availability for the HPC environment. The parameters like security, checkpoint/restart, preemption, introspection, portability of virtualization issues and trends in the hardware virtualization emphasis values in HPC environments. The execution environment takes the advantage of kernel-level facilities to improve performance which is hard to achieve in the traditional process due to excessive overhead & lack of support. In contrast to the UNIX family, hypervisors provide dedicated environment, and load balancing can be facilitated on top of the host OS. The concept of mini-OS [4] is used, where the mini-OS shown in Fig. 1 [4] completely relies on the hypervisor to access Xen network, block, console, frontend and backend. On top of mini-OS kernel, the newlib C library & the IP stack, any application can be embedded. The mini-OS performs scheduling, memory

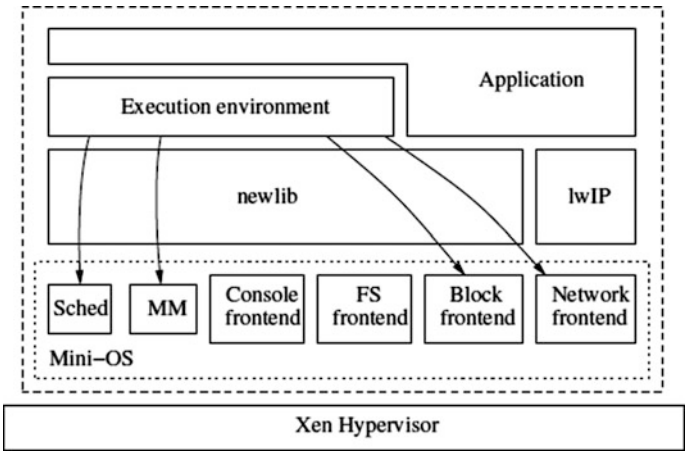


Fig. 1 Light weight guest domain

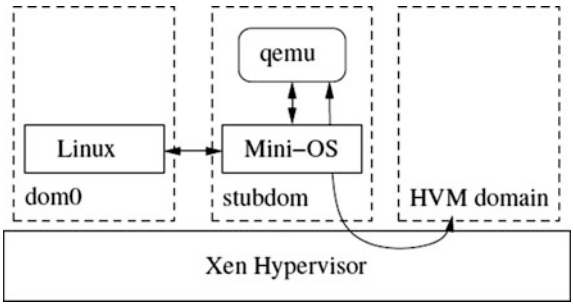


Fig. 2 Qemu based stub domain

management, Disk Input/Output operations for better result. The performance can be achieved by using qemu based stub domain with mini-OS as shown in Fig. 2 [4] and also with reduced overhead.

Experimental results proves that HPC applications are poor on general OS compared to Hypervisors for the reasons like load balancing and the deployment.

3 Process Migration

Process migration [11] is a procedure of transferring task or process from one machine to other. This enables dynamic load distribution, fault resilience, data access locality and improved system administration and also offer check-point approaches. In process migration two instances called source instance and destination instance are considered as shown in Fig. 3 [11]. Process migration is

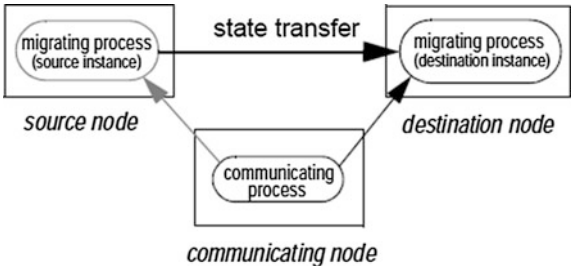


Fig. 3 Overview of process migration

introduced based on multiple processes in distributed systems to improve the performance, processing power and communication channels. The architectures like Non-Uniform Memory Access (NUMA) has different access time depending on memory and the distance of the processor. The advantage of Massively Parallel Processors (MPP) architecture is to repartition of the processors by multiple users. Migration of process could be done using different services of the operating system like IPC (Inter Process Communication), kernel-level migrations or individual generic applications. Process Migration can use different distributed policies like sender-initiated, receiver-initiated, symmetric or random approach. Some of the issues faced while migrating processes are stability, adaptability, complexity, scalability, heterogeneity, transparency, fault resilience, and performance.

There are different algorithms [1] used for process migration namely: (a) Total copy- which transfers the complete state of process before process resumes at the destination. (b) Pre-copy-transfer of process address space in parallel with the execution of process. (c) Lazy copy-which transfers the states for process execution based on the demand made at the destination. (d) Flushing—which keeps the global backup of the states that can be accessed by both source and destination. A new migration algorithm called generic process migration algorithm [11] is described, which is a combination of pre-copy, post-copy and flushing algorithm which consists of three cycles (i) Pre-migration- transfers the address space before migration to the destination. (ii) Migration- transfers the process state from source to destination. and (iii) Post-migration- transfers the address space after migration of a process. This generic algorithm presents an arbitrary behavior in different situations, and overlaps with other algorithms, like Pre-copy, Post-copy and Total-copy. Therefore this generic approach benefits from significant features of the existing algorithms providing better performance. A video process migration [12] is proposed by using sender-initiated dynamic process migration which uses post-copy mechanism with Common Internet File System (CIFS) protocol for balancing the load. The different phases consideration here are (a) Activation phase-to determine time of migration, (b) Selection phase—to select the victim process to migrate, (c) Location phase-to select the destination node, and (d) Migration phase-to migrates the process. Based on the calculation and estimation of load, CIFS protocol used to migrate the video process from source to destination accomplishes

less downtime and less delay and improves the performance. The parameter considered is frames per second (fps) which has improved thus plays the video smoothly on the migrated machine.

4 Load Balancing and Genetic Algorithm

Load balancing is the process of distributing the computations among different computing machines for fair allocation of load on to the nodes. Load balancing [9, 13] can be static or dynamic. Static load balancing [13] includes some potential algorithms like Round Robin, recursive bisection, and more which balances load based on some predetermined data. Whereas dynamic [13] is based on runtime variation of load, including master-slave and fully distributed scheduling which improves the performance. Other load balancing schemes include local and global scheduling methods, optimal and sub-optimal methods, approximate and heuristic scheduling, distributed and centralized scheduling. Each load balancing algorithm has some challenges based on limitations of a given scheme which helps in finding new algorithm to overcome the challenges. The challenges with respect to static load balancing are accuracy, communication delay and cost. Similarly the challenges with respect to dynamic load balancing are rebalancing of each processor load and the communication cost after rebalance. It can be observed that static load balancing algorithms simpler to implement considering parameters like process migration time, overhead, scalability, availability. But generally the dynamic load balancing approach is better. Now a days, streaming applications [14] are on demand which causes traffic in the network load due to user's expectation in the quality of audio or video. To overcome the imbalance of the load on client and server, a generalized framework [14] is proposed in which duplication of the recent accessed files are buffered on different systems and also supports for live streaming. The files are compressed and uploaded on client and downloaded from the server respectively, so that the delay is less. The repeated process of handling video file and migrating if the load is more on the current node increases the performance.

Genetic algorithm (GA) [3, 9] which is used to balance the load among different nodes. It is a search algorithm to find the destination machine to transmit on which the process is resumed. GA begins with the process of chromosome encoding, population initialization, fitness evaluation, selection and use of genetic operator. GA is used to select the correct destination node by using the selection function, which selects the node based on the estimation of load and other parameters of load indicators. GA consists of fitness function to evaluate the fitness of individual solution obtained for each of the iteration. Cross-over generates a new generation by combining individual population traits randomly for better fitness. The cross-over operation of GA may be implemented as single-point cross over or multiple-point cross-over. Mutation operator are used to swap the obtained result of cross over in algorithm depending on the calculated probability.

5 Analysis

Process migration provides [11] the effective solution to migrate from one node to the other. Among different algorithms [1] it's been proved that combination of different migration algorithms results in good performance. During process migration different migration stages are also considered to achieve optimization. Multimedia files [12] can also be migrated from different nodes dynamically using the same above defined approaches based on CPU usage.

Xen paravirtualization [6] provides excellent system maintenance, management, customization for HPC applications [10]. But still the investigation of different techniques is going on to achieve high performance, migration of full systems to balance the load, system isolation error management and also to reduce downtime. Virtualization [6, 9] concept is also introduced based on such hypervisors for fast migration of process's on various OS platforms. Using Xen, VM's behave as if these are running the process using physical resources, thus memory management, resource management are done efficiently and effectively.

Load balancing [13] is the process where load distributed which either uses static or dynamic improves the performance and resource utilization. Streaming applications [14] also improve the user experience and response time by employing load balancing on the server and client end.

6 Conclusion

This paper explains the different approaches of process migration and load balancing that gives the analysis of the efficiency of process migration. Process migration performs better with pre-emptive algorithms. Load balancing needs to consider the applications running on the node to choose the appropriate technique to distribute load. All these considerations can improve the performance comparatively and gives better optimization with low-overhead.

References

1. Zarrabi, A.: Generic process migration algorithm. Int. J. Distrib. Parallel Syst. (IJDPS) **3**(5), September (2012)
2. Li,, Z., Dong, Y.-M., Huang, C.-Y.: A study of link load balancing based on improved genetic algorithm, 6th International Symposium on Computational Intelligence and Design (2013)
3. A survey of parallel genetic algorithms by Erick Cantú-Paz, Department of Computer Science and Illinois Genetic Algorithms Laboratory, University of Illinois, Urbana-Champaign
4. Thibault, S., Deegan, T.: Improving performance by embedding hpc applications in lightweight xen domains. In: 2nd International Workshop on Virtualization Technology in Distributed Computing (VTDC), IEEE (2006)

5. Ning, G., Sun, Y.: Research and implementation of resource management system based on xen virtual machine. In: International Conference on Computer Science and Network Technology (2011)
6. Youseff, L., Wolski, R., Gorda, B., Krintz, C.: Evaluating the performance impact of Xen on MPI and process execution for HPC systems. 2nd International Workshop on Virtualization Technology in Distributed Computing (VTDC), IEEE (2006)
7. Xen Hypervisor Deployment, Management, and Cloud Computing Tools, Todd Deshane and Patrick F. Wilbur Clarkson University
8. Liu, F., Zhou, W., Zhou, M.: A Xen-based data sharing & access controlling method. 3rd International Symposium on Intelligent Information Technology Application (2009)
9. Ranjitha, K., Sandhya, S., Cauvery, N.K.: A survey on: pre-emptive migration of a video process using genetic algorithm on virtual machine. Int. J. Eng. Comput. Sci. ISSN:2319-7242 **3**(5), pp. 5897-5900 may (2014)
10. Mergen, M.F., Uhlig, V., Krieger, O., Xenidis, J.: Virtualization for High-Performance Computing, IBM T. J. Watson Research Center, Yorktown Heights
11. Milojicic, D.S., Douglis, F., Paindaveine, Y., Wheeler, R., Zhou, S.: Process Migration, HP Labs, AT&T Labs–Research, TOG Research Institute, EMC, University of Toronto and Platform Computing, August 10 (1999)
12. Sandhya, S., Cauvery, N.K.: An improved and optimized approach for pre-emptive migration of video process, 978-1-4799-6629-5/14/$31.00 IEEE (2014)
13. Mcheick, H., Mohammed, Z.R., Lakiss, A.: Evaluation of load balance algorithms. 9th International Conference on Software Engineering Research, Management and Applications (2011)
14. Sandhya S., Cauvery, N.K.: Dynamic load balancing for video processing system in cloud I.K. Sethi (ed.), Computational Vision and Robotics, Advances in Intelligent Systems and Computing 332, DOI 10.1007/978-81-322-2196-8_22, Springer India 2015

An Architecture for Detection of Land Mines Using Swarm Robots

Rajesh Doriya, Krishna Mohan Srivastava and Priyanka Buwe

Abstract Coordination and communication of multiple robots in an unknown environment is a great challenge to deal with. In this paper an approach for this particular problem is proposed, in which efficient hybrid architecture is created combining two different approaches for communication and coordination each. Multiple robots are treated as ad hoc mobile nodes and the phenomenon of swarm intelligence is used for them to coordinate, while they use Cluster Head Gateway Switch Routing (CGSR) protocol for communication. Various experiments such as goal reaching and collision avoidance are performed using this architecture. Using this particular framework an application of "detection of land mines in an uneven terrain" is worked upon. The swarm robots communicate and coordinate with each other and the environment to detect the land mines using the proposed architecture, and store its position in a table. This application thus has its focus on saving human lives while exploring the environment consisting of land mines, which has already been explored by robots and thus it becomes safe for them. The hybrid of two algorithms results in better performance of robots in an uneven terrain.

Keywords Multi-robot communication · Multi-robot coordination · Swarm robotics · Particle swarm optimization · Cluster head gateway switch · Routing protocol

R. Doriya (✉) · K.M. Srivastava · P. Buwe
Department of Information Technology, National Institute of Technology, Raipur 492010, India
e-mail: rajeshdoriya.it@nitrr.ac.in

K.M. Srivastava
e-mail: krishna.sri23@hotmail.com

P. Buwe
e-mail: priyankabuwe53@gmail.com

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_30

1 Introduction

Now a days, robots have been widely used in many diverse areas such as homes, medical, logistics (such as AGV), space, military fields, nuclear plants etc. The list of such robots in various applications is shown in Fig. 1. The application outfitted in the paper is "detection of land mines in an uneven terrain" as stated above. Motivation towards choosing this application lies in the fact that robots are widely used in defence field. It deals with the detection and storing of the coordinate position of land mines in the war field, using multiple swarm robots with simple functionalities. The swarm robot used here is well equipped and is able to perform the task efficiently using the functionality of its own. For instance collision avoidance, target reaching as well as detection of mines and storing its position in an array. The reason behind employing multiple simpler robots to achieve a task is both efficient and cheaper than a single powerful robot. The degree of fault tolerance is yet another advantage incurred from multiple robots. Instead of deploying a single powerful robot, multiple simple robots is comparatively nominal.

In the recent times, coordination among multi-robots has been a major area of study for researchers of the field, amongst which Swarm robotics has been the most studied approach for coordination among robots. It draws inspiration from the behaviour of social insects or species with the most remarkable feature of robots in a swarm is the accomplishment of a common task through cooperation. Authors have proposed two strategies for coordination in robotic swarm: centralized and distributed. In the centralized strategy, a single robot acts as the coordinator that monitors coordination among robots in its team while in the distributed (decentralized) approach, each robot coordinates its own movement without the supervision of some single robot. Each strategy has its own pros and cons. As far as communication is concerned, most of the communication strategies proposed or developed are based on ad hoc networks; those rely on infrastructure-less architectures. The routing mechanisms are difficult to implement because robots change

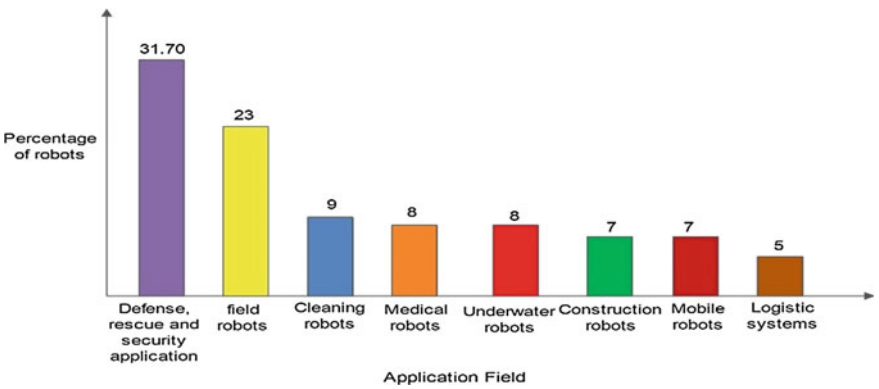


Fig. 1 Use of robots in various application

their positions frequently, thus altering the topology of the network. For robot coordination, robots need to be equipped with communication abilities. Hence, we deployed a hybrid architecture combining the virtue of both the swarm intelligence and communication abilities.

2 Related Work

Several approaches have been proposed for communication among robots. In 1973, a research was initiated on Packet Radio Networking (PRNET) [1]. PRNET provides interconnection between geographically separated users. A table-driven approach 'Destination Sequenced Distance Vector Routing' (DSDV) [2] was introduced where nodes are modeled as routers. Each node maintains a routing table to record available routes and the hop count to reach the destination. A sequence number is assigned by the destination node, which enables to differentiate between stale and new routes.

In order to reduce the overhead of routing messages, Cluster Head Gateway Switch [3] algorithm was introduced which aggregates nodes into clusters. Chen et al. introduced Global State Routing Protocol (GSR) [4], based on Link State Routing (LSR). It provides global knowledge of the entire network topology and avoids flooding of routing messages. Each node maintains a list of neighboring nodes, topology table, next hop table and distance table. Upon reception of new routing messages, a node updates its topology table and reconstructs the routing table and forwards the routing information to its neighboring nodes. Iwata et al. [5] proposed Fisheye State Routing (FSR) in which update message contains information about closer nodes rather than all the nodes. Thus it reduces the size of update messages and saves considerable amount of network bandwidth. He also introduced Hierarchical State Routing (HSR) based on multilevel hierarchical clustering with location management. In this approach nodes are organized into clusters each having its own cluster head. These clusters are then organized into hierarchical levels. The cluster head gathers information from nodes within its cluster and forwards it to the upper level cluster head. Park et al. [6] discusses Temporally Ordered Routing Algorithm for highly dynamic multi hop networks. Control messages are localized only to less number of nodes. The protocol proceeds in three phases: route creation, route maintenance and route erasure.

Johnson et al. [7] proposed an on-demand routing protocol Dynamic Source Routing (DSR). Each node maintains a route cache with the source routes it is aware of. Cache gets updated when new routes are discovered. In later research a significant algorithm called as Ad Hoc On-demand Distance Vector Routing (AODV) [8] was introduced. In this algorithm, routes are created only when needed, thus minimizing the number of broadcasts. A route request packet is broadcasted by the source node to find a path to the destination node. This packet is passed from a node to its neighbor till it reaches the destination. Toh [9] highlighted in their research, associativity based routing which selects routes based on

Table 1 Comparison of various routing algorithms

Routing algorithms	Time complexity	Computation complexity	Routing	Loop free	Multicast capability
DSDV	O (d), d	O (x = N)	Flat	Yes	No
CGSR	O (d), d	O (x = N)	Hierarchical	Yes	No
WRP	O (h), h	O (x = N)	Flat	Yes	No
AODV	O (2d)	O (2 N)	Flat	Yes	Yes
DSR	O (2d)	O (2 N)	Flat	Yes	No
TORA	O (2d)	O (2 N)	Flat	Yes	No
ABR	O (d + z)	O (2 N + y)	Flat	Yes	No
SSR	O (d + z)	O (2 N + y)	Flat	Yes	No

associativity state. Upon receiving a beacon, a node updates its associativity table. It aims to find longer lived routes for ad hoc networks. Table 1 demonstrates the analysis of various routing algorithms with time complexity and communication complexity as their prime metrics.

In order to achieve any task in an efficient manner, coordination among robots is of utmost importance. Aspects of coordination among robots have been studied extensively in the past. Kennedy and Eberhart [10] proposed Particle Swarm Optimization (PSO) to achieve co-ordination among robots discussed in next section of the paper. Estrin et al. [11] discussed localized algorithms for coordination in sensor networks. Hu et al. [12] presents coordination among robots through communication. The paper focuses on co-operative behavior of multiple mobile robots. A model for shared experience learning among robots has been presented. Chen et al. [13] introduces Span- an energy efficient algorithm for coordination among multiple nodes. Each node may turn its receiver off if it does not wish to participate in the network topology. A node may become a coordinator if it notices that two of its neighbours cannot directly communicate with each other. Asama et al. [14] presents collision avoidance algorithm for robots based on rules or negotiation. Collisions among robots are avoided based on rules. If no rules satisfy, the robots negotiate by letting one go. Yamauchi [15] proposed extending frontier-based exploration to multiple mobile robots. Mobile robots are directed to areas between unexplored and explored regions. When a robot moves to a frontier cell, the information of the new region is added to its map. Depth first search was used for navigation. Burgard et al. [16] proposed collaborative multi-robot exploration which assigns different locations to individual robots so as to explore entire space in collaborative fashion. Upon choosing of a target point by a robot, the utility value associated with it gets reduced.

A research [17] provides an extensive research of swarm robotics. Inspiration from the biological world has been briefed along with the types of communication possible in swarm. The two control approaches- centralized and distributed have been discussed and compared. Another paper [18] provides a detailed review of swarm robotics and swarm engineering. The paper states the design and analysis

methods to develop collective behaviors in swarm. The common collective behaviors among robots in a swarm have been explained. Turgut et al. [19] presented self-organized flocking in robotic swarms. The paper presents a mobile robot platform—Kobot specially developed for swarm robotics. Bonani et al. [20] introduced MarXbot-a mobile robot designed for collective robotic research. The paper details the architecture of the miniature robot along with the energy management mechanism used. Schwager et al. [21] proposed a control strategy to allow robots to position themselves so as to reduce sensory information. This information then drives the network to a stable configuration. Fong et al. [22] presents a survey of socially interactive robots. Some of the common robot characteristics along with other associated aspects of design and emotions have been briefed.

Franchino et al. [23] present a distributed architecture for robot coordination. Global positioning system (GPS) is simulated to achieve robot localization. Abase station executes a task by sending commands to the robots. The paper also gives a communication infrastructure to allow robots to exchange messages with limited delays. MAC protocol has been used for accessing the radio channel. Coordination in multi-robots can also be achieved by deploying a cloud, which acts as a decision making server [24] for assisting low cost robots. This approach makes it easy to use and reduces the over-head of communication. Robotic components can also be offered as a service in the cloud over web [25]. Our future work also includes the implementation of cloud service on the existing architecture for better performance.

3 Swarm Robotics

The particle swarm paradigm, which was a unique topic a decade ago, is now no more an unexplored one. This approach was inspired from the biological study of ants another insect's demeanour of collective migration and interaction with their surroundings. These insects coordinate using decentralized control and self-organization. Various researchers conceptualized their demeanour into algorithms which were further being optimized with time, such as the Ant Colony Optimization technique (ACO) [26]. This approach has drawn its concept from the ant's behaviour of taking smaller paths while searching food using pheromone updation. Another concept is particle swarm presented by Kennedy and Eberhart which was focused on producing computational intelligence exploiting the collective nature of birds. In this approach the mobile nodes act just like the birds flocking in groups, rather than implementing their individual cognitive abilities. Originally these simulations were influenced by the work of Heppnerand Grenander's [27] which involved a bird flock searching for corn. This idea was soon developed into this very effective optimization method named as the Particle Swarm Optimization (PSO) [10]. In PSO a number of particles migrate in the search space in search of their target and evaluates the search space in terms of their current location. Each particle then determines its navigation path firstly according to the history of its own current and best (fitness) locations, secondly according to

the position of particle located nearest to the target with some random perturbations. Velocity is also updated as shown in the following equation:

$$v_i = v_i + \cup(0, \emptyset_1) \otimes (p_i - x_i) + \cup(0, \emptyset_2) \otimes \left(p_g - x_i\right) \tag{1}$$

$$x_i = x_i + v_i \tag{2}$$

As explained, each iteration is thus updated accordingly with the optimum as the fitness function until the whole flock is moved close to the target (food in case of bird flock). Swarm robotics means the interaction of multi robot system with the environment and among them to coordinate their behavior collectively. Basically, swarm robotics is implementing swarm intelligence into multi robot systems, considering simple physical robots as swarm particles. Thus these simple robots now behave as swarm particles to coordinate and move. In this paper we use this particular optimization method for robots to coordinate.

4 Proposed Architecture for Communication and Coordination of Swarm Robots

In this section, we present an architecture where multi-robots will be coordinating and communicating to perform a task efficiently. Based on the survey report [28] where multiple coordination and communication algorithms are compared on the basis of their performances, time and space complexities we concluded that CGSR and PSO for communication and coordination respectively are better than other approaches. CGSR works based on the formation of cluster of the robots which falls in the range of each other. Cluster is formed with the help of distributed clustering algorithm. Two types of Distributed clustering algorithms can be deployed over here, lowest ID algorithm and highest connectivity (degree) algorithm. We are using the later one where within a cluster the robot with highest connectivity among its neighbors is selected as cluster-head. Communication to other clusters occurs only through this cluster-head robot. Now the challenge here is smearing the coordination algorithm once nodes are communicating. The coordination algorithm includes the application of PSO which is explained above in Sect. 3. Figure 2 roughly depicts the architecture of the proposed framework, three clusters are depicted here. Firstly looking at the cluster level inside these clusters, there is a cluster-head to decide the basic motion strategy and topology of the individual cluster now on the bigger level only then the cluster-heads will coordinate with each other to plan the overall motion scheme of the clusters.

In *algorithm* 1, every robot broadcasts its connectivity, i.e. how many robots in the network it is connected to. Next is an identifier, i.e. a tiebreaker which will decide the coordinator, if connectivity of the two robots (which are the nominee for coordinator) is same. Last is the position of robot in coordinate system. Each robot is initialized with the default values, now for each robot its neighbors list is created.

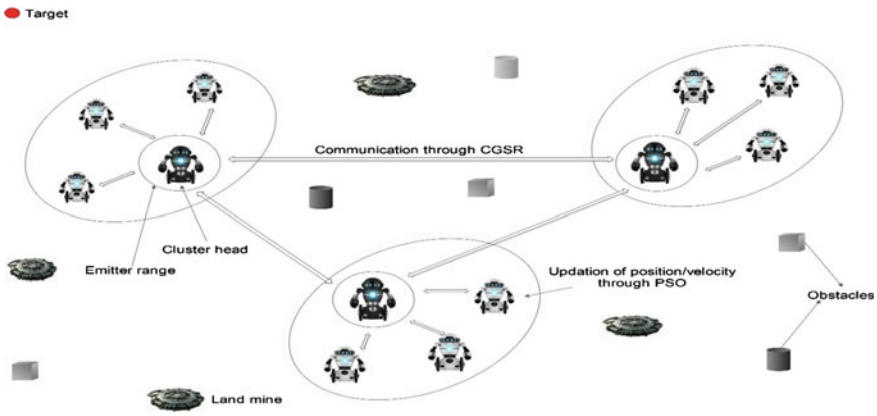


Fig. 2 Hybrid architecture for coordination and communication of robots for land mine

If a robot falls in the neighbour list of the other robot and if the Euclidian distance between them is less than a certain range based on the numbers of the robot in the neighbor list, connectivity to each node is allocated. Thus each robot is traversed one by one.

Algorithm 1: Proposed Algorithm

Every node in the cluster broadcasts its connectivity, identifier and position in the coordinate system after particular amount of time. Every node has a different identifier.
α := connectivity of the identifier
β:= identifier (*tie breaker*)
***Step 1*: Initialize Each node:**
Mark: = false
Is_head: = false
Connectivity: = empty // *to store position of all the neighbours*
***Step 2*: Find neighbours**
if Euclidean_distance < Range
Edge Exists
else
No Edge Exists, set the connectivity
***Step 3*: Select any node** //*formation of cluster*
Mark := true
Compare its connectivity with neighbours
Select the node with highest connectivity
Is_head := true // *Set it cluster head*
***Step 4*: Find neighbours such that**
Euclidean_distance > threshold & Mark := false// *searches unvisited node outside cluster*
Select the node with highest connectivity
Is_head := true // *cluster head for other clusters*
***Step 5*: Repeat step 4 until all the nodes are covered**
Find final set of cluster Nodes
Update the velocity and position according to the equation:

$$v_i = v_i + \cup(0, \emptyset_1) \otimes (p_i - x_i) + \cup(0, \emptyset_2) \otimes (p_g - x_i)$$

$$x_{i=} x_i + v_i$$

5 Land Mines Detection

Ongoing concern of defence unit in any country is to reduce the loss of lives during a war. Using swarm robots in defence instead of human force is the most befitting solution of this concern and need of the hour. Our motivation towards choosing this particular application is also the reason that robots are mostly used in defence sector of our society. This application deals with the detection of mines in an unfamiliar environment where the robots detect a mine placed in the environment. After detecting the mine, they record its position and send its position to all other robots which are in the network, so that they may also avoid those mines. This avoids repetitive detection and position updation of a particular mine. A table for the recorded position is created having the coordinate positions of detected land mines. Different robots are given different target positions to reach and in the course of their navigation whenever they come across a particular land mine, updation process of position starts. Their coordinating approach minimizes the overall effort of robots since they do not have to cover whole area of navigation, which being another advantage of multi robots. They thus can explore wider area simultaneously with detection. Figure 3 illustrates the flow of overall working of the architecture

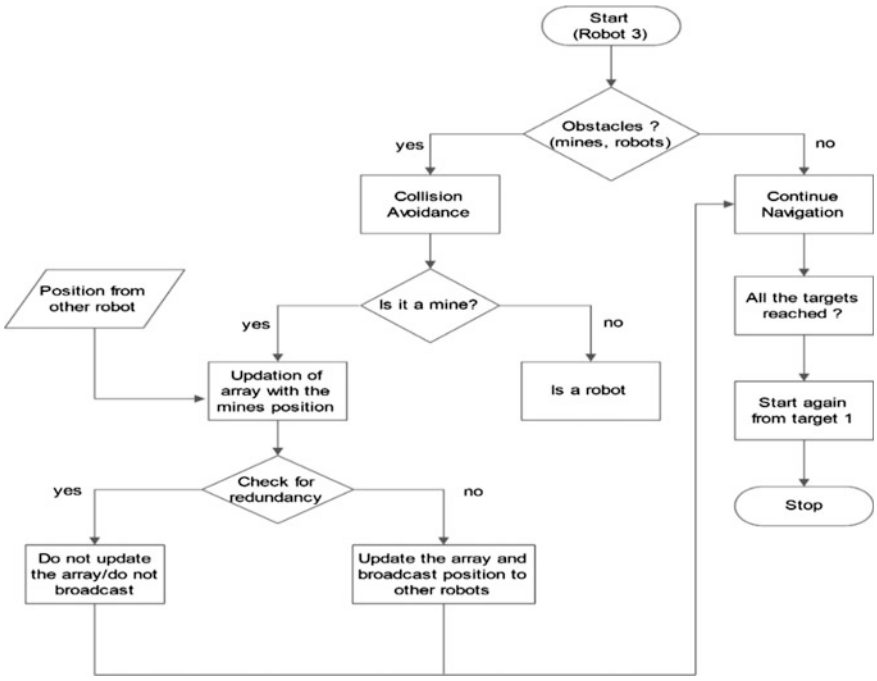


Fig. 3 Flow diagram depicting the working of application

for land mine detection with swarm robots. Functionalities of participating swarm robots and its simulation details are presented in the following subsections.

5.1 Functionalities of Swarm Robots

Robots are equipped with simple functionalities which thus helps in the accomplishment of our simulation. Some of the major ones are listed below:

- **Collision Avoidance**: This task is accomplished by the collision avoidance function, pseudo code of which is depicted in *algorithm* 2.
- **Target Reaching**: Here the GPS and compass values are used to calculate robot's position and the direction of the target respectively so that robot may move to the target by following a proper path coordination via the proposed architecture.
- **Communication**: In the application we build, communication between the robots is used immensely and for this task we have used emitter and receiver for each robot.

5.2 Details of Simulation

Software used for the simulation of our approach is Webots version 6.4, which is a real time fast prototyping software used for simulation of mobile robots through a virtual space. An environment of a square boundary in virtual space consisting of four walls of unit 15 is used for robots to navigate. These are bounding objects such that grid provides an arena for robots to move and avoid the walls of grid. Land Mines is a cubical object which acts as an obstacle for robots, it has a point light source in it which is detected by the light sensors of the robots. Robots we used in this application are simple low cost having many features. Six servo motors are fixed into each robot for its controlled and efficient walk, named as FL (front-left), ML (middle-left), BL (back-left), FR (front-right), MR (middle-right), BR (back-right) servo motors. Figure 4 represents the communication between two robots and an overview of simulation. Setup is also shown in this particular figure where three robots communicate and coordinate through the proposed architecture along with detection of land mines. Communication is shown in the console box.

Communication is done between the clusters using CGSR by making clusters as shown in Fig. 5a and they coordinate via PSO and a robot detecting land mine is shown in Fig. 5b.

Figure 6 shows the simulation in which land mines are detected and coordinate positions are stored in the table. Two infra-red distance sensors in each robot both at

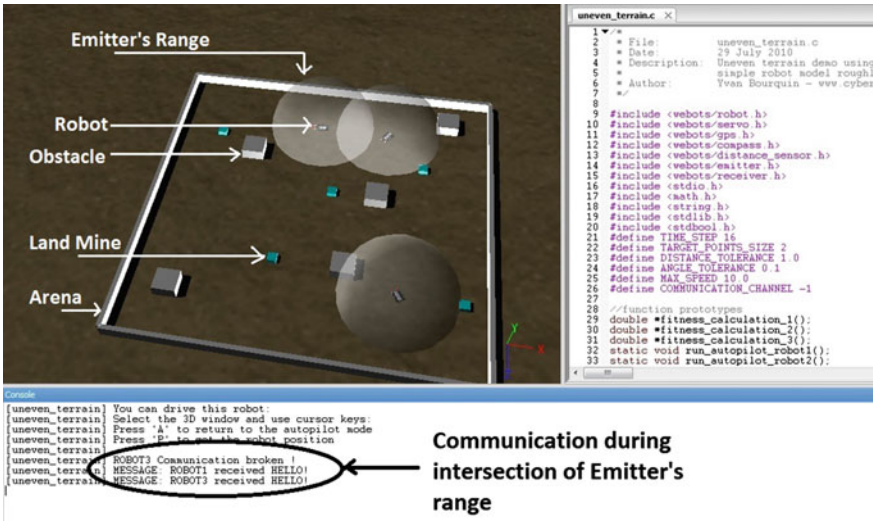


Fig. 4 Simulation results of communication and coordination of swarm robots

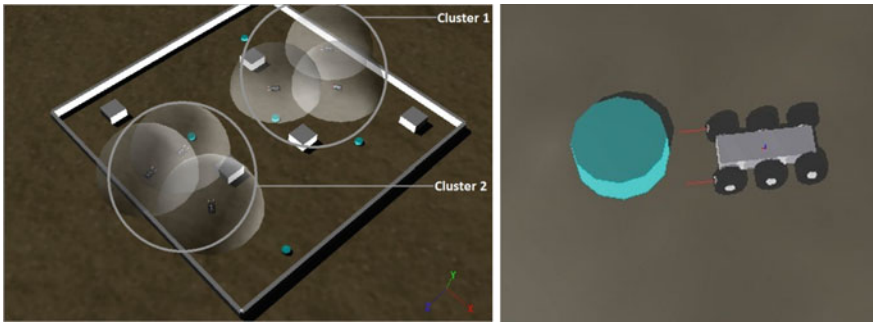


Fig. 5 **a** Cluster formation of swarm robots. **b** A swarm robot detecting a land mine

the left and right side on the head of bots are being used, named as IR0 and IR1. Emitter is used for robots to send out the messages within a particular range at a specific channel, which can be picked by any robot having receiver moving within emitter range. To receive these messages, a receiver in each of our robots is built. GPS and compass are additional devices to send/calculate the current location of the robot in the coordinate system and angle between directions of motion of robot from the virtual north (the virtual north direction assumed during simulation) respectively. The overall process is presented in *algorithm* 2.

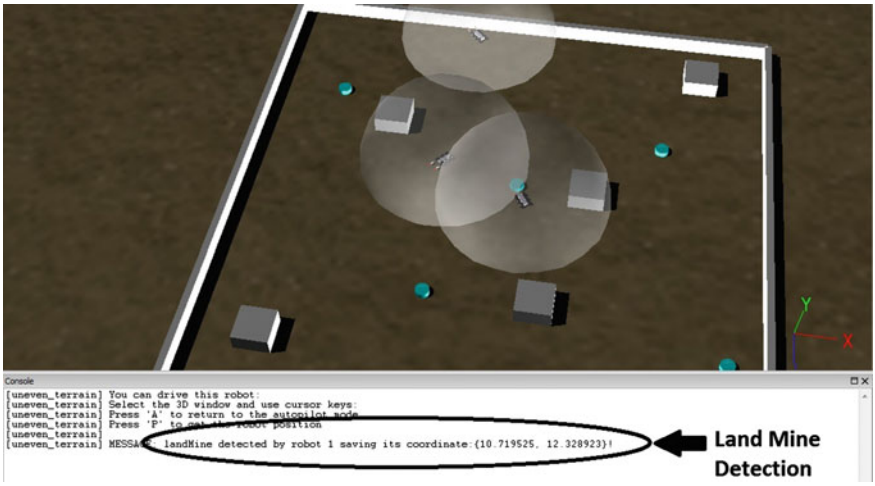


Fig. 6 A swarm robot 1 has detected a land mine and it is broadcasting its coordinate to other swarms robots via cluster head robot

Algorithm 2 : Obstacle avoidance algorithm for mobile robot

```
if IR1_VALUE > 500 then
   if IR0_VALUE > 500 then
      Left Speed = -MAX SPEED
      Right Speed = -MAX SPEED=2// Obstacle detected by both sensors. Move backwards
else
Left Speed = -IR1_VALUE=10
Right Speed = (IR0_VALUE=10) + 5 // turn proportionally to the sensor values
   end if
else if IR0 V ALUE > 500 then
Left Speed = (IR1_VALUE=10) + 5
Left Speed = -IR1_VALUE=10
else
   Left Speed = MAX SPEED
   Right Speed = MAX SPEED // nothing has been detected
end if
```

6 Conclusion

Saving human life in war field has been a prime agenda for all the militaries. The human lives can be saved by deploying service robots into the war field. In our paper we have proposed and implemented an architecture, where robots can be deployed in war field to detect land mines. In the architecture, the concept of swarm robots has been utilized so that larger area with less time can be covered with low

cost robots. For better communication among the robots, a cluster head gateway switch routing protocol is used whereas coordination is achieved with PSO. The architecture is validated with simulation of three swarm robots in an artificially created war field of land mines, where robots are passing the position of detected land mines to other robots so that they can avoid the detected land mines.

References

1. Kahn, R.E.: The organization of computer resources into a packet radio network. IEEE Trans. Commun. **25**, 169–178 (1977)
2. Perkins, C.E., Hagwat, P.B.: Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In: ACM SIGCOMM Computer Communication Review, pp. 234–244 (1994)
3. Chiang, C.-C., Wu, H.-K., Liu, W., Gerla, M.: Routing in clustered multihop, mobile wireless networks with fading channel. In: Proceedings of IEEE SICON, pp. 197–211 (1997)
4. Chen, T.-W., Gerla, M.: Global state routing: a new routing scheme for ad-hoc wireless networks. In: 1998 IEEE International Conference on Communications, 1998. ICC 98. Conference Record. pp. 171–175 (1998)
5. Iwata, A., Chiang, C.-C., Pei, G., Gerla, M., Chen, T.-W.: Scalable routing strategies for ad hoc wireless networks. IEEE J. Sel. Areas Commun. **17**, 1369–1379 (1999)
6. Park, V.D., Corson, M.S.: A highly adaptive distributed routing algorithm for mobile wireless networks. In: Proceedings of the IEEE INFOCOM'97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution, pp. 1405–1413 (1997)
7. Johnson, D., Hu, Y., Maltz, D.: The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4. RFC 4728, February 2007
8. Perkins, C.E., Royer, E.M.: Ad-hoc on-demand distance vector routing. In: Proceedings of the WMCSA'99. Second IEEE Workshop on Mobile Computing Systems and Applications, pp. 90–100 (1999)
9. Toh, C.-K.: Long-lived ad-hoc routing based on the concept of associativity. Internet-Draft, draft-ietf-manet-longlived-adhoc-routing-00. txt, Work in progress (1999)
10. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Encyclopedia of Machine Learning, pp. 760–766. Springer, New York (2010)
11. Estrin, D., Govindan, R., Heidemann, J., Kumar, S.: Next century challenges: scalable coordination in sensor networks. In: Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking, pp. 263–270 (1999)
12. Hu, H., Kelly, I., Keating, D.A., Vinagre, D.: Coordination of multiple mobile robots via communication. In: Photonics East (ISAM, VVDC, IEMB), pp. 94–103 (1999)
13. Chen, B., Jamieson, K., Balakrishnan, H., Morris, R.: Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. Wireless Netw. **8**, 481–494 (2002)
14. Asama, H., Ozaki, K., Itakura, H., Matsumoto, A., Ishida, Y., Endo, I.: Collision avoidance among multiple mobile robots based on rules and communication. In: Proceedings IROS'91. IEEE/RSJ International Workshop on Intelligent Robots and Systems' 91.'Intelligence for Mechanical Systems, pp. 1215–1220 (1991)
15. Yamauchi, B.: Frontier-based exploration using multiple robots. In: Proceedings of the Second International Conference on Autonomous agents, pp. 47–53 (1998)
16. Burgard, W., Moors, M., Fox, D., Simmons, R., Thrun, S.: Collaborative multi-robot exploration. In: Proceedings of the ICRA'00. IEEE International Conference on Robotics and Automation, pp. 476–481 (2000)

17. Yogeswaran, M., Ponnambalam, S.: Swarm Robotics: An Extensive Research Review. Sciyo (2010)
18. Brambilla, M., Ferrante, E., Birattari, M., Dorigo, M.: Swarm robotics: a review from the swarm engineering perspective. Swarm Intell. **7**, 1–41 (2013)
19. Turgut, A.E., Çelikkanat, H., Gökçe, F., Şahin, E.: Self-organized flocking in mobile robot swarms. Swarm Intell. **2**, 97–120 (2008)
20. Bonani, M., Longchamp, V., Magnenat, S., Rétornaz, P., Burnier, D., Roulet, G., et al.: The MarXbot, a miniature mobile robot opening new perspectives for the collective-robotic research. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4187–4193 (2010)
21. Schwager, M., McLurkin, J., Rus, D.: Distributed coverage control with sensory feedback for networked robots. In: Proceedings of the Robotics: Science and Systems (2006)
22. Fong, T., Nourbakhsh, I., Dautenhahn, K.: A survey of socially interactive robots. Robot. Auton. Syst. **42**, 143–166 (2003)
23. Franchino, G., Buttazzo, G., Facchinetti, T.: A distributed architecture for mobile robots coordination. In: 10th IEEE Conference on Emerging Technologies and Factory Automation, 2005. ETFA, pp. 8–156 (2005)
24. Doriya, R., Chakraborty, P., Nandi, G.: Robot-Cloud': a framework to assist heterogeneous low cost robots. In: 2012 International Conference on Communication, Information and Computing Technology (ICCICT), pp. 1–5 (2012)
25. Doriya, R., Chakraborty, P., Nandi, G.: Robotic services in cloud computing paradigm. In: 2012 International Symposium on Cloud and Services Computing (ISCOS), pp. 80–83 (2012)
26. Dorigo, M., Birattari, M.: Ant colony optimization. In: Encyclopedia of Machine Learning, pp. 36–39. Springer (2010)
27. Heppner, F., Grenander, U.: A Stochastic Nonlinear Model for Coordinated Bird Flocks. American Association for the Advancement of Science, Washington (1990)
28. Doriya, R., Mishra, S., Gupta, S.: A brief survey and analysis of multi-robot communication and coordination. In: IEEE International Conference on Computing, Communication and Automation (ICCCA2015), India (2015)

A Study of Challenges and Solutions for Smart Phone Security

Kunjal Gajjar and Aniruddhsinh Parmar

Abstract It is lucid that smart phones have become a must needed device in today's world. With such large pool of people using smart phones ignorantly increases the danger of security extensively. Survey in the paper proves that most of the cyber attacks are possible due to ignorance and negligence of the user. Mistakes made by users like granting permissions to apps without understanding them and without knowing the source, avoiding security features available in the device and storing critical information without encrypting it are major cause of security breach. The paper also provides proactive and reactive solutions to Information security issues.

Keywords Smart phone security · Cyber attacks · Malware · App permissions · Malicious apps · Remote wiping high

1 Introduction

Smart phones have become vital part of Human life. Consumers and organizations use smart phones for personal as well as business use, there by storing important and critical information on it. In today's technology driven society, smart phones provides a great helping hand to customers as well as criminals. Also smart phones provide a very luring services of staying connected to the world, allowing advance

Aniruddhsinh Parmar Author contributed equally in this research work.

K. Gajjar (✉)
CHARUSAT and Faculty, School of Computer Studies AU,
Ahmedabad 380009, India

A. Parmar
C.U. Shah University and Faculty, School of Computer Studies AU,
Ahmedabad 380009, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_31

computing features and impressive web connectivity making it more useable and vulnerable.

Mobile addiction is growing at an alarming rate and majorly teens, college students and middle-aged parents fall victim to this. As of March, 2014 there were 176 million addicts. With such increase in the number of users who mostly fall in the category of youth, attackers have large pool of sources to attack [1]. The survey results demonstrated in the paper has been generated from the responses of smart phone consumers who fall in the age group of 18–30 years.

According to the statistics of global market share held by smart phone operating systems, android is leading the smart phone market in 2014 with 80.7 % thereby increasing the target set to a great size [2]. This huge target is difficult for the attackers to resist. Android accounts for 97 % of all mobile malware which comes from small, unregulated third party app stores predominantly in the Middle East and Asia [3]. Due to this scenario, the paper focuses on the apps and security features of Android platform.

1.1 Types of Attack in Smart Phones

- Unauthorized SMS messages—It is the most common Android malware attack that installs the fake app on the device which secretly sends the expensive messages to the premium rate SMS services [3].
- There are malware infected editions of popular games available on the unofficial or third party Android market. These software use the trick known as Ginger break exploit to gain root access and install malicious code [3].
- Few malwares attack the device with the intention of capturing the messages which contain the information of bank services, transactions and authentication codes [3].
- Potentially Unwanted Applications (PUA) is Android App that may not be considered as malware, but they may introduce risk and security threat. Such apps link to the aggressive advertising networks can track their devices and locations and may even capture contact data [3].
- New attacks use the smartphone cameras, speakers and microphones as the medium of attack. Unlike Bluetooth access, ultrasonic sound does not require permission and can emit sound to anything regardless of a pairing process [4].
- Researchers have demonstrated security threat of stealing the key strokes using the front and back cameras [4].
- Advanced Persistent Threats have increased in number with intention of stealing the sensitive data and generally the attack is aimed at specific individuals to gain access to information.
- Blended Threats, a type of attack which uses combination of email, SMS messaging, instant messaging conversation and mining of social network relationships to send spam, hence using multiple techniques to attack a system.

1.2 Users Unknowingly Helping the Attackers

Unknowingly, by providing the permissions, consumers are helping the attackers to get unauthorized access to the data easily. Following are the permissions which are most commonly used by malicious app and should be granted carefully to any app: Network based location, GPS Location, View Wi-Fi state, Retrieve Running apps, Full alternate access, Read phone state and Identity, Automatically start at boot, Control Vibrator, Prevent from sleeping, Modify/Delete SD card contents, Send SMS Messages, Read contacts stored on the device [5].

2 Related Work

A study done by Ameya Nayak on Android mobile platform security and malware shows that consumers are adopting the technology and are increasing the amount of personal data transmitted and stored on these devices [6]. The survey illustrates rapid rate of malware attacks due to extensive use of smart phones for shopping, banking, e-mail, and other activities that require passwords and payment information. Also storage of important credentials in the device makes cell phones a target for credential theft.

Keng Siau studied challenges to mobile commerce and concluded that due to its inherent characteristics mobile commerce promises businesses unprecedented market potential [7].

According to mobile consumer report of February 2013 Mobile Devices are next power houses and in India more than 80 % people are using mobile phone. Smartphone users are inclined towards the look of device and number of applications it supports rather than its security features which is a serious issue for Information security [8].

Yan Huang pointed out in their study that smart phones are becoming some of our most trusted computing devices and people are storing their sensitive information like password, medical records and financial details on smart phones and that makes smart phones an essential platform for privacy preserving applications [9]. Also Adobe mobile consumer performed survey among the mobile consumers and found that more than 88 % people are using their mobiles for banking and other financial activities and nearly 88 % respondents are using mobile wallet for payment. They specified that people are spending 2.5 times more time on web using smart phones [10].

According to the survey of Dimensional Research 82 % Security Professionals are expecting increase in incidents of mobile security, nearly 64 % participants felt that Android is riskiest for their Corporate data security and nearly all IT professionals are concerned about impact of mobile security incidents. It has also revealed that cost of recovering from mobile security incident is going to increase significantly [11].

According to report of US-CERT smart phones popularity and relatively less security have made them attractive target for attackers. Many smart phone users have failed to use the security facilities available with their phones effectively resulting in sending of important information to the attackers without the knowledge of user [12].

Saghar Khadem mentioned in his thesis that since the manufacturers focus on development of new features and fast release of new services more than security issues and it brings Internet Security threats. He also specified various possible attacks on smart phone like Phone blocking, identity theft, spoofing and SMS spoofing etc. in his thesis [13].

Since on Internet, Smart phones may be the receiver or sender of the message, they can be infected by various types of malicious codes like viruses, worms and Trojan horses [14].

3 Visual Summary of Responses Received from the Survey

The visual summary is prepared considering 313 responses. It showcases the challenges of Information security in mobile devices. Based on the response, we have recommended important points to improve awareness about securing mobile devices by various ways. The response summary is as under:

Que. 1 what type of apps do you prefer for downloading? (Free/Paid)
Que. 2 Do you prefer to install the mobile app from the unknown sources?

Alarming facts coming out of the summary from survey is that nearly 97 % of the participants are downloading free applications. Out of those 97 % percentage, nearly 95 % participants are also downloading it from unknown sources. Free Applications and that is also particularly from unknown sources are the biggest threat for the security of the mobile devices and this behavior of the participants is the biggest challenge to reduce the number of cyber crimes related to mobile devices (see Fig. 1.)

Nearly 32 % of participants prefer to install mobile applications from the unknown sources. 32 % is significant figure for the security of the mobile devices. Unknown sources provide the opportunity to the crackers to offer their malicious applications to the users and that compromises the security of the user's device (see Fig. 2).

Nearly 46 % participants are storing critical information on their device, 14 % of them are prefer to download free apps also from unknown sources and only 8 % of them encrypt their device (see Fig. 3).

Those who prefer to download free applications from unknown sources among them 75 % are storing their important information on the device and 42 % people are not encrypting their device (see Figs. 4 and 5).

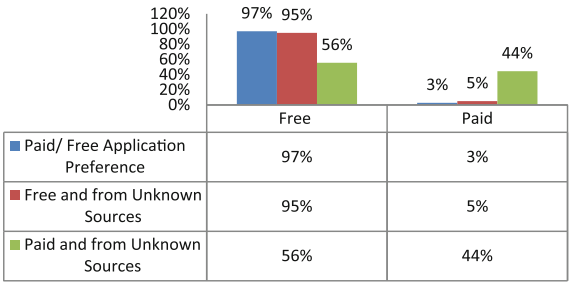


Fig. 1 Participant preferences related to mobile applications

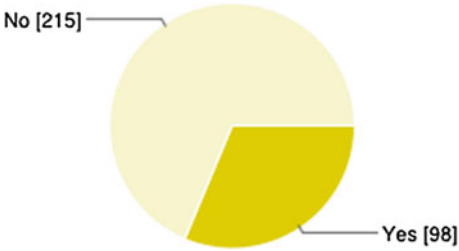


Fig. 2 App download from unknown sources

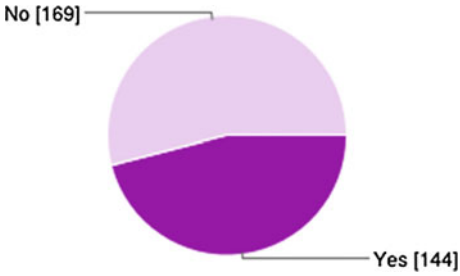


Fig. 3 Important data on device

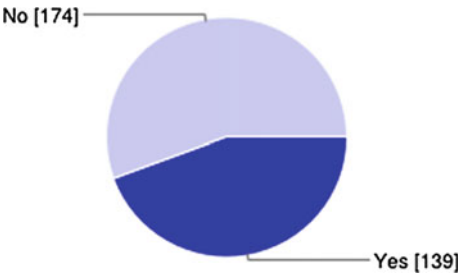


Fig. 4 Encrypted device

Nearly 41 % participants are not aware about Remote wiping functionality available with their smart phones and 37 % participants are not using it so in total 78 % participants are not using it (see Fig. 6) and we also find that only 30 %

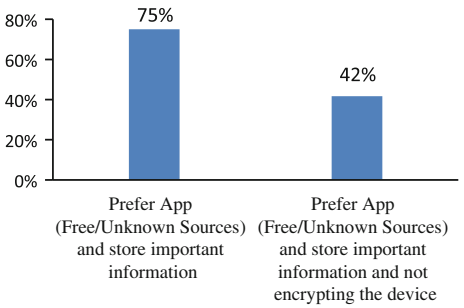


Fig. 5 Encrypted device and free app

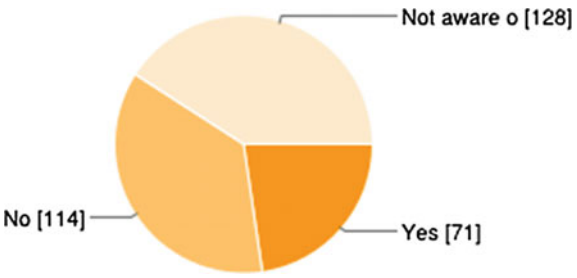


Fig. 6 Usage of remote wiping

Fig. 7 Usage of SIM lock

participants are using a SIM Lock feature which is very common and easy to use feature in the smart phones (see Fig. 7).

From Fig. 8 we can find out that nearly 61 % participants are downloading the applications without paying much attention to the required permissions by that application and Fig. 9 shows the negligence of users as 28 % of users download the app even after getting warning message. That signifies that they are not very serious about the security of their device.

Figure 10 demonstrates that 27 % participants use personal information like birth date, age and pet names as passwords which can be easily guessed by the attackers and help them to compromise the security of smart phone. Figure 11 states that 54 % users don't have any kind of software in their devices to track it in case of theft or loss.

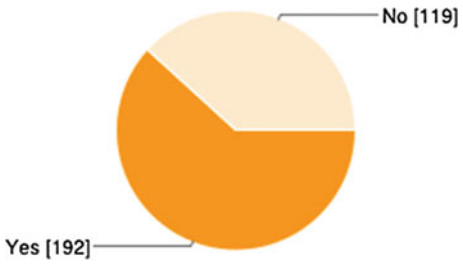


Fig. 8 User reads permissions?

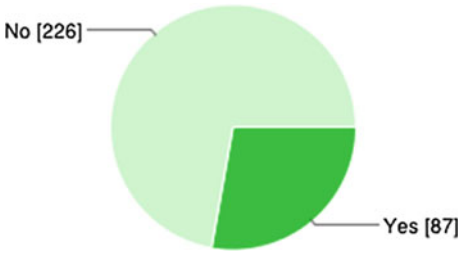


Fig. 9 Downloading app inspite of getting warning?

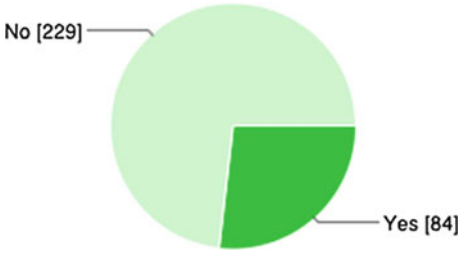


Fig. 10 Consider personal details as passwords?

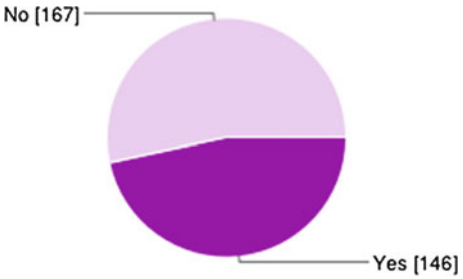


Fig. 11 Tracking application in case of theft or loss

From Fig. 1, we can identify that 97 % of users download free app and Figs. 12 and 13 clearly states that there are large number of such users who avoid warning messages of device security mechanism and also do not read the permissions granted to the app there by providing open platform to the attackers.

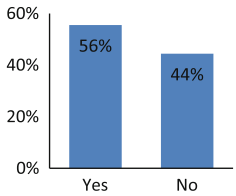


Fig. 12 Prefer free app from unknown sources on device having important information despite of warning

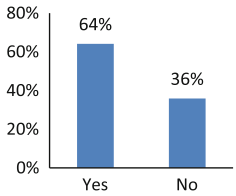


Fig. 13 Downloading free app from unknown sources on device without reading about permissions

4 Solutions

People behind the security attacks are becoming more equipped and hence the biggest mistake made by consumers is to assume that they cannot be attacked or there is no reason for them being attacked. Online frauds have increased in large number and it has become the primary means for cyber criminals. So when it comes to data security, it is advisable to be proactive rather than reactive.

4.1 Proactive Measures for Smart Phone Security

- Give importance to the security rather than look of device while purchasing.
- Use screen lock feature to avoid any unauthorized access to the data.
- Avoid downloading the apps and content from unknown source and verify the app before downloading it.
- Check the permissions that device will be granting to the app before downloading it. Also cross check if the permissions asked are required by the application.
- Avoid storing sensitive and critical information without encryption.
- Install powerful security software for the device.
- Update the device time to time and Backup the data regularly.
- Turn off the Bluetooth, Wi-Fi, GPS or Infrared service when not in use. And authenticate the device before connecting your device it to.

- Use encrypted Wi-Fi networks that require a password and ensure your device does not automatically connect to any new network.
- Record the IMEI (International Mobile Equipment Identifier) of your device which can be useful for further actions in case of loss or theft.
- Enable the remote wiping feature.
- Avoid using online banking or payment gateways over the public Wi-Fi.
- Be careful while using social networking apps and logout from the apps and sites if not in use.
- Avoid rooting or jail breaking the device as it means you enable software to acquire full Administrator privileges.
- Limit exposure of your mobile phone number as it can be a medium for the attackers to breach the device security and achieve the unauthorized access.
- Scan the device and networks in order to find vulnerabilities.
- Apply filtering and use firewalls to prevent unauthorized access.

4.2 Reactive Measures for Smart Phone Security

- Report a stolen device.
- In case of loss or theft, lock or erase a lost device.
- Format the device if any malware is suspected in the device.
- Use SIM Lock feature.
- Report the incident of impersonation.

5 Conclusion

The survey results indicate that large number of smart phone users face information security issue due to their negligence in usage of smart phones and security measures. The participants even after being aware about the scenario of cyber crime and its increasing rate in mobile market are careless for data security of their device. This also raises a serious thought about the ignorance of non IT domain users. Majority of the users are storing critical information on the device and in spite of which they prefer downloading free apps from unknown sources which increases the chances of attacks on the device. So we conclude that there is need of awareness amongst the smart phone users about common threats and proactive actions to deal with it.

References

1. TechEdge: Mobile addiction growing at an alarming rate, CNBC, 24 April 2014. http://www.cnbc.com/id/101608691. Accessed 25 Feb 2015
2. Statista: Global market share held by smart phone operating systems from 2009 to 2014. http://www.statista.com/statistics/263453/global-market-share-held-by-smartphone-operating-systems. Accessed 1 March 2015
3. Sophos: Security threat report 2013. http://www.sophos.com/en-us/medialibrary/pdfs/other/sophossecuritythreatreport2013.pdf. Accessed 17 Feb 2015
4. Darlene Storm: Security is sexy, computer world, 20 Aug 2014. http://www.computerworld.com/article/2598704/mobile-security/new-attacks-secretly-use-smartphone-cameras–speakers-and-microphones.html. Accessed 15 Feb 2015
5. 12 Most abused android app permisions: Trend Micro. http://about-threats.trendmicro.com/us/library/image-gallery/12-most-abused-android-app-permissions. Accessed 5 March 2015
6. Nayak, A., Priety, T., Alshamlan, M., Yen, K.: Android mobile platform security and malware survey. Int. J. Res. Eng. Technol. **2**(11) (2013)
7. Siau, K., Lim, E.-P., Sheu, Z.: Mobile commerce : promises. challenges and research agenda
8. Nielsen: The mobile consumer report. http://www.nielsen.com/in/en/insights/reports/2013/mobile-consumer-report-february-2013.html. Accessed 14 Feb 2015
9. Huang, Y., Chapman, P., Evans, D.: Privacy-preserving applications on smartphones. In: 6th USENIX Workshop on Hot Topics in Security (HotSec 2011), San Francisco, 9 Aug 2011
10. Adobe: Mobile consumer survey 2014. http://offers.adobe.com/en/na/marketing/landings/_54458_amc_mobile_survey_report.html. Accessed 22 Feb 2015
11. Dimensional Research: The impact of mobile devices on information security a survey of IT professionals. http://www.checkpoint.com/downloads/products/check-point-mobile-security-survey-report.pdf. Accessed 17 March 2015
12. Foote, J., Ruggiero, P.: Cyber threats to mobile phones (2011)
13. Khadem, S.: Security issues in smartphones and their effects on the telecom networks (2010)
14. Guo, C., Wang, H., Wenwu, Z.: Smart-Phone Attacks and Defenses: ACM SIGCOMM HotNets, Association for Computing Machinery, Inc., Nov 2004

Software Architecture Validation Methods, Tools Support and Case Studies

Niranjan Kumar

Abstract Foundation of any software system is its architecture. However, intended architecture may lose capability when the implementation does not conform to the designed architecture. In this paper, a set of methods and tools is presented to perform architecture validation for maintaining consistency between defined architecture and implementation. This paper compares three proposed tools on various parameters by assessing its capabilities to detect architectural violations (i.e., deviations between the intended architecture and the implemented architecture).

Keywords Architecture validation · Architecture compliance · Layer validation · Architecture consistency checking · Architecture conformance rules

1 Introduction

Software architecture is the blue print of software and guides the development and evolution of the software [1]. Software application starts with good architecture and design but its quality gradually degrades as application evolves. The defined architecture is not being followed during course of software evolution or maintenance and that leads to drift between architecture and implementation. Software architecture eventually deviates from the original intent and degrades when implementation does not conform to the designed architecture. To ensure architectural compliance, regular checks are essential. Architecture compliance check cannot be done manually because of shier size and complexity of software applications but it requires adequate tool support. This problem may be solved by continuous integration of architecture validation or architecture compliance check during software implementation or application build process. The architecture validation ensures that the code and the architecture are inline during the development process.

N. Kumar (✉)
Siemens Technology and Services Pvt. Ltd, Bangalore, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_32

Features of Architecture Validation [2]:

1. Discover conflict between existing code and defined architecture.
2. Assists in understanding the impact of code change in application.
3. Communicate between existing architecture and existing code base.
4. Establish rules so that developers can maintain compliance to the desired architecture during code development.

The paper presents methods and tools for architecture validation of software application and mechanism to prevent architectural violation during application build process. The paper presents a comparative study done on different methods and tools that realizes architecture validation and gives insight to try out the methods of architecture validation in your project. The paper explains steps by steps process guidance in validating architecture and analysis during software application development. Section 3 presents tools and static methods for validating software architecture and that can be implemented to software intensive system. Section 4 compares these tools capabilities to detect architecture violations on various standards.

2 Purposes and Needs

Why to Validate the Application architecture? The earlier you find the problem in you project the better you off you are. The cost to fix an error during early development phase is less than same error found post development. The foundation of any software system is its architecture. A little deviation in architecture will precipitate disaster on a project and lead to architectural decay. Architectural decay has been defined as non-compliance between conceptual and concrete architecture; deviation from original design architecture. Architecture validation is a way to avoid disaster and keep architect intact during course of development. Architecture validation does not say anything about defined system architecture that it is right or wrong but it certainly detect architecture violation or the conflicts between defined architecture and the existing code base.

Architecture compliance checking is a sound instrument to detect architectural violations (i.e., deviations between the intended architecture and the implemented architecture). Architecture compliance can be checked statically or dynamically. This paper explain the three static approach followed by tools to identify the architecture violation in software application project [3].

2.1 Reflexion Model

In Reflexion model, Architect build a conceptual high level model of intended architecture of the system. It compares two models of a software system against

each other, an architectural model (the planned or intended architecture) and a source code model (the actual or implemented architecture [4–6]. In this method mapping of high level components with low level source code model is performed manually. Visual Studio apply reflexion model to check architecture validation.

2.2 Relation Conformance Rules

Relation conformance rules enable specifying allowed or forbidden relations between two components [7, 8]. In this method, mapping of intended architecture and source code model is performed automatically. NDepend tool uses Source code query languages (SCQL) to check architecture conformance.

2.3 Dependency Structure Matrices

Lattix Inc's Dependency Manager (LDM) includes a simple language to declare architecture conformance rules that must be followed by the target system implementation. A can-use B and A cannot-use B, indicating that classes in the set A can (or cannot) depend on classes in B. Violations in architecture and design rules are automatically detected and visually represented in the extracted DSM [9]. NDepend and Lattix apply Dependency Structure Matrices (DSM) method to automatically detect architecture rule violations. Table 1 summarizes the different validation methods.

Table 1 Comparative study of architecture compliance/validation methods

Comparative study of architecture validation methods			
Criteria	Methods		
	Reflexion method	DSM	Component access rule
Ease of tool usability	Medium (Architect manually building intended architecture)	High (Based on inbuilt DSM algorithms)	Medium (Based on customize and default code query)
Scalability	Low	Medium	Medium
Architecture refactoring support	Medium (Highlight component dependencies violations)	High (Support for architecture restructure)	Low (No support for architecture refactoring)
Manual intervention	High (To build intended architecture and mapping)	Low (Discover architecture based on DSM algorithm)	Low (To Create design queries for architecture violation)
Probability of false +ve	High (Because of manual intervention)	Low (Automatically based on proven DSM algorithm)	Low (Based on code query rule)

3 Tools and Methods for Architecture Validation

3.1 Visual Studio 2010—Layer Diagram and Layer Validation

Layer Diagram and layer diagram validation are the main architecture feature introduced in Visual Studio 2010. Developers can also make use of Layers diagram to verify the application architecture. Layer Diagrams are used to describe the application architecture. The major advantages of these diagrams are to verify that the code satisfies the high level application architecture. This validation ensures that the code and the design architecture are inline during the development process.

Architecture validation occurs on Layer Diagrams in Modeling Projects. Add a Modeling Project to your solution and add a Layer Diagram under that Modeling Project. Once you define the layers of the solution, you can link the layers with the modules or projects in the solution. Once you define the layers and relationships, then you can validate the solution, whether it follows the architecture or not. Layer Diagram performs validation against the defined architecture or application layers to verify that the source code still conforms to the desired architecture.

As shown in Fig. 1, it depicts a case study of a web application, which is structured as a multilayered application consisting of different layers (Business, Data, Data Access, Presentation, Host and Services).Presentation layer uses the functionalities of Business layer and Business layer uses Data Access layer to

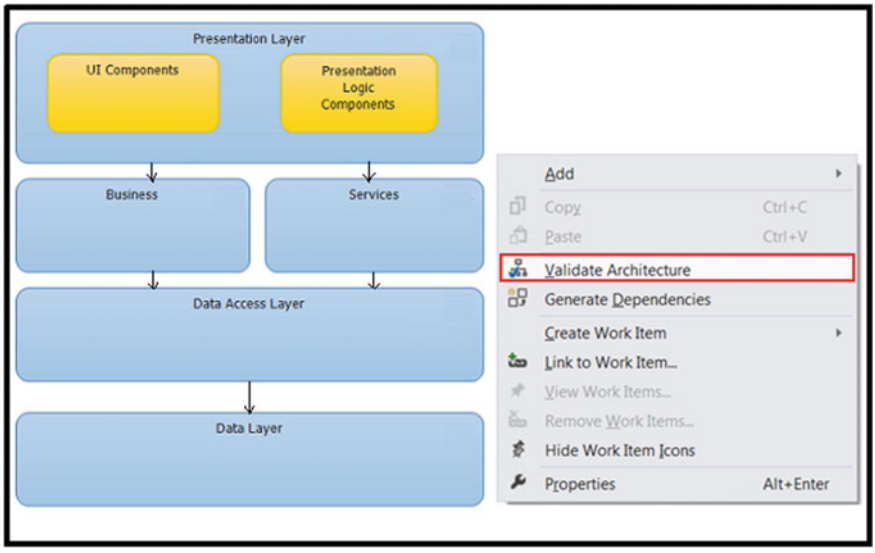


Fig. 1 Architecture validation using visual studio 2010

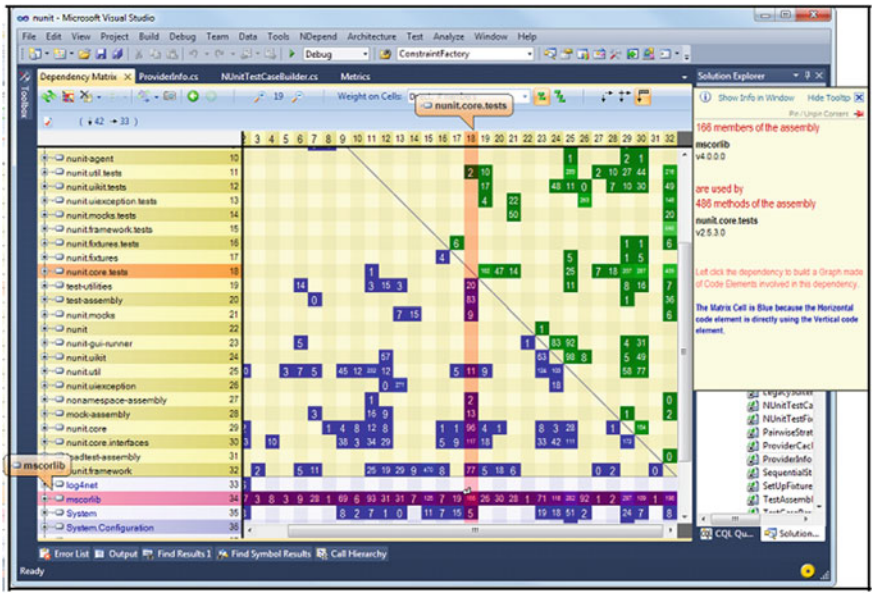


Fig. 2 NDepend architecture validation window

access the data. A developer add a code to access the Data Access Layer directly from the presentation layer that leads to architecture violation [10].

3.2 NDepend—Validating Architecture Through LINQ Query

NDepend integration with Visual Studio is very simple. A new menu item "Ndepend" appears with multiple option in it. It allows you to attach a new NDepend project to current solution and open up list of assemblies available for analysis by project as shown in Fig. 2. Architecture pattern could be actually validated through a Code Query LINQ rule. To do any of the architectural validation at compile time, need to write custom CQLinq queries. Code Language Query, or CQL, is NDepend's own query language for inspecting your code. It's syntactically very similar to SQL.

For Example, A use case of a web application, where architecture defines that presentation layer should not use Data Layer. It does not mean that Presentation layer does not interact with Data Layer but actually it interacts through using mediator Data Access Layer.

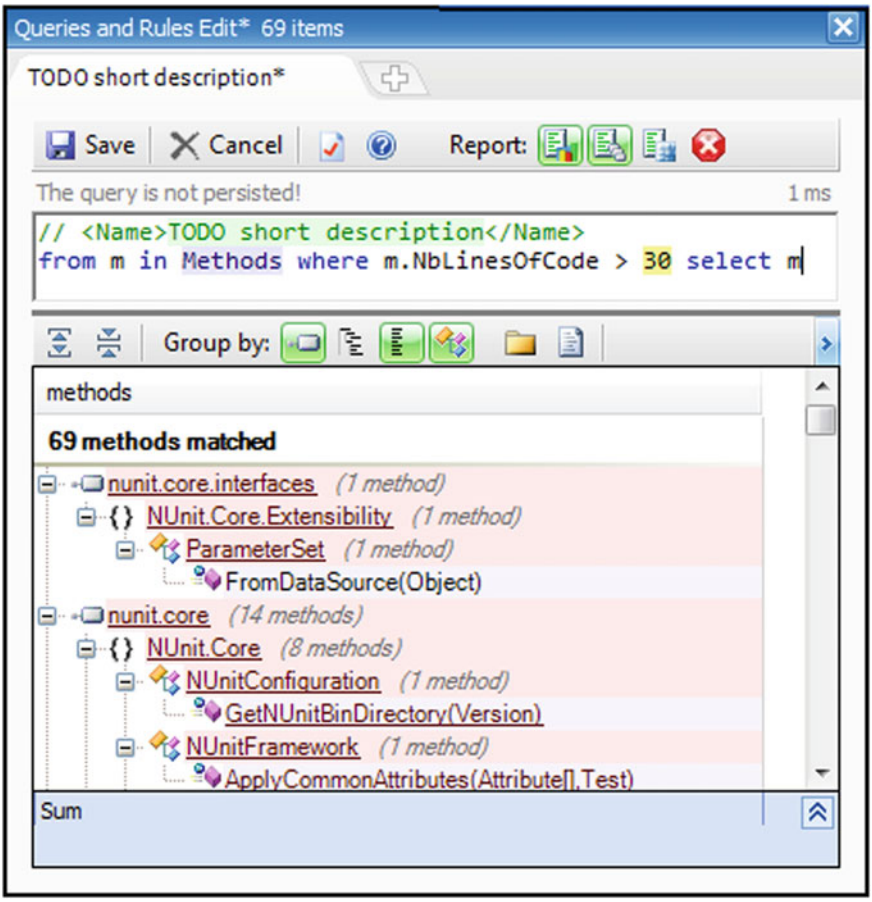


Fig. 3 NDepend CQLinq queries and rules editor

As shown in Fig. 3, CQLinq query can be edited live in the NDepend UI (standalone or in Visual Studio). The query is executed automatically as soon as it compiles. CQLinq is fast and is designed to run hundreds of queries per seconds against a large real-world code base [11]. Rules for single dependencies to ensure or forbid are made very easily by opening the dependency matrix, clicking with the right mouse button on any dependency and selecting "Generate a code rule from this". For example, the following design rule specifies that UI layer shouldn't use directly DB types [12].

```
// <Name>UI layer shouldn't use directly DB types</Name>
warnif count > 0

// UI layer is made of types in namespaces using a UI framework
let uiTypes = Application.Namespaces.UsingAny(Assemblies.WithNameIn("PresentationFramework", "System.Windows", "System.Windows.
Forms", "System.Web")).ChildTypes()

// You can easily customize this line to define what are DB types.
let dbTypes = ThirdParty.Assemblies.WithNameIn("System.Data", "EntityFramework", "NHibernate").ChildTypes().Except(ThirdParty.Typ
es.WithNameIn("DataSet", "DataTable", "DataRow"))
let dalTypes = Application.Namespaces.UsingAny(dbTypes).ChildTypes().ToHashSet()

from uiType in uiTypes.UsingAny(dbTypes)
let dbTypesUsed = dbTypes.Intersect(uiType.TypesUsed)
select new { uiType,
// Empty dbTypesUsed empty means, uiType part of DAL
 dbTypesUsed }
```

For development teams, Default Queries and code rules in NDepend helps them keeping implementation on track and warn them when code deviation start to occur.

(1) Integrating NDepend in TFS Build Process
NDepend can be integrated with build process and generate report after nightly builds. Auto generated report a way to track architecture violation for project under development. Any architectural rule violation breaks the build process and reports it. NDepend closely watches development progress and alert developers to any deviations from defined architecture rules during development process.

3.3 Lattix-DSM with Rules for Layer

Dependencies are extracted from the code by a conventional static analysis, and shown in a tabular form known as the 'Dependency Structure Matrix' (DSM). Lattix LDM provides a new lightweight approach of utilizing dependency models for automated analysis and enforcement of architectures. Design Rules can be established in the DSM model for compliance check of architecture during software application development.

The user enters 'design rules' that distinguish dependencies that violate architectural assumptions. As the system evolves over time, the rules are checked in subsequent analyses to flag deviations from the architecture. Continuous checking of design rules can be used to keep a code base in sync with its design. DSM provides a convenient way to input design rules, by having the user click on cells to identify allowed or forbidden relationships [13].

As shown in Fig. 4, conceptual architecture of Ant has been deduced, an open source regression testing framework. This DSM was created by applying the partitioning algorithm where Tasks subcomponent depends on Ant Framework and Util but AntFramework does not depend on Task component.

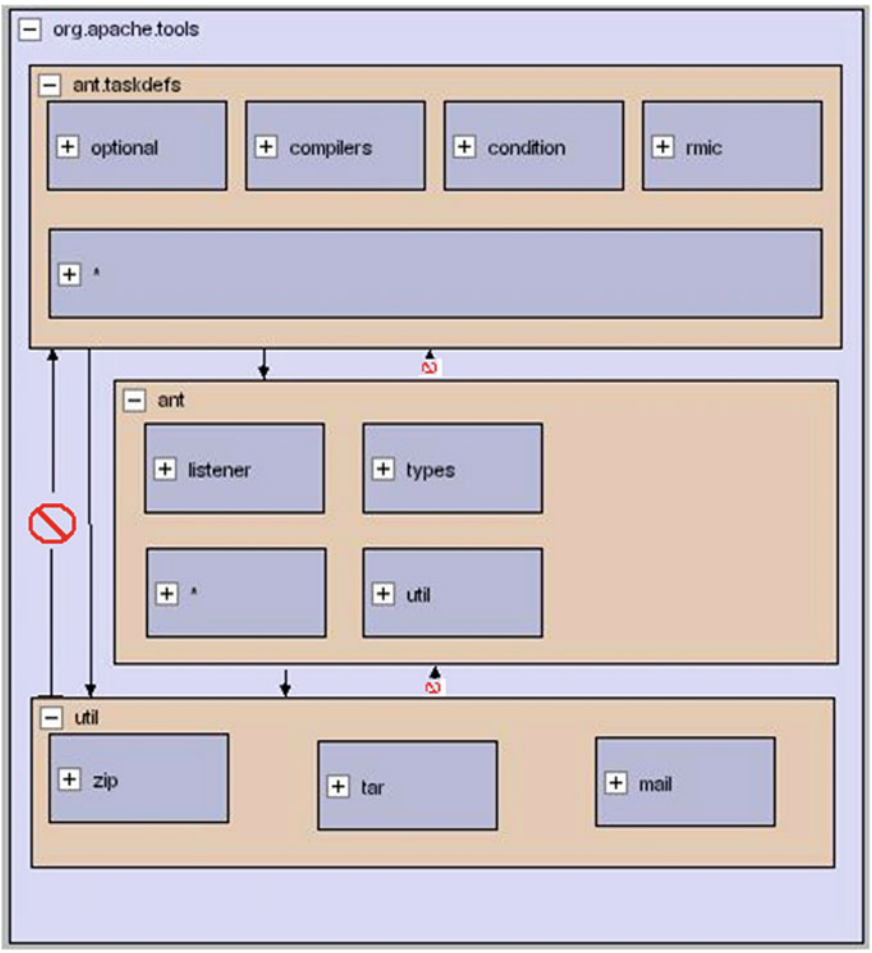


Fig. 4 Ant LDM conceptual architecture

Design Rules for Patterns: Design rules come in two forms, indicating that S1 can and cannot depend on S2

S1 can-use S2
S1 cannot-use S2

Design rule for Ant system can be specified as follows:

$Ant.taskdef can-use Ant.taskdef,$Ant,$Util
$Ant cannot-use Ant.taskdef
$Util cannot-use Ant, Ant.taskdef,$Ant
$Ant can-use Util

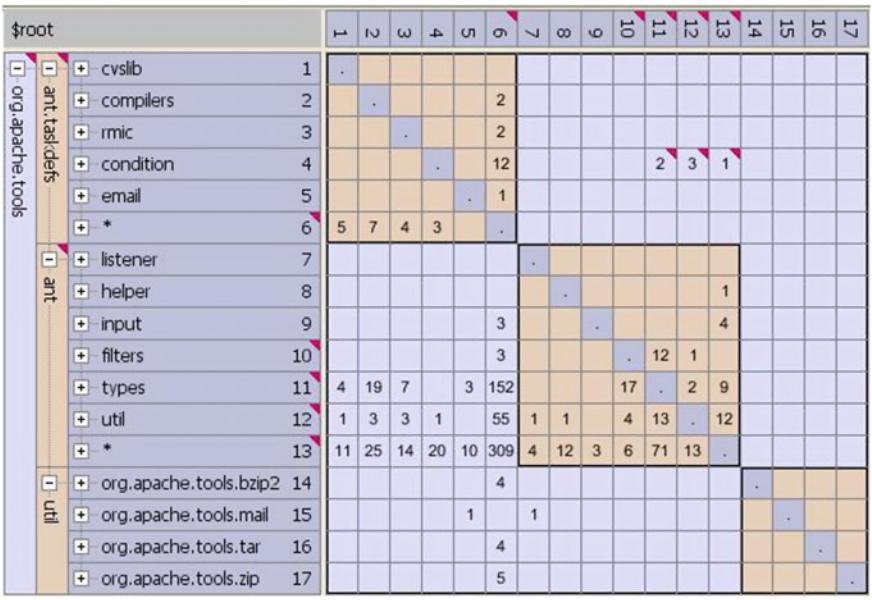


Fig. 5 Ant LDM architecture violation

$$Ant.taskdef can-use Ant,Util
$Util can-use $Util

This rule specifies that $Ant which represents testing task interface of the system, cannot access services provided by $Ant.taskdef. Next rule specify that $Ant.taskdef can access services provided by $Ant and $Util.

Design rule violation is displayed by marking its corner with red as evident in Fig. 5. It means dependency in a cell governed by a cannot-use rule is a design rule violation. So dependency model is powerful model to manage software architecture of software applications.

4 Comparative Evaluation

Following table provide summarized details of Architecture Validation feature of Lattix, NDepend and Microsoft Visual Studio 2012 Ultimate. Although approach to validate the architecture is different for these tools like Visual studio follow Line and Box approach where as Lattix and NDepend use DSM approach. Tool comparison as in Table 2 is derived, based on outcome of evaluation and case study done.

Table 2 Tool comparison

Factors	Architecture validation tools		
	NDepend	Microsoft visual studio 2012	Lattix
Ease of application usability	Complex to use, learning curve high	Simple to use, learning curve low	Simple to use
Version control	Support TFS	TFS support inbuilt	No
Licensing cost	Less	High	High
Configuration and set up	Complex	Easy	Easy
Manual step	Mapping of layer diagram elements to code	Creating CQLinq rules for validation	Layering of system for DSM
Architecture visualization	Dependencies diagram and metrics	Layer diagram and dependencies	DSM matrix/conceptual architecture view
Integration with build process	Complex	Easy	LCM can be easily integrated
Extensibility	CQLinq can be extended and customized	Validation rule is extensible	Rules are extensible
Evaluation performance	Faster	Slow	Fast
Language support	.Net/C++/Java	.Net	C, C++, Java, C#, Delphi or Ada
Reporting	HTML report	No reporting	HTML, XML report

5 Summary

This paper clearly defines the different methods to inculcate the architecture validation in software development process which helps to maintain the application's original architecture by preventing the architectural drift that can occur during course of software evolution. By integrating above proposed tools for architecture validation in application build process makes sure that code doesn't conflict with its design during the software development process. In this paper, Reflexion model, Relation Conformance Rules and DSM Rules are presented to validate architecture statically and tools that support these methods. It is recommended that a goal and context driven approach to decide on which tools and methods to apply for architecture validation in their projects. The goal driven selection approach totally based on comparative study. These tools have limited support so they could not adequately check all aspects of the intended architecture.

References

1. Riaz, M., Sulayman, M., Naqvi, H.: Architectural decay during continuous software evolution and impact of 'design for change' on software architecture. Adv. Softw. Eng. Commun. Comput. Inf. Sci. **59**, 119–126 (2009)
2. Application Architecture: http://msdn.microsoft.com/en-us/library/ee658104.aspx
3. Knodel, J., Popescu, D.: A Comparison of Static Architecture Compliance Checking Approaches. Fraunhofer Institute for Experimental Software Engineering (IESE), Germany
4. Murphy, G.C., Notkin, D., Sullivan, K.J.: Software reflexion models: bridging the gap between (2001)
5. Koschke, R., Simon, D.: Hierarchical reflexion models. In: 10th Working Conference on Reverse Engineering (WCRE 2003), 13–16 Nov 2003, Victoria, Canada, pp. 36–45 (2003)
6. Knodel, J., Lindvall, M., Muthig, D., Naab, M.: Static evaluation of software architectures. In: 10th European Conference (2006)
7. van Ommering, R., Krikhaar, R., Feijs, L.: Languages for formalizing, visualizing and verifying software architectures. Comput. Lang. **27**(1/3), 3–18 (2001)
8. Postma, A.: A method for module architecture verification and its application on a large component-based system. Inf. Softw. Technol. **45**(4), 171–194 (2003)
9. DSM Knowledge: http://www.dsmweb.org/en/dsm-knowledge.html
10. Application Architecture: http://msdn.microsoft.com/en-us/library/ee658104.aspx
11. Define a Gated Check-In Build Process to Validate Changes: http://msdn.microsoft.com/en-us/library/vstudio/dd787631.aspx
12. NDepend Default UI Code Query Rule: http://www.ndepend.com/DefaultRules/Q_UI_layer_shouldn't_use_directly_DAL_layer.html
13. Sangal, N., Jordan, E., Sinha, V., Jackson, D.: Using dependency models to manage complex software architecture

A Novel Approach for Performance Analysis and Optimal Selection of Players in Indian Premier League Auction

Mayank Khandelwal, Jayant Prakash and Tribikram Pradhan

Abstract In the Indian Premier League (IPL), team owners build their cricket team by buying players in the IPL auction. Before the auction begins, the teams have the liberty to retain some of its previously auctioned players in the past IPL season. The rest of the players are available for selection via auction. Initially, all the owners of the teams have the same limited amount of funds to build their team. Naturally, the more players an owner retains, the lesser funds the owner would have to enter into the auction. Therefore, the decision of retaining players has to be perfect for an optimal selection of retaining players as well as selection of players in the auction. We analyze the requirement of the structure of the team, based on voids created due to the remaining players after the selective retaining process. For an optimal decision making in the auction, we define the size and type of voids clearly, which helps the owner select the best combination of players in the auction. Our proposed method attempts to ensure that the owner will be aware of his next steps clearly, he or she buys a player in the auction and direct their funds to buy specifically those players that will fill the voids in the team. We compute *Most Valuable Player (MVP)* by using player's batting points, bowling points and player experience. After obtaining the MVP values, we classify the players by using decision tree approach. Further, we try to find out the players responsible for success of the team and how any two players tend to play in an IPL match.

Keywords MVP · Decision tree · Correlation analysis · Symmetric measure

1 Introduction

BCCI debuted the Indian Premier League (IPL) which is a Twenty20 (T20) cricket extravaganza in 2008. It is held in the month of April–June on an annual basis. As of 2015, IPL consists of eight teams which represent eight cities of India: Chennai

M. Khandelwal (✉) · J. Prakash · T. Pradhan
Department of I&CT, Manipal University, Manipal 576104, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_33

Super Kings (CSK), Delhi Daredevils (DD), Kings XI Punjab (KXIP), Kolkata Knight Riders (KKR), Mumbai Indians (MI), Rajasthan Royals (RR), Royal Challengers Bangalore (RCB) and Sunrisers Hyderabad (SRH). All eight teams are owned and managed by franchises. IPL is the most popular T20 league in the world since it was the first sporting event to be broadcasted on YouTube. Franchises select cricketers through the IPL auction, thus building their team. The team winning the tournament is awarded a prize money of Rs. 150,000,000, runner-up is awarded Rs. 100,000,000 and Rs. 75,000,000 each for 3rd and 4th team. The last four teams get no prize money.

2 Objective

We attempt to identify the *Most Valuable Player* among all the participants in an IPL auction. MVP is dynamic in nature. That means, after every player selection there will be a variation in team's batting and bowling requirements. Based on the requirement value, the MVP value will also vary. Initially, we use decision tree for the classification of players of a particular team with class A, B, C and D respectively. We also analyze the requirement of the owner and suggest which players would have more value if added to the team based on the type of already selected players. Further, we analyze the contribution of individual players for a particular team using the concept of *Correlation Measure using Lift*. Finally, the similarity can be measure among players of a team by using the concept of *symmetric & asymmetric binary variables*.

3 Literature Survey

P. Kansal, P. Kumar, H. Arya and A. Methaila in [1] suggested a method for estimating base price of a player based on his past performance and predicting his selection. This can help the decision making authorities to set price for the players. The authors have used Näive Bayes Theorem, Multilayer Perceptron and J48 Algorithm to compare and arrive at the results. They arrived at the conclusion that Multilayer Perceptron gives the best results.

S. Singh, S. Gupta and V. Gupta in [2] proposed an integer programming real-time model for optimal strategy for binding processes. Spreadsheets were used to document and calculate the results. Spread sheets was the optimal choice considering that flexibility for more weight-age based on recent performance of a player can be easily incorporated to evaluate the final outcome.

S. Singh in [3] uses Data Envelopment Analysis to measure how effective teams are in IPL. The author calculates awarded points, total run rate, profit and returns by determining that total expenses including the wage price of players and staff as well as other expenses. Efficiency score is usually directly related to the performance of

the player in the league. On decomposing the inefficiencies into technical and scale inefficiency, it is realize that the inefficiency is primarily due to un-optimized scale of production and un-optimized transformation of the results and the considered data.

P. Kalgotra, R. Sharda and G. Chakraborty in [4] develop predictive models which aid managers to select players for a talented team in the least possible price. This is calculated on the basis of the player's past performance. The author uses SAS Enterprise Miner 7.1 to build the models. The optimal model is selected on the basis of the rate of validation data misclassification. This model helps in the selection of players by aiding in the author's bidding equation. This research also facilitates the managers to set the salaries for players.

F. Ahmed, K. Deb and A. Jindal [5] use NSGA-II algorithm to propose a new representation scheme and a multi-objective approach for selecting players in a limited budget considering the batting and bowling strengths along with the team formation. Factors such as fielding further optimize the results. The dataset to define performance is taken from IPL–2011 Edition. The author shows analysis in real-time auction events, selecting players one-by-one. The author argues that the methodology can be implemented across other fields of sports such as soccer etc.

S.K. Rastogi and S.Y. Deodhar [6] attempt to find out relevant attributes and their relative valuations. The author uses bid and offer curve concept. This concept is adapted from hedonic price analysis and establish a relation between the bid amount and player characteristics econometric-ally for IPL (2008).

Sonali B and Shubhasheesh B [7] focus on how teams strategically decide on the final bid amount based on past player and team performance in IPL and formats similar to IPL. The authors also shed light on how personalities of players can affect team performance. They analyze the possible factors based on which bidders decide and build a predictive model for pricing in the auction. The analysis is done individually for all the teams.

P.K. Dey, D.N. Ghosh and A.C. Mondal [8] propose that the contribution of each cricketer to team performance can be quantified and performance evaluation of the cricketers is a vital issue. The study measures the performance of bowlers and compute rankings bases on performances using AHP and TOPSIS methods. Performances are computed using AHP-TOPSIS and AHP-COPRAS providing the ranking.

J.M. Douglas and N. Tam [9] take the success in relation with batting, bowling and fielding variables associated with into consideration. The authors compare batting & bowling attributes of the winning and losing teams by analyzing the differences magnitude using Cohen's Effect Size concept. They suggested that the primary indicators for success were losing less number of wickets during the powerplay, having a high 'runs per over' score, score more runs in middle 8 overs and maximizing the number of dot balls which are bowled. They concluded by stating that teams should focus on maximizing 50 + run partnerships, batsmen who hit boundaries, taking wickets and delivering maximum dot balls.

H. Saikia and D. Bhattacharjee [10] classify performances of all-rounders into 'Performer', 'Batting-All Rounder', 'Bowling-All Rounder' and 'Under-performer'. Further, they suggest and consider independent variables that influence an all-rounder's performance by using Step-wise Multinomial Logistic Regression (SMLR). The independent variables are used to predict the class of an all-rounder player using Naive Bayes Classification concept.

P.K. Dey and D.N. Ghosh [11] propose a methodology'AHP-ANOVA-TOPSIS' by identifying the attributes for consideration. The authors then assign weights to the attributes in order to create a decision matrix. The overall contribution of all the decision attributes is computed, after which the total eight for each attribute is calculated. Finally, after computing overall assessment measures for all alternatives, the alternatives are ranked.

4 Methodology

4.1 MVP Calculation

In this section, we need to find out the player's batting points (PBT), player's bowling points (PBW) and player's experience (PEX). In order to find out the above three formulae's, we need to consider the following parameters: Player's Batting Average, Player's Batting Strike Rate, Number of centuries and half-centuries, Bowling Average, Bowling Strike Rate, Economy, Number of 4-wicket and 5-wicket haul and Number of Matches Played. We define the 'Most Valuable Player' (MVP) as the single parameter that can be used to compare any type of player in the auction. MVP is decided on the basis of requirement of type of player selected by the owner. For this, we need the 'Requirement Points' (minimum required in the team) for batting(BARP), bowling(BORP) and experience(ERP). 'Total Requirement' (TRP) is the sum of all requirement points (Batting + Bowling + Experience) i.e.

$$\mathrm{TRP} = \mathrm{BARP} + \mathrm{BORP} + \mathrm{ERP}$$

$$\begin{aligned}\mathbf{PBT} = (((&\mathrm{BattingAverage} * 0.3) + (\mathrm{BattingStrikeRate} * 0.4) \\ &+ (\mathrm{floor}\ (\mathrm{NumberofHundreds})\ * 0.1) + (\mathrm{Number\ of\ Fifties} * 0.2))/10)\end{aligned} \tag{1}$$

$$\begin{aligned}&\textbf{If that the bowler must have bowled minimum 100 bowls in his IPL career, then, PBW} = \\ &(((300/\mathrm{BowlingAverage}) + (200/\mathrm{BowlingStrikeRate}) + (300/\mathrm{Economy}) \\ &+\mathrm{floor}(\mathrm{Numberof4} - \mathrm{wicketshaul}) * 0.1 + \mathrm{floor}\ (\mathrm{Numberof5} - \mathrm{wicketshaul}) * 0.1)/10)\end{aligned} \tag{2}$$

$$\textbf{PEX} = (\text{Number of Matches Played}/\text{Total Number of Matches in IPL so far}) \quad (3)$$

$$\textbf{If PBW} = \textbf{0}\,\text{then, MVP} = (8 * \text{PBT} * (\text{BARP}) + (\text{PBW} * \text{BORP}) + (\text{PEX} * \text{ERP}))/(\text{TRP} * 10) \quad (4)$$

$$\textbf{If } \frac{\textbf{PBT}}{\textbf{PBW}} > = \textbf{2}\,\text{then, MVP} = (7 * \text{PBT} * (\text{BARP}) + (2 * \text{PBW} * \text{BORP}) + (\text{PEX} * \text{ERP}))/(\text{TRP} * 10) \quad (5)$$

$$\textbf{If } \frac{\textbf{PBW}}{\textbf{PBT}} > = \textbf{2}\,\text{then, MVP} = (2 * \text{PBT} * (\text{BARP}) + (7 * \text{PBW} * \text{BORP}) + (\text{PEX} * \text{ERP}))/(\text{TRP} * 10) \quad (6)$$

$$\textbf{Otherwise}, \text{MVP} = (9 * \text{PBT} * (\text{BARP}) + (9 * \text{PBW} * \text{BORP}) + (2 * \text{PEX} * \text{ERP}))/(\text{TRP} * 20) \quad (7)$$

4.2 Decision Tree

Decision Tree is powerful decisive tool used for Classification and Prediction. Every node is bonded with rules that help the data to be classified according to the nature defined by the rules. It is basically used in Data Warehouse for Knowledge Discovery.

Following are the features of a Decision Tree:

- There must be finite number of distinct attributes for classification.
- Target values of data used for classification should be discrete.
- There should not be any missing data which are important for classification.

Following are the components of a Decision Tree:

- **Decision Node** A non-leaf node used to make a decision according to the relevant data taken into consideration for the classification.
- **Leaf Node** Represents the final classification container holding the data post operations occurred at the Decision Node.
- **Path** It represent the result used for classification of the data from the decision node.

In Decision Tree Data is classified starting from the root node using top down approach till the leaf node is encountered. We have used decision tree to classify the

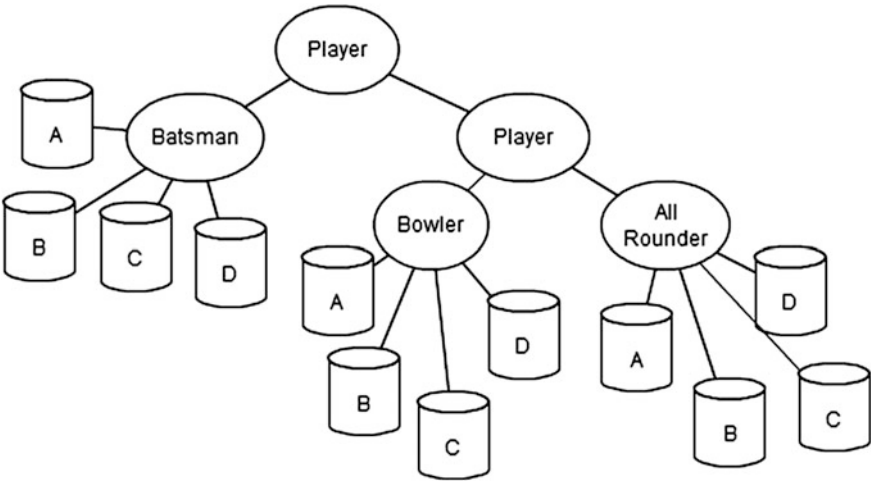


Fig. 1 Classification of players using decision tree

players into the type and class of player, as shown in Fig. 1. We follow *Algorithm 1* to calculate the type and class.

4.3 Correlation Analysis Using Lift

Lift is a correlation measure which suggests that the occurrence of *A* is **independent** of *B* if $P(A \cup B) = P(A)P(B)$. Otherwise, *A* and *B* correlation exists between *A* and *B*. We define *Lift* as follows:

$$lift(A, B) = \frac{P(A \cup B)}{P(A)\,P(B)} \tag{8}$$

If *lift* (A,B) < 1, then the occurrence of *A* is *negatively correlated* with the occurrence of B. If *lift* (A,B) > 1, then the occurrence of *A* is *positively correlated* with the occurrence of B.

If *lift* (A,B) = 1, then the occurrence of *A* is *independent* of the occurrence of B and there exists no correlation.

4.4 Computing Similarity Between Players Using Symmetric and Asymmetric Binary Variables

A symmetric binary variable has both its states (positive/negative) as equally viable and carry the same weight. There is no preference assigned to the outcome. The symmetric binary dissimilarity measure indicates the dissimilarity between objects i & j. For the values based on Table 1, we obtain:

$$d(i, j) = \frac{b + c}{a + b + c + d} \tag{9}$$

An asymmetric binary variable does not have both of its states as equally important. For instance, two positives may be given preference over two negatives. These binary variables are also known as "monary".

A binary variable is asymmetric if the outcomes of the states are not equally important. Given two asymmetric binary variables, the agreement of two 1s is considered more significant than that of two 0s. Therefore, such binary variables are often considered "monary". The dissimilarity based on such variables is called asymmetric binary dissimilarity, where the number of negative matches, t is considered unimportant and thus is ignored in the computation.

$$d(i, j) = \frac{b + c}{a + b + c} \tag{10}$$

Complementarity, we can measure the distance between two binary variables based on the notion of similarity instead of dissimilarity.

The coefficient similarity(i,j) is called Jaccard Coefficient. For example, the asymmetric binary similarity between the objects i and j is:

$$similarity(i, j) = \frac{a}{a + b + c} = 1 - d(i, j) \tag{11}$$

Table 1 A contingency table for binary variables

Object j				
	1	1A	0b	suma + b
object i	0 **sum**	c**a + c**	d**b + d**	c + d**e**

Algorithm 1 Pseudocode for Classification of Players using Decision Tree

```
procedure CLASSIFICATION(PBT, PBW)
    if (PBW = 0) or ( PBT / PBW ) >= 4 ) then
        if PBT>=7.0 then
            Batsman A
        else if PBT>=6.0 then
            Batsman B
        else if PBT>=5.0 then
            Batsman C
        else
            Batsman D
        end if
    else
        if ( PBW/PBT )>=1.25 then
            if PBW>=7.0 then
                Bowler A
            else if PBW>=6.0 then
                Bowler B
            else if PBW>=5.0 then
                Bowler C
            else
                Bowler D
            end if
        else
            if (PBT+PBW/2)>=6.5 then
                All Rounder A
            else if (PBT+PBW/2)>=6.0 then
                All Rounder B
            else if (PBT+PBW/2)>=5.5 then
                All Rounder C
            else
                All Rounder D
            end if
        end if
    end if
end procedure
```

5 Case Study

In the paper we are comparing all types of players on the basis of a single computed parameter called MVP value. The parameter is well calculated taking account of many sets of attributes that define a player's performance. In addition to player performance the MVP value also takes account of the current requirement of the team in the form of Batting Requirement points, Bowling Requirement points and Experience Requirement points. These Requirement points are decided by the owners who need to purchase players from the auction after retaining players of the team from the previous tournament. In other words, these points depict the expectation of owners from the auction. For Mumbai Indians, we calculate the MVP values and type of player in Table 2. For illustration purposes, we have defined the Batting, Bowling and Experience requirement for the IPL teams in Table 3.

In Table 4, the values of Gurinder Sandhu for *SRH* and *KXIP* are 1.8611 and 2.4413. This is because of different requirements of different teams in terms of bowling, in this case. The similar concept applies to other players & teams. More value indicates high desirability for selection as shown in Fig. 2. As the Auction proceeds and players are bought in the auction the ratio of requirement points are

Table 2 Mumbai Indians Player details with classification

Player name	Matches	Average	Stike rate	100s	50s	Balls bowled	BBM	Average	Eco	Strike rate	4w	5w	PBT	PBW	PEX	MVP	Type of player
Rohit Sharma	112	32.25	129.59	1	21	332	4 for 6	29.33	7.95	22.13	1	0	6.67	5.71	9.41	6.51	All rounder A
Aaron Finch	36	26.11	123.84	0	6	0	0	0	0	0	0	0	5.86	0	3.03	4.99	Batsman C
Kieron Pollard	77	27.18	144.31	0	6	1076	4 for 44	29.03	8.58	20.3	1	0	6.71	5.53	6.47	6.15	All rounder A
Ambati Rayudu	81	26.71	125.18	0	10	0	0	0	0	0	0	0	6.01	0	6.81	5.49	Batsman B
Harbhajan Singh	96	15.51	147.93	0	0	2037	5 for 18	24.79	6.71	22.14	1	1	6.38	6.6	8.07	6.65	All rounder A
Lasith Malinga	83	4.68	87.2	0	0	1929	5 for 13	17.66	6.53	16.21	3	1	3.63	7.57	6.97	6.72	Bowler A
Parthiv Patel	79	20.75	109.37	0	5	0	0	0	0	0	0	0	5.1	0	6.64	4.74	Batsman C

Table 3 Player requirement table for IPL teams

Team name	CSK	MI	SRH	DD	RCB	RR	KKR	KXIP
Fund (Rs)	50,000,000	100,000,000	208,500,000	400,000,000	210,000,000	130,000,000	130,000,000	118,000,000
Batting requirement (BARP)	13	22	46	72	24	17	14	9
Bowling requirement (BORP)	10	19	42	69	34	23	21	20
Experience requirement (ERP)	9	16	43	59	21	20	13	15

Table 4 List of players participating in IPL auction

Team name	PBT	PBW	PEX	MVP(CSK)	MVP(MI)	MVP(SRH)	MVP(DD)	MVP(RCB)	MVP(RR)	MVP(KKR)	MVP(KXIP)
Yuvraj Singh	6.22	6.20	7.06	2.2075	2.2084	2.1088	2.1784	2.2388	2.0979	2.2282	2.0814
Dinesh Kartik	5.90	0.00	8.91	2.1667	2.0705	1.9486	1.9608	1.6697	1.6333	1.6170	1.2685
Kevin Pieterson	6.74	5.83	2.69	2.1270	2.1200	1.9939	2.0758	2.1216	1.9543	2.1048	1.9043
Hashim Amla	6.03	0.00	6.22	2.1344	2.0363	1.8979	1.9199	1.6307	1.5739	1.5753	1.1986
Mike Hussey	6.52	0.00	4.62	2.2493	2.1432	1.9836	2.0144	1.7077	1.6322	1.6467	1.2246
Aaron Finch	5.86	0.00	3.03	1.9886	1.8934	1.7446	1.7760	1.5039	1.4284	1.4485	1.0615
Chris Morris	6.57	6.53	6.22	2.2933	2.2958	2.1850	2.2623	2.3290	2.1721	2.3171	2.1532
Kane Williamson	5.60	5.99	6.22	2.0403	2.0450	1.9526	2.0200	2.0903	1.9540	2.0821	1.9521
Irfan Pathan	5.55	5.61	8.24	2.0362	2.0377	1.9577	2.0141	2.0654	1.9510	2.0572	1.9403
Zaheer Khan	4.00	6.08	9.08	1.8413	1.8613	1.8070	1.8595	1.9655	1.8612	1.9677	1.9212
Murali Vijay	6.13	0.00	6.30	2.1703	2.0704	1.9295	1.9520	1.6579	1.6001	1.6016	1.2183
Rahul Sharma	3.66	6.24	3.70	1.6509	1.6758	1.6003	1.6711	1.8074	1.6666	1.8093	1.7395
David Wiese	7.13	6.26	6.47	2.3655	2.3587	2.2419	2.3174	2.3586	2.2041	2.3429	2.1567
Gurinder Sandhu	3.06	7.13	1.43	1.8476	1.9391	1.8611	1.9834	2.3708	2.1332	2.3996	2.4413

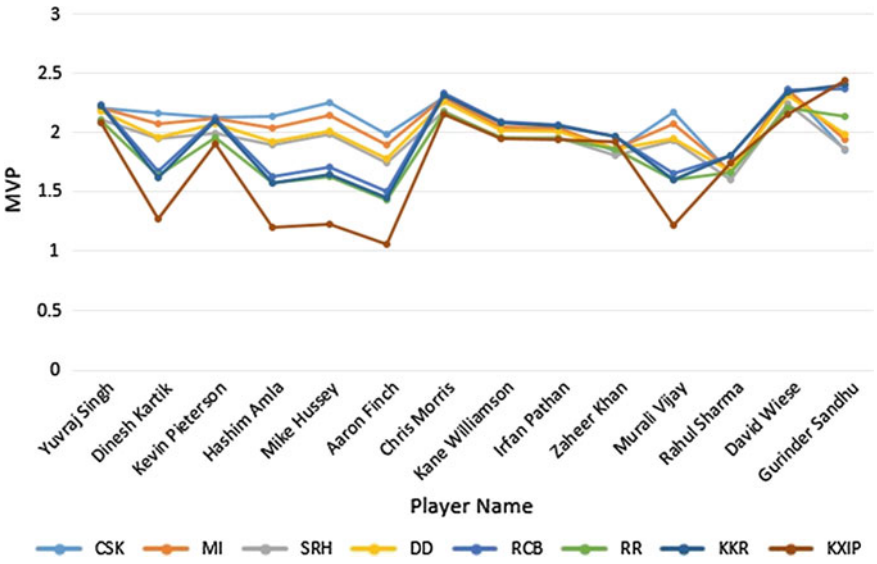


Fig. 2 MVP analysis of different players for different teams

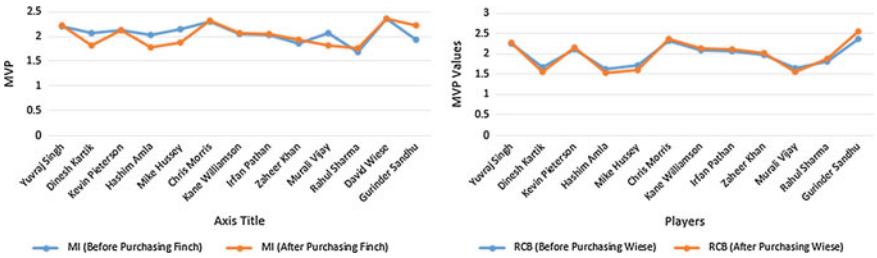


Fig. 3 MVP variation of players: Finch and Wiese before and after auction

changed and so the MVP value of other players change accordingly for the team who has made a transaction. The respective points of the player have been deducted from the corresponding requirement points of the team who purchased him. To explain this change we have taken a dataset and showed three cases where a player is bought and the MVP values got changed. The variation can be seen in Fig. 3 and Fig. 4.

- **Suppose Aaron Finch is bought by Mumbai Indians** When Aaron Finch (Batsman) is bought by Mumbai Indians
  Franchise he has PBT = 5.86, PBW = 0 and PEX = 3.03. These are deducted from Requirement Points. So owners has comparatively less requirement for a Batsman. So all Batsman and Batting All-rounder remaining will have

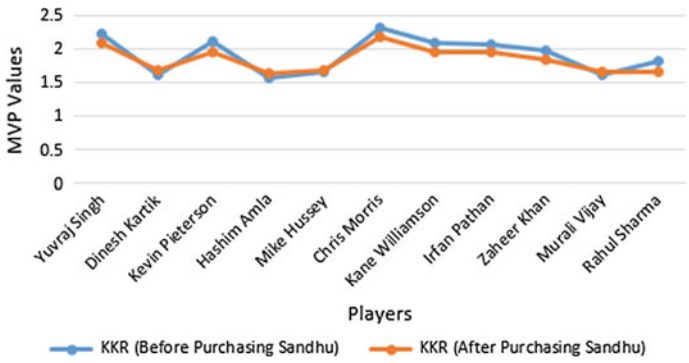


Fig. 4 MVP variation of KKR before and after auction

comparatively less MVP Value and Bowler and Bowling All Rounder will have comparatively more MVP. Batsman like Mike Hussey, Murali Vijay and Dinesh Kartik MVP value reduced for Mumbai Indians. Also Bowler like Zaheer Khan, Rahul Sharma and Gurinder Sandhu MVP Value increased for Mumbai Indians.

- **Suppose David Wiese is bought by Royal Challengers Bangalore** When David Wiese (Batting All Rounder) is bought by RCB, its PBT, PBW and PEX is deducted from RCB team requirement points. So MVP value of players for RCB changes. The Revised values show that RCB require less Batsman and Batting All Rounder as their MVP value has decreased. Also there is a slight increase in MVP in case of bowlers.
- **Suppose Gurinder Sandhu is bought by Kolkata Knight Riders** When Gurinder Sandhu (bowler) is bought by KKR, its PBT, PBW and PEX is deducted from KKR team requirement points. MVP value of players for KKR changes. The revised values show that Bowler and Bowling All Rounder importance decreases and batsmen MVP value increases before Sandhu was bought.

After analyzing the dynamic behavior of MVP, we are focusing on the classification among players of different teams by using decision tree approach. We are classifying players into four kinds of classes: A, B, C and D with the help of PBT, PBW and PEX. In this step we have to we have to follow the same method for all teams participating in IPL. In Table 2, we have shown the classes of Mumbai Indians for 7 players. Table 5 depicts how the MVP absolute values change before and after buying a player.

In Table 6, we have given the detailed results of performance of individual players of Mumbai Indians for the IPL–2014 Edition.

Mumbai, played 15 matches in the IPL–2014 edition, against other teams. Here, we are representing

Table 5 Change in MVP Values

Player name	MI (before Finch)	MI (after Finch)	Player name	RCB (before Wiese)	RCB (after Wiese)	Player name	KKR (before Sandhu)	KKR (after Sandhu)
Yuvraj Singh	2.2084	2.2312	Yuvraj Singh	2.2388	2.2805	Yuvraj Singh	2.2282	2.0918
Dinesh Kartik	2.0705	1.8225	Dinesh Kartik	1.6697	1.5643	Dinesh Kartik	1.6169	1.6798
Kevin Pieterson	2.1200	2.1256	Kevin Pieterson	2.1216	2.1614	Kevin Pieterson	2.1048	1.9511
Hashim Amla	2.0363	1.7858	Hashim Amla	1.6307	1.5287	Hashim Amla	1.5753	1.6208
Mike Hussy	2.1432	1.8747	Mike Hussy	1.7077	1.6016	Mike Hussy	1.6467	1.6822
Chris Morris	2.2958	2.3207	Chris Morris	2.3289	2.3753	Chris Morris	2.3171	2.1654
Kane Williamson	2.0450	2.0769	Kane Williamson	2.0903	2.1353	Kane Williamson	2.0821	1.9462
Irfan Pathan	2.0377	2.0580	Irfan Pathan	2.0654	2.1002	Irfan Pathan	2.0572	1.9457
Zaheer Khan	1.8613	1.9289	Zaheer Khan	1.9655	2.0197	Zaheer Khan	1.9677	1.8470
Murali Vijay	2.0704	1.8157	Murali Vijay	1.6579	1.5542	Murali Vijay	1.6016	1.6477
Rahul Sharma	1.6758	1.7615	Rahul Sharma	1.8074	1.8780	Rahul Sharma	1.8093	1.6485
David Wiese	2.3587	2.3629	Gurinder Sandhu	2.3708	2.5494			
Gurinder Sandhu	1.939	2.2138						

- **1** : If a player played well
- **0** : If a player did not play well
- **-** : If a player did not play the match

After representation, we are finding support and confidence for individual players of a team to calculate the contribution of each player. In Table 6, we have shown only the support and confidence of Mumbai Indians. Similarly, we have to find out the support and confidence for all players of all the teams participating in IPL.

To know the match winner player of a team, we are applying the concept of correlation using lift. In Table 7, we have shown the example of two players, Simmons and R. Sharma with their corresponding contribution for Mumbai Indians. We have calculated the team result with Simmons and without Simmons, and also applied the same for R. Sharma.

$$\text{P (Simmons Performed)} = 7/8 = 0.875$$

$$\text{P (Mumbai Won)} = 5/8 = 0.625$$

$$\text{P (Simmons Played} \cup \text{MI Won)} = 4/8 = 0.5$$

$$\text{Lift (MI Won, Simmons Performed)} = 0.5/(0.875 * 0.625) = 0.9143$$

Since the value is less than 1, we conclude that Simmons' Performance and Mumbai Indians are negatively correlated.

$$\text{P (R. Sharma Performed)} = 12/15 = 0.8$$

$$\text{P(MI Won)} = 7/15 = 0.46$$

$$\text{P(R. Sharma Performed} \cup \text{MI Won)} = (7/15) = 0.46$$

$$\text{Lift(MI won, R. Sharma Performed)} = 0.46/(0.46 * 0.8) = 1.25$$

Since the value is greater than 1, we conclude that R. Sharma's Performance and Mumbai Indians are positively correlated.

Similarly, we can find out the list of consistent performers, along with the match winner category for all the teams which are participating in Indian Premier League–2015.

$$\text{a} = 11; \text{b} = 10; \text{c} = 01; \text{d} = 00$$

$$\text{Dissimilarity (Tare, Rayudu)} = \frac{3+6}{3+3+6+3} = 0.6$$

Table 6 Resultant table of individual performances of players for Mumbai Indians (IPL-2014)

Matches	Team	Hussey	Tare	Rayudu	Rohit	Pollard	Anderson	Harbhajan	Gautam	Malinga
Match 1 KKR	0	0	1	1	1	0	0	0	0	1
Match 2 RCB	0	1	1	1	0	0	1	1	–	1
Match 3 CSK	0	0	1	0	1	0	1	1	1	0
Match 4 DD	0	0	0	1	0	1	1	0	1	1
Match 5 SRH	0	–	0	1	0	1	1	1	1	1
Match 6 KXIP	1	–	1	0	1	1	1	1	1	1
Match 7 RCB	1	–	0	0	1	1	0	1	1	1
Match 8 CSK	0	–	0	1	1	0	1	0	0	1
Match 9 SRH	1	–	0	1	1	0	0	0	0	1
Match 10 KKR	0	–	0	1	1	0	1	1	0	1
Match 11 RR	1	1	0	0	1	0	–	1	–	–
Match 12 DD	1	0	0	1	1	–	–	–	–	–
Match 13 SRH	1	1	1	0	1	0	–	1	–	–
Match 14 KXIP	1	1	1	1	1	0	1	0	–	–
Match 15 RCB	0	1	0	0	1	0	1	1	–	–
Total matches	15	9	15	15	15	14	12	14	9	10
Performed	7	5	6	9	12	4	9	9	5	9
Won matches		4	7	7	7	6	4	6	3	3
Performance/won		3	3	3	7	2	2	4	2	3
Support		0.5556	0.4	0.6	0.8	0.2857	0.75	0.6429	0.5556	0.9
Confidence		0.75	0.4286	0.4286	1	0.3333	0.5	0.6667	0.6667	1

Matches	Zaheer	Ojha	Bumrah	Dunk	Suyal	Simmons	Praveen	Santokie	Lange	Gopal
Match 1 KKR	1	0	–	–	–	–	–	–	–	–
Match 2 RCB	1	0	0	–	–	–	–	–	–	–

(continued)

Table 6 (continued)

Matches	Zaheer	Ojha	Bumrah	Dunk	Suyal	Simmons	Praveen	Santokie	Lange	Gopal
Match 3 CSK	0	1	–	–	–	–	–	–	–	–
Match 4 DD	0	0	–	–	–	–	–	–	–	–
Match 5 SRH	1	0	–	1	–	–	–	–	–	–
Match 6 KXIP	0	–	0	0	–	–	–	–	–	–
Match 7 RCB	–	–	1	0	0	–	–	–	–	–
Match 8 CSK	–	–	0	–	–	1	1	–	–	–
Match 9 SRH	–	0	0	–	–	1	–	–	–	–
Match 10 KKR	–	0	0	–	–	1	–	–	–	–
Match 11 RR	–	1	0	–	–	1	–	1	–	–
Match 12 DD	–	0	1	–	–	1	1	1	–	–
Match 13 SRH	–	0	0	–	–	1	–	–	1	1
Match 14 KXIP	–	0	1	–	–	0	–	–	–	0
Match 15 RCB	–	1	0	–	–	1	0	–	–	–
Total matches	6	12	11	3	1	8	3	2	1	2
Performed	3	3	3	1	0	7	2	2	1	1
Won matches	1	5	7	2	1	5	1	2	1	2
Performance/won	0	1	3	0	0	4	1	2	1	1
Support	0.5	0.25	0.2727	0.3333	0	0.8750	0.6667	1	1	0.5
Confidence	0	0.2	0.4286	0	0	0.8	1	1	1	0.5

Table 7 Analysis of match winner player for Mumbai Indians (IPL-2014)

Player/team	MI won	MI lost	Row total	Player/team	MI won	MI lost	Row total
Simmons performed	4	3	7	R. Sharma performed	7	5	12
Simmons not Performed	1	0	1	R. Sharma not performed	0	3	3
Column total	5	3	8	Column total	7	8	15

Table 8 Similarity Among Mumbai Indian Players

X, Y	b + c	a + b + c + d	Dissimilarity(X,Y)	Similarity(X,Y)
Rayudu, Rohit	3 + 6	6 + 3 + 6 + 0	0.6	0.4
Tare, Rayudu	3 + 6	3 + 3 + 6 + 3	0.6	0.4
Tare, Rohit	1 + 7	5 + 1 + 7 + 2	0.534	0.466
Anderson, Harbhajan	2 + 1	5 + 2 + 1 + 2	0.3	0.7
Harbhajan, Malinga	1 + 4	5 + 1 + 4 + 0	0.5	0.5
Anderson, Malinga	1 + 3	6 + 1 +3 + 0	0.4	0.6

$$\text{Similarity}\ (\text{Tare}, \text{Rayudu}) = 1 - \text{Dissimilarity}\ (\text{Tare}, \text{Rayudu}) = 0.4$$

$$\text{Dissimilarity}\ (\text{Tare}, \text{Rohit}) = \frac{1+7}{5+1+7+2} = 0.534$$

$$\text{Similarity}\ (\text{Tare}, \text{Rohit}) = 1 - \text{Dissimilarity}\ (\text{Tare}, \text{Rayudu}) = 0.466$$

$$\text{Dissimilarity}\ (\text{Rayudu}, \text{Rohit}) = \frac{3+6}{6+3+6+0} = 0.6$$

$$\text{Similarity}\ (\text{Rayudu}, \text{Rohit}) = 1 - \text{Dissimilarity}\ (\text{Tare}, \text{Rayudu}) = 0.4$$

This shows the similarity of batting is comparatively high for Tare and Rayudu as well as Rayudu and Rohit. Dis-similarity is comparatively high for Tare and Rohit.

After finding the match winning capability of individual players, we are trying to find out the similarity among different players by the concept of symmetric and asymmetric binary variables. In Table 8, we have given the example of similarity among Rayudu, Rohit and Tare along with the similarity among Anderson, Harbhajan and Malinga. From the above table, we are getting a higher value for Tare and Rohit in batting performance and a higher value for Anderson and Harbhajan in bowling performance. So, most likely Tare & Rohit are similar in batting as well as Anderson & Harbhajan are similar in bowling for Mumbai Indians.

6 Conclusion and Future Work

In our paper we are using MVP value concept to compare players and their dynamic change during an auction to show the effect of each transaction to a team. It aims at preparing a balanced team for any franchise. To distinguish players during an auction we are using Decision Tree concept to classify players according to their role they can get in a team. We are using correlation analysis using lift to identify match winner players according to their performance in the previous held tournament. This analysis will really identify those players whose play is a deciding factor in matches. We have also calculated similarity and dissimilarity between players on the basis of their performances in the past tournament according to Symmetric Asymmetric Binary Variables. This can correlate player to the task given in a team. Using the results of such cases, a team can find various patterns and predict various results of the efforts made by the team. This may give a new dimension for team management for better results.

References

1. Kansal, P., Kumar, P., Arya, H.: Player valuation in indian premier league auction using data mining technique. In: 2014 International Conference on Contemporary Computing and Informatics (IC3I), pp. 197–203. sJCE, Mysore, India, 27–29 November 2014
2. Singh, S., Gupta, S., Gupta, V.: Dynamic bidding strategy for players auction in IPL. Int. J. Sports Sci. Eng. **5**(1), 3–16 (2011)
3. Singh, S.: Measuring the performance of teams in the Indian Premier League. Am. J. Oper. Res. **01**, 180–184 (2011)
4. Kalgotra, P., Sharda, R., Chakraborty, G.: Predictive modeling in sports leagues: an application in Indian Premier League. In: SAS Global Forum 2013, pp. 019–2013
5. Ahmed, F., Jindal, A., Deb, K.: Cricket team selection using evolutionary multi-objective optimization, 71–78
6. Rastogi, S.K., Deodhar, S.Y.: Player pricing and valuation of cricketing attributes: exploring the IPL Twenty20 vision. Vikalpa **34**(2) (2009)
7. Bhattacharya, S., Bhattacharya, S.: Auction of players in Indian Premier League: the strategic perspective. Zenith Int. J. Multi. Res. **2**(2) (2012), ISSN 2231 5780
8. Dey, S.P.K., Ghosh, D.N., Mondal, A.C.: A MCDM approach for evaluating bowlers performance in IPL. J. Emerg. Trends Comput. Inf. Sci. **2**(11) (2011), ISSN 2079-8407
9. Douglas, J.M., Tam, N.: Analysis of team performances at the ICC World Twenty20 Cup 2009. Int. J. Perform. Anal. Sport **10**(1), 47–53(7) (2010)
10. Saikia, H., Bhattacharjee, D.: On classification of all-rounders of the Indian premier league (IPL): a Bayesian approach. VIKALPA **36**(4), 51–66 (2011)
11. Dey, S.P.K., Ghosh, D.N., Mondal, A.C.: Statistical based multi-criteria decision making analysis for performance measurement of batsmen in Indian Premier League. Int. J. Adv. Res. Comput. Sci. (IJARCS) 3(4) (2012)

Supporting Linked Databases in Keyword Query Searching Using Density Inverted Indexes

Disha M. Bilimoria, Pratik A. Patel and Manali Singh Rajpoot

Abstract Empowering the user to access databases using simple keyword search can be a relief for the user to learn SQL and to understand the schema. Keyword Search has also received a lot of attention in database as it is an effective technique to retrieve the query without knowing the underlying schema. Researches done in this area mostly deal with keyword search in single database then what for when the user is dealing with more than one database, to overcome this problem we are introducing an algorithm of SIL (Searching in Linked Databases). The other challenge in keyword processing is storing the keyword in the table i.e. via Inverted Index, so we introduce a novel technique to DeINIX (Density INverted IndeX) which reduce the memory storage space henceforth the pre-processing time is also reduced and the answer displayed will be First-10 answers. Our empirical result shows that this method answers queries more precisely, takes lesser time and memory space to retrieve a quality answer.

Keywords Keyword query search · Relational databases · Searching in linked databases · Deinix

D.M. Bilimoria (✉) · P.A. Patel
Gujarat Technological University, Ahmedabad, Gujarat, India
e-mail: dishabilimoria@gmail.com

P.A. Patel
e-mail: patelpratik1@live.in

M.S. Rajpoot
Maulana Azad National Institute of Technology, Bhopal, India
e-mail: manali.rajput89@gmail.com

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_34

1 Introduction

Database is a collection of organized data and relational database is a database structured having relationship between the stored items. Keyword Search is promptly an emerging field in the database search. Recent work in this area deals with single database. The more the relationship among the databases the complexity to fire a query increases which is the problem in traditional database; keyword query searching is problem to the solution. Keyword Search facilitates the user to access the database quickly without SQL or database schema. Keyword Search can sometimes answer some difficult and unexpected search for which SQL fail. The following is explained by illustration below. Recent work have studied that keyword search in relational database [1–3] search is focused mainly on single database. Now here we will focuses search on Linked databases.

Let's take an example of the service department. Suppose a manger wants to send an employee "Nimesh" to extend the service for a company "L&T". So what the manager does it take the COOPERATE SERVICE DEPARTMENT DATABASE (Fig. 1a) and COOPERATE DEPARTMENT (Fig. 1b) database he issues a keyword "Nimesh Ria" from the two databases. The query answer whether Nimesh was has dealing with Ria Company for any resolution of any complaint.

Outline of paper in Sect. 2 provides related work. The problem and solution are sketched in Sect. 3. The methodology of searching in single database is summarized in Sect. 4 while Sect. 5 shows how it to search in linked databases. Empirical results are provided in Sect. 6 before we conclude in Sect. 7 (Fig. 2).

2 Related Work

Many researchers have been proposed for keyword search in single databases. In paper [4] the existing approaches is divided into two major categories. The first approach is the Tuple Based approach. The data graph is taken, node here represents the tuple identifier and an edge represents the foreign key references between two tuples. These types of approach work in only one phase i.e. the system analysis

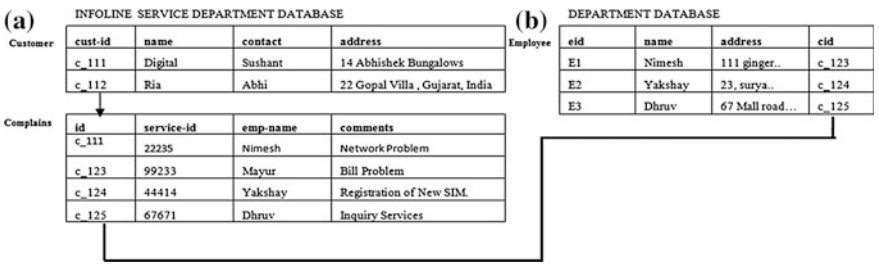


Fig. 1 Keyword search over linked databases

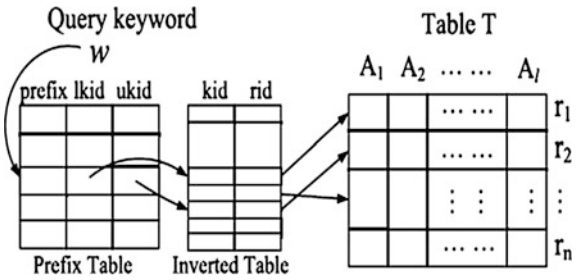


Fig. 2 Using prefix table and inverted index to support search-as-you-type [5]

the graph and build the tree and join the tuples. BANKS, BANKS-II, DPBF and BLINKS falls in this category. BANKS [1] falls under this category; it emulates distance network heuristic by running single source shortest path and uses best first search strategy around the edge in backward direction.

The second one is Schema based approach. The schema is considered as a graph where node represents data relationships and edge represents foreign keys to primary keys dependencies. Working of this approach is as follow firstly the relationship that contains keyword and schema graph is taken as input then all the possible path is discover afterwards in second part appropriate queries retrieved as the tuple DBXPLORER, DISCOVER, DISCOVER-II, PRECIS, STAR, SOAK, KEYMANTIC and KEYRY falls in this category, the algorithm uses candidate network to join the tuple and rank them according to their join. Amongst the two approaches schema based is used more as to deal with schema; which is rather easier to deal with than tuples. The algorithm proposed till date deal with single database, so in this paper we will try to focuses on more than one databases.

Guoliang Li, Jianhua Feng and Chen Li proposed a new concept in keyword search of Search-as-you-type using SQL language [5] which answers the query "on-the-fly". It means user types the keyword query and the relevant answer is retrieved. They have proposed an INDEX BASED method for a single keyword search where they have created a prefix index table for keyword search. Fuzzy search uses INCREMENTAL COMPUTING of a keyword where a character-level incremental method to find similar prefix. For multi-keyword search WORD LEVEL INCREMENTAL COMPUTATION for exact search and fuzzy keyword search is proposed. The disadvantage of this [5] is that it deal with single database and poor ranking strategy, moreover creating a prefix table and inverted index table will consume more space in memory.

3 Problem Statement

The problem defines of the keyword search over linked databases. The keyword search in single database is explained in detail in Sect. 4. The exact keyword search for the single database [5] can be done by LIKE predicate which is already

predefined in SQL and also by creating a User-Defined function in SQL. These were the two method which works without index. The exact keyword search for indexing method is explained in Sect. 4. Both the methods describe deals with single database which is the drawback. We propose a keyword search technique which searches in more than one database. For keyword searching the inverted index table is used to store the keywords but this method will consume more memory space. So, we will remove both the drawbacks by proposing a new approach SIL using DeINIX which will decrease the time complexity as well as the space complexity.

4 Searching in Linked Databases

In [5] searching for single-keyword queries in single database to support search-as-you-type is done using SQL by two different methods:

4.1 No-Index Method

1. Using Like predicates: In SQL there is an in-built predicate LIKE which allows matching the keyword and retrieving the matching record.
2. Using User-Defined Function (UDF): We can add the functions into the databases to check whether a record contains the query keyword or not.

The above two method are less space consuming but it takes more time as it need to scan all the record in table.

4.2 Index Based Method

In [5] an auxiliary table is constructed for indexing to facilitate the prefix search. The Oracle and SQL have inbuilt facilities use prefix search. But this method is limited to some databases hence a method introduced in [5] which construct Inverted index and Prefix table for searching more efficiently.

For constructing an inverted index table, firstly an unique id is assign to keywords in alphabetic order, which are in Table T. Afterwards inverted index table I_T is created with records in form <kid,rid> where kid is the id of keyword and rid is id of record. For a table T, Prefix table P_T is created with <p,lkid,ukid> where p represent the prefix of keywords, lkid the largest string id of prefix p and ukid smallest string id of p.

Let's take an example; Table 2 of inverted index is created from Table 1. The inverted index has tuple <k_8,r_6> the k_8 has keyword "effective" is in record r_3. The prefix table constructed is ("ef",k_3,k_4), now k_3 has "efficient" with minimum length and k_4 has "effective" with maximum length. Hence id will in range <k_3,k_4>. For a

Table 1 dblp: a sample table on publication (about "keyword")

ID	Title	Authors	Publications	Year
r1	An effective fuzzy keyword search scheme in cloud computing	He Tuo, Ma Wenping	IEEE	2013
r2	A framework for processing keyword-based queries in relational databases for exact record	Anshu Kumar Dwivedi, Ashok Kumar Sharma	IJSCE	2014
r3	An efficient approach for ranking based fuzzy keyword searching in XML data	Santhoshi Matha, B. Rakesh	IJDSCT	2008
r4	A system of adaptive keyword search in XML	Nan Li, Weidong Yang, Hao Zhu, and Guansheng Zhu	IEEE	2013
r5	Efficient keyword search for smallest LCAs in XML databases	Y. Xu, Y. Papakonstantinou	SIGMOD	2005
r6	Effective keyword search for valuable LCAs over XML documents	G. Li, J. Feng, J. Wang, L. Zhou	VLDB	2007
r7	Fast nearest neighbour search with keywords	Yufei Tao, Cheng Sheng	IEEE	2014
r8	Keyword search on spatial databases	De Felipe, Hristidis V, Rishe N	IEEE	2008
r9	Secure ranked keyword search over encrypted cloud data	Cong Wang, Ning Cao, Jin Li, Kui Ren	IEEE	2010
r10	Effective keyword search for valuable LCAs over XML documents	Guoliang Li, Jianhua Feng, Jianyong Wang, Lizhu Zhou	ACM	2007

Table 2 (a) Keyword table (b) inveterd index table (c) prefix table

(a) Keyword table		(b) Inverted index table		(c) Prefix table		
Kid	Keyword	Kid	Rid	Prefix	Lkid	Ukid
k1	Ieee	k1	r4	i.e.	k1	k2
k2	Sigmod	k2	r5	si	k5	k5
k3	Efficient	k3	r3	ef	k3	k4
k4	Effective	k4	r1	x	k5	k5
k5	Xml	k5	r3	key	k6	k5
k6	Keyword	k6	r1	i	k1	k5

given Partial Keyword w, we first extract the range of prefix table $P_{T,}$ then find the record in the following range which is sown in Fig. 3. The SQL query for the same is written as below:

SELECT T.* FROM $P_{T,}$ $I_{T,}$ T WHERE P_T.PREFIX ='*W*' AND P_T.ukid ≥ I_T.lkid AND $I_{T.}$rid = T.rid. [5].

With the above SQL we can retrieve the same for "ef" keyword

SELECT dblp.* FROM $P_{dblp,}$ $I_{dblp,}$ dblp WHERE P_{dblp}.prefix = 'ef' AND P_{dblp}.ukid ≥ I_{dblp}.kid AND $I_{dblp.}$rid = dblp.rid

Fig. 3 Algorithm for SIL using DeINIX

Input: Query Keyword K
Output: DeINIX Table D; Set of answer of Keyword K
For all K in [1, n] do
Let rk be the first interval list of R_k
Insert lb (rk) and ub (rk) to L and U
While U≠NULL do
Let l be top (max) element in L
Let u be top (min) element in U
If l≤u then Add [l, u] to I the interval list
If l ≠ u
Let i € Ij be consecutive interval Let i' be the next interval (if any) in Ij
Then l (i') and u (i') to I_k
Return D
Determine the set of query keyword K from G
Join the Foreign-Key FK if same for two tables
Return Keywords

The drawback of this it works on single database more over the Keyword Table, Inverted Index and Prefix table which are created consume more space and time too. The second one is that it uses SQL, for that we need to learn an underlying schema to search a Keyword. So we propose an efficient algorithm searching in CO-Related Databases with DeINIX which will answer the keyword queries without knowing the schema of database on top takes less time and less space (Fig. 2).

5 Work Flow of the Proposed Work

This section explains the general architecture and work flow of the PROPOSED work. The PROPOSED work consists of the two components one is the Pre-Processing and the other is Query Processing. The pre-processing and the query processing are explained in detail.

5.1 Pre-processing

Firstly the database is collected and then table of Inverted Index is created. The Inverted Index table is created by using DeINIX. In DeINIX traditional Inverted Index technique, keyword occurring in the database has been assigning different IDS (Table 3). Let me explain by an example in Table 3 the word “xml” is occurring in r3,r4,r5 and r6 these index are stored individual in Inverted Index table (Table 3a) which result in more memory utilization, hence we propose a technique

Table 3 Traditional inverted index (a) and DeINIX (b)

(a) Inverted index		(b) DeINIX	
Keyword	Rid	Keyword	Rid (range)
Ieee	r1,r4,r7,r8,r9	Ieee	<r1,r1>,<r4,r4>,<r7,r9>
Sigmod	r5	Sigmod	<r5,r5>
Efficient	r3,r5	Efficient	<r3,r3>,<r5,r5>
Effective	r1,r6,r10	Effective	<r1,r1>,<r6,r6>,<r10>
Xml	r3,r4,r5,r6,	Xml	<r3,r6>
Keyword	r1,r2,r3,r4,r5,r6,r7,r8,r9,r10	Keyword	<r1,r10>

DeINIX to support keyword search using range list. Instead of individual we set them in range intervals <r3,r6> which reduces the memory space of the Inverted Index table and make the performance better and fast. Again a question arises what if a we have keyword like "sigmod" is occurring only once in database, the solution for that is we have to write <r5,r5> for occurrence of the keyword once. Other condition is if we can't have continuous interval then we can write as <r1,r1>, <r6, r6> and <r10,r10> for "effective" keyword. After construction of table then we processed towards query processing.

5.2 Query Processing

After the pre-processing is done then the user enters the keyword then it proceeds towards following steps:

Step 1: Keyword Matching

The basic idea of keyword matching is to find the answer for the keyword query. User types the keyword and fetches the answer by retrieving form the DeINIX table and applies Searching in Linked Databases.

Step 2: Searching in Linked databases

The main challenge is to process the keyword over Linked databases. In this paper algorithm is propose to link the databases via foreign keys. In order to join the foreign keys we need to first of all identify the foreign keys in the individual database. The two databases must be sharing approximate common attribute field then only they will be joinable. The next step is to link up two databases with help of foreign keys. Now the databases are link the selection of the keyword is based on DeINIX table. After the searching is completed it displays the result and will go to next step 3.

Step 3: First-10 answers

After the query is generated, multiple result tuples are produced. Now the question arises which answer is relevant to the keyword search for the end user. In this paper we display only first-10 answer from several tuples displayed in the list.

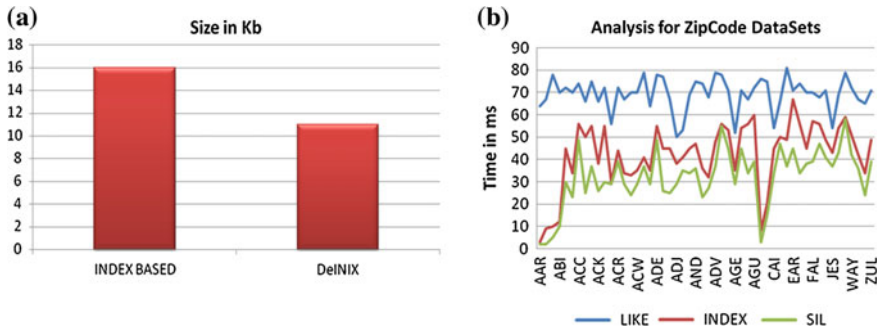


Fig. 4 **a** Storage size of INDEX based and DeINIX method. **b** Analysis for zipcode datasets for different prefix word

6 Empirical Evaluations

We have implemented the existing and propose work on real datasets. The data set is taken on Usa Zipcode.

We have implemented the techniques in Java for the proposed technique of keyword searching in Linked databases and DeINIX for constructing Inverted Index. We use MYSQL for the database storage. We will compare LIKE Method, Index based search and SIL using DeINIX different technique.

The result analysis of the propose technique is shown in Fig. 4a, b with help of graph. We compare the result of LIKE method, Index-based method with prefix table and Searching in Linked Databases with DeINIX for keyword search for USA ZIPCODE dataset. The empirical evaluation for dataset can be seen in graph. The graph clearly shows that LIKE method takes more time compared to Inverted Index Table, as the LIKE search keyword from scratch every time. All this searching technique works for single database. The new technique proposed works on Linked databases and the indexing method propose saves memory space too as shown in Fig. 4a. The comparison of 3 techniques with USA ZIPCODE dataset is shown in Fig. 4b from that we can conclude overall that LIKE take more time compared to 3 techniques and our propose technique take least time. The time here displayed is in milliseconds.

7 Conclusion and Future Work

The keyword search over relational database works on the single database, we have proposed an algorithm SIL using DeINIX which answers the keyword queries in Linked databases with linking the two databases via common foreign key and DeINIX is a new inverted index technique which reduces the memory space. Empirical results is taken on three different dataset and lager number to records and

it shows that the proposed algorithm work efficiently compare to the existing work in its scope. Also the result states that the Like method takes more query execution time compared to all the three techniques. In existing work the inverted index table created was static in our proposed work it's not so, with change in database we can update it dynamically. The SIL with DeINIX answer keyword queries more efficiently.

In Future work one can focus on Fuzzy Keyword search and the ranking techniques in Linked databases which can search more efficiently. Even one can focus on Heterogeneous Databases by applying same exact keyword query search.

References

1. Bhalotia, G., Hulgeri, A., Nakhe, C., Chakrabarti, S., Sudarshan, S.: Keyword searching and browsing in databases using BANKS. In: IEEE (2002)
2. Agrawal, S., Chaudhuri, S., Das, G.: DBXplorer: a system for keyword-based search over relational data bases. In: ICDE (2002)
3. Hristidis, V., Papakonstantinou, Y.: Discover: keyword search in relational databases. In: VLDB (2002)
4. Bergamaschi, S., Guerra, F., Simonini, G.: Keyword Search Over Relational Databases: Issues, Approaches and Open Challenges. Springer (2014)
5. Li, G., Feng, F., Li, C.: Supporting search-as-you-type using sql in databases. In: IEEE (2013)

Composite Analysis of Different Jammers in Cognitive Radio Networks Under Denial of Service Attack

Sushant Singh and Akwinder Kaur

Abstract Cognitive Radios are intelligent radios which possess awareness of their surroundings and bandwidth availability in the existing network. Parameters such as location, nearby transmitters, time of day are some special functionalities that Cognitive Radios dynamically tune for the available spectrum. Cognitive Radios use intelligence to discover if any part of the available spectrum is utilized or not. In wireless networks Denial of Service attacks, such as jamming causes significant performance degradation to the existing network and thus needs to be detected at an early stage. This becomes more important in a Cognitive wireless network employing Dynamic Spectrum Access, where it is simpler for the attackers to dispatch Denial of Service attacks. The proposed work shows the detection of Denial of Service attacks as an early detection mechanism that detects the abrupt changes in some parameters which result in comparatively.

Keywords Cognitive radio networks · Deceptive jammer · Reactive jammer · Denial of service

1 Introduction

Cognitive Radio (CR) is an intelligent radio which has the ability to be configured and programmed dynamically [1, 2]. The transceiver changes its transmission or reception parameter according to the best available channel automatically detects free channels in wireless spectrum that allows high concurrent wireless communication to radio in a prescribed spectrum band at one location [3]. The general architecture of cognitive radio network (CRN) is given in Fig. 1, it shows the unlicensed user called as secondary user which uses the cognitive base station to make communication over the network. CR is the extension of Software Defined Radio in which components are implemented in hardware using software on an

S. Singh (✉) · A. Kaur
Department of CSE, Chandigarh University, Mohali 140413, Punjab, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_35

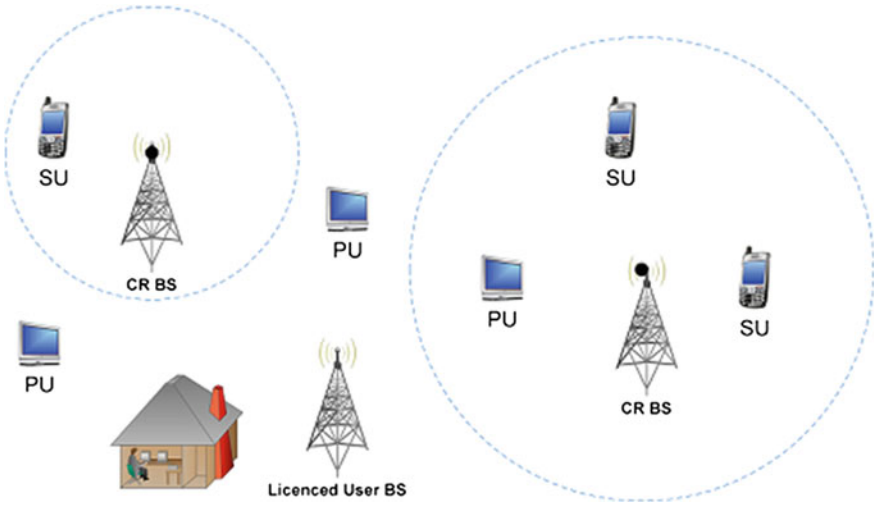


Fig. 1 Architecture of CRN

embedded system. Full CR is when all the parameters are considered that a wireless network can be aware of whether spectrum sensing CR is which detects the opportunity in the available spectrum.

CR learns from its environment and performs dynamic activities. So CR can also learn things by malicious elements of their environment hence perform some abnormal behaviour which affects the performance of primary users [4]. To prevent CRs from these type of activities there are some security issues [5] which need to be considered to protect them from various attacks.

The cognitive agent is able of configuring radio system criterion in acknowledgment to the commands accustomed by the operator. These criterions of parameters cover waveform, protocol, operating frequency, and networking. This action is a free assemblage in the communications ambiance which exchanges the advice about the ambiance with the arrangement it accesses and added CRs. From reading the outputs generated by the radios a CR monitors his act moderately; it again uses this data to actuate the channel conditions, Radio Frequency environment, link performance etc. and accommodate the radio settings to bear the adapted quality of service accountable to an adapted aggregate of user requirements, viable constraint, and authoritative constraints.

CRNs are meant to provide Internet access to under-served areas called Wireless Regional Area Networks by opportunistically getting the analog TV bands made available by FCC for unlicensed used. In CRN, devices are needed to sense the spectrum occasionally and leave the spectrum band if any Primary User wants to communicate. CRN performs a two-stage spectrum sensing method: namely a method called fast sensing and fine sensing. Fast sensing usually takes a few microseconds and uses simple ways of doing things such as energy detection, and

therefore can only report the existence of a signal on the available spectrum. On the other hand fine sensing employs fancy or smart ways of doing things for identification of signal types on the spectrum and may take up to 160 ms in other words the entire length of time of a super frame also called the Channel Detection Time (CDT). This two stage method is meant to strike a balance between the disagreeing goals of proper protection of PU's signals and best QoS for CRN's SUs.

In wireless network several attacks are exists, from which considerable attacks can be classified. Which go through four atop layer. These layers are physical layer, link layer, network layer and transport layer. The three main attacks in the physical layer are PUE attack, Objective function and jamming attacks. Attacks in data link layer consist of Spectrum sensing data falsification (SSDF) attack, control channel saturation dos (CCSD) attack, and selfish channel negotiation (SCN) attack. Important attacks in network layer are the sinkhole attack, Routing attack and HELLO Flood attack. Transport Layer concentrates on the Lion Attack. Jamming attack may be done on physical layer or MAC layer. As an example, the attackers may put on a show to be an authorized or licensed primary user, and the Primary User Emulation attacks can be implemented. The attackers can conduct smart jamming and may also explore the spectrum themselves. Deceptive and reactive jammers are two types of jammers which can be used by attackers to perform jamming attacks. These attacks usually appear at anonymous time and are capricious due to the abridgement of above-mentioned ability of the attackers.

Jamming attack [6] is defined as an attack in which the attacker forwards packets to thwart sending or receiving data for accepted user associated in a transmission session that is called a Denial of Service (DoS) situation. There are four types of jammers: constant Jammer, deceptive Jammer, random Jammer, and reactive Jammer.

The deceptive jammer attacks the legitimate users in the network by sending out packets continuously then other users can only receive the signals. In this state they can only identify a continuous stream of coming packets. The deceptive jammer attempts start a new transmission immediately followed by the previous transmission that make the channel busy and no one else can use the channel and since the packets not detected as malicious all the packets are successfully received and there is no packet errors. The reactive jammers are harder to detect because they do not transmit all the time. They only transmit the jamming signals whenever it senses communication within the channel. The constant jammer forwards the packets of abstracts steadily after any attention to MAC-layer protocols. It doesn't wait for the channel to be idle. The random jammer hold gaps between the signals of jamming, and when performing the attack that is called jamming state it may react as constant or deceptive jammer. It takes some time off to reserve energy in case the jammer doesn't have unlimited power supply.

This paper is divided into the various sections: Sect. 2 describes the related work in the area of the CRNs and Sect. 3 presents the proposed system to simulate and execute the proposed scenario. The simulation results are discussed in Sect. 4. Further conclusion and the future scope are stated.

2 Related Work

In this study the analysis of different types of jammer has been performed for CRNs to evaluate the consequence of DoS attack. In practical jamming attacks are transmitting unwanted signals on the wireless channels which interfere in radio frequency without following the proper standard of Medium Access Control. Such attacks make able to block the transmission of data across the channel and impact the quality of the exist network some of the previous related work with CR are given.

Amjad et al. [7] proposed that IEEE 802.22 CRNs employ a two-stage quiet period mechanism based on a mandatory Fast Sensing and an optional Fine Sensing stage for Dynamic Spectrum Access during every super frame. In this 'DS3' A Dynamic and Smart Spectrum Sensing algorithm, this minimizes the effects of jamming as well as noise on the fast sensing phase of DSA. It improves system's efficiency by striking a balance between spectrum utilization by SUs and delay in the detection of Primary Users' presence on the spectrum, using a dynamic fine sensing decision algorithm with minimal overhead.

Akin et al. [8] proposed the information theoretic security by modelling a CRs wiretap channel under QoS constraints and interference power limitations inflicted on PUs. We initially define four different transmission scenarios regarding channel sensing results and their correctness. Author has provided effective secure transmission, obtain the effective secure capacity which describes the maximum constant buffer arrival rate under given QoS constraints. Show the performance levels and gains obtained under different channel conditions and scenarios. And emphasize the significant effect of hidden-terminal problem on information-theoretic security in CRs.

Wang et al. [9] proposed that in CRNs, CR's nodes adaptively access the spectrum aiming to maximize the utilization of the scarce resource. Crucial to the successful deployment of CRNs, security issues have begun to receive research interests recently. This article surveys the research advances in CRNs. First, the fundamentals of CRNs including the basic components. OSA (Opportunistic spectrum Access), and the inherent issues in CRNs are reviewed. Then they present the issues in security for CR.

Weifang [10] proposed Security in communication is the success key. In wireless network DoS attack is one of the severe attacks. Due to their own characteristics CRNs are prone to DoS attack. And author analysed the structure of CRN. And the possible different DoS attacks in ad hoc CRNs in distinct protocol layers also discussed in the paper.

Balogun and Krings [11] investigated through simulation the performance of CRNs operating in the presence of jamming attacks that are capable of introducing value faults. The purpose of this investigation was to be able to come up with a suitable anti-jamming strategy that is capable of mitigating jamming in CRN. Using Packet Delivery Ratio (PDR) measurements it was observed that these jammers are very effective in their operations as they are capable of bringing down the entire

CRN when their jamming rate is just about 30 %. As a result, author investigated different anti-jamming strategies and analysed their suitability in combating CR operating as jammers with the intent of proposing a solution that is capable of handling all scenarios presented in these models.

Kim et al. [12] developed a mechanism to estimate SINR under reactive jamming. The estimated SINR information of each wireless link was then used to determine the jamming-averse directivity (JAD) of packets, which improved the routing performance of the victim network and validate the proposed mechanism with a simulation study, showed that the proposed JAD escorted (JADE) routing dramatically improved routing path discovery performance including path discovery probability, path length, and elapsed time for path discovery, retransmission attempts, and path quality under reactive jamming. Among the 200 route requests at 10 different configurations in their simulation, the reactive jammer disrupts the 77.5 % of total requests. However, JADE routing decreases the route discovery failure rate to 7.5 % by saving the 96.7 % of failed requests.

Wang et al. [13] investigated the security method in which Jamming attack have faced by secondary users, and in a cognitive radio proposed a framework for anti-jamming defense called as stochastic game framework. The SUs changed their methodology on the most proficient method to save and switch in the middle of control and information channels, per their instance of observe concerning the spectrum handiness, quality of channels and also the movement of attackers. And additionally outlined the minimax-Q learning, By minimax-Q learning, SUs will slowly grasp the most effective strategy, which makes as big as possible and the myopic learning policy compared with it that solely make the payoff maximum at every phase and the environment dynamics and also the attackers' cognitive capability, and a random defence policy are not considered.

Balogun and Krings [14] investigated the performance of CSS CRN operating in the existence of CR constant jammers that are capable of introducing transmissive and omissive value faults. They discovered that existing anti-jamming solutions don't seem to be ready to handle these fault varieties and consequently projected a hybrid FEC code capable of mitigating the complete category of faults known during this adversarial model. They have simulated and investigated the performance of the proposed solution in NS-2. The result of the analysis showed that proposed solution was efficient and robust against any level of DoS caused by constant jammers in a CSS CR network.

Marttinen et al. [15] proposed techniques for combating source selective jamming attacks in tactical cognitive MANETs. For facilitating tactical operations Secure, reliable and seamless communications are important. A significant security threat to the operations of wireless tactical MANETs created by Selective jamming attacks since selective methods possess the potential to fully isolate some of the network from different nodes without giving a transparent indication of the issue. The proposed mitigation techniques used the concept of address manipulation that dissent from other techniques presented in open literature since the techniques

employ de-central architecture instead of a centralized framework and therefore proposed techniques don't need any further overhead. Experimental results showed that the proposed techniques enable communications in the source selective Jamming attacks presence. Once the presence of a source selective jammer blocks whole transmissions, implementing a proposed flipped address mechanism would increases the expected range of needed transmission attempts solely by one in such simulation model.

3 Proposed System

In the proposed system impact of jamming attack/DoS attack is studied. If the slot of primary user is free, the slot is known as opportunity slot in which malicious user sends the jam over the network. In the proposed system routing protocol in CR will introduce a method which decides whether the node is primary and secondary and the dynamic sensing is fine or not. CRN scenario is created using the CRCN architecture to implement the Cognitive phase with the help of CRCN using appropriate routing scheme it make the network with the available Cognitive functionality. Then after creating the network sensing will be done. This sensing is called fine sensing in which sensing has been performed based on few parameters and it makes the detection of opportunistic band which is also known as free band of primary user and sensing in which simple parameters are used to detect the white space in the spectrum i.e. based on energy detection called fast sensing it takes few milliseconds. DoS attack is transmitted by malicious user during free slot of primary user by sending short jam signal over the network. To overcome this issues a proper utilization of free bands is done. The analysis between reactive jammer and deceptive jammer has been performed. Parameters under consideration are given in Table 1.

Table 1 Simulation parameters

Parameters	Values
Channel required	Wireless channel
Protocol	DSR
Number of nodes	10
Window size	32 bits
Packet size	512 bits
Maximum packet in interface queue	500
Transmission time	20 s
Simulation time	60 s
Random motion	250 s
Number of channels in cognitive radio	2
Number of interfaces in cognitive radio	2

4 Experimental Analysis

Various scenarios have been developed using cognitive radio bands. In this spectrum sensing is done by using the Dynamic Spectrum Access approach. The bands have been used for spectrum sensing, the primary and secondary users use these spectrum bands. The primary user is authenticating user for band usage and secondary user use free bands for spectrum sensing. After locating the nodes divide them into the users of two types. First is the primary user and the other is secondary user. The difference between these users is that the primary users have priorities to do the task firstly after that secondary users can do their tasks. Primary and secondary users have been used for the data transmission through spectrum sensing. If the band is free it checks the authentication user then it will allocate this band for data transfer. The secondary users perform the spectrum sensing using Dynamic Spectrum Access for checking of available free spectrum bands. That is also known as opportunity band. In the opportunity band when the secondary user starts the transmission of data from one point to another then the user transmits the short jamming signal, this signal has been transported over the network using the different fast sensing signals. This puts up extra load on the network and causes denial of service attack over the network. This in turn leads to loss of data due to the occurrence of the attack on the primary user so it starts loose data and transmit the signal from wrong spectrum band and that causes excess traffic over the network. Moreover in this, two type of jamming i.e. Deceptive jamming and Reactive jamming occur. The deceptive jammer tricks the legitimate users. It sends out packets continuously making the other users switch into a receive state and remain as they detect a continuous stream of receiving data packets. The reactive Jammer senses the channel at all times, and whenever it senses communication in the channel it starts transmitting the jamming signals. This jammer is harder to detect because it's not transmitting the data packets all the time.

In Fig. 2 Packet Loss is represented. Number of packets lost when packets are transmitted over the network is called packet loss. Grey line in this figure is used to represent Loss without attack and red line is used to represent the Loss with attack that is performed by deceptive jammer. Blue line shows the huge loss that results in the network for packet Loss and is performed due to reactive jammer.

In Fig. 3 the Throughput has been depicted. The number of packets delivered successfully over the network called throughput. Dark red line is used to represent the Throughput with attack performed by deceptive jammer. Grey line shows the throughput whenever the DoS attack has been performed by a reactive jammer.

In Fig. 4 is used to represent the Packet Delivery Ratio. Packet Delivery Ratio is derived as the number of packets delivered with respect to time. Dark line is use to represent the Packet Delivery Ratio with attack performed by deceptive jammer. Grey line shows the Packet Delivery Ratio that is loss in the sent packet due to reactive jammer.

The simulation result shows that the performance degrades as the values shown above in the Table 2 with respected parameters of evaluations.

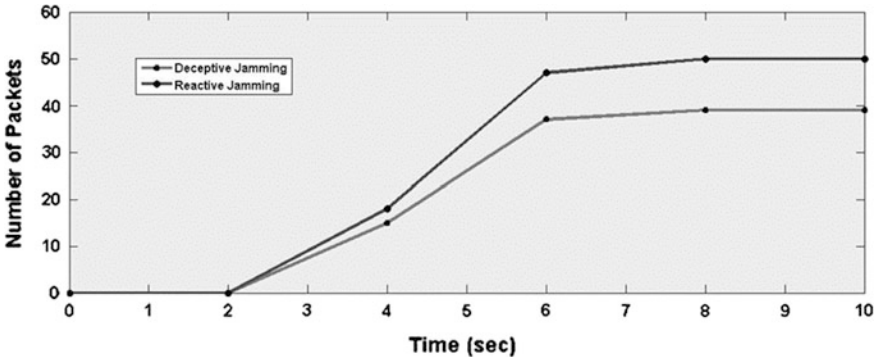


Fig. 2 Packet loss

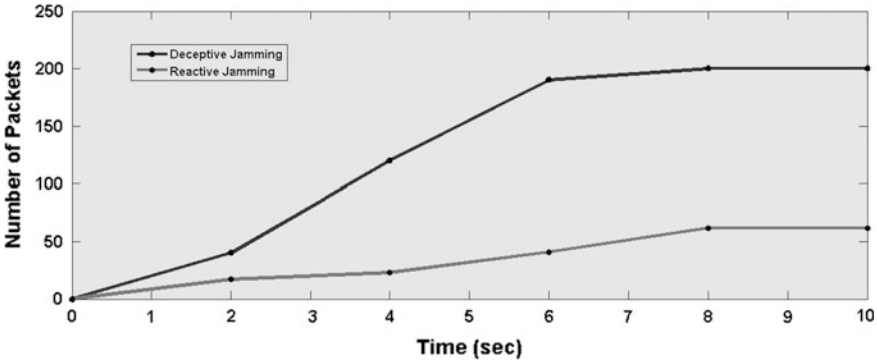


Fig. 3 Throughput

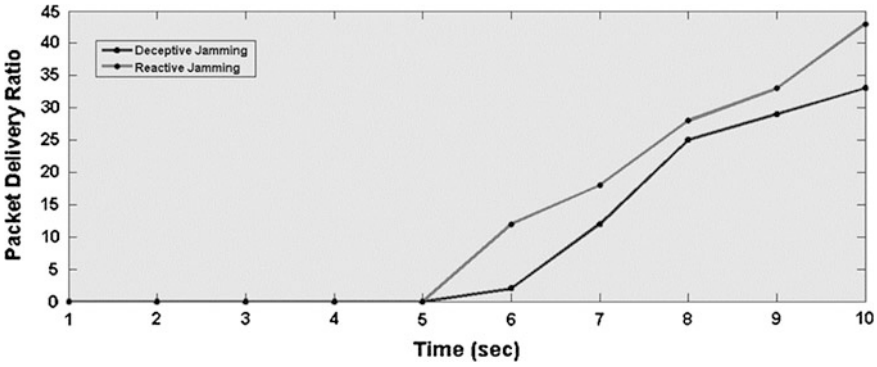


Fig. 4

Table 2 Simulation results

Parameters	Deceptive	Reactive
Generate packets	7428	4300
Receive packets	4061	857
Packet delivery ratio	54.6715	19.7907
Dropped packets	3367	3449
End-to-end delay	18.5603	22.2328

5 Conclusion

In this work simulation of CRN has been done with comparison of two different jammers, Reactive jammer and Deceptive jammer and evaluate the impact of both type of jammer in the cognitive radio networks. Jamming in the CRN has main impact on the service of the electronic network that may cause the available table service to primary exploiter and directly cause harm to the line utilization by protecting the network from jamming plan of attack which makes the network reliable. This work shows the impact of both reactive jammer and deceptive jammer in cognitive radio network which affects the throughput of the channel and leads to the quality degradation. Jamming attack may be detected using different techniques. This work can be further explored by preventing the network from such type of jammer.

References

1. Wang, B., Ray Liu, K.J.: Advances in cognitive radio networks: a survey. IEEE J. Sel. Top. Sign. Proces. **5**(1), 5–23 (2011)
2. Zeng, Y., Liang, Y.C., Hoang, A.T., Zhang, R.: A review on spectrum sensing for cognitive radio: challenges and solutions. EURASIP J. Adv. Sign. Proces. **2**, (2010)
3. Akyildiz, I.F., Lee, W.Y., Vuran, M.C., Mohanty, S.: A survey on spectrum management in cognitive radio networks. IEEE Commun. Mag. **46**(4), 40–48 (2008)
4. Chen, R., Park, J.-M., Reed, J.H.: Defense against primary user emulation attacks in cognitive radio networks. IEEE J. Sel. Areas Commun. **26**(1), 25–37 (2008)
5. Mathur, C.N., Subbalakshmi, K.P.: Security issues in cognitive radio networks. In: Cognitive Networks: Towards Self-aware Networks, pp. 284–293 (2007)
6. Li, H., Han, Z.: Dogfight in spectrum: jamming and anti-jamming in multichannel cognitive radio systems. In: Global Telecommunications Conference, GLOBECOM, IEEE (2009)
7. Amjad, M.F., Aslam, B., Zou, C.C.: DS3: a dynamic and smart spectrum sensing technique for cognitive radio networks under denial of service attack. In: Global Communications Conference (GLOBECOM), IEEE. IEEE (2013)
8. Akin, S.: Security in cognitive radio networks. In: Information Sciences and Systems (CISS), 48th Annual Conference on. IEEE (2014)
9. Wang, F.: Cognitive Radio Networks and Security: A Survey, pp. 1–5. IEEE (2014). ISSN:978-1-4799-2321-2
10. Weifang, W.: Denial of service attacks in cognitive radio networks. In: Environmental Science and Information Application Technology (ESIAT), pp. 530–533 (2010)

11. Balogun, V., Krings, A.: An empirical measurement of jamming attacks in CSS cognitive radio networks. In: Electrical and Computer Engineering (CCECE), IEEE 27th Canadian Conference on. IEEE (2014)
12. Kim, Y.S., DeBruhl, B., Tague, P.: JADE: jamming-averse routing on cognitive radio mesh networks. In: Communications and Network Security (CNS), IEEE Conference on. IEEE (2014)
13. Wang, B., Wu, Y., Liu, K.R., Clancy, T.C.: An anti-jamming stochastic game for cognitive radio networks. IEEE J. Sel. Areas Commun. **29**(4), 877–889
14. Balogun, V., Krings, A.: Mitigating constant jamming in cognitive radio networks using hybrid FEC code. In: Advanced Information Networking and Applications (AINA), IEEE 28th International Conference on, pp. 704–711 (2014)
15. Marttinen, A., Wyglinski, A.M., Jantti, R.: Moving-target defense mechanisms against source-selective jamming attacks in tactical cognitive radio MANETs. In: IEEE Conference on, pp. 14–20. Communications and Network Security (CNS) (2014)

IDLS: Framework for Interactive Digital Learning System with Efficient Component Modelling

D. Pratiba and G. Shobha

Abstract With the increasing demands of educational knowledge, adoption of modern learning management system is found to gain a faster pace in modern educational system. However, it is also found that existing mechanism of digital library system pertaining to educational domain suffers from significant pitfalls that are detrimental for quality education system. Hence, the proposed paper introduces a framework termed as Interactive Digital Learning System (IDLS) which potentially addresses the flaws in existing Massive Open Online Courses (MOOC) by mechanizing a novel and simple component modelling schema. The paper discusses the design principles of IDLS and compared its outcome with existing MOOC, where the proposed IDLS is found to excel better compatibility with cloud based learning management system with respect to assessment quality, assessment technique, contents, interaction, feedback, IP security, and features.

Keywords Component modelling · Digital library · Learning management system · Online learning

1 Introduction

In the current and future era of digitization, there is no functional domain of the human society which is left untouched to be digital. Educational sector also is being recommended by various authorities and standard setting bodies to be completely digital so that, education becomes as a global service in true sense. Every process involved into educational system will be an e-process. The result of such system will be exponential generation of structured, semi-structured and unstructured data including databases, web-logs, xml, text, images, video etc. The unevenness into the rate of generation of data and availability of the storage has leaded every

D. Pratiba (✉) · G. Shobha
Department of Computer Science & Engineering, RVCE, VTU,
Bangalore, Karnataka, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_36

organization (educational) to setup their facilities on cloud infrastructure. In general the formulation of educational digital libraries requires referencing of design theory from the view point of (i) Reference tasks, (ii) Design experiments, and (iii) Design genres. These approaches support to visualize the design constructs for building knowledge repository and sharing knowledge. The purpose of any knowledge repository and sharing platform is to construct finer grained models, which can correlate key design features, cognitive as well as the social dimensions of the context of use. Summer et al. [1] have described, three different model in their work titled "Digital libraries and educational practice: a case for new models" namely Digital library (DL) as a cognitive tool, Component repositories and Knowledge networks and analyzed its utility into educational digital library. There has been an extensive review work on the usability and impact of digital libraries and its design/methodology/approach [2].

Approaches of an open access digital libraries is widely adopted in the recent years as to have a global connectivity of knowledge sharing mechanism among the publishing, economic and scientometrics dimensions. This kind of open access libraries provides easiness of use and learns ability, as well as functionalities commonly met in these systems, affects user interaction and satisfaction [3]. The wide collaboration of institution and globalization of education system require developing a holistic model for various levels of digital library. Zhang [4] has evaluated important criteria from heterogeneous stakeholder groups and developed a model using a three-stage research approach: exploration, confirmation, and verification. A holistic DL evaluation model was constructed using statistical techniques. Eventually, the verification stage was devised to test the reliability of the model in the context of searching and evaluating an operational DL. The proposed model fills two lacunae in the DL domain: (a) the lack of a comprehensive and flexible framework to guide and benchmark evaluations, and (b) the uncertainty about what divergence exists among heterogeneous DL stakeholders, including general users. The digital library objective are not limited just to have a replacement of physical to digital storage, whereas it's optimal objective should be to achieve knowledge extraction and sharing and building application using the extracted knowledge of the reposits. Information processing paradigm which is still fundamental to digital library engineering is not adequate for the way in which information and informational objects—as stored in Digital Library systems is the core fundamental topic to be discussed and researched. Hobohm [5] has studied the information processing paradigm as Shannon/Weaver information theory concept and provides an insight that how Digital Libraries could be conceptualized in a way that has not yet been implemented in standardized formal frameworks such as the DELOS or the 5S-Framework [5].

This paper presents a framework for the purpose of addressing the pitfalls of existing Massive Open Online Courses and illustrates the significant advantages from its outcomes that will pave a significant guideline to futuristic learning management system. Section 2 discusses about the prior research work followed by discussion of drawbacks of existing system in Sect. 3. Section 4 introduces the

proposed IDLS framework. Research methodology is discussed in Sect. 5. Section 6 discusses about the performance analysis of the proposed IDLS with existing MOOC. Finally, Sect. 7 summarizes the paper.

2 Related Work

Various literatures have been proposed for the purpose of designing a digital library considering a specific language set. One of such studies has been proposed by Huang and Wu [6] where the authors have focused on Chinese based digital library system for generating corpora, dictionary and speech database. The study mainly focuses on implementing online e-Learning system for spoken Chinese based on the technology of natural language processing. However, the study didn't furnish substantial evidence of benchmarking about the effectiveness of their framework with the growing data management. The study also didn't address any measures of mitigating unstructured data, which makes the framework not much applicable while working on Big Data. Similar studies on same direction in Arabic language were presented by Abdelkader et al. [7].

An enhanced study for digital library design considering sign languages was also attempted by various authors. One of such study was discussed by Prinetto et al. [8], where the large parallel corpus between Italian and Italian Sign Language were presented. The author has introduced the Lexicon Video sign bank, a new and expanding public dataset that contains high-quality animated video sequences of hundreds of distinct signs. However, such framework is again associated with scalability issues are the dataset which the authors have used is composed of very limited corpus. It was also found that studies are quite limited to specific corpus. The consideration of e-learning on multimedia environment was also seen in the work of Haipeng and Lijing [9] who discussed the computational application of vocabulary learning, especially the application of corpus database. However, the scalability issues remain the same just like the work of [7, 8].

The next phase of the research work has seen the evolution of personalized recommendation model to optimize the effective learning environment for the learner. One of such significant work was carried out by Nagori and Aghila [10]. The study has presented a personalized integrated model for e-learning systems that consists of using Latent Dirichlet Allocation topic modeling technique to make topic analysis on corpus and introducing a similarity measurement to content-based recommendation approach. However, the study lacks efficient benchmarking and encounters the similar issues as in [7].

In order to overcome the above mentioned issues, literatures have also witnessed certain work that increases the contribution of learning based digital library system. Works extracting the existing requirements of users and providing rapid response was new trend in research on digital library system. One of such work was seen in the literature of Egusquiza and Izkara [11] who have provided the infrastructure for the development of short-term collaborative spaces that require a fast response to

emerging needs of a business opportunity. However, the study framework is based on offline data and undergoes same scalability issues if the framework is subjected to massively growing and unstructured data.

The digital library system focusing on e-learning architecture has witnessed multiple contributions from research community for adding various sorts of enhancements mainly to design better corpus based approaches. Studies in similar direction for enhancing the learning experience in existing digital library system was presented by Li and Hu [12], Alqadoumi [13], and Zhang et al. [14]. Such studies have however stick to a more effective service discovery and utility of teaching and learning in the area of English as a second language, especially as a professional foreign language of Chinese and Arabic students mainly. The studies [13] have also identified that building local corpora can be time and effort consuming because of the complex processes involved in selecting texts and tagging them. Another issues in this regard is that corpus linguistics cannot stand alone as a teaching method because it only covers linguistic and verbal features of language and does not, for example, cover the nonverbal side of language [13, 14, 16]. The most recently observed trend of building e-learning based digital library system is to use web service discovery with effective corpus design for the purpose of enhancing e-learning task. The recent study presented by Zu and Zhang [15] has put forward such evidence. The authors build a platform for dissertation academic information extraction, the key to obtain various literature information, can form updating database retrieval, convenient for user to quickly and widely understand scientific researcher grasp its research in the field of the latest knowledge and information.

Hence, it can be seen that majority of the studies carried out in formulating digital library on e-learning system are focused on corpus design with large adoption of linguistic based approaches using static or offline data. None of the studies were found to consider massively growing data, which are highly unstructured posing a significant challenge against all the prior frameworks being proposed. Another significant research gap explored is none of the studies focus on potential security aspects on the concepts being presented using exclusive cloud environment. Hence, a significant research gap is found in the literatures pertaining to design of digital library system which doesn't integrate security measures along with consideration of unstructured massive streams of data, and lacks scalability with no potential benchmarking.

3 Drawback of Existing LMS

Among learning methodologies existing today, online learning offers a significant focal points over conventional/ordinary classroom-based learning. Its most eminent preferences are diminished expenses, as a physical environment is no more obliged and it can subsequently be utilized whenever and spot at the comfort of the understudy. Also the quantity of scholars that can take after the class is not restricted by the site of a physical classroom. Moreover, the learning material is not

difficult to keep overhauled and the educator might likewise consolidate sight and sound substance to give an inviting system and to encourage the understanding of the ideas. At last, it can be seen as a learner-focused methodology which may address the distinctions among instructors, so that every one of them may contrast their material in place with assess and re-use normal regions of information. Then again, there are a few detriments that must be tended to before the full incorporation of e-learning into the scholarly schema. At present, e-learning frameworks are frail on adaptability at the base level. A few resources must be sent and allotted for particular job with the goal that when accepting high workloads, the framework needs to include and arrange new resources of the same sort, making asset procurement and administration extremely lavish.

This prominent problem is additionally identified with the effective usage of these resources. Case in point, in an average college situation, PC research center resources e.g. (programming, equipment) that can suit all the scholars in one educational stream, might in some cases stay under-used because of low request in different educational streams. An alternate hitch is dynamic rates of utilization: there is an appeal for utilization of these resources towards the end of a semester, while it is little toward the start of the semester and essentially non-existent amid breaks. In this way, the physical machines are held actually when they are idle, wasting their maximum capacity. The computing foundation of Learning Management System (LMS) does incorporate machines (equipment) as well as common framework, i.e., structures, climatisation, security gadgets, et cetera. Also, there is likewise an expanded expense in regards to the product permitting with comparing expenses of establishment, upkeep and specialized backing.

4 Proposed System

The globalization has made entire world as a village and it has been predicted that there will be massive opportunities for online digital learning systems. The design should aim for large-scale participation with every possible scope of security aspects along with provisioning of experimenting analytics on the stored data which are used for learning. In our previous work [17], a robust security mechanism for the authentication for the legitimate user has been proposed. The current paper discusses the design approach of an Interactive Digital Learning System (IDLS) keeping a future integration of analytics and mining approach on the data, which could be of the nature of Big Data and the entire framework should be synchronal to the cloud architecture.

The proposed IDLS is designed using component modelling. The features of the IDLS are compared characteristically with existing other IDLS or MOOC, with reference of Authentication, Interaction and Mining Approach to make it interactive to improvise teaching and learning mechanism. The proposed IDLS is designed using multiple components for ensuring the optimal archival of heterogeneous and massive data pertaining to Learning Management System (LMS) over pervasive

and distributed environment. The system considers developing an interface specific to the adopted digital interactive learning system, where it consist of 3 important components e.g. (i) scholar component, (ii) coach component, and (iii) supervisory component. A generalized interface of proposed IDLS is designed which will be accessed using all the three prime components.

With an aid of Fig. 1, the proposed design methodology for IDLS is illustrated. The discussions of the design pattern of the individual components are as follows:

4.1 Scholar Component Modelling (SCM)

A component is designed for meeting the educational knowledge capturing requirements of the scholar in our proposed IDLS termed as Scholar Component Modelling (SCM) and the schema for Scholar Component Modelling is illustrated in Fig. 2. The design of the SCM is done is such a way that it facilities certain advance interactive privilege compatible with cloud environment. The initiation of the SCM is done with enrolment of the scholar, where the private details of the scholar are captured by the sub-components. All the private details are then fetched to IDLS which resides in cloud server. Various string-based attributes like name, mobile phone number, contact details, and email-ids of the scholar are considered

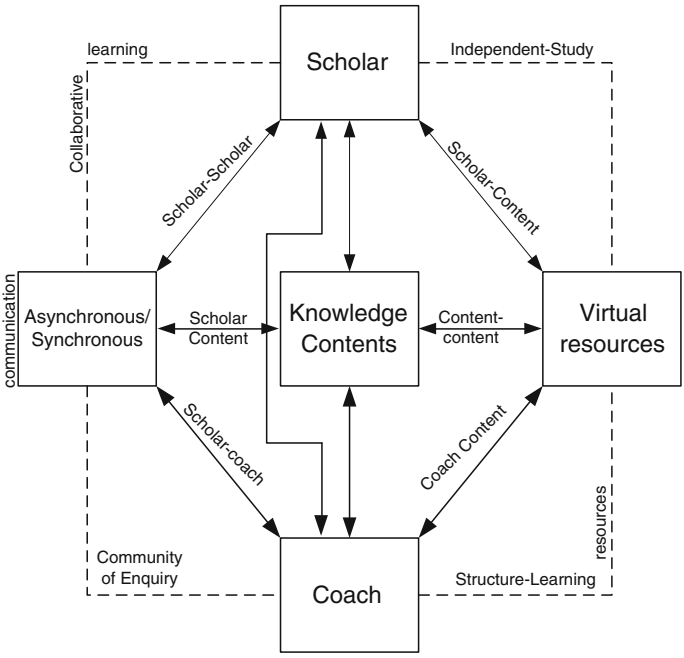


Fig. 1 Proposed schema for IDLS

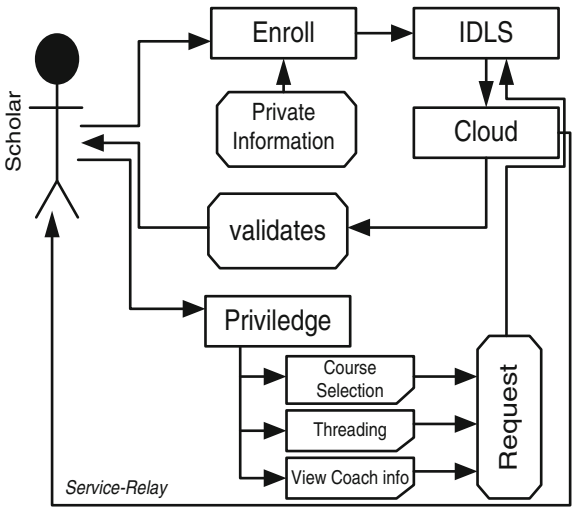


Fig. 2 Schema for scholar component modelling (SCM)

for the purpose of enrolment phase. The system uses My-SQL to store and access such information. We apply the S-DILS algorithm to ensure fail-proof authentication of legitimate user in proposed IDLS.

After the successful validation of the legitimate user, the phase of the component design permits to provide the account related privilege to the scholar. Using this privilege, the scholar will be able to select their course from the list of available courses. Hence, various sub-components were built e.g. course selection, threading mechanism, and visualizing couch details. Based on the selection of the particular component, an event is generated that is specific to particular sub-components. For an example, if the scholar chooses to opt for course materials, the request is sent to the proposed IDLS, which in turn relay the specific services based on the request. The next part of the component design is on threading mechanism. As the proposed IDLS focuses on interactive system, hence, it is felt that in proposed system (i) one scholar can either comment on particular thread, (ii) scholar can even comment on some other scholar's thread in the same interface of proposed IDLS, and (iii) scholar can also create their own thread which are expected to be answered by the other users in the same community. Finally, student can visualize the necessary details about their coach which will assist them to get aware with the professional skills, experience, and qualification of a coach, the type of the courses to be carried out by them, scheduling details etc.

4.2 Coach Component Modelling (CCM)

The modelling of the coach actor is one of the critical components modelling phase in proposed IDLS. Basically, it is meant for an individual like coach in the proposed

IDLS who impart, share, and appraise scholars using the latest ICT approaches. The CCM is designed in such a ways that even a coach will also required to get themselves registered in the proposed application of digital interaction learning system in the same way just like the scholar module.

The important privilege of coach is that they can create their own syllabus and necessary contents required to be pursued by the scholar. Hence, a coach has the rights to upload syllabus or course materials in the form of Microsoft Word, Excel, PowerPoint presentation, PDF, etc. The system also provides flexibility for the coach to upload Frequently Asked Question (FAQ). However, there are certain operations/priviledge of the instructors that are controlled by Administrator e.g. extent of files to be uploaded, type of file formats to be uploaded etc.

4.3 Administrative Component Modelling (ACM)

The proposed system design have both coach and administrator component with higher degree of inter-relationship. The sole motive of the ACM is to control and supervise the entire operations as well as the ongoing process that are used by other components (SCM and CCM). As the entire system runs on cloud environment and posse's higher degree of heterogeneity in the users as well as the resources, hence it is quite imperative to implicate superior and skillful administration to the applications. To be very specific, the design of policy makers are endowed with exclusive control policies objects and attributes that will assist certain supervision over the other components (SCM and CCM) e.g. (i) ACM can set the privilege to the extent of permissible uploads and downloads of the user as transactional policies, (ii) ACM can set privilege to upload the specific types of file-formats in the various syllabus as well as course materials and all types of transactional files to be communicated between the SCM and CCM. Although there is no enrollment process being designed for the ACM but the user of ACM will also need to be authenticated them using their ID and static password. The integrated schema for CCM and ACM is illustrated in Fig. 3.

5 Research Methodology

The proposed paper introduces a framework called IDLS that furnishes the interactive platform for learning management system hosted over cloud. The prime purpose is to provide an enriched set of operational flexibility to capture big data so that performing mining over it becomes less complex computationally. In the present case scenario, the proposed system provides enough supportability to capture higher dimensionality of massive and heterogeneous data that can be termed as Big Data to be hosted in the cloud server, where learning and knowledge sharing is done online with self-controlled pace.

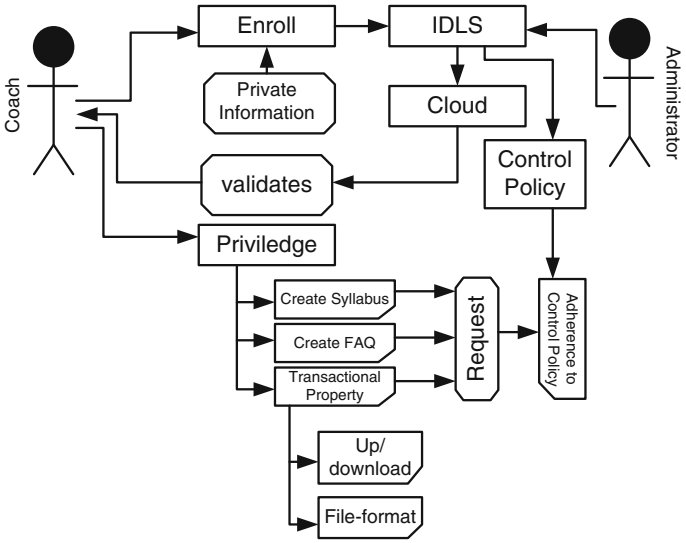


Fig. 3 Integrated schema for CCM and ACM

The IDLS framework considers a set of nodes that are located remotely and highly interconnected to each other. We follow a standard assumption of multi-relaying the online classes with massive number of scholars. As the proposed IDLS is more on component modelling so we give less emphasize on graphical user interface design, as we strongly believe that if the internals of the components are designed effectively than various upcoming applications as well as interfaces can be suitably integrated with the specific needs of any clients. The proposed IDLS system is also provided with multithreading mechanism that permits the user to generate or follow the existing thread or can create a new thread. This mechanism highly enriches the feedback-based mechanism that can be used for understanding or rate the quality of the interactions provided by the coach. The coach can perform interaction with every nodes and can also share multiple resources enough to generate massive heterogeneous data to be reposited on cloud server.

The proposed IDLS design not only supports the online interaction but also supports offline interactions. Scholars can review their request attributes as well as perform the events of transactional operation that allows them to archive various syllabus as well as study material uniquely meant for specific scholars. Figure 4 shows the operational design principle of IDLS where it can be seen that both archival as well as digital repository are utilized to enhance the quality of the modern educational system.

The flexibility in the design of IDLS is quite high that permits the user to use various tools and technologies in future to enhance the educational system. The coach can adopt various types of document resources like PDF, images, Microsoft office files, videos, and many more file format for sharing with scholar users. As the

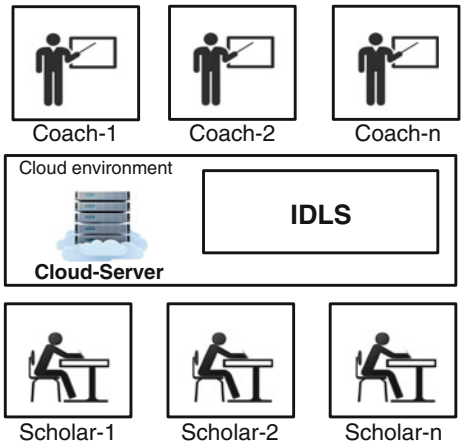


Fig. 4 IDLS framework principle

application is meant to be hosted over cloud platform, so storage and bandwidth usage is never a problem to affect knowledge sharing services.

The potential threading mechanism provided by IDLS system is one of the biggest boon for existing LMS. Various threads can be created, edited, followed by the legitimate users. For an effective LMS system, it is essential that feedback based mechanism is supported in the design principle. Hence, we choose to formulate the threads as the platform from which various forms of feedback mechanism can be supported in future while working on data mining principle. Various problems, issues, personal perceptions of the students can be shared on the thread in proposed IDLS that makes the system completely ready to be used.

6 Performance Analysis

The proposed system bears a strong resembles with Massive Open Online Course (MOOC). Hence, we choose to compare our IDLS framework with MOOC. The entire evaluation of performance comparative analysis was done based on inherent charecteristics of both IDLS and MOOC. Considered characteristics are e.g. (i) Assessment Method Quality, (ii) Assessment method Technique, (iii) Course Contents, (iv) Interaction with Coach, (v) Feedback, (vi) IP security, and (vii) Features. MOOC has various advantages e.g. (i) it provides easy access to global resources, (ii) enhances cross-cultural relationship, (iii) knowledge sharing in discussion forum, (iv) peer evaluation. However, our IDLS also provides all the similar functional advantages that are provided by MOOC, but the proposed IDLS system basically differs from MOOC as follows.

Table 1 highlights the prime difference in proposed IDLS and MOOC. The assessment quality of MOOC cannot be said to be an effective, as it is done using automated machine based. It is already known that machine based approaches uses

Table 1 Performance comparison between IDLS and MOOC

Parameters	MOOC	IDLS
Assessment method quality	Not effective	Highly effective
Assessment method technique	Automated machine-based assessment	Physically assessed by coach
Course contents	Highly simplified	Highly flexible
Interaction with coach	NIL	Highly interactive
Feedback	Doesn't support active feedback	Supports active feedback
IP security	Less	Very high
Features	Static	Highly flexible

data mining features as well as training algorithms to assess the online test in MOOC. It is also known that data mining in MOOC based application are quite NP-hard is the datasets are highly massive in volume. Multiple scholars with multiple subjects from different geographic region are something that makes the MOOC not ready to perform effective assessment. However, proposed IDLS system have the better assessment quality as compared to MOOC, as we are mainly focusing on component designing model, where the proposed IDLS is successful in generating a massive and heterogeneous archival of data, which is now ready to be investigated with novel data mining algorithms. In order to attract higher degree of audience, MOOC offers course contents with very simplified language. Although, it is good for average scholars, but certain scholars may think that MOOC only provides basic and elementary knowledge. However, in proposed system coach can built up any form of course materials based on suitability of the student's requirements. MOOC doesn't facilitate interaction with the coach, where proposed IDLS system does. Also MOOC doesn't support active feedback, whereas in our case the threading mechanism supports quality active feedback mechanism that can be further enhanced in future. Even from security viewpoint, MOOC doesn't offer enough security to its Intellectual Property (IP), whereas, we have already introduced a fail-proof and cost efficient authentication mechanism to provide standard security of IP compliant in cloud environment. The interface features in MOOC is only static, whereas IDLS offers enrich components to support various online upgrades to offer more functional features to enhance the quality of knowledge sharing process.

7 Conclusion

The present paper has introduced a framework called as IDLS that focuses mainly on archiving massive and heterogeneous data. The prime contribution of the study is its component modelling that allows the rich design principle for which reason performing data mining as well as various analysis on the archived data will become less computationally complex. We have tested this IDLS on various

real-time data and found that archiving the data as well as indexing the data doesn't have much adverse effect on the relayed services to the scholars. Moreover, we have also designed a threading mechanism that will assist to give better feedback mechanism for the online services being offered. Our future work will completely focus on performing analysis on data and hence come up with new data mining algorithm that can extracts knowledge on massive data in IDLS in a matter of seconds or minutes in maximum. Usually, performing data mining on Big data that resides in cloud is still an unsolved area in research community. It is said that extracting knowledge from Big Data takes a large amount of time. Therefore, our IDLS will solve the data archival process for massive and voluminous data over cloud. Hence, we strongly believe that our IDLS model will provide a robust test-bed to do so and thereby enhance the quality of educational system.

References

1. Sumner, T., Marlino, M.: Digital libraries and educational practice: a case for new models. In: Proceedings of the Joint ACM/IEEE Conference, pp. 170–178 (2004)
2. Chowdhury, S., Landoni, M., Gibb, F.: Usability and impact of digital libraries: a review. Online Inf. Rev. **30**(6), 656–680 (2006)
3. Tsakonas, G., Papatheodorou, C.: Exploring usefulness and usability in the evaluation of open access digital libraries. Inf. Proces. Manag. **44**(3), 1234–1250 (2008)
4. Zhang, Y.: Developing a holistic model for digital library evaluation. J. Am. Soc. Inform. Sci. Technol. **6**(1), 88–110 (2010)
5. Hobohm, H.C.: Can digital libraries generate knowledge? In: Historical Social Research/Historische Sozialforschung, pp. 218–229 (2012)
6. Huang, F.-L., Wu, J.-Y.: Study and implementation of intelligent e-learning system for modern spoken Chinese. In: Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, vol. 5, pp. 2968–2974 (2009)
7. Abdelkader, A., Boumiza, D.S., Braham, R.: An online arabic learning environment based on IMS-QTI. IEEE Int. Conf. Adva. Learn. Technol. **5**, 2968–2974 (2010)
8. Prinetto, P., Shoaib, U., Tiotto, G.: The Italian sign language sign bank: using word net for sign language corpus creation. In: International Conference on Communications and Information Technology (ICCIT), Aqaba, pp. 134–137 (2011)
9. Haipeng, H., Lijing, D.: English vocabulary e-learning on multimedia environment. In: Internet Technology and Applications (iTAP), International Conference, pp. 1–4 (2011)
10. Nagori, R., Aghila, G.: DA based integrated document recommendation model for e-learning systems. In: International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), pp. 230–233 (2011)
11. Egusquiza, A., Izkara, J.L.: H-KNOW: advanced infrastructure for knowledge based services for buildings restoring. In: 18th International Conference on Virtual Systems and Multimedia (VSMM), pp. 261–267 (2012)
12. Li, J., Ru, D.: AN engineering english learning program on a campus LAN. In: 8th International Conference on Computer Science & Education (2013)
13. Alqadoumi, O.M.: Using corpus linguistics as a tool for reform in english language teaching and learning. In: IEEE Fourth International Conference on e-Learning, pp. 246–252 (2013)
14. Zhang, J., Zeng, G., Zhang, J.: An online english e-learning framework based on web service discovery. In: 6th IEEE Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Vol.1, pp. 482–485, 2011

15. Xu, M., Wang, Y., Zhang, J.: The intelligent extraction of academic information based on web service discovery. In: IEEE 9th International Conference on Mobile Ad-hoc and Sensor Networks, pp. 399–404 (2013)
16. Masseroli, M., Marchente, M.: X-PAT: a multiplatform patient referral data management system for small healthcare institution requirements. IEEE Trans. Inf Technol. Biomed. **12**(4), 424–432 (2008)
17. Pratiba, D., Shobha, G.: S-DILS: framework for secured digital interaction learning system using keccak. Springer-Softw. Eng. Intell. Syste. Adv. Intell. Syst. Comput. **349**, 175–187 (2015)

A Vertical and Horizontal Segregation Based Data Dissemination Protocol

Shubhra Jain, Suraj Sharma and Neeraj Bagga

Abstract A base station and multiple sensor nodes are two main components of the wireless sensor network. Main task of sensor nodes is to aggregate and forward the collected data to the base station. The main objective of wireless sensor network is to maximize the network lifetime and providing as low latency delay as possible. To seek to accomplish this, we are proposing a Vertical and Horizontal segregation based data dissemination protocol. This protocol considers mobile base station that helps in improving the network lifetime by collecting the data from each and every node by moving to provide high throughput. Mobile base station requires informing all the sensor nodes about its location. This consumes large energy. Our proposed protocol saves this energy by informing few sensor nodes, in place of informing all the sensor nodes.

Keywords Sensor node · Sink node · Hotspot problem · Data dissemination

1 Introduction

A Wireless Sensor Network (WSN) mainly composed of two entities: an enormous number of sensor nodes and a mobile sink. Sensor nodes are tiny devices with limited sensing, processing and transmitting capability. They are of low cost, have low storage and computational capacity with a limited energy supply. It is not practically possible to recharge or change their batteries in some scenario. They should be handled and used in such a way to ensure their untended operation of the network for a long enough period of time. On the other hand, a base station or sink is a more resource rich device, i.e. a device with large storage and capacity power and with unlimited power. Sink node is the end point of the accumulated data. A sink may be either immobile (static) or mobile. Depending upon the application

S. Jain (✉) · S. Sharma · N. Bagga
Department of CSE, IIIT, Bhubaneswar 751003, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_37

various data dissemination protocols are proposed. For, e.g. directed diffusion protocol use static sink and Line Based Data Dissemination, Two Tier Data Dissemination and Ring routing protocol, etc. are based on mobile sink. In many scenarios, static sink is untrustworthy for security restraint, because retrieving location information will be easy to detect for attacker in the case of static sink. At the same time, mobile sink provides certainty and improves the lifetime of a network by avoiding excessive transmission overhead on the nodes that are close to the sink that would occupied by a static sink. Because of this overhead these nodes lose their batteries before other nodes, this problem is commonly known as hotspot problem [1]. Mobile sink avoids this problem. In the case of mobile sink, sink has to periodically broadcast their location to all the sensor nodes, so that sensors will know where the sink is. Sink mobility is useful in many applications such as target tracking and habitat monitoring. The communication between sensors and sink takes place in many-to-one paradigm. Tens, hundreds or thousands of sensor nodes can be networked together for a variety of applications for instance environmental monitoring, seismic, structural analysis, traffic monitoring and to measure and monitor ambient conditions in the surrounding environment. Typical sensing tasks are heat, pressure, temperature, motion, light, sound, vibration, pressure of objects, etc. [2]. Data dissemination is a mechanism, used by sensor nodes to propagate the accumulated data or information towards the sink with minimum number of hops to provide high data delivery ratio with minimum latency delay.

Considering these matters, we propose a Vertical and Horizontal Segregation Based Data Dissemination Protocol. In this protocol, flooding of the sink's location is not needed; instead sink will periodically update their location to a scanty of the sensor nodes called Spine nodes. It saves energy and enhances the network lifetime. Also, it uses the concept of backbone creation and tree construction. The rest of this paper is organized as follows: next section briefly analyses the related works. In the Sect. 3, our proposed data routing protocol is discussed. In the Sect. 4, we theoretically analyze this paper. In the Sect. 5, we conclude this paper.

2 Related Work

To expand the network's lifetime and improve the efficiency of the network many wireless sensor network data dissemination protocols have been proposed. In this section we concisely describe few disseminating protocols: Directed Diffusion, TTDD, HPDD, LBDD, Honeycomb based architecture and Ring Routing protocol. At the beginning, diffusion based routing algorithm "Directed Diffusion" [3], consider the static sink, constructs the gradients designating to the sink by broadcasting interests across the sensor field. Once a setup is established, we cannot change those gradients. To incorporate mobile sink, one solution is to update the routes by amending the gradients before mobile sink moves far from the previous location and set up a new structure after sink stops. Two-Tier Data Dissemination

(TTDD) [4] is the first routing protocol that uses the concept of mobile sink. It follows grid-based approach. In TTDD each source node starts constructing the grid with its own location. To get the data, sink floods its query within a local cell. A node which first receives the query from the sink known as an immediate dissemination node for the sink. This node propagates the query until it reaches to the source node which constructed the grid. Then source node sends the data by following the reverse path until it reaches to the sink's immediate dissemination node. Every time source node needs to call a grid construction, it is not feasible when data is generated by numerous sensor nodes and may limit the network lifetime. Hexagonal Path Data Dissemination (HPDD) [5] is same as TTDD, sensors are arranged uniformly as in TTDD, but instead of rectangular grid it uses hexagonal grid. Line Based Data Dissemination (LBDD) [6] defines a vertical line of nodes, which partitioned the sensor field into two equal halves. This virtual line act as a rendezvous area for query and data. Nodes inside this line are called in-line nodes and remaining nodes are called ordinary nodes. When any ordinary sensor node sensed some data, it sends to the virtual line and the first inline node that encountered, stores the data. Then for specific information, a sink sends a query to the virtual line in a perpendicular direction, then first node which receives the query, propagate it in both the direction along the line until it reaches to the inline node which possesses the data and the next time sink will send queries, directly to the inline node. The occurrence of any node failure may degrade the performance and the nodes that are close to the boundary of the network may result in high latency delay. Honeycomb based architecture [7] is based on hexagonal grid structure. In this architecture, sensor field is divided into six parts called Hextant, by virtual lines known as Borderlines. The lines which intersect the predefined center of the sensor network worked as a rendezvous area for data storage and lookup. Sensor nodes send their data to the center cell through the closest border line. Replication and storage of the data take place on the borderline nodes. The sink allows following the same strategy to propagate their query towards the center cell. When a query reaches to the border line node, which has data corresponding to the query, it is sent to the sink by following the reverse path from where the query is received. In Ring Routing [8] ring is a closed strip of sensor nodes, which is one node wide. In order to construct a ring, an initial radius is determined. Nodes closer to the radius play the role of ring node. To build a ring, we start from a node and select the next node in a greedy manner until we reached to the starting node and closed loop is formed. When sink moves an AN (anchor node) is being selected from its neighbor nodes. AN manages the communication between the sink and the sensor nodes. When a node has some data to send, it first obtains location information of AN from the ring. After getting the location of AN, the source node sends its data towards it. Since the ring node process the AN position information, hence they consume more energy. A ring node dies quickly as they handle more traffic. To inhibit these nodes from dying, they exchange their roles with regular nodes intermittently.

3 Proposed Work

In this section, we will discuss our proposed data dissemination protocol (Vertical and Horizontal segregation based protocol) for wireless sensor network with mobile sink. This data dissemination protocol works in five phases: backbone creation, neighbor discovery, tree construction, sinks location information, sink location recovery and data transmission.

Network Assumptions:

(i) All sensor nodes are homogeneous; they have equal power of computation, memory and capacity.
(ii) Sensor nodes are distributed in an even way in the sensor field and location aware as well.
(iii) All sensor nodes are immobile by nature.
(iv) Sink node is mobile and can move freely in the whole sensor field.
(v) Structure of sensor field is square shaped.

3.1 Backbone Creation

First, this protocol divides the sensor field in almost four equal parts by a vertical and horizontal stripe. Both of these stripes are the backbone of this protocol. Nodes inside these stripes from henceforth be known as Spine sensor nodes, and nodes outside these stripes will be known as Regular sensor nodes. We will create the Backbone as follows:

Suppose b and l is the width and length of the sensor fields.

W_h and W_v is the horizontal and vertical backbone selection with width W respectively.

$$W_h = ((b/2) - 1) \text{ to } ((b/2) + 1)$$
$$W_v = ((l/2) - 1) \text{ to } ((l/2) + 1)$$

Figure 1 shows the backbone creation in the Vertical and Horizontal Segregation Based data dissemination protocol. The region where horizontal and vertical stripes intersect each other is called a square region and nodes in this region are called square nodes. From now on we will call the right section of the horizontal backbone from the square region, h_r; similarly left section will be known as h_l. Upper section of the vertical backbone will be named as v_u, and the lower section is named as v_l. After creation of the backbone, we will again divide the sensor field into eight regions which are named as octants, with an angle of 45°. Each of eight octants is marked with roman numerals in the Fig. 1.

3.2 Neighbor Discovery

After constructing the backbone, neighbor discovery is performed by spine sensor nodes and regular sensor nodes both. Regular sensor nodes will broadcast a NBR packet along with its own id, ID and location, LOC to know about their neighbors. NBR_MSG is used for neighbor discovery process and NBR_REPLY is a reply of NBR_MSG.

For neighbor discovery in backbone region, spine nodes also broadcast the NBR packet, and apart from its own id and location, it will also send its RSSI value (Received Signal Strength Indication). This process continues till root node become neighbor of spine nodes of all the four backbone region.

Algorithm I: Neighbor discovery

```
1.  Begin
2.  any node y receives NBR_MSG from node x
3.   if (y ∈ spine sensor nodes && x ∈ spine sensor nodes) then
4.          br_l <NBR_MSG, IDx, LOCx, RSSIx>
5.          if (IDx∉ NBR(y)), then                          // if node y does not contains IDx previously
6.              NBR(y) ← {NBR(y)} ∪ {IDx, LOCx, RSSIx}
7.              if (NBR_MSG_SENT == false)
8.                br_l<NBR_MSG, IDy, LOCy>
9.                NBR_MSG_SENT←True;
10.             end if
11.              if (NBR(y)==R), then
12.             Stop neighbor discovery procedure in backbone region
13.          else
14.            drop the packet
15.         endif
16.   elseif (y ∉ spine sensor nodes && x ∉ spine sensor nodes), then
17.           br_l<NBR_MSG, , IDx, LOCx>
18.          if (IDx ∉ NBR(y)) then
19.            NBR(y) ← {NBR(y)} ∪ {IDx, LOCx}
20.          goto step 7
21.   endif.
22.  end
```

3.3 Tree Construction

The tree construction is performed after a neighbor discovery process. End nodes of the four sections of backbone h_r, h_l, v_u, v_l start broadcasting T_MSG with its id and selects a node having a moderate RSSI (Received Signal Strength Indication) value, i.e. whose value lies between 0.2 and 0.4. The minimum RSSI value indicates more distance. After picking such node as a next node of the tree, the sender node establishes a link. This process continues till we reach to the root node. Root node is the node close to the centre of network area. We are considering the minimum RSSI value so that we can reach to the root node with as minimum spine node as possible. Since we are not considering all the nodes of the backbone region in tree construction, it facilitates to replace the spine tree nodes with other spine nodes when

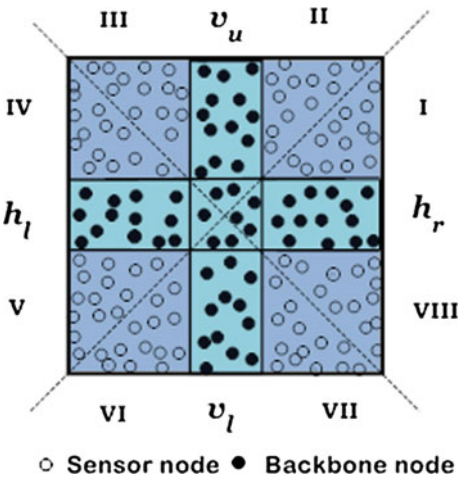


Fig. 1 Backbone creation in vertical and horizontal segregation based protocol

their energy is falling down from a minimum threshold value. It results in an extended network lifetime. Another reason for this is spine tree nodes are designed to keep the sink location information only, neither for data storage nor data forwarding. The Fig. 2 demonstrates how we are constructing tree. Algorithm II shows this procedure.

Algorithm II: Tree Construction

```
1.  Begin
2.  R←min_dist (C, nodes)              //R is the root node or centre node and C is the centre of the sensor network
3.  for each backbone region i, repeat
4.     parent_node_m←End_id_m                                   // end node m's id
5.       for selecting the node with moderate RSSI value i.e. 0.2<RSSI_n<0.4
6.       br_l (T_Msg, ID_m, NBR_m)                              // node m broadcasts T_MSG
7.       While receiving T_MSG by any spine node n, then
8.        if (ID_n∈ NBR_m)                                      // check whether n∈ NBR(m) or not
9.          if (T_Msg_SENT == false)      // node m chooses the neighbor having RSSI, 0.2<RSSI<0.4
10.         br_l (T_Msg, ID_n, ID_m)
11.         T_Msg_SENT← True;
12.         endif
13.      endif
14.      While receiving T_MSG reply by node m, then
15.      parent_node_n← ID_m
16.      child_node_m ← ID_n
17.      if (Child node_m == R)                                 // R denotes a root node
18.         stop the Tree construction procedure
19.      else
20.      goto step4
21. end for
22. end
```

3.4 Sensor Node Region Discovery

After dividing the sensor field into octant, now we have to decide the specific region between the two of backbone, to which sensor nodes will communicate to obtain the location of the sink. We can do this by following the Algorithm III. For example, the nodes belonging to I and VIII octant will communicate to the h_r backbone. Similarly, II and III octant sensor node can get sink location information from v_u backbone and so on. Figure 3 depicts this scenario.

Algorithm III: Sensor node region discovery

```
1.  Begin
2.  α = 45°                                          // α is the angle of the center.
3.  R ← (u, v)                                       // R is the center.
4.  for any node (x, y)
5.  New coordinates (A,B) ← (x-u, y-v)               // evaluate (A,B) corresponding to the center R.
6.  Calculate (rsinθ, rcosθ)                         // sinθ, cosθ are polar coordinate
7.  Calculate θ = tan^-1|B/A|
8.  if ( A>0  && B>0), then
9.          ∝←θ                                      // because (A, B) is in 1st quadrant.
10.         if α lies between 0°- π/4, then
11.         (x, y) belongs to I octant and (x, y) can get sink location information from hr.
12.         else if ∝ lies between π/4 – π/2, then
13.                 (x, y) belongs to II octant and (x, y) can get sink location information from vu.
14. else if ( A<0 && B>0), then
15.         α← π-θ                                   // because (A,B) is in 2nd quadrant.
16.         if  α lies between π/2– 3π/4, then
17.                 (x, y) belongs to III octant and (x, y) can get sink location information from vu.
18.         else if  α lies between 3π/4– π,then
19.                 (x, y) belongs to IV octant and (x, y) can get sink location information from hr.
20. else if ( A>0 && B<0), then
21.         α← π+θ                                   // because (A,B) is in 3rd quadrant.
22.         if α lies between  π- 5π/4, then
23.                 (x, y) belongs to V octant and (x, y) can get sink location information from hr.
24.         else if  α lies between 5π/4- 3π/2, then
25.                 (x,y) belongs to VI octant and (x,y) can get sink location information from vL
26. else if ( A<0 && B<0), then
27.         α← π-θ                                   // because (A,B) is in 4thquadrant.
28.         if α lies between 3π/2- 7π/4, then
29.                 (x, y) belongs to VII octant and (x, y) can get sink location information from vL
30.         else if α lies between 7π/4- 2π, then
31.                 (x, y) belongs to VIII octant and (x, y) can get sink location information from hr.
32. end if
33. end
```

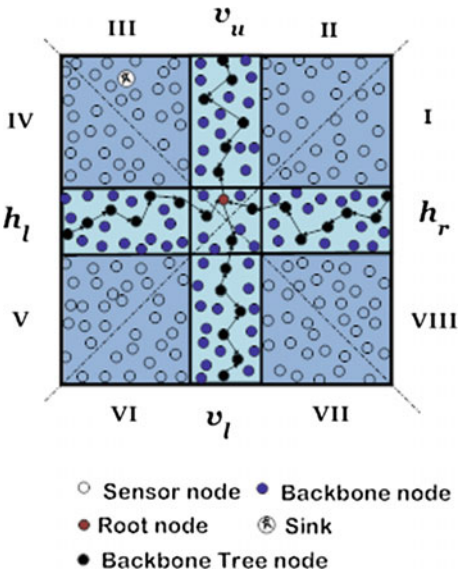


Fig. 2 Tree construction in vertical and horizontal segregation based protocol

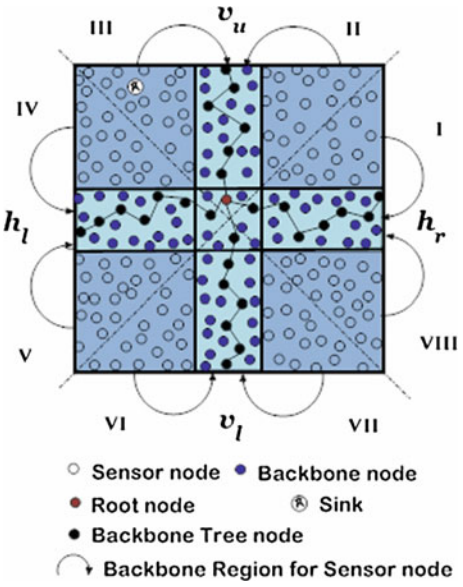


Fig. 3 Sensor region node discovery in vertical and horizontal segregation based protocol

3.5 Sink Location Information

In this phase, sink will broadcast a HELLO_MSG. Any nearest node that receives this message will send a HELLO_REPLY packet including its ID and its location, LOC. Sender of this reply packet will become the gateway node for the sink. Then

sink will send a S_MSG which also contains sink's location, LOC_{sink} and sink's Pause_time. Pause_time is the lifetime of the sink node at that location. Then gateway node will keep forwarding this information until it reaches to any spine tree node. Once it reaches to spine tree node, then that node will broadcast this information to other tree nodes with the help of root node. Every node will keep this information till the timer expires. Figure 4 illustrates how sink informs about its location to the spine nodes. Algorithm IV illustrates it as follows:

Algorithm IV: Sink location information

```
1. Begin
2.  Sink broadcasts the HELLO_MSG
3.  While receiving HELLO_MSG by any sensor node x.
5.          br_l< HELLO_REPLY, IDx , LOCx>
6.        While receiving HELLO_REPLY by sink from node x,
7.          x ← gateway.
8.         br_l<S_MSG, LOCsink, Pause_time>              //Send S_MSG to node x
9.        x forwards S_MSG until it reaches to any spine tree node.
10.     if (S_MSG reaches to the any spine tree node)
11.       flooding the S_MSG to all the spine tree nodes
12.     endif
13. end.
```

3.6 Sink Location Recovery and Data Transmission

When a regular sensor node have some data or information to transmit to the sink. It will send a SLQ (Sink Location Query) packet to the nearest spine tree (ST) node with its ID, to know the location of sink. After receiving the SLQ packet that tree node will reply with SLR (Sink Location Reply) packet which includes sink's location and its pause_time from where it received SLQ packet. After obtaining the sink's location, the source node (node having data) will send its data in a multihop manner, since each node is location aware, so it will choose a path to reach to the sink with minimum number of hops. Figure 5 shows this phase. Algorithm V elaborates this step by step.

Algorithm V: Sink location recovery and data transmission

```
1. Begin
2.    for any node x, who want to forward the data to sink
3.    if (NBRx∈ ST_node)
4.       br_l< SLQ, IDx, LOCx>
5.       while receiving SLQ message by any spinal tree node
6.       br_l< SLR ,IDST-node, LOCST_node>
7.      while receiving SLR by node x
8.       forwards the data to sink by node x, using SLR packet.
9.    else
10.      x ← NBRx
11.       goto step 3, until ST_node ∈ NBRx
12.   endif
13  end
13. end.
```

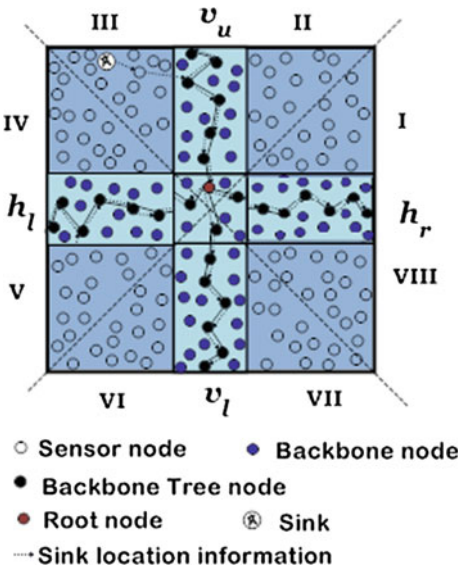


Fig. 4 Sink location information in vertical and horizontal segregation based protocol

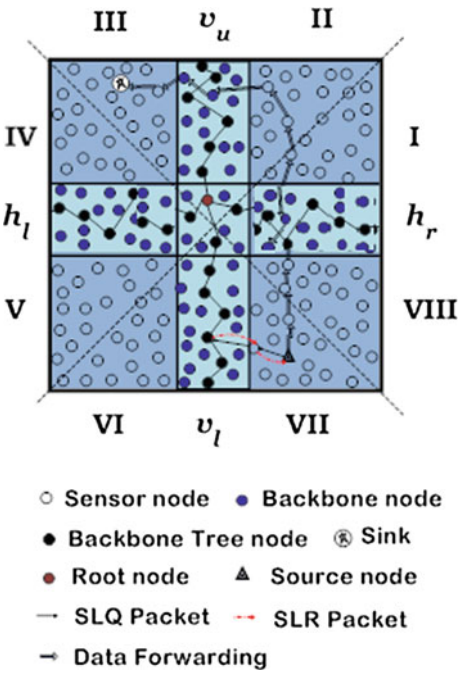


Fig. 5 Sink location recovery information and data transmission in vertical and horizontal segregation based protocol

4 Theoretical Analysis

In the Ring routing protocol, it is easy to cope up when ring node loses their energy. Getting updated sink location information from the ring is easy, but every time for sending the data sensor nodes need to first obtain the location of the sink from ring node, it may lead to wastage of time and may increase the delay especially when sink did not move from its previous location. And for a large sensor network ring construction will be exigent and complex. The sensor nodes, which are away from ring node have to travel a long path to get the location information of the sink, it also increases the latency delay. On the other hand, Vertical and Horizontal segregation based Tree data dissemination protocol provides high data delivery ratio and decreases end to end delay as regular sensor nodes does not require to obtain the location information of the sink every time when they have some data to send since along with the location information they also get its pause time, so they did not require to send SLQ packet till the timer expires.

In honeycomb based architecture, border lines which plays the role of rendezvous area responsible for responding the sink queries and data forwarding, in this situation border line nodes will lose their energy earlier and may cause bottleneck of area. Whereas, in vertical and horizontal segregation based tree data dissemination protocol only few of the spine nodes are responsible for storing and sending the sink location information.

5 Conclusion

Our proposed Vertical and Horizontal segregation based tree data dissemination protocol, combines the best features of Honeycomb architecture and LBDD, as it also has spine nodes, but the role of these nodes is only to keep the location information of the sink, instead of storing the data and forwarding responsibility, as in Honeycomb architecture. We create the tree by connecting the few spine nodes, so energy consumption is greatly reduced in comparison to the LBDD. Hence, it results in prolonging network lifetime. Since regular sensor nodes can get sink location within a short period of time, it helps in reducing end-to-end delay and increases data throughput.

References

1. Tunca, C., Isik, S., Donmez, M.Y., Ersoy, C.: Distributed mobile sink routing for wireless sensor networks: a survey. IEEE Commun. Surv. Tutorials **16**(2), 877–897 (2012)
2. Shah, R.C., Roy, S., Jain, S., Brunette, W.: Data mules: Modeling and analysis of three-tier architecture for sparse sensor networks. Ad Hoc Netw. **1**(2), 215–233 (2003)

3. Intanagonwiwat, X., Chalermek, X., Govindan, R., Estrin, D., Heidemann, J., Silva, F.: Directed diffusion for wireless sensor networking. IEEE/ACM Trans. Netw. **11**(1), 2–16 (2003)
4. Luo, H., Ye, F., Cheng, J., Lu, S., Zhang, L.: TTDD: two-tier data dissemination in large-scale wireless sensor networks. Wirel. Netw. **11**(1–2), 161–175 (2005)
5. Shon, M., Kong, C., Choo, H.: Hexagonal path data dissemination for energy efficiency in wireless sensor networks. In: Information Networking, International Conference, pp. 1–5 (2009)
6. Hamida, E.B., Chelius, G.: A line-based data dissemination protocol for wireless sensor networks with mobile sink. In: Communications, IEEE International Conference, pp. 2201–2205 (2008)
7. Erman, A.T., Havinga, P.: A fault-tolerant data dissemination based on honeycomb architecture for mobile multi-sink wireless sensor networks. In: Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), pp. 97–102. IEEE, Dec 2010
8. Tunca, C., Donmez, M.Y., Isik, S., Ersoy, C.: Ring routing: an energy-efficient routing protocol for wireless sensor networks with a mobile sink. In: Signal Processing and Communications Applications Conference (SIU), pp. 1–4. IEEE (2012)

Application of Data Mining Techniques on Heart Disease Prediction: A Survey

Ritika Chadha, Shubhankar Mayank, Anurag Vardhan and Tribikram Pradhan

Abstract Globally, the medical industry is presumably "information rich" and "knowledge poor". KDD, i.e. knowledge discovery from data is hence, applied to extract interesting patterns from the dataset using different data mining techniques. This massive data available is essential for the extraction of useful information and generate relationships amongst the attributes. The aim of this paper is to compile, tabulate and analyze the different data mining techniques that have been implied and implemented in the recent years for Heart Disease Prediction. Each previous paper exhibits a set of strengths and limitations in terms of the data types used in the dataset, accuracy, ease of interpretation, reliability and generalization ability. This paper strives to bring out stark comparisons and put light to the pros and cons of each of the techniques. By far, the observations reveal that Neural Networks performed well as compared to Naive Bayes and Decision Tree considering appropriate conditions.

Keywords Heart disease · Decision tree · Naive bayes · Classification · Neural networks · Genetic algorithm

1 Introduction

Heart disease is solely the largest cause of death in developed countries and one of the main contributors to disease burden in developing countries. Due to the shortage of doctors and experts and neglect of the patients' symptoms frequently calls for data mining that serves as an analysis tool to discover hidden relationships and patterns in HD (Heart Diseases) medical data. Pre-requisites required for detecting a disease are the numerous tests that a patient has to go through. Added to this is the large amount of complex data about patients, hospital resources, disease diagnosis, electronic patient records, medical devices etc. To prevent this cost-consuming,

R. Chadha (✉) · S. Mayank · A. Vardhan · T. Pradhan
Department of ICT, Manipal University, Manipal 576104, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_38

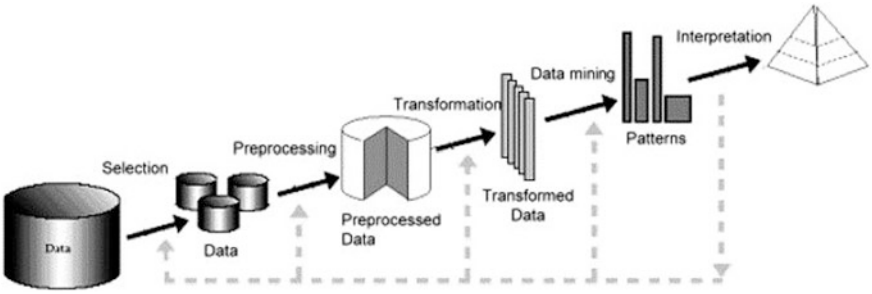


Fig. 1 KDD

cumbersome task, data mining technique comes into play that is efficient and cost-effective. Data mining techniques are the result of a long process of experimenting and R&D. It is divided into two tasks-Predictive Tasks and Descriptive Tasks. Data mining involves few steps from raw data collection to some form of interesting pattern. The process which takes place in iteration includes-Data Cleaning, Data Integration, Data Selection, Data Transformation, Data Mining, Pattern Evaluation, and Knowledge Discovery Process as shown in Fig. 1.

Our work presents an overall view all such tasks that are performed for the extraction of data to be made possible in order to increase the prediction rate of the heart disease by the application of various essential data mining techniques and processes.

2 Related Work

In this paper, [1] Nidhi Bhatla et al., have performed an experiment in their work **An Analysis of Heart Disease Prediction using Different Data Mining Techniques** using the data mining tool Weka 3.6.6. This research results in an accuracy of Neural networks of 100 % compared to 99.62 % and 90.74 % in Decision tree and Naïve Bayes respectively. The method of **Fuzzy Logic and Genetic Algorithm** is used that amalgamates the genetic algorithms for feature selection and fuzzy expert systems by experimenting in Matlab using fuzzy tool. Two kinds of algorithms are used. Earlier, 13 attributes were used for this prediction but this research work reduced the number of attributes to six only using Genetic Algorithm and Feature Subset Selection. This prototype **IHDPS**, Intelligent Heart Disease Prediction System had been developed using techniques such as, Decision Trees, Naive Bayes and Neural Networks. In this paper, the analysis shows that Neural Network has the highest accuracy i.e. 100 % so far. On the other hand, Decision Tree has also performed well with 99.62 % accuracy by using 15 attributes.

The work of Amin et al. [2], **Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors** developed an intelligent data mining system based on genetic algorithm. To transform data into useful form, encoding was done between a range [−1, 1]. **Neural Network Weight Optimization by Genetic Algorithm** system uses back-propagation algorithm for learning and training the neural network on algorithm. The disadvantages were removed in this paper by optimizing the initial weights of neural network. For this, a genetic algorithm which is specialized for global searching was used. The accuracy of prediction of heart disease on the training data was calculated as 89 % and accuracy on validation data was 96.2 %.

In this paper [3] **HDPS: Heart Disease Prediction System**, AH Chen et al. have applied the technique of ANN by using an LVQ system to represent the feature space of observed data using prototypes W = (w(i), …, w(n)). The application of winner-take-all training algorithm is used where the position of the so-called winner is moved closer if it correctly classifies the data point or swayed away if the choice is not apt. The accuracy of classification is near 80 % as well as 85 % sensitivity and 70 % specificity. Their approach consists of three steps namely-selection of 13 important clinical features-age, sex, chest pain type, trestbps, cholesterol, fasting blood sugar, resting ecg, max heart rate, exercise induced angina, old peak, slope, number of vessels colored, and thal. 80 % prediction rate is obtained by developing an artificial neural network algorithm. The next step includes a user-friendly heart disease predict system (HDPS) that generates prediction results using artificial neural network (ANN) techniques on C and C# environment.

This paper [4] **Heart Disease Prediction using Lazy Associative Classification** by M. Akhil Jabbar, lazy data mining approach for heart disease classification is applied. Information centric attribute measure, PCA is applied to generate class association rules. This class association rules is used to predict the occurrence of heart disease. This approach has improved 10.8 % against J4.8 and 19.8 % improvement over naïve Bayes for non-medical data sets. This approach reached 10.26 % improvement over J4.8 and 8.6 % improvement against naïve Bayes respectively for heart disease data set.

In this paper [5], **Early Prediction of Heart Diseases Using Data Mining Techniques**, three classifiers as ID3, CART and DT are used wherein CART is the most accurate with 83.49 % and 0.23 s are engaged to build the model. The most important attributes for heart diseases are cp (Chest pain), slope (The slope of the peak exercise segment), Exang (Exercise induced angina), and Restecg (Resting electrocardiographic). These attributes were found using three tests for the assessment of input variables: Chi-square test, Info Gain test and Gain Ratio test.

Paper [6] **Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques** by Chaitrali S. Dangare includes two more input attributes obesity and smoking to improve the overall prediction rate. Decision trees, Naive Bayes and Neural networks are used which results in Neural Networks providing more accurate results as compared to Decision trees and Naive Baye where Neural Networks has a rate of 99.25 % as opposed to Naive Bayes (94.44 %) and Decision Tree (96.66 %).

In this paper [7] **Predictive Data Mining for Medical Diagnosis**, by incorporating techniques like—ANN, Time Series, Clustering and Association Rules, soft computing approaches etc., Jyoti Soni et al. concluded that Decision Tree outperforms and sometimes Bayesian classification is having similar accuracy as of decision tree but other predictive methods like KNN, Neural Networks, Classification based on clustering are not performing well. Also, after the application of genetic algorithm, the accuracy of the Decision Tree and Bayesian Classification further improves.

Paper [8] **Intelligent Heart Disease Prediction System Using Data Mining Techniques**, Sellappan Palaniappan et al. have used three data mining techniques. Extraction of hidden knowledge from a historical heart disease database, building and accessing models through DMX query language and functions and the training and validation against a test dataset. Effectiveness is accounted for by using Lift Chart and Classification Matrix. The most effective model to predict patients with heart disease appears to be Naïve Bayes followed by Neural Network and Decision Trees.

In the work of Hlaudi Daniel Masethe [9] **Prediction of Heart Disease using Classification Algorithms**, an experiment was performed for the prediction of heart attacks and comparison to find the best method of prediction. This can act as an important tool for physicians to predict risky cases in the practice and advice accordingly. The predictive accuracy determined by J48, REPTREE and SIMPLE CART algorithms suggests that parameters used are reliable indicators to predict the presence of heart diseases.

The work of K. Sudhakar et al. [10], **Study of Heart Disease Prediction using Data Mining** presents the different techniques that are deployed in the recent years for calculating the prediction rate in heart disease. These techniques include-ANN, BN, Decision Trees and Classification Algorithms.

In the work of Aditya Sundar et al. **Performance Analysis of Classification Data mining Techniques Over Heart Disease Data Base** [11], after experimentation, a prototype has been described using data mining techniques namely, Naïve Bayes and WAC (weighted associative classifier). It creates a bridge between significant data and knowledge e.g. patterns, relationships between the medical symptoms. It serves as a tool to train nurses and medical interns to treat patients with heart diseases.

The work of Abhishek Taneja [12] **Heart Disease Prediction System Using Data Mining Techniques** deals with the conduction of 4 experiments by employing selected classification algorithms on a full training dataset containing 7339 instances. KDD has been used in order to develop a prediction model that can predict heart disease cases based on calculations done.

In this paper **Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks** [13] by K. Srinivas et al., the data mining techniques such as-Rule Based, Decision Tree, Naïve Bayes, and Artificial Neural Network to massive volumes of medical care data.

3 Objective

Our paper brings into limelight all the advantages and disadvantages of using the different data mining techniques for the prediction of heart diseases. It also accounts for the prediction rate for different techniques hence, bringing out the comparison between each of them.

4 Methodology

The main methodology used for our work was by examining the publications, journals and reviews in the field of computer science and engineering, data mining and cardiovascular disease in recent times.

4.1 Data Mining and Neural Networks

An artificial neural network (ANN), also known as in short "neural network" (NN), is a mathematical model or computational model based on the neural network found in human anatomy. In this work, it is observed that the Heart Disease Prediction System has been developed using 15 attributes for a 100 % accuracy. However, in few papers, 13 attributes have also been used. For the calculations of the required neural network figure, Weka 3.6.6 is used for experimenting along with few of the researchers implementing heart disease classification and prediction trained via ANN using C as a tool. A Multi-layer Perceptron Neural Networks (MLPNN) with Back-propagation is used. The structure of MLPNN is as shown in Fig. 2. Framework of ANN model: It maps a set of input data onto a set of appropriate

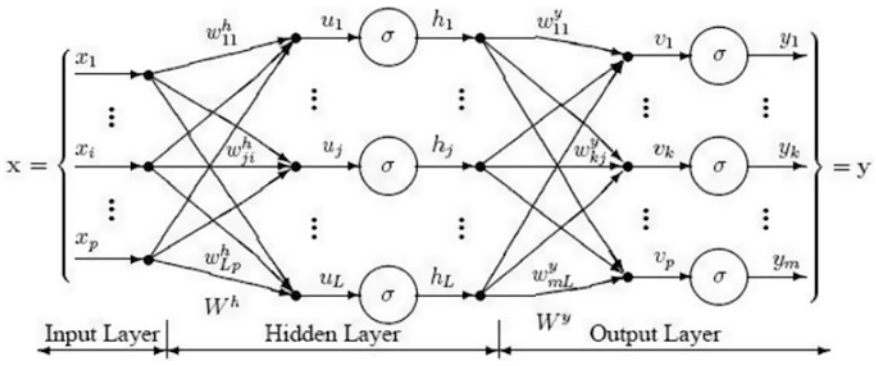


Fig. 2 Structure of MLPNN

output data. It consists of 3 layers namely -input layer, hidden layer & output layer. Weights are allotted to each connection or branch from that particular neuron.

4.2 Genetic Algorithm

Genetic Algorithm (GA) is a heuristic that imitates the process of Darwin's natural evolution as cited in Fig. 3. This algorithm is used to generate optimized solutions. It is inspired by techniques like- inheritance, mutation, selection, and crossover. For instance, in the heart disease prediction, using Feature Subset Selection, GA is used for the reduction of the number of attributes. This includes a set of input values that are routinely considered through the application of fitness function which is nonetheless, flexible expression of modelling criteria.

4.3 Decision Tree (DT)

Decision tree using the classification or regression techniques are built in the form of a tree structure as shown in Fig. 4. It segregates a dataset into smaller subsets. The final outcome is a tree with decision nodes and leaf nodes. A decision node represents the branches while the leaf node is the result of the decision-making process. The topmost decision node in a tree is known as the root node. Each leaf is assigned to one class representing the apt target value. The leaf may hold a probability vector too. Top-down approach is implemented where navigation is done from the root to the leaf according to the result of the tests along the path.

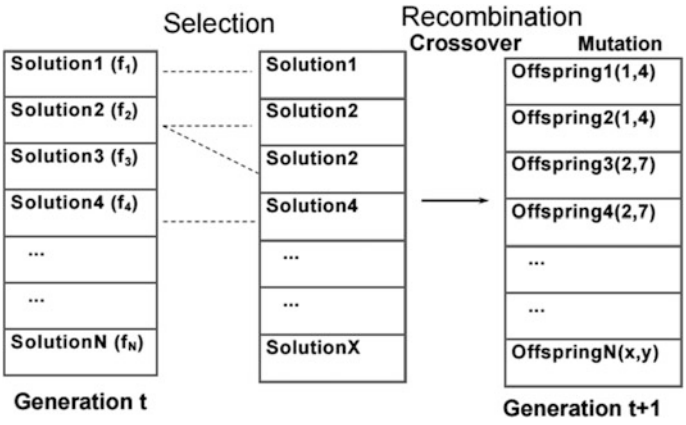


Fig. 3 Overall model of genetic algorithm

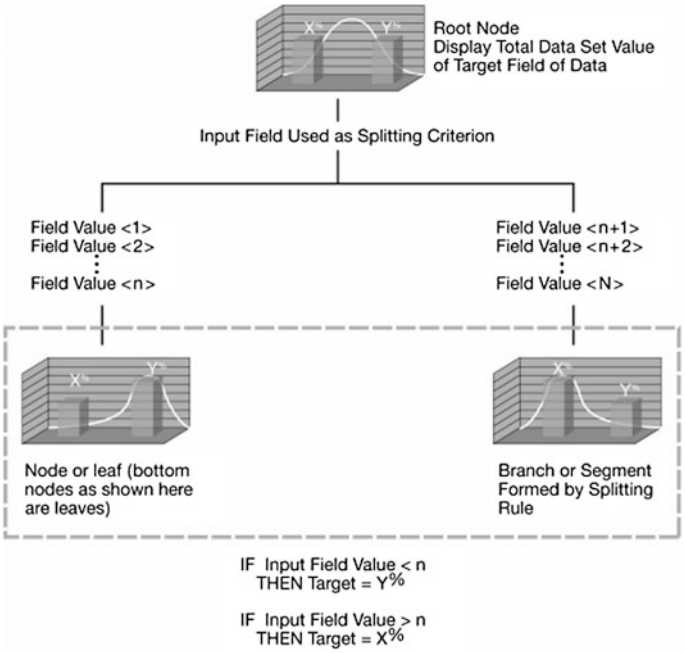
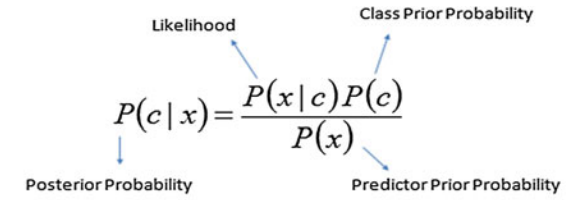


Fig. 4 Structure of a decision tree

4.4 Naive Bayes

This data mining classifier is based on the mathematical model called Bayesian theorem and is perfect for use when the dimensionality of the inputs is very high. Despite it being simple, Naive Bayes can outperform more sophisticated, complex classification methods. Bayes theorem provides a method for the calculation of posterior probability $P(c|x)$, from $P(c)$, $P(x)$, and $P(x|c)$. This assumption is called class conditional independence. The theorem states as follows:



$$P(c \mid \mathrm{X}) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Table 1 Comparison of the prediction rate using data mining techniques in recent years

Title	Author	Year	Dataset	Type	Advantages	Disadvantage	Prediction result
An analysis of heart disease prediction using different data mining techniques	Nidhi Bhatla, Kiran Jyoti,	2013	The dataset from UCI machine learning repository is used.	Neural networks, fuzzy logic and genetic algorithm, supervised machine learning, genetic algorithm, IHDPS	Effective classification	ANN requires more fine tuning. GA's are slow.	Naive Bayes 96.5 % Decision Trees 99.62 % Neural Networks 100 % KNN 45.67 % Classification via Clustering 88.3 %
Genetic neural network based data mining in prediction of heart disease using risk factors	Syed Umar Amin, Kavita Agarwal, Dr. Rizwan Beg,	2013	The data for 50 people was collected from surveys done by the American Heart Association.	Data analysis and encoding, neural network weight optimization by genetic algorithm, neural networks	Data analysis was needed for correct data preprocessing. ANN requires less formal statistical training,	Back propagation algorithm is very slow, and "black box" nature of ANN.	The accuracy of prediction of heart disease on the training data was calculated as 89 % and accuracy on validation data was 96.2 %. The least mean square error (MSE) achieved was 0.034683
HDPS: heart disease prediction system	AH Chen, SY Huang, PS Hong, CH Cheng, EJ Lin	2011	heart disease data from machine learning repository of UCI. We have total 303 instances of which 164 instances belonged	Artificial neural network	One benefit of LVQ is that it creates prototypes that are easy to interpret for experts in the	Its "black box" nature, greater computational burden, proneness to overfitting, and the empirical	The accuracy of classification is near 80 % as well as 85 % sensitivity and 70 % specificity. To confirm the

(continued)

Table 1 (continued)

Title	Author	Year	Dataset	Type	Advantages	Disadvantage	Prediction result
			to the healthy and 139 instances belonged to the heart disease		respective application domain	nature of model development.	goodness of this model, a ROC curve is also displayed in Fig. 4.
Heart disease prediction using lazy associative classification	M. Akhil Jabbar Dr B.L Deekshatulu, Dr. Priti Chandra,	2013	N.A.	Associative classification, principle component analysis, lazy associative classification method	It is possible to build more accurate classifier. Reduced complexity in images' grouping with the use of PCA	The covariance matrix is difficult to be evaluated in an accurate manner. Lazy classifiers typically require more work to classify all test instances.	Accuracy of classification is 90 % for the proposed system.
Early prediction of heart diseases using data mining techniques	Vikas Chaurasia, Saurabh Pal	2013	Heart disease data set available at http://archive.ics.uci.edu/ml/datasets/Heart+Disease. The data set has 76 raw attributes.	CART, ID3, decision tree	CART is easily accessible to beginning users and does not require a high level of technical expertise to operate.	Trees can be extremely sensitive to small perturbations in the data:	83.49 % in CART, 72.93 % in ID3, 82.50 % in Decision tree
Improved study of heart disease prediction system using data mining	Chaitrali S. Dangare, Sulabha S. Apte, Ph.D.	2012	The publicly available heart disease database is used. The Cleveland heart	Decision trees, Naive Bayes, ANN	More powerful for classification problems. Easy to implement.	Trees can be extremely sensitive to small perturbations in the data: "black box" nature of ANN,	Decision Tree: 96.66 % for 13 attributes, 99.62 % for 15 attributes Naive Bayes: 94.44 % for 13

(continued)

Table 1 (continued)

Title	Author	Year	Dataset	Type	Advantages	Disadvantage	Prediction result
classification techniques			disease database is used.			greater computational burden, proneness to overfitting, and the empirical nature of model development.	attributes, 90.74 % for 15 attributes Decision Trees: 99.25 % for for 13 attributes, 100 % for 15 attributes
Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction	J yoti Soni Ujma Ansari Dipesh Sharma Sunita soni	2011	Total of 909 records with 15 medical attributes (factors) were obtained from the Cleveland heart disease database.	Data mining and artificial neural network, genetic algorithm, association rule discovery	It is very comfortable and efficient way of problem solving!	"Black Box" nature of ANN	Accuracy of ANN is 85.53 %, for decision trees it 89 %, and 86.53 % for Naive bayes
Applications of data mining techniques in healthcare and prediction of heart attacks	K. Srinivas B. Kavihta Rani Dr. A. Govrdhan	2010	N.A.	Rule set classifiers, decision trees, ANN, neuro fuzzy, Bayesian Network structure discovery.	Easy to interpret • easy to generate • can classify new instances rapidly • decision trees are powerful classification problems	They can be extremely sensitive to small perturbations in the data, and "black box" nature of ANN	

(continued)

Table 1 (continued)

Title	Author	Year	Dataset	Type	Advantages	Disadvantage	Prediction result
Heart disease prediction system using data mining techniques	Abhishek Taneja	2013	The patient data set is compiled from data collected from medical practitioners in South Africa.	Decision tree classification, Naive Bayes, ANN	more powerful for classification problems. Naive Bayes is easy to implement	Trees can be extremely sensitive to small perturbations in the data:	J48 unpruned with all attributes 94.29 % J48 pruned with all attributes 95.41 % J48 unpruned with selected attributes 95.52 % J48 pruned with selected attributes 95.56 % Naive Bayes with all attributes 91.96 % Naive Bayes with selected attributes 92.42 % Neural Network with all attributes 93.83 % Neural Network with selected attributes 94.85 %
Intelligent heart disease prediction system using data mining techniques	Ms. Ishtake S.H, Prof. Sanap S.A.	2013	A total of 909 records with 15 medical attributes (factors) were obtained from the Cleveland Heart Disease database	Decision Tree Classification, Naive Bayes, ANN	More powerful for classification problems. Naive Bayes is easy to implement	Trees can be extremely sensitive to small perturbations in the data:	Accuracy is 94.93 % for decision trees, 95 % for Naive Bayes, 93.54 % for artificial neural networks.

(continued)

Table 1 (continued)

Title	Author	Year	Dataset	Type	Advantages	Disadvantage	Prediction result
Performance analysis of classification data mining techniques over heart disease data base	N. Aditya Sundar1, P. Pushpa Latha2, M. Rama Chandra3	2012	N.A.	Naive Bayes, weighted association classifier, a priori algorithm:	Easy to implement Good results obtained in most of the cases easily parallelized easy to implement	Apriori can be very slow.	78 % for Naive Bayes and 84 % for Weighted Association classifier.
Study of heart disease prediction using data mining	K. Sudhakar, Dr. M. Manimekalai	2014	N.A.	Neural Networks, Decision trees, Naive Bayes, associative classification	more powerful for classification problems, Easy to implement	Trees can be extremely sensitive to small perturbations in the data: Black Box nature of ANN	
Prediction of heart disease using classification algorithms	Hlaudi Daniel Masethe, Mosima Anna Masethe	2014	Compiled from data collected from medical practitioners in South Africa	Decision tree	More powerful for classification problems, easy to implement	Trees can be extremely sensitive to small perturbations in the data: Black Box nature of ANN	J48: 99.0741 Reptree: 99.0741 Naive Bayes: 97.222 Bayes net: 98.1481 simple cart: 99.0741

5 Comparison of the Recent Years

On studying different papers written in recent years, Table 1 has been constructed. This table bring out a stark contrast in the prediction rate on using different techniques.

6 Conclusion and Future Work

For clear understanding, results/prediction rate for each of the papers are summarized in a tabular form and the best prediction rate obtained in each of the techniques/methodologies is summarized by studying, analyzing and performing a survey on all of the recent papers. It is perceived from our observations/experiments that in few cases, the same classifier produces different accuracy for different data mining techniques based on the number of attributes chosen and the kind of algorithm that is applied. Several classifiers are analyzed for the required prediction.

We need to consider more varied parameters in the dataset for a complete accuracy of the prediction system. The intent is to develop more intelligent heart disease prediction models that employs more of the data mining techniques.

References

1. Bhatla, N., Jyoti, K.: An analysis of heart disease prediction using different data mining techniques. Int. J. Eng. Res. Technol. (IJERT)
2. Amin, S.U., Agarwal, K., Beg, R.: Genetic neural network based data mining in prediction of heart disease using risk factors. In: IEEE Conference on Information and Communication Technologies (2013)
3. Chen, A.H., Huang, S.Y., Hong, P.S., Cheng, C.H., Lin, E.J.: HDPS: heart disease prediction system. In: Computers in Cardiology Conference
4. Jabbar, M.A., Deekshatulu, B.L., Chandra, P.: Heart Disease Prediction using Lazy Associative Classification, 2013 IEEE (2013)
5. Chaurasia, V., Pal S.: Early prediction of heart diseases using data mining techniques. Caribb. J. Sci. Technol.
6. Dangare, C.S., Apte, S.S.: Improved study of heart disease prediction system using data mining classification techniques. Int. J. Comput. Appl. (0975–888) **47**(10), June 2012
7. Soni, J., Ansari, U., Sharma, D., Soni, S.: Predictive data mining for medical diagnosis. Int. J. Comput. Appl. **17**(8), 0975–8887 (2011)
8. Ishtake, S.H., Sanap, S.A.: Intelligent heart disease prediction system using data mining techniques. Int. J. Healthc. Biomed. Res. **1**(3), 94–101 (2013)
9. Masethe, H.D., Masethe, M.A.: Prediction of heart disease using classification algorithms. In: World Congress on Engineering and Computer Science 2014 Vol II WCECS 2014, San Francisco, USA, 22–24 Oct 2014
10. Sudhakar, K., Manimekalai, M.: Study of heart disease prediction using data mining. Int. J. Adv. Res. Comput. Sci. Softw. Eng.

11. Sundar, N.A., Latha, P.P., Chandra, M.R.: Performance analysis of classification data mining techniques over heart disease data base. [IJESAT] Int. J. Eng. Sci. Adv. Technol.
12. Taneja, A.: Heart disease prediction system using data mining techniques. Orient. J. Comput. Sci. Technol.
13. Srinivas, K., Kavihta Rani, B., Govrdhan, A.: Applications of data mining techniques in healthcare and prediction of heart attacks. (IJCSE) Int. J. Comput. Sci. Eng. **02**(02), 250–255 (2010)

An Efficient Approach for Constructing Spanning Trees by Applying BFS and DFS Algorithm Directly on Non-regular Graphic Sequences

Prantik Biswas, Abhisek Paul, Ankur Gogoi and Paritosh Bhattacharya

Abstract Realization of graphic sequences and finding the spanning tree of a graph are two popular problems of combinatorial optimization. A simple graph that realizes a given non-negative integer sequence is often termed as a realization of the given sequence. In this paper we have proposed a method for obtaining a spanning tree directly from a degree sequence by applying BFS and DFS algorithm separately, provided the degree sequence is graphic and non-regular. The proposed method is a two step process. First we apply an algorithm to check whether the input sequence is realizable through the construction of the adjacency matrix corresponding to the degree sequence. Then we apply the BFS and DFS algorithm separately to generate the spanning tree from it.

Keywords Spanning tree · Graph · Algorithms · BFS · DFS · Graphic realization · Degree sequence · Adjacency matrix

1 Introduction

Spanning trees problems forms the core of a numerous set of problems in graph theory. It has a wide range of application in various fields of science and technology ranging from computer and communication networks, wiring connections, VLSI circuits design to travelling salesman problem, multi-terminal flow problem, etc. Recently problems in biology and medicine such as cancer detection, medical imaging, and proteomics, and national security and bioterrorism such as detecting the spread of toxins through populations in the case of biological/chemical warfare are analyzed with the aid of spanning trees.

P. Biswas (✉) · A. Paul · A. Gogoi
Department of CSE, NIT-Agartala, Agartala 799046, India

P. Bhattacharya
Department of Mathematics, NIT-Agartala, Agartala 799046, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_39

A spanning tree of a connected undirected graph G = (V, E), is defined as a tree T consisting of all the vertices of the graph G. If the graph G is disconnected then every connected component will have a spanning tree T_i, the collection of which forms the spanning forest of the graph G. A graph may have many spanning trees.

Although a large variety of algorithms exists, that can compute the spanning tree from a given graph, determining it from a given degree sequence has not yet been tried. Popular algorithms were proposed by Kruskal [18] and Prim [17] that can successfully compute the minimal spanning tree of a given graph. In this paper we have applied a variant of Prim's [17] algorithm to determine the spanning tree directly from a given degree sequence of non-regular simple graphs.

A finite sequence d: d_1, d_2, d_3, …, d_n of nonnegative integers is said to be graphical if there exists some finite simple graph G, having vertex set V = {v_1, v_2, v_3, ..., v_n} such that each v_i has degree d_i ($1 \leq i \leq n$). Although it is quite easy to determine the degree sequence of a given graph, the converse procedure is potentially difficult. This problem is closely linked with the other branches of combinatorial analysis such as threshold logic, integer matrices, enumeration theory, etc. The problem also has a wide range of application in communication networks, structural reliability, stereochemistry, etc.

Two necessary conditions for a sequence to be graphical are: (1) $d_i < n$ for each i, and (2) $\sum_{i=1}^{n} d_i$ is even. However none of these are sufficient condition for the sequence to be graphic. The first known solutions were proposed independently by Havel [1] and Hakimi [2] in the mid 20th century. Although they provided separate proofs yet their work is jointly known as the Havel-Hakimi theorem.

With the advent of time, well known necessary and sufficient conditions were published. Erdős and Gallai [3], Ryser [4], Berge [5], Fulkerson et al. [6], Bollobás [7], Grőnbaum [8] and Hässelbarth [9] independently proposed the sufficient condition for a degree sequence to be graphic. Sierksma and Hoogeveen [10] listed the seven well known characterizations of a degree sequence and their equivalence. The works of Havel-Hakimi were further extended by Kleitman and Wang [14], and that of Erdős and Gallai by Eggleton [16] and Tripathi and Vijay [11]. Dahl and Flatberg [12] proposed a direct way of obtaining Tripathy and Vijay's result from a simple geometrical observation involving weak majorization. Tripathy and Tyagi [13] provided two elegant proofs of Havel-Hakimi and Erdős and Gallai.

In this paper we have taken the problem of determining a spanning tree from a given degree sequence provided the degree sequence is graphic and non-regular. Here the term non-regular implies that the given sequence is graphic and its realization gives a non-regular graph. We first check whether the input sequence satisfies the basic criteria for having a realization. Next we design a sufficient condition by constructing the adjacency matrix corresponding to the input sequence. Finally we apply BFS and DFS algorithm separately on the matrix to obtain a spanning tree.

In Sect. 2 we have given a brief definition of the problem. Section 3 describes our proposed approach while Sect. 4 illustrates the working of the proposed method. Section 5 discusses the results obtained. We draw a concise conclusion in Sect. 6.

2 Problem Definition

The proposed problem can be divided into two sub-problems. The first part is to determine whether the degree sequence is graphic and non-regular and to construct the corresponding adjacency matrix if the given degree sequence is graphic and non-regular. The second part is to construct the spanning tree from the resultant adjacency matrix.

- *Determination graphic sequences*

Let G = (V, E) be a finite simple graph with vertex set V and order n. Let each vertex v_i has a degree d_i where $1 \leq i \leq n$. Then the finite sequence d: $d_1, d_2, d_3, \ldots, d_n$ of nonnegative integers is called a degree sequence of the graph G. The problem statement can be formally stated as follows.

Given a finite degree sequence d: $d_1, d_2, d_3, \ldots, d_n$ *of non negative integers, whether there exists a graph G of order n with vertex set V such that each vertex* v_i *has a degree* d_i *where* $1 \leq i \leq n$.

- *Construction of spanning tree*

Let G = (V, E) be given graph on order n and size m. A spanning tree T of G is defined as a connected graph spanning all the vertices of the vertex set V with exactly $n - 1$ edges belonging to the edge set E. The construction of spanning tree requires us to eliminate $m - n + 1$ edges from the given graph G in order to obtain a sub-graph T of G such that it has order n and size $n - 1$.

3 Proposed Approach

In this section we demonstrate our proposed approach to solve the problem. In the first part, we present an algorithm that determines whether a given degree sequence is graphic by applying some basic criteria and constructing the adjacency matrix corresponding to the input sequence. In next part, we present how the BFS and DFS algorithm can be applied to obtain the spanning tree from the adjacency matrix obtained in the earlier part.

- *Determination graphic sequence through the construction of adjacency matrix*

Here we propose an algorithm that determines whether the given degree sequence is graphical by constructing the adjacency matrix corresponding to the given degree sequence. The degree sequence is stored in a vector *degree*[n] of length n. The vector *Allocated*[n] of length n helps us to determine whether a given position in the adjacency matrix can be allocated 1, based on checking the number of positions allocated in the jth column of the ith row with respect to the degree of vertex j. The $n \times n$ vector *Adj*[n][n] is the adjacency matrix that we construct using the algorithm. Once the entire matrix is constructed, the algorithm then checks whether the

resulted matrix is symmetric. The input to the algorithm is the given degree sequence in a non-increasing order and its output is the decision (i.e. graphic or non-graphic). If the decision of the algorithm is graphic, then we return the adjacency matrix *Adj*[*n*][*n*] thus constructed for further operations.

```
ALGORITHM-ADJ-CONST(d: d1, d2, d3, . . . , dn)
for i ← 1 to n
      Allocated[i] ← 0;
      degree[i] ← di;
   r ← degree[0];
   sum ← 0, flag ← 0, rflag ← 0;
   for i ← 1 to n
      if degree[i] ≥ n
         return Non-Graphic;
      if degree[i] ≠ r
         rflag ← rflag + 1;
      sum ← sum + degree[i];
   if sum mod 2 ≠ 0
      return Non-Graphic;
   if sum mod 2 = 0 and rflag = 0
      return Graphic and regular;
   for i ← 1 to n
      k ← degree[i];
      for j ← 1 to n
         if k > 0
            if i = j
               Adj[i][j] ← 0
            else
               if Allocated[j] < degree[j]
                  Adj[i][j] ← 1;
                     Allocated[j] ← Allocated[j] + 1;
                     k ← k – 1;
               else
                     Adj[i][j] ← 0;
          else
            Adj[i][j] ← 0;
      if k > 0
         return Non-Graphic;
   for i ← 1 to n
      for j ← 1 to n
         if Adj[i][j] ≠ Adj[j][i]
            flag ← flag + 1;
   if flag > 0
      return Non-Graphic;
   else
return Graphic and Adj[n][n];
```

- *Construction of the spanning tree from the adjacency matrix*

In this part we show how to apply BFS and DFS algorithm separately to obtain the spanning tree from the adjacency matrix *Adj*[*n*][*n*] obtained in the previous section. The algorithm takes the adjacency matrix as input and processes it to obtain the spanning tree.

3.1 Breadth First Search Algorithm for Spanning Tree Construction

Step 1. Add the starting vertex (say V_1) to T, insert it in the queue and mark it as "visited".
Step 2. If the queue is empty, then we are done. Otherwise let v be the vertex in the front of the queue.
Step 3. For each vertex V' of G that has not been visited yet and is adjacent to v (there might be none) taken in order of increasing subscripts, add vertex V' and edge (V, V') to T, insert V' in the queue and mark it as "visited".
Step 4. Delete V from the queue.

3.2 Depth First Search Algorithm for Spanning Tree Construction

We define a recursive process P for spanning tree construction through DFS.

Step 1. Add vertex V to T and mark it as "visited".
Step 2. If there is no vertex V' that is adjacent to V and has not been visited yet, then return. Otherwise, let V' be the first non-visited vertex that is adjacent to V.
Step 3. Add the edge (V, V') to T.
Step 4. Apply P to V'.
Step 5. Go to step 2 (*backtrack*).

4 Illustration

In this section we illustrate how the algorithm generates the spanning tree. Suppose that we enter the degree sequence: 4, 4, 2, 2, 2.

In the first part, the algorithm generates the adjacency matrix shown in Fig. 1.

The graph corresponding to the above matrix is shown below in Fig. 2.

Figure 3 shows the initial step of applying BFS and DFS algorithm separately to the above matrix (or graph). We assume that both the algorithms select A as the starting vertex.

Figures 4, 5, 6, 7 demonstrates how the BFS algorithm generates the spanning tree.

The execution of the DFS algorithm is shown below. Figure 1 represents the initial step when the algorithm chooses vertex A as the starting vertex. Starting from A, edge AB and vertex B are added (see Fig. 8). Then vertex D is added to the tree

	V_1	V_2	V_3	V_4	V_5
V_1	0	1	1	1	1
V_2	1	0	1	1	1
V_3	1	1	0	0	0
V_4	1	1	0	0	0
V_5	1	1	0	0	0

Fig. 1 Adjacency matrix

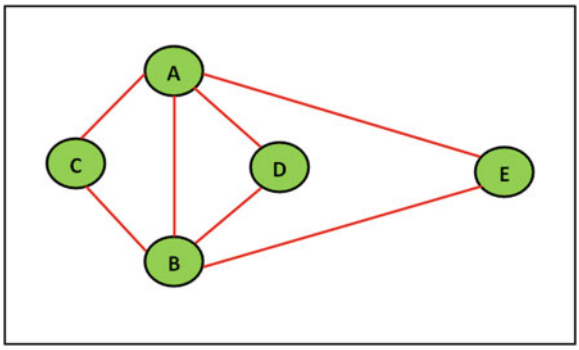


Fig. 2 Graph corresponding to Fig. 1

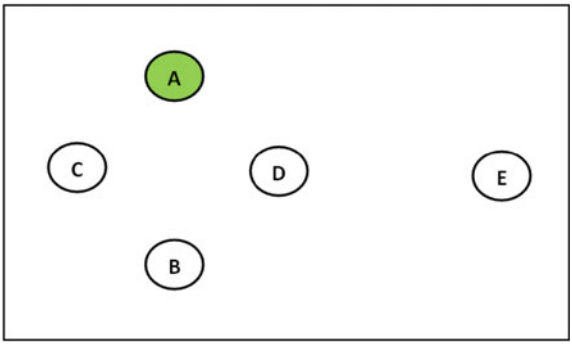


Fig. 3 Initial step of BFS

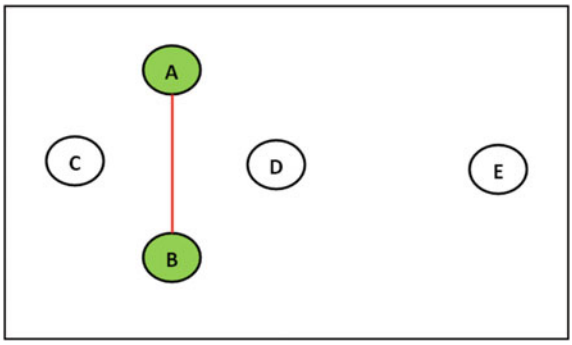


Fig. 4 Edge AB added

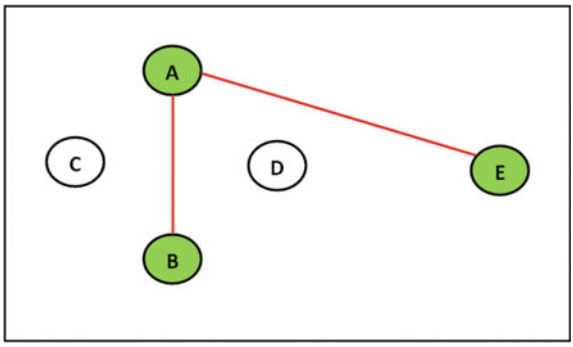


Fig. 5 Edge AE added

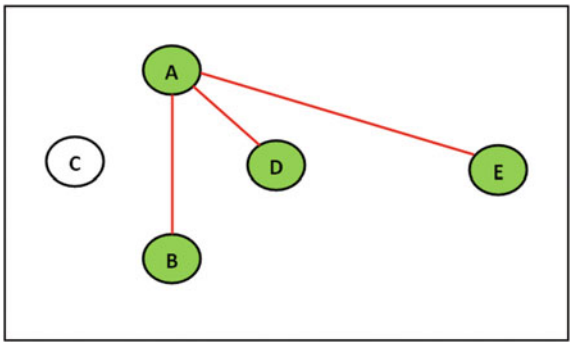


Fig. 6 Edge AD added

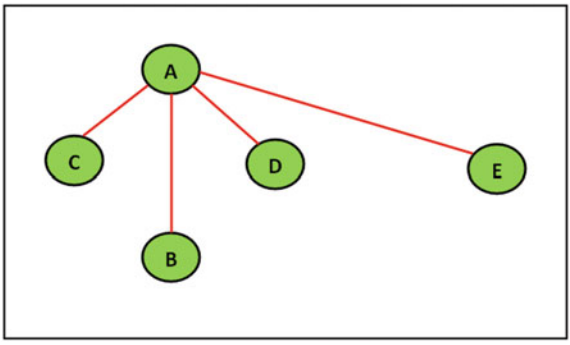


Fig. 7 Edge AC added

in a similar fashion (Fig. 9). However, at D, the edge DA cannot be added as it forms a cycle. Hence we backtrack to vertex B from where a path to vertex E emerges as shown in Fig. 10. Similarly from E, we again backtrack to B and take the new path to add C as shown in Fig. 11. At this point, the spanning tree is created and the algorithm terminates.

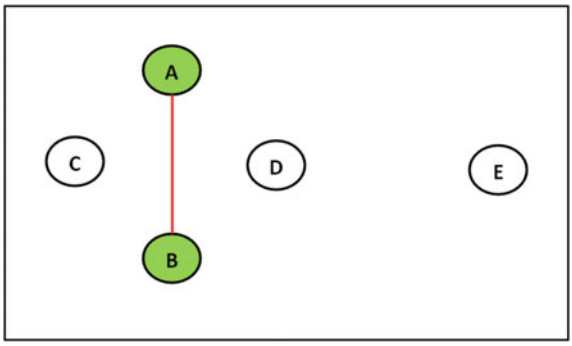


Fig. 8 Edge AB added

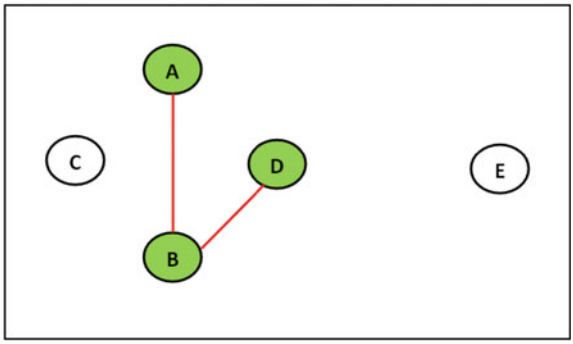


Fig. 9 Edge BD added

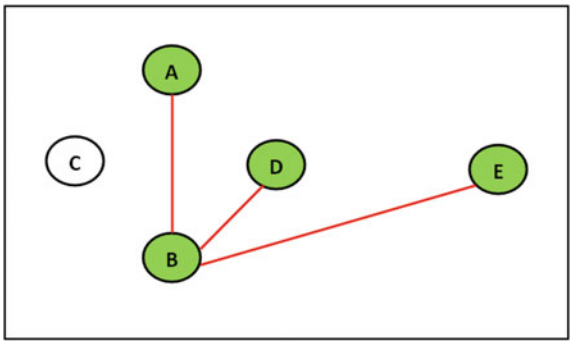


Fig. 10 New path BE

5 Results and Discussion

The output of the algorithm when applied to the test sequences of Table 1, clearly justifies the correctness of the algorithm. For every non-regular graphic sequence, we can obtain the spanning tree.

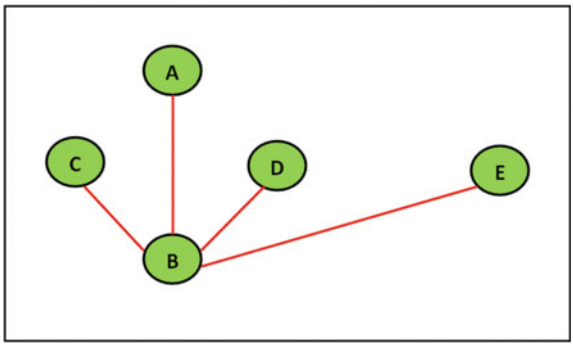
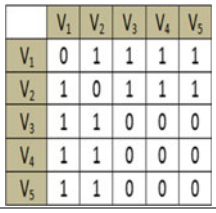
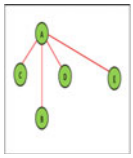
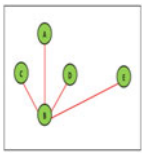
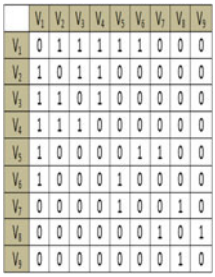
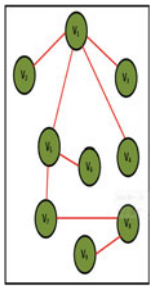
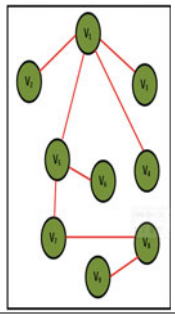


Fig. 11 New path BC

Table 1 Results of the proposed work

Test sequence	Matrix	BFS	DFS
4, 4, 2, 2, 2			
5, 3, 3, 3, 3, 2, 2, 2, 1			
7, 7, 4, 3, 3, 3, 2, 1	Non-graphic sequence	NA	NA

However an important observation is that the spanning tree generated by both the algorithms may not be unique. This is a well known fact and does not deviate from the problem statement. Another point about the algorithm is that the matrix generated in the first step can be linearly permuted over rows and columns to obtain different adjacency matrices thereby varying the spanning trees thus constructed. But in all the scenarios, the algorithm ends up giving the spanning tree corresponding to the final matrix fed to the second part of our approach.

6 Conclusion

From the above discussions, we can easily conclude that the algorithm is an efficient approach to construct a spanning tree directly from the degree sequence of non-regular graphs. It is easy to understand and simple to code. Further we can also obtain all the spanning trees corresponding to a given input sequence by applying the algorithm repeatedly over the various linearly permuted matrices and by choosing a new starting vertex. Hence, we can infer that it is a good approach to determine the spanning tree directly from the degree sequence of a graph.

References

1. Havel, V.: A remark on the existence of finite graphs. (Czech.) Casopis Pest. Mat. **80**, 477–480 (1955)
2. Hakimi, S.L.: On the realizability of a set of integers as degrees of the vertices of a graph. SIAM J. Appl. Math. **10**, 496–506 (1962)
3. Erdős, P., Gallai, T.: Graphs with prescribed degree of vertices. (Hungarian) Mat. Lapok. **11**, 264–274 (1960)
4. Ryser, H.J.: Combinatorial properties of matrices of zeros and ones. Can. J. Math. **9**, 371–377 (1957)
5. Berge, C.: Graphs and Hypergraphs. Elsevier, New York (1973)
6. Fulkerson, D.R., Hofman, A.J., McAndrew, M.H.: Some properties of graphs with multiple edges. Can. J. Math. **17**, 166–177 (1965)
7. Bollobås, B.: Extremal Graph Theory. Acedemic Press, New York (1978)
8. Grŏnbaum, B.: Graphs and complexes. Report of the University of Washington, Seattle, Math. 572B (1969) (private communication)
9. Hässelbarth, W.: Die Verzweighteit von Graphen. Match. **16**, 3–17 (1984)
10. Sierksma, G., Hoogeveen, H.: Seven criteria for integer sequences being graphic. J. Graph Theory **15**(2), 223–231 (1991)
11. Tripathi, A., Vijay, S.: A note on theorem on Erdős and Gallai. Discret Math. **265**, 417–420 (2003)
12. Dahl, G., Flatberg, T.: A remark concerning graphical sequences. Discret. Math. **304**, 62–64 (2005)
13. Tripathi, A., Taygi, H.: A simple criterion on degree sequences of graphs. Discret. Appl. Math. **156**, 3513–3517 (2008)
14. Kleitman, D.J., Wang, D.L.: Algorithms for constructing graphs and digraphs with given valences and factors. Discret. Math. **6**, 79–88 (1973)
15. Bondy, J.A., Murty, U.S.R.: Graph Theory with Applications. Elsevier, New York (1976)
16. Czajkowski, R., Eggleton, B.: Graphic sequences and graphic polynomials: a report. Infin. Finite Sets. **1**, 385–392 (1973)
17. Prim, R.C.: Shortest connection networks and some generalizations. Bell Syst. Technic. J. **4**, 53–57 (1957)
18. Kruskal, J.B.: On the shortest spanning subtree of a graph and the travelling salesman problem. Proc. Am. Math. Soc. **7**, 48–50 (1956)

A Novel Approach for Data Transmission Technique Through Secret Fragment Visible Mosaic Image

Asawari Chavan and Amrita Manjrekar

Abstract A new method is created secret fragment visible mosaic image for secure data communication. This mosaic image is created by composing small fragments of a given secret image and selected cover image. The cover image is arbitrarily selected and uses of this image to hiding of the secret image. Secret image and cover image is split into tiny fragments called tile image and target block respectively. Color variation process is used to hiding a tile image in a similar target block. Color transform algorithm is used for transforming color characteristic of each tile image to corresponding target blocks. The extraction of the secret image may be retrieved nearly lossless. If variable size of the secret image and cover image, then resizing the cover image. Customized metafile (CMF) is created to store the required information for retrieving the secret image. Hash value should be calculated to providing integrity of secret message. The implemented method is possible for big volume and different size of the secret image and cover image. The proposed method is feasible for secure and nearly lossless data embedding technique.

Keywords Color transformation · Secure data transmission · Secret fragment visible mosaic image · CMF

1 Introduction

In the globe of internet, information security is an important issue of storage, transmission of multimedia data. The globe increases availability and transmission of multimedia data in digital form. These digital data enclose confidential or secret information so that they can be protected from leakages during transmissions.

A. Chavan (✉) · A. Manjrekar
Department of Technology, Shivaji University, Kolhapur Vidya Nagar, Kolhapur 4164004, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_40

Many methods have been proposed for secure data transmission, out of which two common approaches are image encryption and data hiding. Image encryption makes use of the natural property of an image, such as high redundancy and strong spatial correlation. The encryption technique protects unauthorized access of the data. The encrypted image is a noisy image such that no one can obtain the secret image data without the correct key.

The steganography is a type of data hiding technique that provides protection of digital multimedia data transmission. The steganography contains hidden a digital image into another cover multimedia data such as image and video. Steganography technique is used when encryption is not acceptable. The purpose of steganography embeds secret data in reselected image.

Mosaic recounting a "process of creating pictures or attractive patterns by cementing simultaneously small pieces of stone, glass, or other hard materials of various colors". A mosaic is composed of smaller pieces or tiles, traditionally made of something like stone, rock or colored glass. Invented in ancient time, they are still used in many applications. Mosaic image is shown in Fig. 1. When the artwork is viewed from up close, the individual smaller elements are visible, yet when viewed at a distance the collection of strokes or tile blends together to yield the overall picture. The term mosaic contains a number of small images called "Tile" and then placed these two single images called "Target Image".

In proposing a system selected by the target, image is arbitrary. Compare the size of given secret image and selected target (cover) image. Variable size of the secret image and the target (cover) image then it is the resize of the target image. Divide the secret image and target image into tiny rectangular shape fragments called tile image and target block, respectively. Sort the tile images and target blocks in ascending order in based on the average standard deviation. Fit tile image into a target block based on similarity of color variations and transformed color characteristics of the tile image to target block. The generated secret fragment visible mosaic image is visually similar to target image. Hash value should be calculated for generating mosaic image. MD-5 hash algorithm uses to compute the hash value. Metafile created, which store the hash value and secret image recovery information about that image.

All recent techniques do not attend to the problem of the size of the secret image and target image. Basically, this proposed method is flexible for variable size of

Fig. 1 Mosaic image

secret image and target image and also calculate the hash value. These methods support the consistency, confidentiality and integrity of the secret image data.

The rest of this paper contains a related work in Sect. 2. Section 3 contains the idea of the secure data transmission. Section 4 contains the algorithm of proposed work. Section 5 shows the experimental results of implemented work and last Sect. 6 contain conclusion of proposed work.

2 Related Work

Battiato et al. [1] Proposed sorting of mosaic images into four types, which includes crystallization mosaic, ancient mosaic, and photo-mosaic and puzzle image mosaic. The first two types are obtained from stale a source image into tiles with different colors, sizes, and rotations. Reconstructing the image is properly painted the tiles. Hence, both the types of mosaic called tile mosaics. The further two types are mosaics obtained by appropriate images from a database to cover the image as an assigned source image. Hence, both are called multi-picture mosaics.

Crystallization mosaic image created using Voronoi diagram by Haeberli. [2] Voronoi diagram generated by placing blocks at random sites. Blocks are filled colors based on the contents of the original image.

"Jigsaw image mosaics" is developed by Kim and Pellacini [3]. Generation of puzzle image mosaic called jigsaw image mosaic, created by many random shapes of tiles chosen from a database.

Battiato et al. [4, 5] created ancient mosaic images with gradient vector flow algorithm. Which track for the majority of vital edges in the original image and capitalize on the covered mosaic area. Battiato and Puglisi [6] also proposed "3D ancient mosaics" image.

Lai and Tsai [7] create a new type of art, image, called Secret-fragment-visible mosaic image. This is a type of the mosaic image by composing tile image and similar target block. The selection of the target image is used, large size of the image database.

An alternating method is developed by Ya-Lin Lee and Wen-Hsiang Tsai. They proposed "A new secure image transmission technique via secret-fragment-visible mosaic images of nearly reversible color transformations" [8]. This proposed method used the same size of the secret image and target image and transform color characteristics using reversible color transform technique.

3 Ideas of the Secure Data Transmission

The Proposed method creates a secret fragment visible mosaic image using color transformation and customized metafile (CMF). Mosaic image is complete using tiny fragments of secret image and target image. The system includes two main

Stages: (1) Mosaic image generation (2) Secret image recovery. Figure 2 is given away to generating mosaic image and Fig. 3 is given away to extraction of the secret image.

In stage, I create a mosaic image using secret image and target image. Secret image is embedded into the target image. The stage I consists four steps: (1) Resizing the target image. (2) Mosaic image generation. (3) Metafile creation. (4) Hash value calculation.

In stage II, extract nearly lossless of the secret image from mosaic image. In stage II includes two steps: (1) Compare hash value. (2) Extraction of the secret image with the help of the metafile.

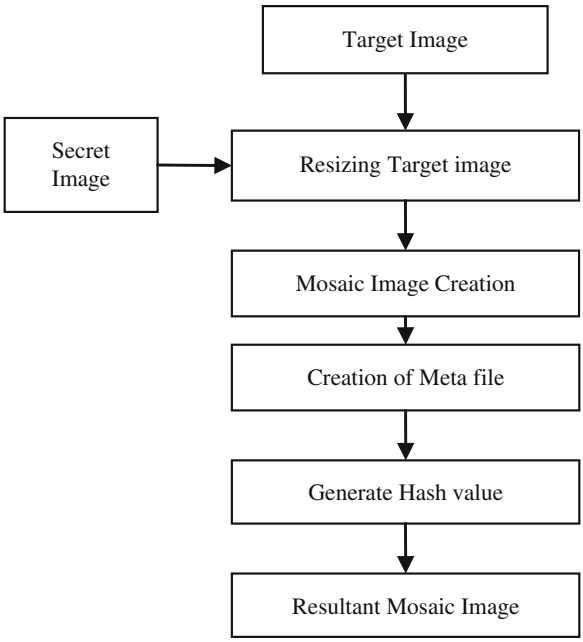


Fig. 2 Embedding the secret image and mosaic image creation

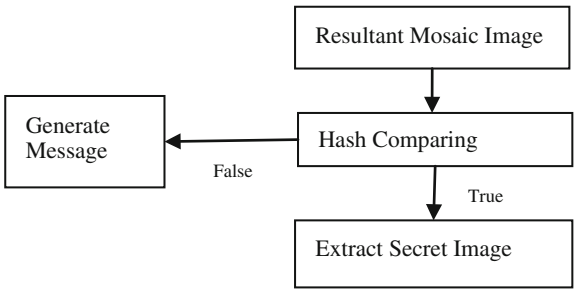


Fig. 3 Extraction of secret image recovery information and secret image

3.1 Generation of Secret Fragment Visible Mosaic Image

Use of secret image S and target image T for creation of the mosaic image.

3.1.1 Resizing Target Image

Secret image and target image are different size images, and then resize the target image.

3.1.2 Color Transform Technique

Color transformation technique, is used fitting the each tile image into a target block. This technique proposed by Reinhard et al. [9]. First secret image and target image are divided into number tile image $\{T_1, T_2,..., T_n\}$ and target blocks $\{B_1, B_2 ..., B_n\}$. Each T_i and B_i contain the number of pixels $\{p_1, p_2 ..., p_n\}$ and $\{p'_1, p'_2 ..., p'_n\}$ respectively. Each pixel contains three color values (r, g, b). Tile image pixel denotes (r_i, g_i, b_i) and target blocks pixel denote $(r'_i, g'_i b'_i)$. Calculate the mean of the tile image and target blocks for each color channel of the pixel using formula (1).

$$\mu_c = \frac{1}{n}\sum_{i=1}^{n} c_i \quad \mu'_c = \frac{1}{n}\sum_{i=1}^{n} c'_i \tag{1}$$

Calculate the standard deviation of each T and B using formula (2)

$$\boldsymbol{\sigma_c} = \sqrt{\frac{1}{\boldsymbol{n}}\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}}(\boldsymbol{c_i} - \boldsymbol{\mu_c})^2}\boldsymbol{\sigma'_c} = \sqrt{\frac{1}{\boldsymbol{n}}\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}}(\boldsymbol{c'_i} - \boldsymbol{\mu'_c})^2} \tag{2}$$

c_i and c'_i represent the C-channel standards of pixels p_i and p'_i respectively.

$$c''_i = q_c(c_i - \mu_c) + \mu'_c \tag{3}$$

where $q_c = \sigma'_c/\sigma_c$ is the standard deviation quotient and c = r, g or b. c''_i Denotes a new pixel value, which is a camouflage of tile image into the target block.

$$c_i = \frac{1}{q_c}\left(c''_i - \mu'_c\right) + \mu_c \tag{4}$$

Above (4) formula is used for recovering the original image

3.1.3 Creation of Metafile

Metafile stored as in binary sequences. First binary sequences contain tile image to target block assignment with their block number, mean of all three color channels and average standard deviation with attached followed by q_c for three colors of the tile image. This is continuing for all sorted tile and assigned target blocks.

In the second part of binary sequences contains counting table entries. Each record of counting table entry contains residual pixel information. This information contains X and Y position of residual pixel and color code.

Finally, metafile binary string appended by 128-bit hash value of embedded image of achieving integrity or consistency of the image.

4 Algorithm of Proposed Work

Following algorithms are used for mosaic image generation and secret image recovery.

Stage I algorithm: *Mosaic image generation*
Input: Secret image *S*, target image *T*.
Output: Mosaic image *F*, Metafile *MD*.

Step 1: Compare the size of the target image *T* and secret image *S*. If the size of target image is differing from the secret image *S*, change the size of *T*.

Step 2: Divide the secret image *S* into *n* tile images $\{T_1, T_2, \ldots, T_n\}$ as well as the target image *T* into *n* target blocks $\{B_1, B_2, \ldots, B_n\}$.

Step 3: Compute the means and the standard deviations of each tile image *Ti* and each target block *Bj* for the three color channels according to (1) and (2) and compute consequently the average standard deviations for *Ti* and *Bj*, respectively.

Step 4: Sort the tile images in the set $S_{tile} = \{T_1, T_2, \ldots, T_n\}$ and the target blocks in the set $S_{target} = \{B_1, B_2, \ldots, B_n\}$ based on the calculated average standard deviation values of the blocks. Robust the blocks in the sorted S_{tile} to the sorted S_{target} in a 1-to-1 approach; and rearrange the mappings according to the indices of the tile images, resulting in a mapping sequence *L* of the form: $T_1 \rightarrow B_{j1}$, $T_2 \rightarrow B_{j2}$, …, $T_n \rightarrow B_{jn.}$

Step 5: construct a mosaic image *F* by appropriate the tile images into the corresponding target blocks according to *L*.

Step 6: Create a *counting table TB* with 256 entries, each an index value contain residual pixel information. Information contains an index of residual pixel and the value of that pixel.

Step 7: Create a metafile. Metafile contains each index of assigning of the tile image into target image $T_i \rightarrow B_{ji}$ in the binary sequence. Binary string represent the mean μ_c and μ'_c of T_i and B_{ji} respectively. Next bit represents the standard deviation quotient $q_{c,}$ appearing in (3).

Step 8: Concatenate all bits string M_i of all T_i in *metafile* sequentially. Finally, appended 128 bit hash value.

Stage II algorithm: Secret image recovery
Input: A mosaic image F and the metafile MD.
Output: the secret image S.

Step 1: Extract information from metafile MD. The binary string split into an array. Each string is read from an array according to the assignment of indexing. In extraction of string obtain the following data items: (1) Hash value (2) Assignment index of tile to target block (3) The calculate means of each T_i and B_i and the related standard deviation quotients of all color channels; and 4) The overflow/underflow residual pixel values in T_i decoded by the counting table.

Step 2: Recover one by one indexing array the tile images T_i, $i = 1 \ldots n$, of the preferred secret image S by the following steps: (1) Use the extracted means and related standard deviation quotients to recover the original pixel values in T_i according to (4). (2) Use the extracted residual pixel index and pixel to assign the value of underflow/overflow pixels (3) Get the results as the final pixel values, resulting in a final tile image T_i.

Step 3: Compose all the final tile images to form the desired secret image S as output.

5 Experimental Results

In this section experimental results have shown in following figure. Figure 5 show is a secret image which in embedded into target image in Fig. 4. Figure 6 is shown as generated secret fragment visible mosaic image with 4 * 4 tiny fragments. The

Fig. 4 Target image

Fig. 5 Secret input image

Fig. 6 Mosaic image 4 * 4

Fig. 7 Mosaic image 8 * 8

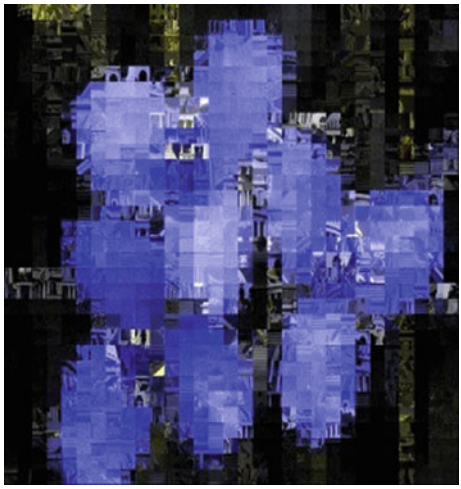


Fig. 8 Mosaic image 16 * 16

Fig. 9 Extracted from 4 * 4

Fig. 10 Extracted from 8 * 8

Fig. 11 Extracted from 16 * 16

Fig. 9 shows the extraction of the secret image from Fig. 6. We have also conducted an experiment by varying the size of tiny fragments. Figures 7 and 8 generated mosaic image using a secret image in Fig. 5 and target image in Fig. 4 with 8 * 8 and 16 * 16 tiny fragments respectively. Figures 10 and 11 are shown extraction of the secret image from Figs. 7 and 8 respectively.

6 Conclusion

Our system has developed a new secure image transmission method with a creation of secret fragment visible mosaic image. Secret fragment visible mosaic image is generated using color transformation and metafile code execution techniques. Extracted image is same as that of the secret image. This method is possible for huge data and variable size of the secret and target image. The system provides confidentiality and integrity of the secret image.

References

1. Battiato, S., Di Blasi, G., Farinella, G.M., Gallo, G.: Digital mosaic framework: an overview. Eurograph.—Comput. Graph Forum **26**(4), 794–812 (2007)
2. Haeberli, P.: Paint by numbers: abstract image representations. In: Proceedings of SIGGRAPH, Dallas, TX, 1990, pp. 207–214
3. Pellacini, J.K.F.: Jigsaw image mosaics. In: Proceedings of SIGGRAPH, San Antonio, TX, pp. 657–664, July 2002
4. Battiato, S., Guarnera, C., DiBlasi, G., Gallo, G., Puglisi, G., Bubak, M. (ed) et al.: A novel artificial mosaic generation technique driven by local gradient analysis. In: Proceedings of ICCS, Crakov, Poland, vol. 5102, pp. 76–85, June 2008

5. Battiato, S., Blasi, G.D., Gallo, G., Guarnera, G.C., Puglisi, G.: Artificial mosaic by gradient vector flow. In: Proceedings of Eurographics, Crete, Greece, pp. 53–56, April 2008
6. Battiato, S., Puglisi, G.: 3D ancient mosaics. In: Proceedings of ACM International Conference on Multimedia, Florence, Italy, pp. 1751–1753 Oct (2010)
7. Lai, I.J., Tsai, W.H.: Secret-fragment-visible mosaic image—a new computer art and its application to information hiding. IEEE Trans. Inf. Forens. Secur. **6**(3), 936–945 (2011)
8. Lee, Y.-L.., Student Member, IEEE, Tsai, W.-H., Senior Member, IEEE: A new secure image transmission technique via secret-fragment-visible mosaic images by nearly reversible color transformations. IEEE Trans. Circuits Syst. Video Technol. **24**(4), April 2014
9. Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE Comput. Graph. Appl. **21**(5), 34–41, Sept–Oct 2001

Edge Preserved De-noising Method for Medical X-Ray Images Using Wavelet Packet Transformation

B. Rajith, Monika Srivastava and Suneeta Agarwal

Abstract X-ray image is one of the prominent modality of medical imaging used in medical diagnosis. This may be corrupted with Gaussian noise due to thermal fluctuations during its acquisition. For reducing these noises a method is applied which combines the Anisotropic Diffusion filter with an edge preserved Wavelet Packet Transformation. Here the edges are detected in each sub-band using Sobel edge detection operator and preserved by excluding these edge coefficients during hard thresholding. This method has proposed a new technique to calculate a threshold value for each sub-band of Wavelet Packet. The quality metrics SNR, RMSE, SSIM, Precision, Accuracy and etc. are used to measure the performance of this method and shows that this approach is better as compared to other noise reduction methods like Adaptive Median Filter, simple Anisotropic Diffusion and simple Wavelet Packet Transformation.

Keywords Wavelet packet transformation · Anisotropic diffusion · Edge detection · Conduction coefficient · Wavelet coefficient · Precision

1 Introduction

Now a day's medical imaging is a promising technology used in medical diagnosis. X-ray, Computed Tomography (CT), SPECT (Single Photon Emission Computed Tomography), PET (Positron Emission Tomography), Scintigraphic, MRI (Magnetic Resonance Image), Ultrasound images etc. are the components of

B. Rajith (✉) · M. Srivastava · S. Agarwal
Department of CSE, MNNIT, Allahabad, India
e-mail: rajitha@mnnit.ac.in

M. Srivastava
e-mail: Srivastava.monika14@gmail.com

S. Agarwal
e-mail: suneeta@mnnit.ac.in

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_41

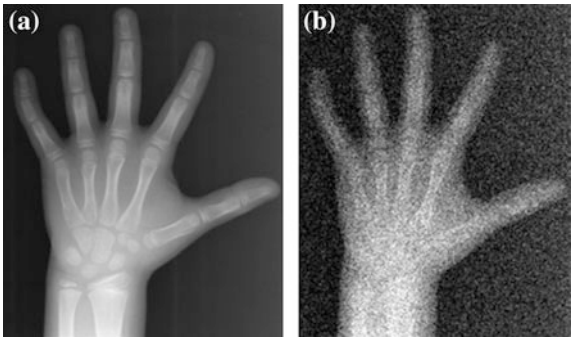


Fig. 1 **a** Original X-ray image of palm. **b** Noisy image corrupted with white noise of zero mean and 0.01 variance

medical imaging. X-ray images are used to diagnose mainly arthritis, bone tumors, fractures, inherited bones abnormalities and many more diseases in human beings. X-ray imaging is done by projecting diverse beam of X-rays, generated by X-ray generators (like; betatrons, synchrotron etc.) towards the human body [1]. This imaging modality is prone to noise because of thermal instability of many unified electronic components used in X-ray generators and detectors during acquisition process. X-ray images may be severely corrupted with Gaussian noise, also known as AGWN (Additive Gaussian White Noise). It is an additive noise, which is consistently spread over the entire image. Here each pixel is composed with actual pixel intensity and random Gaussian noise value. The Gaussian noise has Probability Distribution Function (PDF) as stated in Eq. (1) [2]. Since the presence of noise makes appearance of an image grainy, mottled and unclear. Due to overlapping of noise term over image data important information may be lost. Whereas for better diagnosis image should be properly object identifiable and for this image must be noise free. The appearance of noise is shown in Fig. 1b. So for doing better diagnosis we need to reduce noise from digitized medical X-ray images. There are many conventional noise reduction methods like Mean, Median, Adaptive wiener filters, and Gaussian low pass filters etc. for reducing Gaussian noise. These filters are good enough in reduction of noise but, with a loss in edge pixels, which play a vital role in medical diagnosis. Hence it is necessary not only to reduce noise from X-ray images but also preserve edge.

$$P(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-\mu)^2/2\sigma^2} \tag{1}$$

Ye Fang et al. have proposed a technique WCLA, to identify noise distribution in high frequency sub-bands using Multi Resolution Analysis (MRA) and then Bayesian Rule is applied to reduce noise [3]. This method has very low SNR value. Luo et al. [4] has proposed a combination of mean and median filter with thresholding to reduce white noise mixed with salt and pepper noise from x-ray image. The median filter is improved with a threshold calculation. Some researchers have been observed that the second order Partial Differential Equations (PDEs) are very

prominent technique for noise reduction. These PDEs are considering local homogeneity for doing noise suppression in homogeneous region while performing sharpening at edges using spatial gradients, like nonlinear Anisotropic Diffusion filter proposed by Perona-Malik [5]. This method is further improved by [6]. From past two decades it has been observed that Multi Resolution Analysis is a good choice for de-noising X-ray images as it analyses data at different levels of resolution. There are many narratives about image de-noising using wavelet transformation like ProbShrink, BayesShrink, VisuShrink, SureShrink and etc. [7].

Many researchers have combined the two methods PDEs based methods and Wavelet based algorithms due to their great performances in de-noising the image with edge preservation. Such as B. Hong et al. has proposed a method using Regularized P-M Diffusion in wavelet domain [8]. A method called Wavelet Embedded Anisotropic Diffusion (WEAD) was proposed by Rajan and Kaimal [10]. Here the P-M method is applied at every detail coefficients of Wavelet sub-bands (HL-horizontal, LH-vertical and HH-diagonal). They have used soft thresholding by minimising the Bayesian Risk. The Bayesian Shrinkage has been applied over nonlinearly diffused signals. It has been observed that the performance of these two algorithms is good but can further increased by replacing the Wavelet transformation with Wavelet Packet Transformation (WPT). Such as an approach proposed by Mahalakshmi and Anand [9] used the adaptive Wavelet Packet transformation for reducing Gaussian noise with an improvement in the calculation of adjacent window size for every subblock.

This paper's objective is to improve the performance of WPT by doing edge preservation in Wavelet Packet's subblocks with the help of edge detection operators such as Sobel, Prewitt, Roberts etc. In this paper Sobel operator is used for detecting edges in HL, LH and HH sub-bands. Then a hard thresholding with improved threshold calculation method applied at each non edge coefficients for each sub-band. For better edge detection at wavelet packets initial image must be noise free. For which Anisotropic Diffusion Method is used. In Sect. 2 these techniques are discussed separately. Section 3 explains proposed work and Sect. 4 discuss the experimented results and performance comparison with different methods of de-noising.

2 Proposed Approach

The proposed work was aimed to improve performance using wavelet packet transformation which not only reduces the noise but preserves the edge as well. In this technique firstly noise is suppressed by applying Anisotropic Diffusion filter and the image is decomposed using image using WPT up to third level. After decomposition thresholding is applied to reduce noise from each sub-band with edge preservation. As it's known that through WPT edges can be obtained from 3 sub-bands HL, LH and HH. The coefficients of HL sub-band gives horizontal edges, coefficients of LH sub-band gives vertical edges and the coefficients of HH

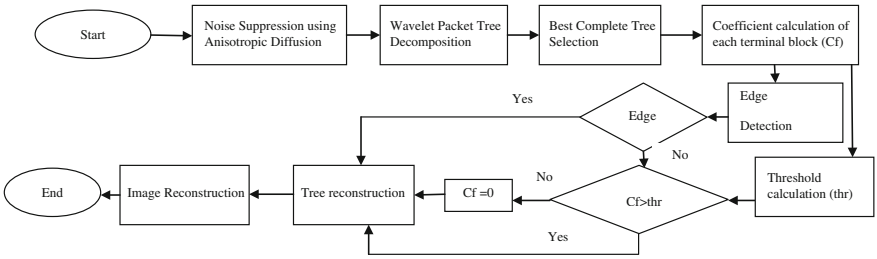


Fig. 2 Flow chart of proposed approach

sub-band gives diagonal edges of an image. Now these are being used for edge preservation in each sub-band by smoothing noise terms. An edge detection technique is used to identify edges and their locations in these sub-bands and thereafter thresholding is applied at non-edge locations. Such edges can be obtained by Sobel, Prewitt and Canny. Here Sobel edge detection operator is used to detect edges with respective directions and sub-band like, Sobel with horizontal direction in HL sub-band, vertical direction in LH sub-band and diagonal direction in HH sub-band. The Fig. 2 shows flow chart of above stated process.

2.1 Proposed Algorithm

Step 1. Load image and converted it in gray scale image for further processing.
Step 2. Suppression of noise using Anisotropic Diffusion Filter: The proposed edge detection requires wavelet packet decomposition where in the image needs to be noise free hence anisotropic diffusion is applied initially.
Step 3. Wavelet Packet Decomposition is done up to third level which results in a tree structure by using Coiflets (coif2) wavelet family. This is used because of its symmetrical and bi-orthogonal nature. As it simplifies the edge detection process and provides better fault tolerance in decomposition.
Step 4. Selection of Best possible Complete Sub Tree: Here Matlab in-built function is used for choosing best complete sub-tree.
Step 5. Computation of Terminal Nodes of the tree because all processing will be done at terminal nodes.
Step 6. Thresholding: Here a new thresholding method is defined for calculating threshold value.
Step 7. Edge detection at each node: Computed coefficients of each node are used for edge detection using sobel operator for horizontal, vertical and diagonal sub-bands. This results in a binary matrix (E).
Step 8. Noise Reduction at non edge location: From the result of step 7 if location E (i, j) is edge point then Hard Thresholding is applied to that coefficient of corresponding sub-block. Else it is left as such.

Step 9. Reconstruction of wavelet Packet Tree: Reconstruct the wavelet packet tree with approximation and modified coefficients of each node.
Step 10. Reconstruction of Image: The de-noised image is reconstructed from new WP Tree by using wavelet Packet reconstruction method.

2.2 Detailed Description of Proposed Algorithm

2.2.1 Anisotropic Diffusion Filter

The Perona-Malik's second order Partial Differential equation for image F is given by Eq. (2) Where *div* is divergence operator, ∇ is gradient and Δ is laplacian operator and c(•) is an conduction coefficient, which is set to 1 if point (x, y) is in interior region, otherwise it is 0 if (x, y) is the boundary point. The edges are identified by taking the gradient of brightness function, which are the intensity differences of $F(x, y)$ with its 4 neighborhoods in north, south, east and west directions. And it is observed that if these differences are zero than the point (x, y) is an interior point otherwise it is an edge point. Now the conduction coefficient can be defined as in Eq. (3). Now to calculate c(•) for each direction north, south, east and west the magnitude of intensity differences are computed as is described in Eq. (4).

$$F_n = div(c(x,y,n)\nabla F) = c(x,y,n)\Delta F + \nabla c \cdot \nabla F \tag{2}$$

$$c(x,y,n) = g(\|\nabla F\|) = e^{(-(\|\nabla F\|/K)^2)} \tag{3}$$

$$\left.\begin{aligned} \nabla F(x,y)_N &= F(x-1,y) - F(x,y) \\ \nabla F(x,y)_s &= F(x+1,y) - F(x,y) \\ \nabla F(x,y)_E &= F(x,y+1) - F(x,y) \\ \nabla F(x,y)_W &= F(x,y-1) - F(x,y) \end{aligned}\right\} \tag{4}$$

Here K is a constant which can be fixed manually or can be computed by taking histogram of magnitudes of gradients and then K is set to 90 % value of integrals of these histograms. Now the Intensity of pixel F(x, y) is calculated by taking a variation of Eq. (2) described in Eq. (5). Where λ is used for stabilizing this method and it ranges from $0 \leq \lambda \leq 0.25$ [5].

$$\begin{aligned} F(x,y)_{n+1} = F(x,y)_n + \lambda[&C_N \cdot \nabla F(x \cdot y)_N + C_S \cdot \nabla F(x \cdot y)_S + C_E \cdot \nabla F(x \cdot y)_E \\ &+ C_W \cdot \nabla F(x \cdot y)_W] \end{aligned} \tag{5}$$

2.2.2 Wavelet Packet Transformation

Simple Wavelet Transformation may provide less advantageous results in some cases where important information is available at higher-level resolution because the decomposition is done only approximation coefficient at every level which is having low-frequency. Hence it may not be able to extract necessary information from image. This can be achieved by applying Wavelet Packet Transformation (WPT). WPT is generalized from of orthogonality, which decomposes signal into low-frequency and high-frequency components which will further decomposed in low-frequency and high-frequency coefficients by using analysis and synthesis filters, respectively [9]. The low-pass and high-pass filter pair h(n) and g(n) can be applied to determine the approximation and detail coefficients described Eqs. (6–9).

$$A(u,v) = \sum_{x,y=-\infty}^{\infty} h(x-2u)h(y-2v)F(x,y) \tag{6}$$

$$D_H(u,v) = \sum_{x,y=-\infty}^{\infty} h(x-2u)g(y-2v)F(x,y) \tag{7}$$

$$D_V(u,v) = \sum_{x,y=-\infty}^{\infty} g(x-2u)h(y-2v)F(x,y) \tag{8}$$

$$D_D(u,v) = \sum_{x,y=-\infty}^{\infty} g(x-2u)g(y-2v)F(x,y) \tag{9}$$

$$where\ u = 1 \ldots N\ and\ v = 1 \ldots M$$

where $a(u,v)$ is approximation coefficient driven by down-sampling the 2D Signal $F(x, y)$ and the $D_H(u, v)$, $D_V(u, v)$, and $D_D(u, v)$ are detail coefficients in three directions horizontal, vertical and diagonal obtained by up-sampling the signal $F(x, y)$. Further decomposition may be applied on both approximation and detail components. The reconstruction can be done by using Eq. (10). WPT can be represented in tree structure form and it is required to extract best complete sub tree for selecting and classifying coefficients for signal analysis. Figure 3 shows full decomposed WP Tree up to second level [11].

$$\begin{aligned} F(u,v) = &\sum_{x,y=-\infty}^{\infty} h(u-2x)h(v-2y)A(x,y) + \sum_{x,y=-\infty}^{\infty} h(u-2x)g(v-2y)D_H(x,y) \\ &+ \sum_{x,y=-\infty}^{\infty} g(u-2x)h(v-2y)D_V(x,y) + \sum_{x,y=-\infty}^{\infty} g(u-2x)g(v-2y)D_D(x,y) \end{aligned} \tag{10}$$

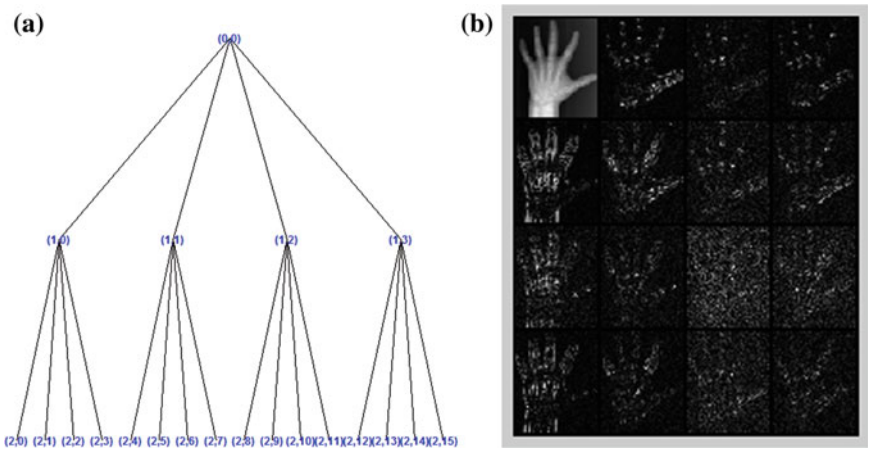


Fig. 3 **a** Complete wavelet packet tree of palm X-ray image up to second level. **b** All decomposed sub-bands of this image

2.2.3 Edge Detection Technique

Edges are abrupt changes in pixel intensity which can be identified by taking gradient of image. There are many edge detection operators available such as Sobel, Prewitt, Canny, Roberts and Laplacian etc. Here Sobel operator is used to retrieve edges in horizontal and vertical directions by convolving the image with 3 × 3 convolution kernels pair G_x and G_y, shown in Fig. 4a, b. These two are applied separately to the image to find edges vertically and horizontally [12].

2.2.4 Calculation of Threshold Value

Threshold value is calculated by considering a 3 × 3 window of coefficients (shown in Fig. 4c) in each sub-block. The steps used for threshold selection is as follows:

(a)

-1	0	+1
-2	0	+2
-1	0	+1

(b)

+1	+2	+1
0	0	0
-1	-2	-1

(c)

ncf(1)	ncf(2)	ncf(3)
ncf(4)	cf (i,j)	ncf(5)
ncf(6)	ncf(7)	ncf(8)

Fig. 4 **a** G_x, **b** G_y, **c** 3 × 3 window mask of centered coefficient cf(i, j)

i. Calculate mean of eight neighborhoods of centered coefficient located at cf(i, j)position in 3 × 3 window.

$$m(i,j) = 1/8 \sum_{p=1}^{8} ncf(p) \tag{11}$$

where ncf(p) corresponds to neighborhood coefficient of cf(i, j), where p ranges in [1, 8].

ii. Then a variance of coefficient cf(i, j) is calculated by using Eq. (12).

$$v(i,j) = \sqrt{\frac{1}{8}\sum_{p=1}^{8}(ncf(p) - m(i,j))^2} \tag{12}$$

iii. Now the difference between the values of coefficient cf(i, j) with its eight neighbors are estimated.

$$d(p) = cf(i,j) - ncf(p), \tag{13}$$

iv. Calculation of average differences and its deviation with each difference d(p) by using Eqs. (14) and (15).

$$d_{ave} = \frac{1}{8}\sum_{p=1}^{8}|d(p)| \tag{14}$$

$$dev = \frac{1}{8}\sum_{p=1}^{8}|d(p) - d_{ave}| \tag{15}$$

v. Then the threshold value for each 3 × 3 window is calculated as follows,

$$T(i,j) = (d_{ave} + dev) * \{m(i.j)/m(i,j) + v(i,j)\} \tag{16}$$

vi. This T(i, j) calculation is repeated for each cf(i, j) by a 3 × 3 overlapping mask.

vii. Now the final threshold value *Thr* for the entire subblock is taken as maximum from the set of T(i, j)'s.

$$Thr = \max(T) \tag{17}$$

This threshold value is applied to reduce noise from HL, LH and HH sub-bands.

3 Result Analysis and Performance Measures

The proposed approach is compared with Adaptive Median Filter (AMF), Simple Wavelet Packet Transformation (WPT), Simple Anisotropic Diffusion (AD) and the Combination of AD and WPT.

3.1 Dataset Collection

The X-ray images are collected from licensed dataset provided by IRMA (Image Retrieval in Medical Applications). They provide different datasets on non-commercial purpose for various imaging research. Among them for this research the image dataset comprising the five different body parts namely hands, chest frontal view, chest lateral view, head and pelvis were gathered. These images were collected from Department of Diagnostic Radiology at the RWTH Aachen University Hospital. In this paper the results were evaluated on three categories of images like Palm, Chest and Pelvis images of about 50 samples.

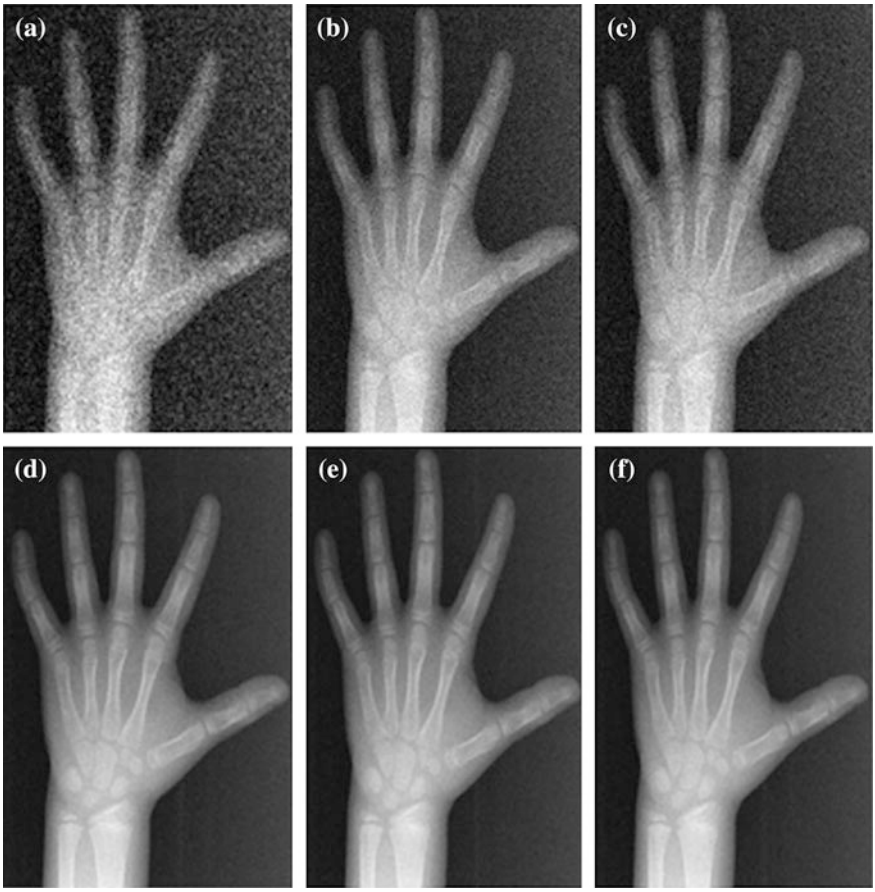


Fig. 5 **a** X-ray image of palm corrupted with Gaussian noise with zero mean and 0.01 variance, **b** de-noised image with adaptive median filter, **c** de-noised image with simple WPT, **d** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25), **e** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25) and WPT, **f** de-noised image using proposed approach

3.2 Result Analysis

The proposed approach was simulated using MATLAB. The Fig. 5a is the noisy X-ray image of palm which was corrupted with White Noise of zero mean and 0.01 variance. The AMF is applied to this noisy image that reduces very less noise and also blurred the edges as shown in Fig. 5b. The resultant image of simple WPT is shown in Fig. 5c, from this it is clear that the image is still having lots of noise. The AD filter is applied 10 times with K = 120 and λ is having value 0.25, whose result is shown in Fig. 5d which is better than the results of AMF and WPT but it is less clear than the proposed method. The Fig. 5e is the result of combination of AD and WPT, where AD is applied before WPT with K = 120, λ = 0.25 and iteration = 10, from this it is clear that the resultant image is less clear than AD. From Fig. 5f. It is clearly observed that the proposed approach have removed almost all the noise from noisy image by applying AD (with K = 120, λ = 0.25 and iteration = 10) and improved edge preserved WPT. Figures 6 and 7 are providing similar interpretation for X-ray images of Chest and Pelvis. Similarly, Figs. 8, 9 and 10 are showing the de-noised images of palm, chest and pelvis of the corrupted image with White noise of zero mean and 0.05 variance, respectively.

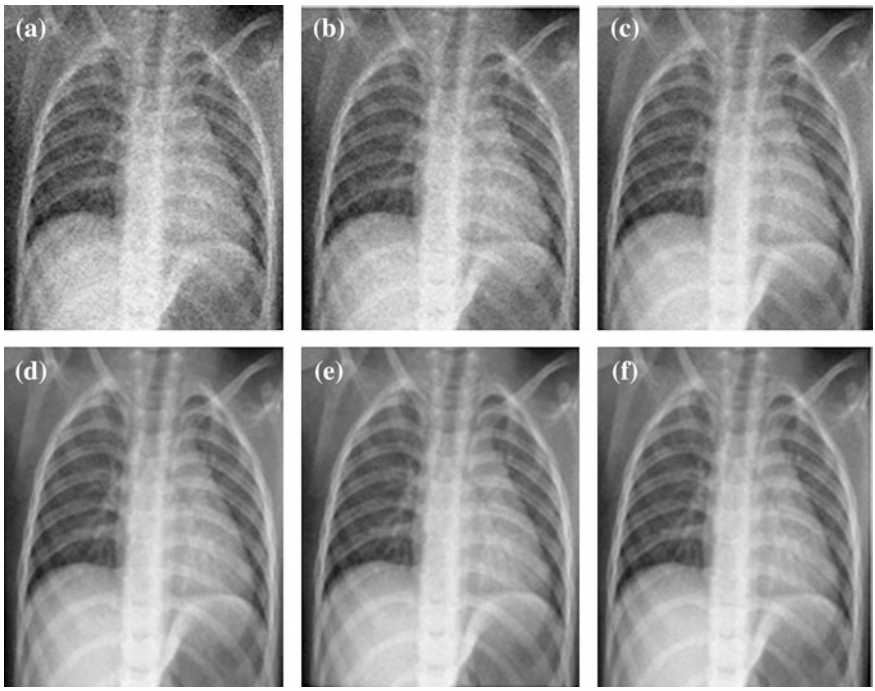


Fig. 6 **a** X-ray image of chest corrupted with Gaussian noise with zero mean and 0.01 variance, **b** de-noised image with adaptive median filter. **c** de-noised image with simple WPT (d) de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25), **e** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25) and WPT, **f** de-noised image using proposed approach

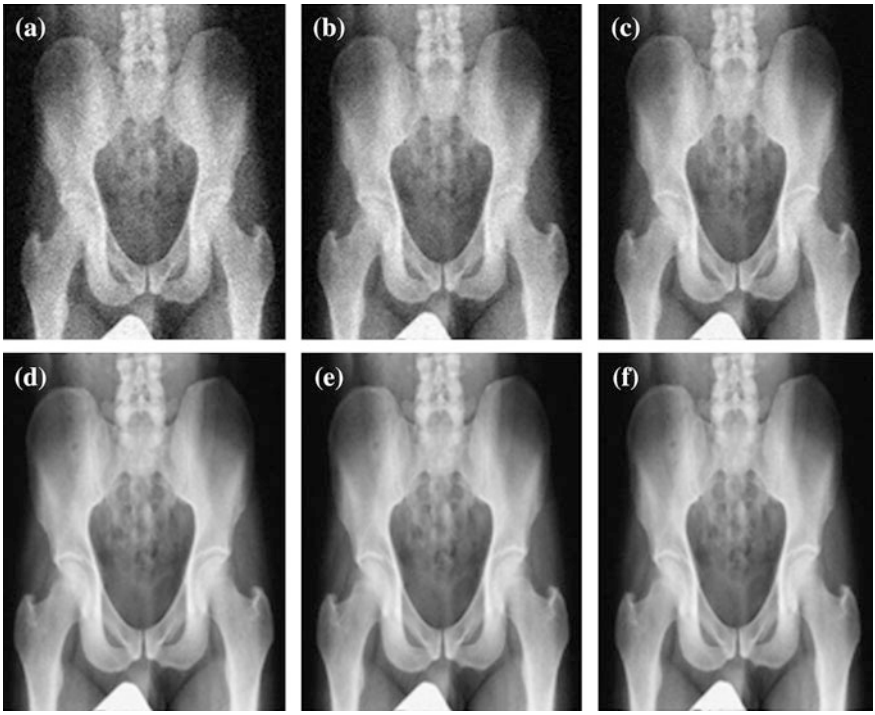


Fig. 7 **a** X-ray image of pelvis corrupted with Gaussian noise with zero mean and 0.01 variance, **b** de-noised image with adaptive median filter, **c** de-noised image with simple WPT, **d** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25), **e** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25) and WPT, **f** de-noised image using proposed approach

3.3 Performance Analysis

The performance of proposed approach is measured in terms of image quality metrics SNR, RMSE, SSIM, Precession, Recall, Accuracy and Error-Rate [13–15]. These metrics are calculated by using Eqs. (18–21).

(i) **Signal-to-Noise Ratio (SNR)**. It is the proportion of image signals to the noise. It can be compute it by using Eq. (18) this is expressed in decibels db. Here, f (i, j) is the original noise free image and A (i, j) is a de-noised image.

$$SNR = 10\log_{10}\frac{\sum_1^m \sum_1^n f(i,j)^2}{\sum_1^m \sum_1^n (f(i,j) - A(i,j))^2} \tag{18}$$

(ii) **Mean Square Error (MSE)**. The mean square error is the error between original signal f(x, y) and the reconstructed signal $\hat{f}(x, y)$. It can be expressed as in Eq. (19)

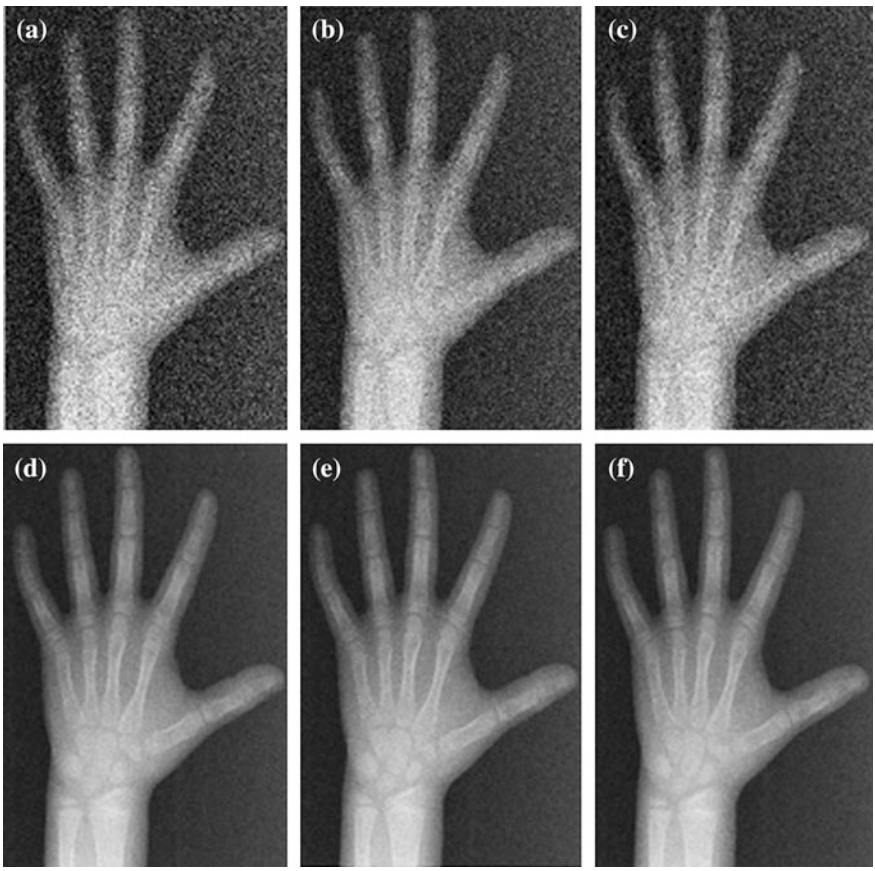


Fig. 8 **a** X-ray image of palm corrupted with Gaussian noise with zero mean and 0.05 variance, **b** de-noised image with adaptive median filter, **c** de-noised image with simple WPT, **d** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25), **e** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25) and WPT, **f** de-noised image using proposed method

$$MSE = \frac{1}{mn}\sum_{x=1}^{m}\sum_{y=1}^{n}\left[f(x,y) - \hat{f}(x,y)\right]^2 \tag{19}$$

(iii) **Root Mean Square Error (RMSE)**. The square root of Mean Square Error, can be computed by using Eq. (20),

$$RMSE = \sqrt{\frac{1}{mn}\sum_{x=0}^{m}\sum_{y=0}^{n}\left[f(x,y) - \hat{f}(x,y)\right]^2} \tag{20}$$

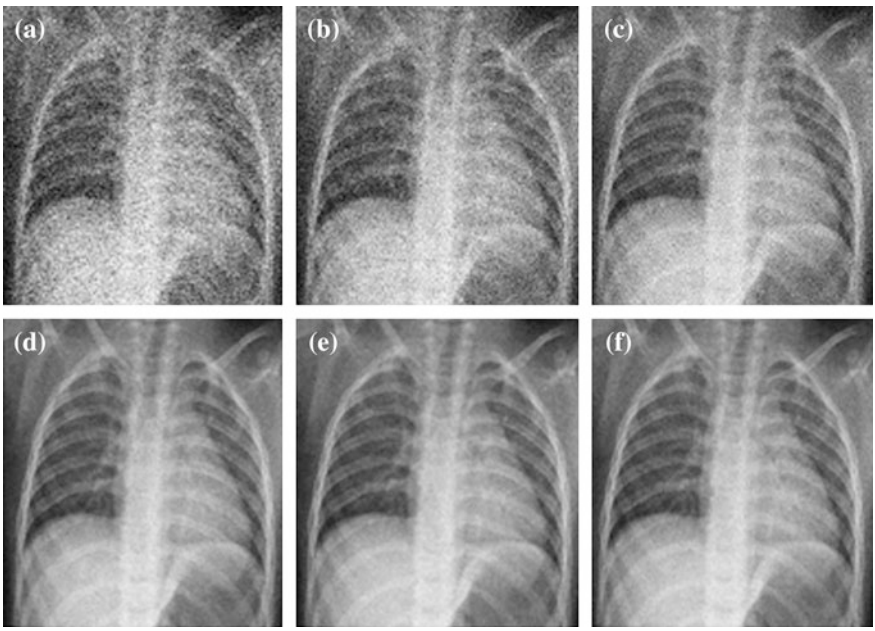


Fig. 9 **a** X-ray image of chest corrupted with Gaussian noise with zero mean and 0.05 variance, **b** de-noised image with adaptive median filter. **c** de-noised image with simple WPT, **d** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25), **e** de-noised image using anisotropic diffusion (iteration = 10, K = 120, λ = 0.25) and WPT, **f** de-noised image using proposed method

(iv) **Structural Similarity Index Map (SSIM)**. It is giving similarity between original and de-noised images computed by Eq. (21):

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{21}$$

Where x = (0…n − 1) and y = (0….n − 1), $c_1 = 2.55$ and $c_2 = 7.65$ are two constants. The value of SSIM should be in the range of 0–1. For a better reconstruction, each method should maximize SNR, SSIM and should minimize MSE, RMSE.

(v) **Precision**. It is the ratio of number of significant pixel identified to the total number of insignificant and significant pixel identified from the reconstructed image. It can be calculated by using Eq. (22).

$$Precision = \frac{Total\ number\ of\ actual\ pixels\ identified}{Total\ number\ of\ pixels} \tag{22}$$

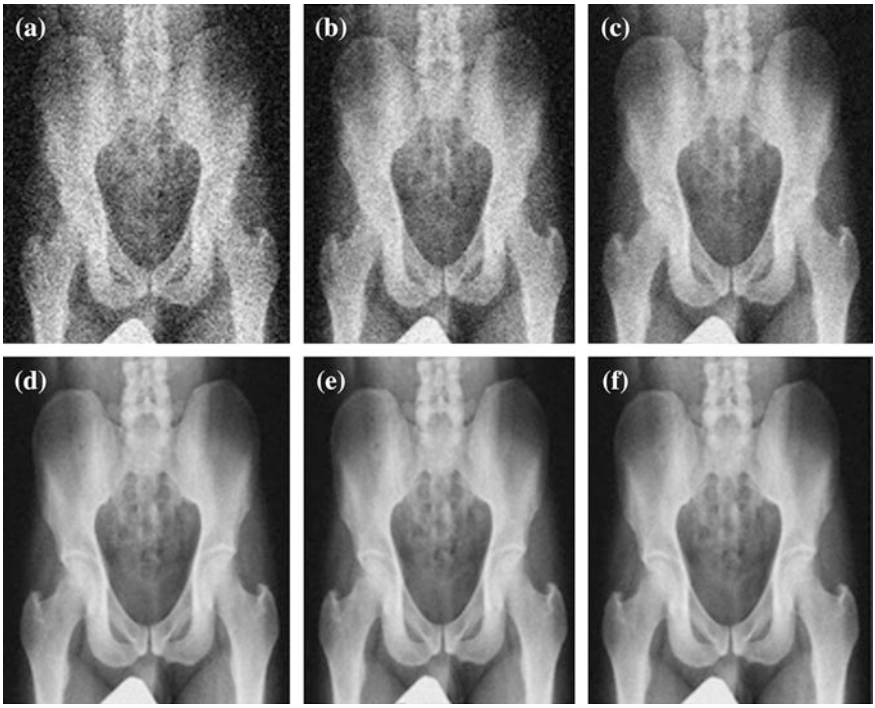


Fig. 10 **a** X-ray image of pelvis corrupted with Gaussian noise with zero mean and 0.05 variance, **b** de-noised image with adaptive median filter, **c** de-noised image with simple WPT, **d** de-noised image using anisotropic diffusion (iteration = 10, K = 120, $\lambda = 0.25$), **e** de-noised image using anisotropic diffusion (iteration = 10, K = 120, $\lambda = 0.25$) and WPT, **f** de-noised image using proposed method

(vi) **Recall**. It is the ratio of number of significant pixel identified to the total number of significant pixels in the image. This can be computed by Eq. (23).

$$Recall = \frac{Total\ number\ of\ actual\ pixels\ identified}{Total\ number\ of\ actual\ pixels\ in\ the\ image} \tag{23}$$

(vii) **Accuracy**. It is the ratio of total number of correct pixels identified to the total number of pixels.

(viii) **Error rate**. It is the ratio of total number of wrong pixels identified to the total number of pixels.

Now, the performance measures for the methods are shown in Table 1. This table comprises 5 methods of de-noising AMF, WPT, AD, Combnation of AD and WPT and the proposed one. The 3 X-ray images (palm, chest and pelvis), corrupted with Gaussian noise of zero mean and 0.01 variance are used for comparision by

Table 1 Performance measures for X-ray images corrupted with Gaussian noise of zero mean and 0.01 variance

Methods	Images	SNR	RMSE	SSIM	Precision	Recall	Accuracy	Error rate
Adaptive median filter	palm	15.4087	16.9470	0.8468	0.9970	0.9942	0.9913	0.0087
	chest	18.5910	17.2832	0.8540	0.9951	0.9918	0.9870	0.0130
	pelvis	18.2231	16.5228	0.8738	0.9960	0.9939	0.9900	0.0100
Anisotropic diffusion	palm	25.9037	5.0622	0.9368	0.9976	0.9981	0.9957	0.0042
	chest	27.7138	6.0462	0.9429	0.9955	0.9976	0.9931	0.0069
	pelvis	27.4992	5.6791	0.8968	0.9965	0.9967	0.9933	0.0067
WPT	palm	17.8932	12.7312	0.9105	0.9970	0.9954	0.9925	0.0075
	chest	20.9207	13.2172	0.9164	0.9948	0.9942	0.9891	0.0109
	pelvis	20.5253	12.6758	0.8883	0.9965	0.9957	0.9923	0.0077
Anisotropic diffusion + WPT	palm	26.4983	4.7273	0.9445	0.9977	0.9980	0.9957	0.0043
	chest	27.7966	5.9889	0.9452	0.9953	0.9976	0.9930	0.0070
	pelvis	27.7643	5.5084	0.8985	0.9964	0.9966	0.9931	0.0069
Proposed method	palm	**26.7210**	**4.6076**	**0.9465**	**0.9977**	**0.9981**	**0.9958**	**0.0042**
	chest	**28.1784**	**5.7313**	**0.9425**	**0.9948**	**0.9974**	**0.9922**	**0.0078**
	pelvis	**27.9518**	**5.3908**	**0.8994**	**0.9967**	**0.9970**	**0.9937**	**0.0063**

Table 2 Performance measures for X-ray images corrupted with Gaussian noise of zero Mean and 0.05 Variance

Methods	Images	SNR	RMSE	SSIM	Precision	Recall	Accuracy	Error rate
Adaptive median filter	palm	9.2739	34.3428	0.5817	0.9968	0.9775	0.9746	0.0254
	chest	12.0741	36.5991	0.5859	0.9941	0.9717	0.9663	0.0337
	pelvis	11.6979	35.0225	0.7002	0.9954	0.9771	0.9728	0.0272
Anisotropic diffusion	palm	20.1169	9.8556	0.7834	0.9976	0.9979	0.9955	0.0045
	chest	23.7582	9.5338	0.8056	0.9952	0.9971	0.9924	0.0076
	pelvis	20.8679	12.1855	0.7820	0.9960	0.9960	0.9921	0.0079
WPT	palm	11.7677	25.7717	0.7361	0.9969	0.9780	0.9752	0.0248
	chest	14.6818	27.1072	0.7596	0.9942	0.9792	0.9738	0.0262
	pelvis	14.0488	26.7179	0.7590	0.9954	0.9845	0.9802	0.0198
Anisotropic diffusion + WPT	palm	20.6087	9.3131	0.8091	0.9976	0.9978	0.9955	0.0045
	chest	24.5904	8.6627	0.8442	0.9947	0.9970	0.9917	0.0083
	pelvis	21.3097	11.5812	0.7934	0.9960	0.9960	0.9920	0.0080
Proposed method	palm	**21.0568**	**8.8449**	**0.8270**	**0.9977**	**0.9980**	**0.9957**	**0.0043**
	chest	**24.8066**	**8.4498**	**0.8422**	**0.9946**	**0.9970**	**0.9917**	**0.0083**
	pelvis	**21.3757**	**11.4936**	**0.7944**	**0.9961**	**0.9964**	**0.9926**	**0.0074**

Table 3 Performance measures for X-ray images corrupted with Gaussian noise of zero Mean and 0.1 Variance

Methods	Images	SNR	RMSE	SSIM	Precision	Recall	Accuracy	Error Rate
Adaptive median filter	palm	6.7846	45.7406	0.4395	0.9967	0.9666	0.9636	0.0364
	chest	9.6936	48.1384	0.4449	0.9938	0.9637	0.9581	0.0419
	pelvis	9.2337	46.5107	0.5903	0.9945	0.9670	0.9620	0.0380
Anisotropic diffusion	palm	16.0576	15.7270	0.6592	0.9974	0.9975	0.9949	0.0051
	chest	20.7021	13.5541	0.6949	0.9950	0.9966	0.9917	0.0083
	pelvis	17.4476	18.0659	0.6994	0.9956	0.9956	0.9914	0.0086
WPT	palm	3.6301	65.7694	0.4563	0.9965	0.9568	0.9538	0.0462
	chest	12.3092	35.6217	0.6526	0.9940	0.9709	0.9654	0.0346
	pelvis	11.4898	35.8714	0.6708	0.9952	0.9754	0.9709	0.0290
Anisotropic diffusion + WPT	palm	16.4751	14.9891	0.6933	0.9973	0.9973	0.9947	0.0053
	chest	21.6931	12.0926	0.7579	0.9942	0.9965	0.9908	0.0092
	pelvis	17.6284	17.6937	0.7043	0.9957	0.9957	0.9915	0.0085
Proposed method	palm	**16.7490**	**14.5238**	**0.7240**	**0.9974**	**0.9976**	**0.9951**	**0.0048**
	chest	**21.7429**	**12.0234**	**0.7612**	**0.9941**	**0.9966**	**0.9908**	**0.0093**
	pelvis	**17.9254**	**5.6695**	**0.7199**	**0.9956**	**0.9958**	**0.9915**	**0.0085**

using above stated metrics. From the table it is observed that the SNR value (26.7210) of proposed approach is more than the SNR value of all other methods (AMF-15.4087, AD-25.9037, WPT-17.8932 and AD + WPT-26.4983). This also shows that the accuracy (0.9958) of proposed method is better than others (as AMF-0.9913, AD-0.9957, WPT-0.9925 and AD + WPT-0.9957). It is clear from the table that metrics RMSE, SSIM, Precision, Recall and Error rate are also improved for proposed method. The results for Chest and Pelvis are also improved through proposed method as compared to other methods. Similarly, from the Tables 2 and 3 it can be observed that the performance of proposed method is much better than other 4 methods for all the three X-ray images (Palm, Chest and Pelvis) corrupted with White Noise of variance 0.05 and 0.1 respectively.

4 Conclusion

This paper has proposed an approach to reduce Gaussian noise from X-ray images with a good edge preservation technique. Through this method fine image details are preserved by reducing noise. The increased value of SNR shows that proposed approach is good enough to reduce noise than other methods. The accuracy shows that presented approach is more accurate than other methods. The SSIM value shows that the de-noised image is structurally similar with original image and its value is also increased as compared to others. The Precision and Recall values are showing that the proposed method generates actual image data as it was in original image and their values are also increased than other methods. All these metrics shows that the performance of proposed method is much better than AMF, AD, WPT and combination of AD and WPT.

References

1. Health risks from exposures to low levels of ionizing radiation; BEIR VII Phase 2. Published by National Research Council. Medical Radiation Studies. Chap. 7, pp. 155–157
2. Ravi Garg, Er., Abhijeet Kumar, Er.: Comparison of various noise removals using Bayesian framework. Int. J. Mod. Eng. Res. **2**(1), 265–270 (2012)
3. Fang, Y., Zhou, Y., Ge, D., Zhou, Z.: De-noising based on wavelet analysis and Bayesian estimation for low-dose X-ray CT. In: 9th International Conference on Electronic Measurement & Instruments, 2009. ICEMI 09, pp. 2- 944–947 (2009)
4. Luo, B., Sun, Z., Xue, M., Liu, H.: 23, Nov 2013 Improved Noise Reduction Algorithms for Medical X-ray Images. IEEE Int. Conf. 22-pp. 359–362
5. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**, 629–639 (1990)
6. Catté, F., Lions, P.L., Morel, J.M., et al.: Image selective smoothingand edge detection by nonlinear diffusion, SIAM. J. Numerical Anal. **29**, 182–193 (1992)
7. Pizurica, A., Philips, W.: Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising. IEEE Trans. Image Process. **15**(3), 654–665 (2006)

8. Zhiming, W., Hong, Bao, Li, Zhang: Image denoising by anisotropic diffusion in wavelet domain. Third Int. Conf. Meas. Technol. Mechatron. Autom. (ICMTMA), IEEE Trans. **2**, 359–362 (2011)
9. Mahalakshmi, B.V., Anand, M.J.: Adaptive wavelet packet decomposition for efficient image denoising by using Neighsuren shrink method. Int. J. Comput. Sci. Inf. Technol. **5**(4), 5003 (2014)
10. Rajan, J., Kaimal, M.R., Image denoising using wavelet embedded anisotropic diffusion (WEAD). In: IET International Conference on Visual Information Engineering, 2006, pp. 589–593 sept 2006
11. Rosenthal, A., Jetzfellner, T., Razansky, D., Ntziachristos, V.: Efficient framework for model-based tomographic image reconstruction using wavelet packets. IEEE Trans. Med. Imaging **31**(7), 1346–1356 (2012)
12. Maini, R., Aggarwal, H.: Study and comparison of various image edge detection techniques. Int. J. Image Process. (IJIP) **3**(1), 1–11 (2009)
13. Srivastava, R., Singh, S.K., Shukla, K.K.: Research developments in computer vision and image processing: methodologies and applications. Information Science References Publishing Company, Chap. 5, pp. 76–82 (2013)
14. Jizba, R.:Measuring Search Effectiveness (2007)
15. Shah, V., Bhat, P., Dasarath, M., Gupta, S., Bhavsar, J.: Content based image retrieval using combination of Kekre transform and HSV color segmentation, IACSIT Int. J. Eng. Technol. **3** (2011)

Visual Tracking on Riemannian Space via Standard Deviation Based Model Updated Method

Anuja Kumar Acharya and Biswajit Sahoo

Abstract In this paper we proposed a new simple standard deviation based model update method for tracking the object of interest in the successive frame of the video. Each candidate object in the searching frame is first represented with a matrix of features and then transferred into the low dimensional covariance matrix for the purpose of matching the target object. Suitable candidate object is judged on the basis of the minimum Forster distance between the candidate and the reference object. A Model updation method is adopted in this proposed work to track longer trajectory when the object undergoes deformation in shape and size. This model update is carried out by updating each observation vector of the referenced model by adding the vector mean difference of standard deviation between the referenced object and the detected objects to the referenced model. A covariance matrix is then extracted from this updated feature image which can be used for the next set of future reference frame. In the proposed model also we use the Kalman filtering to effectively handle the background clutter and temporary occlusion. Simulation result shows the current method is robust for real time tracking.

Keywords Feature matrix · Covariance · Riemannian · Subspace

1 Introduction

The ability to track and retain information about an object by finding the correspondences of the previously detected objects in the current frame has several vision applications. The main challenging issue of tracking is to handle effectively the intrinsic and extrinsic appearance variation of the object. The former is linked to the variation of object which undergoes change in the shape and pose change. The letter related to the variation in the camera motion, illumination change and occlusion.

A.K. Acharya (✉) · B. Sahoo
School of Computer Engineering, KIIT University, Bhubaneshwar 751024, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_42

Many methods where previously developed on vision tracking such as appearance based and fixed template method. Tracking with fixed templates can be reliable over short durations, but it copes poorly with appearance changes over longer durations that occur in most applications. Reliability can be improved with the use of subspace models of appearance [3]. Black and Jepson proposed an algorithm using a pre-trained view-based eigenbasis representation and a robust error norm [1]. This method mainly focus on the motion estimation between the frames and also it undergoes the learning a set of view based eigen bases prior to the task of tracking. but these are object specific and often require training prior to tracking. Frey [2] proposed a tracker with image templates that model the mean and the variance of each pixel during tracking. Yu and Wu [6] present an differential tracking based spatial appearance model tracking. Ho et al. [7] developed a tracking algorithm based on the appearance of the object. In his approach the subspace is updated at real time and tracking depend on the previous best approximation of searching the object. Hypothesis based multivariate approach [11] also give significant result for the tracking. But this method is not suitable for the object undergoes continuous appearance variation. Li et al. [4] propose an incremental PCA algorithm for subspace learning. In Skocaj and Leonardis [5], a weighted incremental PCA algorithm for subspace learning is presented. The above two process is quite robust for tracking but it lacks with capturing the spatial and statistical property of the object. This cause the method not fertile towards the illumination and change in variation of the object. In order to tackle those issues, Porikli and Tuzel [8] proposed appearance based covariance matrix descriptor for capturing both the statistical as well as the spatial property of the object window. As the mean covariance matrix is possessing the property of symmetric positive definite and also lie on the Riemannian manifold so the statistics for covariance matrices of image features may be computed through Riemannian geometry. Appearance based model using the Riemannian geometry is robust only when the deformation of appearance is minimal. Many different type of tracking algorithm has been developed in the recent past need on this geometry, Tuzel et al. [9] present a new algorithm for human detection using Riemannian manifolds. Lim et al. [10] present a human tracking framework using the identification of system dynamics and nonlinear dimension reduction technique. Although log euclidean distance can be used for matching the object but the methods are lacking to obtain the optimal mean covariance of the detected region needed for the model update.

Based on this log euclidean Riemannian metric [12], we propose a Visual Tracking on Riemannian Space via Standard deviation Based Model Updated Method. In the proposed method, the object of interest is first represented with set of spatial and statistical feature. From this feature matrix then a low dimensional covariance matrix is constructed. A riemannian space based Forster distance is used to find the best candidate region in a frame. Each observation of the referenced model is updated on the basis of finding the average difference of deviation of each features of the referenced object from the series of detected object.

The main aim of our work is to update the referenced model which in terms depends on finding a mean covariance matrix from the set of detected covariance region. As the covariance factor is mainly depends on the deviation of each feature

sample from its mean. So the key idea behind is to updated the each sample feature of the feature referenced image with a small value 'δ' which is the mean difference of standard deviation (SD) between the referenced and detected object. This leads to minimize the variance structure between the updated referenced model and the detected region.

Contribution of our Work:

- The target object window is scale down to smaller feature image for faster computation of covariance matrix.
- This tracked object window size is always fixed one and its observation which is described through 'd' dimensional space is updated with addition of newly detected matching object.
- Observation vector of target object is updated on the basis of the deviation of the object from their mean dimension.

Our work is organized as follows. Representation of the target window in the form of set of feature vector is discussed in the Sect. 2. Followed to this, Sect. 3. Discuss the frame work of the proposed model and its updated strategy. Section 4 discuss the performance analysis of our proposed model.

2 Covariance Matrix Representation

Using Tuzzel et al. [8]. Covariance matrix representation, The observed Image I of M × N one dimensional or three dimensional color image is converted into d dimensional feature image of size M × N × d. Mathematically this mapping can be written as

$$\mathrm{F(n,p)} = \varnothing(\mathrm{I,r,c}) \tag{1}$$

Here the function Ø is a mapping function and it can be any mapping such as color, image gradient or edge magnitude. For a given rectangular region R **C** I, denote the {fi}i = 1 … n as the d dimensional feature point obtained by Ø within R. In this paper each pixel of the object is defined through the set of feature as

$$f_k = [\mathrm{x\ y\ I(x,y)\ I}_X\ \mathrm{I}_Y\ \sqrt{I_{x^2} + I_{y^2}} \cdot\ \arctan\frac{I_X}{I_Y}].$$

where x, y, are the spatial coordinate and I(x, y), Ix, Iy are the intensity value and its deviation of the pixel along x and y. Consequently, the image region R can be represented as $d \times d$ covariance matrix.

$$S_r = \frac{1}{(mn-1)}\sum_{i=1}^{n}\left(y_{ij} - \bar{y}_i\right)\left(y_{ij} - \bar{y}_i\right) \tag{2}$$

The covariance matrix descriptor of a grayscale or color image region is a 7 × 7 or 17 × 17 symmetric matrix. The pixels' coordinates are involved in the computation of the covariance matrix in order to include the spatial information about the image region and the correlations between the positions of the pixels and the intensity derivatives into the covariance matrix. It is possible to compute covariance matrix from feature images in a very fast way using integral image representation [13].

3 Framework for Visual Tracking

Our proposed frame work includes two stage. As shown in the Fig. 1, In the first stage the object of interest is represented into a low dimensional feature covariance matrix and in the second stage it uses the Kalman filter to predict the future location and estimate the current state from the given all the previous observations. Followed to this the reference feature image is updated by adding a small value δ. This value is the mean difference in the deviation of the so far detected region from the referenced object.

3.1 Object Representation

In this proposed appearance based model, each pixel of the object is treated as the sample or observation defined with set of spatial and statistical attributes or features.

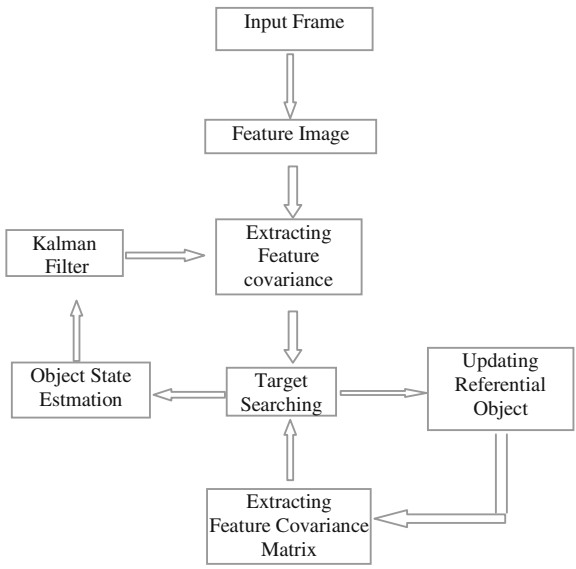


Fig. 1 Framework of the proposed model

Thus a feature image can be constructed from the one dimensional intensity or three dimensional color image. Using Eq. (2) this feature image matrix is extracted to low dimensional covariance matrix.

3.2 Matching of Candidate Region

To Find the best candidate region in the target frame, one needs to compute the effective distance between the candidate covariance matrix Ci and the reference covariance matrix Cj. As the covariance matrix are not lie on the euclidean space so a log euclidean based Forster distance [7] is employed for the said purpose for computing the similarity between the two matrix.

This method uses the sum of the squared logarithms of the generalized eigenvalues to compute the dissimilarity between covariance matrices as

$$\rho(C_i, C_j) = \sum_{k=1}^{d} \left(\sqrt{\log^2 \lambda_k(C_i, C_j)} \right) \tag{3}$$

where $\lambda_k(C_i, C_j)$ are the generalized eigenvalues of (C_i, C_j) which is derived from the equation

$$\lambda_k C_i X_k - C_j X_k = 0 \tag{4}$$

here Xk represent the generalized eigen values. The distance measure ρ satisfies the metric axioms, positivity, symmetry, triangle inequality, for positive definite symmetric matrices.

Minimum is the distance between the matrix maximum closer is the similarity factor between the two object.

3.3 Model Update Strategy

As the object under the region of similarity undergoes change in shape size and appearance with respect to time, to adapt this variation it is necessary to update the appearance model. Let R is the referential feature image region with size m × n × d. This region which is represented with d dimensional feature or attribute is updated. This model updation is carried out with estimated object region R_1, R_2, … R_J. We keep set T of previous covariance matrices [C_1, C_2 … C_T] where Ci denotes the current covariance matrix of the detected region. From this set we compute sample mean covariance matrix.

Although a straight forward solution may be to take the aggregate covariance of all the detected covariance matrix. But as the covariance matrix do not conform to euclidean space and it is confronted to a riemannian space. It can be possible to find the mean of detected covariance matrix by using the Riemannian mean. But the computational cost for this linearly grows as time progresses.

An average covariance structure can be constructed from the detected regions of the previous frame. A small value δ which is the mean difference of the standard deviation between the reference model and the detected region, is added with each sample or observation of the original reference model. A low dimensional covariance structure is then derived from this resulting updated model which can be used for the next set of successive frame. Following steps are consider for the model updation.

Step 1. Find the mean feature vector and the variance of the reference feature image as

$$\mu_{r1}\mu_{r2}\mu_{r3}\ldots\mu_{rd}$$

and

$$\sigma_{r1}\sigma_{r2}\ldots\sigma_{rd}$$

Step 2. similarly find the mean and variance of the detected region as $\mu_{d1}\mu_{d2}\mu_{d3}\ldots\mu_{dd}$ and $\sigma_{d1}\sigma_{d2}\ldots\sigma_{dd}$

Step 3. Find out the difference of rate of deviation and variance of each dimension from its mean of the referenced model and detected region.

$$\sigma_j = \sigma_{rj} - \sigma_{dj}$$

$$\delta_j = \left(\sqrt{(\sigma_{rj})} - \sqrt{(\sigma_{dj})}\right)/2$$

Step 4. Update each observation Yi of the referenced model with the value δ

i. $\forall Y_{ij} = Y_{ij} + \delta_j$ if $Y_{ij} > \mu_{r1}$
ii. $Y_{ij} = Y_{ij} - \delta_j$ if $Y_{ij} < \mu_{r1}$

Step 5. compare the variance of the each respective dimension of original reference object and the detected one and update the original model if

- If $(\sigma_j > 0)$ then update the each observation of the reference object with value and $Y_{ij} = Y_{ij} + \delta_j$ if $Y_{ij} > \mu_{r1}$

Once the original referenced region is updated then extracted its covariance matrix of this referenced model which can be used for the next sequence of candidate object in the video.

ALGORITHMS

1. *Initialize* $\hat{C}=C_1, Y=Y_1$
2. *for* t =1..T do
3. - *Compute* $\sigma_{ij}=\left(\sqrt{(\sigma_{1j})}-\sqrt{(\sigma_{dj})}\right)$
4. *compute* … $\delta=\frac{1}{2T}\sum_{t=1}^{T}\sigma_t$
5. *if* $\left(Y_{ij}>\mu_j\right)$
6. $Y_{ij}=Y_{ij}+\delta_j$
7. *else*
8. $Y_{ij}=Y_{ij}+\delta_j$
9. $\hat{C}=COV(Y)$

3.4 Kalman Filter

The main aim of using Kalman filter is to detect and identify the state of the object in the presence of the occlusion. From the set of past observation the future state of the object can be predicted. Once the true state information is available this predicted measurement is undergoes the update phase.

The Kalman filter estimates the position (x, y) in the frame of the object to be tracked. Since the standard Kalman filter uses one correct measurement, data association should be considered to classify true measurement and false measurements. Each Kalman filter is configured as follows

$$X_k = AX_{k-1} + W_k \tag{5}$$

$$Z_k = HX_k + V_k \tag{6}$$

where

$$X = \left[P_x P_y V_x V_y\right]^T$$

Px Py represent the center position of x-axis, and y-axis. Vx, Vy are the velocity of x-axis and y-axis. Matrix A represents the transition matrix, matrix H is the measurement matrix, and T is the time interval between two adjacent frames. Wk, Vk are the Gaussian noises with the error covariances Qk and Rk. Process of the Kalman filter is as follows.

The Kalman filter state prediction Xt and state covariance prediction Pt are defined by:

$$X_{k|k-1} = AX_{k-1|k-1}$$

where Xt and Pt denotes the estimated state vector and error covariance matrix respectively at time t. Then the Kalman filter update steps are as follows:

$$K_k = P_{k|k-1}H^T\left(HP_{k|k-1}H^T + T\right)^{-1}$$

$$X_{k|k} = x_{k|k-1} + K_k\left(Z_k - Z_{k|k-1}\right) \tag{7}$$

$$P_{k|k} = (1 - KH)P_{k|k-1} \tag{8}$$

where K is the Kalman gain and this is depend on the accuracy of the measurement. KF algorithm starts with initial conditions with K0 and P0. Kt is the Kalman gain, which defines the updating weight between the new measurements and the prediction from the dynamic model.

3.5 Time Complexity

The total time complexity of our proposed algorithm is depend on three factor. Computation time for updating the reference feature matrix is O(mnd). The time computation for finding the low dimensional covariance matrix using integral covariance representation is $O(mnd^2)$ and the Forster distance measurement to find the similarity between two covariance matrix is $O(mnd^3)$. So the total computational complexity is $O(mnd + mnd^2 + mnd^3)$. Asymptotically this complexity can be expressed as $O(n^3)$.

4 Experimental Analysis

In order to find the performance of the proposed tracking method, experiment are carried out using the opencv-2.4.2 on Ubuntu platform. Seven number of feature where considered for each pixel of the object A 9×9 neighborhood window size is considered for the searching of candidate object. The experiment are carried out on different types of vision video dataset Fig. 2 we show the simulation result of our approach without and with model update taking into account. Also we compare the sequence with the output of covariance based model update method [8]. Figure 3 we show the simulation result of our proposed algorithm applied on two video dataset. The first one is the tracking on the moving person and the last one on tracking a pedestrian in a shopping mall. We also analyzed the number of trials to find the correct estimation. This is based on ordering the search regions according

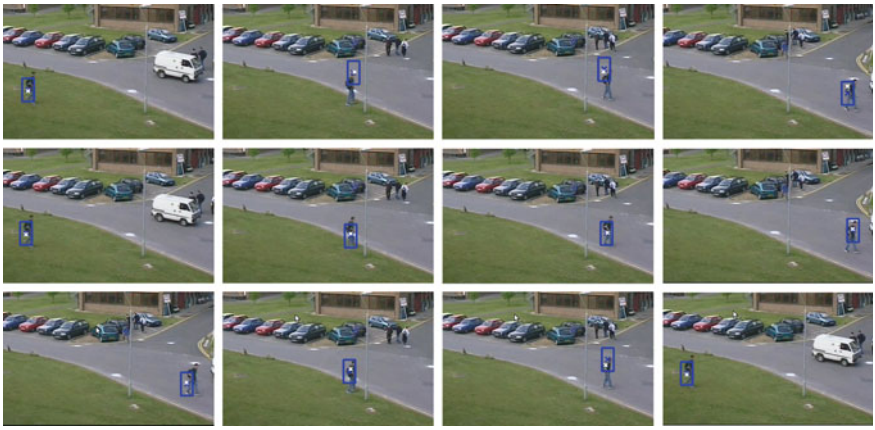


Fig. 2 First row and 2nd row showing the result of our proposed method without and with model update and 3rd row showing the simulation result of existing covariance based model update method

Fig. 3 1st row tracking the person from stationary camera of a video 2nd row tracking the occluded human body in a shopping mall

Table 1

Input video\parameter	Detection rate (%)	No of trial
Moving man video	96.28	0.0012
Pedestrian tracking	94.43	0.0915

to the match scores until we find the correct estimation. We defined the metric as the ratio of the total number of trials to the total number of possible regions.

We show in Table 1, the detection rate and the no of trials for moving object tracking on moving man video and pedestrian video. The simulation results shows the robustness of the algorithm.

5 Conclusion

In this paper we have presented a Visual Tracking on Riemannian Space via Standard deviation Based Model Updated Method. The object of interest is represented with set of features. The object similarity is measured with log euclidean based Forster distance. Moreover deviation based sample update is used to find the covariance of the updated model. Kalman filter is also used to handle the occlusion and background clutter for tracking. Experimental result shows the proposed method is promising for real time tracking.

References

1. Black, M.J., Jepson, A.D.: Eigen tracking: robust matching and tracking of articulated objects using a view based representation. IJCV **26**(1), 63–84 (1998)
2. Frey, B.: Filling in scenes by propagating probabilities through layers into appearance models. In: Proceedings of IEEE CVPR I, pp. 185–192, Hilton Head (2000)
3. Hager, G.D., Belhumeur, P.N.: Efficient region tracking with parametric models of geometry and illumination. IEEE Trans. PAMI **27**(10), 1025–1039 (1998)
4. Li, Y.: On incremental and robust subspace learning. Pattern Recogn. **37**(7), 1509–1518 (2004)
5. Skocaj, D., Leonardis, A.: Weighted and robust incremental method for subspace learning. In: Proceedings of Ninth IEEE Int'l Conference Computer Vision, vol. 2, pp. 1494–1501, Oct 2003
6. Yu, T., Wu, Y.: Differential tracking based on spatial-appearance model (SAM). Proc. CVPR **1**, 720–727 (2006)
7. Ho, J., Lee, K., Yang, M., Kriegman, D.: Visual tracking using learned linear subspaces. In: Proceedings of IEEE Conference Computer Vision and Pattern Recognition, vol. 1, pp. 782–789 (2004)
8. Porikli, F., Tuzel, O.: Covariance tracking using model update based on lie algebra. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 1, pp. 728–735, June 2006
9. Tuzel, O., Porikli, F., Meer, P.: Human detection via classification on riemannian manifolds. In: Proceedings of IEEE Conference Computer Vision and Pattern Recognition, pp. 1–8, June 2007
10. Lim, H., Morariu, V., Camps, O.I., Sznaier, M.: Dynamic appearance modeling for human tracking. In: Proceedings of IEEE Conference Computer Vision and Pattern Recognition, vol. 1, pp. 751–757 (2006)
11. Acharya, A.K., Sahoo, B. Swain, B.R.: Object tracking using a new statistical multivariate hotelling's T2 approach. In: IEEE International on Advance Computing Conference (IACC), pp. 969–972 (2014)
12. Li, X., Hu, W.M., Zhang, Z.F., Zhang, X.Q., Zhu, M.L., Cheng, J.: Visual tracking via incremental log-Euclidean Riemannian subspace learning. In: Proceedings of IEEE Int'l Conference Computer Vision and Pattern Recognition, pp. 1–8, June 2008
13. Tuzel, O., Porikli, F., Meer, P.: Region covariance: a fast descriptor for detection and classification. In: Proceedings of ECCV, pp. 589–600 (2006)

Kannpos-Kannada Parts of Speech Tagger Using Conditional Random Fields

K.P. Pallavi and Anitha S. Pillai

Abstract Parts Of Speech (POS) tagging is one of the basic text processing tasks of Natural Language Processing (NLP). It is a great challenge to develop POS tagger for Indian Languages, especially Kannada due to its rich morphological and highly agglutinative nature. A Kannada POS tagger has been developed using Conditional Random Fields (CRFs), a supervised machine learning technique and it is discussed in this paper. The results presented are based on experiments conducted on a large corpus consisting of 80,000 words, where 64,000 is used for training and 16,000 is used for testing. These words are collected from Kannada Wikipedia and annotated with POS tags. The tagset from Technology Development for Indian Languages (TDIL) containing 36 tags are used to assign the POS. The n-gram CRF model gave a maximum accuracy of 92.94 %. This work is the extension of "Parts of Speech (POS) Tagger for Kannada Using Conditional Random Fields (CRFs).

1 Introduction

Parts of Speech (POS) tagging is a process in which each word of a sentence is tagged with appropriate syntactic label such as noun, verb, adjective, preposition, and so on [1]. The syntactic label that represents these lexical categories is known as a tag. For example, NN label represents a common noun tag, and VMF label represents a finite verb tag.

Example Kan: rAma (N_NNP) shAlege (N_NN) hOda (V_VMF).
En: Ram (N_NNP) went (V_VMF) to (PSP) school (N_NN).

A tool that tags appropriate POS of each word in a given document is known as POS tagger. A POS tagger can be developed using various techniques like rule based techniques, transformation based techniques, Machine Learning

K.P. Pallavi (✉) · A.S. Pillai
Department of Computer Applications, Hindustan University, Chennai 603103, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_43

(ML) techniques, and example based techniques and so on. Rule based techniques have been written using syntactic rules, whereas Machine Learning techniques use probabilistic models and stochastic grammar. In this paper, a supervised machine learning technique called Conditional Random Fields (CRFs) is discussed in developing POS tagger for Kannada.

Kannada is one of the south Indian Languages belonging to the Dravidian family. It is the native language of the Karnataka state and spoken mostly in the southern parts of India. Even though Kannada has 50.8 million speakers[1] all over the world and is the official language of the Karnataka[2] state, research in computational linguistics in Kannada is still lagging. One of the main reasons for research to lag in computational linguistics for Kannada is due to its rich morphological and agglutinative nature [2]. It is an agglutinating language with suffixes and nominative-accusative syntax. It also follows Subject-Object-Verb (SOV) constituent order [3]. A word in Kannada could be comprised of 8 suffixes, where the word could either be a single word or combination of words consisting of root words and suffixes. It could also be made up of two or more words, which leads to ambiguity. Pre-processing rules to split those kinds of words have been written according to orthography of a language. Most of the other ambiguities are solved by using a fine grained TDIL hierarchical tagset. Tagset is one of the key elements of text processing techniques like Chunking, Named Entity Recognition, POS. POS tagset have been proposed for few Indian Languages by Technology Development for Indian Languages (TDIL). For Kannada POS, several tagset exist like TDIL [4], IIIT Hyderabad,[3] Bhuvaneshwari C. Melinamath [5], Vijayalakshmi patil's [6], Shambhavi [2] and Antony [7]. Along with the tagset, corpus, a collection of large raw data, plays an important role in training and testing POS tagger. For effective Natural Language Processing (NLP) results, corpus should be from the same domain where NLP application will be prefer. For example, the biomedical POS trained corpus gives better accuracy only for the biomedical datasets. Similarly, the generic area/domain trained corpus gives good results for general datasets. The POS tagger used a large corpus of 80,000 tokens collected from Kannada Wikipedia was trained using Conditional Random Fields (CRFs).

Conditional Random Field is a Supervised Machine Learning (SML) technique. SML techniques are widely being used for text processing, which considers labelled datasets for training. CRFs offer a unique combination of properties by discriminating trained models for sequence segmentation and labelling [8]. Labelled dataset used for training this POS tagger contains manually annotated POS tags with pre-processed corpus words and some extracted linguistic features of those words. This model achieved a competitive accuracy.

Second section highlights the related work done in this field. In the third and fourth sections, challenges and proposed parts of speech tagger for Kannada are

[1] http://timesofindia.indiatimes.com/india/Indiaspeak-English-is-our-2nd-language/articleshow/5680962.cms?referral=PM.

[2] http://en.wikipedia.org/wiki/Kannada.

[3] http://ltrc.iiit.ac.in/nlptools2010/files/documents/POS-Tag-List.pdf.

discussed respectively. Finally, in the last two sections, the paper presents results achieved and concludes with the future work.

2 Related Work

Parts of Speech tagger was developed enormously in English and other European languages. Earlier POS taggers were mostly based on rule based and supervised machine learning techniques. One among those English POS taggers by Toutanova et al. [9] set the benchmark by achieving an accuracy of 97.25 %. After this, the researchers turned their head towards unsupervised machine learning techniques. An unsupervised grammar induction task was demonstrated experimentally that the universal POS categories generalize well across language boundaries and gave competitive accuracy without relying on gold POS tags [10]. Apart from these kinds of POS taggers, Dipanjan et al. [11] approached with token and type constraints for cross-lingual part-of-speech tagging for 8 languages, which reduced error rate by 25 % when compared to the prior state of the art [12]. These taggers were not working well for social media data like twitter data. Kevin et al. [13] has developed a POS tagger for twitter data which gave 90 % accuracy and they believe that their approach can be applied to address other linguistic analysis needs as they continue to rise in the era of social media and it's rapidly changing linguistic conventions.

Recently, A HMM based POS tagger for Hindi and Marathi was developed by Nisheeth Joshi et al. [14] and Jyothi Sing [15]. Hindi POS tagger achieved 92 % accuracy, using Indian Language (IL) POS tagset. Jayabal Ganesh et al. [16] achieved a precision of 87.74 % for Tamil POS tagger using only a small set of POS labelled suffix context patterns. Biswa Ranjan Das et al. [17] has got an accuracy of 81 % for Odia POS tagger using artificial neural networks. A rule based Graphical User Interface tool has been developed by using Netbeans IDE 7.1 for Manipuri by Kh Raju Singha et al. [18]. The 35 rules were framed using 1500 lexicons and achieved an accuracy of 92 %. A Malayalam POS tagger using statistical approach with the Hidden Markov Model following the Viterbi algorithm is described by Jisha P Jayan et al. [19].

Antony et al. [7] proposed a POS tagger for Kannada using Support Vector Machine (SVM) using hierarchical tagset consisting of 30 tags with a compatible accuracy of 86 % and Shambhavi et al. [2, 20] randomly tried with Maximum Entropy (Maxent), Hidden Markov Model (HMM) and CRF. CRF proved better than Maxent and HMM with an accuracy of 84.58 % for 51,269 training data and around 2932 of test data tokens collected from Enabling Minority Language Engineering (EMILLE) corpus. Siva Reddy et al. [21] and Mallama et al. [22] developed POS taggers which gave competitive accuracies for Kannada using Hidden Markov Model and Decision Trees.

In the above Kannada POS papers there are many challenges which were not attempted. Some of the challenges are described in next section.

3 Challenges

All Indian languages are agglutinative in nature [3] where many suffixes or other morphemes are added to the base of a word. Kannada is one such south Indian Languages with 15 vowels, 25 consonants. Developing a Parts Of Speech tagger for Kannada is challenging due to its rich morphological and highly agglutinative nature. Rich morphology occurs with word inflections and are called word level ambiguities. For example, "AguMtukanaMte ((ಆಗುಮತುಕನಮತೆ))" is a word where nouns "AguMtuka" joined with demonstrative "aMte". Consider another example "kELugara ((ಕೇಳುಗರ))" is a word where verb "kEL", it becoming noun after joining with "ugara". Similarly, many ambiguities exists at sentence level also. A sentence level ambiguity includes word tagged as adjectives, postpositions, nouns and adverbs, depending on the POS. For example, ugra ((ಉಗ್ರ)). The method to overcome these kind of challenges are proposed in the following section.

4 Proposed Kannpos

Kannada POS tagger was developed using CRFs for the corpus contained various articles. The corpus was annotated manually using TDIL tagset. Tagset, CRFs, corpus used and annotation method are explained below.

4.1 Tagset

TDIL tagset is unique when compared with the other tagsets. TDIL [23] POS schema is based on W3C XML Internalization best practices and one to one mapping table for all the labels used. It consist of 11 main categories Noun, Pronoun, Demonstrative, Verb, Adjective, Adverb, Postposition, Conjunction, Particles, Quantifiers and Residuals of POS which are classified into 29 level 1 and 5 level 2 sub categories. It proposed different levels of verb category, where other tagsets missed this part. For example, the level 1 verb tags are—main, verbal and auxiliary. Level 2 verb tags are—finite, non-finite, infinitive and gerund. We used TDIL tagset for tagging corpus.

4.2 Corpus

The source of data used in this POS tool was collected from Kannada Wikipedia articles from different domains like sports and eminent personalities. Around

80,000 words were collected, annotated manually, trained and tested in 80:20 ratio. The POS tagger is tested on this corpus and, also on newspaper data.

4.3 Pre-processing Rules

The corpus was pre-processed manually. Pre-processing rules were written to separate symbols and punctuation marks from words. In Indian languages, the symbols and punctuations are written with words as in Manipuri [18]. The words are also written together in Kannada, those words were orthographically split using python script.

Examples

1. Kan: mAyeyeMdhare → mAye + eMdhare
   En: illusion means → illusion + means
2. Kan: rajyadalliruva → rajyadalli + iruva
   En: In the state → state

In example 1, the noun "mAye" suffixed to the demonstrative "eMdhare" and in example 2, the non-finite verb "iruva" suffixed to the common noun "rajyadalli".

4.4 Annotation

Annotation is a complex task and consumes enormous time, compared to corpus collection. Corpus was tokenized through python programs before annotation. Tokenization is splitting the sentences into each word called as tokens. Tokens were arranged in column wise and one blank line was given between sentences before annotation. The 6700 words were annotated manually. Another 73,300 words were tagged using that base engine. Tagged words were validated, and any incorrectly tagged words were corrected manually. Annotated corpus was trained using CRFs.

4.5 Learning Techniques

Conditional Random Fields is a supervised machine learning technique. According to Lafferty et al. [8], "It is a framework for building probabilistic models to segment and label sequence data. Conditional random fields offer several advantages over hidden Markov models and stochastic grammars for such tasks, including the ability to relax strong independence assumptions made in those models. Conditional random fields also avoid a fundamental limitation of maximum entropy

Markov models (MEMMs)". One unique feature of CRF is, it enables the incorporation of arbitrary local features in a log-linear model [13].

In this work, CRFs has been used to develop POS tagger. Since, CRFs is a supervised learning method, rules were written to training the model. Rules were designed using unigram and bigram for tagging. Unigram rules were based on the list look up approach, where bigrams generate conditional probability [24] rules based on the given features to overcome the difficulties faced from Unigram. Bigram rules determined the correct tag when the same word is tagged with different tags in training data. The rules were constructed using words $\{w_{-2}, w_{-1}, w_0, w_1, w_2\}$, features $\{f_0, \ldots, f_{12}\}$. Although the rules were framed using words and features, the annotated tags were used to predict the tag of each word in a training data set. The sequence of words, features extracted from the words and annotated tags were used as input to training CRF model, whereas the output was the POS tags tagged for each word in a sequence of sentences. The features used to frame rules are briefed below.

4.6 Features

Features plays a very important role for CRFS. Identified features includes words and extracted linguistic information of words. Here are the features in detail.

Word: Parts Of Speech exists within the context boundary limit of 5 words. So word feature with window size 5 was used. That means, if the current word was w0, then previous two words were w − 1, w − 2 and next two words were w + 1, and w + 2 are considered.

Case markers: In Kannada, there are 8 case markers for nouns. They are listed with examples in Table 1 and those were identified by comparing with word suffixes. The case markers are not limited to nouns. These can be the inflections of pronouns, verbal nouns and verbs as well. In the training set, we identified 978 words coming with case markers, but they were not nouns. Those words are listed below in Table 2 and some examples of those kind of words are given in Table 3.

Last character: Many tokens/words occur with the same set of last characters. Words with same last character were identified.

Prefix: Almost all the pronouns starts with similar kind of letters. So, prefix features helped in identifying the pronouns easily. Apart from pronouns many other words starts with similar kind of characters. Those were identified easily with the combination of other features.

Tense: Tense markers play a very important role in verbs. There are 60 finite verbs occurring with case markers. This ambiguity was solved by identifying tense and

Table 1 The list of case markers of Kannada

Vibhakti (Case)	Pratyagalu (Markers)	Examples
prathama (Nominative) dvithiya (Accusative) thrithiya (Instrumental) chaturthi (Dative) paMchami (Ablative) shasti (Genitive) saptami (Locative) Sambodana (Vocative)	ು (u) ನ್ನು(annu) ಇಂದ (iMda) ಗೆ (ge) ಕೆ (ke) ಇಕೆ (ike) ದೆಸೆಯಿಂದ (deseyiMda) ಅ (a) ಅಲ್ಲಿ (alli) ಏ (E) ಓ (O) ಆ (A)	Huduganu pustakavannu batteyiMda tayige ramanige Danakke mudukana deseiMda akkana americadalli gurugaLE siteyOramA

Table 2 List of tags with the number of tags which are nouns, but coming with case markers

Tags names	Number of tags
Demonstratives	251
Verbal nouns	188
Pronouns	173
Postpositions	137
Particles	88
Finite verbs	60
Conjunctions	29
Quantifiers	19
Adverbs	18
Adjectives	13
Auxiliary verbs	2

png markers for verbs. In Kannada grammar, we have 3 tenses past, present and future. The case marker for past tense is "da", future tense markers are "uva" and "va" and the past tense marker is "utta". Sometimes the present tense marker is used as future tense. So, according to the modern linguistics/grammar, Kannada morphology has only two tenses, past and non-past [3]. Other than these, some tense markers which occurred frequently in our training set are listed in Table 4.

PNG (Person, Number and Gender markers): PNG marker gives the information on nouns from verbs. These are always coming with verbs. Common PNG markers in Kannada are: Lu, nu, ru, ge and gi.

Table 3 Examples for words which are not nouns, but coming with case markers

Words with Vibakthi's but not nouns	Vibakthis	Tags of the words	Words with Vibakthi's but not nouns	Vibakthis	Tags of the words
ಆ	ಆ	DM_DMD	ಬೇರೂರಬೇಕೆ	ಕೆ	DM_DMQ
ಇಂದ	ಇಂದ	DM_DMR	ರಿಗೆ	ಗೆ	DM_DMD
ಗಳಿಗೆ	ಗೆ		ರೊಂದಿಗೆ	ಗೆ	
ಹೇಗೆ	ಗೆ	DM_DMD	ಅಗಾಗ್ಗೆ	ಗೆ	DM_DMD
ಎಂಬುದಕ್ಕೆ	ಕೆ	DM_DMD	ಈಬಗ್ಗೆ	ಗೆ	DM_DMD
ಅನ್ನು	ಅನ್ನು		ನಿಗೆ	ಗೆ	
ಯಾಕೆ	ಕೆ	DM_DMD	ನಮ್ಮೊಂದಿಗೆ	ಗೆ	DM_DMD
		DM_DMD	ಬಾರಿಗೆ	ಗೆ	DM_DMD
		DM_DMQ			PR_PRP
					PSP

Table 4 Tense markers with examples

Tense markers	Examples
"ುತ್ತೇ"	ಉತ್ತರಿಸುತ್ತೇವೆ
"ುತ್ತಿ"	ಬಳಲುತ್ತಿದ್ದ
"ುತ್ತಾ"	ಬರೆಯುತ್ತಾರೆ
"ುವ"	ಬರೆಯುವ
"ಿದ"	ನೋಡಿದ
"ಗೊಂಡು"	ಕೈಗೊಂಡು
"ಕೊಂಡು"	ತೆಗೆದುಕೊಂಡು
"ತಿದ್ದ"	ಎದುರಿಸುತ್ತಿದ್ದ
"ಕ್ಕೀಡು"	ಬಂಧನಕ್ಕೀಡು

Examples Kan: ashwini (N_NNP) shalege (N_NN) hodaLu(V_VM_VF)
En: Ashwini (N_NNP) went (V_VM_VF) to (PSP) school (N_NN)

In the above example, Lu is the feminine marker which identified Ashwini. L gives information for a second person and also person is in one number.

Negation markers: The verbs which are ending with illa and alla belong to the negation category. Verbs with negation markers were treated as negations, along with negation markers. Some nouns include negation information. Those were separated from negation markers.

Examples thinnuvudilla, maduvudilla

Digits: Most commonly, English number symbols are found in Kannada script. But in few cases only Kannada number used. So, to avoid the confusion both English and Kannada number symbols were grouped to form a Digit Feature. The suffixes attached with the number like raMdu, ralli, nE makes the task easy. Suffixes were treated as demonstrators.

Examples Kan: 12raMdu
En: On 12th

Here, the preprocessed rule was written to separate 12 and raMdu. Then 12 tagged as digit and suffix raMdu tagged as demonstrator.

Symbols: Symbol feature helped in identifying the symbols as well as surrounding words.
Punctuations: The comma, full stop, question mark and colon together forms punctuation feature. The sentence breakers occurs after a full stop, colon and question mark. Including comma, all punctuation helps in finding surrounding words.
Foreign words: Non-Kannada script comes into this group.

All the above features were played important role in increasing the precision of POS tagger. It is reported in results section.

5 Results

The POS tagger tool was experimented on sentences taken from various articles. These sentences consisted of 80000 words, out of which first 64000 were taken as training data and the remaining 16000 words were taken as test data. The test data contained 40 % of words that were not part of training data. The POS tagger tool tagged 16000 test data and the results are presented below in measure of precision. The precision is a ratio of the number of tagged words (*True Pasitive*) and the number of positive responses (*True Pasitive* + *False Pasitive*) as shown below.

$$Precision(P) = \frac{TruePositive}{TruePositive + FalsePositive}.$$

Table 5 Fivefold validation results

Folds	Precisions
1st	91.45
2nd	87.56
3rd	90.57
4th	90.8
5th	88.24
Average	89.724

Table 6 Randomized fivefold validation results

Folds	Precisions
1st	92.14
2nd	92.1
3rd	91.94
4th	92.1
5th	91.8
Average	91.998

Table 7 Sample results

Words	Tags
ಕೇಂದ್ರ	N_NN
ಸರಕಾರದ	N_NN
ಯೋಜನೆಗಳನ್ನು	N_NN
ರಾಜ್ಯಗಳ	N_NN
ಮೇಲೆ	PSP
ಹೇರುವ	V_VM_VNF
ಬದಲು	V_VM_VINF
ಪ್ರತಿಯೊಂದು	N_NN
ರಾಜ್ಯದ	N_NN
ಅಗತ್ಯಕ್ಕೆ	N_NN
ಅನುಗುಣವಾಗಿ	N_NN
ಮತ್ತು	CC_CCD
ಅವುಗಳಿಗೆ	PR_PRP
ಸರಿಹೊಂದುವ	V_VM_VNF
ಯೋಜನೆಗಳನ್ನು	N_NN
ರೂಪಿಸುವುದಕ್ಕೆ	V_VM_VNF
ಪ್ರಧಾನಿ	N_NN
ನರೇಂದ್ರ	N_NNP
ಮೋದಿ	N_NNP
ಒಲವು	N_NN
ವ್ಯಕ್ತಪಡಿಸಿದ್ದಾರೆ	V_VM_VF
.	RD_PUNC

where, True Positives means the system tag is same tag as the gold tag.

False Positives means the system tag is different tag compared to the gold tag.

POS tagger achieved a precision of 87.78 % for 16000 tokens. These tokens were divided into 80:20 ratio for fivefold validation. The fivefold experiment gave the precision between 87 and 91 % shown in Table 5.

The sentences were randomized to training and test again. We have achieved the precisions between 91.8 and 92.94 % for fivefold validation and it is shown in Table 6. We experimented on a new data set of 2,000 words, which is collected from the Kannada daily newspaper and gave precision of 94.7 %. The sample results are shown in Table 7.

6 Conclusion

In this paper, an approach for Kannada POS tagger with hierarchical tagset has been proposed. It has been trained on 64,000 words, and tested on 16,000 unseen words. These 80,000 words were collected from Kannada Wikipedia. They were annotated manually with the help of base engine which was developed on 6,700 words. An annotated words were added with the 12 linguistic features and saved in a training file along with annotated POS tags. The features were designed based on case markers, prefixes, suffixes, tense markers, verb suffixes, png markers, numbers, symbols, foreign words and punctuations based on the linguistic of the language. The linguistic rules were constructed based on permutations and combination of those words with 5 window size. With this tagger, we achieved a competitive accuracy of 92.94 % compared to the other existing works.

This work can be extended in developing NLP applications like Chunker and Named Entity Recognizers for Indian languages. It can be extended by applying unsupervised algorithms to identify the hidden features in the corpus.

References

1. Pallavi., Pillai, A.S.: Parts Of Speech (POS) Tagger for Kannada using conditional random fields (CRFs). In: National Conference on Indian Language Computing (NCILC 2014) 1st to 2nd Feb 2014
2. Shambhavi, B.R., Kumar, R.: Kannada Part-Of-Speech Tagging with Probabilistic Classifiers. Int. J. Comput. Appl. **48**(17), 0975–888, June 2012
3. Sridhar, S.N.: Modern Kannada Grammar. Lordson Publishers Pvt. Ltd., Delhi (2007) (First published in 1990 by Routledge, England under the title *Kannada* in the Descriptive Grammars Series edited by Bernard Comrie. Reprinted in 2007 by Manohar Publishers and Distributors, 4753/23 Ansari Road, Daryaganj, New Delhi 110 002. ISBN 81-7304-767-7)
4. Unified Parts of Speech (POS) Standard in Indian Languages—Draft Standard–Version 1.0, Department of Information Technology Ministry of Communications and Information Technology, Govt. of India

5. Melinamath, B.C.: Improvement over IL-POST tagset for Kannada. Int. J. Comput. Sci. Eng. (IJCSE) **3**(3), 179–186 May 2014
6. Patil, V.F.: Designing POS Tagset for Kannada, Linguistic Data Consortium for Indian Languages (LDC-IL), Organized by Central Institute of Indian Languages, Department of Higher Education Ministry of Human Resource Development, Government of India, March 2010
7. Antony, P.J., Soman, K.P.: Kernel based part of speech tagger for Kannada. In: International Conference on Machine Learning and Cybernetics (ICMLC), vol. 4, pp. 2139–2144, IEEE (2010)
8. Lafferty, J., McCallum, A., Pereira F.C.N.: Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the International Conference on Machine Learning (ICML-2001)
9. Toutanova, K., Klein, D., Manning, C.D., Singer, Y.: Feature-rich part-of-speech tagging with a cyclic dependency network. In: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, vol. 1, pp. 173–180, Association for Computational Linguistics (2003)
10. Spitkovsky, V.I., Alshawi, H., Chang, A.X., Jurafsky, D.: Unsupervised dependency parsing without gold part-of-speech tags. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1281–1290, Association for Computational Linguistics (2011)
11. Täckström, O., Das, D., Petrov, S., McDonald, R., Nivre, J.: Token and type constraints for cross-lingual part-of-speech tagging. Trans. Assoc. Comput. Linguist. **1**, 1–12 (2013)
12. Li, S., Graça, J.V., Taskar, B.: Wiki-ly supervised part-of-speech tagging. In: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp. 1389–1398, Association for Computational Linguistics (2012)
13. Gimpel, K., Schneider, N., O'Connor, B., Das, D., Mills, D., Eisenstein, J., Heilman, M., Yogatama, D., Flanigan, J., Smith, N.A.: Part-of-speech tagging for twitter: annotation, features, and experiments. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers, vol. 2, pp. 42–47. Association for Computational Linguistics (2011)
14. Joshi, N., Darbari, H., Mathur, I.: HMM based POS tagger for Hindi. In: Proceeding of 2013 International Conference on Artificial Intelligence, Soft Computing (AISC-2013) (2013)
15. Bagul, P., Mishra, Archana, Mahajan, Prachi, Kulkarni, Medinee, Dhopavkar, Gauri: Rule Based POS Tagger for Marathi Text. Proc. Int. J. Comput. Sci. Inf. Technol. (IJCSIT) **5**(2), 1322–1326 (2014)
16. Ganesh, J., Ranjani Parthasarathi, T. V. Geetha, and J. Balaji. "Pattern Based Bootstrapping Technique for Tamil POS Tagging." In *Mining Intelligence and Knowledge Exploration*, pp. 256–267. Springer International Publishing, 2014
17. Das, B.R., Patnaik. S.: A novel approach for odia part of speech tagging using artificial neural network. In: Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2013, pp. 147–154. Springer International Publishing (2014)
18. Singha, Kh.R., Singha, Ksh.K.B., Purkayastha, B.S.: Developing a part of speech tagger for Manipuri. Int. J. Comput. Linguist. Nat. Lang. Process. **2**(9) Sept 2013
19. Jayan, Jisha P., Rajeev, R.R.: Parts of speech tagger and chunker for malayalam: statistical approach. Comput. Eng. Intell. Syst. **2**(2), 68–78 (2011)
20. Shambhavi, B.R., Ramakanth, K.P., Revanth, G.: A maximum entropy approach to Kannada part of speech tagging. Int. J. Comput. Appl. **41**(13), 9–12 (2012)
21. Reddy, S., Serge S.: Cross language POS taggers (and other tools) for Indian languages: an experiment with Kannada using Telugu resources. Cross Ling. Inf. Access **11** (2011)
22. Reddy, M.V., Hanumanthappa, M.: POS Tagger for Kannada Sentence Translation. Int. J. Emerg. Trends Technol. Comput. Sci. (IJETTCS) **1** (2012)

23. Department of Information Technology Ministry of Communications & Information Technology Govt. of India. Unified Parts of Speech (POS) Standard in Indian Languages—Draft Standard–Version 1.0
24. Che, W., Wang, M., Manning, C.D., Liu, T.: Named entity recognition with bilingual constraints. In: HLT-NAACL, pp. 52–62 (2013)

Image Processing for Soybean Disease Classification and Severity Estimation

Saylee Gharge and Priyanka Singh

Abstract In agriculture, Plant disease is one of the major congestion for increasing productivity and quality of food. False diagnosis of plant disease causes excessive use of pesticides which in turns affects the quality of crop. In this paper, we proposed algorithm for detection of Soybean disease and its Severity. The research paper focuses on classification and infected area estimation of Frogeye, Downy mildew and Bacterial Pustule disease of Soybean. In this proposed approach, Image enhancement technique for enhancing the image quality is used. Then k-means segmentation algorithm is applied to separate infected cluster from leaf. Neural Network is used to classify Frogeye, Downy mildew and Bacterial Pustule. Accuracy of 93.3 % is achieved for 30 images. After classification area estimation of infected area is performed.

Keywords Soybean leaf images · Decorrelation stretch · k-means · Neural network · Region properties

1 Introduction

Soybean is one of the most important crops in India. Soybean is the rich source of protein and vegetable oil. In India, it is cultivated mainly in Madhya Pradesh, Uttar Pradesh, Maharashtra and Gujarat. Soybean grows well in warm and moist climate. A temperature of 26.5–30 °C appears to be the optimum for its cultivation in India [1]. The main cause of reduction of Soybean production is Disease. There are several diseases of Soybean; this paper mainly focuses on Frogeye Disease, Downy Mildew and Bacterial Pustule. Plant/Crop Diseases are easily identified by their symptoms. In Frogeye disease, small, irregular to circular lesions having gray with reddish brown borders appear on upper surface of leaf, Downy mildew pale green

S. Gharge (✉) · P. Singh
Department of Electronics and Telecommunication, VES Institute of Technology, Mumbai University, Mumbai, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_44

to yellow spots are seen on upper surface of leaf and in Bacterial pustule lesion appear on upper part of leaves are surrounded by greenish yellow halo. Accurate disease detection and management requires continuous monitoring of Soybean crop and this is not possible. Also availability of fewer experts and large number of Soybean farms increases the demand of automatic detection of diseases.

2 Related Work

Researchers have been used several approaches such as image processing, machine learning and crowd sourcing [2] for fast and accurate detection of crop diseases. Meunkaewjinda et al. [3] used Machine learning algorithms to detect Grape Diseases, SOM is applied for segmenting leaf from background, Genetic algorithm for selecting number of cluster and SVM is used for classification purpose. Al Bashish et al. [4] tries to identify six different plant diseases using k-means for separating infected part form leaf then BPNN is consider to detect the disease. According to author one of the clusters corresponds to one of the diseases. Sanjeev et al. [5] proposed image processing technique for detecting and classifying two Grape diseases Downy Mildew and Powdery Mildew. To remove noise anisotropic diffusion is used, and then k-means is applied for segmenting infected part. Finally neural network is applied for classification of Grape Diseases. Zulkfli et al. [6] aims to identify disease in chili plants. They used color clustering for separating leaf from background. Then infected leaf and healthy leaf is classified based on colour features.

The main contribution of this paper is detecting frogeye disease, Downy mildew and Bacterial Pustule of Soybean fast and accurately using k-means and neural network. Then, infected area estimation is performed using properties of image. Disease severity can be estimated by taking area into consideration. The main goal of the paper is to detect the disease and its severity in order to reduce the pesticide spray. The paper is organized as follow proposed algorithm has been presented in Sect. 3, we present results and conclusion in Sects. 4 and 5.

3 Material and Method

In this paper, algorithm for detection and classification of disease using Soybean leaf images is proposed. Images taken as input are downloaded from internet [7]. Images are captured from digital camera and stored in standard jpeg format. In this paper, for Soybean diseases identification includes various steps like image enhancement segmentation. Flowchart of proposed algorithm is given in Fig. 1.

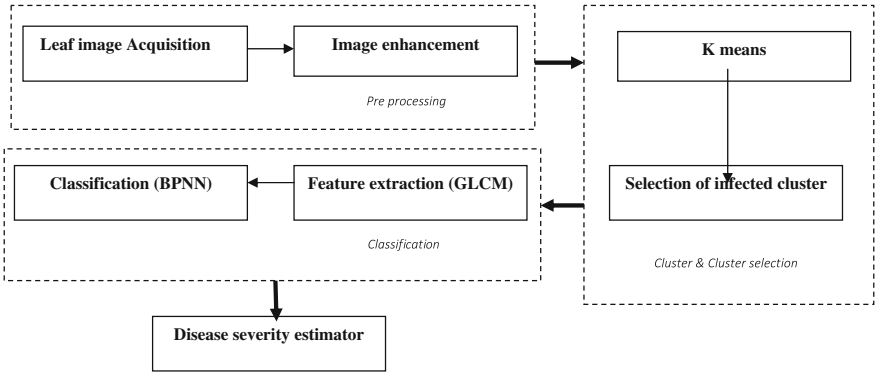


Fig. 1 Flowchart of proposed algorithm

3.1 Pre-processing: It Includes Two Main Steps

a. Leaf image Acquisition: Soybean leaf image database used for training is taken from net. Total 30 leaf images of Soybean are used, 10 each for frogeye disease, Downy mildew and Bacterial pustule. Input images are of resolution 768 * 512 pixels.
b. Image Enhancement: For improving the results of disease analysis or to increase the visual interpretation of image decorrelation stretch image enhancement technique is used in this paper. In these colors values in R, G & B channels are mapped to new colors values with wide range.

3.2 Clustering and Cluster Selection

Clustering is method of partitioning image into small clusters based on internal properties of pixels. This step contains:

a. k-means: k-means [8] is used for image segmentation. k-means makes clusters such that pixels within the cluster are close to each other and far from other cluster. First RGB image is converted into CIELAB color space. LAB is device independent color space. In l * a * b, 'l' describes the brightness value, 'a' and 'b' components gives the color information. So 'a' and 'b' values of pixels are used for clustering process. k-means algorithm proposed by Stuart Lloyd is used. Euclidean distance metric is used for measuring the difference between two colors. The k-means algorithm is composed of following steps:

- Image pixels $x1$, $x2$, … xn are taken as input.
- Each pixel is assigned to only one of the cluster.
- Each pixel is assigned to its nearest centroid. Euclidean distance metric is used.

- After assigning each pixel into one of the cluster. Recalculate the K new centroids by taking mean of assigned pixels.
- Repeat Steps 2 and 3 till convergence is achieved. This produces partitioning of image into k cluster.

k-means is used to divide the Soybean leaves into four clusters. In our case, these clusters contain background, Leaf green part and infected part. Many times infected part appears in more than one cluster. It happens because leaf images are infected by more than one disease or presence of different levels of disease at same time.

Results obtained from k-means clustering are shown in Fig. 2. In Fig. 2a shows the original leaf image of Soybean infected by Disease, (b) and (c) shows the healthy part of leaf, (d) is the background and (e) contains the infected cluster.

b. Selection of Infected Cluster: In this step each infected cluster is selected automatically. This is done by using cluster center. Cluster center contains 'a' and 'b' value of pixels. Experimentally we determined the cluster center of infected part then using conditional logic automatically select the infected cluster. After this, green pixels removed from the selected cluster based on pixels R, G and B value. Pixels with R < G values are retained and for rest of the pixels values of R, G and B channels are set to zero. Automatic selection of infected cluster is shown below (Fig. 3).

3.3 Classification

a. Feature Extraction: The next step is extracting features of infected cluster using GLCM [9]. It generates matrix 'Pij' that counts the number of times a pixel with

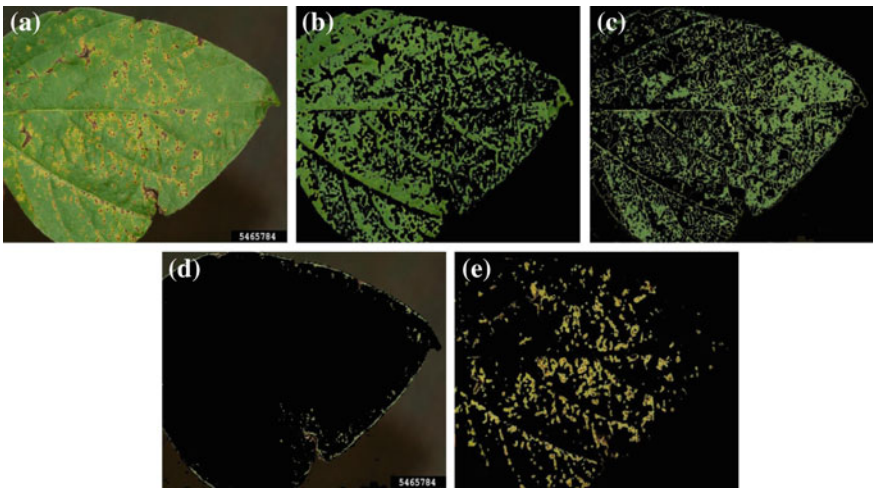


Fig. 2 **a** Original image. **b** First cluster. **c** Second cluster. **d** Third cluster. **e** Fourth cluster obtained after K means

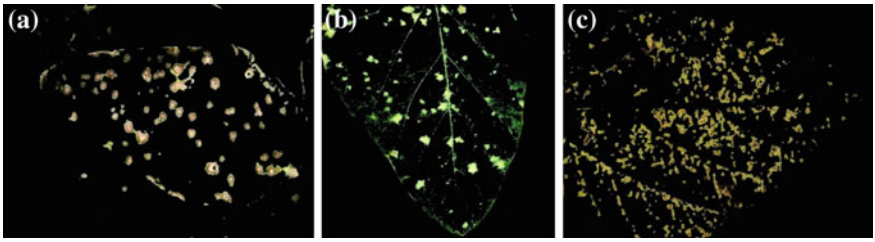


Fig. 3 **a** Infected cluster of frogeye disease. **b** Infected cluster of downy mildew. **c** Infected cluster of bacterial pustule

grey-level 'i' occurs at position p from a pixel with grey-level 'j'. The GLCMs are represented by the function P(i, j, d, θ) where 'i' represents the gray level of the location (x, y) in the image and 'j' represents the gray level of the pixel at a distance 'd' from location (x, y) at an orientation angle of 'θ'. GLCM of hue is used to extract seven features contrast, homogeneity, energy, difference variance, difference entropy, maximum probability and entropy.

b. Neural Network: Three layer i.e. input layer, hidden layer and output layer back propagation neural network [10] is used for classification of Soybean disease. Input layer has seven nodes whereas hidden layer contains 20 neurons. It performs training, validation and testing on features extracted from leaf images.

3.4 Disease Estimator

In this step, first RGB leaf image is converted into binary image (Fig. 4).

For disease severity calculation, area of infected cluster and total leaf image is determined using regionprop of MATLAB toolbox. Then major axis and minor axis are calculated using the regionprop. Mathematical representation:

A_T = Total number of pixels of Leaf image
A_I = Total number of pixels of Infected Cluster
L = Major axis of Leaf Image
W = Minor axis of Leaf Image
R_T = L * W
$R_I = (R_T * A_I)/A_I$
R_I = Area of Infected cluster
Percentage of Infected Area = $(R_I/R_T) * 100$

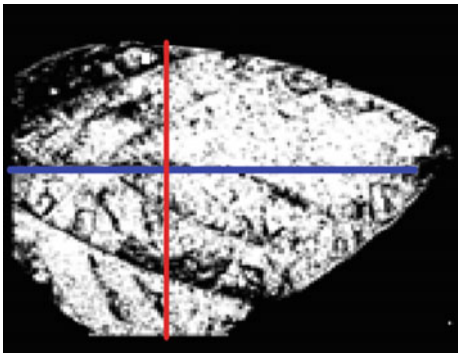


Fig. 4 Binary image of soybean leaf (*blue line* is major axis and red line is minor axis)

4 Result and Disscusion

30 images of Soybean leaf is used for analyzing the performance of proposed algorithm. Feature extracted from the images are used as input in neural network. 25 images are used for training and 2 and 3 images used for validation and testing. In Fig. 5a, Confusion matrix shows the results of classification of frogeye disease (class 1), Downy mildew (class 2) and bacterial pustule (class 3). It shows 93.3 % accuracy i.e. 1 image each of class 1 and class 3 are misclassified as class 2 and Fig. 5b the ROC curve shows the better performance of proposed algorithm.

Disease severity is estimated by calculating area, major axis and minor axis of total leaf image and infected cluster. Table 1 shows the results of Soybean. It represents the leaf image along with Soybean diseases and percentage of infected area for each.

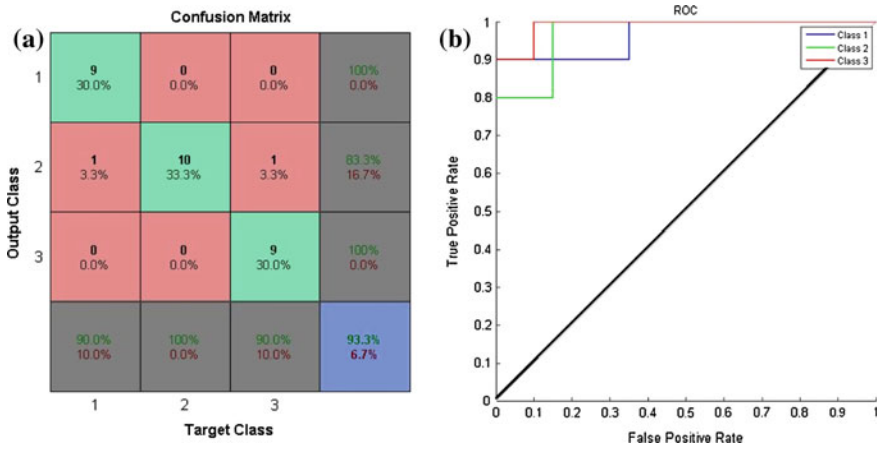


Fig. 5 **a** Confusion matrix of neural network and **b** ROC curve of neural network

Table 1 Results of soybean

	Soybean disease	Leaf image		Infected cluster		% Infected area
		A_T	R_T	A_I	R_I	
Leaf 1	Frogeye disease	256120	4.73E+05	12931	2.39E+04	5.0529
Leaf 2	Frogeye disease	124565	1.69E+05	11832	1.61E+04	9.5266
Leaf 3	Frogeye disease	150194	3.19E+05	9417	2.00E+04	6.2696
Leaf 4	Frogeye disease	144217	2.87E+05	12687	2.52E+04	8.7805
Leaf 5	Frogeye disease	59597	1.49E+05	7447	1.86E+04	12.4832
Leaf 6	Frogeye disease	99976	1.63E+05	16191	2.64E+04	16.1963
Leaf 7	**Downy mildew**	**144787**	**2.41E+05**	**11304**	**1.88E+04**	**7.8008**
Leaf 8	Frogeye disease	79570	2.59E+05	38230	1.24E+05	47.8764
Leaf 9	Frogeye disease	163165	2.36E+05	5846	8.46E+03	3.5847
Leaf 10	Frogeye disease	159799	2.41E+05	21793	3.29E+04	13.6515
Leaf 11	Downy mildew	90478	2.59E+05	28197	8.07E+04	31.1583
Leaf 12	Downy mildew	200085	3.01E+05	54867	8.25E+04	27.4086
Leaf 13	Downy mildew	38107	7.36E+04	14715	2.84E+04	38.587
Leaf 14	Downy mildew	30291	4.73E+04	2997	4.68E+03	9.8943
Leaf 15	Downy mildew	119471	1.78E+05	36936	5.50E+04	30.8989
Leaf 16	Downy mildew	116043	1.99E+05	29437	5.05E+04	25.3769
Leaf 17	Downy mildew	100942	1.41E+05	22898	3.20E + 04	22.695
Leaf 18	Downy mildew	24584	6.98E+04	1488	4.22E+03	6.0458
Leaf 19	Downy mildew	240254	3.15E+05	16890	2.21E + 04	7.0159
Leaf 20	Downy mildew	109519	2.71E+05	18421	4.56E+04	16.8266
Leaf 21	Bacterial pustule	112974	2.11E+05	22848	4.27E+04	20.237
Leaf 22	Bacterial pustule	207571	4.57E+05	21685	4.77E+04	10.4376
Leaf 23	Bacterial pustule	117496	2.76E+05	23498	5.52E+04	20
Leaf 24	Bacterial pustule	283818	3.99E+05	29852	4.20E+04	10.5263
Leaf 25	Bacterial pustule	37571	1.31E+04	22156	7.73E+03	59.0076
Leaf 26	**Downy mildew**	**99534**	**1.50E+05**	**22023**	**3.32E+04**	**22.1333**
Leaf 27	Bacterial pustule	59863	1.01E+05	24746	4.18E+04	41.3861
Leaf 28	Bacterial pustule	71279	1.73E+05	11649	2.83E+04	16.3584
Leaf 29	Bacterial pustule	30095	4.15E+03	4786	659.9734	15.903
Leaf 30	Bacterial pustule	26151	4.46E+03	10344	1.76E+03	39.4619

Highlighted part in table represents the incorrectly classified disease of Soybean. Based on this the accuracy is as follows:

$$\begin{aligned}\text{Accuracy} &= (\text{Correctly classify images}/\text{Total number of images}) * 100 \\ &= (28/30) * 100 \\ &= 93.3\%\end{aligned}$$

5 Conclusion and Future Scope

The main aim of this paper is to increase the productivity and reduce the usage of pesticide by identifying and grading the soybean disease accurately. The algorithm consist of steps like segmentation using K means, classification of images using neural network and finally grading of disease using regionprop function of MATLAB. The proposed algorithm achieves good accuracy in classification.

Future work includes disease classification among more than three Soybean diseases. Real implementation of proposed algorithm on Soybean farms.

References

1. Bisaliah, S.: Soybean development in India. In: ICRISAT (1986)
2. Singh, P., Jagyasi, B., Rai, N., Gharge, S.: Decision tree based mobile crowd sourcing for agriculture advisory system. In: Proceedings of Indicon (2014)
3. Meunkaewjinda, A., Kumsawat, P., Attakitmongcol, K., Srikaew, A.: Grape leaf disease detection from color imagery using hybrid intelligent system. In: Proceedings of ECTICON (2008)
4. Bashish, Al., Braik, M., Bani-Ahmad, S.: A framework for detection and classification of plant leaf and stem diseases. In: International Conference on Signal and Image Processing, Chennai: IEEE (2010)
5. Sannakki, S.S., Rajpurohit, V., Nargund, V., Kulkarni, P.: Diagnosis and classification of grape leaf diseases using neural networks. In: Fourth International Conference on ICCCNT, July 2013
6. Zulkfli, H., Yeon, A., Shakakff, M., Abdul, A., Rohani, S.: Feasibility study on plant chilli disease detection using image processing (2008)
7. www.ipmimages.org
8. Kanungo, T., Mount, D., Netanyahu, N., "An efficient k-means clustering algorithm: analysis and implementation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 881–892 (2002)
9. Haralick, S.K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **SMC-3**, 610–621 (1973)
10. Zhang, G.: Neural networks for classification. IEEE Trans. Syst. **30**(4), 245–259 (2000)

Secure Cryptographic Algorithm Using Digital Logic

Bidisha Mandal, Sourabh Chandra and Sk. Safikul Alam

Abstract Today's civilization is inextricably involved with all sorts of data communication. For every organization and business bodies, handling large amount of data with proper confidentiality and security is an important task to do. To achieve the secure transfer of data through the network we have proposed a symmetric key cryptographic algorithm that provides security to the confidential information by encrypting this information twice. This algorithm comes up with the concept of secure key transfer as we use symmetric key concept. This algorithm incorporate binary addition, folding method, logical XOR operation with encryption key and generation of 2's complement of a number represented in 8-bit binary equivalent value.

Keywords Cryptography · Encryption · Decryption · Shared link · Logical XOR · Cipher text

1 Introduction

"Cryptography", the Greek originated word reveals the meaning "Hidden Writing" has becomes an integral part of data security related fields [1, 9]. The competency to securely transfer of confidential information over network has becomes a prominence factor for every organization. Cryptography [3] provides a way to secure data communication. Cryptographic algorithms are providing a secure communication process where information is transfer based on the information security aspects. For example, an internet security protocol called IPSec uses a set of algorithms like DES, HMAC with SHA1 and MD5 to provide a secure communication [4].

Cryptography transmits data as a cipher text so that only the intended receiver can decode and process it. Using two basic method called Encryption and

B. Mandal (✉) · S. Chandra · Sk. Safikul Alam
Department of CSE, Calcutta Institute of Technology, Kolkata, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_45

Decryption Cryptography achieves the security goals [2]—confidentiality (i.e. No one other than the intended receiver can understood the message), integrity (i.e. the information cannot be alter during transmission), authenticity (i.e. only authorized person can read and process the information) and non-repudiation (i.e. sender of message cannot deny after the successful communication).

- Encryption: Plain text + Encryption key + Encryption algorithm = Cipher text
- Decryption: Cipher text + Decryption key + Decryption algorithm = Plain text

Cryptography is of two types based on type of key used, Symmetric key cryptography and Asymmetric key cryptography. Symmetric key cryptography also known as private key cryptography as sender and receiver both uses the same private key for both encryption and decryption [7, 8]. The sender uses a mathematical function that arrange the bits of the plain text using the secret key and receiver use the reverse function to decrypt the message using the same key. Whereas Asymmetric key cryptography known as public key cryptography as sender uses receiver's public key for encryption and receiver uses his own private key for decryption. Symmetric key cryptography requires less memory space and faster than the Asymmetric key cryptographic algorithm. Symmetric key cryptography is a self-certified method to verify the key instead of using Digital signatures for verification of key as Asymmetric key cryptography [5]. As Symmetric key cryptography is a self-certified method, the overall security of this method depends on the length and security of the secret key used for encryption [6]. The well-known Symmetric key cryptography algorithms are AES, DES, 3DES and Blowfish and Asymmetric key algorithms are DSA, ECC, RSA, Digital signature etc.

2 Proposed Algorithm

To achieve the goal of data security we have designed a symmetric key cryptographic algorithm to provide a secure transfer of data through the network. This algorithm encrypt the text twice and encrypt the shared link through which receiver will generate the decryption key. This algorithm incorporate binary addition to generate the first encrypted text, folding method to generate the encryption key and the logical XOR operation with encryption key to generate the final cipher text. Here the 2's complement value of each word length of the plain text and 2's complement of no of word present in this text are send as shared link.

2.1 Encryption Algorithm

Step 1. Read the plain text.
Step 2. Count the length of each word and store it into an array.

Step 3. Add word length with its corresponding letter's ASCII and space after this word i.e. 1st encrypted text.
Step 4. Add all the word lengths and fold the results until it becomes a single digit.
Step 5. Apply the logical XOR operation on the single digit value and all letters ASCII of the 1st encrypted text i.e. cipher text generated.
Step 6. Send the cipher text along with the 2's complement of no of words and 2's complement of each digit of array of word length as the shared link.

2.2 Decryption Algorithm

Step 1. Read the cipher text and the 2's complement of number of words and 2's complement of each digit of array of word length.
Step 2. Generate the number of words from the 2's complement of it and then generate the actual word length from the 2's complement of each elements of the array of word length.
Step 3. Add all digits of the array of world length and fold the results until it becomes a single digit.
Step 4. Apply the logical XOR operation on the single digit value and all letters ASCII of the cipher text i.e. 1st decrypted text.
Step 5. Repeat until all elements of array of word length has traversed.
Step 6. Plain text generated.
Read the word length.

a. If it is not the last word length, then
   - Subtract the word length from each character of the 1st decrypted text from 1st array index to (word length + 1) number of array index.
   - Make (word length + 2) number index as 1st array index.

b. Otherwise, subtract the word length from each character of the 1st decrypted text from 1st index to word length number of array index.

3 Results and Discussion

The overall process of encryption and decryption is demonstrated follows with an example.

3.1 Encryption

3.1.1 Text Encryption

First, read the plain text, which we have to encrypt. Let, the plain text is—"Hello! Bob start 202#". Then count the word length of each word in the plain text and then add the word length with its corresponding letter's ASCII and space after this word. This is the first encrypted text. Now to produce a more secure encrypted text we have encrypted the text second time by applying the logical XOR operation on the first encrypted text and a single digit value produced from the array of word length as follows:

H	E	l	l	o	!	B	o	b	s	t	a	r	t	2	0	2	#
6						3			5					4			

Therefore, here the array of length of each word is:

6	3	5	4

Now, we have added all the digits of the array of word length and applying folding method on these results until it turns into a single digit value.

After addition: 6 + 3 + 5 + 4 = 18

Applying folding method on 18:

- 1st round: 1 + 8 = 9

Single digit encryption key = 9

Final encrypted text: Applying the encryption algorithm on the plain text "Hello! Bob start 202#" as shown in Table 1, we have generate the final cipher text "Gb{{|./L{1 * qpo ∼ p,? = ?."

3.1.2 Key Encryption

In case of symmetric key cryptography, we needs to send the encryption key i.e. the private key with the cipher text. However, if we send the original private key then it is easier for the hacker to generate the original text from this key. Therefore, before we send the key through the network we have encrypted the key by generating the 2's complement of each digit of the array of word length from which we generate the encryption key and 2's complement of number of elements present in this array and then these two as shared link. The process is follows:

Table 1 Character table for encryption process

Plain text characters	ASCII of letters	Word length	ASCII after addition with word length	First encrypted text	Single digit encryption key	ASCII after XOR operation	Final encrypted text
H	72	6	78	N	9	71	G
E	101		107	K		98	B
L	108		114	R		123	{
L	108		114	R		123	{
O	111		117	U		124	|
!	33		39	‘		46	.
	32		38	&		47	/
B	66	3	69	E		76	L
O	111		114	R		123	{
B	98		101	E		108	L
	32		35	#		42	*
S	115	5	120	X		113	Q
T	116		121	Y		112	P
A	97		102	F		111	O
R	114		119	W		126	∼
T	116		121	Y		112	P
	32		37	%		44	,
2	50	4	54	6		63	?
0	48		52	4		61	=
2	50		54	6		63	?
#	35		39	‘		46	.

Therefore, the array of length of each word is:

6	3	5	4

Moreover, number of elements in this array is 4:
2’s complements for 1st shared link:

- 2’s complement of 6 is: 250
- 2’s complement of 3 is: 253
- 2’s complement of 5 is: 251
- 2’s complement of 4 is: 252

2's complement of 4 is: 252

First shared link:

250	253	251	252

Second shared link: 252

3.2 Decryption

Decryption is similar to the encryption process except it is in reverse order. After the cipher text "Gb{{|./L{l*qpo ∼ p,? = ?." and two shared link has sent to the intended receiver the receiver started to decrypt the message as follows:

3.2.1 Key Decryption

Receiver first generate the 2's complement of the second shared link i.e. generate the number of elements in first shared link. Then using this generates the 2's complement of each elements of first shared link that contains the word length. Applying folding method receiver generates the decryption key as follows:

2's complement of 2nd shared link: (generating number of elements presents in the array of word length)

- 2's complement of 252 is: 4

2's complements of each elements of 1st shared link: (generating array of word length)

- 2's complement of 250 is: 6
- 2's complement of 253 is: 3
- 2's complement of 251 is: 5
- 2's complement of 252 is: 4

Array of word length:

6	3	5	4

Now, we have added all the digits of the array of word length and applying folding method on these results until it turns into a single digit value.

After addition: 6 + 3 + 5 + 4 = 18

Applying folding method on 18:

- 1st round: 1 + 8 = 9

Single digit encryption key = 9

3.2.2 Text Decryption Process

Hence, applying the decryption algorithm the receiver decrypts the cipher text and generates the plain text: "Hello! Bob start 202#" as shown in Table 2.

3.3 Final Output

3.3.1 Sender Takes the Plain Text

Sender enters the message to be send and here the message is "Hello! Bob start 202#". After that sender encrypts the text as follows, according to the encryption algorithm we have described earlier. Figure 1 shows the plain text taken by the sender.

Table 2 Character table for decryption process

Cipher text characters	ASCII of letters	Single digit decryption key	ASCII after XOR operation	First decrypted text	Word length for subtraction	ASCII after subtraction	Plain text
G	71	9	78	N	6	72	H
B	98		107	K		101	e
{	123		114	R		108	l
{	123		114	R		108	l
|	124		117	U		111	o
.	46		39	‘		33	!
/	47		38	&		32	
L	76		69	E	3	66	B
{	123		114	R		111	o
L	108		101	E		98	b
*	42		35	#		32	
Q	113		120	X	5	115	s
P	112		121	Y		116	t
O	111		102	F		97	a
~	126		119	W		114	r
P	112		121	Y		116	t
,	44		37	%		32	
?	63		54	6	4	50	2
=	61		52	4		48	0
?	63		54	6		50	2
.	46		39	‘		35	#

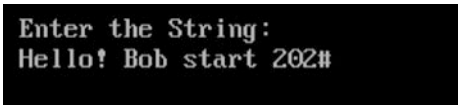


Fig. 1 taking plain text

3.3.2 First Encrypted Text After Addition with the Word Length

Encryption process started with counting the word length of each word in the message and then adding all the characters of the message with their corresponding word length. When a space occurred then adds the length of the word just before the space with ASCII representing space i.e. 32. This is the first encrypted text as shown in Fig. 2.

3.3.3 Final Encrypted Text After XOR Operation with Encryption Key

Before the generation of final encrypted text, we generate the encryption key by applying addition method on the word length of each word of plain text and applying folding method on this result until we get a single digit. Now we apply logical XOR on this single digit with the characters of the first encrypted text and generate the final cipher text. Figure 3 shows the final cipher text after applying all steps of encryption algorithm.

3.3.4 First Decrypted Text After XOR Operation with Decryption Key

The decryption process started with the generation of single digit decryption key from the shared link which is an array containing the 2's complement of the length of each word of the plain text and applying reverse process of key encryption and

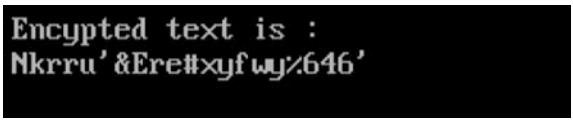


Fig. 2 1st encrypted text

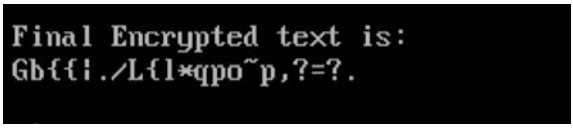


Fig. 3 Final cipher text

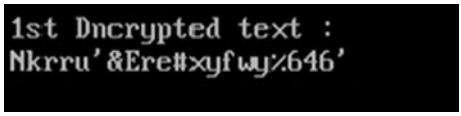


Fig. 4 1st decrypted text

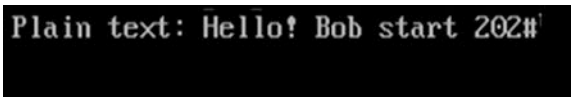


Fig. 5 Plain text

folding method. Then apply logical XOR on this single digit with the characters of the cipher text and generate the first decrypted text as shown in Fig. 4.

3.3.5 Plain Text After Subtraction Operation with Word Length

Now using the information about the word length in the shared link, we subtract the word length from the characters of the first decrypted text and finally the plain text has generated as shown in Fig. 5.

4 Conclusion

A Symmetric key cryptography is how much secure, this rate highly depends on the secrecy of the key being used as the key has to be send with the cipher text. Including this features of Symmetric key cryptography our proposed algorithm provides a way to transfer of key securely as a shared link by encrypting the key. This algorithm generates the cipher text twice to provide more security. Without knowing the proper steps to generating the key from the shared link, it is almost impossible to decipher the cipher text.

References

1. Forouzan, B.A.: Cryptography and Network Security, Special Indian Edition. Tata McGraw-Hill, New Delhi (2007)
2. Kahate, A.: Cryptography and Network Security, 2nd edn. Tata McGraw-Hill, New Delhi (2009)
3. Stallings, W.: Cryptography and Network Security Principles and Practices, 4th edn. Prentice Hall, Pearson Education (2009)
4. Comer, D.E.: Internetworking with TCP/IP Principles, Protocols, and Architectures, vol. 1, 4th edn. ISBN 0-13-018380-6

5. Chandra, S., Paira, S., Alam, Sk.S., Sanyal, G.: A comparative survey of symmetric and asymmetric key cryptography. In: IEEE International Conference on Electronics Communication and Computational Engineering (ICECCE 2014), pp. 83–93
6. Mandal, B., Chandra, S., Alam, Sk.S., Patra, S.S.: A comparative and analytical study on symmetric key cryptography. In: IEEE International Conference on Electronics Communication and Computational Engineering (ICECCE 2014), pp. 131–136
7. S.Chandra, S.Bhattacharyya, S.Paira, S.S.Alam, A study and analysis on symmetric cryptography. In: IEEE International Conference of Science Engineering and Management Research (ICSEMR), IEEE Xplore Digital Library, Print ISBN: 978-1-4799-7614-0 (2014)
8. https://www.princeton.edu/∼achaney/tmve/wiki100k/docs/Symmetric-key_algorithm.html
9. http://searchsoftwarequality.techtarget.com/definition/cryptography

‘Let’s Exercise’: A Context Aware Mobile Agent for Motivating Physical Activity

Saurav Gupta, Sanjay P. Sood and D.K. Jain

Abstract Context-aware Computing, considered as a part of ubiquitous computing, is an upcoming technology that has the potential to be used for improving one’s own health and providing personalized healthcare services. This paper discusses a randomized controlled trial conducted amongst 97 individuals, who were screened for stress and obesity. Out of these, 33 individuals (n = 33) were identified as suffering from both stress and obesity. With the fact that physical activity acts as a catalyst in reducing stress and obesity, the mobile application, ‘Let’s Exercise’ was designed to send context-aware alerts to the users. These alerts motivated and recommended these users to take up physical activity depending upon their operating environment. The 33 users were subject to a four-week observational period, after which a positive behavioral change was observed amongst these individuals. This was due to the increase in the level of physical activity in their daily routines after receiving the contextual alerts. Post the study, the users also showed strong confidence and willingness in the adoption of this technology.

Keywords Context awareness · Computer to physical environment interaction · Computer to human interaction · Ubiquitous computing · mHealth

S. Gupta (✉)
Department of Health Infomatics & Electronics, Engineer, Centre for Development of Advanced Computing, Mohali 160071, India

S.P. Sood
Department of Information Technology, SeMT, Chandigarh Administration, ccChandigarh 160017, India

D.K. Jain
Centre for Development of Advanced Computing, Mohali 160071, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_46

1 Introduction

Context aware computing has recently been gaining grounds in the field of ubiquitous computing due to its characteristic ability to sense and adapt to the changing environment. Context awareness creates a system in which there exists computer to physical environment interaction in real-time and provides decision support to the end-user. There are 03 aspects that are often considered while retrieving contextual information [1]: (i) the computing environment which includes device computing abilities and connectivity options (ii) the user environment which includes location based services and (iii) the physical environment which includes weather, noise etc. Based on this contextual information, the system provides a real-time decision based on a pre-configured criteria. This criterion is defined as rules in the rule library of the system. The rules, when applied on a given contextual information, triggers a defined set of actions which is notified to the user in form of alerts or notifications. This prompts the user to take an appropriate decision.

Globally, there have been proven studies that reflect the use of context aware computing as a solution for improving healthcare services [2–6]. In the Indian landscape, where there have been known issues in the healthcare industry, context aware computing could provide an opportunity to manage and monitor one's individual health on a real time basis [7, 8].

This paper highlights that diseases like stress and obesity, which have been a disturbing problem especially in developing countries like India, can be tackled by sending the users context-aware alerts. These alerts are motivating in nature and recommend certain exercises to the users to encourage them to take up physical activity. Studies have suggested that there exists a common point between stress and obesity [9], that is, both can be reduced by increasing or initiating a physical activity [10]. A look at the physical activity patterns amongst people in India indicates that majority of the populace is living a sedentary lifestyle [10]. Hence the objectives defined as part of this research work were:

- To enable adoption of active and healthy lifestyle amongst the people, while being 'on-the-fly'.
- To bring about a behavioral change amongst the people to take up physical activity.

In Sect. 2 of the paper, the methodology for conducting the research is discussed. It initially identifies the subjects suffering from both stress and obesity. Then, it goes on to discuss the system design and the message database created to motivate people to take up physical activity. Section 3, discusses the results of the study amongst the shortlisted individuals (n = 33) and highlights the trends of the responses. Section 4, concludes the paper and Sect. 5 shows the user interface of the context aware mobile application, 'Let's Exercise'.

2 Methodology

In order to carry out the study, the first step was to identify the subjects who were suffering from both stress and obesity. The second step was to design the context aware system, which would send contextual alerts to the identified subjects. The third step was to create a database of positive motivating messages and a set of physical activities so as to improve adherence rates. The activity in each of these steps are explained as follows.

2.1 Identification of Subjects

For the identification of subjects suffering from obesity and stress, 02 surveys were conducted for 97 randomly selected individuals who were between 25–35 years of age. The first was conducted to identify the individuals suffering from obesity and the second survey was conducted on the same set of individuals to identify those having high degree of stress. Post the surveys, an intersection of both shortlisted sets was done to get individuals suffering from both stress and obesity. These two surveys conducted are explained as follows.

2.1.1 Identification of Subjects Who Were Obese

For identifying people as obese, Body Mass Index (BMI) was used as the criteria. As per World Health Organisation (WHO), "BMI is a simple index of weight-for-height that is commonly used to classify overweight and obesity in adults" [11]. Figure 1 shows the formula for calculating BMI for an individual and Table 1 shows the standard BMI chart.

A total of 97 people were surveyed in the age group of 25–35 years. In this, 52.6 % of the individuals were having normal BMI, 31 people were found to be overweight (i.e. having BMI > 25) and 02 persons were found to be obese (i.e. having BMI > 30). The results of BMI distribution of individuals are shown in Fig. 2 For this survey, individuals who were either overweight or obese were shortlisted. Hence, 33 individuals were shortlisted.

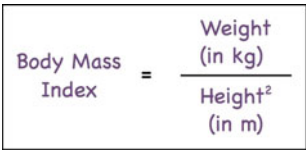
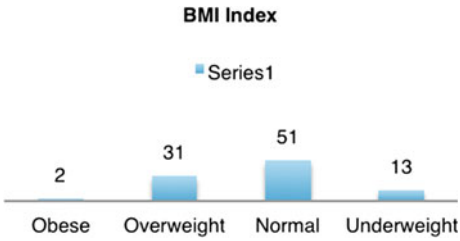


Fig. 1 BMI formula

Table 1 BMI Chart

BMI less than 18.5	Underweight
BMI 18.5–24.99	Healthy weight
BMI 25.00–29.99	Overweight
BMI 30 or more	Obese

Fig. 2 BMI distribution



2.1.2 Identification of Subjects Who Were Under Stress

Globally, there doesn't exist any mathematical relation for calculating stress. However, various survey questionnaires have been designed to identify people suffering from stress and hence, for this study, a pre-defined third party survey was used [12]. The survey questionnaire comprised of 73 questions across 5 sections namely, Physical Indicator, Sleep Indicator, Behavioral Indicator, Emotional Indicator and Personal Habits. A tablet-based application was designed to take inputs from the users. The responses were taken based on a 5-point likert scale. The responses included, 'Strongly Agree', 'Agree', 'Neutral', 'Disagree', and 'Strongly Disagree'. The same 97 individuals, who were screened for obesity, were asked to fill the survey. The responses were compiled and evaluated against the stress calculator [12] as shown in Fig. 3. The results, as shown in Fig. 4a (Physical stress), 4b

Fig. 3 Stress calculator

PERSONAL STRESS LEVELS

	Very Low	Medium	High	Very High	Danger
Physical Indicators					
Point Total	22	30	38	48	54+
Sleep Indicators					
Point Total	5	8	10	12	14+
Behavior Indicators					
Point Total	18	27	36	45	50+
Emotional Indicators					
Point Total	21	29	37	46	55+
Personal Habits					
Point Total	9	15	20	25	30+

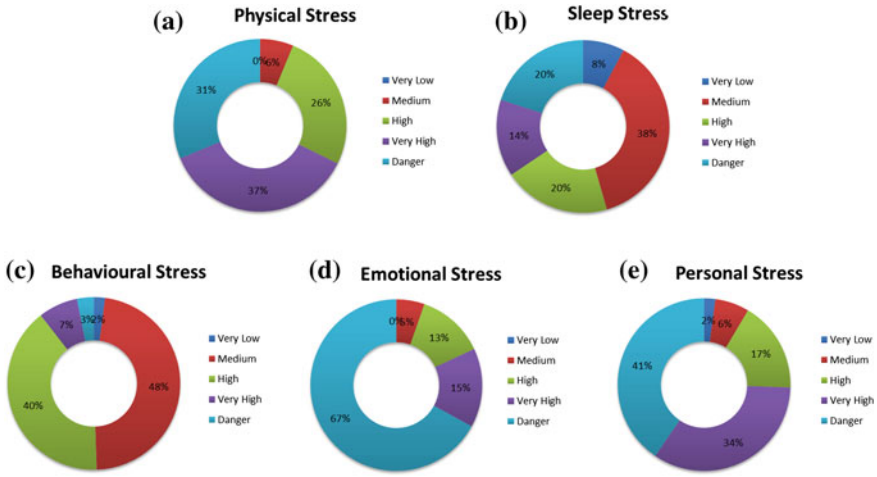


Fig. 4 **a** %Physical stress. **b** %Sleep stres. **c** %Behaviorial stress. **d** %Emotional stress. **e** %Personal stress

(Sleep Stress), 4c (Behavioral Stress), 4d (Emotional stress) and 4e (Personal Stress), highlights the fact that more than 50 % of the people were stressed. Significant percentage of individuals, who were surveyed, had very high degree of 'Emotional' stress (95 %) followed by 'Personal' stress which stood at 92 %. On the other hand, behavioral stress was found to be least significant but still 50 % of the individuals were suffering from it.

After identifying individuals suffering from stress and obesity, an intersection of both surveys was done and subsequently 33 subjects were identified who were suffering from both these health problems. Another observation from this study was that all the individuals who were either overweight or obese were suffering from some form of stress.

2.2 *System Design*

A context aware system, 'Let's Exercise', was designed to motivate the identified individuals, suffering from stress and obesity, to take up physical activity [13]. 'Let's Exercise' is an Android OS based mobile application that uses contextual information for sending alerts to the users. This creative smart application senses and recognizes the contextual environment of the user based on the following sensors: (a) Location, (b) Temperature and (c) Date & Time. The location information, includes latitude and longitude, is fetched from the sensors available on the smart phone namely, GPS, Wi-Fi and using the cell tower (Base Transceiver

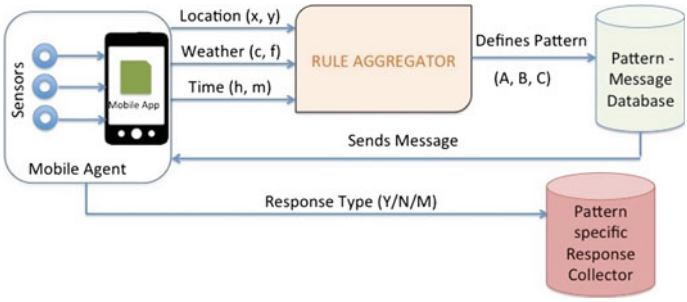


Fig. 5 System architecture

Station) location. Based on the configurations made by the user, the location type of the user was identified as either 'Home', 'Office' or 'Outside'. For sensing the weather parameters of the user's operating environment, the 'Open Weather' API service was integrated with the 'Let's Exercise' mobile application. The parameters included temperature, humidity and forecast. Based on the user options, the temperature was classified as either 'Favorable' or 'Non-favorable'. For sensing date and time, the information from the smart phone was fetched and subsequently, based on user options, an acceptable time zone for receiving the alerts was defined.

In the application, SQL-Lite database was used to store the various activity related messages. This database was synchronized, using web services, with the MS-SQL database at the web server. The system architecture of the context aware system is shown in Fig. 5.

2.3 Creation of a Message Database

It has been a proven fact that sending positive motivational messages to the users results in high adherence rates [14–17]. Hence, a positive message highlighting the benefits of physical activity was appended with the activity related alert. The format of the alert sent to the user was structured as:

Alert = Positive motivational message about Physical Activity + Context-based physical activity recommendation.

A repository of motivational messages was created at the server side and every time an alert was created, it was sent using push notifications. The exercises for enabling physical activity were embedded in the mobile application's database and was retrieved and appended along with the motivational message while generating the alert. The users could configure the 'Let's Exercise' application to receive up to 10 alerts in a day. These alerts were divided according to the acceptable time zone of the user and sent regularly.

3 Results

In a randomized controlled trial, gender distribution of the 33 individuals, identified as both obese and stressed, is shown in Fig. 6. These 33 subjects were asked to install the mobile application, 'Lets Exercise' on their smart phones. The application was configured on smart phones to receive alerts based on users' preferences.

For each alert sent, four response types were provided: (a) Let's do it, (b) Will do but later, (c) Too busy to do it and (d) No, Thanks. The responses of each alert from every individual were recorded in the database. Over an observational period of 04 weeks, users' response was observed. Figures 7a–d below show the week-wise responses of the users. In this, the selection of responses 'Lets do it' and 'Will do but later' signifies the intent of the user to take up the suggested physical activity. Hence, combining these two response-types as 'Yes' (or Positive Response), a response trend was plotted for the 04-week observational period as shown in Fig. 8.

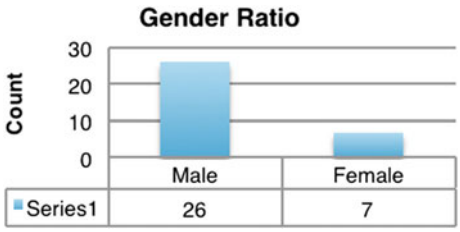


Fig. 6 Gender distribution

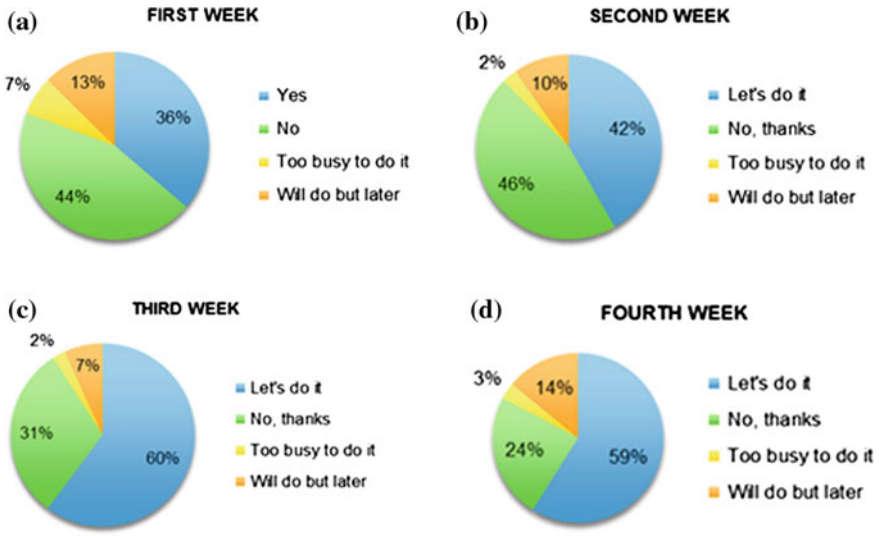


Fig. 7 **a** Week 1 response. **b** Week 2 response. **c** Week 3 response. **d** Week 4 response

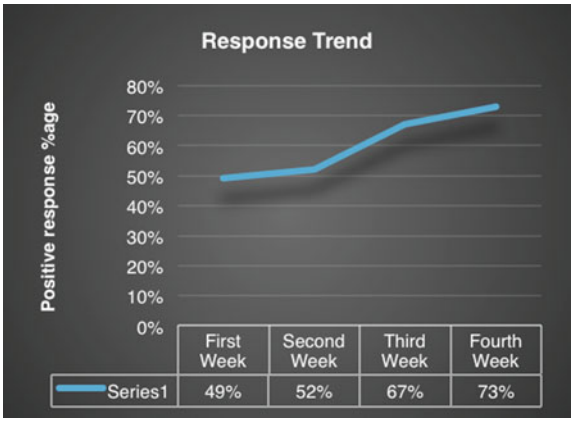


Fig. 8 Response trend for response type 'YES'

During the study, the responses in affirmative increased from 49 % in the first week to 73 % at the end of the 4-week. The response trend depicts a steady increase in the motivational levels amongst the users to take up physical activity. Thus, a behavioral change was observed amongst the subjects with the use of this technology. Along with this study, a survey was conducted about the use of this technology amongst the users. More than 80 % of the users expressed satisfaction and confidence while using this technology. Also, 61 % of the users expressed their willingness for the continued usage of this technology as part of their daily routines.

Recently, the BMI index has been revised in which persons having BMI greater than 25 are considered obese. This has lead to an increase in the prevalence rates for obesity in India [18] and hence, calls for greater adherence to physical activity. Sensitization needs to be created using technology amongst the individuals on the importance of undertaking physical activities.

4 Conclusion

Context-aware computing provides a platform for enabling a computer-to-physical environment interaction in a ubiquitous framework. This new-age style of computing opens up new possibilities for providing personalized healthcare. This paper uses contextual computing to design a mobile agent, which helps users reduce stress and obesity by motivating them to take up physical activity. The trial that was conducted with the shortlisted 33 subjects supports the claim that ubiquitous technologies can be meaningful in improving one's own health. At the end of the study period, a behavioral change was observed amongst the users, which showed increase in adherence to physical activity. This research concludes with a finding that context aware computing is an effective tool for solving real life health problems.

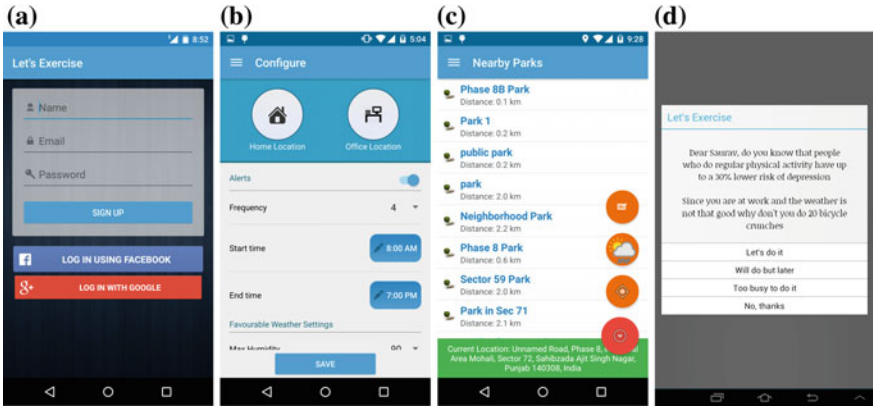


Fig. 9 **a** Login, **b** configuration, **c** nearby parks with info, **d** contextual alert

5 User Interface

The user interface of the mobile application, 'Let's Exercise' is shown in Fig. 9a–d. The interface has been designed based on 'Material Design' released by Google for Android development.

Based on the feedback from the individuals, the mobile application, 'Let's Exercise', has been made available for free download from the Google Play Store for all Indian citizens.

References

1. Vrbaski, M., Petriu, D.: Toward a context awareness framework for healthcare applications. In: IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA), doi:10.1109/MeMeA.2013.6549764, May 2013
2. Gardini, M. et al.: Clariisa, a Context-Aware Framework Based on Geolocation for a Health Care Governance System. In: IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom) (2013)
3. Tobón, D.P., Falk, T.H., Maier, M.: Context awareness in WBANs: a survey on medical and non-medical applications. In: IEEE Wireless Communications, Aug 2013
4. Lin, Y., Jessurun, J., Vries, B., Timmermans, H.: Motivate: towards Context-Aware Recommendation Mobile System for Healthy Living. In: 5th International Conference on Pervasive Computing Technologies for Healthcare (Pervasive Health) and Workshops 2011 (IEEE) (2011)
5. Lin, Y., Jessurun, A.J., Vries, B., Timmermans, H.: A context- aware persuasive system for active living-simulator system design. In: Conference proceeding of 10th International Conference on Design & Decision Support Systems in Architecture and Urban Planning, pp. 1–12 (2010)
6. Khan, M., Lee, S.: Need for a context-aware personalized health intervention system to ensure long-term behavior change to prevent obesity. In: 5th International Workshop on Software Engineering in Health Care (SEHC 2013), May 2013

7. Viswanathan, H., Chen, B., Pompili, D.: Research challenges in computation communication, and context awareness for ubiquitous healthcare. IEEE Communications Magazine, May 2012
8. Seppälä, A., Nykänen, P., Ruotsalainen, P.: Development of Personal Wellness Information Model for Pervasive Healthcare. J. Comput. Netw. Commun. **2012**(596749) (2012)
9. Torres, S.J., Nowson, C.A.: Relationship between stress, eating behavior, and obesity. Nutrition **23**(11–12):887–894 (2007)
10. Anjana, et al.: Physical activity and inactivity patterns in India—results from the ICMR-INDIAB study (Phase-1). Intern. J. Behav. Nutr. Phys. Act. **11**(26), 2014 (2014)
11. WHO. link: http://www.who.int/mediacentre/factsheets/fs311/en/. Accessed 03 Feb 2015
12. Stress Indicator Questionnaire. The Counseling team International. link: www.thecounselingteam.com. Accessed 03 Feb 2015
13. Gupta, S., Sood, S.P.: Context aware mobile agent for reducing stress and obesity by motivating physical activity-a design approach. IEEE 9th IndiaCom, March 2015
14. Krepsemail, G.L., et al.: Development and validation of motivational messages to improve prescription medication adherence for patients with chronic health problems. Patient Educ. Couns. **83**(3), 375–381 (2011)
15. Kim, C.M., Keller, J.M.: Effects of motivational and volitional email messages (MVEM) with personal messages on undergraduate students' motivation, study habits and achievement. British J. Educ. Technol. **39**(1), 2008 (2008). doi:10.1111/j.1467-8535.2007.00701.x
16. Visser, J., Keller, J.M.: The clinical use of motivational messages: an inquiry into the validity of the ARCS model of motivational design. Instr. Sci. **19**(6), 467–500 (1990)
17. Mulgund, P., et al.: Personalized medication adherence motivating and reminding system (PMAMRS). Adv. Impact Des. Sci. Mov. Theory Pract. Lect. Notes Comput. Sci. **8463**, 448–452 (2014)
18. Aziz, N., Kallur, S.D., Nirmalan, P.K.: Implications of the revised consensus body mass indices for asian indians on clinical obstetric practice. J. Clin. Diagn. Res. (2014). doi:10.7860/JCDR/2014/8062.4212

Real Time Smart, Intelligent and Novel Embedded Vehicle Interceptor for Security Applications

H.V. Ravish Aradhya, Aravindkumar Gumtaj and Mohana

Abstract Now a day, in cities so many people may lose their lives in accident due to driving the vehicles over speed or sometimes because of violation of the traffic rules. Currently in all cities there is a speed limit for vehicles in different locations. The vehicles should follow the rule and drive within these limits, if violates it may be penalized by the traffic officials. Present system is camera, radar based and human controlled interceptor. If multiple vehicles violate the rules simultaneously then current system is finding difficult to trap them and it leads more errors. However, once the vehicle crosses the official, once again the vehicle may cross the speed limits. The police official will never notice that, the person has violated the rules. In such cases, it is needed to track the vehicles throughout the range. This paper addresses the problem by giving a smart solution and novel approach to accomplish this task by designing, an efficient low cost embedded system. This will automatically generate the violation report and sends a penalty message to vehicle owner immediately. And it is fully automated and there is no need for the traffic officials to manually monitor the vehicles round the clock. The system is implemented using ARM processor LPC2129, ZigBee transceiver P20 module, microcontroller 89C51, GPS634R and GSM SIM300 modules.

Keywords Vehicle interceptor · ARM processor · Zigbee · GPS · GSM

H.V.R. Aradhya (✉) · A. Gumtaj
Department of Electronics & Communication Engineering,
R.V. College of Engineering, Bangalore 560059, India

Mohana
Department of Telecommunication Engineering,
R.V. College of Engineering, Bangalore 560059, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_47

1 Introduction

In almost all the cities, government has set many traffic rules for the vehicles to avoid congestion and accidents. In current technology the traffic system is camera, radar based and human controlled. If any of the vehicles violate the traffic rules the driver is no expected to cross the speed limit, if done so, the traffic officials patrolling the area will penalize the driver. There is lot of scope to improve the efficiency of the system. The designed system will overcome the current real time problems by automating the vehicle interceptor. The technology is implemented using two control systems. First is intra vehicular module which consists of ARM7 (LPC2129), ZigBee receiver, GPS (global positioning system), and GSM (global system for mobile) [1, 2]. This entire module is placed inside the vehicle. Second control system consists of microcontroller 89C51 and ZigBee transmitter which are placed on the road. Where it will send the signals to all vehicles by using ZigBee transmitter with its range. When the vehicles come within the range the intra vehicular module will capture data for further processing.

Figure 1 shows the complete system model of smart, intelligent and novel embedded vehicle interceptor.

2 System Functions and Composition

This section describes brief theory of the modules used in the system.

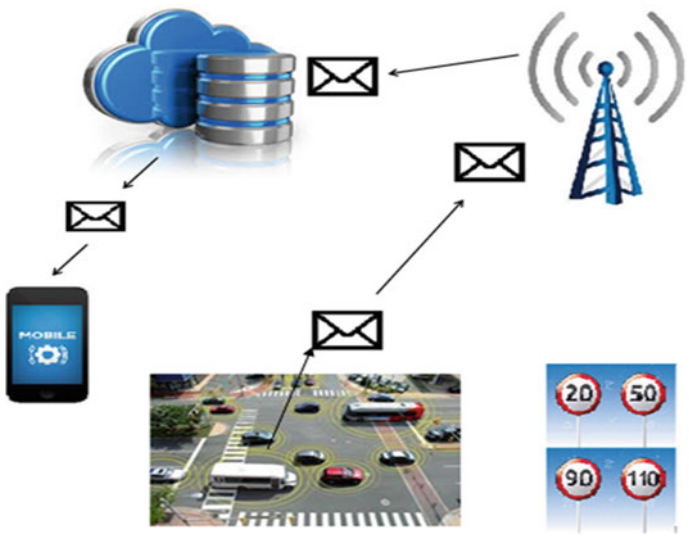


Fig. 1

2.1 Microcontroller and ZigBee Transmitter

Interceptor transmitter module is placed on the road. It consists of microcontroller along with ZigBee transmitter, which sends the signals continuously within the range.

2.2 Intra Vehicular Module

It consists of ARM processor, ZigBee receiver, GPS, GSM module, LCD display.

2.3 ARM Processor

This is the heart of the intra vehicular module placed in all vehicles. This processor is connected with all peripherals such as ZigBee receiver, GSM, GPS and LCD display. Initially it receives the signal from ZigBee receiver for further processing.

2.4 ZigBee Receiver

If the vehicle enters within the range, it receives the signal from ZigBee transmitter.

2.5 GPS

It is used to calculate the speed of vehicle. It gives latitude and longitude of current location of the vehicle. The speed of vehicle information will continuously update in the processor.

2.6 GSM Module

It will send the message to the RTO department, only if vehicle violates the rule. This is controlled by ARM processor.

2.7 LCD Display

It is interfaced with ARM processor. Mainly it is used to display the information.

3 Design and Implementation

This section describes the implementation of all modules used in the designed system.

3.1 Design of Intravehicular Module

Figure 2 shows the block diagram of Intravehicular module. It consists of ARM processor, ZigBee, GSM, GPS system and LCD display. ARM processor LPC2129 is main processing unit in this module. Once the vehicle has started, the ZigBee device will turn on automatically. During the vehicle movement, ZigBee device will continuously waiting for the information. Once the vehicle enters into speed limit region (which is displayed on the road) ZigBee receives speed limit information and its current location. Then it will send the information to ARM processor which will

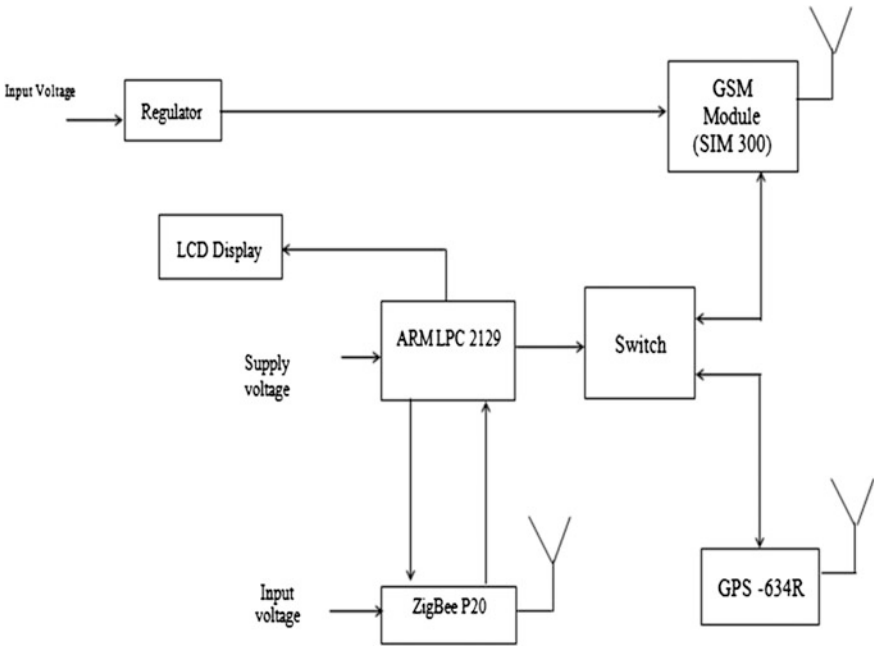


Fig. 2

continuously calculate the vehicle speed through GPS device [3, 4, 5]. And it will compare with the ZigBee data information. If the vehicle violates the speed limit, GSM device will turn on. The processor uses AT commands to communicate with the GSM device and sends message to the GSM device. This device will send offence message to the RTO department. The communication between the processor with other devices is UART communication and its baud rate is 9600 bps.

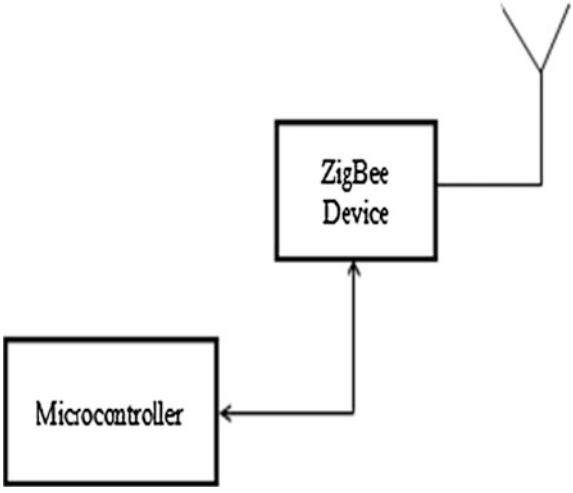
3.2 Design of Embedded Device on the Road

Figure 3 shows the block diagram of embedded device installed on the road. It consists of microcontroller 89C51, which gives the speed limit information to ZigBee transmitter where it broadcasts this information on the road.

3.3 Design of Server Module

Figure 4 shows the block diagram of server module it is installed in RTO department. It consists of database and GSM. The database is designed using MS Access, consists of two tables [6]. First table contains the details of vehicle owners' information such as driving license ID, Name, Address, Mobile number; Email ID etc. second table contains offence details of the vehicles. Communication is established between server and GSM module using serial port communication. User Interface is developed using Visual Basic 6.0 software. The GSM continuously wait for the messages. A message consists of different packets, such as vehicle number, offence place, date and time. The software reads this message by using AT

Fig. 3



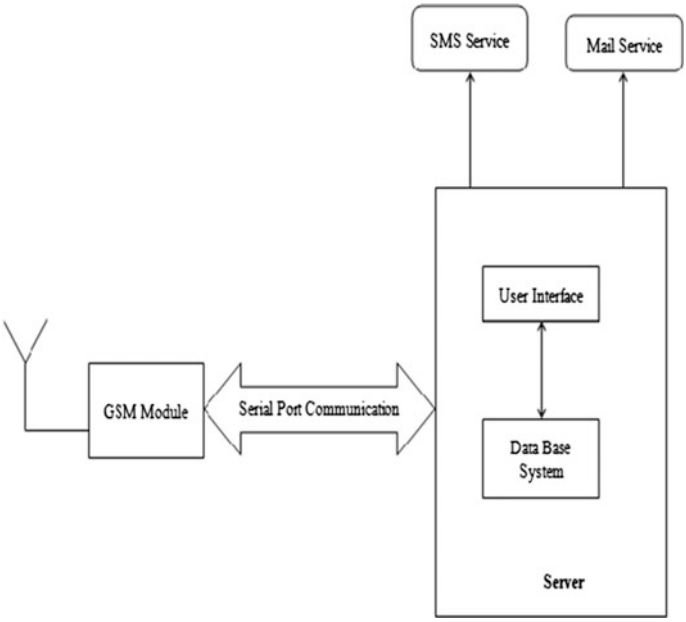


Fig. 4

commands and finds the vehicle details in the database by using vehicle number. Once it gets the details, it generates the penalty message and offence mail. These will be sent to the vehicle owner automatically.

4 Hardware and Software Design Flow

4.1 Design Flow of Intravehicular Module

Figure 5 shows the flow chart for intravehicular module design. Once the vehicle starts the ZigBee device turned on and it waits for the information. When vehicle moves on speed limit region, ZigBee device receives the information of speed limit. Then processor receives this information and turns on the GPS to find out the speed of the vehicle. If the vehicle violating the speed limit, processor immediately activates the GSM and sends violation message to RTO department through GSM module.

4.2 Design Flow of Server Module

Figure 6 shows the flow diagram of server module. An operator start the server, command has sent to the GSM module to activate the serial port communication

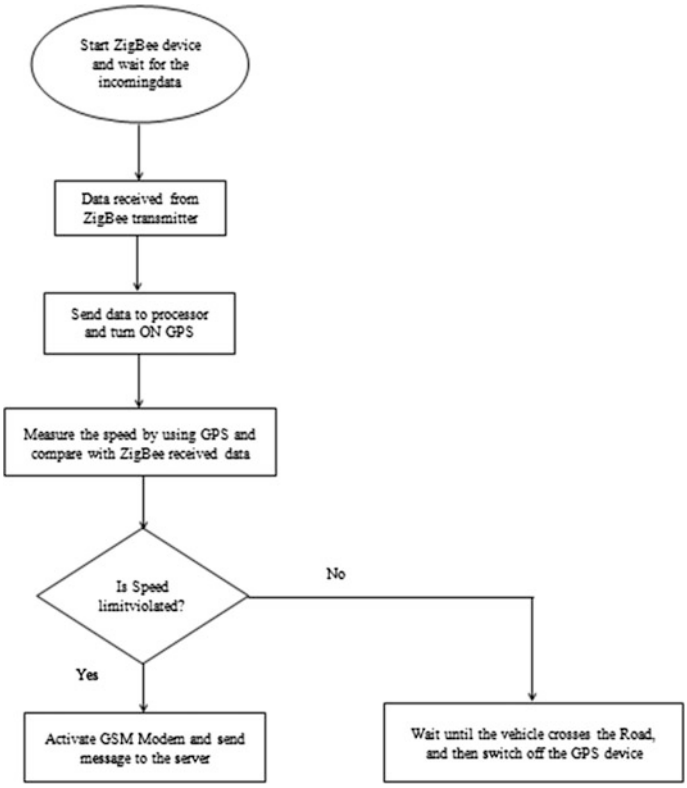


Fig. 5

and read mode operation [7]. GSM module is waiting to receive violated message from the vehicles. Once the message is received, it appears on the user interface. Simultaneously software searches the vehicle details in the database. Server gets the details, the penalty message and penalty mail sends to the vehicle owner.

5 Results

5.1 Results of Intravehicular Module

Figure 7 shows hardware implementation of intra vehicular module. LCD display, displays the complete information of vehicle such as speed, voilated information etc.

Figure 8 shows hardware implementation of embedded device installed on the road. Microcontroller is connected to Zigbee device, it continuously transmits the information on road within its range.

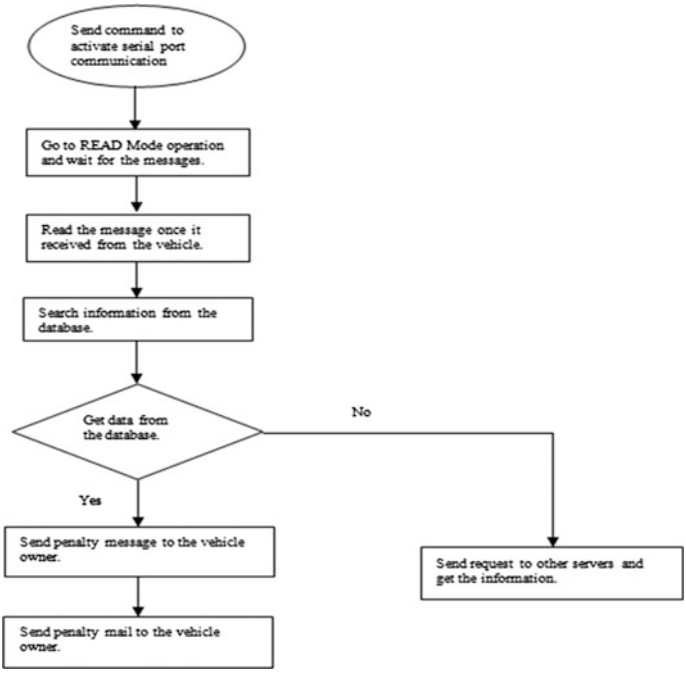


Fig. 6

Fig. 7

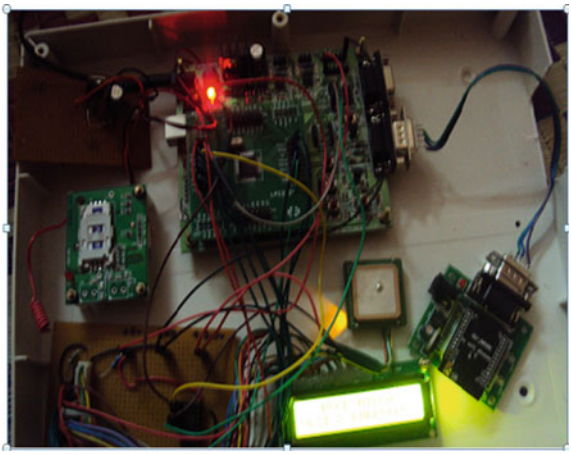


Figure 9 shows the server implementation. Server is connected to GSM module, it receives the voilated messages from vehicles. Software does the further operations after receiving the messages.

Fig. 8

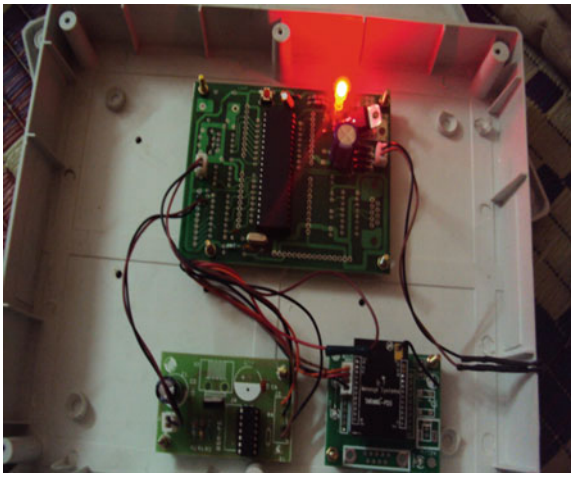
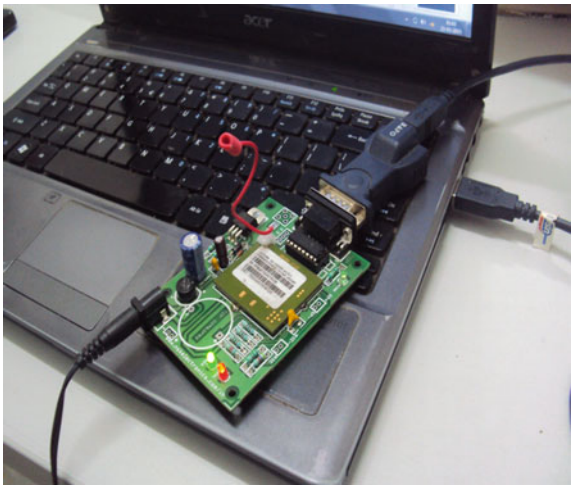


Fig. 9



5.2 *Results of Server Module*

Figure 10 shows the Creation of user interface. It contains port list, start and stop options.

Figure 11 shows the Selection of port number. Operator selects port number and start option.

Figure 12 shows the read mode activation. Once the user selects the start option the serial port communication established between GSM module and server. GSM module enters into read mode.

Figure 13 shows the Waiting for vehicle violation message. Once GSM module enters into the read mode, it starts waiting for the violation message from the

Fig. 10

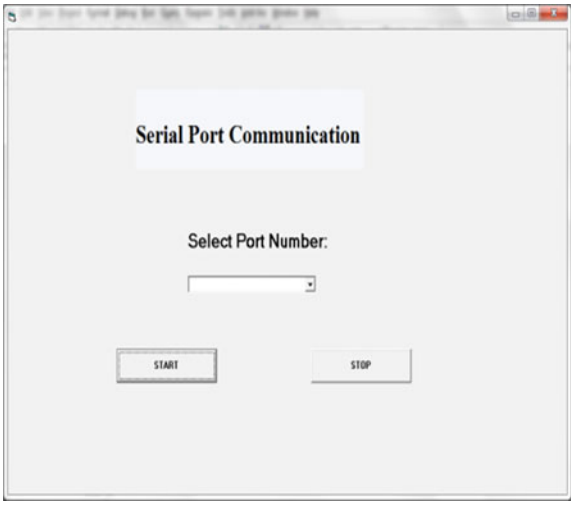


Fig. 11

vehicles. When the GSM module receives the message, it reads the message and it consists of vehicle number, offence location, date and time.

Figure 14 shows the Searching vehicles information in database. Software searches the complete information of the vehicles such as mobile number, mail id etc.

Figure 15 shows the Sending penalty message to vehicle owner. Penalty message contains the violated amount and due date.

Figure 16 shows the sending mail to vehicle owner. The software automatically sends mail to the vehicle owner.

Figure 17 shows the Database of vehicles. It contains vehicle number, owner's details such as name, license id, address, mobile number, mail id etc.

Fig. 12

Fig. 13

Fig. 14

Fig. 15

Fig. 16

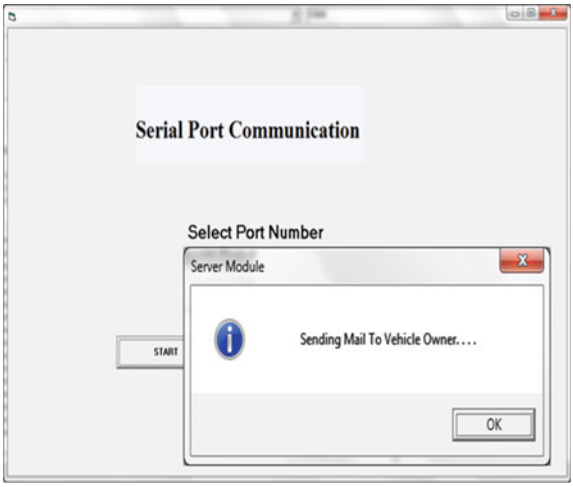


Fig. 17

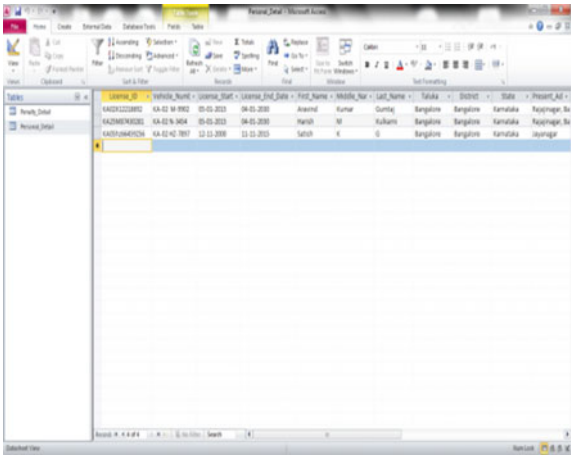


Fig. 18

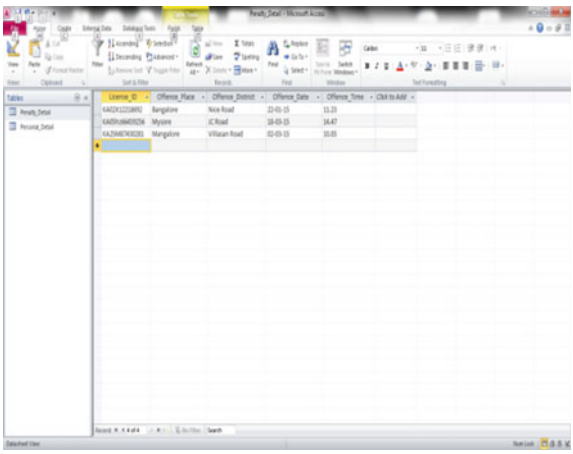


Figure 18 shows the Database of vehicle penalty information. It contains vehicle number, offence location, date and time.

6 Conclusion

In this paper smart, intelligent, novel embedded vehicle interceptor system is designed and implemented using ARM processor LPC2129, ZigBee transceiver P20 module, microcontroller 89C51, GPS634R and GSM SIM300 modules. The designed system is low cost, high performance; it eliminates the human intervention to control the speed limit and automatically monitors the vehicle speed with the specified locations round the clock. The implemented system controls the bribe and the information is automatically updated in the server. It leads to more profitable and helpful to the traffic department and reduces the accidents.

References

1. Dissanayake, S.D.., Karunasekara, P.P.C.R.., Lakmanaarachchi, D.D.., Rathnayaka, A.J.D.., Samarasinghe, A.T.L.K.: Zigbee wireless vehicular identification and authentication system. In: 2008 4th IEEE International Conference on Information and Automation for Sustainability, 12–14 Dec. 2008, Colombo, pp. 257–260. ISBN: 978-1-4244-2899-1, doi:10.1109/ICIAFS.2008.4783998
2. Benjamin, F.P., Yao, N.A., Adjetey, S.R., Appah, B.: Design and development of GPS/GSM based vehicle tracking and alert system for commercial inter-city buses. In: 2012 IEEE 4th International Conference on on Adaptive Science and Technology(ICAST), 25–27 Oct. 2012, Kumasi, pp. 1–6. ISBN:978-1-4673-787-7, doi:10.1109/ICASTech.2012.6381056
3. Lihe, Lvfang, et all.: Design of intelligent mobile vehicle checking system based on ARM7. In: 2010 IEEE 2nd International Conference on Advanced Computer Control (ICACC), 27–29 March 2010, Shenyang, vol. 2. ISBN:978-1-4244-5845-5, doi:10.1109/ICACC.2010.5487192

4. Sattibabu, G., Satyanarayan, B.V.V., Kona, V.V.S.: Automatic vehicle speed control with wireless in-vehicle road sign delivery system using ARM7. Int. J. Technol. Enhancements Emerg. Eng. Res. (IJTEEE) **2**(8), 32–34 (2014). ISSN:2347-4289
5. Lokhande, A., Bahe, S., Kumar, B.: Accident identification using ARM-7, GPS and GSM. Discov. Publ. **18**(53), 81–84 (May 9 2014)
6. Fleischer, P.B., et all.: Design and development of GPS/GSM based vehicle tracking and alert system for commercial intr-city buses. In: 4th IEEE International Conference on Adaptive Science & Technology (ICAST), 25–27 Oct 2012, Kumasi, pp. 1–6. ISBN:978-1-4673-4787-7, doi: 10.1109/ICASTech.2012.6381056
7. Singh, S., Chawla, R., Singh, H.: Intelligent speed violation detection system. Int. J. Eng. Tech. Res. (IJETR) **2**(1) (January 2014). ISSN:2321–0869

1-Wire Communication Protocol for Debugging Modem Chipsets

T. Ninikrishna, Soni Sahu, Rathina Balan Thalaiappan, Siri Haricharan and R.K. Manjunath

Abstract 1-Wire communication protocol was basically designed for short distance communication between very close by devices. This paper presents how the 1-wire protocol can be used for debugging an ARM based master Modem chipset through a slave device. The master and slave devices here are the Modem chipsets of mobile phones, of which master can be any new Modem in mobile platforms and the slave is another Modem chipset connected to the master using the single wire interface. For a new mobile chipset, getting a debugging environment is a desirable task which is presented in this paper. A minimum debugging environment and a simpler way of logging is desired. For this, a debugging environment is developed that would have a better capability of detecting the faults and errors and correct them in any new bring up Modem board for mobile phone platforms used in Intel. The general purpose input output pins are configured to work in the early boot stage of the phone. The whole implementation allows the slave to keep track of the master in the early boot process of the master. Since at this stage the operating system and the interrupts of the Modem would not be initialized, a slave is used to debug the master through a 1-wire communication protocol. A Digital Storage Oscilloscope is used to check the arrival of bits at the GPIO of the master in the early boot process and slave is checked for the same bits using an external debugger Trace32 and attention commands in Intel standard Tool for Hardware interaction. This implementation is used in Intel Mobile Communications for chipset bring up camp and for chipset validation process.

Keywords 1-Wire · 4G · ARM · Communication protocol · Capture and compare · Debug · External interrupt · Single wire interface (SWI) · Timing · Trace32

T. Ninikrishna (✉) · R.K. Manjunath
Department of TE, R.V. College of Engg, Mysore Road, 560029 Bengaluru, India

S. Sahu · R.B. Thalaiappan · S. Haricharan
Intel Mobile Communications, Core Drivers Team, Bengaluru 560066, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_48

1 Introduction

Single wire interface (SWI) communication protocol between master and slave device allows interaction between them and provides solution for time relationships, control of master and keeping track of the master device [1]. Debugging of master modem chipset is usually done using an external debugger interface (e.g. JTAG interface) of the System-on-Chip (SOC). For debugging, all SOCs have a debug bus and a connector to the external debugging hardware. This bus is used for the control and trace debug of the master chipsets. But in early boot process this procedure cannot be applied. There should be some technique to debug the SOCs in early boot stage of operation. Intel is using an external debugger interface for debugging the bring up mobile chipset in the pre-initialization stage, which is very complex. Configuring the debugger at this level is made easy with the new technology presented in this paper, using 1-wire.

This paper presents a solution for debugging a master Modem through a slave using a 1-wire technology in the early boot stage. In a 1-wire technology, data is encoded in the width (or time) and can be decoded with the knowledge of the time period or width of the waveform. Master and slave can act as transmitter and receiver, respectively, and only one directional transfer can happen at a time. This half duplex transmission is initiated by master. In this work, we use this technology to transfer a byte to the slave via the single wire.

The intention is to keep track of the master Modem-board in the early initialization stage and check on the boot sequence routine in the master. This is done by configuring the General Purpose Input/output (GPIO) pins in the ARM based Modem board. For a new modem, having a debugging environment is a difficult procedure to be implemented. A debugger interface should work and depend on the actual trace or flow of the master. Always a less complicated debugging environment is desired with limited peripherals and easier and faster setup running on a new chip. For this, a debugging environment is to be developed that would have a better capability of detecting the faults and errors and correct them in any new bring up Modem board. This 1-wire interface provides this where trace points are set at known locations. The baseband chipset in Intel Mobile Communications provides this new single wire interface to have an efficient debug environment and have a parallel development in chipset validation along with the use of external debugger like Trace32.

2 Related Work

The initial use of 1-wire technology was to facilitate communication between nearby devices and auxiliary memory on a port. Later, this technology was developed for network applications. Maxim Integrated Products used 1-Wire networks using twisted-pair wire beyond the limits but care needs to be taken on the

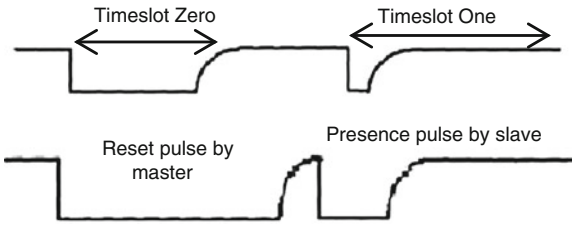


Fig. 1 Dallas 1-wire technology using timeslots for data communication [3]

radius and weight of the network [2]. Studies were made by Maxim Integrated Products for using a microprocessor UART for implementing a master which shows the relationship between 1-wire signals and the UART. Most 1-wire devices give small amounts of data and it is convenient to use GPIO to act as the master [1]. Dallas 1-wire technology uses a single wire interface to communicate the data. Data transfer is initiated by the master device, which initially resets the bus by pulling the bus low and after master releases the line, slave responds to this reset by pulling it for another time duration which shows the presence.

Rick Downs has mentioned in his paper that once master identifies the presence of the slave device, data is communicated between them in time slots. Figure 1 shows this process [3]. Wu Zhuang, in his paper, has mentioned about 1-Wire bus technology using ROM search which is based on binary tree and gives the process details of the search. 1-wire devices should be driven at the correct time by the single wire and the host must be able to distinguish the various slots of the slave [4]. Single wire has the advantages of less I/O ports, low cost, simple structure, easy expansion and maintenance. These characteristics have made it easy to implement the search algorithm inside the 1-wire device by Yu Gui. Standard 1-wire protocol operates in a strict pattern and timing sequence. The protocol has four signal types: reset pulse, write 0 pulse, write 1 pulse, and the read data from the line [5]. Single wire technology has also been used to minimize the number of input/output to control the peripheral devices, which is presented in the paper by Michael Bairanzade [6]. Keeping all these ideas in mind, 1-wire interface protocol is developed for debugging a master mobile chipset using a slave in the early initialization stage of the mobile phone.

3 Process Methodology

This project has few steps in which the work is carried out. The basic sequence can be briefed as: (1) Building a base repository and creating a view of the repository so that the new additions can be implemented onto the repository and committed (2) Compiling the created view of the base which forms an area where object files and symbols are generated (3) Loading the generated files onto the Modem Hardware to confirm the behavior of the Modem as per the repository (4) Interaction

with the hardware using manually generated test commands to check the functionality of the implementation (5) Debugging the Modem using an external debugger interface and correcting the errors in implementation, if any. (6) Check the data at the pins using an oscilloscope. To implement these steps, various tools are used at Intel. We can quickly go through each step and see how the implementation is carried out using them.

3.1 Clearcase Repository and Views

IBM's rational Clearcase is a group of software tools to support for Software Configuration Management (SCM) of the source codes and all the software developments. It forms the base for many businesses and handles projects with many developers. This project is implemented by first creating a dynamic view for the SWI debug process. A successful compilation will generate the object files which can be viewed in details pane. Virtual Network Computing (VNC) Viewer can be used for compilations. VNC is a system of desktop sharing process that is used to control a computer remotely.

3.2 Hardware Behaviour and Interaction

The process by which the compiled code is transferred to the Modem is called as flashing the Hardware. This is accomplished by a tool called Flash Tool. The flash files, with extension. fls, are loaded onto the Hardware using this Flash Tool. Each software release has its recommended Flash Tool version to use for the Modem chipset.

Intel uses a tool to interact with the hardware chipset and can be used as an interface to test the Modem with a set of Attention (At) commands. Hyper terminal or Putty can replace this tool. AT commands are instructions that are used to control a Modem. Sending and Receiving packets can be tested using these commands using the Intel standard API for test interface.

3.3 In-Circuit Debugger (ICD)

Many processors provide and on-chip debugging system. Examples are the JTAG interface for ARM and PowerPC family. The on-chip debugging system lets us view the read/write registers, real time and single-step execution, hardware breakpoints and features of how triggering works, assembler debugging, on-chip breakpoints, Multiprocessor debugging and so on [7]. The tool, Trace32-ICD, is used for this project for carrying out the debug procedure.

3.4 Use of SWI for Debugging the Master

The Debug-SWI protocol is implemented in order to get the execution of Master Modem visible on the slave device. Unlike the generic SWI protocol, which is used to communicate with devices like weather instruments and digital thermometers, the Debug-SWI process is used to monitor the master using the slave. A single wire connected from the Master to the Slave will provide a track of master at the slave side and let the user monitor the master. This driver falls under the Intel's Board Support package (BSP) subcomponent in the Clearcase Versioned Object Base (VOBs). There are many dependencies of SWI driver component and all these dependencies are various components in the VoBs. These dependencies are called as entities and are declared using the Intel standard API (Application Program Interface) layer. When a phone starts, all these entities get initialized and ultimately, when all the modules run without any interruption, the phone boots up. Each time a module is initialized, an entity name corresponding to this module, can be sent to the slave via the single wire. Thus master is monitored through the slave for any sort of failures. Figure 2 shows the final hardware setup which would be used to debug the master using the slave device through the single wire interface. Figure 3 shows the methodology adopted for the transmission.

3.5 Board Support Package (BSP) Subcomponent

This subcomponent, for this project, includes Single Wire Interface (SWI), Pad Control API, Capture and Compare API, Inter-Integrated Circuit (I2C) and so on. Above mentioned components are important for this project.

3.5.1 Intel Pad Control API

This Pad control API driver is used to configure various pads in the Modem. A pad can be a GPIO or connection to one of the many inputs and outputs from chipset's internal blocks. A symbolic pad name defines the function used in the driver and is the identification to the pad exposed. The symbolic pad name for this project is GPIO_SWI_MASTER. One pad can be controlled by one user at a time and this

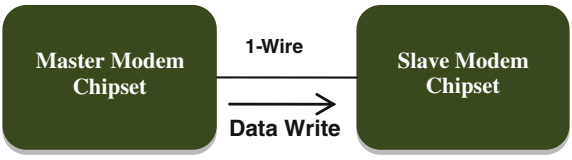


Fig. 2 Final hardware setup

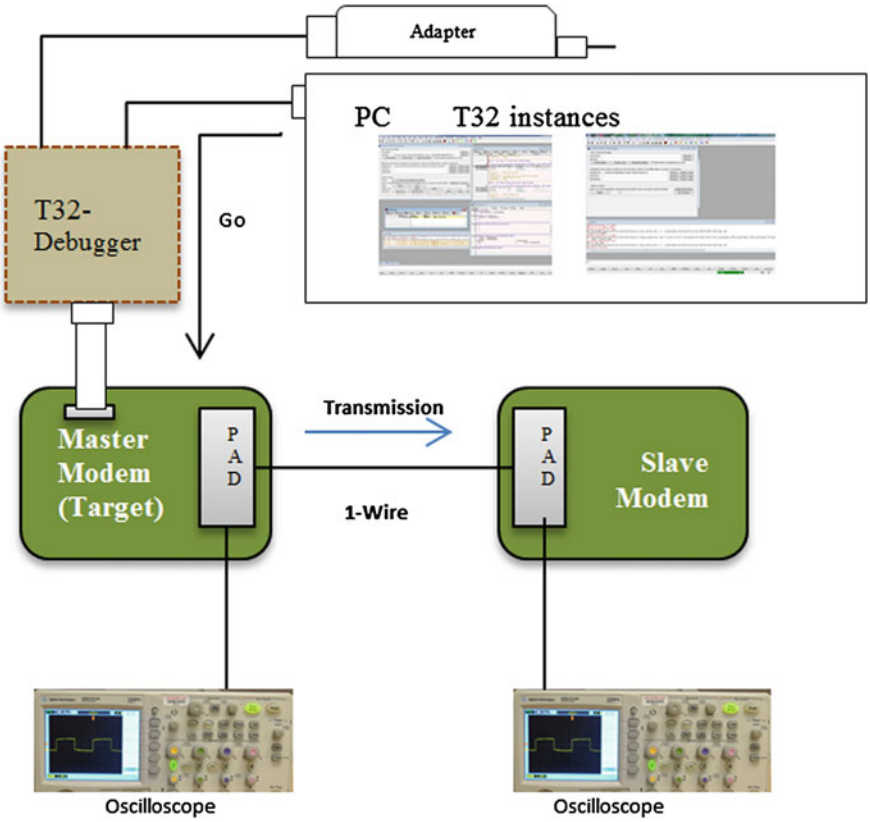


Fig. 3 Hardware setup required for implementation of the protocol

driver takes care of this. To get the pad active, OPER_ACTIVATE operation is chosen to activate the chosen pad. User can request read/write access of GPIO though this API.

3.5.2 Capture and Compare API

This driver is a part of BSP subcomponent and is used for counting pulses. The baseband chipset provides these units for timers, capture and compare features. This driver is responsible for capturing and comparing the hardware configuration. It provides a interfaces for a set of use cases to other modules like time services, and external interrupts. In this project, the slave device uses this timer for its functionality. A software timer would not be accurate and moreover, process may get interrupted by high priority interrupts. So we use a hardware timer which is more accurate.

3.6 *Other Dependencies*

Apart from pad control and capture and compare unit, this implementation needs to depend on many other units. A brief note on those dependencies can be given as below:

3.6.1 Intel Standard API for External Interrupts

This Module is used to provide external interrupt services to other driver components. The functionality provided by this driver allows for allocation, configuring and releasing the external interrupt resource. The function allocate_lisr() from this module allocates the interrupt resource, control_lisr() configures the settings for interrupt for the slave device.

3.6.2 Intel Standard API for Entity Declaration

This module is a layer that supports universal interface to access the debug SWI module through the I/O framework. This interface is exposed to customers too. The framework handles the requests from user and holds the hardware profiles and handles the communication to the hardware block (Fig. 4). The entities which were mentioned in the previous discussion are declared and defined in this layer. Entity is a framework in which driver will register itself and will take care of the boot order of the drivers. The Boot component handles the boot-up and shutdown of the system. This component provides a way to hook various entities into the exact position in the boot-up of the Modem by defining their dependencies to other entities.

The macro ENTITY_DEFINE defines the entity which has to boot up and shutdown with the system. In this way the debug SWI (DSWI) entity is declared and

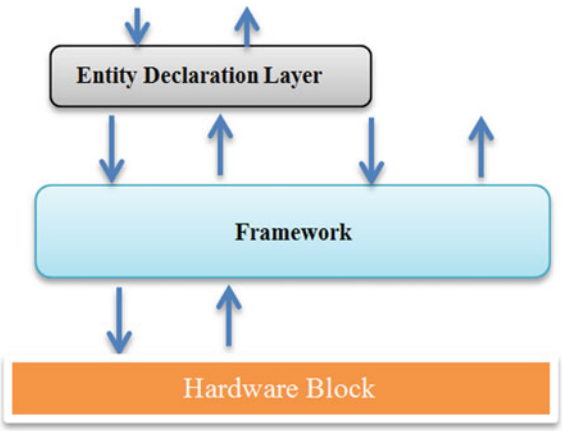


Fig. 4 Layer of Intel standard API for entity declaration

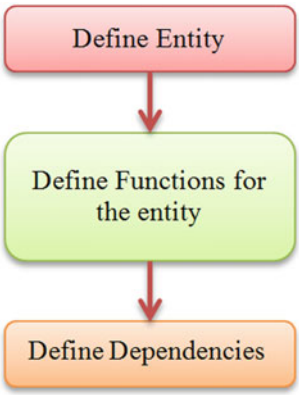


Fig. 5 The software flow in entity definition

defined with its dependencies and hooked at the pre-initialization stage in such a way that this module is initialized prior all the modules so that the entities running after the entity DSWI, will be captured by the single wire protocol implemented here (Fig. 5).

3.6.3 Intel Standard API for Boot Managing

This subcomponent in the system initializes and shuts down all the entities in the exact sequential order on the basis of the dependencies defined in Entity Declaration layer API. Every entity is defined by its boot_init() function and is the point where the entity is initialized. This exact hook gives way to transferring the entity to the 1-wire interface. Thus placing the entity at the correct hook point will provide with transferring the entity name to the desired pad and hence the 1-wire and finally to the slave.

3.6.4 Time Services

The Intel standard API for time delay module which is a module of time services subcomponent is hardware independent and implements a delay function that holds the processor from other tasks until the delay is over. In DSWI implementation, we use this function for a specific amount of time for the transmission of bits 1 and 0.

4 Protocol Flow and Constraints

The debugging procedure needs heavy analysis work to be carried out. On the basis of this analysis it was learnt that in order to get the track of the master Modem and get the sequence of execution of processes in the master, all the modules in the master Modem needs to be known. Every phone will have a sequence in which the

modules of the chipset are initialized. Knowledge of this will lead to framing of a buffer, which would contain all the entities that are initialized before the complete booting process. The intention would be to send an entity to the slave through the single wire interface, which would be one byte long. Hence, if at each step of initialization of a particular module or entity, the corresponding entity name or value is sent to the master pins, then these bits can be carried over by the SWI technology to the slave. Thus, if this is carried out, the slave will get interrupted when master interrupts at any specific execution of modules. Figure 6 shows all the dependencies of the DSWI implementation.

This methodology requires the understanding of how to configure the pads and the external interrupt handlers. Also, care needs to be taken to synchronize the master and the slave properly. In order to synchronize the two Modem boards, we need to set different timings for bit 1 and bit 0 transmission and reception. A major requirement of this work is that the driver for the initialization of the 1-wire should be initialized well before all the initializations happen.

This is because this interface is used to transfer the bits corresponding to the entity at the initialization instant itself. So if the single wire interface is not initialized, then it will never detect the bits at all. For this a detailed analysis is carried out on the dependency graph of the master Modem chipset. From the dependency graph, we can check out which modules are initialized in what exact sequence and with this we decide where to initialize the driver corresponding to SWI. Once the driver for SWI is initialized prior to the all the modules, the module just after the SWI driver initialization, can be made as a dependency to this driver. Thus all the modules will now fall below the SWI driver in the sequence chart.

Now that the initialization of the SWI is accomplished, the next step is to start the implementation of the entity transfer to the pins in the master chipset. For this we follow a specific procedure. Initially the SWI driver initialization starts with configuring the GPIO pad names for the SWI master. The constraint here would be to choose a function that would work at pre initialization phase of the phone. During pre-initialization phase, the external memory and C-runtime system is not

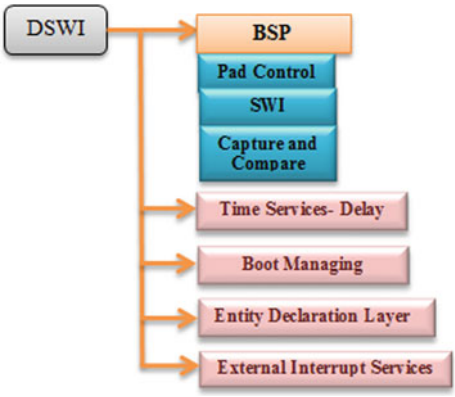


Fig. 6 Modules which DSWI depends on

set up and thus limited functionality needs to be provided to start the system. So pre-initialization functions, need to be used, to configure the system.

The next step is to write the entities to the 1-wire interface. For this, we have a function to write the entity name whenever it is initialized. This function should form the packet with the data length of 1 since the entity is always a single entity. While the packet is formed, it is again taken care that the writing process carried out to the SWI interface follows the pre-initialization function. Here too we use functions that implement the write procedure when the system is not yet ready. This writing is configured for the dedicated pad name for the SWI interface. Figure 7 shows this process in a flow chart. The generic SWI protocol transfers packets with the desired packet length, payload data, packet length and the checksum data. But this project has the same form of data to be sent on SWI and that is the initializing and executing entity names. Hence the payload of the packet is this entity name and the order in which these entities are initialized in the system. The length of the data is always a single entity and the checksum, generated is the sum of all data.

In a regular SWI protocol, callback functions are implemented that call the next user. Users are project specific and can be anyone who is in need of SWI. There is also a callback function to transfer the next byte to the slave. For each callback there is a timer associated, that triggers the callback function after the timer expires. This timer needs to be initialized while the SWI driver is initialized. In this project we do not need a callback function for the next packet transfer, since the user is always the same i.e., the initialization functions and there is always a single byte to be transferred.

Another constraint that this project presents is the use of timers. The timers of the system will not be initialized in the pre-initialization stage. So there should be some procedure by which the bitwise write process in implemented. This is done by using a delay function for writing the bits 1 and 0. There will be certain wait functions which will work in the early initialization stage. These functions needs to studied and analyzed and used to write the bits to SWI. Each time a bit is accepted, it is written to the pad. The pin is made HIGH for bit 1 and for bit 0 transmissions and each time a 1 or 0 is transmitted, a delay time is introduced for bit 1 and bit 0.

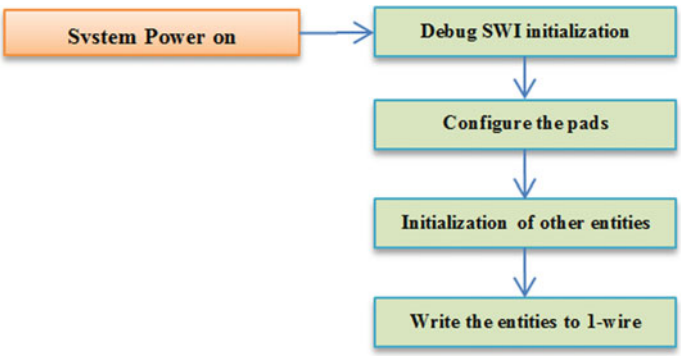


Fig. 7 Initialization process

These timing should synchronize with the timings used in the slave. After every transmitted bit, the procedure is repeated to accept the next bit and then again make the pad HIGH accordingly. At the end of bit transfer, the pad corresponding to SWI is made LOW, which corresponds to the end of transfer (Fig. 8).

Thus, with the above procedural implementation, we write the entity corresponding to the executed module to the SWI whenever it is initialized. And at the completion of the phone boot process, we get a buffer at the slave which will have all the initialized modules of the Modem in the order of execution. In case a master interrupts in between at some point of initialization, the slave also gets interrupted and we get to know at which point exactly the master has interrupted. This gives an efficient debugging environment using SWI protocol. The arrival of the data at the

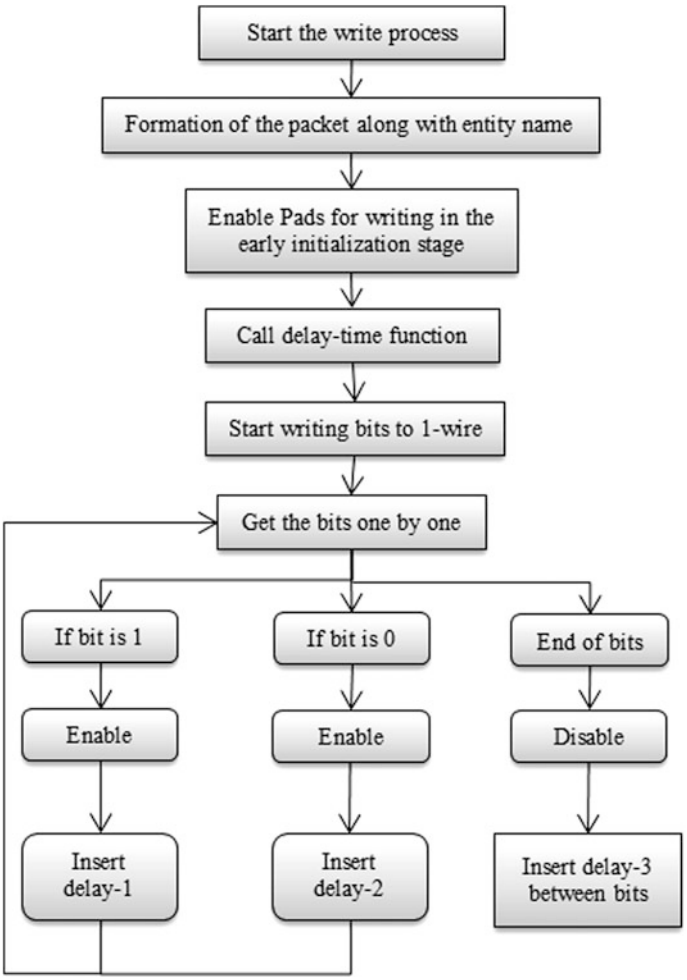


Fig. 8 Process flow in DSWI implementation

pad control API can be verified using an oscilloscope since during early initialization, debugger interfaces will not work.

The slave protocol is the generic protocol which accepts the data at the configured pads. This project uses the same slave for all the master Modems. Any new Master Modem should work with this protocol with the same slave. The pre-initialization stage is applicable only to the master and not to the slave. So slave can use the same pad control functions, external interrupt handlers and exception handlers. The slave side should be able to get the data from the dedicated pad for SWI configured for the same. The function get_hw_function() ensures this implementation. An external interrupt service routine allocate_lisr() function ensures that the interrupt at the slave pin is allocated to the dedicated SWI pad. Proper Care is to be taken to ensure that the timings and delays used in master Modem is the same as slave.

5 Result and Analysis

The protocol was implemented and verified for the successful transmission and reception of data packet which is formed in the protocol. The packet generated at the master side has a version type (0 × 2), a user (0 × 3), length of data (0 × 1 for one entity write), a payload (all the entity values written to the 1-wire where the last entity value written is 0 × D) and the checksum corresponding to higher and lower byte. This can be seen in the screenshot taken from the debugger trace32 instance as in Fig. 9. This packet when formed and transmitted to the slave side at the pre-initialization stage, will give a waveform according to the data transmitted which can be decoded using the widths of the waveform. The waveform captured using the oscilloscope on the slave side during the early boot stage of the master is shown in Fig. 10.

For the waveform captured from the oscilloscope, it can be seen that the data is encoded in the form of widths of the waveform. The delays of the waveform show the actual bit transferred. From Fig. 10a it was seen that smaller delay of 200 µs (corresponding to bit 1) and the larger delay of 1100 µs (corresponding to bit 0) gives the actual data encoded within the wave. This is according to the delay periods given while implementing the protocol. There is a delay of 500 µs between every bit transition as seen in the waveform which is in accordance with the delay-3 as in the flowchart of Fig. 8. This packet can also be viewed by the timing values received at the slave side as seen in Fig. 11. It can be seen that after a reset pulse of

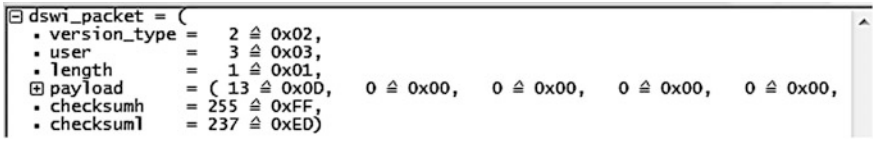


Fig. 9 Packet formed at the master modem chipset

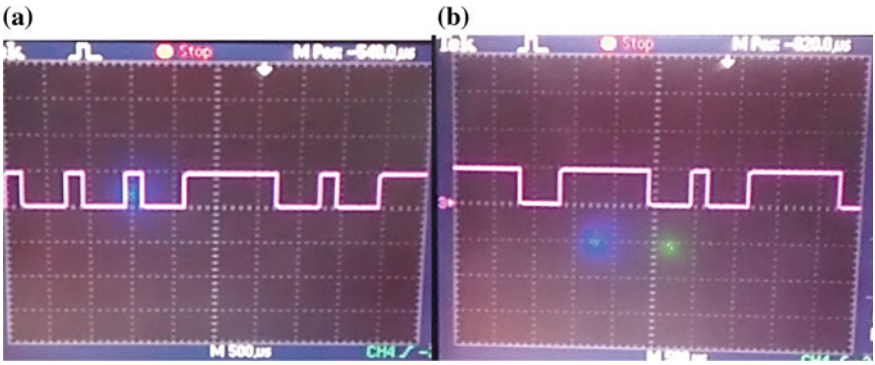


Fig. 10 Captured waveform at the master side during early initialization stage **a** for the entity with value 0 × D = 1101 **b** for the transition from one entity write to next write happening

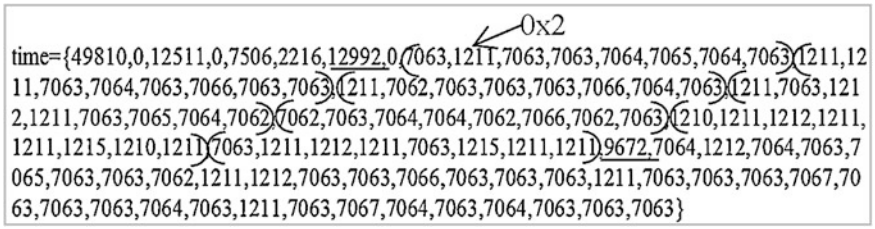


Fig. 11 Timings received at the slave side that correspond to the packet sent by the master

high time period, as shown as the underlined time period, the bits of the packet are transferred, which is as per the data packet sent by the master.

As implemented, the bits inside each bracket show the 8 bit equivalent for the version type 0 × 2 and similarly for the rest of the packet of the master. After the whole packet is transmitted, there is a long delay of 9672 µs which shows the stop of bits transfer. Thus the Pulse Width Modulation (PWM) process by the single wire successfully lets us decode the data packet. This implemented protocol can be used in real time scenarios and projects carried out in Intel Mobile Communications. The main application of this project is for chipset validations in Intel's chip bring up camps.

6 Conclusion and Future Scope

The establishment of 1-wire technology for debugging a master Modem in mobile platforms implemented in this paper provides single step debugging technique in early boot stage of mobile phones. The above process technology gives an efficient

way to implement this and get a proper track of any boot process of a newly introduced chipset. This technique also conveys that 1-wire technology can be used in applications where any networking equipment like Modem malfunctions and needs to be tested and validated for efficient performance. Intel brings up new chipsets in every release and during this stage a validation process is undergone. For this process, debugging of the chipset in the pre-initialization stage is a necessity. The complexities of using external debugger interface, like Trace32, during this stage are thus reduced by the 1-wire interface that helps in debugging the master bring up modem chipset. This protocol, in future, can be made bidirectional by which the slave device can be made to send commands to master to query or read or write the master side peripheral registers and memory in order to analyze the system's post fault.

References

1. Application note 214. Using a UART to implement a 1-Wire Bus Master, Maxim Integrated Products Inc. Sept 2002
2. Application note 148. Guidelines for Reliable Long Line 1-Wire Networks, Maxim Integrated Products Inc. Sept 2013
3. Downs, R.: Using 1-wire I/O for distributed system monitoring. Dallas Semiconductor Corporation, IEEE, 0-7803-5078-2/98/$10.00 0 (1998)
4. Zhuang, W., Fang, G., Yuan, L.: Binary tree for 1-wire technology in the ROM search. In: International Forum on Information Technology and Applications, 978-0-7695-3600-2/09 $25.00 © 2009 IEEE. doi:10.1109/IFITA.2009.21
5. Gui, Y.: 1-Wire search algorithm and its application. In: International Conference on Computer Distributed Control and Intelligent Environmental Monitoring, IEEE, 978-0-7695-4639-1/12 (2012)
6. Bairanzade, M.: Implement the single wire protocol. Application Note, Semiconductor Components Industries, LLC, AND8264/D (2006)
7. Trace-32. In-Circuit Debugger (ICD), Lauterbach Datentechnik GmbH, Germany 1989–1998, Aug 1998

Wavelet Transform Based Multi Rate Filter Bank for Human Auditory System

G.K. Girisha and S.L. Pinjare

Abstract Since Hearing loss vary with frequency, the gain and compression ratios in the hearing aid must also vary with frequency. This is implemented by filtering the signal into different frequency bands and applying separate gain and compression to the signal in each band. Multirate signal processing involves the change of the sampling rate while the signal is in digital domain. The main idea of using multirate filter banks is the ability of the system to divide the signal of consideration into two or more signals or to compose two or more signals into single signal in frequency domain. The work reviews on a design technique for analysis and synthesis filter banks. The analysis bank is to divide the incoming signal into different bands and the synthesis bank to perfectly reconstruct the amplified version of original signal.

Keywords Multirate filter banks · Simulink · System generator

1 Introduction

Multirate signal processing involves the change of the sampling rate while the signal is in digital domain. Multirate filter banks have found various applications in speech coding, signal processing, image processing as well as transmission of many signals through the same channel. The design technique involves designing of analysis and synthesis filters banks.

Figure 1 shows the overview of the filter bank. The audio signal amplitude is matched to the audiogram to provide correct signal to the hearing impaired person. The analysis filter bank is designed to decompose the incoming signal into different frequency sub bands and then the gain processing is done to adjust the hearing loss.

G.K. Girisha (✉)
Department of ECE, Nitte Meenakshi Institute of Technology, Bangalore 560064, India

S.L. Pinjare
Department of ECE, REVA University, Bangalore, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_49

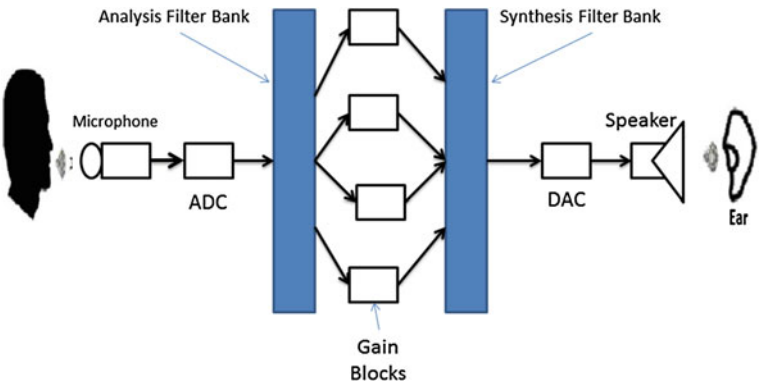


Fig. 1 Overview of filter bank

Next the frequency sub band components are recombined to reproduce selectively amplified version of incoming signal using synthesis filter bank.

2 Filter Bank

A filter bank is set of filters connected with sampling operators and delays. The down samplers are decimators and up sampler are interpolators.

Figure 2 presents a two channel filter bank, consisting of analysis and synthesis stage, which further consists of low pass and high pass filters. Low pass filter is indicated by $H_0(z)$ and high pass filter by $H_1(z)$. Here $G_0(z)$ and $G_1(z)$ are low pass and high pass filter in synthesis stage. We need to choose $H_0(z)$, $H_1(z)$, $G_0(z)$ and $G_1(z)$ such that to get perfect reconstruction. Since the filters are not ideal, the frequency response overlap, there is aliasing in each channel, even there is amplitude and phase distortion. We must choose all the filters in the filter banks in such a way as to cancel the errors, so we need to explore the conditions for perfect reconstruction, so we need to choose bi-orthogonal filter banks, such that synthesis filter banks should be inverse of analysis filter bank.

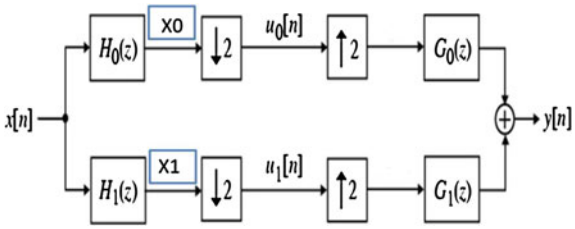


Fig. 2 The two channel reconstruction filter bank

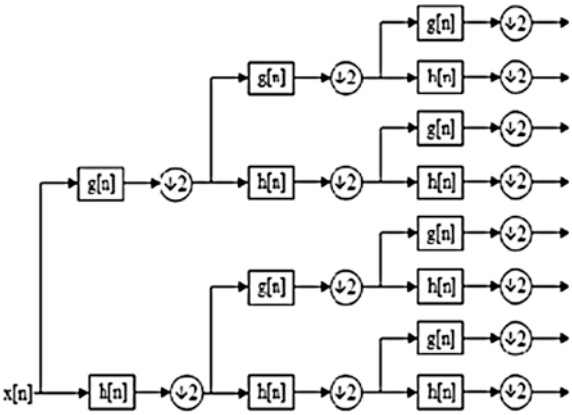


Fig. 3 Discrete wavelet packet transform filter bank architecture

The condition for perfect reconstruction is

$$G_0 * H_0 + H_1 * G_1 = 1$$

Aliasing occurs when frequency component of any signal goes higher than half the sampling rate also known as Nyquist limit. Since the combination of down sampler followed by up sampler, this zeros out the odd numbered components, it keeps only even powers of $H_0(z) * X_0(z)$.

Here the input signal X(z) is split by two filters. Filter $H_0(z)$ splits X(z) into low pass component, filter $H_1(z)$ splits X(z) into high pass component, both the outputs of the filters are decimated by 2, so to reconstruct the signal, the filter in the reconstruction bank should be designed in such a way that output Y(z) is identical to input X(z). This is the condition for perfect reconstruction. We need to choose decimator of factor 2, since it discards every odd components, so at the output of $H_0(z)$ and $H_1(z)$ it sets all odd samples of these signals to zero.

As seen in the Fig. 3 an eight channel analysis filter bank, h[n] and g[n] are low pass and high pass filters. The filter bank in wavelet packet transform is iterated at each level over all frequency bands. Thus the complete structure provides good flexibility for the signal representation to achieve better accuracy. The bandwidth of a filter decreases with increasing level of decomposition.

3 Filter Bank Implementation

The block diagram shown in the Fig. 4 gives the overview of the time domain signal processing using wavelet packet transform. The input audio signal to this system is of 24 kHz Bandwidth and its sampling frequency is 48 kHz. The input audio signal is in time domain and it is double data type.

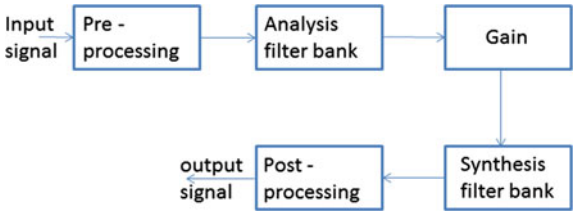


Fig. 4 Block diagram of time domain signal processing using wavelet packet transform

I. **Pre-processing**: The input signal which is of 24 kHz Bandwidth and a sampling frequency of 48 kHz is low pass filtered at a cut off frequency of 8000 kHz and then passed to DWPT filter bank.
II. **Wavelet Analysis**: This block consists of iterations of low pass filters and high pass filters. This consists of levels of Discrete Wavelet Packet decomposition stages. Thus using this architecture we have decomposed the input signal into 8 different frequency components with a uniform bandwidth of 1 kHz for each bands ranging from 1 to 8 kHz.
III. **Gain**: Using the audiogram of particular hearing loss, gain is applied to specific frequency band according to the audiogram. This is done by using shifter. Here left shifting is done to increase the gain.
IV. **Wavelet Synthesis**: After gain processing the decomposed signal of eight bands, signals are applied to synthesis wavelet section. In this structure quadrature mirror filter coefficients are applied to low pass filters and high pass filter, such that aliasing effect is cancelled out and thus satisfies the perfect reconstruction property. Thus at output of synthesis filter bank selective amplified version of incoming signal is obtained.
V. **Post processing**: The output from synthesis filter bank is low pass filtered at cut off frequency of 8000 kHz. This architecture uses efficient implementation of Finite Impulse Response (FIR) Filters using Distributed Arithmetic (DA) technique. FIR filters are used because they have linear response. The Distributed Arithmetic FIR consists of look up table (LUT), shift registers and scaling accumulator. DAFIR filter based architecture provides efficient area-time-power implementation, this involves less latency and area–delay complexity when compared with FIR filters.

4 Results

4.1 Implementation of 8-Channel Filter Bank

Structure of 8-channel filter bank using perfect reconstruction condition is shown in Fig. 5. The incoming audio signal is split into 8 sub bands, the analysis filter bank is

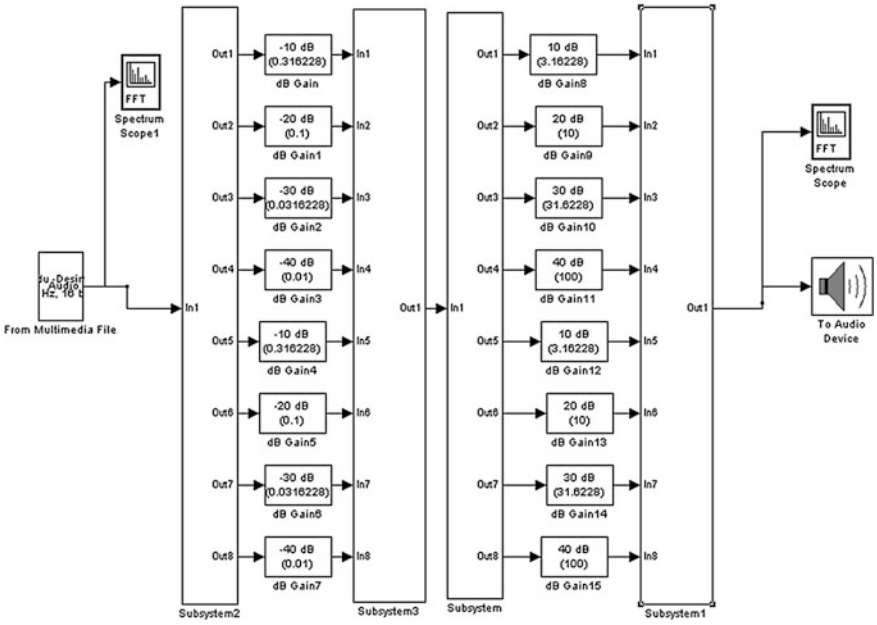


Fig. 5 Structure of 8-channel filter bank using perfect reconstruction condition

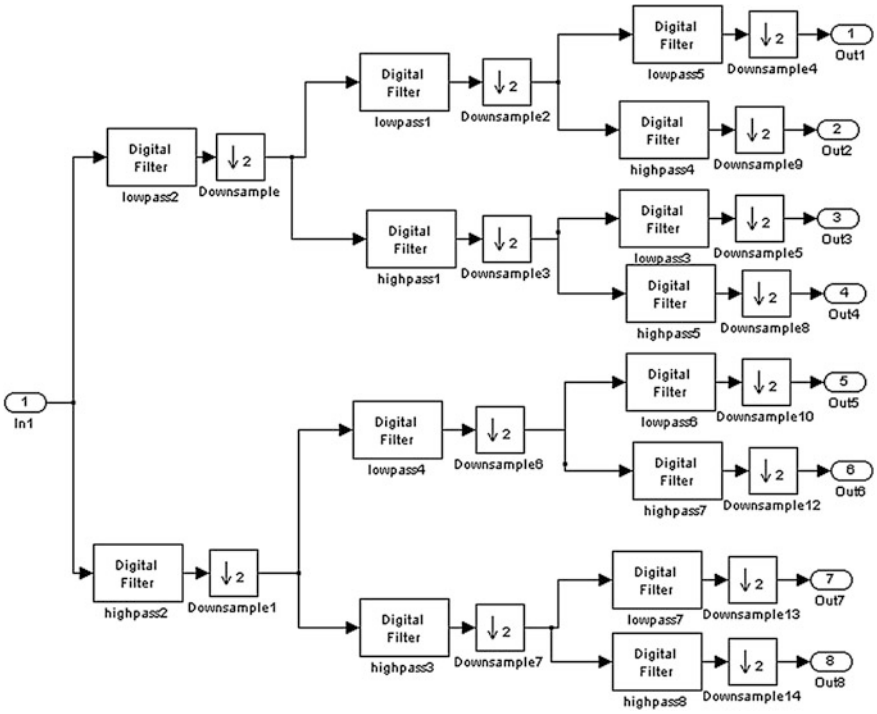


Fig. 6 Analysis block of the filter bank

represented in the Fig. 6 and the signal is synthesised back using the synthesis filter bank as shown in Fig. 7. The input audio file and the amplified version of same are shown in Figs. 8 and 9 respectively.

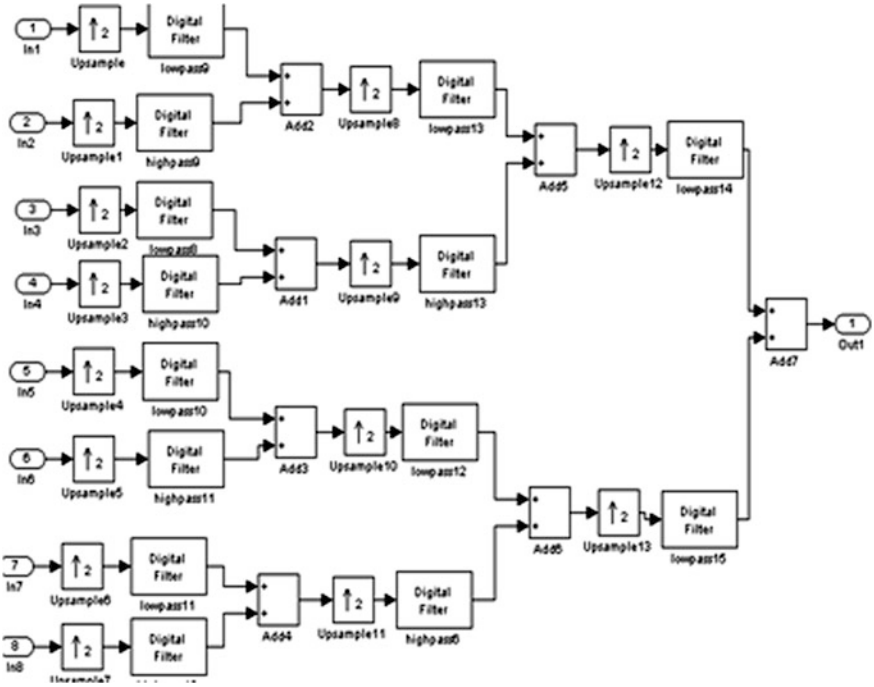


Fig. 7 Synthesis block of the filter bank

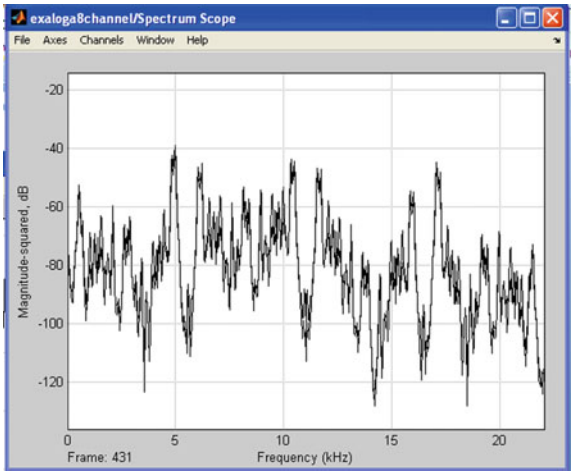


Fig. 8 Sample input wave to 8-channel filter bank

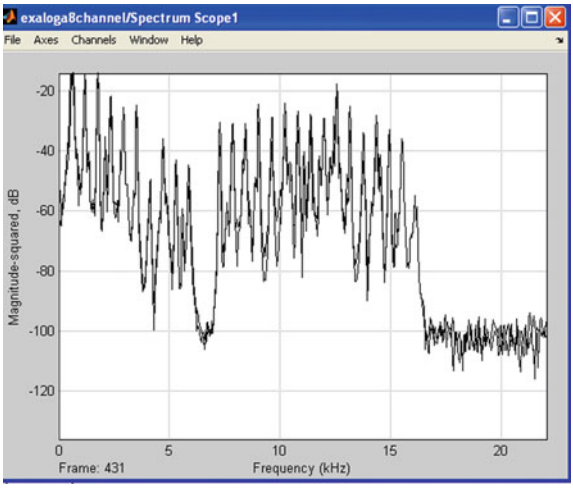


Fig. 9 Output wave from 8-channel filter bank

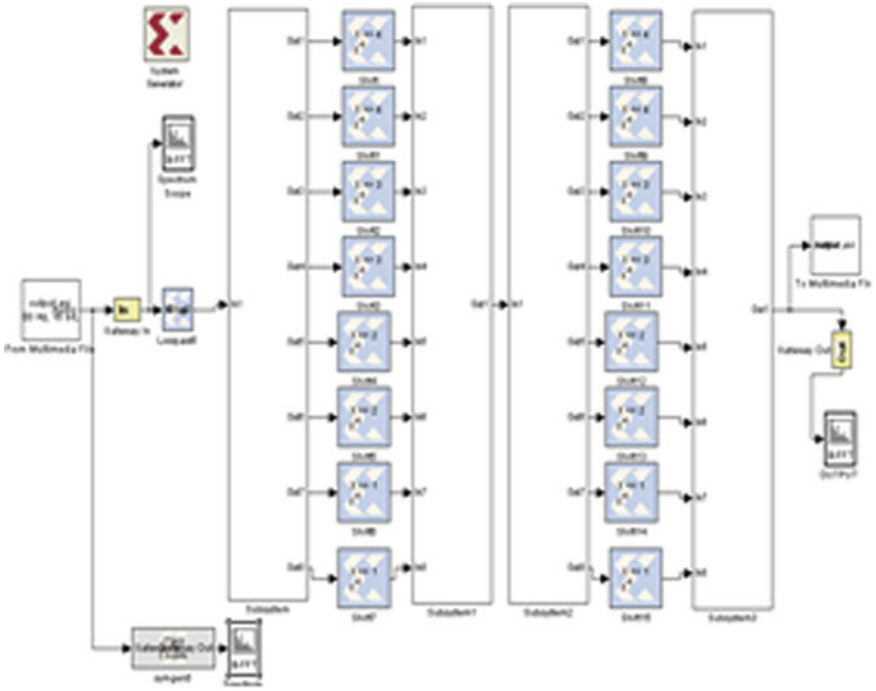


Fig. 10 8-Channel filter bank implemented on FPGA using system generator

4.2 Implementation of 8-Channel Filter Bank Using System Generator and Implementing It on FPGA

Similar filter bank is designed using system generator tool as shown in Fig. 10 and is implemented on FPGA and the results are verified.

5 Conclusion

The paper implements architecture for the Discrete Wavelet Packet Transform. For input audio signal of 24 kHz bandwidth, by using wavelet transform technique, amplification is done to match the audiogram of the hearing aid patient. The proposed architecture is simple, further level of decomposition can be achieved by using same identical filter banks.

Integrating Neural Linguistic Programming with Exergaming

Shyam Sajan and Kamal Bijlani

Abstract The widespread effects of digital media help people to explore the world more and get entertained with no effort. People became fond of these kind of sedentary life style. The increase in sedentary time and a decrease in physical activities has negative impacts on human health. Even though the addiction to video games has been exploited in exergames, to make people exercise and enjoy game challenges, the contribution is restricted only to physical wellness. This paper proposes creation and implementation of a game with the help of digital media in a virtual environment. The game is designed by collaborating ideas from neural linguistic programming and Stroop effect that can also be used to identify a person's mental state, to improve concentration and to eliminate various phobias. The multiplayer game is played in a virtual environment created with Kinect sensor, to make the game more motivating and interactive.

Keywords Exergaming · Neural linguistic programming · Stroop effect

1 Introduction

Current life style trends have seen wide spread relevance of digital media in almost all life aspects. The excessive use of computer and television has caused children to lead a sedentary life-style, thus created several social, mental and physical problems. People started to spend hours with digital media and this new lifestyle creates many health issues. One among the main problem that aroused is obesity.

S. Sajan (✉) · K. Bijlani
Amrita E-Learning Research Lab, Amrita VishwaVidhyapeetham,
Amrita School of Engineering, Kollam 690525, Kerala, India
e-mail: shyam.sajan@gmail.com

K. Bijlani
e-mail: kamal@amrita.edu

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_50

Obesity is identified as a serious health issue that younger population is facing and has recently reached epidemic proportions in many parts of the world, especially in the developed nations [1]. Based on the data from the 2003–2004 US National Health and Nutrition Examination Survey, an estimated 4.2 millions of children between the age of 6 and 11 are overweight [1]. It is also estimated that approximately 1 in every 4 children in the US might become obese by the year 2015 [1].

Regular exercises are the best solution for eliminating obesity or any other health issue caused because of sedentary lifestyle. Considering the addiction of younger population for digital games Exergaming [2] is introduced. This may explained as video games, also form of exercise, functions by capturing various body movements using accelerometers or by gesture recognition methods. The Exergaming has become an attractive research topic because of its impact on population. The earlier researches on Exergaming focus explicitly for improving physical health and no concern for mental health is given that has changed recently.

MeMaPads [3] a multimedia Exergaming and edutainment system that considers both the physical and mental aspects of it's users. MeMaPads is a padding system, consisting of a custom number of tiles, used to play the games, a memory and a math game. In the memory game, children are required to find matching pairs of pictures among a set of images from different category [3]. The system mainly implements math based games and adapts for children. The proposed system is designed for young population with help of Stroop effect and Neural Linguistic Programming. In1970s John Grinder and Richard Bandler (Canada) introduced NLP [4], a technique to model our performance, phobias, depression, habit disorder, learning disorder etc. But in recent days NLP is practiced with help of games called new code games. The new code games for NLP [5] aims at gaining unconscious attention which helps us to attain context free high performance state, enhances the cognitive development of the gamer.

Brain Age [6], is a series of video games developed and published by Nintendo, trains brain with math games like Sudoku. The game "Alphabet" is a new code NLP based game available in android market. These games are proven successful for improving mental health.

In the proposed system we collaborate a new code game based on Stroop effect [7] with Exergaming, to ensure both physical as well as mental wellbeing of the gamer. The system responds with the gamer on screen and feedback is given for improving performance. As the performance improves we could ensure improved physical and mental health.

2 System Design

The proposed system aims to design a game for both mental and physical health. So a new code game of NLP, based on Stroop effect [8] is implemented in virtual environment with feedback mechanism to improve the performance. Using Kinect

camera [9, 10] the movements of user is recorded and is matched with trained set of movements. The game is designed in a way that it ensures physical activity.

2.1 User Interface

The proposed system consists of Kinect camera, display screen and a processing unit. The users have to stand in front of the Kinect sensor [8] and it could identify the performed gestures. User have to coordinate his hand and leg activities according to the instructions displayed on the screen. The scores are marked based on the accuracy of user's performance.

2.2 Neural Linguistic Programming

According to NLP, behavior works purely based on past. All the present day thoughts are based on past experiences. Various fears or phobias, depression, laziness and so on got its origin from bad past experiences. So if we could change our past, then we could overcome various problems due to bad experiences in past, but changing past is an impossible thing. The only way to alter the effect of past in our brain is altering the visual cue.

Classic neural linguistic programming is practiced with techniques like swish [11] where user himself recalls the visual cue and replaces them with positive one. Thus all stories ends happy and bad thoughts will be removed. If the NLP is practiced regularly then the reason of fear or origin of every negative thought can be replaced. This changes the person's involuntary response to the situation.

2.3 Stroop Effect

In 1935 John Ridley Stroop published the paper on experimental psychology [12]. Stroop, in his experiment used three different kind of stimuli to study about interference in verbal reactions. Based on reaction time to respond the stimuli we could understand the speed of processing, attention and our state of performance as well. Two stimuli are given at a time so interference occurs. Training our brain with the game could help us to achieve context free performance state or 'flow'. The Fig. 1 shows the various stimuli that can be used in the game.

In recent years new code NLP is introduced which specializes the role of conscious attention in unconscious mind. Unconscious mind imagines future scenarios and include their consequences to our brain. If we could change unconscious mind in forming outcomes and choosing resources it will reflect in our real life. New code NLP also tries to change the person's state instead of replacing one behavior with

Fig. 1 Shows various stimulus that can be used in game

another. Which means the person achieve flow state or high performance state where one completely gets involved into the activity, all negative thoughts that effect performance gets vanished and the very same time positive energy fills. Context free high performance state means harnessing our own unconscious process to perform high level of conscious awareness.

2.4 The Game

Variations can also use:

1. Colors
   Here we display names of color in different colored alphabets
2. Emotional
   We display various emotions like happy, sad, grief, love etc. are displayed mixing with more neutral words like umbrella, bag, shoe etc. and the words are colored. It's found that depressed individuals take more time to read words with negative meaning.
3. Spatial
   Demonstrates interference between the stimulus locations with the location information in the stimuli. An arrow is displayed around a center point and gamer is made to discriminate the direction in which arrow points irrespective of it's displayed position around the center point. Based on instructions gamer have to coordinate his body movements.

Such as

1. Move left hand to right side and right leg to left side when red color is shown.
2. Move right hand to left side and left leg to right side when blue color is displayed.
3. Both hands up and on toes when green is displayed.

We could change the instruction such as color of the word, meaning of the word, could be made more complex.

The sensitive attention helps psychologists understand how the brain evaluates the information. Reading words was a lot easier task compared to reading color. When words meaning and color is asked to process simultaneously, two different regions of brain is activated. The right side finds the meaning of letters and the left

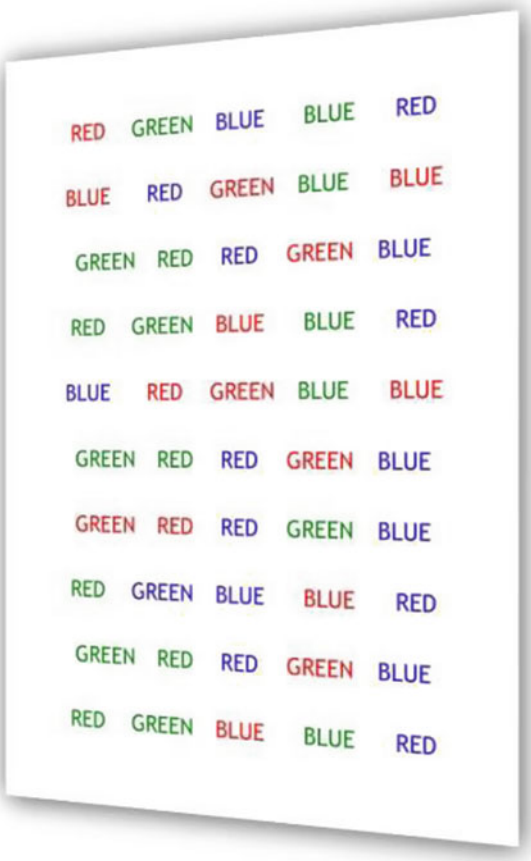


Fig. 2 Two different stimuli are given at a time, color of the word and meaning of the word

side of brain analyses the color. When we ask the person to analyze the color (active) rather than the word meaning (passive) the brain become conflicted and slows down your response time. Where the Anterior Cingulate [13] helps to solve the problem and make correct choice. The Fig. 2 shows the two different stimuli, color of the word and meaning of the word, which can given at a time.

Engaging both the hemispheres of brain at a time helps to attend the context free high performance state. This can be generalized to any activity producing unconscious behavior in any context where individual is challenged. Coordinating body movements help the individual to involve more and the high performance state will be achieved faster. Another aspect is the Exergaming or game play with exercise. Playing the game means coordinating hand and leg movements according to the displayed color instructions.

If the user response is identified then score will be increased. Since the colors or stimulus are displayed in varying time intervals. The time to response decreases as game proceeds based on improving score. The minimum response time is 1.5 s and maximum is 5 s. After reaching 1.5 s interval the game changes to next higher level, where a different stimulus with a different gesture will be used to play the game. This helps in exercising and so the obesity will be eliminated. Thus the game benefits both physically and mentally.

2.5 Kinect Sensor

Using Kinect we could get twenty body-joint positions, each joint with x, y and z positions. Player is made to stand in front of the Kinect sensor at a minimum distance of 1.8 m. With the help of Microsoft SDK 1.8 [14] twelve joint positions are captured for a single frame and the captured x and y coordinate positions are passed to dynamic time wrapping algorithm for finding the best matched gesture.

3 The Game Engine Module

The pattern of stimulus is displayed on view box and the player is asked to respond. The player's response is matched with the trained responses of system, where neural linguistic programming is used for forming the response. Various difficulty levels are set for the gamer. By changing the response time (time in which new stimulus is displayed) and by complicating the response.

Which means customization of the game is possible and displays the graphics, text and audio of game according to the instruction of game engine.

4 Implementation

The proposed system is implemented using Microsoft SDK 1.8, Visual Studio 2013 and Kinect Development SDK 1.8 with help of Kinect sensor. We used Microsoft Speech Engine's API [15] for the verbal spelling of texts and messages. Figure 3 shows the levels in which the system is implemented. Kinect sensor captures all the twenty join positions and the most similar coordinate values will be matched with the trained values using DTW algorithm [16]. If match is found then positive feedback is given to the user, increasing score or moving him to next higher level. Every time the user's gesture is matched with the respective instruction and also with the trained gesture. Where normalization part helps to identify the correct threshold to discriminate between various gestures.

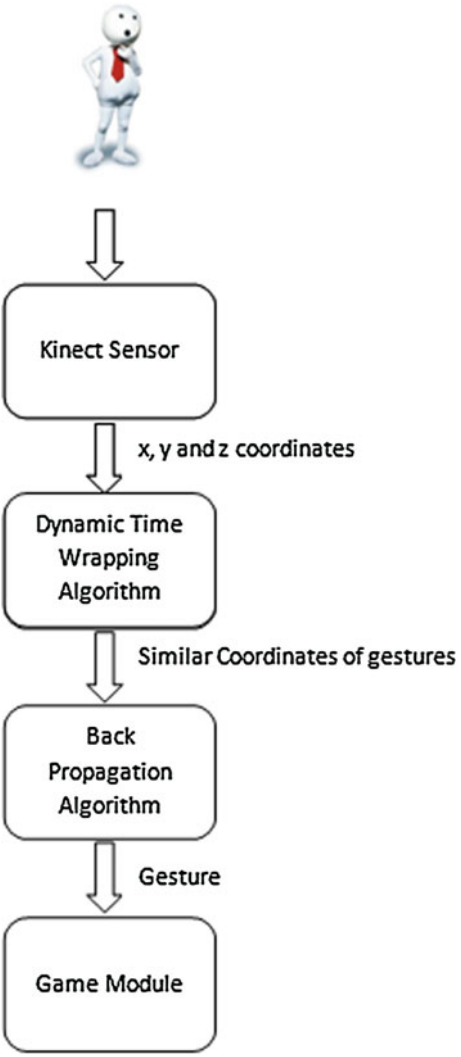


Fig. 3 Shows the block diagram

5 Results

The proposed system performance is evaluated with different players with different heights. The player's from age group of 18–30 and height between the ranges of 156–185 cm. Each player is made to play in a 1 min time slot. The Table 1 shows the efficiency of proposed system to recognize the gestures. As analyzed from the Table 1 data, four players with different physical appearance are asked to perform the gestures on a slot having one second time interval (time for which a stimulus

Table 1 Gesture identification

		Person 1 (156 cm height)	Person 2 (179 cm height)	Person 3(185 cm height)
Right hand rise	Correct	50	52	51
	Incorrect	10	8	9
Left hand rise	Correct	47	45	46
	Incorrect	13	15	14
Both hand rise	Correct	55	53	54
	Incorrect	5	7	6

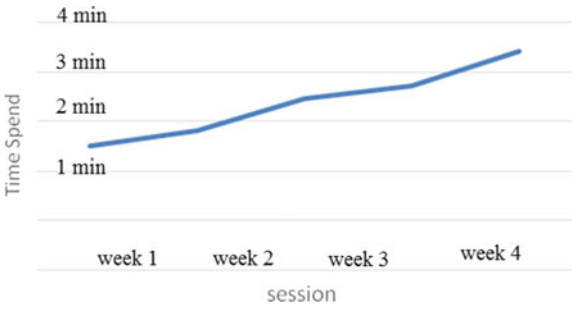


Fig. 4 Time spend by a player on a week's play

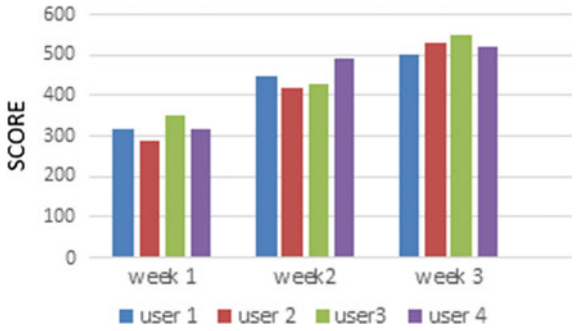


Fig. 5 Shows the average score achieved for four players in three weeks

will be displayed and player is asked to respond). The Table 1 shows the number of correctly identified gestures.

The Fig. 4. is plotted for the time a player spends in different sessions. We can see steady increase on time spend to play the game after each session. This shows how the gamer is motivated by the game.

Four players with different age group and height is asked to play the game continuously on three weeks for 3 min time. The performance is plotted in a Fig. 5. As you can see, in first week all players are scored not well. But in the final week they reach the highest score. From which we can inferred that continuous practice improves the ability to understand instructions and coordinate body movements.

Table 2 Error rate for various time intervals

Number of players	Interval (s)	Error rate
12	1	0.23
8	1.2	0.1
4	1.5	0.04
6	2	0.003

To test the performance of the proposed system the game needs to played in a controlled environment. Here different time intervals are used to measure the error rate. A response time between 1 and 2 s is chosen initially (Table 2). It is observed that in certain cases score is improved even for untrained gestures. Which resulted in a score higher than the maximum allowed score (calculated based on number of gestures in a time slot). For each of the time slot the actual score and error score is measured (similar to Table 1). The difference between the scores is divided by the maximum allowed score of that time slot and error rate is calculated (refer the Eq. (1)). The result shows minimum error rate is guaranteed in between 1.5 and 2 s.

$$Error\ rate = \frac{Maximum\ score - Error\ score}{Total\ score} \tag{1}$$

After analyzing the results we could see that the error rate of the proposed system is below 0.2. Also the response time is a main distraction factor that can be used to complicate the game. So as level improves response time is decreased and the game is made complicated.

6 Conclusion and Future Works

The system implements an Exergame that helps to improve both mental and physical health. Practicing NLP code game in virtual environment helps the user to involve physically for the cognitive development. This is ensured by coding the instructions with neural linguistic programming. The virtual environment for playing the game is created using a Kinect sensor and gesture recognition algorithm which identifies the performed gestures with a minimum error rate of 0.2. The system is tested for players with different physical appearance and for different response time. The results showed that response time itself is a distraction factor and the can be used to complicate the game. Response time between 1.5 and 2 s is used for the game play. Where separate levels are formed based on the response time for a particular stimulus. The increase in player's score after each session shows the improved concentration and physical coordination of player. In future more interesting games can be implemented using the same virtual system. Kinect camera sensor helps to create a 3-dimensional avatar of the user and using the infrared camera, depth can also be sensed. The avatar helps the user to involve

deeply with the game. In future collaborate racing and fighting games can be brought into this environment where the gestures or actions can be coded according to neural linguistic programming.

References

1. Kumanyika, S.K., Obarzanek, E., Stettler, N. et al.: Population based prevention of obesity: the need for comprehensive promotion of healthful eating, physical activity, and energy balance: a scientific statement from american heart association Council on epidemiology and prevention, interdisciplinary committee for prevention (formerly the expert panel on population and prevention science. Journal of the American Heart Association (2008)
2. Brox, E., Hernández, J.E.G.: Exergames for elderly: social exergames to persuade seniors to increase physical activity. In: 5th International Conference on Pervasive Computing Technologies for Healthcare (2011)
3. Karime, A., Khaldi, B.H.A., Aljaam, J.M., Saddik, A.El.: MeMaPads: Enhancing Children's Well-being Through a Physically Interactive Memory and Math Games, IEEE (2012)
4. NLP. Internet: http://www.neurolinguisticprogramming.com/
5. Code games for NLP. Internet: http://www.daryllscott.com/new-code-colors-game/
6. Brain age games, internet: http://brainage.nintendo.com/
7. Stroop Effect. Internet: http://en.wikipedia.org/wiki/stroop_effect, (J Ridley Stroop. Studies of Interference in Series of Verbal Reactions, Journal of Experimental Psychology, Accessed http://psychclassics.yorku.ca/Stroop/)
8. Kinect Sensor. Internet: http://www.microsoft.com/en-us/kinectforwindows/ and http://en.wikipedia.org/wiki/Kinect
9. Introduction to kinect for windows. Internet: http://www.microsoft.com/enus/kinectforwindows
10. Patsadu, O., Nukoolkit, C., Watanapa, B.: Human gesture recognition using kinect camera, IEEE (2012)
11. Swish. Internet: http://www.personal-development-planet.com/nlp-swishpattern.html and http://www.nlp-now.co.uk/nlp_swish.html
12. Experimental Psychology. Internet: http://en.wikipedia.org/wiki/Experimentl_psychology
13. Anterior Cingulate Cortex. Internet: http://en.wikipedia.org/wiki/Anterior_cingulate_cortex
14. MicrosoftSDK. Internet: http://www.microsoft.com/enin/download/details.aspx?id=10121
15. Mirza Climkovic. Neural Networks and Back Propagation Algorithm
16. Priyank, S., Motion learning using the neural network (2012)

Standardization of UAN Simulator Channel Framework

M. Monika and Shanta Rangaswamy

Abstract With the Advent of technology and man's interest in exploring underwater environment made him to develop new aquatic applications. Underwater Acoustic Networks are used to support these new applications. UAN's are made up of sensor nodes, vehicles and other devices, for communication between them many network protocols are designed, after developing any new protocol that has to be validated in the ocean environment which is difficult and much costlier, so the next option is to rely on simulators. Simulator results may not be fully accurate. There are many simulators for wireless communication using radio waves but these can't be used for underwater acoustic communication as they differ greatly. In this paper acoustic jprowler simulator is implemented using jprowler as the base simulator, at acoustic layer an improved channel model is used which considers few ocean dynamics and environmental parameters. Right communication can be achieved by appropriate Noise model, attenuation model at the acoustic layer and a mac protocol is designed using results from the physical layer. Simulation results analyzed and compared against the results from field experiment in Chesapeake Bay in USA.

Keywords Underwater acoustic network · Jprowler · Acoustic layer model · Cooperative mac

1 Introduction

Underwater acoustic networks (UAN) are Ad hoc wireless networks made up geographically distributed sensor nodes in underwater environment communicates through acoustic channels, gateway nodes are placed at the water surface and acts as intermediate between the sensor nodes and operators on sea shore. These networks facilitate to wide varieties of applications in ocean sensing and detection, checking oil spills, cost line protection, protection of marine life, submarine detection,

M. Monika (✉) · S. Rangaswamy
Department of CSE, R.V. College of Engineering, Bengaluru 560059, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_51

tsunami monitoring [1]. Acoustic communication is considered as the best communication because radio waves in water suffer from high attenuation, optical signals highly scattering. Underwater networks are different from ground-networks due to the following reasons; the speed of EM waves in underwater is five times lower than terrestrial networks. Transmission delay also depends upon mobility of sensor nodes. Acoustic channels are also affected by noises produced by shipping activities, aquatic animals, motors, power plants and turbulence noise, multipath effects results in the degradation of acoustic signals, Doppler spread results in continuous cylindrical spreading of frequency in shallow water, transmission is also caused due to the attenuation of signals, primary causes of attenuation are due to viscous friction and ionic relaxation. These issues made researchers to develop new protocols for underwater network their deployment in sea for testing is costlier and system testing is impractical before full system is implemented. Hence researchers went for simulators, there are very few open source simulation tools available for research community, but they very generic channel model which leads to less accurate results this motivated us to develop realistic channel models which gives near accurate results against the sea experiment results. In this paper easily available open source tool Jprowler is considered and physical layer model is implemented within it for UWSN, a cooperative mac protocol using the results from Acoustic layer model.

2 Channel Framework Details

Jprowler simulator is taken as the base tool, it is a discrete-event simulator where protocols developed for wireless ad hoc networks are analyzed [2]. It is implemented in java, open source and extensible as protocols implemented in its library. Its limitation is it only radio models and one mac protocol. Jprowler as the base simulator, the acoustic layer model for the UWSN and a SNR based mac at the mac layer is implemented, system model for acoustic communication in Jprowler is as shown in Fig. 1.

2.1 Acoustic Layer Model

In underwater physical layer model, the standard attenuation model is used [3] that is given by,

$$A(l,f) = l^k * (10^{\frac{\alpha(f)}{10}})^l \quad (1)$$

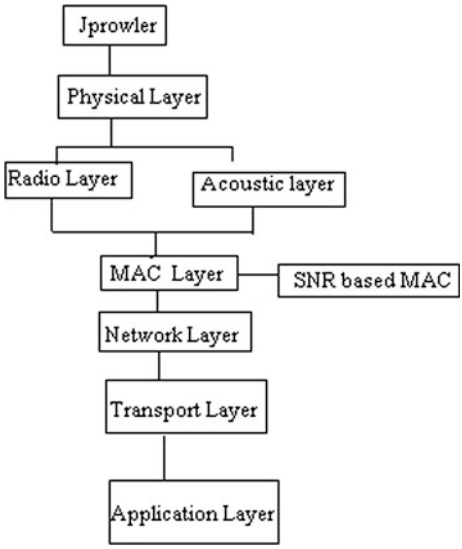


Fig. 1 System model

where, l is the distance between sender and receiver, f is the frequency, k is the practical spreading factor (k = 1.5) [4], f is the frequency in hz.

α(f) is given by Throp's equation,

$$\alpha(f) = \frac{0.11 * f^2}{1 + f^2} + \frac{44 * f^2}{4100 + f^2} + 2.75 * 10^{-4} * f^2 + 0.003 \tag{2}$$

In shallow water to calculate more realistic transmission loss the general equation given in Roger's model [5] is used.

$$\mathrm{TL} = 15 \log_{10} \mathrm{R} + \mathrm{AR} + \mathrm{B} + \mathrm{CR}^2 \tag{3}$$

where R is the transmission range, since sound energy is not perfectly contained in reflection (reflection coefficient < 1) and refraction, the particle spreading coefficient is taken as 15 $\log_{10}$ R, A, B, C are the co efficient different values of A, B, C are given in [5]. Path loss can be calculated by,

$$\mathrm{PL} = \mathrm{l}^{\mathrm{K}} \tag{4}$$

$$\text{Total Attenuation is calculated by, } \mathrm{Attenuation} = \mathrm{TL} + \mathrm{PL} \tag{5}$$

The Acoustic channel is affected by ambient noise which is more in shallow water, so in this paper existing noise model is considered [6]. Four different equations [7] of noises are turbulence, wind, shipping and noise from thermal fluctuations [1, 3],

$$\begin{aligned}
\text{Turbulence noise NSL1(f)} &= 17 - 30\log(\text{f}) \\
\text{Noise due to Shipping activity NSL2(f)} &= 40 + 20(\text{S} - 0.5) + 26\log(\text{f}) \\
&\quad - 60\log(\text{f} + 0.3) \\
\text{Wind noise NSL3(f)} &= 50 + 7.5\text{sqrt(w)} \\
&\quad + 20\log(\text{f}) - 40\log(\text{f} + 0.4) \\
\text{Thermal noise NSL4(f)} &= -15 + 20\log(\text{f})
\end{aligned} \tag{6}$$

Total noise is given by,

$$\text{Noise} = \text{NSL1} + \text{NSL2} + \text{NSL3} + \text{NSL4} \tag{7}$$

Power received when frequency f, power transmitted above distance l can be calculated by,

$$\text{P(r)} = \text{l} * 10.0/\text{R}. \tag{8}$$

where R is the communication range of protocols.

Then SNR can be calculated by,

$$\text{P(r)}/\text{attenuation} * \text{Noise} \tag{9}$$

2.2 SNR Based Co-operative Mac Protocol

The main objective of cooperative mac layer communication is to get better data transmission rate and to enhance the channel reliability. This paper works on the second objective that is to enhance the channel reliability. Cooperation transmission is enabled when data packet sent by the sender is not received by the receiver. Most commonly used protocols for cooperation transmission are cooperative diversity mac [8], and Differentiated cooperative mac. In CD-MAC if the direct transmission is failed, then the sender node and the helper node retransmits the packets simultaneously to the receiving node. But it suffers from error burst and limited signal transmission distance. So differentiated cooperative mac is used, it adopts cooperative ARQ mechanism, if the signal arriving at recipient node is stronger than all other signal that is with high SNR value then node can receive the signal correctly, when signal strength received (SNRi) falls below certain threshold (SNRth) then helper nodes are selected to enable cooperativity. Helper nodes are chosen in such a way that common nodes which are within the communication range of source and destination are chosen to transmit the signal further. So correct helper nodes are chosen to retransmit only in case of failure thus improving the cooperative efficiency. Figure 2 shows the computed SNR in the path from node 34 to gateway or sink node 0.

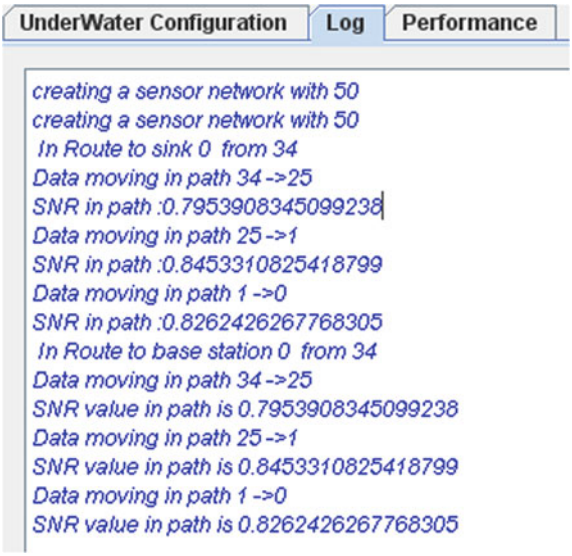


Fig. 2 Log showing SNR calculation in the path from node 34 to sink node 0

Algorithm Next Node with Geographic Distance

1. Calculate euclidian distance of each node i to the target node record distance di is distance of node i to target node
2. For all n nodes repeat the following steps
3. Initialize mindis to infinity.
4. If(di<mindis)
5. Mindis=di
6. Min node=i
7. Return i as nextnode

Algorithm SNR Based Mac

1. While (nextnode!=-1)
2. Calculate SNR using equation 9
3. If(SNRi<SNRth)
4. Enable cooperativity
5. Else
6. Follow direct path

3 Frame Work Evaluation

The accuracy of physical layer within the context of mac protocol is evaluated by choosing the commonly used performance metric to evaluate mac protocol that is packet delivery ratio, throughput compared against the field experiment results from the test in 2011 [1].

Tables 1 and 2 gives details of the simulator environment and parameters used in the simulation.

3.1 Bay Experiment Summary

A field test was conducted in chesapeake bay Maryland USA. In the experiment 8 nodes were considered and they were deployed 1 km apart. Sediment type is clay-slit in the deployment area. Shipping activity was dense, wind speed 8–10 m/s. Average depth of the bay is 14 m [1].

3.2 Simulation Results

This simulation is an average of 19 runs. Figure 3 shows the increase in the packet delivery ratio when SNR based cooperative mac used.

Figure 4 shows Improved simulation results and Field experiment results which is taken from [1].

Slight variations are seen in the simulation results this may be due to the environment settings used remains same throughout the simulation but that is not the case in actual sea environment. The parameters like shipping activity, wind speeds are not constant in sea environment.

Table 1 Jprowler simulator

Base Simulator Used	Jprowler
Implementation language/IDE used	Java/Netbeans IDE
Performance chart	JFreechart
Acoustic layer model	Loss due to absorption is calculated using Roger's model
	Noise model used considers ambient noise
Mac layer protocol	SNR based cooperative mac
Mobility of the nodes	Random–way point

Table 2 Simulation parameters used

Parameters	Value
Wind speed	8 m/s
Shipping activity	Dense = 1
Sediment type chosen	Clayey–slit
Absorption coefficient	A = 0.418, B = 52.57, C = 5.18 for clayey-slit
Frequency	200 Hz

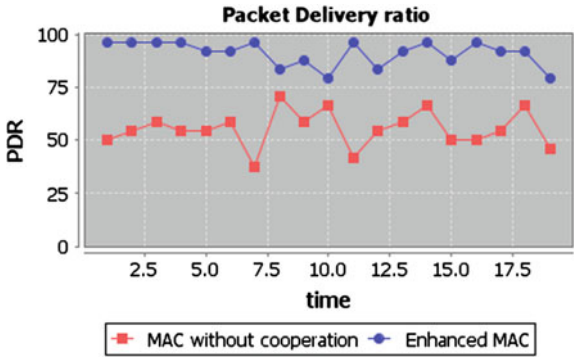


Fig. 3 Packet delivery ratio versus time

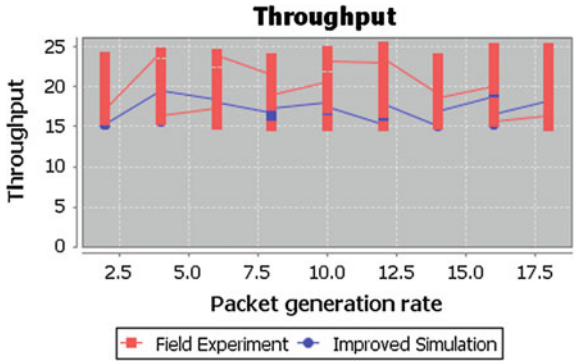


Fig. 4 Field experiment versus simulation results

Figure 5 shows the comparison of original simulation and the different cases by considering the sediment types, wind speeds. In case 1 considered clayey slit sediment type is considered, dense shipping activity and wind speed 25 m/s. In case 2 clay slit, dense shipping activity, wind speed 10 m/s, in case 3 fine sand, no shipping activity, wind speed of 10 m/s. Inference drawn is the channel model does not affect the protocol dynamics.

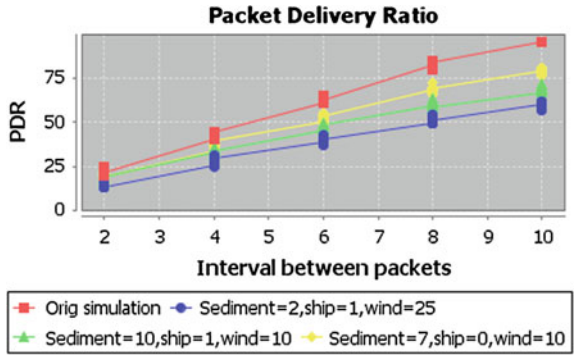


Fig. 5 PDR versus interval between the packets

4 Conclusion and Future Work

In this paper UAN Simulator an extension to existing open source simulator jprowler is presented, extension includes the improved physical layer model with attenuation and noise models and the novel cooperative mac scheme based on SNR is considered as cross layer design of mac and physical layer UAN. The results obtained is found to be encouraging, so this simulator can be used to achieve near accurate results with less resources and in turn less cost.

In this simulator currently flooding is used at the network layer, in future focus is to implement an efficient network layer model for Underwater Acoustic Network to increase the accuracy and efficiency, and also this work can be expanded to allow users to set different environmental parameters in the same simulation run.

References

1. Zuba, M., Jiang, Z., Yang, T.C., Su, Y., Cui, J.-H.: An advanced channel frame work for improved underwater acoustic network simulators. In: MTS (2013)
2. Tzimas, A.: Asynchronous consensus protocols: preliminary results, simulations and open questions. In: Proceedings of the fourth IEEE conference on decision and control (2005)
3. Domingo, M.C.: Overview of channel models for under water wireless communication networks. Phys. Commun. Sept 2008
4. Zorzi, F.: On the effects of node density and duty cycle on energy efficiency in under water networks, Oceans 10 IEEE Sydney
5. Rogers, P.H.: On board prediction of propagation loss in shallow water. In: NRL Report No. 8500, Sept 1981
6. Coates, R.: Underwater Acoustic Systems. Wiley, New York (1989)
7. Mitra, U.: A delay-reliability analysis for multihop underwater acoustic communication. In: Proceedings of the Second Workshop on Underwater Networks-WuWNet 07 (2007)
8. Li, Y., Liu, K., Liu, F., Xu, Z.: A rapid cooperation-differentiated medium access control protocol with piggyback for multi hop wireless network. In: EURASIP Journal on Wireless Communication and Networking vol. 264 (2013)

Analysis of AODV Routing Protocol and Enhancement for Energy Efficiency

N. Apoorva, Lalitha R. Rakshitha, Namrata Purvimath and Spoorthi Pujari

Abstract Mobile Ad Hoc Networks (MANETs) are self-configuring collection of mobile nodes that are designed to function in an infrastructure-less environment. Since the lifetime of mobile nodes is limited by their battery capacity, energy efficiency of routing protocols is essential for sustenance of the network. Two algorithms namely, Energy Mean Value algorithm and Energy Efficient Maximum Lifetime Ad Hoc Routing (EEMLAR) algorithm have been proposed to improve the energy efficiency of Ad hoc On demand Distance Vector (AODV). The conventional AODV protocol has been simulated on QualNet for two different scenarios and the results have been compared.

Keywords AODV · MANETS · EEMLAR · Energy mean value · QualNet

1 Introduction

Mobile ad hoc networks are independent self'-sustaining system of mobile hosts which operate even in the absence of a centralized administration. This property of MANETS plays a vital role in critical applications like military communication, disaster management and search and rescue operations which are characterised by limited coverage area. It also finds ubiquitous applications in commercial fields like sensor networks, Personal Area Networks (PAN) and navigation.

MANETs are plagued by many challenges such as dynamic topology, limited bandwidth and security threats. The constant mobility of the nodes makes it difficult to maintain routing information in the network. Since the data from the sender to receiver is transmitted in a multi hop manner in a wireless media, it is prone to snooping and poses security concerns. But the most serious drawback of MANETs is the limited lifetime of the mobile nodes due to their battery constraint. The network heavily relies on the intermediate nodes to act as router and forward the

N. Apoorva (✉) · L.R. Rakshitha · N. Purvimath · S. Pujari
Department of ISE, RV College of Engineering, Bangalore 560059, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_52

data packets from the source to the destination. Hence the failure of one or more nodes can be detrimental to the working of the network. A number of metrics like number of nodes, packet size, energy of nodes and routing protocol are responsible for the energy efficiency of the network.

This paper focuses on determining the most suitable routing protocol to address the issue of energy constraint in MANETs. Routing protocols for MANETs can be broadly classified as proactive protocols and reactive protocols. The latter maintains a routing table and periodically updates the route information which is an energy consuming process. Reactive protocols on the other hand discover a route only when a message needs to be sent from the source to the destination thereby minimising energy usage. AODV is a reactive routing protocol that works based on broadcasting RREQ (Route REQuest) from source node to its neighbouring nodes, which in turn forward the RREQ to its neighbouring nodes until the destination node is found and optimum route from source to destination is discovered.

In this paper, Energy Mean Value (EMV) algorithm and Energy Efficient Maximum Lifetime Ad Hoc Routing (EEMLAR) algorithm are compared, which enhances the basic AODV protocol to improve the energy efficiency. The effect of varying the metrics mentioned above on the two algorithms is studied to graphically represent the differences between them.

2 Literature Survey

In this section, a general overview of AODV and its working principles are described after which the enhancements to basic routing protocol to improve its energy efficiency are explained.

2.1 AODV

In the paper by Ravneet Kaur [1], the author describes AODV as a reactive protocol that performs route discovery only when a message needs to be sent from the source to the destination. There are two processes involved in AODV.

1. Route discovery: The route discovery process is performed by first checking the route table for an existing route from the source to destination. It is described in the paper by Jin-Man Kim and Jong-Wook Jang [2]. If it does not exist, it involves the source node broadcasting a RREQ to all its neighbouring nodes. This RREQ is further forwarded by the neighbouring nodes to its neighbouring nodes until either the RREQ reaches the destination or till it reaches a node which has a route to the destination. At each hop of the RREQ a Reverse Path towards the source node is set up so that when the RREQ reaches the destination the RREP can be transmitted through the Reverse Path. Also, the Forward Path for transmitting the data packet is constructed using the Reverse Path.

2. Route maintenance [1]: The battery life of the nodes is increased by deactivating the routes when not in use. The RERR (Route Error) packet is used for this process. The record of the predecessor nodes are maintained in the route table. In case a route is broken, the node sends RERR packet to all its predecessor node which eliminates the broken links. The routes must be rediscovered before using them again.

2.2 Enhanced AODV

To improve the energy efficiency of AODV protocol, the parameters that effect the battery life of a node need to be considered. The node in a network can operate in one of the following modes:

1. Transmitting mode: The node is transmitting a packet.
2. Receiving mode: The node is receiving a packet.
3. Idle mode: In this mode the node neither sends nor receives packets. It is in a listening state where it waits for any packet that might be sent to it.
4. Sleep mode: This mode uses minimum battery. In this mode the node can neither transmit nor receive. It is in a dormant state and needs an internal or external trigger to activate it and bring it to one of the above three modes.

3 Algorithms

3.1 Energy Efficient Routing Protocols in Mobile Ad Hoc Network Based on AODV Protocol

The EELMAR algorithm to improve the energy efficiency in AODV is described by Shilpi Jain and Sourabh Jain [3]. This algorithm is constructed by considering different criteria like keeping track of the residual battery power, previously used paths, back-up paths and message overhead. The overhead involved in updating of routing tables is also considered. It takes advantage of the idle and sleep mode of nodes to save the battery life.

A node involved in transmission losses some amount of data which depends on the nature of the packet, the packet size and distance between the nodes. In the basic AODV algorithm the optimum route is selected based on the shortest route from the source to destination. In this case however some routes maybe overused and the battery capacity of the nodes in this route maybe exhausted. Thus, the remaining battery in a node should be considered as a metric to decide the best route.

The optimizing function is used to decide the best route where the energy cost of a route is given by

$$\text{Cost} = \text{time} + \frac{1}{\text{Remaining energy of node in route}} + \frac{1}{\text{no. of hops}}$$

where Time is the duration for the RREQ to reach the destination from the destination, Remaining energy is the sum of remaining energies of all nodes in the considered route, Number of hops is the total number of nodes in the route considered.

The algorithm works in the following way:

1. The routing table is searched to find a route from the source to destination. If the route exists in the route table and if the route has not expire, the particular route is selected.
2. If there is no route from the source to destination in the route table, RREQ is transmitted from the source to its neighbouring nodes. If a valid RREQ is received there are two possibilities
   a. If the neighbouring node receiving the RREQ is the destination node there are two possibilities
      i. The destination node has already sent a RREP to the source. In this case the RREQ from the source is considered as a duplicate and it is rejected.
      ii. If the RREQ is sent for the first time from the source, the destination node sets up a timer t time interval.
   b. The corresponding entry is made in the route table. The destination node waits for t time interval for another valid REQ from the source. After time out the node is moved to idle or sleep mode.
3. If the neighbouring node is not the destination node, the route table is updated and step 2 is repeated
4. When the route is discovered, the data is transmitted from the source to destination. If the current path is not active, the data is transmitted through alternate back up paths.
5. After transmission of the data, the source either sends another RREQ or is deactivated.

3.2 Energy Mean Value Algorithm

The Energy Mean Value Algorithm to improve the energy efficiency in AODV is described by Akhilesh Tripathi and Rakesh Kumar [4]. The paper deals with the enhancement of network lifetime and battery life using Energy Mean Value Algorithm. The algorithm decides the optimum route based on the maximum remaining energy of each node.

Step 1. Choose the sender node in a given network topology.
Step 2. Send an RREQ to all the neighboring nodes.
Step 3. The remaining energy of each node is considered in selecting the nodes for communication in the network. Among the nodes, the one with the maximum remaining energy is selected.
Step 4. If two or more nodes have the same maximum energy, each node sends an RREQ to its neighboring nodes.
Step 5. Among these neighboring nodes, the one with the maximum remaining energy is selected.
Step 6. Repeat steps 2, 3 until the requests reach the destination.

The parameters that are affected by the Energy Mean Value algorithm are:

1. Energy consumed: The total energy consumed by the network is equal to the energy consumed in transmission of the packet and energy consumed in receiving the packet.
2. Packet delivery ratio: It is defined as the ratio between the number of packets received at the destination to the number of packets sent by the source.
3. Communication overhead: It is given by the ratio of the number of control packets to the number of data packets.

The results show that the battery life and throughput is increased when compared to the traditional AODV, because it considers the maximum remaining energy present in each node though the path between the nodes is long. In traditional AODV, the algorithm chooses the shortest path but if a node is deactivated due to the exhaustion of energy, the packet will not reach the destination. Thus, the performance is depleted.

4 Simulation Model

Qualnet simulator has been used to simulate behaviour of AODV protocol. In the simulation model, the two scenarios considered consist of 20 nodes and 40 nodes that are arbitrarily scattered in an area of 1500 m × 1500 m square region with 1 and 2 links respectively. The traffic model used is CBR (Constant Bit Rate) with packet size of 512 bytes and simulation time of 1000 s. To include the dynamic topology of nodes in MANETs, random waypoint is used as the mobility model. The parameters considered in the simulation is shown in the following Table 1.

Figure 1 shows that 20 nodes are placed in random connected to three subnets. Two constant bit rate links are established between four nodes. In the first link, node 4 is the source and node 18 is the destination. In the second link, node 14 is the source and node 34 is the destination.

Figure 2 shows that 40 nodes are placed randomly connected to three subnets. Two constant bit rate links are established between four nodes. In the first link,

Table 1 Simulation parameter

Parameters	Value
Area size	1500 m × 1500 m
Simulation time	1000 s
Wireless propagation model	Two ray
Node placement	Random
Traffic type	CBR
Mobility model	Random waypoint
MAC protocol	MAC 802.11
Network protocol	IPv4
Routing protocol	AODV
Items to send	100
Item size	512 bytes

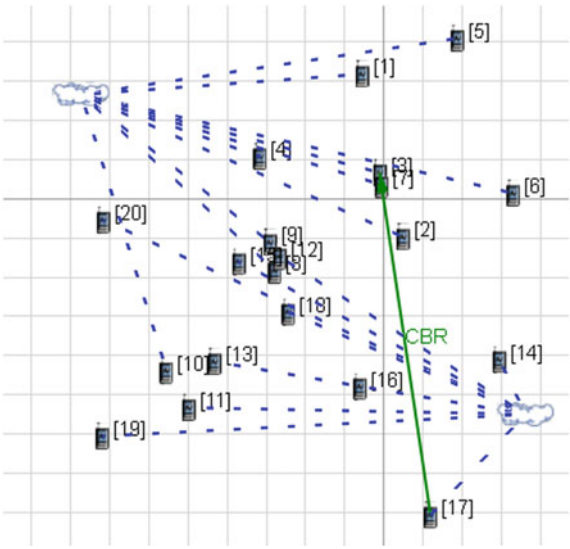


Fig. 1 Simulation with 20 nodes

node 4 is the source and node 18 is the destination. In the second link, node 14 is the source and node 34 is the destination.

5 Results

The simulation results are as shown in Table 2. The effect of variation of number of nodes on the average jitter, average delay and Constant Bit Rate is analysed. As the number of nodes in the network increases, the average jitter is also found to increase. Similarly, the average delay of the network also increases. On the other

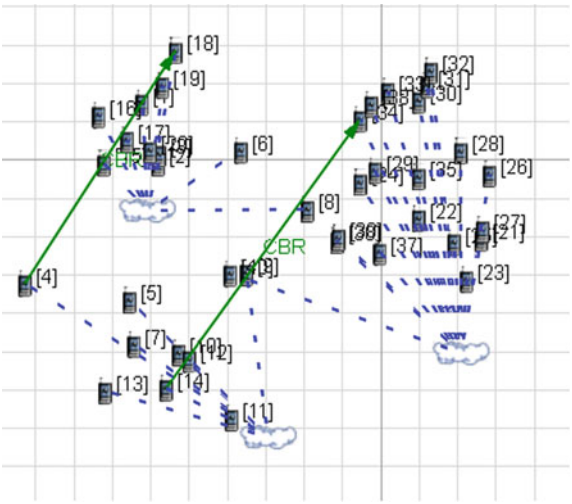


Fig. 2 Simulation with 40 nodes

Table 2 Simaltion Result Summary

Parameter	20 nodes	40 nodes
Maximum average jitter (in seconds)	0.0024	0.0028
Maximum average delay (in seconds)	0.0027	0.0025
Number of unicast messages sent	24	24, 24 (link 1, link 2)
Number of unicast messages received	9	12, 19 (link 1, link 2)
Number of messages dropped	15	12, 5 (link 1, link 2)

hand, the number of packets dropped decreases with the increase in number of nodes in the network.

Figure 3 shows the average jitter for 20 nodes. It can be inferred that the maximum average jitter is 0.0024 for node 3.

Figure 4 shows that the maximum average jitter in a 40 nodes scenario is 0.0028 for node 40.

Figure 5 shows the average delay for 20 nodes. It can be inferred that the maximum average delay is 0.0027 for node 3.

Figure 6 shows that the maximum average delay in a 40 nodes scenario is 0.0025 for node 40.

Figure 7 shows that total unicast messages sent in a 20 nodes scenario is 24 for node 17.

Figure 8 shows that total unicast messages received in a 20 nodes scenario is 9 for node 3.

Figure 9 shows that total unicast messages sent in a 40 nodes scenario is 22 for node 4 and 24 for node 24.

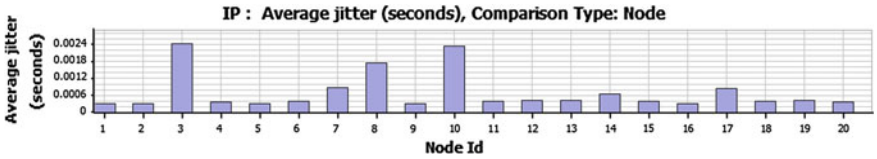


Fig. 3 Average jitter for 20 node scenario

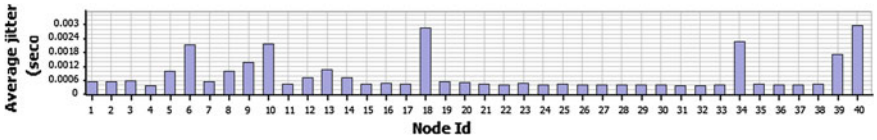


Fig. 4 Average jitter for 40 nodes scenario

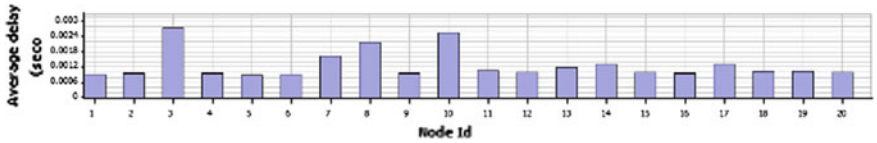


Fig. 5 Average delay for 20 nodes scenario

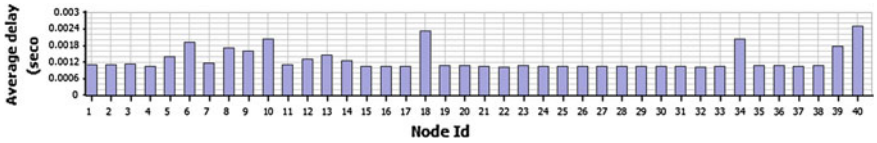


Fig. 6 Average delay for 40 nodes scenario

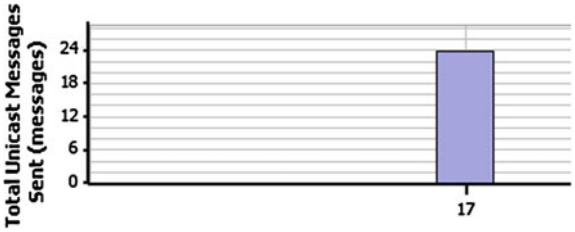


Fig. 7 Total unicast messages sent for 20 nodes

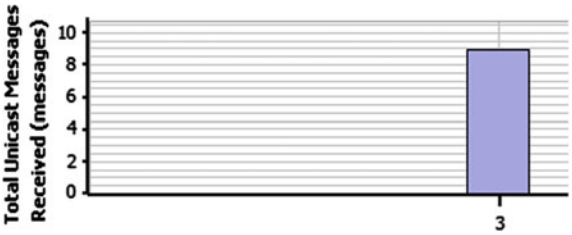


Fig. 8 Total unicast message received for 20 nodes

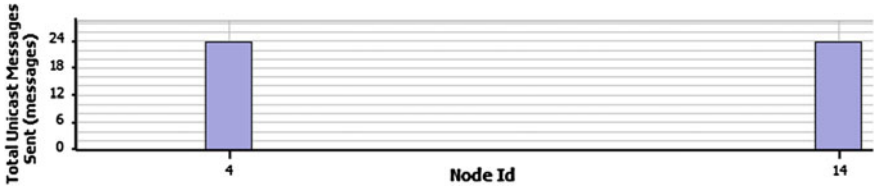


Fig. 9 Total unicast messages sent for 40 nodes

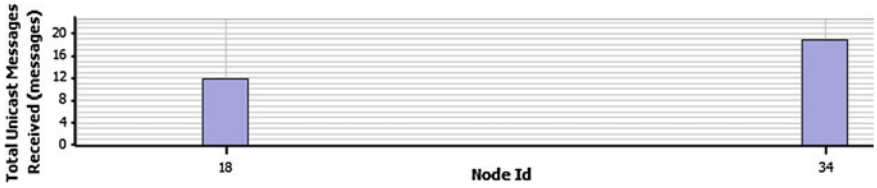


Fig. 10 Total unicast messages received for 40 nodes

Figure 10 shows that total unicast messages received in a 40 nodes scenario is 10 for node 18 and 18 for node 24.

6 Conclusion and Future Work

As seen, the performance of AODV protocol when applied to a scenario with more number of nodes deteriorates. Also, the main drawback with MANETs is the limited lifetime of nodes due to the battery constraints. Hence, to improve the energy efficiency of the nodes in the network thereby improving the network lifetime, algorithms like EELMAR and EMV should be applied to the AODV protocol. Future work should investigate the effect of these algorithms on the energy efficiency of AODV protocol. The performance of these algorithms must be compared and contrasted with that of the conventional AODV protocol. The merits and demerits must be evaluated.

References

1. Kaur, R.: Energy efficient routing protocols in mobile ad hoc network based on AODV protocol. Int. J. Appl. Innov. Eng. Manag. (IJAIEM) **2**(1), 323–324 (2013)
2. Kim, J.-M., Jang, J.-W.: AODV based energy efficient routing protocol for maximum lifetime in MANET. In: Proceedings of International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ICIW) (2006)
3. Jain, S., Deptt, S.J.: Energy efficient maximum lifetime ad-hoc routing (EEMLAR). Int. J. Comput. Netw. Wirel. Commun. (IJCNWC) **2**(4), 451–453 (2012)
4. Tripathi, A., Kumar, R.: MECB-AODV: a modified energy constrained based protocol for mobile ad hoc networks. Int. J. Distrib. Parallel Syst. (IJDPS) **3**(6), 22 (2012)

The Cost-Benefit Analysis of Vulnerability of Web Server Through Investigation

Seema Verma and Tanya Singh

Abstract With the rise of web applications, there has been a paradigm shift as web servers are no longer a technical requirement for any organization, but it is seen as an extension for business propositions. These web servers are designed with the basic function of hosting websites and applications such as gaming, data storage and enterprise applications. This paper attempts to investigate the various vulnerabilities of the web server and the path followed by the attacker to attack. The extensive survey has been done on the various methods of attacks which are an outcome of the loopholes in the design and architecture, deployment of the web server. If the components are designed and modeled taking into considerations the various vulnerabilities explained in the paper, the chances of attacks on the web server is minimized drastically. The paper also discusses the novel formula for the calculation of severity of attacks. This formula helps to assess the attack and is instrumental in the calculation of cost-benefit analysis.

Keywords Vulnerabilities of webserver · Types of attacks · Attack analysis · Severity of attacks

1 Vulnerabilities of a Webserver

The web servers host various applications which are coded wither in PHP or simple HTML. These applications relay all important information regarding the organization for better proliferation in the market. However, due to poor coding design and its deployment, these web servers are vulnerable. The weak areas can be due to poor implementation of network infrastructure design, weak implementation at the time of implementation, operational flaws, or due to poor code design. The hackers

S. Verma (✉)
Department of Electronics, Aim & Act, Banasthali University, Rajasthan, India

T. Singh
Department of Information Technology, Amity University, Uttar Pradesh, Noida, India

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9_53

who attempt to invade inside the network through weak web servers can be of any age group and their basic aim is to deface the website and gain access to a valuable data from the system. Concurrently, they may try to misuse it. The loss can be very high either in terms of monetary loss or losing goodwill and reputation. In case of financial institutions, it may be the loss of revenue and in case of corporate and government systems; it may be the loss of goodwill [1, 2].

As stated by Computer Emergency Response Team, the several security threats to a public web server can be further categorized as the following [3]:

(1) Unauthorized access

   (i) Defacement
   (ii) Content Theft
   (iii) Data Manipulation

(2) Improper Usage

   (i) Launch Pad for external attacks
   (ii) Hosting improper/malicious contents (e.g. Phishing)

(3) Denial of Service
(4) Physical threats

Unauthorized Access is a form of attack where the hackers take the advantage of security flaws such as lack of operational and monitoring control, loopholes in the operating system and application level program, weak password based on dear ones name or birthdays, social engineering attacks. Using these flaws, the hacker can deface the website. Once, the access to the system is gained, the most valued data is either stolen or manipulated.

Improper Usage is a form of attack in which the vulnerable website is completely compromised and can be used for launching attacks on another website. It can also be used to host malevolent contents for phishing attacks to take place. Denial of Service (DoS) is also termed as "availability attack". This attack attempts to limit a system's accessibility and usability. By sending the "ping" command, the user can flood the system with several requests. These requests consume the processor or memory resource. The end result can be the operating systems and applications that may crash when they receive specific strings of improperly formatted data. Generally, it has been seen that the hacker uses IP spoofing to conceal his identity when launching a DoS attack. Physical Threats are a form of attack on the hosted system resources. If these res sources are not protected, it can lead to physical damage, theft of the device or loss of data due to environmental hazards such as fire, humidity etc. [4].

2 Survey of Attacks on the Web Server

The different forms of attacks are not only limited to simple port scans, denial of service attacks (DDoS), spreading of worms, viruses and malware, spamming, phishing, but also exploitation of known vulnerabilities of unpatched web applications to gain sensitive information from the victims. The other lesser forms of attacks are [5]:

• URL Misinterpretation	• Source code Disclosure
• Directory browsing	• Session hijacking
• Reverse proxying	• Buffer overflows
• SQL query poisoning	• Cross—site scripting

2.1 URL Misinterpretation

The Uniform Resource Locator is basically a four field string that comprises the protocol (HTTP, FTP), Domain namespace, Port number (optional), path [6, 7]. Following are the URL misinterpretation attacks in the form of path traversal which has been tabulated in detail in Table 1.

2.2 Directory Browsing

While visiting the web pages, sometimes there is a list of folders available in the explorer to facilitate easier navigation of the web server. These listings if not attended properly are open ports and they lead to security loopholes. The Table 2 explains the various commands through which the directory can be traversed.

2.3 Reverse Proxying

In order to reduce the traffic load on the web server, Reverse Proxy is placed near the server. It retrieves the information from more than one server and returns it to the client. The client feels that the information has originated from the proxy server itself. Hence, it behaves as an intermediate for its associated servers. Also termed as "*Proxy Server*", it provides caching, acceleration, content-filtering, and isolated functions. The web proxy that the browser uses for outbound traffic is set and is controlled by the ISP or home network. Reverse Proxy passes the request from the internet. It also invades a firewall and reaches any private networks which are kept in isolation. Hence, Reverse Proxy allows the internet clients to have direct,

Table 1 Commands for URL misinterpretation

Types of URL misinterpretation	Different ways for URL interpretation	Remarks
URL manipulation	http://abc.com/forum/?cat=2	If the information is deliberately entered by the unauthentic user and is another numeric data or any ASCII which is not a numeric, the website may go into an undefined state and responds in the form of some error message
	http://abc.com/forum/?cat=6	
	http://abc.com/forum/?cat=**********	
Trial and error	http://abc.com/admin/	Redirects the path for administration directory
	http://abc.com/admin.cgi	
	http://abc.com/phpinfo.php3	Redirects and searches the script about information about the Remote system
	http://abc.com/.bak	The backup copy can be easily retrieved by the server
	http://abc.com/.bash_history http://abc.com/.htaccess	Command to find hidden files in the remote system
Directory traversal	http://abc.com/abc/test/ascii.php3	The tree structure path can be travelled back and back till it reaches the root. New commands as shown can be issued which may give all the information stored on the server
	http://abc.com/abc/test/	
	http://abc.com/abc/	
	http://abc.com/../../../../directory/file	
Advanced attacks	http://abc.com/..%2F..%2F..%2Fdirectory/file	The new form of Uniform Resource Locater codes to attack a directory with a Unicode notation
URL manual modification	http://abc.com/cgi-bin/script.cgi?url=index.htm	Verifications are very important. An attacker may modify the URL manually

*Courtesy: http://en.kioskea.net/contents/31-url-manipulation-attacks

unmonitored access to the sensitive servers. The vulnerability of this kind leads to "*man-in the middle attack*" against the network as long as the web—application does not require users to use client-certificates for authentication. Many times, the attacker forces its victim to connect its server to a malicious web server instead of connecting directly to the real server. For example, an "*Open—Relay*" is a misconfigured server that sits between a proxy server and a reverse proxy server [8, 9].

2.4 SQL Query Poisoning

Online business and organizations are protected by both software and hardware firewall solution. The purpose of the firewall is to filter network traffic that passes into and outside the organization's network limiting the use of the network to be

Table 2 Commands for directory traversal

Directory traversal	
Command to move about the server	[www.abc.com//user/public_html/] cd ..
	[www.abc.com/export/home/user/]
If we are browsing the URL, we can traverse deeper URL and it will show "secret output" return if this attack is not protected	www.abc.com/index.html –> maps to /user/public_html/index.html
	www.abc.com/../secrets/showSecrets.php –> maps to /user/secrets/showSecrets.php
An inbuilt application which uses data input to ask the application to retrieve the desired files	www.abc.com/getFile?fileName.txt
PHP code to retrieve the reply	<?**php**
	$filename = $_GET["getFile"]; //get the filename from the URL
	$fh = fopen("/home/user/public_html/files/".$filename, "R"); //open file for reading
	Echo fread("/$fh,filesize($filename)")// read and display contents
	?>
Modify the URL to request a secret file	www.abc.com/getFile?../../secrets/showsecrets.php
Displays the secret file contents to the attacker	www.abc.com/user/public_html/files/../../secrets/showSecrets.php–> maps to /user/secrets/showsecrets.php

*Courtesy: https://www.golemtechnologies.com/articles/directory-traversal

permitted as legitimate users. The SQL code attack takes the advantages of unexpected input vulnerabilities to execute SQL commands with malicious data or data to be executed dynamically through the remote source. The malicious SQL commands can therefore be inserted through a remote user in the server application. The person who has malicious intention executes SQL queries on the database and either try to spoof it or modify the contents [10]. Hackers can easily perform SQL code poisoning. They run queries and known scripts on the web browser to find important tables and field names. Since, it is easier to perform; it becomes the most vulnerable form of web server attack [11].

2.5 Source Code Disclosure

In this form of attack, the users with malicious intention try to access the source code of the server-side application. Once the source code is known, the hacker has the footprint of the complete logic of the server-side of the application and the entire web server becomes vulnerable. The hacker can request for the files which can be an image, macros or ASP, JSP, PHP scripted files. These requested

Table 3 Command for source code disclosure

Source code disclosure	
The session can be kept alive by using this code	http://www.abc.com/example-%61%83%70
ASP source code may be retrieved using this code	http://www.abc.com/showcode.asp?source=/msadc/sample/../../../../../inetpub/wwwroot/default.asp

automated files are code executed on the server leading to the source-code disclosure attack. In this type of attack, the source code of server-side of attacks can be easily retrieved by the attacker. The advantage gained is to change the codes and unleash the logic behind the web application. Hence, he can backdoor into the system very easily. The Table 3 exhibits the command for source code disclosure.

2.6 Session Hijacking

The attacker captures the authenticated session ID either through brute force or the reverse engineering. Port scans, Ping sweep also helps to steal the session ID. Since the HTTP protocol is stateless, hence the challenge faced by the designers is to maintain the record of states visited when more than one connection are attempted from the same user [12]. For every request generated, it is not possible to authenticate it. Hence, to reduce the overhead, the session is stored in *the cookie*. The session id also stores other requests originating from the same user on the web server. The applications automatically generate these IDs to store relevant parameters [13, 14].

Understanding Cookies: The cookies are set to expire upon closing the browser and based on the amount of time needed to expire and the location where they are stored, they are categorized into two types, *Session cookies or Non-Persistent Cookies* that expires as the session is closed and the persistant cookies that is saved on the hard drive in the location [15]. The cookie can be displayed by using the following code in Javascript:

$$\boldsymbol{alert}\ (\boldsymbol{document}.\ \boldsymbol{Cookie})$$

Precaution: The sessions when established should be encrypted using SSL to avoid being predictable. The sessions can be hijacked either by brute force attack, cookie stealing or injecting queries also known as cross-site scripting.

2.7 Buffer Overflow Attack

This form of overflow attack occurs when a hacker purposely calls another malicious program as an exploit through a stack overflow. Every program or process

needs a stack or heap to preprocess the pointer to the program. Hence, the overflow can be in both a stack and a heap. If the exploitation of the user input is caused due to overflow of the stack or heap, such an attack is called either stack based overflow attack or heap based overflow attack respectively [16].

Precaution: The stack has the fixed size. If the programmer takes care to specify the amount of space to store the transaction data from the stack, the buffer overflow attack may be avoided. *Care is taken that the input codes inserted during the application should not be lengthier than the specified space reserved for the execution of the program.* Secondly, the method of stack canary can be implemented to avoid stack overflow attacks. Any random value is chosen and is stored on the top of the stack return pointer. When a malicious program tries to call its address in place of the return address, this canary value is overwritten. Hence, whenever there is a process running, the canary value must be checked to avoid attacks [17]. Heap based buffer flow attacks, floods the memory space allocated for a program. Any vulnerable program uses a command called "strcpy" to copy an input string into the heap buffer. Any input longer than the buffer tries to copy its content on the heap. The hacker now tries to copy executable scripts on the heap and attempts to compromise the server [18, 19]. "EXEC ("sh")" is the command in Linux and after the execution of this command; the command prompt window opens and makes the system very vulnerable.

How the attackers gain the full control of the operating system? Every operating system controls the access level of the user as well as executable programs. Any such programs therefore run in kernel mode. When a stack-overflow attack takes place, the hacker pads the sides of the program with NOP instructions. The hacker generally takes the advantage of NOP instructions to hide its malicious script. Hence, the malicious program runs at the kernel level. However, the running application program believes that the real application program is still going on [16].

2.8 Cross-Site Scripting

The features within the Java Script applications are currently supported by browsers such as HTML tags, scripts, hyperlinks. Cross-site scripting attacks points out to an increase in vulnerabilities and threat which allows malicious exploiting of Java script codes. In this form of vulnerability, an attacker inserts malicious HTML, generally dynamically executed Java script code into a web page of a web application [20, 21].

Cross-site scripting attacks work by the mechanism in which the attacker transfers the information from the trusted URL to a fictitious site and victim has no idea of that. The trick used by an attacker is to hide the fictitious site's URL into the URL bar that appears very genuine to the user [22, 23].

The above paragraphs have shown in depth the various attacks that can take place on the web server. It has deeply analyzed the various loopholes that can take place while code-designing the web server and the network where it has been

Table 4 Proposed algorithm for detection of attacks for honeypot based intrusion detection system

Step I: Analyze existing symptoms on the deployed web server through Wireshark
Step II: Determine the IP address from which the packet was sent and its ownership
Step III: Is the problem within the Trusted Zone?
• If yes: Isolate the trusted ZoneIf no: deep scan the packets for more signatures using ACC algorithm (Pattern Matching)
Step IV: Block the IP address.

deployed. The following sections revisit the attacks and its effect on the web server. Based on the severity of attacks, the paper has tried to do the cost-benefit analysis of the system.

3 Cost-Benefit Analysis of Attacks (Based on the Experimental Results)

To determine the severity of attacks, it is required to *gather the symptoms* through the log files using active and passive monitoring. Every activity does not happen for once. There has to *be a sequence of failed attempts* which is very important to monitor and can be checked through the regular monitoring of the log files. The methodology used to evaluate the attack on the web server includes monitoring the network through active network monitoring tool, called as wireshark *v* 1.10.6. The detailed analysis of the wireshark window reveals the IP address from which the malicious data has been send. Further, detailed analysis of the IP address reveals if the data was sent from the trusted zone or from outside the network.

Still, there are chances that the malicious data leaks into the network. Hence, every packet is matched for signatures using pattern matching algorithm using Aho-Corasic–Cache Deterministic Finite Automata (ACC) Algorithm. The Table 4 exhibits the step which can be taken by the Honeypot based Intrusion Detection and Response System to monitor and analyze the attacks and classify the IP address from which the attack has taken place. Honeypot based IDS has the characteristics to retain the newly found signatures in its database. It adds to the old signature based traffic. The attacker keeps on looking for the vulnerabilities and tries all the various commands. It tries to confuse the TCP by sending ping commands at a very short interval such that the TCP shows half-open connection, look for open ports, looks for the vulnerabilities in the application level programs.

4 Formulae for the Severity of Attack

All the traffic that enters into the network is first directed to the client honeypot. The honeypot is a tool designed that checks for the state changes and registers the changes. Any change in state signifies an attack is taking place. This raises an alarm. The complete behavior of packers that are sent to honeypot can be calculated through the formulae as shown below. This formulae also helps to monitor and identify an attacks.

Assumption 1 The traffic is flowing constantly, τ

Assumption 2 Let us assume that there is no hostility and the traffic is having no intrusion. Let this condition be '***P***i'.

$$\textbf{\textit{If there is an hostile activity}},\ \text{the state changes from}\ P_i \rightarrow P_{i+1} \tag{1}$$

The change in the state change remains for some time '*t*'.

$$\text{Hence, the total effect of the state change is}\ (\boldsymbol{P_{i+1} \cdot t}) \tag{2}$$

$$\text{Assume the attack monitored through the passive attack remains for the time}'N' \tag{3}$$

$$\text{Combining equation}\ 1, 2\ \text{and}\ 3,\ \text{the total effect of the attack is}\ 'N.P_{i+1} \cdot t' \tag{4}$$

In case, there is several different forms of attacks, say ***'Q'***, (as can be seen from Wireshark window), then the severity of attack can be explained by the formulae as in the equation given below:

$$\text{Severity of Attack} = (\text{N} \cdot \text{Pi} + 1 \cdot \text{t}) \cdot \text{Q}$$

4.1 Formulae for Severity of Attack

The novel formula is used as a tool to realize to what extent the enterprise has been exposed to the intruders post attack. This formula classifies the attack into old signatures and new attack and then categorizes them.

5 Performance Testing of Formulae on Various Platforms

A questionnaire was developed and sent to the number of network administrators. They have been asked to rate their understanding for the newly developed proposed formulae for the assessment of cost-benefit analysis. Based on the feedback of the administrators who have visited the lab and those who are managing their own lab

Table 5 Comparison of performance of the conventional IDS and proposed IDS on various parameters

Installation	IDS with conventional system	IDS with formulae
Ease of deployment	Yes	Yes
Policy definition	No	Refined
Authentication and access control	No	Yes
Maintainability of attack signatures	No	Yes
Research and respond to new attack signatures	No	Yes
Capability to Auto fix certain Vulnerabilities	Yes	Yes
SNMP policy to isolate the network	No	Yes
Log file/database maintenance	Yes	Yes
Limitation and restrictions on enterprise-wide alerting and reporting	Manual	Automated

with traditional system, the tabular report has been generated. The sample size taken was 37. The Table 5 highlights the effectiveness of the assessment of cost-benefit analysis of the Intrusion detection system by comparing the performance of the conventional IDS and proposed IDS on various platforms.

With the help of the formulae the given tool assesses the severity of attacks. For the 3 months, the tool monitored the log activity with the help of wire shark. The results were stored in the window after clicking the stop window. The IP address of the hackers could be traced from various locations. If there were series of 'SYN' and the three—way handshake for the TCP protocol wasn't complete. It showed the attack is "Denial of Service". The database for the IP addresses was fed into the pool list so that the system can capture the traffic, scan it and then forward it to the internal network. All details for the same was stored in the traffic list at ACL. The new forms of attacks were detected as the tool scans for the signatures. The deep packet scanning also takes place as the packets are scanned heavily. The new signatures are updated in the database for future purpose and the log files of are also maintained. Certain attacks such as virus, worms, spam are automatically taken care of. Log Monitoring was done through scanning the lines through the tool. Based on the logs, following were the statistics as shown in the graph depicted in Fig. 1.

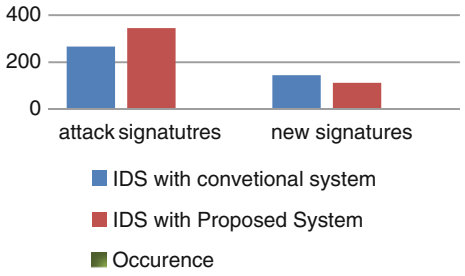


Fig. 1 Performance of conventional system and proposed system that categorizes the attacks based on the logs

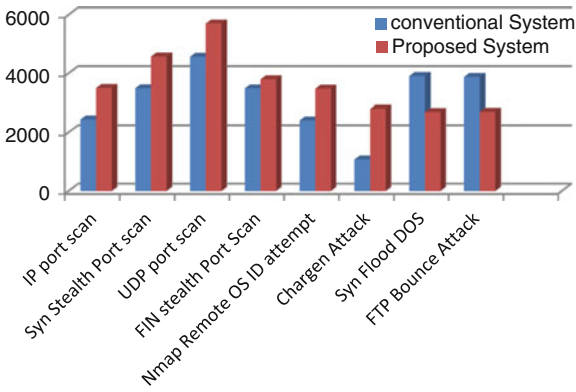


Fig. 2 Response to the different form of attacks

Since, the tests cannot be done live because of ethical issues. Hence, in the next test, using Scapy in Python, all malformed Packets were generated using suitable commands. All these malformed packets that are types for IP Port scan, SYN stealth Port scan, UDP port scan, FIN stealth port scan, Chargen attack, SYN Flood DoS, FTP Bounce attack. These malformed packets are then aimed at a single machine on the target subnet VLAN Guest. In order to maximize the impact, the rogue machine is kept on the same VLAN Guest. Then again the series are repeated for number of attacks simultaneously. There is no load on the network and no IP Fragmentation. The observation of the Log table of the machine shows that it can handle the attacks and easily recognize it and then generates the alerts. The log analysis also shows that the attack is low on FTP bounce attack and Syn Flood attack. The further analysis shows that the tool defers the traffic which are meant for FTP bounce attack and DoS attacks at the initial stage only. So, the tool is low on capturing the same. The conventional IDS are strong in the detection of DoS which raises alerts only and do not automatically defers it as is done by the proposed tool. The proposed tool when implemented has categorized the several attacks and has shown the response to the various types of attacks as shown below. So, the tool is low on capturing the same whereas conventional IDS are basically detecting DoS which are common.

The Fig. 2 explains the various forms of attacks that can be detected by the tool such as IP port scan, Syn Stealth Port Scan, UDP port scan, NMAP Remote OS ID attempt, Chargen Attacks and many more as well as their responses.

6 Conclusion and Future Scope

The severity of attack as discussed in the Sect. 3 is used to assess and analyze the impact of cost-benefit impact of implementing the honeypot based Intrusion Detection System. The proposed formulae assess the severity of attack through classification and categorization. If every form of attack is known, it can be easily

classifies, categorized into various forms of attacks and can be also used for comparison. Section 2 discusses in detail how we can categorize the various forms of attacks. The administrator needs to check and monitor the Command URLs are not injected into the system or stored in the repository for pattern matching. The attack may take any one form or combinations of all or many forms. Depending on how strong the monitoring is done and alerts are raised, the attacks can be minimized or can also be deferred. The monitoring of attacks can also be done effectively when there is an alert. Since the cost is included in the implementation of the security policies, strategies, implementation of the design assist in filtering packets at all the levels, the benefits of implementing the tool is how successful it is to defer an attack or minimize it. When the packet is reached from outside the network to inside the network, it is further filtered using Access Control Lists (ACLs) and port mirroring technique. The formulae is novel and can be applied in those organizations where an IDS is implemented for detection and response for the security of Cloud scenarios, SCADA networks and Security of Infrastructure.

References

1. Anderson, K.E.: International intrusions: motives and patterns., In: Proceedings of the Bellcore/Bell South Security Symposium (1994)
2. Chapter 13: Security technologies: internetworking technologies. Handbook, pp 1–12. http://docwiki.cisco.com/wiki/Security_Technologies (2003)
3. CERT: http://www.cert.org/resiliency/index.html (2004)
4. Verma, S., Singh, T.: Vulnerability assessment of web servers using honey pots: perspectives, ethical issues, legal implications. Int. J. Comput. Appl. **72**(22), (2013)
5. Verma, S., Singh, T.: Automated multilevel defence model to investigate packets for web interface. In: The 9th IEEE International Conference, WOCN, pp 1–6 (2012)
6. Ollman, G.: URL Embedded Attacks: Attacks using the Web Browser. http://www.technicalinfo.net/papers/URLEmbeddedAttacks.html
7. Hu, J., et.al.: Measuring botnet size by using URL and collaborative mail servers. In: The IEEE Fifth International Conference on Networking and Services, pp 161–164 (2009)
8. Wang, L., et.al.: Forwarding requests among reverse proxies. In: AT&T Labs. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.5.5603&rep=rep1&type=pdf
9. Mao, Z., Herley, C.: A robust link translating proxy server mirroring the whole web. ACM SIGAPP Appl. Comput. Rev. **11**(2), 30–42 (2008)
10. Furmanyuk, A., Karpinskyy, M.. Borowik, B.: Modern approaches to the database protection. In: Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 4th IEEE Workshop on, pp. 590–593 (2007)
11. Janczewski, L.J., Colarsk, A.M.: SQL code poisoning. In: Cyber Warfare and Cyber Terrorism, Information Science Reference Publisher (2008)
12. Barth, A., Berkeley: HTTP state management mechanism. In: Request for Comments: 6265, April 2011
13. Nikiforakis, N., et.al.: Session Sheild: lightweight protection against session Hijacking. In: ESSOS 2011, LNCS 6542, pp, 87–100 (2011)
14. Balachander, K., Jennifer, R.: Web protocols and practice: HTTP/1.1, networking protocols, caching and traffic measurement. In: AT& T Labs (2001)
15. Miyazaki, A.D.: Online privacy and the disclosure of cookie use: effects on consumer trust and anticipated patronage. J. Pub. Policy Market. 19–33 (2008). Spinger

16. Posey, B.: Buffer-overflow attacks: how do they work? http://searchsecurity.techtarget.com/news/1048483/Buffer-overflow-attacks-How-do-they-work (2005)
17. Toth, T., Kruegel, C.: Evaluating the Impact of Automated Intrusion Response Mechanisms. http://www.auto.tuwien.ac.at/∼chris/research/doc/2
18. Berger, E.D., et al.: Composing high-performance memory allocators. ACM SIGPLAN Notices **36**(5), 114–124 (2001)
19. Berger, E.D., et al.: Reconsidering custom memory allocation. In: Proceedings of the 17th ACM SIGPLAN Conference on Object-oriented Programming, Systems, Languages, and Applications, pp. 1–12. ACM Press, New York (2002)
20. Bua, J., et.al.: State of the art: automated black- box web application, vulnerability testing. In: IEEE Symposium on Security and Privacy, pp 332–345
21. Kirda, E., et.al.: Noxes: a client side solution for mitigating Cross site scripting attacks. In: 21st Annual ACM Symposium on Applied Computing, SAC'06, pp. 330–337 (2006)
22. Huang, Y.W., et.al.: Web application security assessment by fault injection and behavior monitoring. In: Proceedings of the 12th ACM International World Wide Web Conference (WWW 2003), pp. 148–159, May 2003
23. Huang, Y.W., et.al.: Securing web application code by static analysis and runtime protection. In: Proceedings of the 13th ACM International World Wide Web Conference (WWW 2004), pp. 40–52, May 2004

Author Index

N.R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, DOI 10.1007/978-81-322-2553-9