International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 637 - 640, 2013, ISSN: 2319–8656

Classification of The Insured On The Basis Of The Customers Loyalty Through Using Data Mining Techniques

Fateme Sadat Mousavi
Department of Computer Science and Research
Islamic Azad University
Guilan, Iran

Maryam Khademi
Department of Applied Mathematics
Islamic Azad University
Tehran, Iran

Hamid Reza Ahmadifar
Department of Computer
Guilan University
Guilan, Iran

Abstract: Insurance companies ask for too much information about their customers and the information was saved in their data base for many years, customers and prediction of their requirements as well as estimation of future customers were all classified on the basis of experimental inferences and guessing of the insurance companies. Therefore, most of the insurance companies carried out new and intelligent researches and investigations about their own customers through using information technology tools.
The main purpose of this research is to identify the loyal customers who have bought fire insurance policy and have extended it. The data available in information bank system of fire insurance in a private company has been investigated in the country during four years.
The obtained results involve classification and grouping of fire insurances on the basis of decision tree by using Rapidminer software. After that, the test data are separated from train data randomly and it is determined that 70% of data are used for training and the remaining 30% of data are expended for testing.

Keywords: Customers Classification and Grouping, Insurance Industry, Data Mining, Decision Tree, Rapidminer

1. INTRODUCTION

One of the important issues affecting on the management of dynamic companies such as insurance companies is related to customer relationship management. Using communicative and information technologies and tools as well as obtaining improved customer relationship management provide the continuous relationship with the customers and create their satisfaction and loyalty. In addition, it provides profits for insurance companies, and results in obtaining competitive advantages in comparison to other competitors. Hence, customer relationship management has been considered as the most important and strategic in organizations' decision making.According to obtained experiences, those organizations that consider the concepts of customer, goods, market, sales, shopping, competition, propaganda and quality as traditional concepts fail to be successful, and lose their investments as well [1].

2. NECESSITY OF CUSTOMER CLASSIFICATION

In recent years, the relationship between companies and customers has considerably changed. The organizations have found out that their success depends on proper understanding of the customers, prediction and considering their requirements and wants [5]. Therefore, classification of the current customers with the aim of presenting improved services and obtaining customers' loyalty has great and special importance. In addition, this classification allows the insurance companies to recognize and identify the aim of markets and potential customers according to behaviors patterns of the current customers. Also, it allows them to take into account marketing activities.

3. APPLICATIONS OF DATA MINING IN INSURANCE INDUSTRY

One of the problems of statistical analysis is that there are great data in insurance companies. The most important characteristics of data mining analyze data among great data or data storage. Due to this issue, analysts can analyze data easily. Nowadays, data mining has considerable role in the analysis carried out by insurance companies [3].

3.1 Customer Relationship Management

Nowadays, in competitive business, strong and stable management of customer relationship allows the companies to identify the profitable customers, and to create long-term plans for retention of their loyalty. The main and key competitive factors that should be taken into account by insurance companies are as follows: identifying profitable customers, retention of their loyalty for as long time and developing relationships. Through using data mining, insurance companies can determine the value of the customers and can predict their future behaviors. In this way, they can make certain decisions.

3.1.1. Dimensions of Customer Relationship Management

Four dimensions of CRM include:

- Customer Identification
- Customer Attraction

- Customer Retention
- Customer Development

These four dimensions can be considered as closed loop of customer relationship management.

3.1.1.1 Customers Identification

Customer relationship management begins with identification of the customers. This stage involves recognizing population who will be probably our customers or will be profitable. The elements of customer identification contain the analysis of target customers and customers grouping and classification. The analysis of customers includes searching profitable parts of the customers on the basis of basic characteristics of the customers. Customers' grouping and classification consist of dividing all customers into smaller groups, and involve relatively similar customers in each group.

In this research, according to above mentioned dimensions of customer relationship management, classification of current

3.1.1.2 Customers Attraction

After the stage of customer identification, this stage begins.Then based on identifying the potential customers, organizations can target the attempts and resources of direct attraction of the customers. The element of customer attraction is direct marketing. Direct marketing includes development process, and stimulates the customers to order through various channels. For example, direct email and/or distribution of coupons are direct marketing [6].

3.1.1.3 Customer Retention

This stage is the main topic of CRM. The main condition of customer retention is customer satisfaction, and this indicates the customers' expectations, wants and understanding their satisfaction. The elements of customer retention include individual marketing, personal marketing that is supported by analysis, identifying and predicting the changes occurring in the customers. Loyalty programs involve supporting activities with the aim of establishing long-term relations with the customers, credit ranking, and the quality of services or satisfaction of loyalty programs [6].

3.1.1.4 Customers Development

This stage consists of developing abundant transactions, value of transactions and recalling the customers. The elements of developing customers involve the analysis of customers' life period value and cross and up, and indicate promotion activities with the aim of increasing the related services that a customer presents. The main goal of market basket is to increase abundant transactions of the customer and to reveal the procedures in customers' behaviors.

The aim of CRM dimensions is to provide deep understanding of the customers, and to maximize customers' value for a long time. Indeed, the analysis of customers relationship management, behaviors and characteristics of the customers not only allows the managers to make future decisions about the company but also presents a competitive strategy for attraction and retention of the current customers as well as getting new customers.

3.2 Attraction of New Customers and Increasing Market Share

One of the important issues related to attraction of new customers is the way of presenting the rate. In traditional methods, increasing customers' attraction is as same as increasing the attempts of sales unit, but if these attempts are directed by the achievements obtained through extracting data, better results will be obtained [2].

3.3 Finding Potential Customers

Potential customersrefer to those customers changing into a good customer if they behave the customers in a proper way. Frequently, data mining has an important role in finding the customers. The most important results of this studyobtained data mining in customer service are as follows:

- Recognizing and identifying good customers
- Selecting the suitable channel to reach to potential customers
- Selecting the suitable message for various groups of customers

3.4 Determining Rate of Insurance Premium

One of the important issues affecting on the survival and retention of the customers and attending to them is the way of presenting the goods price and rate or presenting services by the organization. Hence, the way of determining the cost has an important role in retention of market share and the profitability of insurance companies. The insurance prices should be enough attractive for the customers, and should have competitive advantage. Also, it should have the advantage of compensation of damages and company prices [4].

In insurance industry, like other industries, the customer can select its own insurer. Some factors are involved in this decision such as price, the quality of services and

personal inclinations. In such environments, it is important to identify the customers leaving the company, and to determine the reason for taking preventive actions and retention of the customers.

4. IMPLEMENTATION OF SUGGESTED METHOD

Output model of decision tree classification has been demonstrated in figure 1.4 in terms of customers' loyalty. In this model, depth of the tree is equal to fourteen, and reliability method is X-Validation. Insurance record has been selected as dependent variable.

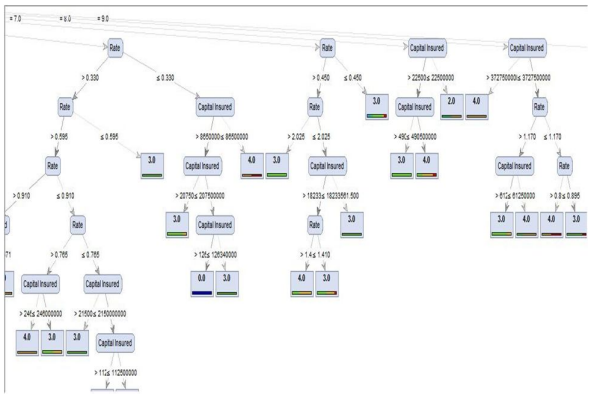


Figure 1: Part of Graphic output of tree model in terms of customers' loyalty

The obtained results show that, for instance, insurance policies belonging to the class of risk 1, and their rate is less that 1.62 and more than 1.37, and also their capital is more than 160,000,000 with the share of 1% are placed in class 0. That is, these customers do not have any insurance records in Alborz insurance. Insurance policies belonging to the class of risk 8, and their rate is more than 1.17, and also their capitals are more than 61,250,000 and less than 372,750,000 are placed in class 3 with the share of 0.9%. This means that these customers have insurance records for three years, and this indicates the customers' loyalty for three years.According to figure 2, the rate and amount of proper prediction of data in tree model is about 82.69%.

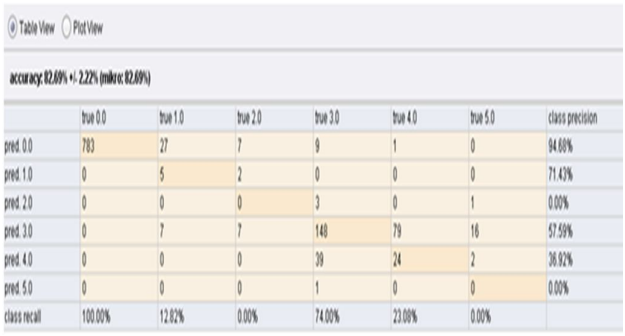


Table View Plot View

accuracy: 82.69% +/- 2.22% (mikro: 82.69%)

	true 0.0	true 1.0	true 2.0	true 3.0	true 4.0	true 5.0	class precision
pred. 0.0	783	27	7	9	1	0	94.68%
pred. 1.0	0	5	2	0	0	0	71.43%
pred. 2.0	0	0	0	3	0	1	0.00%
pred. 3.0	0	7	7	148	79	16	57.59%
pred. 4.0	0	0	0	39	24	2	36.92%
pred. 5.0	0	0	0	1	0	0	0.00%
class recall	100.00%	12.82%	0.00%	74.00%	23.08%	0.00%	

Figure 2: Precision of tree model

5. CONCLUSION

The aim of this research is to investigate and analyze the insured classification in fire insurance of a private insurance. Data that has been extracted through using decision tree method has been analyzed and investigated. According to the obtained results of data analysis related to insurance customers, these customers have insurance records for one year in fire insurance. Decision tree obtained from data mining shows that insurance policies belonging to the class with less risk don't indicate that they have more records. Insurance policies belonging to class with the risk of less than three years show one year record. In addition, sixty nine percent of the persons have bought insurance policy for the first time, while insurance policies belonging to class with risk of more than three have insurance records for three years. Almost 67% of the insured have insurance records for more than three years.

Based on the collected information, it has been demonstrated that those customers who have insurance policy with more risk show more loyalty. These persons extend their own insurance policy every year because they feel danger in terms of their own properties.

5.1 Suggestions for Future

According to the obtained experimental results, the customers having insurance policy with less risk show less inclination toward extending insurance policy in the company. Probably, one of the reasons is the less danger of these insurance policies. Since insurance policies having less risk are less damaged thanothers, and are more profitable,retention of the customers has great importance. The concentration of marketing activities is on the sales of insurance policies with less risk; therefore, damages will decrease, and profitability will increase for insurance companies.

6. REFERENCES

[1] Chalmeta, R. " Methodology for customer relationship management", The Journal of System and Software, 2005, 15:192-201.

[2] Gayle, S. 2000. Data Mining in the Insurance Industry. SAS Institute Inc.

[3]Gharehnezhad, S. (2010) The Necessity of Insurance Customer Retention Through Using Data Mining Techniques and Tools. Insurance Research, New World of Insurance, 150-151:2.

[4] Hosseinzadeh, L. 2007. Classification of Target Customers in Insurance Industry Through Using Data Mining. M.A. Thesis. TarbiatModarres University.

[5] Catler, F. 2003. Marketing management of Analysis, Planning, Executing and Controlling. Translated by Forouzandeh, B. Tehran, Taripat

[6] Ling, R., Yen, D. C. "Customer relationship management: an analysis framework and strategies", Journal of Computer Information Systems, 20001, 41, 82-97.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 641 - 644, 2013, ISSN: 2319–8656

Estimation of Beacon Update interval for GPSR Protocol in MANETs

Archana.S.S
R.V.S College of Engg. & Tech
Coimbatore, TamilNadu, India

Sathaya. B
Maharaja Engineering college
Avinashi, TamilNadu, India

Priyanka.S
Kathir College of Engg. & Tech
Coimbatore, TamilNadu, India

Soundaryadevi.M 4
R.V.S College of Engg. & Tech
Coimbatore, TamilNadu, India

Barath Kumar.S.S 5
Sri Shakthi Institute of Engg. & Tech
Coimbatore, TamilNadu, India

Abstract: In MANET, each node is needed to maintain its location information with the neighbor nodes for geographic routing. For this purpose they transmit the beacon packet periodically, which is not efficient in geographic routing due to packet collision. Also it has its effects in terms of update cost. Further the inaccuracy in local topology has its impact on the performance geographic routing. To overcome this in this paper we introduce the prediction scheme for Beacon Update and a prediction Scheme for Link Expiration Time. For geographic routing we use Greedy Perimeter Stateless Routing (GRPS) Protocol with greedy techniques, which select its next hop based on the optimization criteria. When greedy fails in selecting the next node nearest to destination we use perimeter rule which search over the perimeter of the node for destination node. Using the Mobility prediction scheme for Beacon technique in GPSR we try to decrease the update cost and increase accuracy of local topology.

Keywords: *beacons, broadcast, geographic routing, GPSR, link time estimation.*

1. INTRODUCTION

In MANET each node is free to move in an arbitrary manner. Hence it's necessary for nodes to maintain updated position information with the immediate neighbor. Also there will be frequent changes in the topology of the mobile nodes. In geographic routing, the destination node and the node in the forwarding path can be mobile. In such case it is necessary to reduce the effects caused by the changing topology, which is a difficult task in geographic routing to reconstruct the network topology frequently in presences of changing topology. Various routing schemes have been proposed which uses the location information such as LAR.

To obtain the location of node's neighbor, each node exchanges its location information (Using GPS) with its neighbor by periodic broadcast of beacons. Using this information each node build local map of the node, which is known as local topology. This periodic beaconing is not fair in terms of update cost and for the possible for collision. To overcome this drawback, in this paper we propose an efficient beacon scheme, which dynamically adjust the frequency for beacon update based on nodes mobility. Also we enhance the location information to estimate the expiration time of the link between two nodes.

2. LITERATURE SURVEY

Greedy Perimeter Stateless Routing (GPSR) [2] uses the neighbor list of node and destination location for forwarding decision. In Greedy forwarding strategy the next hop is selected based on the optimal path that is it always selects a node which is closest to the destination. Nodes broadcast beacon to immediate neighbors periodically for maintaining local topology.

In [4] showed that the inaccuracy of location information has a significant impact on the performance of geographic routing protocols. They applied a mobility prediction scheme as [3] to GPSR and studied its impact of on the performance. However, they only use the prediction scheme to compute current position of neighbors and still employed periodic beacon updates.

Heissenbuttel et al[5] has shown that periodic beaconing can cause the inaccurate local topologies in highly mobile ad-hoc networks, which leads to performances degradation. They proposed several simple optimizations that adapt beacon interval to node mobility or traffic load, including distance-based beaconing (DB), speed-based beaconing.

In the distance-based beaconing, a node transmits a beacon when it has moved a given distance d. The node removes an outdated neighbor if the node does not hear any beacons from the neighbor after a maximum time out. However, this approach has two problems. First, a slow node may have many outdated neighbors in its neighbor list since the neighbor time-out interval at the slow node is longer. Second, when a fast moved node passes by a slow node, the fast node may not detect the slow node due the infrequent beaconing of the slow node, which reduces the perceived network connectivity.

In the speed-based beaconing, the beacon interval is dependent on the node speed. A node determines its beacon interval which is inversely proportional to its speed. Nodes piggyback their neighbor time-out interval in the beacons. A receiving node compares the piggybacked time-out interval with its own time-out interval, and selects the smaller one as the time-out interval for this neighbor. In this way, a slow node can have short time-out interval, which eliminate first problem. But it still suffer from the second problem i.e. a fast node may not detect the slow node existences.

In [4] has shown that there is LLNK problem which occur in mobile ad hoc network due to mobility of nodes. For routing the selected hop may not exists in the expected location even though its in the neighbor list. Since due to mobility nature of the network that hop may moves out of the radio range. This situation is defined as lost link (LLNK) problem and can be caused by one of the following two reasons: Node mobility, Asymmetry in a communication link.

3. EXISTING METHODS

3.1 Adaptive Position Update (APU)

Initially, each node broadcasts a beacon to its neighbors to inform its presences stating its current location and velocity. Following this each node periodically broadcasts its current location information. This position information of beacons is stored at each node. Each node continuously updates its neighbor list based on its transmission range, current location and the position updates received from its neighbors. Neighbors which are outside the nodes communication range are not considered for data forwarding. Hence beacon is very important in building the local topology.

Instead of periodic beaconing, APU adapts the beacon update intervals to the mobility of the nodes and the amount of data being forwarded in the neighborhood of the nodes. APU uses two principles 1) nodes that are frequently changing its position are updated frequently 2) nodes which are in forwarding path are updated.

3.1.1 Mobility Prediction (MP) Rule

This rule adapts the beacon generation rate to the mobility of the nodes. Nodes that are highly mobile need to frequently update their neighbors since their locations are changing dynamically. On the contrary, nodes which move slowly do not need to send frequent updates. A periodic beacon update policy cannot satisfy both these requirements simultaneously, since a small update interval will be wasteful for slow nodes, whereas a larger update interval will lead to inaccurate position information for the highly mobile nodes.

In our scheme, upon receiving a beacon update from a node *i*, each of its neighbors, denoted by the set *N(i)*, records its current position and velocity and continues to track node *i's* location using a simple prediction scheme (discussed below). Based on this position estimate the neighbors *N(i)*, check whether node *i* is still within their transmission range and update their neighbor list accordingly. The goal of the MP rule is to send the next beacon update from *i* when the error between the predicted location in *N(i)* and *i's* actual location is greater than an acceptable value. To achieve this, node *i*, must track its own predicted location in its neighbors, *N(i)*.

$$X_p^i = X_l^i + (T_c - T_l) * V_x^i$$

$$Y_p^i = Y_l^i + (T_c - T_l) * V_y^i$$

We use a simple location prediction scheme based on the physics of motion to track a nodes current location. Note that, in our discussion we assume that the nodes are located in a two-dimensional coordinate system with the location indicated by the x and y coordinates. However, this scheme can be easily extended to a three dimensional system.

Let (X_a, Y_a), denote the actual location of node i, obtained via GPS or other localization techniques. (X_p, Y_p) denotes the predicated position of the node I at current time. Node i then computes the deviation $D^i devi$ as follows:

$$D_{devi}^i = \sqrt{(X_a^i - X_p^i)^2 + (Y_a^i - Y_p^i)^2} \qquad (1)$$

If the deviation (Obtained in (1)) is greater than a certain threshold, known as the *Acceptable Error Range (AER)*, it acts as a trigger for node i to broadcast its current location and velocity as a new beacon. The AER threshold is an important parameter that can affect the performance of the APU scheme. The MP rule, tries to maximize the effective duration of each beacon, by broadcasting a beacon only when the position information in the previous beacon becomes inaccurate. This extends the duration of the beacon for nodes by the following way: For low mobility node, reduces the number of beacons, for highly mobile node frequent beacons for the support of the rapidly changing topology.

3.1.2 On-Demand Learning (ODL) Rule

The MP rule solely may not be sufficient for maintaining an accurate local topology. Consider the example illustrated in Fig. 1, where node 1 moves from P1 to P2 at a constant velocity. Now, assume that node 1 has just sent a beacon while at P1. Since node 2 did not receive this packet, it is unaware of the existence of node 1. Further, assume that the AER is sufficiently large such that the MP rule is never triggered. However, as seen in Fig. 1 node 1 is within the communication range of 2 for a significant portion of its motion. If either 1 or 2 was transmitting data packets, then their local topology will not be updated and they will exclude each other while selecting the next hop node.

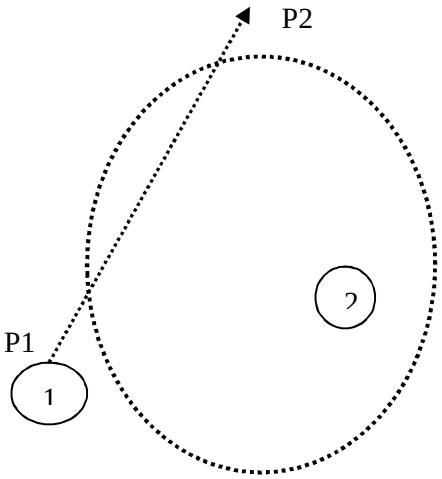
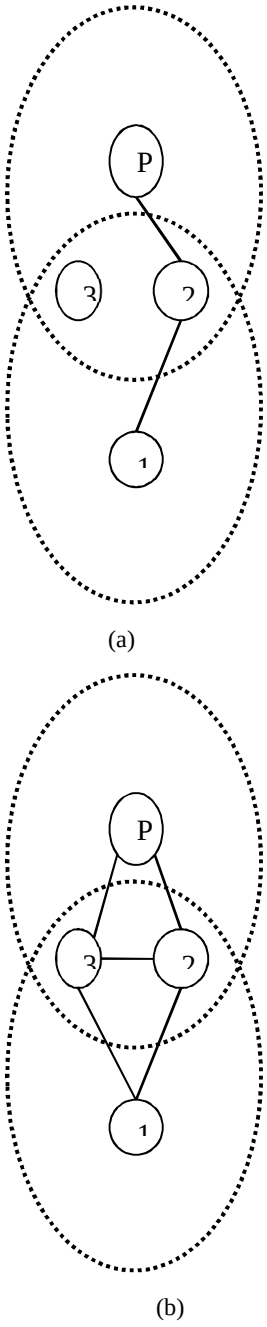


Fig 1. An example illustrating a drawback of the MP rule

Hence, it is necessary to devise a mechanism which will maintain a more accurate local topology in those regions of the network where significant data forwarding activities are on-going. This is precisely what the *On-Demand Learning (ODL)* rule aims to achieve. As the name suggests, a node broadcasts beacons *on-demand*, i.e. in response to data forwarding activities of that node. According to this rule, whenever a node overhears a data transmission from a *new* neighbor, it broadcasts a beacon as a response. The ODL rule allows active nodes that are involved in data forwarding to enrich their local topology beyond this basic set. Thus the rich list is maintained only at the active nodes and is built reactively in response to the network traffic. All inactive nodes simply maintain the basic neighbor list. By maintaining a rich neighbor list along the forwarding path, ODL ensures that in situations where the nodes involved in data forwarding are highly mobile, alternate routes can be easily established without incurring additional delays.

Figure. 2(a) illustrates the network topology before node 1 starts sending data to node P. The solid lines in the figure denote that both ends of the link are aware of each other. The initial possible routing path from 1 to P is 1-2-P. Now, when source 1 sends a data packet to node 2, node 3 receive the data packet from 1. Since node 3 is in communication range of node 1. As 1 is a new neighbor of 3, according to the ODL rule, node 3 will send back beacons to 1. As a result, the link 1-3 will be discovered. Further, based on the location of the destination and their current locations, node 3 will discover that the destination P is within their one-hop neighborhood. Similarly when 2 forward the data packet to P, the link 2-3 will be discovered. Fig. 2(b) reflects the enriched topology along the routing path from 1 to P.



(a)

(b)

Fig 2. An example illustrating the ODL rule

3.2 Prediction of Link Expiration Time

We introduce our mobility prediction method utilizing the location and mobility information provided by GPS. We assume a free space propagation model, where the received signal strength solely depends on its distance to the transmitter. We also assume that all nodes in the network have their clock synchronized (e.g., the GPS clock itself). Therefore, if the motion parameters of two neighbors (e.g., speed, direction, radio propagation range, etc.) are known, we can determine the duration of time these two nodes will remain connected. Assume

two nodes i and j are within the transmission range r of each other. The predicted time is the link expiration time (LET) between the two nodes. Let (X_i,Y_i) be the coordinate of mobile host i and (X_j,Y_j) be that of mobile host j. Also Let V_i and V_j be the speeds, and θ_i and θ_j be the moving directions of nodes i and j respectively. Then, the amount of time two mobile hosts will stay connected, D_t is predicted by:

$$D_t = \frac{-(ab + cd) + \sqrt{(a^2 + c^2)r^2 - (ad - bc)^2}}{a^2 + c^2}$$

Where

$$a = V_i \cos\theta_i - V_j \cos\theta_j$$

$$b = X_i - X_j$$

$$c = V_i \sin\theta_i - V_j \sin\theta_j$$

$$d = Y_i - Y_j \qquad (2)$$

Using (2), the routing protocol always selects routes with the largest LET for data forwarding. However they used this scheme only for link expiration by implementing periodic beacons.

4. PROPOSED METHOD

In [1], they used the Adaptive Position Update Scheme for dynamically adjust the frequency for beacon update by using GPSR protocol for geographic routing. However in GPSR, the selected next node for forwarding the packet may get lost in the forwarding path due to the mobility of node out of the range, which decreases the performances of routing protocols. In this paper we propose our work by using the dynamic adjustment of the frequency for beacon update interval and estimation of the expiration time of the link between the mobile nodes. Based on the link expiration time the next hop is selected for forwarding the packets. By selecting the next hop with greater link expiration time we can reduce the effects of LLNK problem, which occur due to mobility of node and asymmetric link communication.

5. CONCLUSIONS

In this paper, we proposed the Prediction scheme for beaconing scheme and the mobility prediction rule to estimate the link expiration time of the nodes to increase the performances of geographic routing protocol. Using java swan(Scalable wireless Ad hoc network) simulator the paper is under analysis. Future work includes the analysis of various geographic routing protocols using the above prediction scheme for beaconing and routing in Mobile Ad hoc Networks and to find optimal protocol parameters.

6. REFERENCE

[1] Q. Chen, S.S. Kanhere, and M. Hassan, “Adaptive Position Update for Geographic Routing in Mobile Ad Hoc Networks,” IEEE Transactions on Mobile Computing, vol. 12, no. 3, pp-489-501, March 2013.

[2] B. Karp and H. T. Kung,”GPSR: Greedy Perimeter Stateless Routing for Wireless Networks,” in Proceedings of MOBICOM 2000, Boston, MA, USA, 2000, pp. 243-254.

[3] W.Su, S.J. Lee, and M. Gerla, ”Mobility prediction and routing in ad hoc wireless networks”, International Journal of Network Management, 2001.

[4] D.Son, A.Helmy, B.Krishnamachari. ”The Effect of Mobility-Induced Location Errors on Geographic Routing in Mobile Ad Hoc and Sensor Networks: Analysis and Improvement Using Mobility Prediction”, in IEEE Transactions on Mobile Computing, vol. 3, pp. 233-245, July 2004.

[5] M. Heissenbuttel, T. Braun, M. Walchli, and T. Bernoulli, “Evaluating of the Limitations and Alternatives in Beaconing,” Ad Hoc Networks, vol. 5, no. 5, pp. 558-578, 2007.

INTERNATIONAL JOURNAL OF COMPUTER APPLICATIONS TECHNOLOGY AND RESEARCH
VOLUME 2– ISSUE 6, 645 - 649, 2013, ISSN: 2319–8656

Effects of using RZ and NRZ modulation formats for TDM-PON system on Transmission Characteristics for Downstream Signals

Nimrat Kaur
Sai Institute of Engg. & Tech,
Amritsar, Punjab, India.

Malti Sarangal
Sai institute of Engg. & Tech,
Amritsar, Punjab, India.

Abstract: In this paper, the simulative analysis of 8 channels TDM-PON system has been carried out by using NRZ and RZ modulation formats for downstream signals. The performance of TDM-PON system has been evaluated by varying input Pin (0 - 25) dBm for both modulation formats at a constant bit rate of 622Mbps.It has been observed that the TDM-PON system gives the optimum performance at power 20dBm and it is found that RZ is superior to NRZ. The transmission performance characteristic of TDM-PON for both the modulation formats has been evaluated in terms of Q-factor and BER. The System shows the repeaterless transmission up to the distance of 130 km.

Keywords: PON, TDM, NRZ, RZ, OLT, ONU.

1. Introduction

Telecommunication's future access network is increasing day by day with a need to meet greater bandwidth and reliable data transmission. Optical access network is the most demanding network for industry after the deployment of optical fiber in metro areas, where the use of optical fiber technology by reducing the CO(Central Office) equipments and maintenance cost of these systems with provision of higher data transmission operating over repeater less P2P networks. The overall aim is to provide better solutions with higher scalability to service providers to make fiber reach to end user. To accomplish it, Network needs to be promising improved for high data rates and long reach results by implementing optical fiber using Passive Optical networks [1].

Passive Optical Network is considered to be one of the most widely used networks in this scenario because they provide the cheapest way to provide fiber to home. A Passive optical network is a point to multipoint network consists of a central office node, called an optical line terminal (OLT), one or more user nodes, called optical network units (ONUs) and the fibers and splitters between them. The cost-effectiveness of PON comes from the two facts, one is other than passive element (splitter), and no other active element is present in the link. Secondly, a transmitter, the feeder fiber and remote splitter is shared among many subscribers.PON uses a single wavelength to transmit the data in both directions i.e in Downstream direction and upstream direction. Different flavors of PON are emerged in recent years; Ethernet-PON (EPON) which is a recent version, standardized in the IEEE 802.3, Broadband-PON (BPON), and Gigabit-PON (GPON).

2. BROADBAND PASSIVE OPTICAL NETWORKS (BPON)

Fast evolution access networks become evidence under the pressure of demand for broadband internet services, which are requested both by the residential and corporate users, today's predominant broadband access solutions deployed are digital subscriber line(DSL),community antenna television based networks, which can support up to 50Mb/s in downstream direction.

Broadband passive optical network is applied to support triple play services such as voice, video, data economically and effectively. BPON technology transmits 622Mb/s in downstream direction and 155Mb/s for upstream direction. This network uses the ATM technology as a data transport mechanism. ATM protocol provides support for different types of service by means of different adaptation layers .The small size of ATM cells and the use of virtual channels and links allow the allocation of available bandwidth to end users with a fine granularity.BPON operates on principle of Time Division Multiplexing, (TDM-PON), since PON delivers services with different multiplexing techniques, such as TDM, Hybrid-TDM, WDM (Wave division multiplexing) [2- 3].

2.1 TDM-PON

Time division Multiplexing PON systems proposed so far are mostly based on Bit interleaving technique for Time division multiplexing, This is because some users demands different speeds depending on subscribed services. This raised the need to delivered services in data packets or asynchronous transfer mode [4].The Data transmission is carried out after a particular time delay at transmitter side and same delay is utilized at receiver side in order to reduce traffic.

3. LITERATURE REVIEW

Shaukat et al.[5] analyzed the system on the basis of Data Rate, Fiber length, Coding technique, number of users, wavelengths and their effects on Bit Error rate (BER) as the key performance parameter using OptiSystem Simulator version 7. It had been found that, in downstream direction, higher data rates correspond to degrade system performance over differing fiber lengths provided the number of users was kept constant.

Bock et al. [6] they had proposed WDM/TDM-PON architecture by using free spectral range FSR periodicity and arrayed waveguide grating AWG. A shared tunable laser, photo receiver, DBA and remote modulation were used for Tx and Rx. Transmission test showed correct operation at 2.5Gb/s to a 30 km. by mean of optical transmission test, they demonstrated that this architecture was feasible and offered good performance with low optical losses as compared to other PON architectures.

Calabretta et al. [7] proposed an innovative architecture to realize a single feeder bidirectional WDM/TDM-PON on modified NRZ (DPSK) downstream signals at 20Kb/s and narrowband AWG.AWG was used as the channel distributer and demodulator for all the DSPK Channels. The experimental results showed that a system employing 8 * 20 Gb/s downstream TDM channels with a 1 : 4 power splitter can serve up to 32 ONUs at average traffic of 5 Gb/s (downstream) and 250 Mb/s (upstream).

Rajniti et al. [8] analyzed 2.5Gb/s eight channel bidirectional WDM-PON over transmission distance of 70km.The Performance of system was analyzed by varying the extinction ratio from 2 to22 dB, and concluded that with the increase in ER of the modulator from 2 to22 dB, BER (Bit error rate) decreases and Q value improves for WDM-PON system up to 20dB beyond which it saturates. Q value decreases sharply with increase in transmission distance. Finally concluded that NRZ is superior to RZ and system gives optimum performance at input power 10dB.

4. EXISTING WORK

With the standardization of Time-division-multiplexing passive optical networks (TDM-PONs), a cost-effective

Components	Parameters	
	Type	Value
Light Source	Frequency	193.1THz
	Power	5dBm
PRBS generator	Bit rate	2-5Gb/s
Photo detector	Responsivity	1A/W
	Dark current	10nA
Modulator	Modulation format	RZ
Time Delay	OLT1	0sec
	OLT2	1/(Bit Rate)*1/4 sec
	OLT3	1/(Bit Rate)*2/4 sec
Optical Fiber	Fiber length	20-100

Table1: The parameters and their values for TDM PON Network

access technology based on optics has been developed. However, further development needs to be carried out in order to fully exploit the benefits of optical fiber technology. The Existing technologies achieved performance with TDM PON and WDM PON using different Coding Schemes by the use of 16 users in Optical Network Unit (ONU) receiver side. They analyzed the performance achieved by TDM/WDM-PON for Manchester, non-return-to-zero (NRZ), return-to-zero (RZ) modulation formats. The Fig.1 shows the transmitter consists of pseudo generator and the bit sequence is modulated using a RZ/NRZ and Mach Zehnder modulator. Optical signal from each user is then combined using a power Combiner and is sent through the SMF. At the receivers end a Power Splitter is used, which directs the optical data signals to each one of the ONU and synchronized according to time intervals. The Simulation Set up, parameters used (with values) were as follows:

It was revealed that Manchester coding showed better performance through Min BER, Max Q factor, and Eye Diagram when compared with RZ and NRZ coding techniques.

5. PROPOSED WORK

In the proposed work, the simulation of 8 channels TDM-PON has been carried out by varying input Power by using RZ and NRZ modulation of formats. The Fig.2 shows at transmitter side, Optical transmitters (Bit rate 622Mb/s) operating at frequency 193.1, with variable input power (0 to 25 dBm) has been used by varying modulation formats.

Since the data is sent in form of packets, optical delay of 1 sec is used.

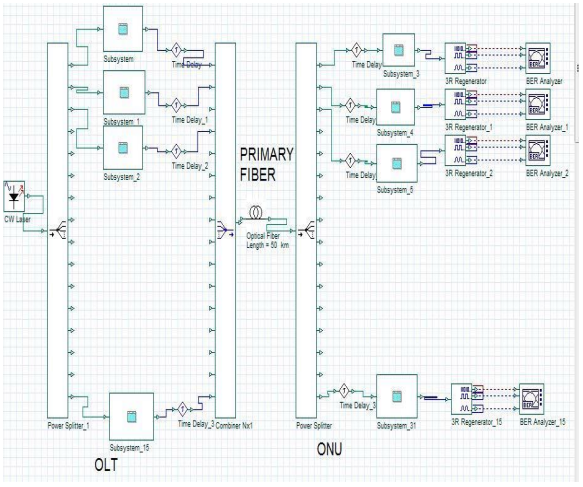


Fig.1: TDM-PON Simulation Set up

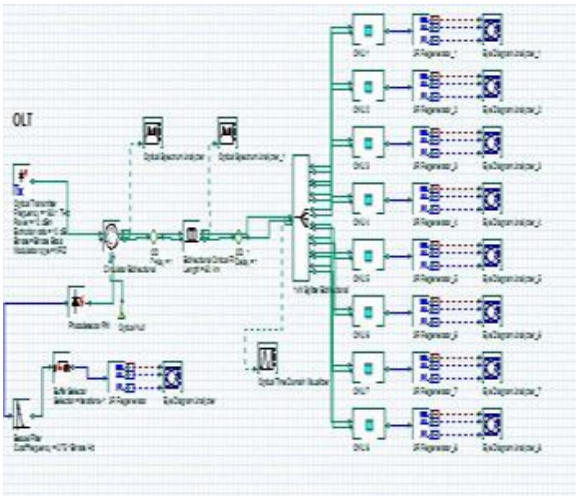


Fig.2: TDM-PON Simulation Set up

PIN photo detector detects the electrical signal. It has responsitivity of 1A/W and Dark current=10nA.The Optical Spectrum Analyzers and time domain visualizers are used to evaluate the input/output spectrum and time domain respectively. The Signal from Optical transmitter is sent through the bidirectional optical fiber includes attenuation affect=0.2dB/Km, Dispersion

Slope at=0.075ps/km-nm and Dispersion at=16.75ps/Km-nm.

At Receiver side, a Power Splitter is used to direct the signal to each of 8 ONU's, followed by 3R regenerator, which is used to regenerate the electrical signal. Each 3R regenerator is further directly connected to BER analyzer, which we used to obtain the graphs and to measure the performance of

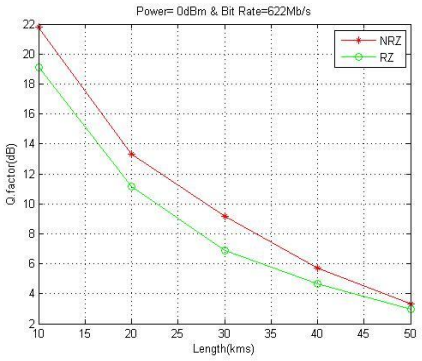
Table 2: The parameters and their values for Proposed TDM PON Network.

Components	Parameters	
	Type	Value
Light Source	Frequency	193.1THz
	Power	0 to 25 dBm
Optical transmitter	Bit rate	622Mb/s
Photo detector	Responsivity	1A/W
	Dark current	10nA
Modulator	Modulation format	RZ/NRZ
Time Delay	OLT	1sec
Optical Fiber	Fiber length	10-150km
Bessel Filter	Filter order	4
	Insertion loss	0dBm

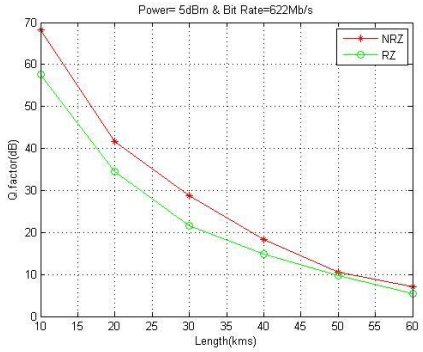
eye diagram, Min BER, Q-value and Eye Height.

6. RESULTS AND DISCUSSION

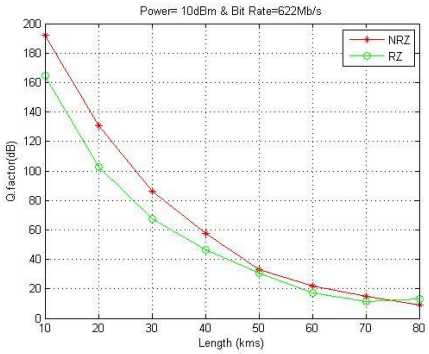
To estimate the performance, we have compared two modulation formats by varying by input power and Length of fiber. The Fig.3 (a), 3(b), and 3(c) shows the graphical representation of Q value as a function of transmission distance at varied input power (0-25) dBm and with constant Bit Rate of 622Mbps for NRZ and modulation formats for downstream direction respectively.



(a)



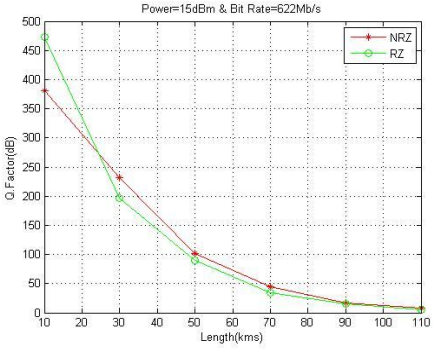
(b)



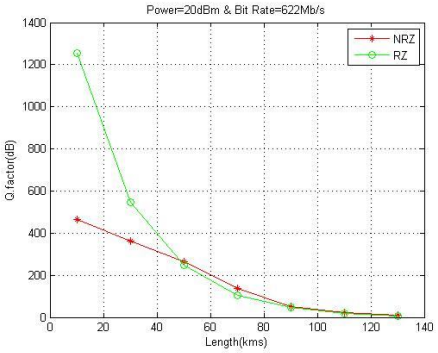
(C)

Fig.3: Length vs. Q-factor for NRZ, RZ modulation formats at Bit rate 622Mb/s (a) Pin=0dBm (b) Pin=5dBm (c) Pin=10dBm.

It is cleared from graphs with increase in Power and Length of fiber, system performance also increases in terms of NRZ Data format. This can be understood from the fact that for low powers, the performance of NRZ system degraded. However at higher powers, after Pin=10dB, the performance of RZ improved and gives better system performance as compared to NRZ.



(a)



(b)

Fig.4: Length vs. Q-value for NRZ, RZ modulation formats at Bit rate 622Mb/s (a) Pin=15dBm (b) Pin=20dBm

The fig.4 (a), 4(b) shows the graphical representation of Q-value as a function of transmission distance for varied Pin (0-25) dBm for NRZ and RZ modulation formats for downstream data respectively. We have evaluated the system performance by varying power rates up to 25dBm.It is clearly indicated from graphs as the increase the power, an abrupt change has been observed in system performance. It has been observed that with increase in higher power rates, RZ gives the better performance as compared to NRZ.RZ gives the optimum performance at power=20dBm than NRZ.

Fig.4 shows the eye diagrams for NRZ and RZ in downstream direction respectively. As the power rates increases, and reaches up to 10dBm, it is clearly visible from fig.4.1 that the performance of NRZ modulation format goes on decreasing due to signal distortion or XPM (Cross Phase modulation) effects i.e. decreasing the Q-value.

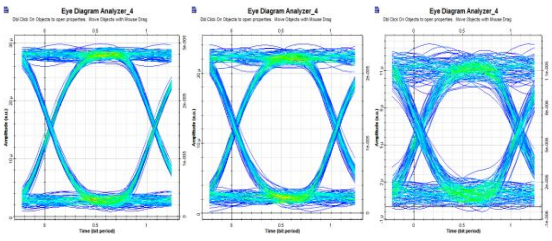


Fig.4.1: Eye diagrams for NRZ at Pin (a) 0dBm, (b) 5dBm, and (c) 10dBm at distance 10km, 40km, 80km respectively.

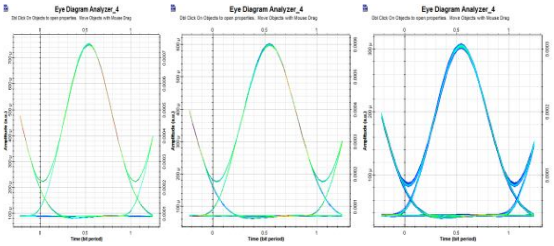


Fig.4.2: Eye diagrams for RZ at Pin (a) 15dBm, (b) 20dBm, and (c) 25dBm at distance 10km, 40km, 80km respectively.

7. CONCLUSION

We have simulated eight channels TDM-PON system over a transmission distance of 130km, with varying input power and length of fiber. It has been observed that with increase in power up to power =10dBm, and with increase in length of fiber, Q-value also increases for NRZ. But after higher power rates at 15dBm, 20dBm, and 25dBm respectively, it starts decreasing and RZ starts increases, and faithful transmission distance covered is 150 km. Further, it has been concluded that RZ gives optimum performance at Power 20dBm than NRZ.

8. REFERENCES

[1] David Nowak, John Murphy "FTTH: The Overview of Existing Technologies".

[2] Frank J. Effenberger,Hiroshi Ichibangase, "Advances in Braodband Passive optical Networking Techniques", IEEE Communication Magazine, December 2001.

[3] Brian Ford, Motoyuki Nakamura, "Experience of deployment of broadband optical access system and its future vista".

[4] Chang-Joon Chae, Member, IEEE, and Nam-Hyun Oh, "WDM.TDM-PON system employing a wavelength- selective filter and a continuous- wave shared light source", IEEE photonics technology letters, vol. 10, no. 9, September 1998.

[5] S.F. Shaukat, U. Ibrahim, "Monte Carlo Analysis of Broadband Passive Optical Networks", World Applied Sciences Journal 12 (8): 1156-1164, 2011, ISSN 1818-4952.

[6] Carlos Bock, Josep Prat and Stuart D.Walker, "Hybrid WDM/TDM-PON using AWG FSR and featuring centralized light generation and dynamic bandwidth allocation", journal of light wave technology, vol. 23, no. 12, December 2005.

[7] N.Calabretta, M.Presi, R.Proietti, G.Contestabile and E.Ciaramella, "A bidirectional WDM/TDM-PON is using DPSK downstream signals and a narrowband AWG", IEEE photonics technology letters, vol. 19, no. 16, august 15, 2007.

[8] Malti, Meenakshi Sharma, Anu Sheetal, "Comparison of CSRZ, DRZ and MDRZ Modulation Formats for High Bit Rate WDM-PON system using AWG", International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459, Volume 2,Issue 6,June 2012).

[9] Redhwan Q. Shaddad, Abu Bakar Mohammad, "Emerging Optical Broadband Access Networks from TDM PON to OFDM PON," PIERS Proceedings, Malaysia, March 27-30, 2012.

[10] Rajniti,Anita Suman, "Comparison of RZ and NRZ data Formats for 2.5Gb/s Bidirectional WDM/TDM-PON using Narrowband AWG" ,International Journal of VLSI and Signal Processing Applications,VOL.1,Issue 2,May 2011,(95-101),ISSN 2231-3133.

[11] Rajalakshmi, Ankit Srivastava, "Analysis of TDM and WDM PON using Different Coding Schemes for Extended Reach", International Journal of Computer Science and Network Security, Vol.11 No.7, July 2011.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 650 - 659, 2013, ISSN: 2319–8656

Reviewing Cluster Based Collaborative Filtering Approaches

F.Darvishi-mirshekarlou
Department of computer, Shabestar branch, Islamic Azad University
Shabestar, Iran

SH.Akbarpour
Department of computer, Shabestar branch, Islamic Azad University
Shabestar, Iran

M.Feizi-Derakhshi
Department of computer engineering, Tabriz University
Tabriz, Iran

Abstract: With regard to rapid development of Internet technology and the increasing volume of data and information, the need for systems that can guide users toward their desired items and services may be felt more than ever. Recommender systems, as one of these systems, are one of information filtering systems predicting the items that may be more interesting for user within a large set of items on the basis of user's interests. Collaborative filtering, as one of the most successful techniques in recommender systems, offers some suggestions to users on the basis of similarities in behavioral and functional patterns of users showing similar preferences and behavioral patterns with current user. Since collaborative filtering recommendations are based on similarity of users or items, all data should be compared with each other in order to calculate this similarity. Due to large amount of data in dataset, too much time is required for this calculation, and in these systems, scalability problem is observed. Therefore, in order to calculate the similarities between data easier and quicker and also to improve the scalability of dataset, it is better to group data, and each data should be compared with data in the same group. Clustering technique, as a model based method, is a promising way to improve the scalability of collaborative filtering by reducing the quest for the neighborhoods between clusters instead of using whole dataset. It recommends better and accurate recommendations to users. In this paper, by reviewing some recent approaches in which clustering has been used and applied to improve scalability , the effects of various kinds of clustering algorithms (partitional clustering such as hard and fuzzy, evolutionary based clustering such as genetic, memetic , ant colony and also hybrid methods) on increasing the quality and accuracy of recommendations have been examined.

Keywords: Information overload, Recommender systems, Collaborative filtering, Clustering, Fuzzy clustering, Evolutionary based clustering

1. INTRODUCTION

Due to expansive growth of the World Wide Web as well as increasing the amount of available information for each person, some problems have appeared for users in identifying useful, required and interesting information for each person. In most cases, people are faced with choices and very large data volumes, and searching all of them is out of user' capability. This problem is called information overload [1]. With regard to increasing high volume of data on the Internet, users have encountered with the problem of finding the right product at the right time. Finding final data on the basis of users' needs has became complicated and a time consuming process. In response to this growing epidemic, especially e-commerce, recommendation systems have been proposed .These systems are personalized technology for filtering information. In daily life, often some suggestions and comments of friends are used in choosing something. In addition, use our previous experience is used in selecting a specific item. Sometimes, it has happened that, on the basis of the suggestion of friends, we have bought a book or a particular product, or we have watched a film or have listened to music.

There are lots of e-commerce businesses in which one or more variations of recommender system technology is utilized in their web sites like Amazon.com (book recommendation system), movielens (move recommendation system), ringo (music recommendation) and etc. These systems use knowledge of the user's interests. This knowledge is obtained from searching the web pages for finding their favorite items or pages.Recommender systems are classified into three fundamental groups, namely, content based model, collaborative filtering and

hybrid systems. Content based systems save content Information and product specifications, and then provide some suggestions for users on the basis of the similarity between customer's purchase history and other items. Collaborative filtering (CF) act is based on the experience obtained from the purchase history of similar customers [4]. Hybrid systems try to combine above models in different ways to overcome the problems that have appeared because of using both content and collaborative filtering. Also these systems improve recommendation performance [26, 27, 28]. A brief review of collaborative filtering and its challenges are mentioned in section 2 and some of recent collaborative filtering approaches that are based on clustering are reviewed in section 3. In section 4 we give the conclusion of this work.

2. COLLABORATIVE FILTERING

Collaborative filtering, as one of the most successful techniques, is based on the assumption that people who have similar interests in terms of some items, they will have the same preferences in other items. So the goal of collaborative filtering is to find the users who have similar ideas and preferences or to find the nearest neighbor of them. This method is carried out in three steps: preprocessing, similarity computation and prediction / recommendation generation.

In preprocessing step, user-item matrix is built. This matrix contains the ratings that represent the expression of user's preferences. These preferences are explicitly obtained by rating the product in (1-5) scales, or implicitly by their purchase history [3].

Table1. A sample of user-item rating matrix

rating	Item1	Item 2	Item 3	Item 4
User 1	5	5	1	1
User 2	4	5	1	2
User 3	1	1	5	5
User 4	2	1	5	4
User 5	1	1	1	3

In similarity computation step, statistical techniques are used to find similar users with active user on the basis of their similar past behaviors. It reflects distance, correlation, or weight between two users. There are many different similarity measures to compute similarity or weight between users or items such as Cosine Vector, Pearson Correlation, Spearman Rank Correlation, Adjusted Cosine and etc. Pearson correlation, as the most common measure, has been mentioned.

$$W_{u,v} = \frac{\sum i \in I \ (r_{u,i} - \bar{r}_u)\ (r_{v,i} - \bar{r}_v)}{\sqrt{\sum i \in I \ (r_{u,i} - \bar{r}_u)^2} \sqrt{(r_{v,i} - \bar{r}_v)^2}} \quad (1)$$

$W_{u,v}$ is the similarity between two users. u and v. I is the set of co-rated items by tow users. $r_{u,i}$ and $r_{v,i}$ stand for the rating that has presented by user u and v in item i. $\bar{r}_u$ and $\bar{r}_v$ refer to the average rating of the co-rated items of the u th and v th users respectively.

In prediction step, weighted aggregate of similar user's ratings are used to generate predictions for active user. Finally, after predicting rating for items that have not been observed by active user, recommendation has been generated, and a list of items with high rating has been recommended to user.

$$P_{a,i} = \bar{r}_a + \frac{\sum u \in U \ (r_{u,i} - \bar{r}_u) \ . w_{a,u}}{\sum u \in U \ | . w_{a,u} |} \quad (2)$$

$P_{a,i}$ is the predicted rating for the active user, a, on a certain item, i. $\bar{r}_a$ and $\bar{r}_u$ stand for average ratings for the user a and user u in all other rated items, and $w_{a,u}$ is the weight between the user a and user u.

2.1 Memory/model based CF

Collaborative filtering is grouped into two general classes, namely, neighborhood-based (memory-based) and model-based methods. In Memory-based CF systems, the whole user-item rating dataset is used to make predictions. This system can be performed in two way known user-based and item-based recommendations. User-based collaborative filtering predicts an active user's rating in an item, based on rating information from similar user profiles, while item-based method look at rating given to similar items.[2]

Model-Based system contains building models on the basis of extracting information from datasets that can predict without using whole dataset. In these systems, the complete dataset is merged into train and test dataset. Train set is utilized to train the model using various algorithms like Bayesian network, clustering algorithm, regressions, matrix factorizations etc. Then, the trained model is used to generate recommendations for active user in the test set. The test set must be different and independent from the training set in order to obtain a reliable estimation of true error. Management of low density of data set is one of the most important advantages of model-based system that improves the scalability of big data sets. Because of the off-line building of models, the response speed time will be decreased, and less memory is used. The high cost of building these models is disadvantage of model-based systems [2].

2.2 Challenges of CF

There are three fundamental challenges for collaborative filtering recommender systems such as Data sparsity, Cold Start and Scalability.

Data sparsity: this issue is take place when the user-item matrix is extremely sparse, that is, users rate only a small number of items, so accuracy of recommendation will be decreased.

Scalability: with development of e-commerce and growing the number of users and items in such systems, the Scalability will increase, and ultimately the prediction calculations will be prolonged. Dimensionality reduction, clustering and item-based collaborative filtering are more common ways to alleviate this challenge.

Cold-start: when a new user or new item enters the system, it is difficult to find similar ones, because there is not enough information about his/her history in system. To overcome this issue, the hybrid system is used commonly. In this system, both rating and content information are used for users or items for prediction and recommendation.

2.3 Evaluation metrics

Several metrics have been proposed in order to evaluate the performance of the various models employed by recommender systems. Statistical Accuracy metrics and Decision support metrics are two major evaluation metrics. Statistical Accuracy metrics measure how much the predicted rate is close to the true rating that is expressed by user. Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) are the most common error metrics [6].

In Decision support metrics, the accuracy of decisions in recommender system is measured. Recalling and precision are most common metrics obtained by classifying the recommendation according to table 2.

Table2. Recommendation outputs classification

Correct / obtained	Recommend	Not recommend
Recommend	True-Positive(TP)	False-Positive(FP)
Not recommend	False-Negative(FN)	True-Negative(TN)

$$\text{Precision}=\frac{tp}{tp+fp} \quad (3)$$

$$\text{Recall}=\frac{tp}{tp+fn} \quad (4)$$

$$\text{Accuracy}=\frac{tp+tn}{tp+tn+fn+fp} \quad (5)$$

Precision is the fraction of the recommended items that are interesting to users, and recall is the fraction of the items having recommended higher ratings [6].

2.4 Datasets

Most often used and freely available datasets for collaborative filtering are EachMovie (EM), MovieLens (ML) and JesterJoke datasets.

EachMovie is a movie rating data set collected by the Compaq Systems Research Center over an 18 month period beginning in 1997. The base data set contains 72916 users, 1628 movies and 2811983 ratings. Ratings are on a scale from 1to 6. **MovieLens** is also a movie rating data set. It was collected through the ongoing MovieLens project, and is distributed by GroupLens Research at the University of Minnesota. MovieLens contains 6040 users, 3900 movies, and 1000209 ratings collected from users who joined the MovieLens recommendation service in 2000. Ratings are on a scale from 1 to 5**. Jester Joke** data set is collected by Goldberg et al. The data set is much smaller than above datasets that containing 70000 users, but only 100 jokes.

3. CLUSTER BASED CF APPROACHES

Clustering, as one of the common techniques to grouping the object, is a promising way to improve the scalability of collaborative filtering approaches by reducing the search for the neighborhoods in the preference space, and generates some recommendations for users without using whole dataset. It means that after Clustering, users similar to active users are chosen from his own cluster instead of choosing them among all users. Fig1 shows the impact of clustering on reducing user-item matrix dimensions. By applying similarity calculation against clustering, time complexity reduces from O(N) to O(k) where k is number of clusters.

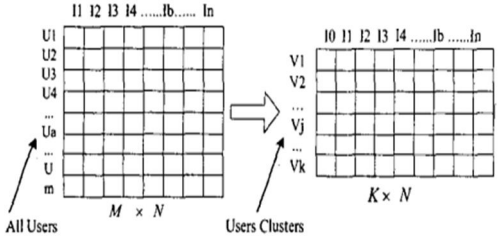


Figure 1. user clustering for collaborative filtering

There are many collaborative filtering based on partition clustering (hard/fuzzy) and evolutionary based clustering. In this paper, some recent approaches have been reviewed.

3.1 CF based on partition clustering

Clustering algorithms such as K-means assigns each user to a unique cluster. It is believed that this is too restrictive for most of the real world scenarios. In reality, users often have more diverse preferences. For example, one user may prefer both 'action movie' and 'comedy movie', or even more. By this consideration, it is more reasonable to allow a user to be assigned to more than one group. For this reason, fuzzy clustering algorithms are applied. so that users are clustered into different groups, and each user may be located in more than one group.

3.1.1 CF based on Fuzzy Clustering

An approach is proposed to improve Item based method in [7] employing FCM algorithm in item based collaborative filtering. In this method, items are partitioned into several clusters, and prediction is accomplished against clusters. By using clustering technique in Item-based, partitioning the data will be even less, so computational cost decreases. This approach shows that it overcomes scalability, cold start and sparsity. It saves more than 99% computational time, and does not change the prediction quality and eventually real-time prediction.

A new fuzzy algorithm is proposed in [8] employing Entropy based FCM in item based collaborative filtering. Despite of well accomplished IFCM [7], prediction results obtained from all clusters may include too much noise, and the accuracy would be affected. In Entropy based IFCM, higher value of degree of membership in objective function will exponentially increase, and it is distinct from lower value. Due to ability of using a wider range of fuzzifier values in IFCME, IFCME is more flexible than IFCM. The results show that IFCME improve MAE by 3.2% and 13.4% in single rate items in comparison to IFCM. Another fuzzy algorithm is presented in [9] who has formulated objective function with Exponential equation (XFCM) in order to increase the capability of assigning degree of membership. By the way, noise filtering is incorporated in XFCM, and noisy data are clustered differently in comparison to other Fuzzy Clustering. Thus, the centroid is strong in the noisy environment. The experiments show that centroid produced by XFCM becomes robust through improvement of prediction accuracy 6.12% over (FCM) and 9.14% over Entropy based FCM (FCME). Although using fuzzy clustering is more convenient to allow data to locate in multiple clusters [7, 8], this is not enough to make accurate recommendations because irrelevant data could be assigned to the clusters and overwhelm the rating predictions. In general, the ratings should be computed by using only ratings from relevant items. To overcome this issue, a new clustering algorithm is offered in [10] who reformulates the clustering's objective function with an exponential equation. This equation locates data to clusters by aggressively excluding irrelevant data. In this method, data that are farther from cluster, negative degree of membership is allocated. The negative membership degree indicates very low correlation of data and cluster. On the other hand, if the membership degree goes beyond 1, data truly belongs to cluster. Thus, these properties are used to filter out irrelevant data when the degree of membership is negative. The MAE results show that XFCM outperforms FCM by 5.2∼9.8%, FCME by 1.0∼6.1% and the item-based method by 2.7∼6.9%.

An improved FCM algorithm is presented in [11] who strengths item clustering by injecting FP-Tree approach to it. The reason of using FP-Tree approach is to calculate means of nominal and set type data. This means that the E-Commerce data always have many kinds of data types, like numerical, nominal, set and etc. For example, a product contains a number of features, brand, main function and price that are respectively nominal, set and numerical. Traditional FCM algorithms are not able to handle these types of data. Therefore, in order to overcome this problem and handle these types of data, FP-Tree is used. In this way, in the repeated part of FCM, FP-Tree algorithm is used to category data feature. Then, average of numerical feature is calculated, and average is considered as mean. Finally, the mean of one cluster is obtained. After clustering items, cluster-based smoothing technique [24] is employed to estimate probability of unseen term and fill the missing values in data set. The goal of using smoothing method is to handle the sparsity problem in collaborative filtering. At the end, neighbors of active user are selected and prediction for an item is calculated. Experiment results show that this framework works well, and can efficiently improve the accuracy of prediction.

A framework is proposed in [12] to extend traditional CF algorithms by clustering items and user simultaneously. This is approximately like Co-Clustering (COC) problem [25] in which each user or item can only belong to a single cluster, but the main difference is that each user or item can be located in multiple clusters (MCOC). For example, in a movie web site, a user may be interested in multiple topics of movies, and a movie can be interesting for different groups of

users from different aspects. So, FCM algorithm is used to obtain this goal. To combine subgroups with existing collaborative filtering in this approach, collaborative filtering algorithm is used in each subgroup, and then prediction results of each subgroup are unified. Finally Top-N recommendation is performed. Experimental results show that, when more subgroups are used, each subgroup becomes more density. Hence, data sparsity problem reduces in some CF methods. Also, short runtime of MCOC demonstrates its good efficiency. Using subgroups is a promising way to further improve the top-N recommendation performance for many popular CF methods.

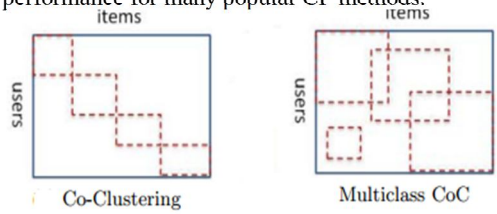


Figure 2. Comparison of COC and MCOC

A novel approach is suggested in [13] to alleviate the sparsity and scalability challenges in collaborative filtering recommendation. This approach firstly converts the user-item ratings matrix to user-class matrix. The user-classification matrix represents frequency expression of user preferences in terms of classification of items. Since the number of classifications is far less than that of the items, increase density of the data increases considerably in the resulted matrix. After converting matrix, FCM partitioning algorithm is applied to divide users into clusters. Finally, a list of Top-N recommendations is presented for each user. Computational experiments show that this approach is more effective and efficient than other well-known CF approaches, such as user-based CF and K- means-based CF.

New collaborative filtering method is proposed in [14] taking into account the impact of time on calculation of the users' similarity. It means that similarity of users is calculated in the same time or in similar periods. In this way, weight of time is assigned for each rating, and the weights of the most recent ratings demonstrate the latest ratings given by the users, and reflect the current interests of them, so the nearest neighbor will be found accurately. After finding similarity, FCM algorithm is used to clusters users. After clustering the user, fuzzy cluster-item rating matrix is constructed. It shows the rating given by a user cluster to an item. After calculating the dense user cluster-item rating matrix, similarities of items are calculated by choosing similarity measure. Finally, in recommendation generation, top N items that have the highest ratings are recommended to the user. Results show that the dimensions of the sparse user-item rating matrix are reduced, and improved algorithm can effectively raise the accuracy of recommendations.

3.1.2 CF based on Hard Clustering

A method is offered in [15] using cluster ensemble for collaborative filtering recommendation. In cluster ensemble result of some different clustering algorithms or result of several independent runs of one clustering algorithm are combined.

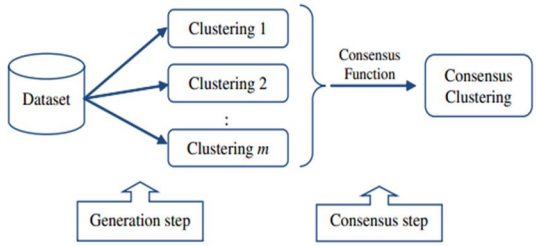


Figure 3. Clustering ensemble process

The main target in cluster ensemble is to have better performance than the single clustering algorithms. The goal of this approach is to show that employing cluster ensemble, in comparison to using single clustering, can improve recommendation accurately. In this approach, two base clustering algorithm, k-mean and SOM (Self Organized Map), are used. In this way, each clustering algorithm is run several times, and then the result of these runs is individually combined in census function. After clustering step, recommendation generation is performed once with SOM ensemble as well as k-mean ensemble individually. Finally, their results are compared with single k-mean and single SOM. Experiment results show that, by using cluster ensemble, recommendation accuracy and precise is high.

A hybrid method is proposed in [16, 17] overcoming cold-start problem in user and item respectively. The proposed method combines clustering and decision tree in which both the rating and content information are used. In this way and in the first step [16], items are clustered by k-mean clustering on the basis of ratings. K cluster obtained from clustering contains the items that are interesting for users with same preferences. One of the important content attributes of items in recommender systems is movie genres, such as action, comedy, romance, and etc. In the second step, decision tree is constructed in order to associate the new items with existing items. In this decision tree, genre information is used as attrib-

utes of tree, and clustering number obtained in the first step are used as the result to build decision tree. After building decision tree, in the third step, new item is classified. When new item without rating is entered to the system, the algorithm gains its attribute, and the item enters to decision tree. By following down the tree and answering the question correctly, the cluster number of new item will be achieved. Finally, in the last step, rating prediction is calculated, and recommendation is generated on the basis of this assumption that new item will be preferred by users who prefer the items in the cluster in which the item has been placed in the classifying procedure.

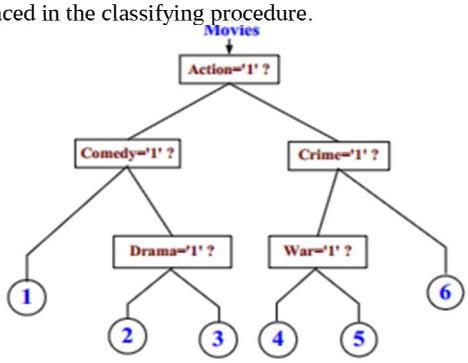


Figure 4. Example of decision tree in item side

According to [17], the proposed algorithm is used by user. Clusters are firstly used on the basis of ratings, and after that, decision tree is built by user demographic information (such as gender, age, occupation) and clustering numbers. The rest of algorithm is similar to [16], but it is used by user. Experiment results show that prediction accuracy is quite high in item and user's cold-start condition.

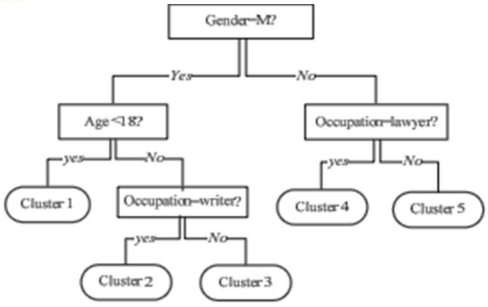


Figure 5. Example of decision tree in user side

A method has been proposed by [18] to alleviate sparsity problem in collaborative filtering in which item clustering and user clustering are used to generate recommendation simultaneously. This approach is performed in three phases. In the first phase, item based collaborative filtering is used for rating smoothing in which a full user-item matrix is obtained without non-ratings, and sparsity issue is resolved.

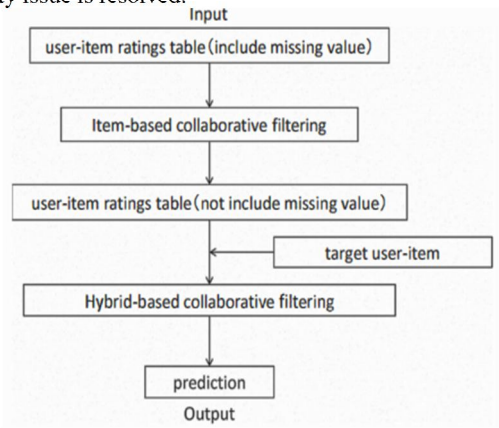


Figure 6. Stream of approach method

After rating smoothing, in the second phase, hybrid clustering technology is used to predict rating for target user. In hybrid clustering, users and items are firstly clustered in full user-item matrix. Prediction is calculated in the last phase. At first, users are found in the target user's neighborhood by user clustering result, and likewise items are found by item clustering result.

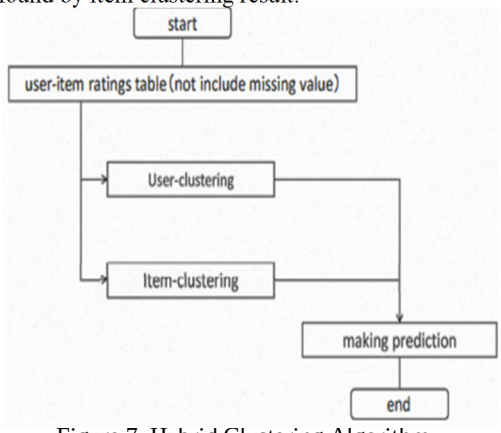


Figure 7. Hybrid Clustering Algorithm

Secondly, the user and item ratings are obtained from the full matrix. Finally, through those ratings, the item is recommended to target user. According to the experiment, it is concluded that the proposed collaborative filtering provides accurate prediction, and recommendation is more appropriate for the users because, in this method, the target user's neighborhood and target item's neighbor-

hood are used simultaneously for recommendation.

3.1.3 CF based on evolutionary clustering

On of drawback that classical clustering algorithms are faced with it is falling in local optima, and in this case, evolutionary clustering algorithms are recommended to resolve it.

A hybrid clustering has been used by [19] for recommendations in which k-mean clustering and genetic clustering have been combined. The most important pivot point of the k-mean clustering is picking up the right initial seeds. An unsuitable choice of clusters may outcome poor results, so quality of clusters is dependent on initial cluster centers. In this approach, Genetic algorithm is used to pick up suitable initial seeds for k-means clustering. In fact, the objective of this approach is to pick up optimal initial seeds to produce high quality recommendations. After clustering, movies are recommended to target user which are interesting for users in the cluster. The experiment results show that, in this hybrid method, percentage of correct prediction is better than simple k-mean, and provide better recommendations to the users.

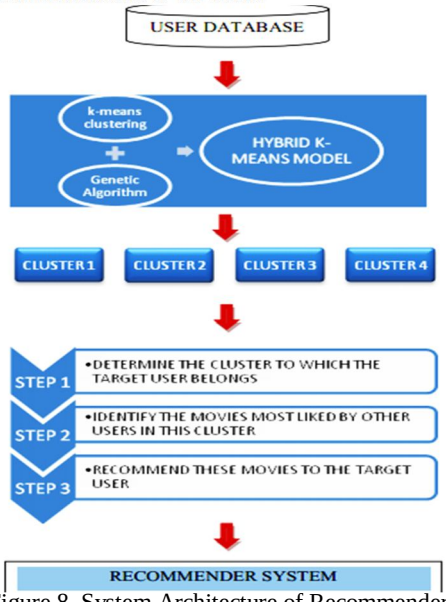


Figure 8. System Architecture of Recommender System

A method has been presented by [20] using Memetic clustering algorithm in collaborative filtering recommendations. Though traditional K-means clustering algorithms are simple and take less time in clustering, falling in local optima is more probable. Therefore, in order to overcome this problem, Memetic clustering algorithms are used in this approach. Memetic algorithms are different from Genetic algorithms. In this way, cultural evolution is constructed in Memetic algorithms by meme instead of gene, and before they get involved in evolutionary process, local search and quest has been embedded in the process to refine the solutions, so local search optimization is an elementary part of Memetic algorithms. In this approach, Collaborative filtering is performed in two phases. In the first phase, a model is extended on the basis of Memetic Clustering algorithm, and in the second phase, trained model is used to predict recommendations for active user. Experiments demonstrate that predictive accuracy of the proposed systems is clearly better than traditional collaborative filtering techniques.

Genetic and Memetic clustering algorithms are compared by [21] for collaborative recommendation. It has been concluded that accuracy of Memetic algorithm is better than the genetic algorithm in terms of prediction.

Ant based clustering is used by [22] for collaborative filtering in which recommendation process is based on the behavior of ants. In the natural world, ants release a certain amount of pheromone while walking for finding food and other ants follow a path which is rich of pheromone, and eventually find food. Since collaborative filtering is based on behavior of people with similar preferences, there is a similarity between ant behaviors and collaborative filtering. It means that recommendation is generated on the basis of similar users (ants) that are in same path (opinion). This approach contains two steps, namely, Clustering and recommendation. In clustering step, the rating matrix is clustered into K clusters through using ant based clustering, and Clusters data and centroid are stored for recommendation. Finally, Pheromone Initializing is performed for each cluster .In the second step, recommendation is generated. In this phase, at first, suitable clusters are selected on the basis of density of cluster and similarity with active user profile. Then, rating quality of items in each selected cluster is computed on the basis of average rating of the item in selected cluster and Variance of the ratings given by individual users for the item in the chosen cluster. Then, ratings of items are calculated by rating quality and average rating of the item in the chosen cluster. Later, the pheromone updating strategy is performed. It helps to increase the best solution pheromone, and in this way, the best clusters are chosen, and quality of recommendations improves. Finally TOP-N recommendation is generated for active user, and to sort the Pheromone Information for Future Experiment results of recommendations show that this

approach works better for large dataset in comparison to traditional collaborative filtering.
A method has been presented by [23] combining ant based clustering and FCM for clustering users, and locating users in suitable classes and providing best clusters of users with similar concerns. The initialization is the most important part of FCM. In this approach, calculation of initial clusters' centers is based on Ant colony algorithm. In this way, the ants move initially toward each user to form heaps. The centroids of these heaps are taken as the initial cluster centers, and the FCM algorithm is used to refine these clusters. In the second stage, the users obtained from the FCM algorithm are hardened according to the maximum membership criteria to build new heaps. After clustering, the adaption distinct between active user and most similar use clusters is calculated, and then appropriate recommendation are provided for user. Ant based Algorithm helps to provide optimal solutions, and FCM considers the uncertainty in user's interests. The results show that better recommendations will be suggested through using the proposed method for clustering users.

4. CONCLUTION

Recommender systems are considered as a filtering and retrieval technique developed to alleviate the problem of information and products overload. Collaborative filtering is the most popular and successful method that recommends the item to the target user. These users have the same preferences and are interested in it in the past. Scalability is the major challenge of collaborative filtering. With regard to increasing customers and products gradually, the time consumed for finding nearest neighbor of target user or item increases, and consequently more response time is required. There are some ways to overcome this drawback such as dimension reduction algorithms, item based collaborative filtering algorithm, clustering algorithms and etc. Clustering algorithms are the most effective techniques to overcome the scalability challenge. Clustering techniques partition the user or items on the basis of rating or other features, and then finding neighbors are performed within clusters instead of within whole dataset. Too many researches have been carried out in collaborative filtering, and so many approaches have been proposed for improving scalability and recommendation accuracy by applying clustering. Various kinds of clustering algorithms have been used by researchers to resolve scalability problem in collaborative filtering. In this paper, some recent approaches have been collected and surveyed. Fuzzy clustering is one of the common clustering algorithms in which different ways have been used and applied in collaborative filtering. In this paper, some of them have been mentioned. Fuzzy clustering allows users to be member of several clusters according to their variety of interests. Due to this issue, in many cases, better results are obtained in recommendations qualities by using fuzzy clustering. Also with regard to changes occured in objective function in fuzzy clustering such as using entropy function instead of Euclidean distance or using exponential function, better results are obtained. Fuzzy clustering is improved with FP-tree if that there are many kinds of data. In this case, better results are obtained in comparison to using k-mediod clustering. Using time factor in finding nearest neighbor before fizzy clustering is one of the factors considered in resolving scalability and sparcity problem. Clustering item and user fuzzy simultaneously is one promising way to offer better recommendations. Evolutionary based clustering algorithms along with resolving local optima of classical clustering algorithms can be used in collaborative filtering and to improve scalability. Different types of these algorithms have been mentioned in this paper such as genetic, memtic and ant colony clustering. Comparing genetic, memetic and k-mean clustering algorithms results in the same data set, and it is concluded that memetic algorithm has a lower prediction error (MAE) in comparison to genetic and k-mean clustering. Ant based clustering as a common way used in collaborative filtering has better results in comparison to k-mean clustering in large datasets. Through using hybrid algorithms such as a combination of genetic and k-mean clustering or a combination of ant colony and fuzzy clustering, better results are obtained in comparison to above mentioned algorithms.

5. REFERENCES

[1] Felfering, A., Friedrich, G. and Schmidt-Thieme, L. 2007. Recommender systems. IEEE Intelligent systems,pages 18-21.

[2] Taghi, M., Khoshgoftaar, Xiaoyuan Su 2009. A Survey of Collaborative Filtering Techniques. Hindawi Publishing Corporation Advances in Artificial Intelligence. Volume 2009, Article ID pages 421-425.

[3] Christian Desrosiers, George Karypis, 2011. A comprehensive survey of neighborhood-based recommendation methods.

[4] Daniar Asanov, 2011. Algorithms and Methods in Recommender Systems. Berlin Institute of Technology, Berlin, Germany.

[5] J. Ben Schafer, Joseph Konstan, John Riedl,1999. Recommender systems in e-

commerce. Proceedings of the 1st ACM conference on Electronic commerce

[6] Matteo Matteucci, Eugenio Lentini, Paolo Cremonesi, Roberto Turrin, 2010. An evaluation Methodology for Collaborative Recommender Systems.

[7] Chuleerat Jaruskulchai, Kiatichai Treerattanapitak, 2009. Item-based Fuzzy C-Means Clustering for Collaborative Filtering. NCCIT , volume 5,issue No10, pages 30-34.

[8] Chuleerat Jaruskulchai, Kiatichai Treerattanapitak,2009. Entropy based fuzzy C-mean for item-based collaborative filtering", In Proc. the 9th Int. Symposium on Communication and Information Techonology, pages 881-886

[9] Chuleerat Jaruskulchai, Kiatichai Treerattanapitak,2010. Membership enhancement with exponential fuzzy clustering for collaborative filtering. In Proc. the 17th Int. Conf. Neural Information Processing.volume part1, pages 559-566.

[10] Chuleerat Jaruskulchai,Kiatichai Treerattanapitak, 2012. Exponential Fuzzy C-Means for Collaborative Filtering. journal of computer science and technology, volume 27, issue No3, pages 567-576.

[11] Yanhui Ma, Yonghong Xie 2010. A Novel Collaborative Filtering Framework Based on Fuzzy C-Means Clustering Using FP-Tree. International Conference on Computer Application and System Modeling. IEEE,volume12, pages 4-9.

[12] Bin Xu, Chun Chen, Deng Cai, Jiajun Bu 2012. An Exploration of Improving Collaborative Recommender Systems via User-Item Subgroups. International World Wide Web Conference Committee, france, pages 21-30.

[13] Nai-Ying Zhang, Jing Wang, Jian Yin2010. Collaborative Filtering Recommendation Based on Fuzzy Clustering of User Preferences. Seventh International Conference on Fuzzy Systems and Knowledge Discovery, volume 4, pages 1946-1950.

[14] DongYue Ling Zheng, Xinyu Zhao, Shuo Cui 2010. An Improved Personalized Recommendation Algorithm Based On Fuzzy Clustering. 3rd International Conference on Advanced Computer Theory and Engineering, volume 5, pages 408-412.

[15] Chih-Fong Tsaia, Chihli Hung, 2012. Cluster ensembles in collaborative filtering recommendation, journal of Elsevier, volume12, issue No 4, pages1417-1425.

[16] Dongting Sun, Zhigang Lou, Fuahi Zhang 2011. A novel approach for collaborative filtering to alleviate the new item cold-start problem. the 11th international symposium on communication and information technology, pages 402-406.

[17] Dongting Sun, Zhigang Lou, Gong Li 2011. A Content- Enhanced approach for cold-start problem in collaborative filtering. IEEE, pages 4501-4504.

[18] Hideyuki Mase, Hayato Ohwada 2012. Collaborative Filtering Incorporating Hybrid-Clustering Technology. International Conference on Systems and Informatics, pages 2342-2346.

[19] anupriya Bhao, Prof. Dharminder Kumar, Dr. Saroj 2011. An Evolutionary K-means Clustering Approach to Recommender Systems", International journal of Computer Science and its Applications, pages 368-372.

[20] HemaBanati, ShikhaMehta 2010. MEMETIC Collaborative filtering based Recommender System. Second International Conference on Information Technology for Real World Problems, pages 102-107.

[21] Hema Banati, Shikha Mehta 2010. a multiperspective evaluation of MA and GA for collaborative filtering recommender system. International journal of computer science & information Technology, volume 2, issue No5, pages 103-122.

[22] H. Kaur, P. Bedi, R. Sharma 2009. Recommender System Based on Collaborative Behavior of Ants. Journal of Artificial Intelligence, pages 40-55.

[23] Ayoub Bagheri, Mohammad. H. Saraee, Mohammad Davarpanah, Shiva Nadi1 2011. FARS: Fuzzy Ant based Recommender System for Web Users. International Journal of Computer Science Issues. January,volume 8, issue No 1, pages 2-3-2009.

[24] Chenxi Lin, Gui-Rong Xue, Qiang Yang, WenSi Xi, Hua-Jun Zeng, Yong Yu,Zheng Chen 2005. Scalable Collaborative Filtering Using Cluster-based Smoothing. ACM SIGIR'05 Conf, pages114-121.

[25] Thomas George, Srujana Merugu 2005. A scalable collaborative filtering framework based on co-clustering. In Proceedings of the IEEE ICDM Conference.

[26] P. Melville, R. J. Mooney, and R. Nagarajan 2002. Content-boosted collaborative filtering for improved recommendations. in Proceedings of the 18th National Conference on Artificial Intelligence (AAAI '02), pages 187-192.

[27] M. Balabanovi, Y. Shoham, 1997. Content-based, collaborative recommendation. Communications of the ACM, volume 4,issue No 3, pages 66-72.

[28] L. Roy, R. J. Mooney 1999. Content-based book recommendation using learning for text categorization. in Proceedings of the Workshop on Recommender Systems: Algorithms and Evaluation (SIGIR '99). Berkeley. Calif, USA

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 660 - 662, 2013, ISSN: 2319–8656

Applying Data Mining Techniques on Soil Fertility Prediction

S.S.Baskar
Department of computer Science
St. Joseph's college
Trichy, India

L.Arockiam
Department of computer Science
St. Joseph's college
Trichy, India

S.Charles
Department of computer Science
St. Joseph's college
Trichy, India

Abstract: The techniques of data mining are very popular in the area of agriculture. The advancement in Agricultural research has been improved by technical advances in computation, automation and data mining. Now a days ,data mining is being used in a vast areas .The products of data mining system and domain specific data mining application soft ware's are available for tailor made use, but data mining in agricultural on soil datasets is a relatively a young and contemporary research domain. Larger volume of data are harvested along the with the crop harvest in agriculture. Inferring the knowledge from huge volume of data is virtually a difficult task in the current scenario. This research uses the data mining techniques for analysis of soil dataset. This data mining algorithms are used for analysing the soil datasets for classification purposes. The various techniques of data mining is used and compared in this research.

Keywords: data mining, classification, regression, Naive Bayes,J48

1. INTRODUCTION

Data Mining is a modern research domain in past. The techniques are useful to draw out significant and utilizable knowledge which can be perceived by many individuals. Data mining programs consists of diverse methodologies which are predominantly produced and used by commercial enterprises and biomedical researchers. These techniques are well disposed towards their respective knowledge domain. In general, the statistical techniques are time consuming and highly expensive. Efficient techniques can be developed and tailored for solving complex soil data sets using data mining to improve the effectiveness and accuracy of the Classification of large soil data sets [1]. A soil test is the analysis of a soil sample to determine nutrient content, composition and other characteristics. Tests are usually performed to measure fertility and indicate deficiencies that need to be remedied [2]. The soil testing laboratories are provided with suitable technical literature on various aspects of soil testing, including testing methods and formulations of fertilizer recommendations [4]. It helps farmers to decide the extent of fertilizer and farm yard manure to be applied at various stages of the growth cycle of the crop. In a research carried out by Leisa J. Armstrong, comparative study of current data mining techniques such as cluster analysis and statistical methods was carried out to establish the most effective technique. They used a large data set extracted from the Western Australia Department of Agriculture and Food (AGRIC) soils database to conduct this research. The experiments analysed a small number of traits contained within the dataset to determine their effectiveness when compared with standard statistical techniques [3]. In our approach, we have developed an automated system for soil classification based on fertility. After obtaining the fertility class labels with the help of automated system, we carried out a comparative study of various classification techniques with the help of data mining tool known as WEKA. The dataset used, was collected from one of the soil testing laboratories in Trichirapalli District (Tamil Nadu, India).Rest of this paper focuses on the prediction of untested attributes. This research has implemented a very sound practical application of linear regression technique by forecasting an obscure property of the soil test. The outcome of this research will result into substantial diminution in the price of these tests,

which will save a lot of efforts and time of Indian soil testing laboratories.

2. RESEARCH METHODOLOGY

2.1. Dataset Collection

The dataset is part of surveys which are carried out regularly in Trichirappalli District. Primary data for the soil survey are acquired by field sampling. These samples are then sent for chemical and physical analysis at the soil testing laboratories; hence this dataset was collected from a private soil testing lab in Trichirappalli. It contains information about number of soil samples taken from 3 regions of Trichirappalli district (Manapparai , Vayampatti, and Manikandan Block). Dataset has 9 attributes and a total 2188 instances of soil samples. Table1 describes data collected for each soil sample.

Table-1 Attribute Description

Field	Description
N	Nitrogen
P	Phosphorous
K	Potassium
Fe	Iron conent in the soil
Mn	Manganes content in the soil
Zn	Zinc content in the soil
Cu	Copper content in the soil
pH	pH
EC	Electrical conductivity (dS/m)
OC	Organic Carbon percentage

2.2. Automated System

Soil classification system is essential for the identification of soil properties. Expert system can be a very powerful tool in identifying soils quickly and accurately .Traditional classification systems include use of tables, flow-charts. This type of manual approach takes a lot of time, hence quick, reliable automated system for soil classification is needed to make better utilization of technician's time [9]. We propose an automated system that has been developed for classifying soils based on fertility. Being rule-based system, it depends on facts, concepts, theories which are required for the implementation of this system. Rules for soil classification were collected from soil testing lab. The soil sample instances were classified into the fertility class labels as: Very High, High, Moderately High, Moderate, Low, and Very Low. These class labels for soil samples were obtained with the help of this system and they have been used further for comparative study of classification algorithms.

3. A COMPARATIVE STUDY OF SOIL CLASSIFICATION

The classification of soil was considered critical to study because depending upon the fertility class of the soil the domain knowledge experts determines which crops should be taken on that particular soil and which fertilizers should be used for the same. The following section describes Naive Bayes, J48, JRip algorithms briefly.

3.1. Naive Bayes

Naive bayes classifier is very simple and strong classifier based on applying Bayes' theorem with strong independence assumptions. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. An advantage of the naive Bayes classifier is that it only requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification [5].

3.2. J48 (C4.5)

J48 is an open source Java implementation of the C4.5 algorithm in the Weka data mining tool. C4.5 is a program that creates a decision tree based on a set of labeled input data. This decision tree can then be tested against unseen labeled test data to quantify how well it generalizes. This algorithm was developed by Ross Quinlan. It is an extension of Quinlan's earlier ID3 algorithm. C4.5 uses ID3 algorithm that accounts for continuous attribute value ranges, pruning of decision trees, rule derivation, and so on. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier [6].

Table-2 Comparison of different classifiers

Classifier	Naïve Bayes	JRip	J48
Correctly Classified Instances	855	1998	2065
Incorrectly Classified Instances	1345	202	135
Accuracy	38.86%	90.81%	93.86%
Mean Absolute Error	0.324	0.0313	0.0283

4. PREDICTION OF ATTRIBUTES

Using regression algorithms like Linear Regression, Least Median Square, Simple Regression different attributes were predicted. According to these results the values of Phosphorous attribute was found to be most accurately predicted and it depends on least number of attributes. When all attributes are numeric,

3.3. JRip

This algorithm implements a propositional rule learner, Repeated Incremental Pruning to Produce Error Reduction (RIPPER), which was proposed by William W. Cohen as an optimized version of IREP. In this paper, three classification techniques (naïve Bayes, J48 (C4.5) and JRip) in data mining were evaluated and compared on basis of time, accuracy, Error Rate, True Positive Rate and False Positive Rate. Tenfold cross-validation was used in the experiment. Our studies showed that J48 (C4.5) model turned out to be the best classifier for soil samples.

linear regression is a natural and simple technique to consider for numeric prediction, but it suffers from disadvantage of linearity. If data exhibits non-linear dependency, it may not give good results .In this case, least median square technique is used. Median regression techniques incur high computational cost which often makes them infeasible for practical problems [8]. Several regression tests were carried out using WEKA data mining tool to predict untested numeric attributes. Linear-Regression test for predicting phosphor gave the best and accurate results. These predictions can be used to find out phosphor content without taking traditional chemical tests in soil testing labs, and this will eventually save a lot of time. Statistical results of these tests are given in Table3. There were very limited variations amongst the predicted values of phosphor attribute. Though the Least Median of Squares algorithms is known to produce better results, we noticed that the accuracy of linear regression was relatively equivalent to that of least median of squares algorithm.

Table-3 Comparisons of Regression Algorithms

Algorithm	Linear Regression	Least Median Square Regression
Time taken to build the model	0.18 s	10.84 s
Relative Absolute Error	10.63%	10.08%
Correlation Coefficient	0.9710	0.9905

Table-4: Predictions on test data

Actual Value Using Soil Testing	Predicted Value Using Linear Regression	Error
10.3	10.661	0.361
7.7	7.431	-0.269
4.6	4.653	0.053
9.5	8.478	-1.022
2.9	3.035	0.135
5.1	4.915	-0.185
15.3	15.667	0.367
7	7.402	0.402
18.4	18.743	0.343
4.4	4.388	-0.012
13.5	13.438	-0.062

Here the Relative Absolute Error is nearly same for both the prediction algorithms. Even though Least Median Square regression gives better numeric predictions but the time taken to build the model is 67 times that of Linear Regression, hence computational cost used by Linear Regression is much lower than that of least median square technique.

5. CONCLUSION

In this paper, we have proposed an analysis of the soil data using different algorithms and prediction technique. In spite the fact that the least median squares regression is known to produce better results than the classical linear regression technique, from the given set of attributes, the most accurately predicted attribute was "N" (Nitrogen content of the soil) and which was determined using the Linear Regression technique in lesser time as compared to Least Median Squares Regression. We have demonstrated a comparative study of various classification algorithms i.e. Naïve Bayes, J48 (C4.5), JRip with the help of data mining tool WEKA. J48 is very simple classifier to make a decision tree, but it gave the best result in the experiment. In future, we can plan to build Fertilizer Recommendation System which can be utilized effectively by the Soil Testing Laboratories. This System will recommend appropriate fertilizer for the given soil sample and cropping pattern.

6. REFERENCES

[1] A. Kumar & N. Kannathasan, (2011), "A Survey on Data Mining and Pattern Recognition Techniques for Soil Data Mining ", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3,

[2] "Soil test", Wikipedia, February 2012

[3] L. Armstrong, D. Diepeveen & R. Maddern, (2004), **"**The application of data mining techniques to characterize agricultural soil profiles"

[4] "Methods Manual-Soil Testing in India", Department of Agriculture & Cooperation Ministry of Agriculture Government of India, 2011

[5] "Naïve Bayes", Wikipedia, February 2012

[6] "C4.5 (J48)", Wikipedia, February 2012

[7] W. Cohen, (1995)," Fast Effective Rule Induction, in Twelfth International Conference on Machine Learning.

[8] I. Witten & F. Eibe, (2005), "Data Mining: Practical Machine Learning Tools and Techniques",2nd Edition, San Francisco: Morgan Kaufmann,

[9] A. Al-Rawas, S. Al-Alwai, (1998), A. Bisma & Y. Al-Alwai, "Soil Classification Decision Support System Using An Expert System Approach", Engineering Journal Of University of Qatar, Vol. 11, p. 103-115.

[10] S. Cunningham & G. Holmes, (1999), "Developing innovative applications in agriculture using data mining", Department of Computer Science University of Waikato Hamilton, New Zealand, Technical Report.

[11] P. Gruhn, F. Goletti, & M. Edelman, (2000) " Integrated Nutrient Management, Soil Fertility, and Sustainable agriculture: Current Issues and Future Challenges", International Food Policy Research Institute, N.W. Washington, D.C. U.S.A, technical Report.

[12] R. Vamanan & K. Ramar, (2011), "Classification Of Agricultural Land Soils A Data Mining Approach", International Journal on Computer Science and Engineering, ISSN: 0975-3397, Vol.3

[13] P. Bhargavi & S. Jyothi, (2011), "Soil Classification Using Data Mining Techniques: A Comparative Study",International Journal of Engineering Trends and Technology

[14] J. Han & M. Kamber, (2006), "Data Mining: Concepts and Techniques" Second Edition, San Francisco: Morgan Kaufmann Publishers,

[15] S. Diplaris, G. Tsoumakas, P. Mitkas & I. Vlahavas, (2005), "Protein Classification with Multiple Algorithms", P. Bozanis and E.N. Houstis (Eds.): pp. 448 – 456

[16] A. Sharma, (2011), "A Comparative Study of Classification Algorithms for Spam Email Data Analysis"; International Journal on Computer Science and Engineering, ISSN: 0975-3397 Vol. 3

[17] M. Kumari & S. Godara, (2011), "Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction", International Journal of Computer Science and Technology,, Vol. 2, and ISSN: 0976

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 663 - 665, 2013, ISSN: 2319–8656

Probabilistic Opportunistic Access of Cognitive Radio Network using Partially Observable Markov Decision Process

Atul Singh Jamwal
Department of Computer Science and Engineering,
SVIET, Banur
India

Nidhi Bhalla
Department of Computer Science and Engineering,
SVIET, Banur
India

Abstract: Growing demands of spectrum utilization make us to think about cognitive radio. Cognitive radio is smart enough to adjust its parameters according to spectrum use thereby increasing the efficiency of spectrum access. It means it should have the ability to make optimistic decisions dynamically in a multi operated network. Partially observable markov decision process is the technique of making optimistic decisions among various alternatives. MDPs are useful for studying a wide range of optimization problems solved via dynamic programming and reinforcement learning. In this paper, we demonstrate how POMDP is useful while using cognitive radio network.

Keywords: Cognitive Radio, POMDP, spectrum heterogeneity, white Spaces

1. INTRODUCTION

Inefficient spectrum utilization is one of the main concerns in the field of spectrum access as most of the spectrum bandwidth remains unutilized. Moreover growing demand of spectrum access makes the situation worse. The breakneck proliferation of wireless devices and rapid growth of wireless services continue to stretch the limited spectral resource. In fact, most spectrum bands suitable for terrestrial wireless communication have already been allocated by the regulatory agencies to existing licensees. Cognitive radio networks have been considered as the viable technology to improve spectral efficiency. With primary licensee's consent, secondary users equipped with cognitive radios may be allowed [1] to transmit on primary bands when the Primary users are inactive. Serious challenges must be resolved in order for the cognitive radios to be acceptable. First, secondary users must not be disruptive to primary user communications. Secondly, an access mechanism is required to reduce contention between secondary users to efficiently share spectrum opportunities.

One of the most difficult thing but important problem when designing an Open Spectrum [3] Access Network is how the unlicensed users decide when and which channel they should sense and access without conflicting the communications among Primary Users. Also, coordination and synchronization among secondary users may be limited due to the decentralized nature of secondary user access, particularly if secondary users of different networks coexist. In addition, each secondary users may not be able to sense all channels due to the limitation on hardware or/and sensing capability. Open Spectrum allows unlicensed [5] (secondary) users to share spectrum with legacy (primary) spectrum users, thereby "creating" new capacity and commercial value from existing spectrum ranges. Based on agreements and constraints imposed by primary users, secondary users opportunistically utilize unused licensed spectrum on a non-interfering or leasing basis. Open spectrum system designs must also deal with *spectrum heterogeneity*, where spectrum available to secondary devices fluctuates with both location and time due to movement and traffic variations of primary users. A user seizing spectrum without coordinating with others can cause harmful interference with its surrounding neighbours, thus reducing available spectrum. In open spectrum systems, primary users' mobility and traffic variations result in the fact that the available spectrum observed by secondary devices fluctuates with both location and time. We call this property *spectrum* heterogeneity [4]. In addition, the interference constraint and the reward (i.e. throughput, connectivity) obtained on each spectrum band could be different due to non-uniformly partitioned spectrum bands, differences in power constraints and associated technology.

1.1 Markov Decision Process

Markov Decision Process is a mathematical tool which provides a mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker. In simple terms MDP is an automation of decisions where subsequent decisions are the outcome of transition from present state to next state. While transiting from one state to next state there can be two types of outcomes: one is immediate effect and other is future gain and optimistic [9] decision is the the action that makes the right tradeoffs between the immediate rewards and the future gains, to yield the best possible solution. The four components of an MDP model are: a set of states, a set of actions, the effects of the actions and the immediate value of the actions.

- **States:** When making a decision, we need to think about how your actions will affect things. The state is the way the world currently exists and an action will have the effect of changing the state of the world. If we think about the set of every possible way the world could be, then this is would be the set of state of the world. Each of these states would be a state in the MDP.
- **Actions:** The actions are the set of possible alternatives we can make. The problem is to know which of these actions to take in for a particular state of the world.
- **Transitions:** When we are deciding between different actions, we have some idea of how they will affect the current state. The transitions specify how each of the actions change the state. Since an action could have different effects, depending upon the state, we need to specify the action's effect for each state in the MDP. The most powerful aspect of the MDP is that the effects of an action can be probabilistic. Imagine we are specifying the effects of doing action 'a1' in state 's1'. We could say that the effect of 'a1' is to leave the process in state 's2', if there was no question about how 'a1' changes the world. However, many decision processes have actions that are

not this simple. Sometimes an action usually results in state 's2', but occasionally it might result in state 's3'. MDPs allow you to specify these more complex actions by allowing you to specify a set of resulting states and the probability that each state results.

- **Immediate Rewards:** If we want to automate the decision making process, then we must be able to have some measure of an action's value so that we can compare different actions. We specify the immediate value for performing each action in each state.

1.2 Partially Observable Markov Decision Process

A Partially Observable Markov Decision Process (POMDP) is a generalization of a Markov Decision Process. A POMDP models an agent [7] decision process in which it is assumed that the system dynamics are determined by an MDP, but the agent cannot directly observe the underlying state. Instead, it must maintain a probability distribution over the set of possible states, based on a set of observations and observation probabilities, and the underlying MDP.

An exact solution to a POMDP yields the optimal action for each possible belief over the world states. The optimal action maximizes (or minimizes) the expected reward (or cost) of the agent over a possibly infinite horizon. The sequence of optimal actions is known as the optimal policy of the agent for interacting with its environment

2. POLICY STRATEGIES USED IN SPECTRUM SENSING

Decision-making is the cognitive process leading to the selection of action among variations. One-way to automate the decision making process is to provide a model of dynamics [6] for the domain in which the machine will be making decisions. A reward structure can be used to motivate immediate decision that will maximize the future reward. POMDP is an aid in the automated decision-making. POMDP policy informs CR users what action to be executed. It can be a function or a mapping and typically depends upon the channel states. In this section, we provide detail formulation of policy strategy either optimal and sub optimal based on greedy approach for sensing decision.

- *Optimal Strategy*: Channel access based on POMDP is known as an optimal strategy which model the channel opportunity of network system as discrete time Markov chain with number of channel state and formulate as M=2^N states, where N is number of channel. The state diagram for N=2 is described in figure 1 where state (0, 1) indicates the first channel is available and the second channel is busy.
- *Suboptimal Strategy or Greedy Approach:* Due to the complexity of optimal policy computation when number of slot and channel increase, sub optimal protocol based on a greedy approach can be used. Greedy approach follows the problem solving heuristic by making local optimal decision choices with the hope of finding the global optimum. It reduces the dimension of states from exponential to linear by regarding to N, i.e. from M=2^N to N states.
- *Random Approach:* Random Policy approach is useful for many ill-structured global optimization decision problems with continuous and/or discrete variables. Typically random policies sacrifice a guarantee of optimality for finding a good solution quickly with convergence results in probability.

3. COMPARISON BETWEEN DIFFERENT STRATEGIES

We have compared three different policy strategies by inputting number of channels, bandwidth and transition probabilities of each channel and time horizon T and obtained average expected reward per slot for each strategy. In this work we have calculated throughput performance of the greedy sensing strategy, Random sensing strategy and Optimal sensing strategy with perfect spectrum identification in independent channels as shown in figure 1.

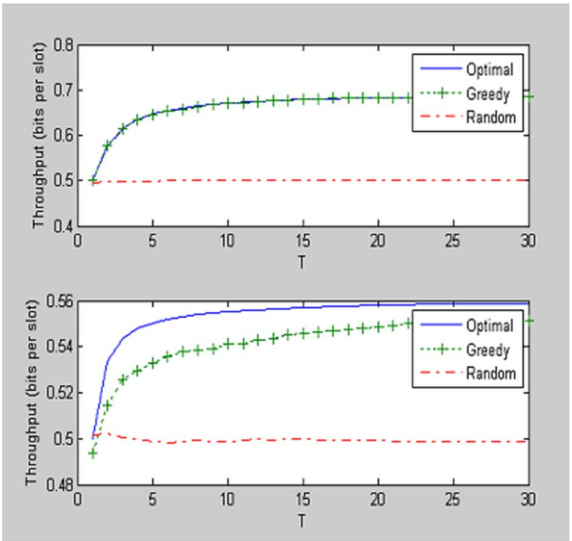


Figure 1: Comparison between different strategies

4. CONCLUSION AND FUTURE SCOPE

With the increasing demand for radio spectrum on one hand and inefficient usage of the licensed bands on the other, a reform of the spectrum access policy seems inevitable. Opportunistic spectrum access is envisioned to resolve the spectrum scarcity by allowing unlicensed users to dynamically utilize white spaces across the licensed spectrum on a priority basis. In this article we have implemented three decision policies to detect and sense the spectrum so as to derive an efficient decision making strategy for the effective utilization of cognitive radio technology.

The strategies discussed in this article can be used to optimize the decision making ability of the network and make it more powerful and efficient while solving dynamic decision problem which is main area of concern to implement Cognitive Radio Technology. An important venue for further research is the interplay of spectrum sensing and higher-layer functionalities to enhance the end user's perceived QoS

5. REFERENCES

[1] Milind M.Buddhikot, Paul Kolodzy , Scott Miller, Kevin Ryans , Jason Evans "DIMSUMNet : New Directions in wireless networking using coordinated dynamic spectrum access", IEEE, WoWMoM05 , June 2005.

[2] I.F.Akyyilidiz , W.Y Lee , M.C.Varun , M.Shantidev "Next Generation/Dynamic Spectrum Access/ Cognitive Radios wireless Netwroks : A Survey" A Computer Network Journal(Elsevier) vol 50, Issue 13 , p2127-2159 , 2006

[3] Chunyi Peng · Haitao Zheng · Ben Y. Zhao K. "Utilization and fairness in spectrum assignment for opportunistic spectrum access" Springer , May 2006 , Pg no 555-576.

[4] S.Gerihofer , L.Tong, B.M.Sadler " Cognitive Medium Access : constraining interference based on Experimental models" IEEE journal , Selected Area communication , vol 26 No.1 , pg 95-105 , 2007.

[5] Raul Etkin, Abhay Parekh, and David Tse "Spectrum Sharing for Unlicensed Bands" IEEE J.Sel area Communication. Vol.25 No.3 pp517-528 , April 2007.

[6] Q.Zha , L.tong , A.Swami , Y.Chen , " Decentralized Cognitive MAC for opportunistic Spectrum Access in Ad Hoc Networks : A POMDP Framework " IEEE J , Selected area , communication Vol.25 No.3 , p589-600 , 2007.

[7] Kyounghwan Lee and Aylin Yener "Throughput Enhancing Cooperative Spectrum Sensing Strategies for Cognitive Radios" IEEE Proc of ACSSC , p2045-2049 , Nov 2007.

[8] I.F.Akyyilidiz , w.Y Lee , K.R.chowdhary " CRAHNs: cognitive Radio ad Hoc Networks" . Journal of Ad Hoc Networks , p 810-836 , Science Direct, 2009.

[9] Ayman A. El-Saleh, Mahamod Ismail, Mohd. A. M. Ali, and Ahmed N. H. Alnuaimy "Capacity Optimization for Local and Cooperative Spectrum Sensing in Cognitive Radio Networks" World Academy of Science, Engineering and Technology 50 2009.

[10] Power Control in Cognitive Radio Systems Based on Spectrum Sensing Side Information Karama Hamdi, Wei Zhang, and Khaled Ben Letaief, Fellow, IEEE.

[11] A.A.El-Saleh, M.Ismail, M.A.M. Ali, A.N.H. Alnuaimy. "Capacity Optimization for Local and Cooperative Spectrum Sensing in Cognitive Radio Networks". Intern. Journal of Electronics, Circuits and Systems, Vol.3, 2009.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 666 - 670, 2013, ISSN: 2319–8656

Skin Tone Based Secret Data Hiding in Images Using Wavelet Transform

Vemula Harini
Department of ECE
Vardhaman college of Engineering
Hyderabad, India

Jeevan Kishor G.
Department of ECE
Vardhaman College of Engineering
Hyderabad, India

Abstract: Biometric feature accustomed implement steganography is skin tone region of pictures. Here secret information is embedded inside skin region of image that may offer a wonderful secure location for information activity. For this skin tone detection is performed exploitation HSV (Hue, Saturation and Value) colour area to boot secret information embedding is performed exploitation frequency domain approach – DWT (Discrete rippling Transform), DWT outperforms than DCT (Discrete cos Transform). Secret data is hidden in one in all the high frequency sub-band of DWT by tracing skin pixels therein sub-band. Four bit secret key and cropping results into associate degree increased securitythan activity information without cropping i.e. in whole image therefore cropped region works as a key at secret writing aspect. This study shows that by adopting an object orienting steganography mechanism, in the sense that, we have a tendency to track skin tone objects in image, we get a higher security

Keywords: Biometrics; Skin tone detection; DWT; DCT; Cropping; secret key; Security; PSNR.

1. INTRODUCTION

In this extremely digitalized world, the web is associate degree important role for knowledge transmission and sharing. However, since it's a worldwide and advertised medium, some confidential knowledge may well be taken, copied, modified, or destroyed by associate degree inadvertent observer. Therefore, security problems become an important issue.

Encoding could be a well know procedure for secured knowledge transmission [2]. Frequently used encoding ways embrace RSA, DES (Data encoding standard). Though encoding achieves certain security effects, they create the key messages unreadable and unnatural or unmeaning.

These unnatural messages typically attract some inadvertent observers' attention. This can be the rationale a replacement security approach known as "steganography" arises. As associate degree example, the quilt text [3]: "I'm feeling extremely stuffy. Emily's medication wasn't sturdy enough while not another medicinal drug." Hides the sentence "Meet American state at nine" If the reader retains the second letter of every word in sequence. In steganography secret message is that the knowledge that the sender desires to stay confidential and may be text, images, audio, video, or the other knowledge which will be represented by a stream of bits. the quilt or host is that the medium during which the message is embedded and serves to hide the presence of the message. The message embedding technique is powerfully keen about the structure of the quilt, and during this paper covers and secret messages are restricted to being digital pictures. The cover-image with the key knowledge embedded is named the "Stego-Image". The Stego-Image should correspond the quilt image underneath casual review and analysis. Additionally, for higher security necessities, we can write in code the message knowledge before embedding them in the cover-image to supply additional protection [4]. For this the encoder typically employs a stego-key that ensures that only recipients WHO apprehend the corresponding secret writing key will be ready to extract the message from a stego-image. For proposed methodology cowl image is cropped interactively and that cropped region works as a key at secret writing facet yielding improved security. There are 2 things that require to be thought-about whereas designing the steganographic system: Invisibility: Human eyes cannot distinguish the distinction between original and stego image. Capacity: The a lot of knowledge a picture will carry better it's. But giant embedded knowledge might degrade image quality considerably. Rest of the paper is organized as follows. Section II presents literature survey and theoretical background. In section III projected methodology is delineated intimately with skin tone detection, DWT, embedding and extraction procedure step by step. Section IV incontestable the experimental results. Finally conclusions ar provided in section V.

2. LITERATURE SURVAY

2.1 Steganography in abstraction Domain

This is a simplest steganographic technique that embeds the bits of secret message directly into the smallest amount significant bit (LSB) plane of the duvet image. During a gray level image, each component consists of eight bits. The essential conception of LSB substitution is to implant the confidential information at the rightmost bits (bits with the littlest weighting) in order that the 2010 International Conference on information Storage and information engineering embedding procedure doesn't have an effect on the initial component worth greatly [5]. The mathematical illustration for LSB is: xi' = xi – xi mod 2k + mi (1) In equation (1), xi' represents the i th component worth of the

stego-image and xi represents that of the initial cover image. mi represents the decimal worth of the i th block in the confidential information. The quantity of LSBs to be substituted is k. The extraction method is to repeat the k-rightmost bits directly. Mathematically the extracted message is represented as: mi =xi mod 2k (2) Hence, a straightforward permutation of the extracted mi offers North American country the original confidential information [6]. This technique is simple and straightforward however this has low ability in-tuned some signal processing or noises. And secret information may be simply purloined by extracting whole LSB plane.

2.2 Steganography In Frequancy

Robustness of steganography may be improved if properties of the duvet image can be exploited. For example it's usually desirable to cover message in strident regions instead of drum sander regions as degradation in smoother regions is a lot of noticeable to human HVS (Human Visual System). Taking these aspects into thought working in frequency domain becomes a lot of engaging. Here, sender transforms the duvet image into frequency domain coefficients before embedding secret messages in it [7]. Totally different sub-bands of frequency domain coefficients give vital data regarding wherever very important and non very important pixels of image resides. These ways ar a lot of complicated and slower than spatial domain methods; but they're more secure and tolerant to noises. Frequency domain transformation may be applied either in DCT or DWT.

2.3 Adgustive Steganography

Adaptive steganography is special case of 2 former methods. It's conjointly referred to as "Statistics aware embedding" [8] and "Masking" [4]. This technique takes applied mathematics world features of the image before making an attempt to plant secret knowledge in DCT or DWT coefficients. The statistics can dictate where to form changes.

3. PROPOSED METHOD

Proposed methodology introduces a brand new methodology of embedding secret information inside skin because it isn't that a lot of sensitive to HVS (Human Visual System) [1].This takes advantage of statistics options like skin tone, instead of embedding information anyplace in image, information are going to be embedded in selected regions. Summary of methodology is concisely introduced as follows. Initially skin tone detection is performed on input image mistreatment HSV (Hue, saturation, value) colour house. Second cowl image is remodelled in frequency domain. This can be performed by applying Haar-DWT, the only DWT on image resulting in four subbands. Then payload (number of bits within which we are able to hide data) is calculated. Before embedding process we have to generate secret key finally secret information embedding is performed in one among the high frequency sub-band by tracing skin pixels in this band. Before activity all steps cropping on input image is performed then in barely cropped region embedding is completed, not in whole image. Cropping results into a lot of security than while not cropping. Since cropped region and secret key works as two keys at secret writing facet. Here embedding method affects solely sure Regions of Interest (ROI) instead of the complete image. Therefore utilizing objects within pictures are often a lot of advantageous. This can be additionally referred to as Object homeward steganography [1]. Next sub-sections briefly introduce skin tone detection and DWT. A. skin colour Tone Detection. A skin detector generally transforms a given element into an acceptable color house then uses a skin classifier to label the element whether or not it's a skin or a non-skin element. A skin classifier defines a choice boundary of the skin colour category in the color house. Though this can be a simple method has tested quite difficult. Therefore, important challenges in skin detection area unit to represent the colour in an exceedingly way that's invariant or a minimum of insensitive to changes in illumination.[9] and Another challenge comes from the actual fact that many objects within the world might need skin-tone colors. This causes any skin detector to own a lot of false detection within the background if the surroundings aren't controlled [10]. The simplest thanks to decide whether or not a element is skin color or not is to expressly outline a boundary. RGB matrix of the given color image are often born-again into totally different color areas to yield distinguishable regions of skin or close to skin tone. There exists many color areas.Mainely 2 kinds of color areas area unit exploited within the literature of biometrics that area unit HSV (Hue, Saturation and Value) and YCbCr (Yellow, Chromatic Blue, Chromatic red) areas It is by experimentation found and in theory tested that the distribution of human skin colour perpetually resides in an exceedingly certain vary inside those 2 color areas [1]. Color house used for skin detection during this work is HSV.

$S_{min=}0.10$, $S_{max}=0.68$, $H_{min}=0^0$, $H_{max}=25^0$

3.1 Separate ripple remodel(Dwt)

This is another frequency domain during which steganography is enforced. DCT is calculated on blocks of freelance pixels, a secret writing error causes` discontinuity between blocks leading to annoying block artefact. This downside of DCT is eliminated victimization DWT. DWT applies on entire image. DWT offers higher energy compaction than DCT with none block whole. DWT splits element into varied frequency bands known as sub bands referred to as

LL – Horizontally and vertically low pass

LH – Horizontally low pass and vertically high pass

HL - Horizontally high pass and vertically low pass

HH - Horizontally and vertically high pass

Since Human eyes square measure way more sensitive to the low frequency half (LL sub band) we will hide secret message in other 3 elements while not creating any alteration in LL sub band [12]. As alternative 3 sub-bands square measure high frequency sub-band they contain insignificant information. Concealment secret information in these sub-bands doesn't

degrade image quality that abundant. DWT utilized in this work is Haar-DWT, the best DWT.

3.2 Embedding Model

Suppose C is original 24-bit color cowl image of M×N Size. It's denoted as: C= 1 ≤ i ≤ M, 1 ≤ j ≤ N, xij, yij, zij Є {0,1,..,255}

Let size of cropped image is Mc×Nc wherever Mc≤M and Nc≤N and Mc=Nc. i.e. Cropped region should be precise sq. as we've to use DWT anon this region. Let S is secret knowledge. Here secret knowledge thought-about is binary image of size a×b. Fig. one represents flow sheet of embedding process.

1) Step 1: Once image is loaded, apply skin tone detection on cowl image. this can turn out mask image that contains skin and non skin pixels.

2) Step 2: raise user to perform cropping interactively on mask image (Mc×Nc). when this original image is additionally cropped of same space. Cropped space should be in a particular square type as we've got to perform DWT later and cropped area ought to contain skin region like face, hand etc since we will hide knowledge in skin pixels of 1 of the sub-band of DWT. Here cropping is performed for security reasons. Cropped parallelogram can act as key at receiving aspect. If it knows then solely knowledge retrieval is feasible. Snooper may attempt to perform DWT on whole image; in such a case attack can fail as we have a tendency to square measure applying DWT on specific cropped region solely.

3) Step 3: Apply DWT to solely cropped space (Mc×Nc) not whole image (M×N). This yields four sub-bands denoted as HLL, HHL, HLH, and HHH. (All four sub-band ar of same size of Mc/2, Nc/2). Payload of image to carry secret knowledge is determined supported no. of skin pixels gift in one amongst high frequency sub-band during which knowledge are going to be hidden.

4) Step 4: Perform embedding of secret knowledge in one amongst sub-band that we have a tendency to obtained earlier by tracing skin pixels in that sub-band. apart from the LL, low frequency sub-band any high frequency sub-band will be hand-picked for embedding as LL sub-band contains important data. Embedding in LL sub-band affects image quality greatly. We have chosen high frequency HH sub-band. Whereas embedding, secret knowledge won't be embedded all told pixels of DWT sub band but to solely those pixels that ar skin pixels. So here skin pixels ar copied victimization skin mask detected earlier and secret knowledge is embedded. Embedding is performed in G-plane and B-plane however strictly not in R-plane as contribution of R plane in coloring is over G or B plane. therefore if we have a tendency to ar modifying R plane constituent values, decoder facet doesn't retrieve knowledge in the slightest degree as skin detection at decoder facet offers different mask than encoder facet. Embedding is finished as per raster-scan order (as shown in Fig.2) that embeds secret knowledge constant by constant in selected sub-band [6], if constant is skin constituent

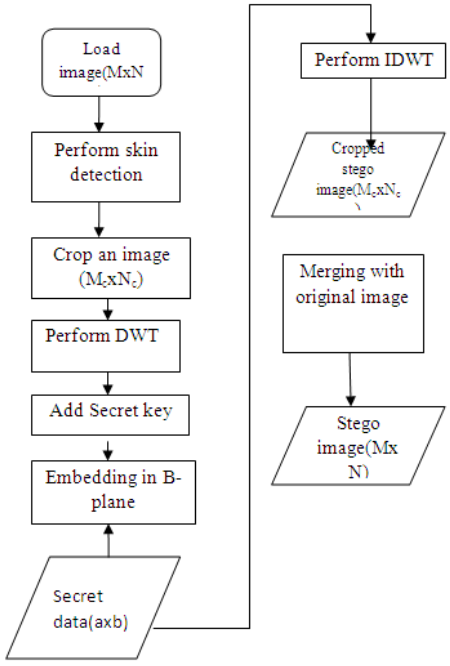


Figure 1: Flow chart for embedding process

5) Step 5: Perform IDWT to mix four sub-bands. 6) Step 6: A cropped stego image of size Mc×Nc is obtained in higher than step (step 5). this could be the same as original image when visual scrutiny however at this stage it's of size Mc× North Carolina, thus we'd like to merge the cropped stego image with original image to urge the stego image of size M×N. To perform merging we have a tendency to need coefficients of 1st and last pixels of cropped space in original image so r calculated. Thus a stego image is prepared for quality analysis.

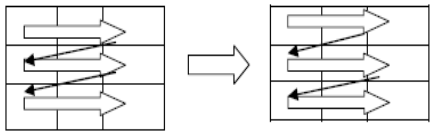


Figure2: Raster scans order

3.3 Extraction Method

Secret knowledge extraction is explained as follows: 24 bit color stego image of size M×N is input to extraction process. We have a tendency to should want price of cropped space to retrieve data.

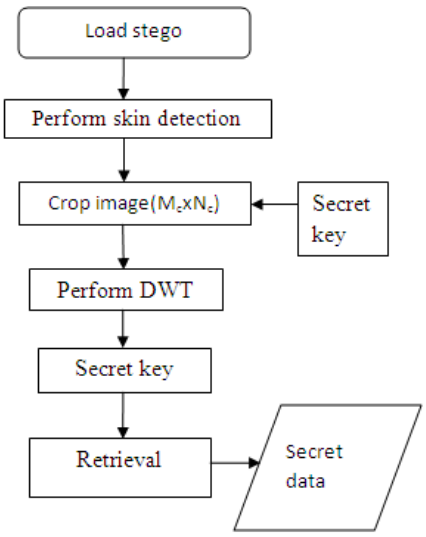


Figure 3. Flow sheet of Extraction method

Suppose four bit secret key keep in 'D' and cropped space price is keep in 'rect' variable that is same as in encoder. Thus this 'rect' and 'D'can act as two keys at decoder facet. All steps of Decoder area unit opposite to Encoder. Care should be taken to crop same size of sq. as per Encoder. By tracing skin pixels in HHH sub-band of DWT secret knowledge is retrieved. Extraction procedure is pictured using flow sheet in Fig. 3

4. SIMULATION

In this section we tend to demonstrate simulation results for proposed theme. This are enforced victimization MATLAB 7.5. A twenty four bit color image is utilized as cover-image of size 256×256, shown in Fig. 4. Fig. five shows sample secret image cowl} within cover image.

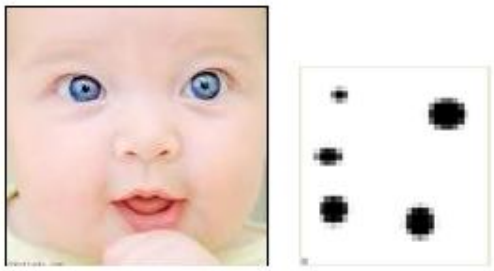


Figure4: Cover image
Figure5: Image to hide

The secret message S is binary image of size 32×32. We use Peak signal to noise quantitative relation (PSNR) to gauge quality of stego image when embedding the key message. The performance in terms of capability and PSNR (in dB) is demonstrated for the strategy within the following subsections. PSNR is outlined as per equivalent weight.3 and Eq.4.

$$PSNR = 10\log10\ (255^2/MSE) \qquad (3)$$

xij and yij represents picture element values of original cowl image and stego image severally. The calculated PSNR sometimes adopts decibel worth for quality judgement, the larger PSNR is, higher the image quality (which means that there's somewhat difference between cowl image and stego image).On the contrary smaller decibel worth means that there's a a lot of distortion. PSNR values falling below 30dB indicate fairly a coffee quality. However, top quality strives for 40dB or a lot.

4.1 Performance of the planned methodology

After embedding secret knowledge in cropped image, resulted cropped stego image is shown in Fig. 6. (Result of step five of embedding process). As this doesn't seem like cowl image merging is performed to get final stego image that's shown in Fig. 9. (Result of step half-dozen of embedding process). For merging co-ordinates of 1st and last pixels of cropped image in original image square measure calculated. once playacting extraction method retrieved image is shown in figure eight. Above methodology uses cropping. Same planned methodology is implemented for while not cropping case. In while not cropping case secret knowledge is hidden in one among the sub-band that is obtained by playacting the DWT on whole image and not only to cropped region. PSNR is calculated for four totally different final stego pictures resulted from a thought of image and three additional sample pictures. This PSNR for various cases is shown in table one. Average PSNR of planned methodology is calculated supported the obtained PSNR. Performing biometric steganography with cropping or without cropping, each square measure having its own benefits and disadvantages. However if methodology is enforced with cropping then it'll guarantee additional security than while not cropping case. As with cropping case we want cropped region at the decoder aspect then solely secret knowledge extraction is feasible. So cropped region works as a key at decoder aspect. For while not cropping methodology interloper could try and perform DWT randomly and may hack secret knowledge from sub-band with trial and error methodology. From the table one it's obvious that PSNR of while not cropping case is over with cropping case. So, this is often trade off that happens if we want additional security.

Fig6:Cropped Skinmask image Figure7:Cropped mask image

Fig8:Mask image Fig9:Stego image after embedding process

TABLE: 1 CAPACITY AND PSNR OF 4 FINAL STEGO IMAGES IN PROPOSED METHOD

Cover image(256x256)	Capacity of cover image		PSNR		Size of logo
	Case A	Case B	Case A	Case B	
Image 1	7173	5294	53.0	50.3	32x32
Image2	1067	1056	51.9	49.5	32x32
Image3	1452	1354	51.1	49.2	32x32
Image4	4850	2572	46.4	45.4	32x32
Average PSNR			50.7	48.5	

Case A- Without Cropping
Case B- With Cropping

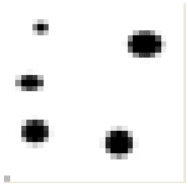


Fig10: Retrieved image

5. CONCLUSION

Digital Steganography may be a fascinating scientific space which falls underneath the umbrella of security systems. In this paper biometric steganography is bestowed that uses skin region of pictures in DWT domain for embedding secret knowledge. By embedding knowledge in mere bound region (here skin region) and not in whole image security is increased. Also image cropping idea introduced, maintains security at respectable level since nobody will extract message while not having worth of cropped region. Options obtained from DWT coefficients square measure utilised for secret knowledge embedding. This additionally will increase the standard of stego as a result of secret messages square measure embedded in high frequency sub-bands that human eyes square measure less sensitive to. in line with simulation results, projected approach provides fine image quality

6. REFERENCES

[1] A. Cheddad, J. Condell, K. Curran and P. Mc Kevitt, “Biometric inspired digital image Steganography”, in: Proceedings of the 15th Annual IEEE International Conference and Workshops on the Engineering of Computer-Based Systems (ECBS’08), Belfast, 2008, pp. 159-168.

[2] Petitcolas, F.A.P.: “Introduction to info Hiding”. In: Katzenbeisser, S and Petitcolas, F.A.P (ed.) (2000) info activity Techniques for Steganography and Digital Watermarking. Norwood: Artech House, INC.

[3] Lin, E. T. and Delp, E. J.:“A Review of information activity in Digital Images”. Retrieved on one.Dec.2006 from laptop Forensics, Cyber crime and Steganography Resources, Digital Watermarking Links and Whitepapers, Apr 1999

[4] Johnson, N. F. and Jajodia, S.: “Exploring Steganography: Seeing the Unseen.” IEEE laptop, thirty one (2): 26-34, Feb 1998.

[5] Fridrich, J., Goljan, M. and Du, R.., (2001). “Reliable Detection ofLSB Steganography in Grayscale and Color pictures.” Proceedings of ACM, session on multimedia system Security and Watermarking, Ottawa, Canada, October 5, 2001, pp. 27- 30.

[6] Po-Yueh subgenus Chen and Hung-Ju designer “A DWT primarily based Approach for Image Steganography”, International Journal of subject area and Engineering, 2006. 4, 3: 275-290

[7] Chang, C. C., Chen, T.S and Chung, L. Z., ”A steganographic methodology based upon JPEG and quantisation table modification,” info Sciences, vol.[4], pp. 123-138(2002).

[8] Provos,N. and Honeyman, P: “Hide and Seek: Associate in Nursing introduction to steganography”. IEEE security and privacy, 01 (3): 32-44,May-June 2003

[9] Abbas Chedda, Joan Condell, Kevin Curran and Paul megacycle Kevitt “A Skin Tone Detection formula for Associate in Nursing adaptational Approach to Steganography” , college of Computing and Intelligent Systems, Faculty of Computing and Engineering, University of Ulster, BT48 7JL, Londonderry, European country, UK,2008

[10] Ahmed E., Crystal M. and Dunxu H.: “Skin Detection-a short Tutorial”, reference of biometry by Springer-Verlag Berlin Heidelberg 2009

[11] Sobottka, K. and Pitas, I.:”Extraction of facial regionsand options using color and form info.” Proc. IEEE International Conference on Image process, pp. 483-486.(1996).

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 671 - 675, 2013, ISSN: 2319–8656

Implementation of Mobile-Cloudlet-Cloud Architecture for Face Recognition in Cloud Computing using Android Mobile

Priya Bhatnagar
R.C.E.W, Jaipur
Rajasthan, India

Abstract: Real time face recognition applications use face recognition algorithms, which extract information such as shape, size, and position of facial features (eyes, nose, ear, lips) and then compare this extracted feature to stored featured database to locate matching images. Face recognition application used in airport security, in passport office, in R.T.O office, in government organization, in development and research department. In this paper we implement a application which recognize a face and minimize computation time using our mobile-cloudlet-cloud architecture named MOCHA. And also develop an algorithm for percent variance. Our experimental result demonstrate that we can communicate anywhere from the world by using jelastic cloud. We will run application on mobile devices using jelastic cloud as the backend server.

Keywords-Face detection, Face recognition, Jelastic cloud, Cloudlet, mobile devices

1. INTRODUCTION

Face recognition applications that automatically identify an individual from captured images or videos are everywhere, for applications such as airport, passport office, R.T.O office, government organization, and research center. Face recognition algorithms analyze images, extract information such as the shape, size and position of the facial features (e.g., eyes, nose, mouth), and then use these extracted features to search a facial database to locate matching images. The rapid adoption of smart phones, net books and tablets in the consumer market, multimedia content is becoming the dominant form of information. These devices are becoming increasingly more powerful, with more processing power, storage, and sensing capabilities. In addition, it is now possible to rent computing, storage, and network resources as needed via cloud computing, in which data is processed and stored remotely at large-scale compute and data centers [1], [2].

This paper describes our work with the design and implementation of the face recognition on the MOCHA (Mobile Cloud Hybrid Architecture) cloud computing platform, which provides a mobile-cloudlet-cloud architecture. And we are using Jelastic cloud server and deploy application so that we can communicate anywhere from the world. And the second goal of this system that we are using GPS coding for checking the communication network and time monitoring and GPS class is monitoring time scheduling for one request and if it cross more than 1sec then send communication error. We have written our own logic where communication happens within 0.2sec approax for one image. And the third goal of this system that We developed our own algorithm which compares the percent matching faces with respect to given face name in the database. It displays %matching face with corresponding face and database id also.

2. LITERATURE SURVEY

With the ubiquity of new information technology and media, more effective and friendly methods for human computer interaction (HCI) are being developed which do not rely on traditional devices such as keyboards, mice, and displays. Furthermore, the ever decreasing price/ performance ratio of computing coupled with recent decreases in video image acquisition cost imply that computer vision systems can be deployed in desktop and embedded systems. The rapidly expanding research in face processing is based on the premise that information about a user's identity, state, and intent can be extracted from images, and that computers can then react accordingly, e.g., by observing a person's facial expression. In the last five years, face and facial expression recognition have attracted much attention though they have been studied for more than 20 years by psychophysicists, neuroscientists, and engineers. Many research demonstrations and commercial applications have been developed from these efforts. A first step of any face processing system is detecting the locations in images where faces are present. However, face detection from a single image is a challenging task because of variability in scale, location, orientation (up-right, rotated), and pose (frontal, profile). Facial expression, occlusion, and lighting conditions also change the overall appearance of faces. Facial expression recognition concerns identifying the affective states (happy, sad, disgusted, etc.) of humans [3]. Evidently, face detection is the first step in any automated system which solves the above problems. It is worth mentioning that many papers use the term "face detection," but the methods and the experimental results only show that a single face is localized in an input image. In this paper, we differentiate face detection from face localization since the latter is a simplified problem of the former. Meanwhile, we focus on face detection methods rather than tracking methods. While numerous methods have been proposed to detect faces in a single image of intensity or color images, we are unaware of any surveys on this particular topic. A survey of early face recognition methods before 1991 was written by Samal and Iyengar . Chellapa et al. wrote a more recent survey on face recognition and some detection methods. Among the face detection methods, the ones based on learning algorithms have attracted much attention recently and have demonstrated excellent results. Since these data driven methods rely heavily on the training sets, we also discuss several databases suitable for this task. A related and important problem is how to evaluate the performance of the proposed detection methods. Many recent face detection papers compare the performance of

several methods, usually in terms of detection and false alarm rates. It is also worth noticing that many metrics have been adopted to evaluate algorithms, such as learning time, execution time, the number of samples required in training, and the ratio between detection rates and false alarms. Evaluation becomes more difficult when researchers use different definitions for detection and false alarm rates. In this paper, detection rate is defined as the ratio between the number of faces correctly detected and the number faces determined by a human. An image region identified as a face by a classifier is considered to be correctly detected if the image region covers more than a certain percentage of a face in the image . In general, detectors can make two types of errors: false negatives in which faces are missed resulting in low detection rates and false positives in which an image region is declared to be face, but it is not. A fair evaluation should take these factors into consideration since one can tune the parameters of one's method to increase the detection rates while also increasing the number of false detections. In this paper, we discuss the benchmarking data sets and the related issues in a fair evaluation. With over 150 reported approaches to face detection, the research in face detection has broader implications for computer vision research on object recognition [5]. Nearly all model-based or appearance-based approaches to 3D object recognition have been limited to rigid objects while attempting to robustly perform identification over a broad range of camera locations and illumination conditions. Face detection can be viewed as a two-class recognition problem in which an image region is classified as being a "face" or "nonface."

Existing system

There exist solutions that provide face recognition in wireless communication using mobile-cloudlet-cloud architecture on cloud. In [1] Tologa Soyata, Rajani muraleedharan and Colin Funai, present a face recognition application on cloud (Amazon web server) using mobile-cloudlet-cloud architecture. And present a algorithm which minimize overall response time. And used SPREAD (Secure protocol for reliable data delivery) scheme to statistically enhance the data confidentiality service in an ad hoc network. SPREAD is based on secret sharing and multi-path routing. Which have some disadvantage like

a) Mobile nodes might be compromised themselves (e.g., nodes be captured in a battle field scenario) and subsequently be used to intercept secret information relayed by them.

b) Data security applied on whole data which can be trapped in between and decrypted the entire data.

c) Data security percent is less.

There are no existing model which utilizes all the three components (mobile, cloudlet, jelastic cloud) working together with specific algorithms, applications and initial result.

3. PROPOSED SYSTEM

In this paper we present a application for face recognition and percent variance in wireless communication using mobile-cloudlet-cloud architecture on jelastic cloud.

With the help of this application user can communicate anywhere from the world.

3.1 Related Work

Android development SDK provides API libraries and developer tools which are necessary to build, test, and debug apps for android. For a new developer, recommends to download the ADT Bundle to quickly start developing apps. It includes the Android SDK mechanism and a version of the Eclipse IDE with built-in ADT(Android Developer Tools).

Eclipse+ADT plug-in

Android SDK tools

Android platform-tools

The latest Android platform

3.2 System architecture

The mobile-cloudlet-cloud architecture broken into three subsystem: 1) mobile devices 2) cloudlet 3) Jelastic cloud

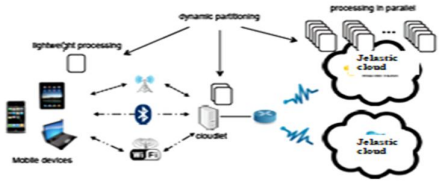


Fig. 1. The MOCHA Architecture: mobile devices interact with the cloudlet and the cloud via multiple connections and use dynamic partitioning to achieve their QoS goals (e.g., latency, cost).

Figure [1] The MOCHA architecture[1]

3.2.1 Mobile device

In our work we use mobile devices, such as Smartphone's. The main task of our mobile device is to capture image and send them to the cloudlet in raw form for processing. After pre processing captured image sent to the cloud to perform real time template matching over a large database located in the cloud.

3.2.2 Cloudlet

The cloudlet is a special-purpose inexpensive compute box with the capability of massively parallel processing MPP. In our approach we can use till 16 cloudlet [4].

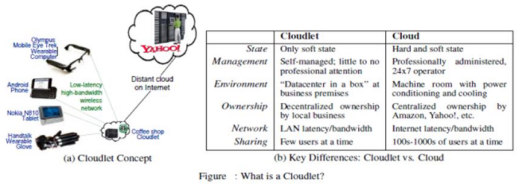


Figure 2. The cloudlet concept

3.2.3 Cloud

Cloud computing provides computing and storage resources in a pay-as-you-go manner. In our architecture, a client program which is running on the cloudlet sends a request to the servers on Jelastic cloud where the actual program runs on virtual instances in parallel and the results are sent back to the mobile device or cloudlet.

3.3 System requirement

3.3.1 Hardware Requirement

- Internet connection required in mobile and laptop

- Android 2.3/4.0 based mobile

3.3.2 Software Requirement

- Jdk1.6
- Android sdk4.0
- Web server : Tomcat 6.0
- IDE : Net Beans 7.0.1

3.4 Face recognition using MOCHA architecture

Face recognition application in wireless communication using mobile-cloudlet-cloud architecture divided in two stages: First is **face detection**: In which we use Viola-Jones and Adaboost algorithm for face detection[1] and percent variance we try all possible rectangle locations, and classify if rectangle contains a face (and only the face) and Select Features i=1..n, Learn thresholds θ_i , Learn weights α_i . Features can be used to represent both the statistically close facial information and sparsely related background data in a sample image. In order to be successful a face detection algorithm must possess two key features, accuracy and speed. There is generally a trade-off between the two. Through the use of a new image representation, termed integral images, Viola and Jones describe a means for fast feature evaluation, and this proves to be an effective means to speed up the classification task of the system. The integral image computes a value at each pixel (x,y) that is the sum of the pixel values above and to the left of (x,y), inclusive. This can quickly be computed in one pass through the image. Let A, B, C, D be the values of the integral image at the corners of a rectangle. Then the sum of original image values within the rectangle can be computed:

$$\text{Sum} = A - B - C + D$$

Only 3 additions are required for any size of rectangle. This is now used in many areas of computer vision.

Adaboost algorithm takes a feature set and a training set of positive and negative images, any number of machine learning approaches could be used to learn a classification function. AdaBoost learning algorithm is used to boost the classification performance of a simple learning algorithm (e.g., it might be used to boost the performance of a simple perceptron). It does this by combining a collection of weak classification functions to form a stronger classifier. In the language of boosting the simple learning algorithm is called a weak learner.

Second stage is **face recognition** where we use Eigen face approach Face Images are projected into a feature space ("Face Space") that best encodes the variation among known face images. The face space is defined by the "eigenfaces", which are the eigenvectors of the set of faces. Acquire the training set and calculate eigenfaces (using PCA projections) which define eigenspace. In eigenface approach we done this steps

- When a new face is encountered, calculate its weight.
- Determine if the image is face.
- If yes, classify the weight pattern as known or unknown.
- (Learning) If the same unknown face is seen several times incorporate it into known faces.

Applications

- AirPort
- Passport Office
- RTO Office
- Government Organization
- Research & Development Center

3.5 Flow diagrams

3.5.1 Face detection/recognition flow diagram

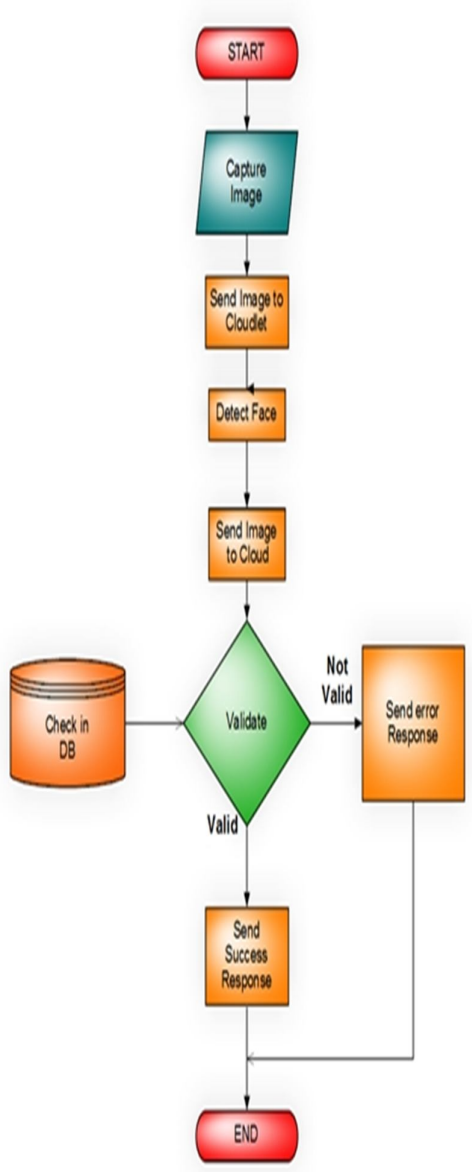


Figure 3. face detection/recognition flow diagram

3.5.2 Flow diagram for cloudlet

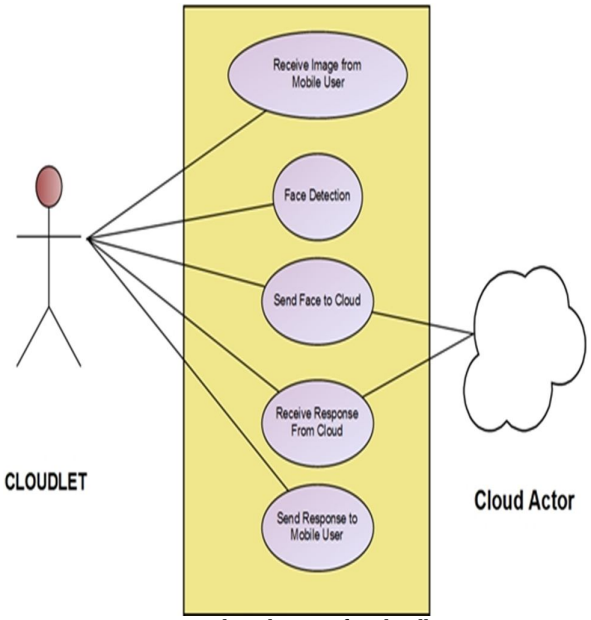


Figure 4. Flow diagram for cloudlet

3.5.3 Flow diagram for cloud

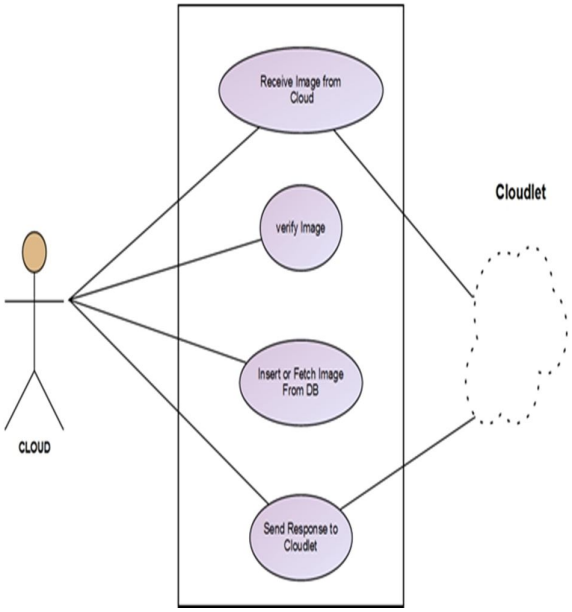


Figure 5. Flow diagram for cloud

3.6 Testing

During software development testing is the major quality control measure employed its basic function is to detect the errors in the software. For this, different levels of testing are used, which performs different tasks on the aim to test different aspects of the system.

This table presents a test case to check the integration of Cloud Vision with Face Recognition in cloud using android phone.

Table 1. Test case to check the integration of cloud vision

Step number	Description	Input	Expected output
1.	Select setting in android application		It should establish the connection with cloud server else display error in communication
2	Enter valid user id and password and click on submit button		Main page should be displayed with different face detection and identification
3	Select face detection and user id		It should identify the face on cloud server and display verify or not
4	Enter valid user id and password and click on submit button		Main page should be displayed with different face detection and identification
5	Select face detection		It interact with cloudlet and detect that the face is existing or not
6	Select face detection		Error in communication if no proper network
7	Select face identification		Error in communication if no proper network
8	Select face identification		It should identify the faces in cloud server and display verify or not

This below table present test functionality of Cloud Server Connection

Table 1. Test case functionality of cloud server

STEP NUMBER	DESCRIPTION	INPUT	EXPECTED OUTPUT
1.	Deploy the application to Jelastic Cloud		it should verify the cloud environment else display the error
2.	Deploy the application to Jelastic Cloud		Cloud server verify the jelastic environment as SAAS and process the request
3.	Deploy the application to Jelastic Cloud		it should verify to port number on which application will be running and deploy the application to server
4.	Select project run as server		Application should start run if all the connection is proper on server

5. FUTURE GOALS

In future more sophisticated synchronization algorithms which permitting cloud to cloud communication has to be extended rather than multiple cloudlet-cloud communication links. And also develop application for 3D face recognition on MOCHA architecture.

6. CONCLUSION

In this paper we presented an approach for face detection which minimizes computation time while achieving high detection accuracy. The approach was used to construct a face detection system which is approximately 15 faster than any previous approach. Preliminary experiments, which will be described elsewhere, show that highly efficient detectors for other objects, such as pedestrians, can also be constructed in this way. New algorithms, representations, and insights where presented which are quite generic and may have broader application in computer vision and image processing. and also present an algorithm for percent variance. We have designed MOCHA to integrate mobile devices, the cloudlet, and using jelastic cloud for better development. In our approach GPS class is monitoring time scheduling for one request and if it cross more than 1sec then send communication error. We have written our own logic where communication happens within 0.2sec approax for one image.

7. REFERENCES

[1] Ming-Hsuan Yang, David Kriegman, and Narendra Ahuja, “Detecting faces in images : A survey,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34–58, 2002.

[2] Paula Viola and Michael J. Jones, “Robust real time face detection, “International Journal of Computer Vision, vol. 57, no. 2, pp. 137–154, 2004.

[3] Matthew Turk and Alex Pentland, “Eigenfaces for recognition,” Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.

[4] Paula Viola and Michael J. Jones, “Robust real time face detection, in Second International Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, Computing, and Sampling, July 2001, pp. 1–25.

[5] Tolga Soyata, Eby G. Friedman, and J. H. Mulligan, “Integration of Clock skew and register delays into a retiming algorithm,” in Proceedings of the International Symposium on Circuits and Systems, May 1993, pp.1483–1486.

[6] Tolga Soyata and Eby G. Friedman, “Retiming with non-zero clock skew, variable register and interconnect delay,” in Proceedings of the IEEE Conference on Computer-Aided Design, Nov 1994, pp. 234–240.

[7] B. Chun and P. Maniatis, “Dynamically Partitioning Applications between Weak Devices and Clouds,” in Proceedings of ACM Mobile Cloud Computing and Services (MCS), June 2010.

[8] P. Stuedi, I. Mohamed, and D. Terry, “Where Store: Location-based Data Storage for Mobile Devices Interacting with the Cloud,” in Proceedings of ACM Mobile Cloud Computing and Services (MCS), June 2010.

[9] A. Miettinen and J. Nurminen, “Energy Efficiency of Mobile Clients in Cloud Computing,” in Proceedings of Usenix Hot Cloud, June 2010.

[10] E. Jung, Y. Wang, L. Prilepov, F. Maker, X. Liu, and V. Akella, “User-Profile-Driven Collaborative Bandwidth Sharing on Mobile Phones,” in Proceedings of ACM Mobile Cloud Computing and Services (MCS), June 2010.

[11] A. Shieh, S. Kandula, A. Greenberg, and C. Kim, “Seawall: Performance

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 676 - 679, 2013, ISSN: 2319–8656

Implementation of Medical Image Fusion Using DWT Process on FPGA

D.Khasim Hussain
P.R.R.M. Engineering College
Shabad, Ranga Reddy, India

C.Laxmikanth Reddy
P.R.R.M. Engineering College
Shabad, Ranga Reddy, India

V.Ashok Kumar
PRRM ENGG COLLEGE
Shabad, Ranga Reddy, India

Abstract: Image fusion is a data fusion technology which keeps images as main research contents. It refers to the techniques that integrate multi-images of the same scene from multiple image sensor data or integrate multi images of the same scene at different times from one image sensor.. Wavelet Transform has good time-frequency characteristics. It was applied successfully in image processing field. Nevertheless, its excellent characteristic in one-dimension can't be extended to two dimensions or multi-dimension simply. Separable wavelet which was spanning by one-dimensional wavelet has limited directivity. The experiments show that the method could extract useful information from source images to fused images so that clear images are obtained. The selection principles about low and high frequency coefficients according to different frequency domain after wavelet. In choosing the low-frequency coefficients, the concept of local area variance was chosen to measuring criteria. In choosing the high frequency coefficients, the window property and local characteristics of pixels were analyzed.

Keywords: Fusion oriented images, FPGA, Application of computer vision, DWT, Wavelet Coefficient maps .

1. INTRODUCTION

The actual fusion process can take place at different levels of information representation [1]: a generic categorization is to consider the different levels as, sorted in ascending order of abstraction: signal, pixel, feature and symbolic level. This site focuses on the so-called pixel level fusion process, where a composite image has to be built of several input images. To date, the result of pixel level image fusion is considered primarily to be presented to the human observer, especially in image sequence fusion (where the input data consists of image sequences). A possible application is the fusion of forward looking infrared (FLIR) and low light visible images (LLTV) obtained by an airborne sensor platform to aid a pilot navigates in poor weather conditions or darkness. In pixel-level image fusion[2], some generic requirements can be imposed on the fusion result. The fusion process should preserve all relevant information of the input imagery in the composite image (pattern conservation) The fusion scheme should not introduce any artifacts or inconsistencies which would distract the human observer or following processing stages .The fusion process should be shift and rotational invariant, i.e. the fusion result should not depend on the location or orientation of an object the input imagery .In case of image sequence fusion arises the additional problem of temporal stability and consistency of the fused image sequence.

The main target in these techniques is to produce an effective representation of the combined multispectral image data, i.e., an application-oriented visualization in a reduced data set [3]–[8].The human visual system is primarily sensitive to moving light stimuli, so moving artifacts or time depended contrast changes introduced by the fusion process are highly distracting to the human observer. So, in case of image sequence fusion the two additional requirements apply. Temporal stability: The fused image sequence should be temporal stable, i.e. gray level changes in the fused sequence must only be caused by gray level changes in the input sequences, they must not be introduced by the fusion scheme itself. Temporal consistency: Gray level changes occurring in the input sequences must be present in the fused sequence without any delay or contrast change.

2. IMAGE FUSION [1]

2.1 Image Fusion Process

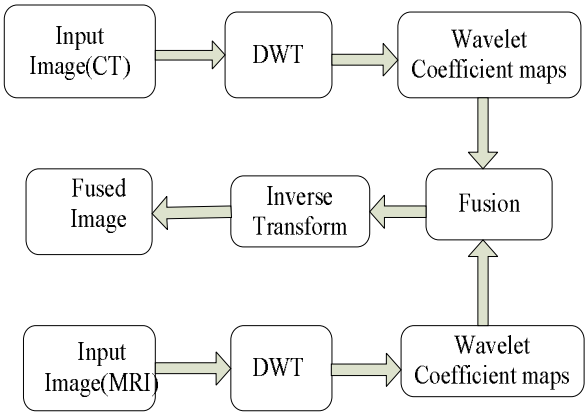


Figure 1: Block Diagram Image Fusion Process

When constructing each wavelet coefficient for the fused image. We will have to determine which source image describes this coefficient better. This information will be kept in the fusion decision map. The fusion decision map has the same size as the original image. Each value is the index of the source image which may be more informative on the corresponding wavelet coefficient. Thus, we will actually make decision on each coefficient. There are two frequently used methods in the previous research. In order to make the decision on one of the coefficients of the fused image, one way is to consider the corresponding coefficients in the source images as illustrated by the red pixels. This is called pixel-based fusion rule. The other way is to consider not only the corresponding coefficients, but also their close neighbors, say a 3x3 or 5x5 windows, as illustrated by the blue and shadowing pixels. This is called window-based fusion rules. This method considered the fact that there usually has high correlation among neighboring pixels.

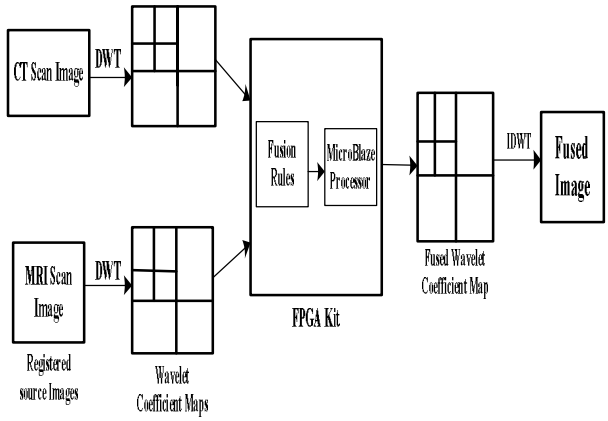


Figure 2 Fusion process on FPGA Kit

In our research, we think objects carry the information of interest, each pixel or small neighboring pixels are just one part of an object. Thus, we proposed a region-based fusion scheme. When make the decision on each coefficient, we consider not only the corresponding coefficients and their closing neighborhood, but also the regions the coefficients are in. We think the regions represent the objects of interest. We will provide more details of the scheme in the following.

2.2 Wavelet Transform

Wavelets are mathematical functions defined over a finite interval and having an average value of zero that transform data into different frequency components, representing each component with a resolution matched to its scale.

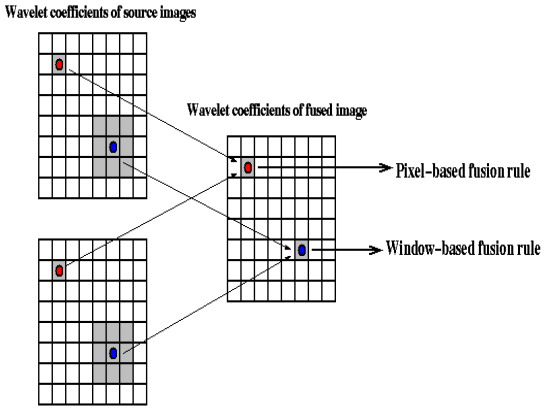


Figure 3: Wavelet Coefficient Representation

2.3 2-D Transform Hierarchy

The 1-D wavelet transform can be extended to a two-dimensional (2-D) wavelet transform using separable wavelet filters. With separable filters the 2-D transform can be computed by applying a 1-D transform to all the rows of the input, and then repeating on all of the columns.

2.4 Discrete Wavelet Transform

Calculating wavelet coefficients at every possible scale is a fair amount of work, and it generates an awful lot of data. If the scales and positions are chosen based on powers of two, the so-called dyadic scales and positions, then calculating wavelet coefficients are efficient and just as accurate. This is obtained from discrete wavelet transform (DWT).

The 2-D subband decomposition is just an extension of 1-D subband decomposition. The entire process is carried out by executing 1-D subband decomposition twice, first in one direction (horizontal), then in the orthogonal (vertical) direction. For example, the low-pass subbands (Li) resulting from the horizontal direction is further decomposed in the vertical direction, leading to LLi and LHi subbands.

Similarly, the high pass subband (Hi) is further decomposed into HLi and HHi. After one level of transform, the image can be further decomposed by applying the 2-D subband decomposition to the existing LLi subband. This iterative process results in multiple "transform levels". In Fig. 4. the first level of transform results in LH1, HL1, and HH1, in addition to LL1, which is further decomposed into LH2, HL2, HH2, LL2 at the second level, and the information of LL2 is used for the third level transform. The subband LLi is a low-resolution subband and high-pass subbands LHi, HLi, HHi are horizontal, vertical, and diagonal subband respectively since they represent the horizontal, vertical, and diagonal residual information of the original image.

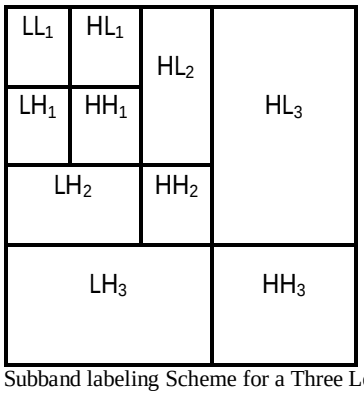


Figure 4: Subband labeling Scheme for a Three Level, 2-D Wavelet Transform

To obtain a two-dimensional wavelet transform, the one-dimensional transform is applied first along the rows and then along the columns to produce four subbands: low-resolution, horizontal, vertical, and diagonal. (The vertical subband is created by applying a horizontal high-pass, which yields vertical edges.) At each level, the wavelet transform can be reapplied to the low-resolution subband to further decorrelate the image. Fig. 5 illustrates the image decomposition, defining level and subband conventions. The final configuration contains a small low-resolution subband. In addition to the various transform levels, the phrase level 0 is used to refer to the original image data. When the user requests zero levels of transform, the original image data (level 0) is treated as a low-pass band and processing follows its natural flow.

Low Resolution Subband

	4	3	Level 2	Level 1

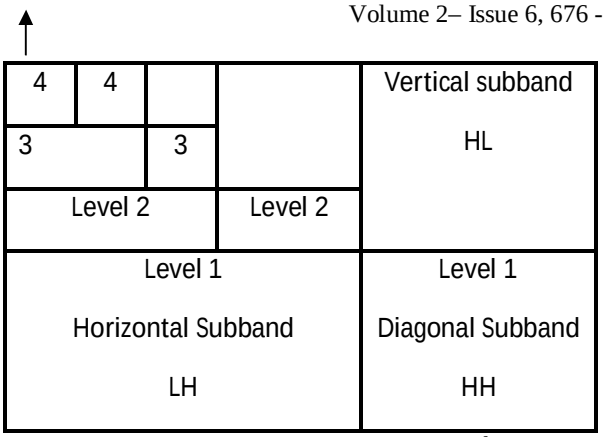
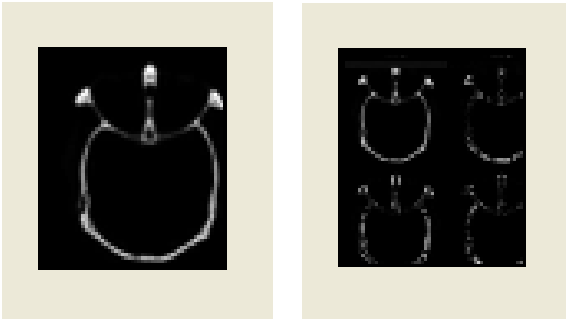


Figure.5 Image Decomposition Using Wavelets

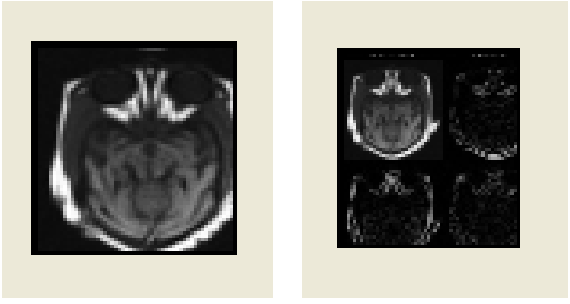
Wavelet transform is first performed on each source images, then a fusion decision map is generated based on a set of fusion rules. The fused wavelet coefficient map can be constructed from the wavelet coefficients of the source images according to the fusion decision map. Finally the fused image is obtained by performing the inverse wavelet transform.From the above diagram, we can see that the fusion rules are playing a very important role during the fusion process.

3. RESULTS

The image fusion process follows by considering System C coding for DWT method to implement fused image using FPGA .



(a) **(b)**



(c) **(d)**

(e)

Figure 6 Fused Images of CT and MRI Scan

Above figure shows the fused and DWT levels of MRI Scan Images are taken from the Xilinx platform studio tool. Registered Source images like CT scan (fig 6.a) and MRI scan(fig 6.c) images are taken for fusing the both resolution images by considering pixel coefficients of the images by using DWT. After the DWT process the resulted image for CT scan and MRI scan are shown in figure 6 (b),(d). By considering the low resolution subband images in fused CT scan and MRI scan images fusion will be done using DWT method. Inverse DWT is applied for the image to get final fused image . It is implemented on Spartan -3 kit by using of Xilinx is tool. It shown in figure 7.

Figure7: Spartan-3 kit

A new approach to 3-D image fusion using wavelet transforms. Several known 2-D DWT fusion schemes have been extended to handle 3-D images have been proposed.

Wavelet transform fusion diagrams have been introduced as a convenient tool to visually describe different image fusion schemes. A very important advantage of using 3-D DWT image fusion over alternative image fusion algorithms is that it may be combined with other 3-D image processing algorithms working in the wavelet domain

4. CONCLUSION

In order to evaluate the results and compare these methods two quantitative assessment criteria Information Entropy and Root Mean Square Error were employed. Experimental results indicated that there are no considerable differences between these two methods in performance. The fusions have been implemented for medical images and remote sensing images. It is hoped that the techniques can be extended for colored images and for fusion of multiple sensor images with memory constraints.

5. REFERENCES

[1] A. Goshtaby and S. Nikolov, "Image fusion: Advances in the state of the art," *Inf. Fusion*, vol. 8, no. 2, pp. 114–118, Apr. 2007.

[2] V. Tsagaris and V. Anastassopoulos, "Multispectral image fusion for improved RGB representation based on perceptual attributes," *Int. J. Remote Sens.*, vol. 26, no. 15, pp. 3241–3254, Aug. 2005.

[3] J. Tyo, A. Konsolakis, D. Diersen, and R. C. Olsen, "Principalcomponents- based display strategy for spectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 708–718, Mar. 2003.

[4] W. Zhang and J. Kang, "QuickBird panchromatic and multi-spectral image fusion using wavelet packet transform," in *Lecture Notes in Control and Information Sciences*, vol. 344. Berlin, Germany: Springer-Verlag, 2006, pp. 976–981.

[5] V. Shah, N. Younan, and R. King, "An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008.

[6] K. Kotwal and S. Chaudhuri, "Visualization of hyperspectral images using bilateral filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2308–2316, May 2010.

[7] Q. Du, N. Raksuntorn, S. Cai, and R. J. Moorhead, "Color display for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1858–1866, Jun. 2008.

[8] S. Cai, Q. Du, and R. J. Moorhead, "Feature-driven multilayer visualization for remotely sensed hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3471–3481, Sep. 2010.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 686 - 689, 2013, ISSN: 2319–8656

Presenting a New Method for Detection and Prevention of Single Black Holes Attack in AODV Protocol in Wireless Ad Hoc Network

Iman Zangeneh
Department of Computer Mathematics,
Science and Research Branch,
Islamic Azad University,
Mahshahr, Iran.

Sedigheh Navaezadeh
Sama Technical and Vocational Training College
Islamic Azad University
Mahshahr Branch,
Mahshahr, Iran

Abolfazl Jafari
Department of Computer,
Science and Research Branch,
Islamic Azad University,
Sari, Iran.

Abstract: There is no infrastructure in wireless ad hoc networks, and nodes independently manage the networks. Therefore, the connection between nodes is provided by the nodes themselves, and these nodes act as a router. In this case, they use routing protocol such as AODV. In order to provide the connections, nodes exchange data and control packages by trusting to each other. Since these networks have unique and special characteristics, they face with too much attack. One of these attacks is black hole attack in which destructive node attracts the network traffic, and destroys the packages. In this paper, black hole attack in AODV routing protocol has been investigated, and some solutions have been suggested. Simulation results indicate that, in proposed method, the rate of package delivery has been considerably increased in comparison with AODV.
Keywords: wireless ad hoc network, simulation, black hole attack, intrusion detection system, NS2 simulator.

1. INTRODUCTION

Wireless ad hoc networks are composed of independent nodes managing the networks without any infrastructures. In wireless ad hoc networks, topology is dynamic, and nodes can freely enter the networks or leave it. In these networks, the connections between the nodes is wireless. Due to this advantage, these networks are available in places where establishing wireless networks is not possible. Wireless ad hoc networks can be used in impassable or mountainous areas and battlefields where the soldiers can communicate with each other. Also, they can be used in natural events such as flood or earthquake. Since there is no fixed infrastructure in these networks, nodes act as a host and router [1; 2; 3], and they use different routing protocols in routing process such as AODV [4; 5]. Finding the route and sending the packages are performed in the network by the nodes themselves on the basis of mutual trusting. Due to the characteristics of wireless ad hoc networks such as lack of fixed infrastructure and the trust of nodes to each other, these networks are exposed to attacks. One of these attacks is black hole attack. In this attack, destructive node uses the vulnerability of routing packages in on-demand protocols like AODV, and attracts the network traffic. Finally, it destroys all packages. In AODV routing protocol, when the source node demands a route toward destination, middle nodes are responsible for detecting the route. In order to do this task, they send route demand packages to neighbors. This process continues until destination node or the node that has found a new route toward destination receives the package of route demand [6; 7; 8]. Destructive node does not do this work; rather, it immediately and falsely responses to the source node through which there is a new route to destination. After receiving this response, the source node sends data packages to black hole. Then, the black hole attracts and receives data packages, and destroys them. In this paper, a method has been suggested, and in this method, black hole node is identified in AODV routing, and then it is removed from routing process. In order to identify black hole attack, fidelity level is allocated to each node. Afterwards, fidelity levels of nodes are stored in a table called fidelity table. The network nodes store this table. This table is updated by the source node, and then it is distributed. In this way, other nodes can update their own fidelity table. Through using this table, collecting the responses in response table and changing the way of selecting AODV protocol, black hole node is identified in proposed method, and then it is removed from routing. Simulation results show considerable improvement of package delivery rate in comparison with AODV.

2. AODV ROUTING PROTOCOL

One of the protocols used in wireless ad hoc network for finding the route is routing protocol on the basis of AODV demand. In this protocol, all nodes cooperate with each other to find and discover the route through control messages such as route request (RREQ), route response (RREP) and route error (RERR). The characteristics of AODV are less overhead and less usage of band width due to small size of these packages. In order to be sure that there are no turns in finding and detecting the route, this protocol uses sequence number of destination for each destination entry. The procedure of finding the route by AODV is as follows: when the source node sends data to destination node, it distributes RREQ message, and then neighbor nodes in the source node receive this message. Each middle node investigates its own routing table by receiving RREQ. If there is no new route to destination node, then RREQ is sent to neighbors. This process continues until destination node or middle node that follows a new route toward destination receives RREQ. When RREQ is received by this node, RREQ message is created and sent inversely to source direction. When RREQ message moves in the network, the number of its steps increases by passing through each node. The node sending RREP expands its own routing table according to the number of steps, and then it updates the sequence number of destination node. Each RREQ has an indicator. When a node receives two RREQs with the same indicator, the newer RREQ is removed. In there are two routes toward receiving destination, then the route having maximum sequence number is selected. If sequence

numbers are same, the message with minimum number of steps is selected [9; 10].

3. BLACK HOLE ATTACK

The node performing black hole attack waits until one RREQ is received from neighbor nodes. After receiving RREQ, it, immediately and without investigating its own routing table, responses to the node sending RREQ by sending a false RREP. Black hole locates maximum sequence number and minimum steps in its own RREP. In this way, it deceives the node requesting the route. When the node sending RREQ receives RREP, it assumes that it has discovered the best route; therefore, it sends data packages to black hole. Black hole destroys all packages. Since black hole does not investigate its own routing table, it responses to the node requesting the route before other nodes. If black hole can attract the network traffic, then it provides prevention of service. There are two kinds of black hole attacks; namely, single black hole and cooperative black hole. In single black hole, there is a black hole node in the network, while in cooperative black hole, there is more than one black hole, and they cooperate with each other [11]. In this paper, single black hole attack is investigated.

4.LITERATURE REVIEW

According to [12], in order to discover single black hole attack, middle node sending RREP should introduce the node of next step. Source node sends a frequent request (FREQ) to the node of next step, and asks about responding node. If the node of next step is not destructive, then the accuracy of responding node can be identified. A method has been proposed by [13]. In this method, a request package of route confirmation is sent to next step of the respondent node. In the next step, by receiving request package of route confirmation, it tries to find out whether its routing table has a route toward destination or not. If there is any route, then it sends response package of route confirmation (CREE) to source node. This package involves route information. In this way, it identifies the accuracy of responding node. Overhead of this method is high due to high operations. As suggested by [14], by using timer, the source node waits until receiving several RREPs. Afterwards, RREPs are investigated. RREPs involving common nodes and steps are valid and reliable, and others are unreliable. Unreliable RREPs are not taken into account. According to [15], subversion of responding node can be identified through using survey packages of neighbors. A method has been proposed by [16] to identify cooperative black hole attack. In this method, data routing table (DRI), frequent request package (FREQ) and frequent response package (FREP) are used. There is an entry for each neighbor in DRI kept by node, and this indicates that whether node has been sent by the neighbor or not. The neighbors through which data has been sent are reliable. A method has been presented by [17] to identify black hole in DSR algorithm. In this method, the concepts of watchdog and route evaluator have been added to DSR algorithm. The duty of watchdog is to identify misbehavior of the nodes and to investigate whether the node has delivered packages to next step or not. This method is not efficient in collision conditions. In these conditions, it has the power of less transferring and removing lots of packages.

5. PROPOSED METHOD

In our method, AODV protocol is changed in a way that it prevents black hole attack. In this method, fidelity level is allocated to participating nodes, and this is the basis of nodes' reliability. This fidelity level is changed by the source node on the basis of loyal participation of the nodes. The source node sends RREQ package to neighbors. Afterwards, by using timer, it waits for some seconds to collect RREPs. These responses are collected until the end of timer time, and then they are stored in a table called table of response storage. Equation (1) is used to select the response.

RF=sequence Number*Node's Fidelity Level (1)

Where Sequence Number is available number of sequence in RREP, and Node's Fidelity Level is fidelity level of respondent node. This equation is calculated for each received response. Finally, the response whose RF (RREP's Fidelity) is higher than others is selected. If RF is same in two or more RREPs, then the response having minimum number of steps is selected. After selecting RREP, source node begins to send data packages. When data package is received, destination node sends ACK to the source node, and in this way, fidelity level of responding node increases. By receiving ACK, the source node can increase fidelity level of responding node because it has been proved that it is secure and reliable. Of the source node does not receive ACK after the end of timer time, then it reduces fidelity level of responding node in terms of identifying black hole attack. Fidelity tables are periodically exchanged among participating nodes. Since black hole nodes do not send packages, the source node does not receive ACK from destination node, so the source node immediately reduces fidelity level and does not use the received responses.

Figure 1, 2 and 3 show the way of collecting responses, general distribution of fidelity table and route selection in proposed
method.

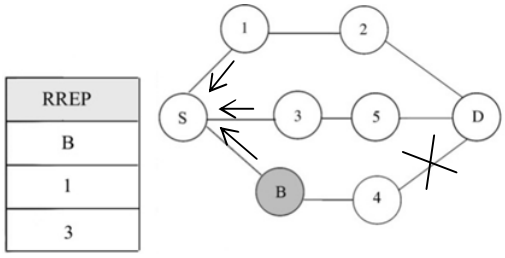


Figure 1: Collecting Responses

As it is observed in figure 1, the source node waits for some minutes after sending RREQ, and received RREPs are stored in response table.

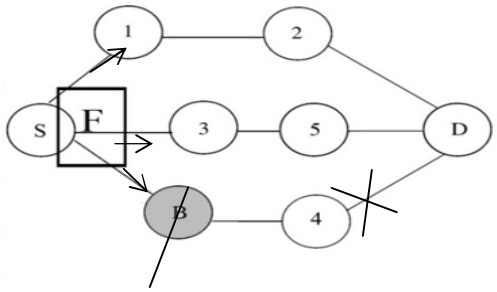


Figure 2: General distribution of fidelity table

Figure 2 shows distribution of fidelity table by the source node. After reducing fidelity level of black hole and distributing fidelity table by the source node, the value of response RF sent by black hole is reduced, and the response sent by black hole is not selected.

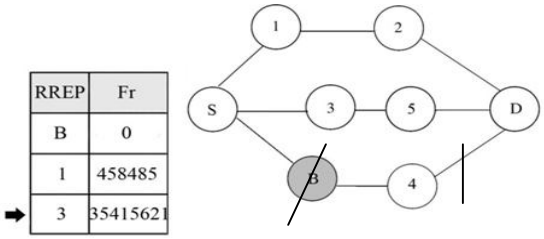


Figure 3: selecting a response

Figure 3 demonstrates the mode in which fidelity level of black hole node as well as its RF is zero. Among the nodes, the response of node 3 whose RF is high is used.

6. SIMULATION RESULTS

NS2 stimulator software has been used for simulation. The measured criteria for evaluating the efficiency of network are as follows: Delivery rate of package: refers to the ratio of the amount of data packages sent by the source node and the number of data packages received in final destination. End-to-end delay average: is delay average between data packages sent by the source node and data packages received by destination. This involves all delays created in the route, frequent delay in MAC layer and etc.

Routing overhead: is the ratio of produced control packages to sent data packages.

The number of nodes in the network is equal to 25 nodes. The perimeter of the network is 700*700 meters. These nodes are located in random places. In the scenario including black hole, one of these nodes is destructive node, and it performs black hole attack. Four traffic currents send data packages in the network with fixed rate. The size of packages is 512 byte. Duration of simulation is 300 seconds. Simulation is performed five times. The speed of nodes' movement toward random destination is different. Simulation results have been shown in the following diagrams. In these diagrams, AODV refers to the network without any black hole node, and routing is performed with AODV protocol. BAODV is a network with a black hole node, and routing is performed on the basis of AODV. FAODV is a network without any black hole node, routing is performed on the basis of proposed method, while BFAODV is a network with one black hole node, routing is performed according to proposed method. In scenario involving black hole node, the performance of proposed method is much better than AODV.

Figure 4: Delivery rate of package with different speeds

Figure 4 shows that when there is no black hole in the network, delivery rate of package in proposed method is lesser than AODV, but when there is a black hole, it has better application.

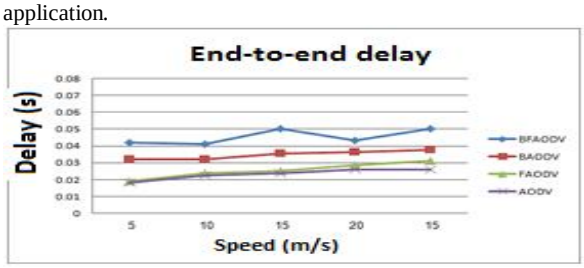


Figure 5: End-to-end delay in various speeds

Figure 5 demonstrates more delay of the proposed method in terms of sending data packages in comparison with AODV. This is due to source waiting to collect RREPs and calculation of the proposed method to select the response.

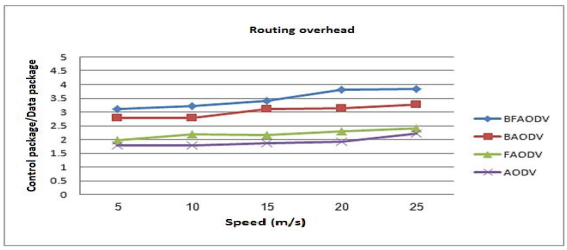


Figure 6: Routing overhead in different speeds

Also, due to much computing and general distribution of fidelity table, routing overhead in the proposed method is more than AODV. Black hole increases overhead due to sending control packages.

7. CONCLUSION

In this paper, a method has been proposed. In this method, according to behavior of black hole, the method of selecting AODV responses changes in a way that the source node ignores the response received from black hole node, and sends data packages from another route. This can be done by allocating fidelity level to network node, changing the way of selecting response, updating and distributing fidelity table by the source node. We simulated five scenarios by NS2 simulator. At first, five scenarios were simulated without a black hole node, and then they were simulated with a black

hole node. The results indicate that our proposed method has increased delivery rate of package from 22,32 percent to 42,34 percent in scenarios involving black hole. In this method, end-to-end delay and routing overhead is more than AODV due to waiting of the source node to collect response packages, more processing in comparison with AODV as well as general distribution of fidelity table.

8. REFERENCES

[1] H. Deng,.W. Li, and D. P. Agrawal, . 2002 Routing security in ad hoc networks, IEEE Communications Magazine, vol. 40, no. 10, pp. 70-75, Oct.

[2] S. Lee, B. Han, and M. Shin, Robust routing in wireless ad hoc networks, in ICPP Workshops, pp. 73, 2002.I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. NewYork: Academic, 1963, pp. 271–350Zachman, John A., "A Framework for Information Systems Architecture", IBM Systems Journal, Vol. 26, No. 3, 1987.

[3] S. Makki, N. Pissinou, H. Huang, 2003 The Security issues in the ad-hoc on demand distance vector routing protocol (AODV), In Proc. of the 2004 International Conference onSecurity and Management (SAM'04), pp.427-432C. E. Perkins, E. M. B. Royer, and S. R. Das, Adhoc OnDemand Distance Vector (AODV) routing, RFC 3561, July.

[4] Y.C. Hu and A. Perrig, 2004A survey of secure wireless ad hoc routing, IEEE Security & Privacy Magazine, vol. 2, no. 3,pp. 28-39, May/June 2004.

[5] M. A. Shurman, S. M. Yoo, and S. Park, 2004. Black hole attack in wireless ad hoc networks, in ACM 42ndSoutheast Conference (ACMSE'04), pp. 96-97,Apr.2004.

[6] Y. A. Huang, W. Fan, W. Lee, and P. S. Yu, Cross-feature analysis for detecting ad-hoc routing anoma-lies, in The 23rd International Conference on Distributed Computing Systems (ICDCS'03), pp. 478-487, May 2003.

[7] Y. A. Huang and W. Lee, Attack analysis and detection for ad hoc routing protocols, in The 7th International Symposium on Recent Advances in Intrusion Detection (RAID'04), pp. 125-145, French Riviera, Sept. 2004.

[8] Latha Tamilselvan, Dr. V Sankaranarayanan, Prevention of Co-operative Black Hole Attack in MANET, Journal of Networks, Vol. 3, No. 5, May 2008.

[9] Satoshi Kurosawa, Hidehisa Nakayama, Nei Kato, Abbas Jamalipour and Yoshiaki Nemoto, Detecting Blackhole Attack on AODV-based Mobile Ad Hoc Networks by Dynamic Learning Method, International Journal of Network Security, Vol.5, No.3, PP.338–346, Nov. 2007.

[10] C. E. Perkins, E. M. B. Royer, and S. R. Das, Ad hoc On-Demand Distance Vector (AODV) routing, RFC 3561,July 2003.

[11] Hesiri Weerasinghe, Huirong Fu, Preventing Cooperative Black Hole Attacks in Mobile Ad Hoc Networks:Simulation Implementation and Evaluation, International Journal of Software Engineering and Its Applications Vol. 2, No. 3, July, 2008.

[12] H. Deng, W. Li, and D. P. Agrawal, Routing security in ad hoc networks, IEEE Communications Magazine, vol.40, no. 10, pp. 70-75, Oct. 2002.

[13] S. Lee, B. Han, and M. Shin, Robust routing in wireless ad hoc networks, in ICPP Workshops, pp. 73, 2002.

[14] M. A. Shurman, S. M. Yoo, and S. Park, Black hole attack in wireless ad hoc networks, in ACM 2nd Southeast Conference(ACMSE'04), pp. 96-97, Apr. 2004.

[15] Mehdi Medadian, M.H. Yektaie and A.M Rahmani, Combatwith Black Hole Attack in AODV routing protocol in MANET, 2009, AH-ICI 2009. First Asian Himalayas International Conference, pp: 1-5, 3-5 Nov. 2009.

[16] Lata Tamilselvan, Dr.V Sankaranarayanan, Prevention of Cooperative Black Hole Attack in MANET , Journal Of networks, Vol. 3, NO. 5, May 2008.

[17] S. Marti , T. J. Giuli, K. Lai, and M. Baker,Mitigating routing misbehavior in mobile ad hoc networks. In mobile Computing and Networking (MOBICOM), pages 255-265,2000.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 690 - 698, 2013, ISSN: 2319–8656

The proposed System for Indoor Location Tracking

Ashraf Mohamed Ali Hassen
Department of Electronics and Communications
Modern Academy in Maadi, Egypt

Abstract: Indoor location tracking systems are used to locate people or certain objects in buildings and in closed areas. For example, finding co-workers in a large office building, locating customers within a shopping mall and locating patients in the hospital are a few applications of indoor location tracking systems. Indoor tracking capability opens up multiple possibilities. To address this need, this paper describes the implementation of a Bluetooth-based indoor location tracking system that utilizes the integrated Bluetooth modules in any today's mobile phones to specify and display the location of the individuals in a certain building. The proposed system aims for location tracking/monitoring and marketing applications for whom want to locate individuals carrying mobile phones and advertise products and services.
Keywords: Bluetooth, RSSI, Indoor tracking, location aware applications, ISM Band, ubiquitous, GPS

1. INTRODUCTION

Today's, Bluetooth functionality is widely included in almost all mobile phones. As a result most people carrying a mobile phone are also carrying a Bluetooth capable device. Hence, using the Bluetooth protocol contributes in minimizing the deployment cost of an indoor localization system. A Bluetooth-based indoor tracking system using low-cost stationary beacons and mobile host devices is described. The system permits low-cost, rapid deploying of applications where room-level location detection is sufficient. To allow the proposed localization algorithm to be applied to a variety of mobile phones, the algorithm is executed on a central server instead of the mobile phone and the location information resides on the central server. The proposed system is composed of a Server Module which is a java application that runs over desktop PC and is used to display the locations of the nearby mobile phones and send location-based advertising/control message/signal. The system also composed of a Locating Modules which written in C and runs over the Bluetooth-enabled boards which used to periodically collect the Bluetooth addresses of the nearby devices/mobiles accompanied with their sensed signal strengths and send them to the Server Module to process, specify, and display the location of the mobile phones. In this experimental study there are four Bluetooth-enabled boards used as prove of concept.

2. RELATED WORK

Indoor location tracking systems deliver usefulness to end users, if designed and maintained appropriately. The interest for assistive technology is increasing. Generally the need for services increases with age [1]. It is important to note that there is no one perfect location system. Each system must be evaluated based on the intended application across a variety of dimensions such as its accuracy, the infrastructure requirements, and the ability to scale.

2.1 Active Badge

The Active Badge is designed to be worn by visitors and employees of an organization to allow a central database to keep track of their location within the building. The badge transmits a unique code via a pulse-width modulated IR signal to networked sensors/receivers deployed throughout a building. The Active Badge uses 48-bit ID codes and is capable of two-way communication. The badge periodically beacons the unique code (approximately every 10–15s), and the information regarding which sensors detected this signal is stored in a central database. Since the IR signal does not travel through wall, the sensors are deployed throughout the space [2]. As with any diffuse infrared system, Active Badges have difficulty in locations with fluorescent lighting or direct sunlight because of the spurious infrared emissions these light sources generate.

2.2 Active Bat

Active Bat [3] is an ultrasonic-based location tracking systems consisting of ultrasound receivers dispersed in a space and location tags that emit ultrasonic pulses. Active Bat tags emit short pulses of ultrasound and are detected by receivers mounted at known points on the ceiling, which measure the time-of-flight of each pulse. Using the speed of sound, the distance from the tag to each receiver is calculated. Given three or more measurements to the receivers, the position of the tag can be determined using trilateration. One key concept of Active Bat is the use of an RF signal to cue the tag to transmit its ultrasonic pulse. The RF cue gives the receivers in the environment a starting point for timing the received ultrasonic pulse. Since the speed of light is significantly faster than the speed of sound, the RF signal delay is negligible and does not need to be considered for calculating the time-of-flight of the acoustical signal. The information about the location of Active Bat tags is managed by a central server. Using ultrasound time of flight this way requires a large fixed-sensor infrastructure throughout the ceiling and is rather sensitive to the precise placement of these sensors. Thus, scalability, ease of deployment, and cost are disadvantages of this approach [2].

3. SYSTEM DESCRIPTION

The proposed system utilizes the integrated Bluetooth modules in any today's mobile phones to specify and display the location of the individuals in a certain building. The proposed system aim for smart home location tracking/monitoring applications and marketing applications for whom want to locate individuals carrying mobile phones and provide context services and/or products advertisements. It is an integrated embedded and desktop system that helps the user to get the location of customers/home inhabitants within a certain region where there are several Locating Modules that are installed as one node per room to cover certain area which periodically investigate/inquiry the Bluetooth signal strength and addresses of the nearby mobile phones and send these data to a central Server Module. The Server Module

processes, specifies, and displays the location of each mobile phone, by means of its MAC address using table and a simplified graphical map. Moreover the administrator can set certain messages to be sent if the mobile phone is near a certain place, where the Server Module can send the settled message to the mobile users.

The proposed approach is motivated by the drawbacks reported by the existing indoor localization approach and by the resource limitations typically imposed on the simple Bluetooth nodes, where a neural network-based scheme is described. The proposed algorithms based mainly on measurement of the received signal strength indicator (RSSI) this measurement is used as input to an algorithm that estimates the location of the mobile node that algorithm utilizes the use of neural network rather than specific measurement values and it targets the same order of simplicity as the I^{3}BM [4] algorithm, but with the high accuracy of the Environment Adaptive [5] [6] algorithm. To do so; it includes a training phase, for modeling the surrounding environment and calculating the internal weights of the neural network, and location determination phase. Room-level location is selected, where rooms refer to physical rooms or specific logical subspaces, as this is the most needed for smart home applications or generally indoor environment.

4. SYSTEM ARCHITECTURE

The system architecture is shown in figure 1

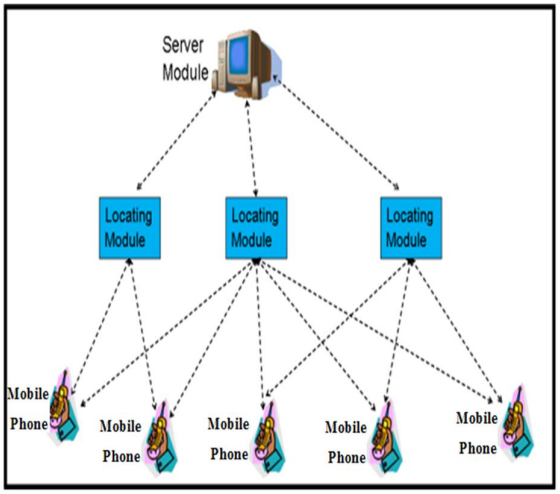


Figure 1 system architecture

4.1 Server Module: It is a JAVA application runs over desktop personal computer and it is used to process, specify and display the locations of the nearby mobile phones and send location-based advertising message. This application utilizes a Java neural network library, by Neuroph, that supports creating, training and saving neural networks.

4.2 Locating Modules: It is a C program that runs over the Bluetooth-enabled Arduino boards, used to periodically collect the Bluetooth addresses of the nearby devices accompanied with their signal strengths and send them to the Server Module to process, specify, and display the location of the mobile phones. Each Locating Module covers a number of nearby mobile phones according to their location and received signal strength. In this study there are four Bluetooth-enabled boards used as a prove of concept. Figure 2 shows rooms layout.

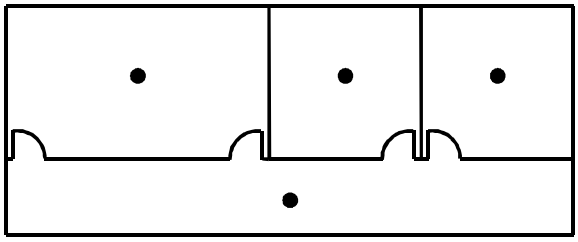


Figure 2 rooms layout

4.3 Mobile phones with Bluetooth service enabled and set to visible. The administrator will just invoke the application and communicate with it appropriately to explore nearby mobile phones.

5. TRACKING ALGORITHM

Indoor location tracking/monitoring approach based mainly on measurement of signal characteristic as the received signal strength indicator (RSSI), the link quality indicator (LQI), the time of arrival (ToA), or the angle of arrival (AoA). This measurement is used as input to an algorithm that estimates the location of the mobile node; in this study we select RSSI as our measure. Measuring RSSI is very simple, it is implemented in all Bluetooth protocol stacks and is available for all Bluetooth platforms. Hence, there are many indoor location monitoring/tracking schemes that utilize this metric as the measurement for location estimation [7]. Among the most popular are the I^{3}BM [4] and the Environment Adaptive [1] [5].

The major challenge, which has addressed by all indoor location monitoring/tracking schemes, is the high sensitivity of the measured signal characteristic to the surrounding environment. The I3BM scheme has addressed this challenge through mechanisms as signal filtering and room adjacency utilization. The scheme has maintained simplicity, but the achieved accuracy has not been that acceptable [4]. The Environment Adaptive has addressed the high sensitivity challenge by modeling the environment in a calibration phase. Accurate location has been obtained at the expense of high computational cost and large memory requirement [5] [6].

5.1 I^{3}BM Algorithm

The I^{3}BM targets room-level position estimation. It positions the mobile node in the room whose beacon node yields the highest RSSI. To reduce the high sensitivity of the measured values to the surrounding environment, the I3BM employs mechanisms as signal filtering and room adjacency utilization. As there are many factors influencing the signal readings, it is irrational to base the positioning decision upon a single measurement, but rather upon the outcome of filtering a group of signal values [4] [8].

Assuming that noises tend mainly to reduce the signal strength, a MAX filter has demonstrated good performance, where the selected RSSI value is the maximum received value for a sequence of packets communicated between the mobile and each beacon node over a pre-specified period of time[8]. Room adjacency utilization also improves accuracy of location estimation where people move according to feasible paths. Hence, only rooms that are adjacent to current location should be considered as possible estimate for the new location. Figure 8 demonstrates the room adjacency relationship for the layout in Figure 2.

The I^{3}BM scheme is simple but it doesn't provide enough accuracy to build a convenient smart home application. This is due to the fact that the signal measurements are highly

sensitive to the surrounding environment and only the information contained in the highest RSSI is utilized for decision making, whereas all the information reported by the other beacon nodes are ignored.

5.2 Environment Adaptive Algorithm

The Environment Adaptive scheme has been developed [5] [6] driven by the goal of having an accurate position estimation scheme that is robust to environmental changes. The scheme has two phases, calibration phase and localization phase. In the calibration phase, environment characteristics are modeled by computing the m × m transformation matrix, T, which relates the distance between the m anchor nodes and the RSSI of the packets communicated between them using

$$T = D\,S^{-1} \quad \text{--- (1)}$$

Where D denotes the m × m distance matrix whose element d_{ij} represents the known Euclidean distance between the i^{th} and j^{th} anchor nodes, and S denotes the m × m strength matrix whose element s_{ij} represents the RSSI value of the packets between the i^{th} and j^{th} anchor nodes.

In the localization phase, given T and the m-dimensional vector I, whose element I_i represents the RSSI value of the packets between the mobile node and the i^{th} anchor node, then

$$V = T\,I \quad \text{--- (2)}$$

Where V denotes the m-dimensional vector whose element v_i is the calculated distance between the mobile node and the i^{th} anchor node. The mobile node position estimate, P(x, y) is then determined using the following gradient descent equation [9] [10].

$$P_{k+1} = P_k + \propto \sum_{i=1}^{m} \left(1 - \frac{V_i}{E(A_i, P_k)}\right)(A_i - P_k) \quad \text{--- (3)}$$

where P_k is the mobile node position estimate at the k^{th} iteration, A_i is the position of the i^{th} anchor node, α is the gradient step size, and $E(A_i, P_k)$ is the Euclidean distance between the i^{th} anchor node and the mobile node position estimate. At the first iteration, the scheme assumes that the mobile node is directly located at the position of the anchor node with the minimum v_i. The estimated mobile node position, (x, y), is finally mapped to a specific room as per the space dimensions of the underlying environment.

The Environment Adaptive scheme has demonstrated good performance and much less sensitivity to variations in the surrounding environment [5]. However, the inverse in (1) and the iterative nature in (3) are computationally expensive and necessitate slow response and large memory requirements. Here, it should be noticed that the scheme can utilize either the RSSI measurements or its LQI alternative; the original description of the scheme [5] employs the RSSI measure.

5.3 The Proposed Neural Network-based Scheme

Neural networks are mostly used for fuzzy, difficult problems that don't yield to traditional algorithmic approaches. Learning in neural networks is particularly useful in applications where the complexity of the data or task makes the design of such functions by hand impractical. The proposed system utilize a neural network which could be described as "4 × 6 × 4 neural network" where it has one input layer with four neurons, one hidden layer with six neurons and one output layer with four output neurons.

5.3.1 Developing Environmental Radio Map

The first step is creating a radio map of the area. Measurement locations were chosen roughly every one meter, as shown in figure 3, excluding the places that were covered by tables or desks. The measurements were taken by a person who was holding the mobile phone, to simulate real usage. Because Bluetooth signals are absorbed by the human body, the direction of the person with regard to the beacon can influence the signal. We therefore measured the signals in four different directions, taking 1 sample in each direction (facing north, east, west and south). Overall measurements were taken on 152 different locations, giving a total of 608 fingerprints.

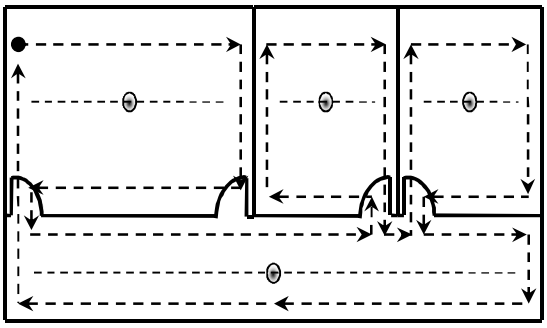


Figure 3 Radio Map for the locations of measurements

These measurements are applied to a neural network within the JAVA program as a training set. Primarily we utilized a tool named "Neuroph" which is a lightweight Java neural network framework to develop common neural network architectures. It contains well designed, open source Java library with small number of basic classes which correspond to basic neural network concepts. Also has nice GUI neural network editor to quickly create Java neural network components. Creating the neural network architecture therefore means coming up with values for the number of layers of each type and the number of nodes in each of these layers.

Every neural network has exactly one input layer; and the number of neurons comprising this layer is completely and uniquely determined as the number of columns in your data which is four in our case. Also additional node as a bias node is included, Bias neuron is very important, and the error-back propagation neural network without Bias neuron does not learn. A bias input always has the value of 1. Without a bias, if all inputs are 0, the only output ever possible will be a zero [11].

Like the Input layer, every neural network has exactly one output layer. Determining its size (number of neurons) is simple; it is completely determined by the chosen model configuration. In this experiment the neural network is used as a classifier, and then it has number of output neurons equal to four which is the number of rooms to classify between them [11].

Finally, concerning the hidden layers since one hidden layer is sufficient for the large majority of problems where there is a consensus is the performance difference from adding additional hidden layers: the situations in which performance improves with a second (or third, etc.) hidden layer are very small. Now to determine the correct number of neurons to use

in the hidden layers several trails is performed that guided by the following rule-of-thumb [11]:

- The number of hidden neurons should be between the size of the input layer and the size of the output layer.
- The number of hidden neurons should be 2/3 the size of the input layer, plus the size of the output layer.
- The number of hidden neurons should be less than twice the size of the input layer.

These three rules provide a starting point for our guided trial and error process, with the help of the visual tool "Neuroph Studio"; we reach to the optimal size of hidden layer to be six neurons. Finally we can describe the utilized neural network as "4 × 6 × 4 neural network" where it has one input layer with four neurons, one hidden layer with six neurons and one output layer with four output neurons.

5.3.2 Procedure of Obtaining a Representative Trained Neural Network

In order to train neural network these data set have to be normalized. Normalization implies that all values from the data set should take values in the range from 0 to 1. For that purpose it would be used the following formula:

$$X_n = \frac{X - X_{max}}{X_{max} - X_{min}}$$

Where:

X, is the value that should be normalized

X_n, is the normalized value

X_{min}, is the minimum value of X

X_{max}, is the maximum value of X

The output is represented by 4 digits of data set represent the target room. For example, "1 0 0 0" represent office room, "0 1 0 0" meeting room, "0 0 1 0" lab room and "0 0 0 1" cafeteria room.

We used the Neuroph [11] wizard to create a Neuroph project, during this wizard there are some settings that we used to create the neural network. Among these setting is the type of training, which is set to supervised training which is the most common way of neural network training. As supervised training proceeds, the neural network is taken through a number of iterations, until the output of the neural network matches the anticipated output, with a reasonably small rate of the error. Figure 4 shows Trained Neural Network Snapshoot.

To achieve the main goal and decide which the best solution for this problem is, several neural networks are created with different settings. Among these settings is the number of neurons in the hidden layer which play a vital role in the whole system performance. Five, six, seven and eight neurons in the hidden layers are tried and the best performance is estimated and presented in the results section.

For testing the neural network, the experimental data were divided randomly into two sets (training and testing dataset). Separation of 70/30 dataset was done for training/testing where 70% of the experimental data used to train the neural network and 30% of it for testing. Simply Neuroph framework is used to test the performance of the trained network.

This is repeated several times within Neuroph framework till reaching to the optimal results then the corresponding parameters are utilized in the final Java application; experimental results are given in what follows.

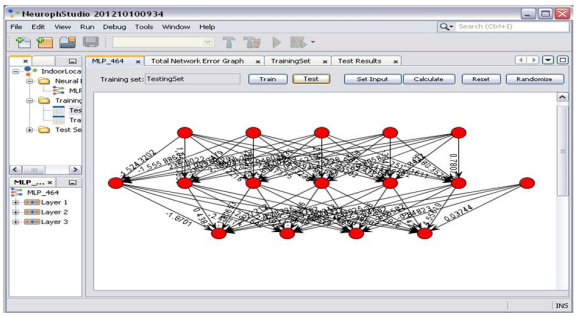


Figure 4 Trained Neural Network Snapshoot

6. SYSTEM CONFIGURATION

6.1 Locating Modules Setup

Setup one locating module in each room, it is recommended to have the locating modules separated by minimum of 4 meters and maximum of 10 meters.

6.1.1 Arduino BT (Bluetooth) Board

The Arduino BT shown in figure 5 is a microcontroller board based on the ATmega328 and the Bluegiga WT11 Bluetooth module[12].

Figure 5 Arduino Board

It supports wireless serial communication over Bluetooth (but is not compatible with Bluetooth headsets or other audio devices). It has 14 digital input/output pins (of which 6 can be used as PWM outputs and one can be used to reset the WT11 module), 6 analog inputs, a 16 MHz crystal oscillator, screw terminals for power, an ICSP header, and a reset button. It contains everything needed to support the microcontroller and can be programmed wirelessly over the Bluetooth connection. Each of the 14 digital pins on the Arduino BT can be used as an input or output, using pinMode(), digitalWrite(), and digitalRead() functions. They operate at 5 volts. Each pin can provide or receive a maximum of 40 mA and has an internal pull-up resistor (disconnected by default) of 20-50 kOhms. In addition, some pins have specialized functions.

6.2 Server Module Setup

A Computer device that has a Bluetooth dongle connected to it through the USB port. The computer device run the developed JAVA application, named Server Module, which

could connect to all Locating Modules to sent command and receive response. Figure 6 shows example of server measurements configuration file.

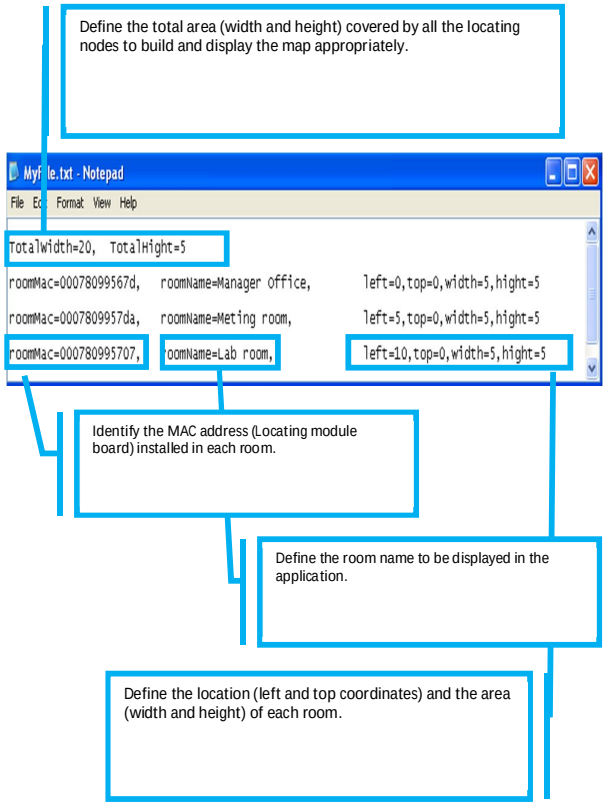


Figure 6 Example of server measurements configuration file

The running java application has a suitable UI interface, which allow for users to easily benefits from different capabilities of the system.

6.2.1 Building the Server Configuration File

The configuration file will be located on the administrator module machine. It is just a text file contains in its first line the total width and total height of the area that the map will displayed in, then each line next the first line contains a complete entry. Any entry will contains room MAC, room displayed name, room top position, room left position, room width and finally room height. The server configuration file is needed to inform the server of the mapping between the locating module and the corresponding room, the following example demonstrates the contents of the file.

6.3 System Hardware Requirements

Table 1. Hardware Resources

Resources Name	Description
Mobile Phone	Any mobile phone that supports Bluetooth and Messaging
Arduino Board	Bluetooth Enabled board
ATMEGA328P Programmer	To burn the Arduino application on the board
Server PC	To deploy/run Server Module
USB Bluetooth Dongle	Attached to the Server PC

6.4 System Software Requirements

Table 2. Software Resources

Resources Name and version	Description
NetBeans IDE	Used for developing the Server Module application
Arduino IDE	Used to develop and compile the application that run on Arduino Bluetooth Board
AVR Studio 4	Used to program the Atmeg328P microcontroller that is on Arduino Bluetooth Board

7. SYSTEM DEPLOYMENT

Figure 7 depicts the layout of the environment used to develop the indoor locating system. As depicted, a single fixed, beacon node is installed in each room at a pre-determined location.

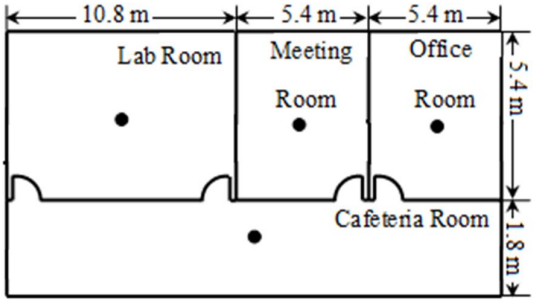


Figure 7 Layout of the overall system

The experiments were carried out using Arduino Bluetooth board[12] as beacons where there is a board installed per each room, and mobile phones from different brands (Nokia 5630 Music Express, BlackBerry 9700, and Samsung D900i) for measuring the Bluetooth signals. There is no software needed to pre-install on the mobile phones just it is required to be Bluetooth enabled and made discoverable. The physical environment consist of four rooms where there is typically three physical rooms titled as Office room, Meeting room, Lab room and finally the corridor along these rooms but virtually we considered it as another room named, Cafeteria room, as shown in figure 7. The office room has size of 5.4x5.4 m2 large, the meeting room has size of 5.4x5.4 m2, the Lab room has size of 10x5.4 m2 and the corridor has size of 21.6x1.8 m2 large. The Locating Module run over the Arduino Bluetooth-enabled boards/beacons and the Server Module runs over a personal computer with a Bluetooth dongle. It should be noticed that beacon nodes are placed such that once the mobile node is moved into a room it becomes closest to that room's beacon node, taking into account the effect of walls separating different rooms. This is rational to probabilistically increase the RSSI value of the signal received from the closest beacon node over the values received from the others.

Server module will create one thread to collect the data from all locating modules (4 modules in this study). By this approach the server module will avoid synchronization problems on shared resources between multiple threads. But will add more code complexity to minimize the refresh rate. Within this thread the server Module will open multiple Bluetooth connections simultaneously with multiple locating

modules using the Bluetooth dongle and periodically iterate on all the boards to send command and receive response.

The server module will iterate periodically on all the boards, each iteration composed of two rounds, in the first round the server module will just send start "S" command to all the boards sequentially. In the second round, the server module might wait only once at the first board to finalize its inquiry time but all other following boards will be ready with its inquiry response and the server module will retrieve its data without waiting more time.

When the locating module receives the start, "S", command it will start the inquiry for nearby mobiles. Upon receiving this inquiry by nearby mobiles, each mobile node will reply that inquiry by an inquiry response that contains the MAC address of that mobile and its corresponding signal strength, RSSI. The RSSI value in the inquiry response is automatically calculated by the Bluetooth MAC layer. The locating module will collect these data (MAC addresses and signal strength) and send it line by line to the server module just upon server module read operation. When the locating modules send all the data to the server module the locating module will be waiting for the next start command. The locating module will close the connection with the server module upon user request to exit/stop the application where the server module will send a stop command "X" to all beacon nodes that upon receiving the stop command will close the connection with the server module.

Now the server module processes the whole information and estimates the locations of all Bluetooth-enabled mobile phones. It then invokes the appropriate services for each mobile user according to the estimated location. Actually the Server module employ a trained neural network where for a certain mobile MAC address there will be a vector R that has the RSSI values of that MAC address for each room where r_i refer to the RSSI value of that mobile as reported by the beacon node in room i. by applying this as an input to the trained neural network and based on the different values of the internal weights between neurons that reflect the environmental effect on the measured value some output neurons are activated referring to the estimated location (room-level). Figure 8 shows room adjacency relationships

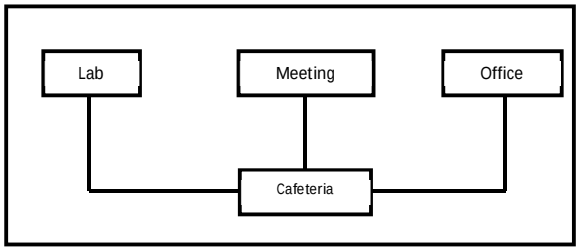


Figure 8 ROOM Adjacency Relationships

To improve the accuracy of position estimation a post processing mechanism as room adjacency is utilized where people move according to feasible paths. Hence, only rooms that are adjacent to current position should be considered as possible estimate for the new position.

To improve the results of the overall system and enhance system stability of the displayed status, the mobile node has to say two consecutive times that it is say in room 1 but this will affect the refresh rate of the displayed status so a confidence factor is calculated and this value is compared to certain confidence threshold, calculated by trial and error, and based on this comparison the displayed status will be updated minimizing the refresh rate and improving the system stability.

8. SYSTEM TEST RESULTS

Figures 9 ,10,11 and 12 show total network error using eight,five,six and seven neurons respectively tables 3,4,5 and 6 Show System Positioning Estimation Accuracy Using Eight ,five ,six and seven Neurons respectively these results are taken through different attempts.

8.1 Attempt 1

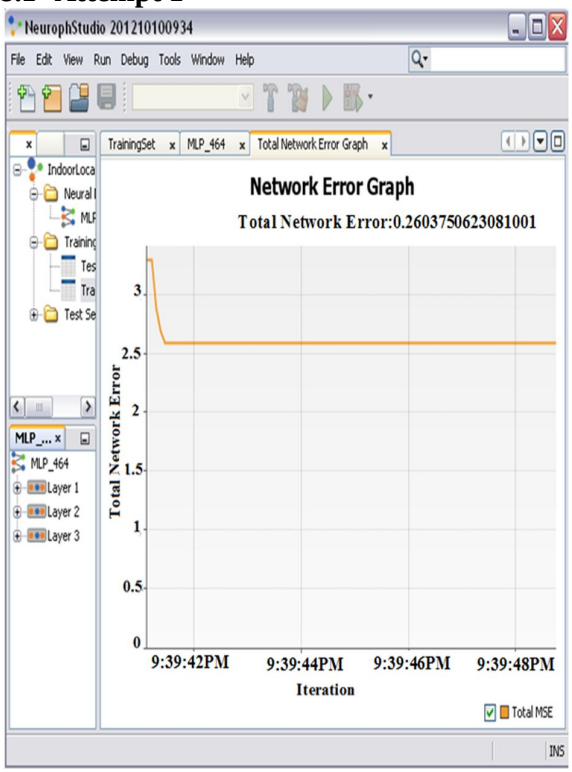


Figure 9 Total Network Error Using Eight Neurons

Table 3. System Positioning Estimation Accuracy Using Eight Neurons

Scheme	Refresh Rate	Accuracy
Neural network utilized	10 sec	73.93%
Room adjacency applied	10 sec	87.23%
Room adjacency applied and twice per room	20 sec	91.48%
Room adjacency and twice per room with confidence threshold applied	10 sec	91.48%

Using Neuroph, We create a neural network with 4 input neurons, 4 output neurons and set the number of neurons in the hidden layer to eight. Using the collected training data samples to train the neural network, and after testing we got that in this attempt the total network error is 26% as illustrated in the figure. Now utilizing this neural network in our java application with these configurations we get the following results.

8.2 Attempt 2

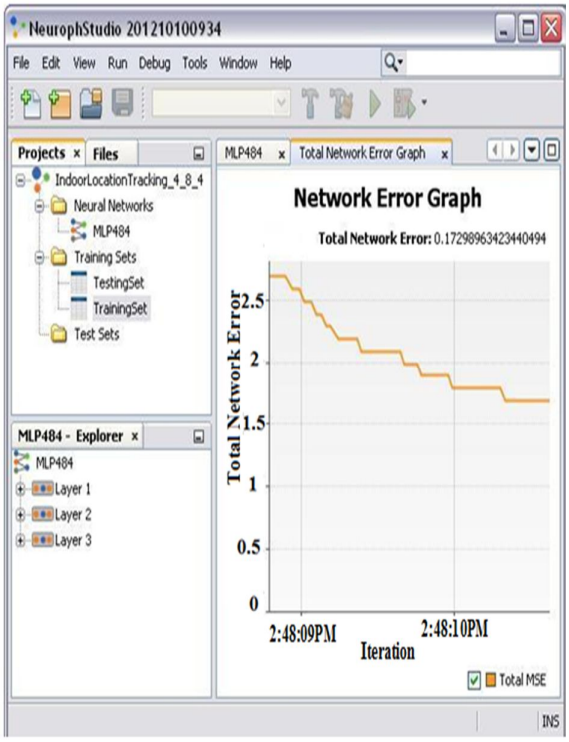


Figure 10 Total Network Error Using Five Neurons

Table 4. System Positioning Estimation Accuracy Using Five Neurons

Scheme	Refresh Rate	Accuracy
Neural network utilized	10 sec	82.97%
Room adjacency applied	10 sec	90.95%
Room adjacency applied and twice per room	20 sec	94.68%
Room adjacency and twice per room with confidence threshold applied	10 sec	94.68%

We create a new neural network with 4 input neurons, 4 output neurons and set the number of neurons in the hidden layer to five. Using the same collected training samples as previous attempt to train the neural network, after testing we got that in this attempt the total network error is 17% as illustrated in the figure. Now utilizing this neural network in our java application with these configurations we get the following results.

8.3 Attempt 3

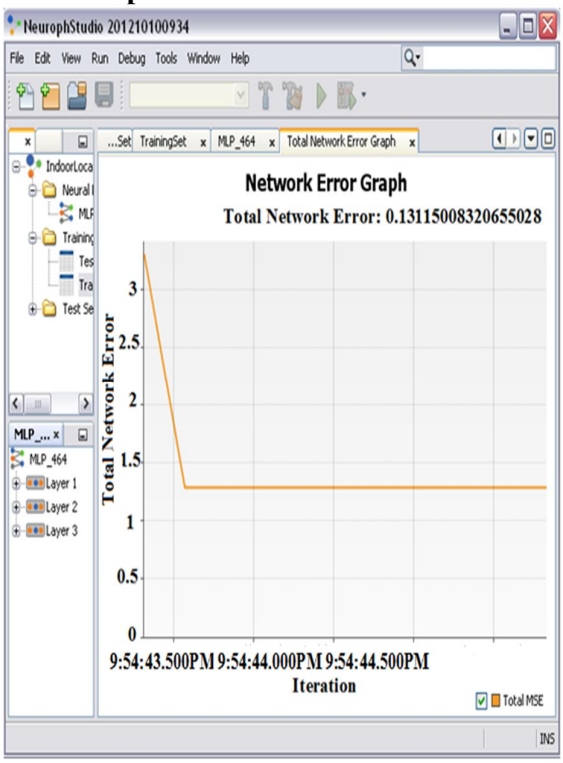


Figure 11 Total Network Error Using Six Neurons

Table 5. System positioning estimation accuracy using six neurons

Scheme	Refresh Rate	Accuracy
Neural network utilized	10 sec	86.70%
Room adjacency applied	10 sec	95.21 %
Room adjacency applied and twice per room	20 sec	97.34%
Room adjacency and twice per room with confidence threshold applied	10 sec	97.34%

We create a new neural network with 4 input neurons, 4 output neurons and set the number of neurons in the hidden layer to six. Using the collected training samples, as in the previous attempts, to train the neural network, after testing we got that in this attempt the total network error is 13% as illustrated in the figure. Now utilizing this neural network in our java application with these configurations we get the following results.

8.4 Attempt 4

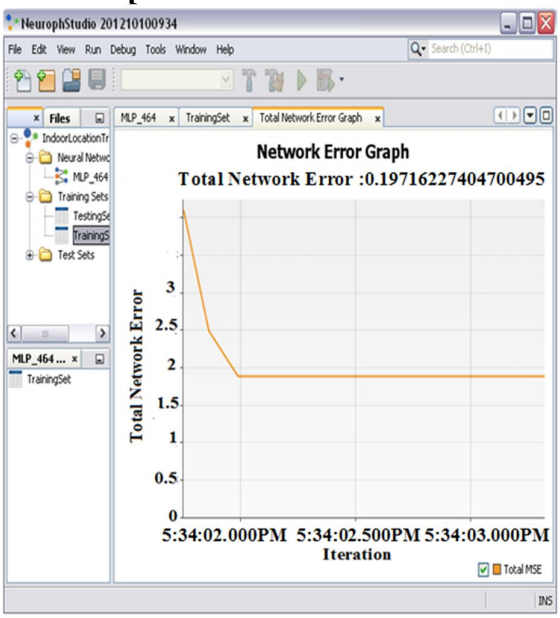


Figure 12 Total Network Error Using seven Neurons

Table 6. System positioning estimation accuracy using seven neurons

Scheme	Refresh Rate	Accuracy
Neural network utilized	10 sec	80.32%
Room adjacency applied	10 sec	89.36%
Room adjacency applied and twice per room	20 sec	92.02%
Room adjacency and twice per room with confidence threshold applied	10 sec	92.02%

Finally, we create a new neural network with 4 input neurons, 4 output neurons and set the number of neurons in the hidden layer to seven. Using the collected training samples, as in the previous attempts, to train the neural network, after testing we got that in this attempt the total network error is 19% as illustrated in the figure. Now utilizing this neural network in our java application with these configurations we get the following results.

9. CONCLUSION

As demonstrated, proper selection of the number of internal neurons in the hidden layer influences the accuracy of the results and from the mentioned trials we could conclude that the optimum size of the hidden layer in this typical environment is six neurons that gave the optimum results. Another point utilizing room adjacency relationships consistently improved estimation accuracy. Also, to minimize the stochastic instability of the RSSI measurements the user location is updated only when the system report two consecutive times the same decision but this effect the system refresh rate and to retain the refresh rate again to its original value a confidence threshold value is applied and by this approach estimation accuracy and system stability are maximized. In a nutshell estimating location is based not only upon a single value reading of the RSSI from different rooms and selecting the highest RSSI, but based on the nature of the underlying environment which implicitly included in the neural network weights calculated in the training phase and applying a post processing fixes to enhance system accuracy and stability.

10. FUTURE WORK

Still, the field of indoor location tracking is not yet mature enough to propose a universal solution that works in any setting. Currently, the choice of technology and positioning algorithm depends mainly on the requirements of the system. One important aspect to consider is how our buildings will be used in the future. Will we perform other or different daily activities other than today? Do we need support from technology in order to perform such new activities?

Regarding the technology, multiple technologies have been proposed to tackle the problem of indoor localization some examples being infrared, ultrasound and Radio Frequency IDentification. Still, most research is dedicated to the usability of two technologies. Large number of papers, e.g., [13] and [14], argue that Ultra Wide Band (UWB) radio offers excellent means to determine one's location with high precision. To summarize, there is no single location technology today that is ubiquitous, accurate, low-cost (in terms of required hardware and installation), and easy to deploy. Although a novel location technology that fits all these parameters might still become available in the future, a more realistic approach is to combine several of the existing technologies into an integrated location system. For instance, integration between GPS and indoor positioning can be made and by this integration positioning signal can be obtained everywhere (ubiquitous).

11. REFERENCES

[1] Molin, G., Pettersson, C., Jonsson, O. and Keijer, U., "Living at Home with Acquired Cognitive Impairment – Can Assistive Technology Help? Technology and Disability", IOS Press, pp 91-101, 2007.

[2] John Krumm, "Ubiquitous Computing Fundamentals", ISBN 978-1-4200-9360-5, September 21, 2009.

[3] Ward, A., Jones, A., and Hopper, A. "A new location technique for the active office", IEEE Personal Communications 4(5), 1997.

[4] Ming-Hui Jin, Chung-Jung Fu, Chih-Hao Yu, Hung-Ren Lai, and Ming- Whei, "I3BM ZigBee Positioning Method for Smart Home Applications," International Journal of Smart Home, Vol. 2, No.2, pp. 127-139, April, 2008.

[5] Janire Larranaga, Leire Muguira, Juan-Manuel Lopez-Garde, and Juan-Ignacio Vazquez, "An Environment Adaptive ZigBee-Based Indoor Positioning Algorithm," IEEE International Conference on Indoor Positioning and Indoor Navigation, Switzerland, pp. 1-8, 15-17 September, 2010.

[6] Miguel Dom´ınguez-Dur´an, D´ebora Claros, Cristina Urdiales, Francisco Coslado, and Francisco Sandoval, "Dynamic Calibration and Zero Configuration Positioning System for WSN," The 14th IEEE Mediterranean Electrotechnical Conference, MELECON, pp. 145-150, 5-7 May, 2008.

[7] Ambili Thottam Parameswaran, Mohammad Iftekhar Husain, and Shambhu Upadhyaya, "Is RSSI a Reliable Parameter in Sensor Localization Algorithms – An

Experimental Study," Department of Computer Science and Engineering, State University of New York at Buffalo, USA, pp. 1-5, 2009.

[8] Y. X. Zhao, Q. Shen, and L. M. Zhang, "A Novel High Accuracy Indoor Positioning System Based on Wireless LANs," Electromagnetics Research C, Vol. 24, pp. 25-42, 2011.

[9] J. Hightower, R. Want and G. Borriello; SpotON: An indoor 3D location sensing technology based on RF signal strength, UW CSE00-02-02, February 2000, http://www.cs.washington.edu/homes/jeffro/pubs/hightower2000indoor/hightower2000indoor.pdf

[10] RadioFrequencyIdentification (RFID) home page, http://www.aimglobal.org/technologies/rfid/, February 2012

[11] http://neuroph.sourceforge.net/, February 2012

[12] http://arduino.cc/en/Main/ArduinoBoardBT?from=Main.ArduinoBoardBluetooth, February 2012

[13] G. Zhang, S. Krishnan, F. Chin, and C.C. Ko., "UWB multicell indoor localization experiment system with adaptive TDOA combination", In Vehicular Technology Conference, IEEE 68th, pages 1-5, 2008.

[14] S. Gezici, Zhi T., G.B. Giannakis, H. Kobayashi, A.F. Molisch, H.V. Poor, and Z. Sahinoglu. "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks", Signal Processing Magazine, IEEE, 22(4):70-84, 2005.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 699 - 707, 2013, ISSN: 2319–8656

Real Time Object Identification for Intelligent Video Surveillance Applications

L.Menaka
R.V.S.College of Engineering and Technology
Karaikal, India

M.Kalaiselvi Geetha
Annamalai University
Chidambaram, India

M.Chitra
R.V.S.College of Engineering and Technology
Karaikal, India

Abstract Intelligent video surveillance system has emerged as a very important research topic in the computer vision field in the recent years. It is well suited for a broad range of applications such as to monitor activities at traffic intersections for detecting congestions and predict the traffic flow. Object classification in the field of video surveillance is a key component of smart surveillance software. Two robust methodology and algorithms adopted for people and object classification for automated surveillance systems is proposed in this paper. First method uses background subtraction model for detecting the object motion. The background subtraction and image segmentation based on morphological transformation for tracking and object classification on highways is proposed. This algorithm uses erosion followed by dilation on various frames. Proposed algorithm in first method, segments the image by preserving important edges which improves the adaptive background mixture model and makes the system learn faster and more accurately. The system used in second method adopts the object detection method without background subtraction because of the static object detection. Segmentation is done by the bounding box registration technique. Then the classification is done with the multiclass SVM using the edge histogram as features. The edge histograms are calculated for various bin values in different environment. The result obtained demonstrates the effectiveness of the proposed approach.

Keywords: object classification, SVM, morphological operator, edge histogram

1. INTRODUCTION

In the present scenario, a large amount of security measures are needed in public areas due to terrorist activities. Manual surveillance is highly impractical. Hence automated video surveillance becomes essential. In that rationale, object identification is the preliminary step. The ability to reliably detect pedestrians from video data has very important applications in many fields like, intelligent transportation, automated surveillance, security, robotics and assistive technology for visually impaired, advanced human machine interfaces, automated driver assistance systems in vehicles etc.

Pedestrian are more vulnerable to accidents and collisions involving pedestrians often produce severe injuries. Accurately detecting pedestrians from a video is one of the most challenging tasks for object detection which attracts most of the researchers working in this field. This paper proposes an approach for conditions both pedestrian, vehicle and tree identification in real time with and without background subtraction technique.

1.1 Difficulties in object identification

The main challenges of vision-based object identification for automated video surveillance applications are listed below

1.1.1 The image background may be cluttered with other stationary object:

This is especially obvious for the urban road scenes. This may complicate the object identification in the video surveillance system.

1.1.2 Images are captured from a moving platform:

The captured images may contain vehicles and other objects at different distances, in a background that constantly changing, when the test vehicle is moving. This makes the background subtraction technique commonly used to detect objects in a video surveillance system unsuitable for this usage.

1.1.3 Variable outdoor illumination conditions:

The illumination of the road environment may change during different times of day and under different weather conditions. This may affect the quality of the captured images.

1.1.4 Occlusion and different poses of objects in a image:

In moving object identification, different poses of pedestrian and vehicle need to be considered for the accurate identification. Consideration of different poses of human and vehicles are difficult in the implementation of video surveillance system. The occlusion is also unavoidable in the case of traffic congestion.

2. RELATED WORK

Vision based vehicle detection is an area of research in the intelligent transportation systems community. In the literature, many studies have been performed on the static images. A support vector machine(SVM) approach was used in [1],built multiple detectors using Haar wavelets, PCA, truncated wavelet features using neural networks and SVM classifiers. Template matching is one of the methods used for vehicle detection and tracking. In this method, Scale Invariant Feature Transform (SIFT) points were extracted which is invariant to image scaling and rotation and partially invariant to change in illumination and 3D rotation. A statistical approach has been used in [3], performing vehicle detection using principle component analysis (PCA) and independent component analysis(ICA) to do classification on a statistical model and increased its speed by modeling the PCA and ICA vectors with the weighted Gaussian mixture model. A review of ecent template matching methods for detection and tracking of vehicle is presented in [4]. The binary classification scheme is

an efficient tool which can be used for object detection and matching which is described in [7].

There are many technologies that are currently being used for pedestrian and vehicle detection such as ultrasonic sensors, Doppler radar sensors piezo-metric sensors etc which is presented in [8]. These sensors while being very effective have various drawbacks ranging from cost effectiveness to durability. The "irregular blobs" can be detected and then cluster the pieces according to the common motion constraint of the extracted features. It reduces the computational costs by limiting the feature analysis only to "irregular blobs" is presented in [9]. Using 2D and 3D models for detection has also been explored by several authors. A 3D model-based detection approach with background subtraction is presented in [11]. The 2D templates has been created from the 3D vehicle models which is used to generate multiple hypotheses for a given foreground mask. This approach use only the template contours, so like the feature based approaches, its performance on noisy, low resolution and crowded scenes is uncertain.

The partial occlusion problem is addressed by a feature based tracking algorithm in [14]. The detection is done based on the "corner" features and then grouped them according to a "common motion constraint". However, both algorithms depend purely on the accuracy of feature detection and matching, which makes them error prone in noisy, low resolution videos. Common motion constraint is not applicable in very crowded scenes, where the vehicles are forced to move at similar speeds. Video based detection emerged as an important aspect of research, as proliferation high performance cameras and faster inexpensive computing systems became assessable.

This paper introduces two methods at detection stage. First method for classifies the objects using height and width. Second method classifies the object by Support Vector Machine (SVM) classifier. The method combines edge histogram and SVM. The training of the SVM classifier is carried out offline.

The proposed classification procedure is useful not only for increasing the speed and reliability of feature matching but also for reducing the computational load associated with the robust estimation of the registration parameters.

3. HEIGHT – WIDTH BASED CLASSIFICATION

3.1. BACKGROUND SUBTRACTION

The main reason of using background subtraction is that it works well if the background is static for long time. In traffic surveillance system, camera often remains static. This approach uses mixture of Gaussian for background subtraction.

The proposed approach uses height-width based classification method as seen in Fig. 1.

3.1.1. Mixture of Gaussians

Mixture of Gaussians method maintains a density function for every pixel and is capable of handling multi model backgrounds as in Eq. (1)

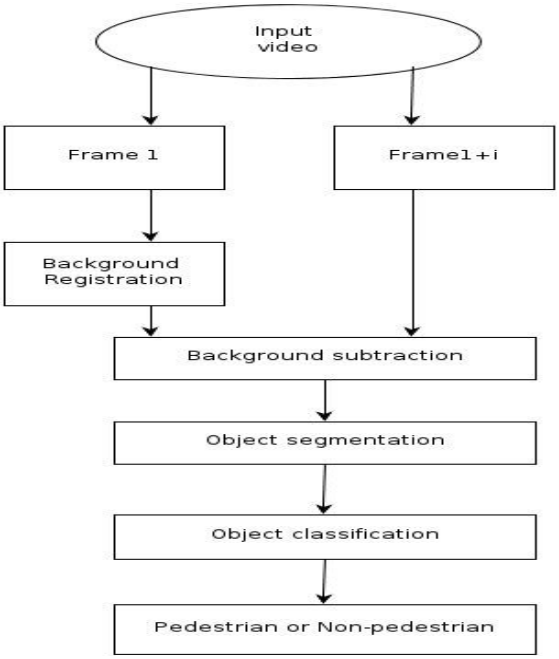


Fig. 1. A model for height-width based approach

Objects must be initially classified as pedestrians or vehicles.

$$p(x|\pi,\mu,\Sigma) = \sum_{k=1}^{K} \pi_k \; N(x|\mu_k \Sigma_k) \qquad (1)$$

Where, x random variable

π_k Mixing coe□cient of the k^{th} Gaussian

μ_k Mean of the k^{th} Gaussian

Σ_k Covariance of the k^{th} Gaussian

Mixture of Gaussians robustly deal with lighting changes, repetitive motions, clutter, introducing or removing objects from the scene and slowly moving objects. It can be updated without having to store large number of frames in buffer hence reducing memory costs. The learning rate is passed into MOG model. The learning rate is the rate at which the model adapts to changes in the video image. Low values correspond to a slowly adapting model. High values make the model adapt quickly to scene changes. An example of Background subtraction is shown in Fig. 2(a) and Fig. 2(b).

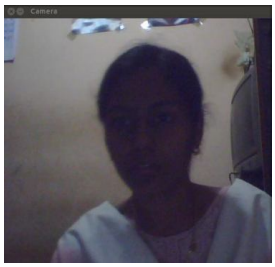


Fig. 2 (a). Input frame

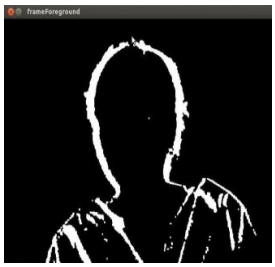
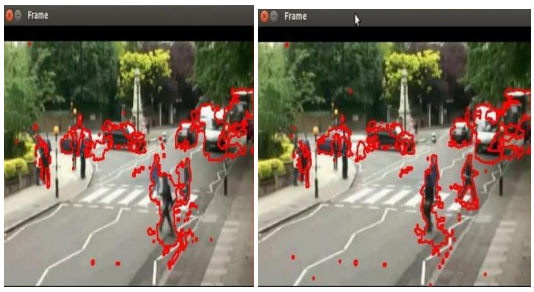


Fig. 2(b). Background subtracted frame

3.2 OBJECT DETECTION

In the proposed approach shown in Fig. 2, the first frame is considered as background. The mixture of Gaussian method is used in background subtraction. Next the background is eliminated and the foreground objects only maintained. At segmentation stage, morphological operators are used. Dilation and erosion are the two morphological operators used for segmenting the object edges here.

Mathematical morphology is used for analyzing object shape characteristics such as size and connectivity, which are not easily accessed by linear approaches. After background subtraction successfully extract out the foreground, it reproduces them into a binary image as seen in Fig. 2(b). To make the objects more recognizable and informative, these objects should be marked and recorded. Contouring is the process of marking out the edge of the object, making it more recognizable and informative. Once background subtraction is done, moving objects are marked using contours. An example of object detection is shown in Fig. 3.



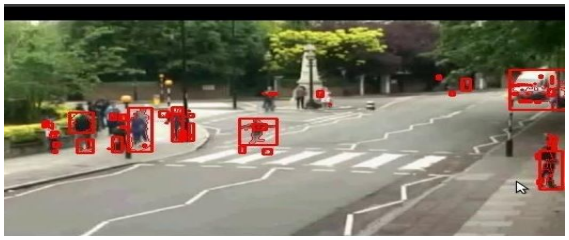
(a)Frame 123 (b)Frame 125

Fig. 3. Moving object such as pedestrian and vehicle detection

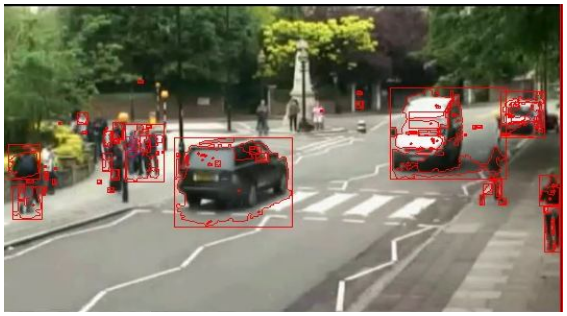
3.3 OBJECT IDENTIFICATION

In computer vision field, object's edge is usually identified as the area of pixels that contrasts with its neighbor pixels or area of moving pixels. After getting the edge of the object, the height of the object can be measured by subtracting the top left pixel position of the rectangle and bottom left pixel position of the rectangle. The width of the object can be measured by subtracting the top left pixel position of the rectangle and top right pixel position of the rectangle. The threshold value set to 1.126 or below for pedestrian and from1.126 to 2.72 for vehicle from the ground truth results. Based on threshold value for the height and width, the object is classified whether it is a pedestrian or vehicle. An example is shown in Fig. 4 (a) and Fig. 4 (b).

Here, some of human are not boxed. These humans are standing for a long time. So these humans are considered as background using the background updating model which is explained in Section 3.



(a)



(b)

Fig. 4. Moving object identification

4. STATIC AND DYNAMIC OBJECT CLASSIFICATION

In this work, object classification is done mainly on 3 classes as pedestrians, vehicles and trees. The proposed approach uses static object detection methods using without background subtraction model is shown in Fig. 5.

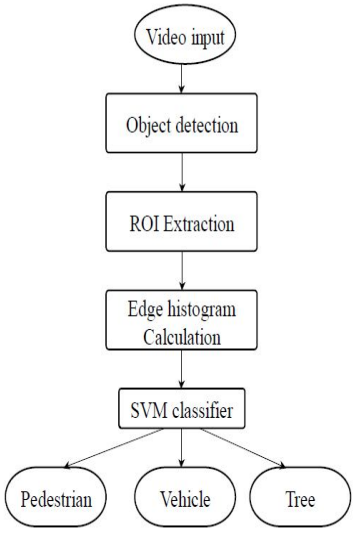


Fig. 5. A model for the static and dynamic object classification

4.1 OBJECT DETECTION

Static Object detection methods such as Edge based detection are used here. By the thresholding intensities in small regions objects are separated from the background.

To aid this process, binary mathematical morphology can be used to aggregate close pixels into a unified object. Thus, the detection of objects can simply consist of the rectangles that enclose the dominant line segments and their neighbors in the image plane. To improve the shape of object regions, extract consistent contour lines and morphological operations to restore small breaks on the detected contours. Thus the presence of objects may be detected by the edge complexity within the road area, which can be quantified through analysis of the histogram in the Section 5.4.

Edge-based object detection is often more effective than other background removal or thresholding approaches, since the edge information remains significant even in variations of

ambient lighting. The various frames for object detection is shown in Fig. 6.

Fig. 6. Static and dynamic object detection from PETS dataset

4.2 ROI EXTRACTION

In this project, the next step is to develop a segmentation strategy to extract the regions of interest (ROI) corresponding to pedestrians, vehicles and trees in the images. This process is called as "bounding box registration", since the ROI is often represented by a bounding box in each video frame. A simple connected-components algorithm is used to label regions in the surveillance video. Each such region

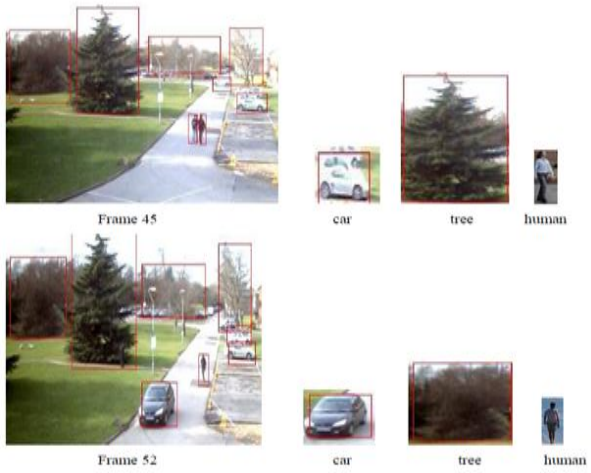


Fig. 7. ROI extraction of car, tree and human

potentially includes both moving and unmoving object. Since, the classification operates on rectangular regions of pixels based on the calculated height and width of the rectangle. The bounding boxes smaller than a given threshold (height < 42 and width < 22 in this experiment) is discarded which as too small to contain any of the target objects.

The OpenCV library function cvFindContours is used for connected-component analysis. For speeding up the tracking process, the region of interest (ROI) is rectangular and corresponds to the bounding box of the object. The example shows that the extracted ROI of pedestrian, vehicle and trees in Fig. 7. The frames 45 and 52 in Fig. 7 are taken from the PETS data set.

4.3. EDGE HISTOGRAM CALCULATION

In this approach, an object view is represented by a discrete set of points sampled regularly along the internal or external contours. For every point, histogram is computed that approximates the distribution of adjacent point locations relative to the reference point. The feature vector can be constructed from edge histogram values of the object. The histograms vectors used are therefore able to capture similarity within a class and at the same time discriminate between different classes.

There is no "best" number of bins, and different bin sizes can reveal different features of the data. Some theoreticians have attempted to determine an optimal number of bins, but these methods generally make strong assumptions about the shape of the distribution. Depending on the actual data distribution and the goals of the analysis, different bin widths may be appropriate. Experimentation is usually needed to determine appropriate bins. In this work, for the same contour edges, various histogram bins such as 6, 16, 32 and 64 are used to calculate the histogram values. The output for these 4 histogram bin values are analyzed and tabulated in Section 5.

4.4. SUPPORT VECTOR MACHINES (SVM)

SVM is based on the principle of structural risk minimization. For linearly separable data, SVM finds the separating hyper plane which separates the data within the largest margin. For, linearly inseparable data, it maps the data in the input space into high dimension space $x \in R^I \mapsto \Phi(x) \in R^H$ with kernel function $\Phi(x)$, to find the separating hyper plane. SVM was originally developed for two class classification problems. The N class classification problem can be solved using N SVMs. Each SVM separates a single class from all the remaining classes (One-vs.-rest approach).

Given a set of frames corresponding to N classes for training, N SVMs are trained. Each SVM is trained to distinguish a class and other classes in the training set. During testing, the class label y of a class x can be determined using:

$$y = \begin{cases} n. \text{ if } d_n(x) + t > 0 \\ 0. \text{ if } d_n(x) + t < 0 \end{cases}$$

Where, dn (x) =max{ di (x)}N i=1, and di (x) is the distance from x to the SVM hyper plane corresponding to frame i, the classification threshold is t, and the class label y=0 stands for unknown. The model for SVM based classification is given in Fig. 8.

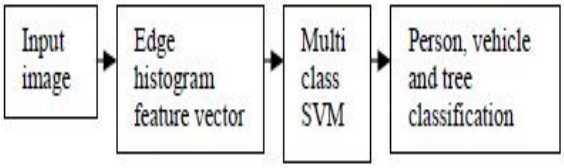


Fig. 8.Work flow of the SVM approach

The Fig. 9 addresses the key issues in the development of SVM trained pattern classifier. The representation and the learning engine. The architectural overview of the system is provided in Fig. 9 as applied to the task of people, vehicle and tree detection and shows the training and testing phases. In the training step, the system takes as input 1) a set of images of the object class that have been aligned and scaled so that they are all in approximately the same position and the same size and 2) a set of patterns that are not in our object class.

An intermediate representation that encapsulates the important information of three object class is computed for each of these patterns, yielding a set of feature vectors. These feature vectors are used to train a pattern classifier to differentiate between in-class and out-of-class patterns. In the testing phase, detecting objects in out-of-sample images are also considered.

5. EXPERIMENTAL RESULTS

5.1. EXPERIMENTAL SETUP

Exhaustive experiments are conducted to evaluate the performance of the proposed approach. Experiments were conducted on Intel Pentium V processor with 2.67 GHz speed with 2GB RAM and the object identification was done in C++ with OpenCV library.

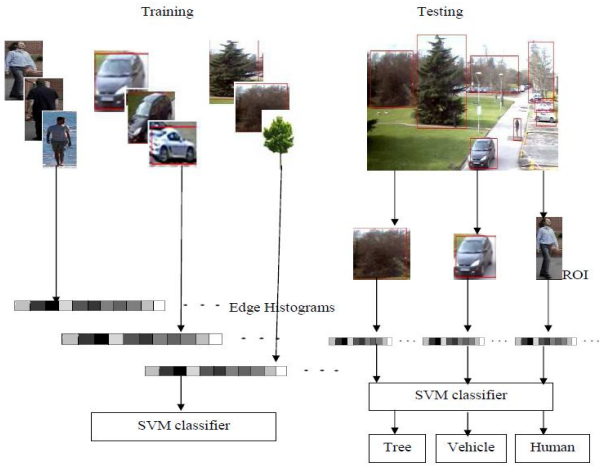


Fig. 9. Training and Testing of SVM classifier

5.2. EXPERIMENTAL DATASETS

To examine the performance of the proposed method, four video set was collected from a PETS dataset is shown in table 1. The video data is split into frames at 20 frames per second with OpenCV library and sampled at a resolution of 320 X 240. The sequences of objects were automatically extracted using background subtraction method and ROI registration method.

Video set	Duration (sec)	Human (sec)	Vehicle (sec)	Tree (sec)
Pedestrian, car and trees	30	526	1121	387
Pedestrian , all type of vehicles and trees	45	734	954	762
Pedestrian , all type of vehicles and trees	60	976	1578	589
Images under illumination changes and different	20	632	584	366

Table 1. Experimental dataset

5.3. PERFORMANCE METRICS

5.3.1. *Performance evaluation of height-width based approach*

In this method, the sample image is first given into height-width based classification method. The aspect ratio is calculated based on the height and width of the bounding box which is given in the Eq. 1. After getting the aspect ratio, which is compared with the threshold. Based on the threshold value, the object is identified as a pedestrian or non-pedestrian. So, care should be taken when selecting the threshold value. The Figure 15, 16 and 17 shows some of the object identified frames.

The aspect ratio is calculated by

$$\text{Aspect ratio} = \frac{\text{width}}{\text{height}} \quad (1)$$

The moving pedestrian alone detected and identified in Fig. 10. Pedestrian which are boxed shows that they are moving in a road. The remaining pedestrians are static for a particular system. So, these pedestrians are not identified by this system. The Fig. 11 contains both the pedestrian and vehicle. The bounding box for the vehicle is considerably high than the bounding box of the pedestrian. From these variations, the classification is done. The height, width and aspect ratio values for both human and vehicle are shown in Fig. 12.

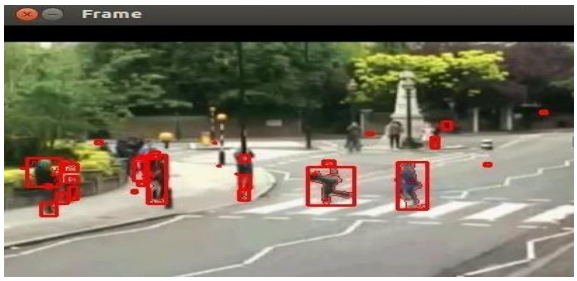


Fig. 10. Height-width based approach contains only pedestrian

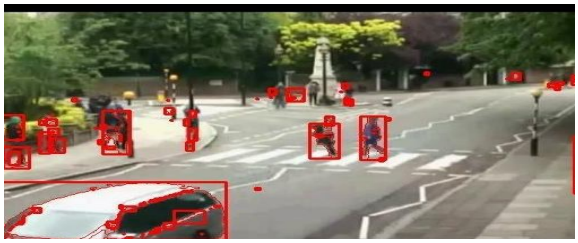


Fig. 11. Height-width based approach contains both pedestrian and vehicle which is nearer to camera

Fig. 12. Height-width based approach contains both pedestrian and vehicle which is far away from camera

This table shows the height, width and aspect ratio (calculated based on the Eq. 1). The threshold value set to below and equal to 1.126 for pedestrian and from 1.126 to 2.72 for vehicle by trial and error after conducting several experiments. The highlighted box shows the threshold values such as 1.8, 1.7419, 1.64255, 1.81667 and 1.5617 which values fall between 1.126 and 2.72. So these boxes are considered as vehicles. Remaining values are below the 1.126 which are categorized as pedestrian.

In this work, from the ground truth results, the performance can be evaluated by true positive rate and false positive rate. This quantity measures recall, precision and localization. Based on the equation (2) and (3) the area of the ROC curve is measured

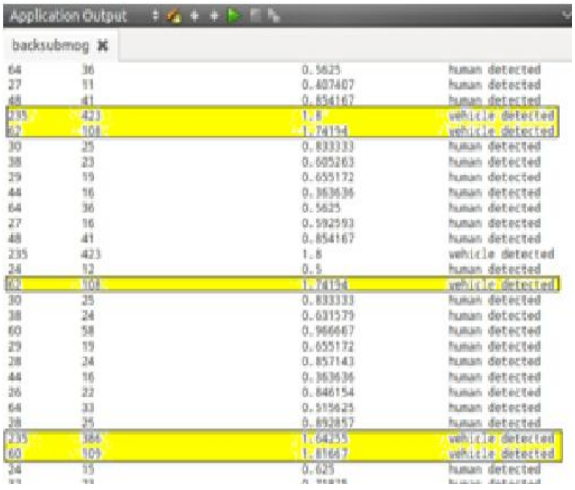


Fig. 13. Object Classification done under height-width approach

The true positive rate TPR is the percentage of non occluded objects in the camera's view that are detected. TPR is assessed by dividing the number of truly detected objects by the total number of objects.

TPR is defined by

$$TPR = \frac{\text{detected objects}}{\text{Total number of objects}} \quad (2)$$

The false positive rate FPR is the proportion of detection that were not true objects. The FPR is assessed by dividing the number of false positives by the total number of detections. This is the percentage of erroneous detection. FPR is a measure of precision and localization. It is defined by

$$FPR = \frac{\text{false positives}}{\text{detected objects} + \text{false positives}} \quad (3)$$

Table 2.Performance of height-width based approach

Height-width based classification method	TPR (%)	FPR (%)
Pedestrian only	91.1	6.8
Both pedestrian and car	90.9	7.8
Both pedestrian and all other types of vehicles	91.3	11.2

The table 2 shows the Average true positive ratio and false positive ratio for 25 frames that containing only pedestrian, 50 frames that containing both pedestrian and only car and 75 frames contain both pedestrian and all types of vehicle.

5.3.2. PERFORMANCE EVALUATION OF STATIC-DYNAMIC OBJECT CLASSIFICATION METHOD

The training and testing is done with the four video set represented with its duration in Table2. Three dataset is generated dynamically at the execution of each video set. Each dataset contains thousands of images. These videos are sampled at a resolution of 320X 240 at 20 frames per second. These video set contains 2034, 2450, 3143 and 1582 images respectively. The number of human, vehicle and trees are tabulated for each video set in Table 1.

5.3.2.1. Training set

The training set consists of 10000 positive images and the class 1, 2 and 3 consists of 3500, 4000 and 2500 respectively. The images were originated from 3000 properly cropped object's front, rear and side view images. Since the vehicles' rear views are approximately symmetrical, the images were horizontally flipped to generate the 10000 positive training samples. The non-symmetric objects were also included in the training set to make the classifier more robust to remove any false positives picked up by the processing stage. The classifier's model created from the training is then tested on the evaluation set.

5.3.2.2. Evaluation Set

This is an independent set of 3 classes of images not used in the classifier training. They are reserved for evaluating the performance of the generated classifier and the estimation of operating points for the precision-recall plot. Video 1 contains pedestrian and vehicle of all types. This set contains 2034 positive images. Video 2 and 3 contains pedestrian, vehicle of all types and trees. This set contains 5593 positive images. Video 4 contains 1582 images under illumination changes and different poses.

5.3.3. PERFORMANCE OF FEATURE VECTOR USING DIFFERENT NUMBER OF HISTOGRAM BINS

The number of histogram bins determines the fineness of the details captured by the features. Having too many bins may capture a lot of irrelevant features originated from the image's noise or the background clutter.

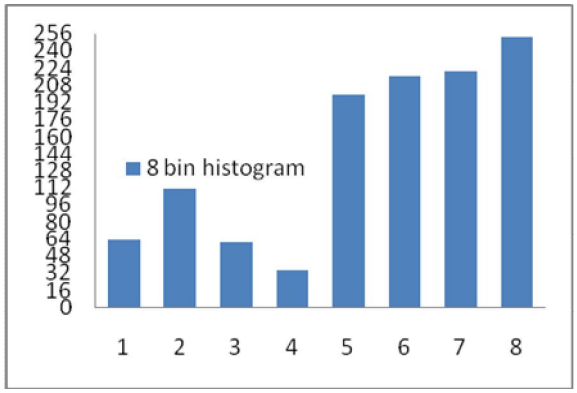


Fig. 14. 8 bin histogram

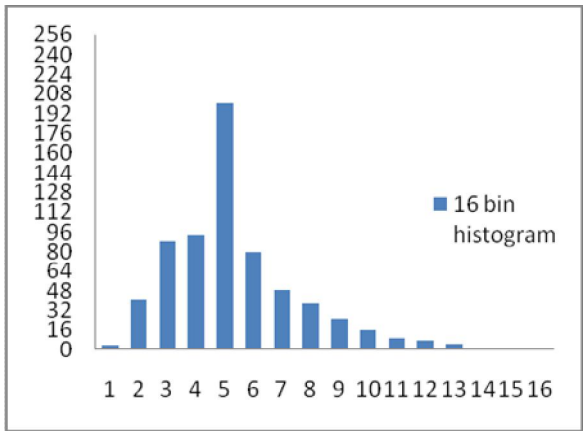


Fig. 15. 16bin histogram

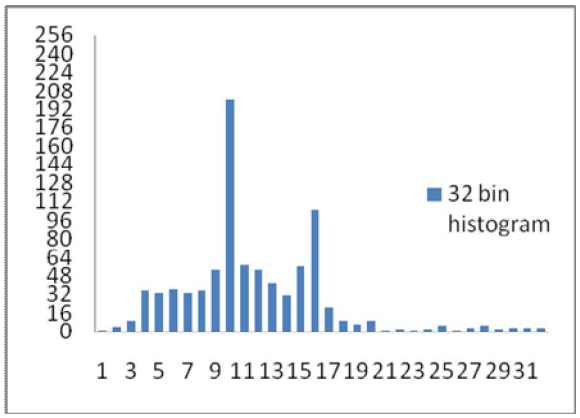


Fig. 16. 32 bin histogram

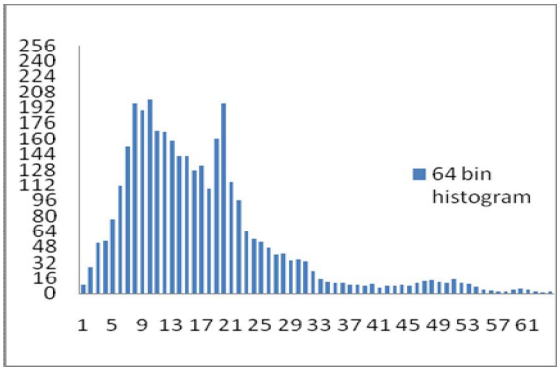


Fig. 17. 64 bin histogram

It will also slow down the classification process. On the other hand, using too few histogram bins may miss out some important object's features and thus degrade the classification performance. In this experiment, the performance of the edge features with 8, 16, 32 and 64 histogram bins were evaluated which is shown in Fig. 14, Fig. 15, Fig. 16 and Fig. 17.

To evaluate the method, True positive (TP), False positive (FP) and False negative (FN) are used. In a classification task, the precision for a class is the number of true positives (i.e. the number of items correctly labeled as belonging to the positive class) divided by the total number of elements labeled as belonging to the positive class (i.e. the sum of true positives and false positives, which are items incorrectly labeled as belonging to the class). Recall in this context is defined as the number of true positives divided by the total number of elements that actually belong to the positive class (i.e. the sum of true positives and false negatives, which are items which were not labeled as belonging to the positive class but should have been).

The terms positive and negative refer to the classifier's prediction (sometimes known as the expectation), and the terms true and false refer to whether that prediction corresponds to the external judgment (sometimes known as the observation). Statistical metrics such as Precision, Recall, Accuracy and F-measure (F) gives an overall performance for this method which is explained in Equation (4), (5), (6) and (7).

Precision (P) is the proportion of the predicted positive cases that were correct. High precision means that an algorithm returned substantially more relevant results than irrelevant.

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Recall (R) is the proportion of positive cases that were correctly identified. High recall means that an algorithm returned most of the relevant results.

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

A measure that combines precision and recall is the harmonic mean of precision and recall, the traditional F-measure or balanced F-score:

$$F_{\alpha} = \frac{2PR}{P+R} \quad (6)$$

Accuracy (AC) is the proportion of the total number of predictions that were correct. The accuracy is the proportion of true results (both true positives and true negatives) in the population.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (7)$$

The above performance metrics can be applied for the evaluation test video using different histogram bins and the performance is tabulated. Table 3 shows that the precision, recall, F-measure value and accuracy for video 1 at different histogram bin. Table 4 shows that the precision, recall, F-measure value and accuracy for video 2 at different histogram bin. Table 5 shows that the precision, recall, F-measure value and accuracy for video 3 at different histogram bin. Table 6 shows that the precision, recall, F-measure value and accuracy for video 3 at different histogram bin. From the values of these four tables, it can be seen that features with higher numbers of bins perform better.

Table 3. Performance of pedestrian, car and trees as seen in the dataset given in Section 6.2 for various histograms

Number of Histogram bins	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
8	89	87.23	88.11	90.25
16	90.12	89.12	89.62	91
32	91.12	90.22	90.67	91.83
64	91.66	90.76	91.20	92.93

Table 4. Performance of pedestrian, all types of vehicles and trees as seen in the dataset given in Section 6.2 for various histograms

Number of Histogram bins	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
8	89.78	87.43	88.589	91.5
16	93.16	91.25	92.198	93.12
32	93.78	92.72	93.24	94
64	95.3	94.1	94.49	95

Table 5. Performance of pedestrian, all types of vehicles and trees as seen in the dataset given in Section 6.2 for various histograms

Number of Histogram bins	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
8	90	89	89.49	92.5
16	94	92	92.98	94
32	96.7	92.8	94.71	94.83
64	98	93.1	94.97	96

Table 6. Performance of pedestrian, all types of vehicles and trees under illumination changes and different poses as seen in the dataset given in Section 6.2 for various histograms

Number of Histogram bins	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
8	69.78	68.54	69.15	67.54
16	72.34	70.23	71.29	72.25
32	75.58	72.65	74.08	75.25
64	84.43	82.12	83.25	84.22

The confusion matrix shows how the predictions are made by the model. The rows correspond to the known class of the data, i.e. the labels in the data. The columns correspond to the predictions made by the model. The value of each of element in the matrix is the number of predictions made with the class corresponding to the column for examples with the correct value as represented by the row. Thus, the diagonal elements show the number of correct classifications made for each class, and the off-diagonal elements show the errors made.

Table 7. Confusion matrix for pedestrian, car and trees as seen in the dataset given in Section 6.2 with the 64 histogram bins

Class	Pedestrian (%)	Vehicle (%)	Tree (%)
Pedestrian	91.67	6.67	1.66
Vehicle	3.181	95.45	1.36
Tree	5	3.33	91.67

Table 8. Confusion matrix for pedestrian, all types of vehicles and trees as seen in the dataset given in Section 6.2 with the 64 histogram bins

Class	Pedestrian (%)	Vehicle (%)	Tree (%)
Pedestrian	95	1	4
Vehicle	3.34	95.34	1.32
Tree	0.66	4.66	94.68

Table 9. Confusion matrix for pedestrian, all types of vehicles and trees as seen in the dataset given in Section 6.2 with the 64 histogram bins

Class	Pedestrian (%)	Vehicle (%)	Tree (%)
Pedestrian	95.625	2.5	1.875
Vehicle	2.245	96.11	1.645
Tree	2.13	1.39	96.48

Table 10. Confusion matrix for pedestrian, all types of vehicles and trees under illumination changes and different poses as seen in the dataset given in Section 6.2 with the 64 histogram bins

Class	Pedestrian (%)	Vehicle (%)	Tree (%)
Pedestrian	84.12	8.46	7.42
Vehicle	8	83.98	8.02
Tree	7.88	7.56	84.56

This table 7, table 8, table 9 and table 10 shows that the confusion matrix for the three classes of the evaluation video sets given in the table 1 with 64 histogram bins. The average accuracy for these videos are 92.93, 95.1, 96.07 and 84.22 respectively. The overall performance of the system is 92.08.

The Fig. 21 shows the graph which is drawn between the F-measure and time at seconds. For each video set mentioned in the table 1, the curve is drawn. In this graph, F-measure value is saturated between 2 and 3 seconds. From the result, the video set above 2 seconds gives better performance of the object identification.

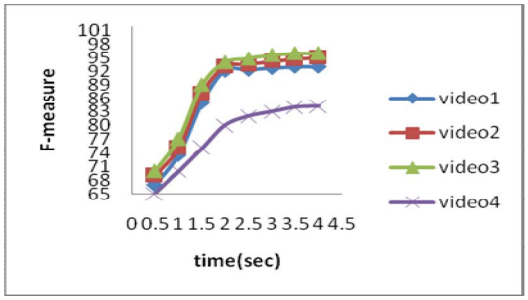


Fig. 21. Performance of the F-measure value for the 4 video sets

6. CONCLUSION

In height-width based classification method, background registration technique, segmentation using morphological operator and identification using height and width of objects are presented. The system effectively combines simple domain knowledge about object classes with time domain statistical measures to identify target objects in the presence of partial occlusions. The computational complexity of this algorithm is linear in the size of a video frame and the number of objects detected.

In static-dynamic object classification method, object detection is done without background elimination, segmentation using bounding box registration technique and the classification is done with the multiclass svm. A system has been developed to detect and identify for both dynamic and static objects on highways efficiently. This paper is mainly used to control the increasing traffic-flow on highways and to meet safety and security standards. For Accident avoidance, static object classification is needed with the dynamic object classification. As we have considered traffic on highways there is no question of shadow of any cast such as trees but sometimes due to occlusions two objects are merged together and treated as a single entity.

REFERENCES

1. Journal Papers

[1] Sayanan sivaraman, mohan manubhai trivedi, " A General Active Learning Framework for on-road vehicle Recognition and tracking" in Transactions on Intelligent Transportation systems, vol no 11, issue 2, pages 267-272,June 2010 IEEE.

[2] P. M. Daigavane and P. R. Bajaj, " Real Time Vehicle Detection and Counting Method for Unsupervised Traffic Video on Highways", International Journal of Computer Science and Network Security, VOL.10 No.8, pages112-117August 2010.

[3]Rajiv Kumar Nath, swapan kumar Deb, " On Road Vehicle /Object Detection and Tracking using Template", on Indian Journal of Computer Science and Engineering vol. No .2 pages 98-107,2009.

[4] C. Wang, J. Lien, "Automatic vehicle detection using local features—A statistical approach," IEEE Transactions on Intelligent Transportation Systems, vol. 9,Issue no. 1, pp. 83–96, Mar. 2008.

[5] A.Elgammal R.Duraisamy,D.Harwood and Larry S.Devis ," Background & Foreground modeling using non parametric Kernal density estimation for visual surveillance",on proceeding of the IEEE, vol.90,no 7, pages 1151 - 1163 july 2002.

2. Text Book

[6] The OpenCV Reference Manual Release 2.4.2 published on july 2012.

3. Conference Proceedings

[7]F.Dornaika,san Sebastian, Bilbao, P.Chakik,"Efficient object detection and Matching using Feature classification on Pattern Recognition",20th International Conference on Pattern Recognition(ICPR), 2010 pages 3073 – 3076.

[8] S.Schraml, A.N.Belbachir, N.Brandle,"A Real-Time Pedestrian Classification Method for Event-Based Dynamic Stereo Vision" in IEEE 2010.

[9] G. Jun, J. K. Aggarwal, and M. Gokmen, "Tracking and segmentation of highway vehicles in cluttered and crowded scenes", in IEEE Workshops on Applications of Computer Vision, pages 1-6,2008.

[10] Zhenjun Han, Qixiang Ye, and Jianbin Jiao," Online feature evaluation for object tracking using Kalman Filter" in 19th International Conference on Pattern Recognition, 2008. (ICPR2008), pages 1 –4, 2008.

[11] X. Song and R. Nevatia," A model-based vehicle segmentation method for tracking" IEEE Proceedings of the 10th International Conference. on Computer Vision, pages 1124 - 1131 Vol. 2 2005.

[12] N Dalal, B Triggs,"Histograms of Oriented Gradients for Human Detection" in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 708 - 713, 2013, ISSN: 2319–8656

Occlusion and Abandoned Object Detection for Surveillance Applications

M. Chitra
RVS college of Engineering and Technology
Karaikal, India

M.Kalaiselvi Geetha
Annamalai University
Chidambaram, India

L.Menaka
RVS college of Engineering and Technology
Karaikal, India

Abstract: Object detection is an important step in any video analysis. Difficulties of the object detection are finding hidden objects and finding unrecognized objects. Although many algorithms have been developed to avoid them as outliers, occlusion boundaries could potentially provide useful information about the scene's structure and composition. A novel framework for blob based occluded object detection is proposed. A technique that can be used to detect occlusion is presented. It detects and tracks the occluded objects in video sequences captured by a fixed camera in crowded environment with occlusions. Initially the background subtraction is modeled by a Mixture of Gaussians technique (MOG). Pedestrians are detected using the pedestrian detector by computing the Histogram of Oriented Gradients descriptors (HOG), using a linear Support Vector Machine (SVM) as the classifier. In this work, a recognition and tracking system is built to detect the abandoned objects in the public transportation area such as train stations, airports etc. Several experiments were conducted to demonstrate the effectiveness of the proposed approach. The results show the robustness and effectiveness of the proposed method.

Keyword: Occlusion, Histograms of Oriented Gradients descriptors, Support Vector Machine, mixture of Gaussians techniques, Blob, abandoned object.

1. INTRODUCTION

Surveillance is monitoring the behavior, activities for the purpose of influencing, managing, directing, or protecting. One of the main advantages of video surveillance is that it could be used both as a preventive mechanism and a forensic tool for crimes. Unfortunately, constant surveillance of public domains is challenging and labor-intensive. Therefore, improved automated abandoned object detection and occlusion detection systems are increasingly in demand.

Robustness of object detection is affected by occlusion in the presence of multiple objects. Occlusion is the hiding of an object by another during multiple object tracking. In video sequences Fig 1(a), occlusions create several challenges to the tracking algorithm. The task of reliable detection and tracking of multiple objects becomes highly complex for crowded scenarios. The detection of suspicious object is one of the most important task in video surveillance. Suspicious objects are generally unattended packages left in public places. This work aims to detect occlusions and abandoned object detection which captures most of the researches working in video surveillance. Object detection in monocular image sequences still suffers from a lack of robustness due to temporary occlusions, objects crossing and changing lighting conditions.

Abandoned objects are detected by difference image. An example for occluded scene and abandoned object/baggage is shown in Fig. 1(b).

(a). Occluded frame (b). Abandoned object

Figure. 1 Occluded and abandoned frames

The rest of the paper is arranged as follows: Section 2 discusses a few approaches explored by other researchers to solve the problem. Section 3 explains the technical details of our method, Section 4. Section 5 wraps up the paper with an experimental result and conclusion of this paper.

2. RELATED WORKS

Although many algorithms have been proposed in the literature, the problem of multiple interacting objects tracking in complex scene is still open. Methods to solve the occlusion problem in multiple interacting objects tracking have been previously presented in [8, 12, 15, 16]. Sensor based occlusion tracking methods [8, 18] cannot handle the difference between moving and nonmoving obstacles. In general, solutions to both pedestrian and vehicle occlusion problems extend from spatial to temporal domains can be broadly classified into many types. Some types like active contour –based [2], stereo vision-based, region based, model based [16], feature based [12] are also seen in the literature.

Object tracking in [17] overcome occlusion by fusing multiple camera inputs, but it cannot handle complete occlusion.

Several methods using color, texture and motion [13, 14, 19]. [14] Performs robust tracking that deals some instances of occlusion but it cannot handle illumination changes.

Histograms of Oriented Gradients (HOG) and Local Binary Pattern (LBP) [9] are used to detect partial occlusion but it cannot handle the articulated deformation of people. Multi-cue model widely used in tracking and detection systems [14] has achieved considerable success in object tracking because of its simplicity and robustness. However, color features cannot give good performance when an object and its background have the similar colors. Contour-based object detection and recognition method proposed in [2, 4]. [7,10] uses filter based detection and tracking.

Most of the proposed techniques for abandoned object detection rely on tracking information [24, 25] to detect drop-off events, while fusing information from multiple cameras. As stated by Porikli [26], these methods are not well suited to complex environments like scenes involving crowds and large amounts of occlusion. Stauffer and Grimson [27] present an event detection module that classifies objects, including abandoned objects, using a neural network, but is limited to detecting only one abandoned object at a time. In [28] a multi-camera video surveillance system is proposed to detect the owner of each abandoned object is determined and tracked using distance and time constraints and multiple cameras with overlapping fields of view are exploited to cope with occlusion of various types.

3. OCCLUSION AND ABANDONED OBJECT DETECTION

3.1 Occlusion Detection

Manual surveillance requires the system to be monitored by personnel and is expensive and inconvenient. Hence, efforts have been put into automatic surveillance. One of the main components of automatic video surveillance is detecting occluded objects. Fig.2 shows the proposed system architecture for occlusion detection.

Occlusion is the main cause for performance degradation in surveillance systems. Under occlusion, the objects will become overlapped and may be found moving together in a scene. Occlusion can be of three types. Self occlusion, inter-object occlusion, background occlusion. The goal of our work is to develop a general framework to detect and track objects with persistent and occasional heavy occlusion. Background subtraction is carried out by a mixture of Gaussians techniques (MOG) is in Fig.3. Blob detection for detecting occluded object is carried out in OpenCV. Multiple occluded object detection involves computing the Histogram of Oriented Gradients descriptors (HOG) and linear Support Vector Machine (SVM) as a classifier.

3.1.1 Mixture of Gaussian Algorithm(MOG)

Pixelprocesses–

$$\{X_1,\ldots\ldots,X_T\} = \{I(x_0, y0, i): 1\leq i \leq t\},$$

Where, I- image sequence, t- time, {x0,y0}- the history of pixels.

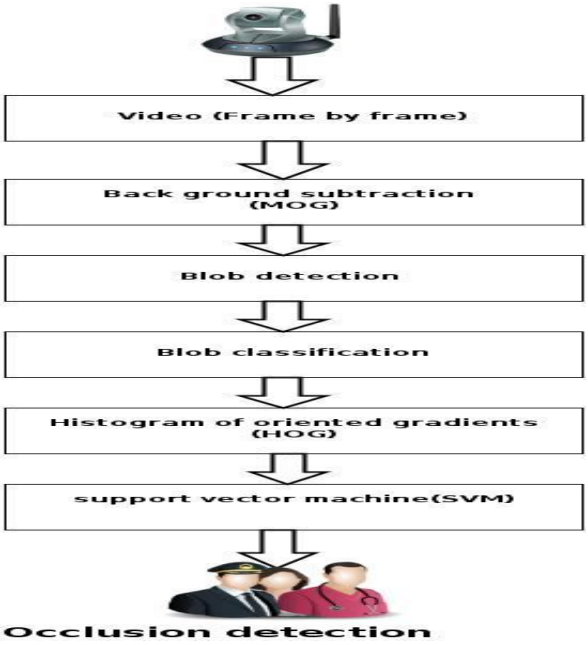


Figure. 2 Occlusion detection

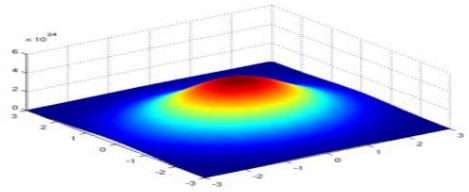


Figure. 3 Mixture of Gaussians algorithms

Model the values of a particular pixel as a mixture of Gaussians. We determine which Gaussians may correspond to background colors-Based on the persistence and the variance of each of the Gaussians [23]. Pixel values that do not fit the background distributions are considered foreground until there is a Gaussian that includes them. Update the Gaussians. Pixel values that do not match one of the pixel's "background" Gaussians are grouped using connected components.

Background modeling - constructs a reference image representing the background.

Threshold selection- determines appropriate threshold values used in the subtraction operation to obtain a desired detection rate.

Pixel classification - classifies the type of a given pixel, i.e., the pixel is the part of background (including ordinary background and shaded background), or it is a moving object.

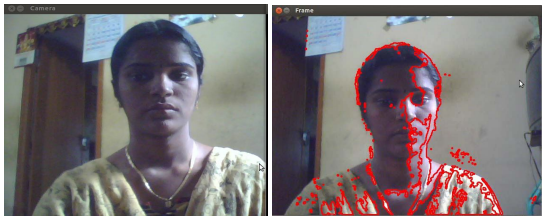


Figure. 4 Background subtraction using MOG

3.1.2 Blob Detection

Once the background subtraction is done with MOG, each foreground object is detected as a blob. Each module is applied on a bi-level image obtained though a mixture of Gaussian background subtraction and some basic filtering performed. Using OpenCV [22]. A tracked blob is considered to be occluded if its major region is covered by foreground and it should continue to be tracked if either it is occluded or its area and centroid is matched with any of the blobs. When two objects pass close to each other, they are detected as a single blob. Fig.5 shows the blob detection.

Often, one object will become occluded by the other one. One of the challenging problems is to maintain correct labeling of each object after they split again. Fig.6 demonstrates the blob merging and splitting at the time of occlusion.

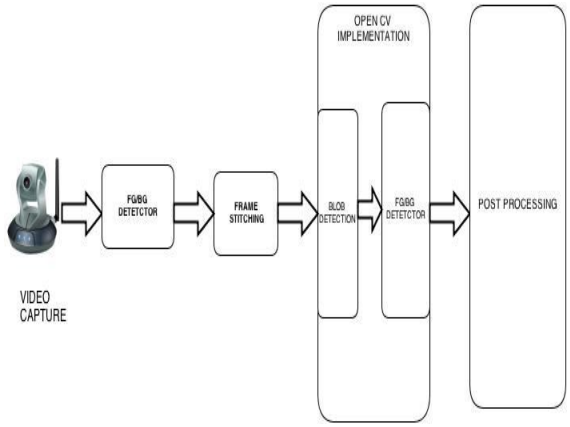


Figure. 5 Blob Detection and Tracking

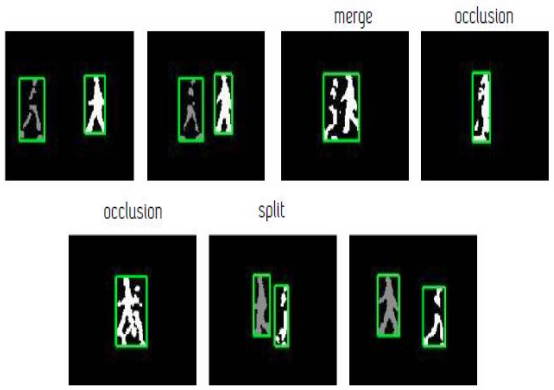


Figure. 6 Blob Merging and splitting at the time of occlusion

3.1.3 Histogram of Oriented Gradients descriptors (HOG)

Histograms of Oriented Gradients (HOG) are feature descriptors used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy. Fig.7 shows orientation of each gradient sample rotated relative to key point orientation.

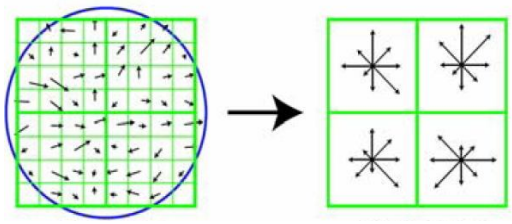


Image gradients Key point descriptor

Figure. 7 Histograms of Oriented Gradients

For normalization process all blocks are typically overlapped and rotation invariance. It is typically combined with SVM classifier.

3.1.4 SVM Classifier

The Support Vector Machine classifier is a binary classifier which looks for an optimal hyper plane as a decision function. Once trained on images containing some particular object, the SVM [20] classifier can make decisions regarding pedestrians. SVM is based on the principle of structural risk minimization. For linearly separable data, SVM finds the separating hyper plane which separates the data within the largest margin. For, linearly inseparable data, it maps the data in the input space into high dimension feature space $x \in \mathbb{R}^I \rightarrow \Phi(x) \in \mathbb{R}^H$ With kernel function $\Phi(x)$, to find the separating hyper plane. SVM was originally developed for two class classification problems. The N class classification problem can be solved using N SVMs. Each SVM separates a single class from all the remaining classes (One-vs.-rest approach).Given a set of frames corresponding to N classed for training, N SVMs are trained. Each SVM is trained to distinguish a class and other classes in the training set. During testing, the class label y of a class x can be determined using:

$$y=\begin{cases} n, \text{if } d_n(x) + t > 0 \\ 0, \text{if } d_n(x) + t \leq 0 \end{cases}$$

Where, $d_n(x) = \max\{d_i(x)\}^N_{i=1}$, and $d_i(x)$ is the distance from x to the SVM hyper plane corresponding to frame i, the Classification threshold is t, and the class label y=0 stands for unknown.

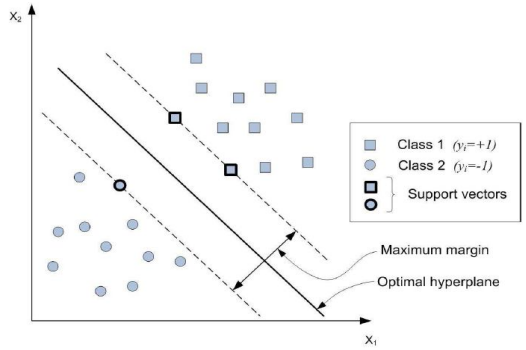


Figure. 8 Blob Optimal separating hyper plane and margin for a two dimensional feature space

3.2 Abandoned Object Detection

The detection of suspicious (dangerous) object is one of the most important task in video surveillance. An abandoned object not belonging to the background scene and remaining in the same position for a long time. Our system for abandoned object detection was designed to assist operators surveilling indoor environments, such as airports, railway, or metro stations etc. Fig 9 shows the abandoned object detection.

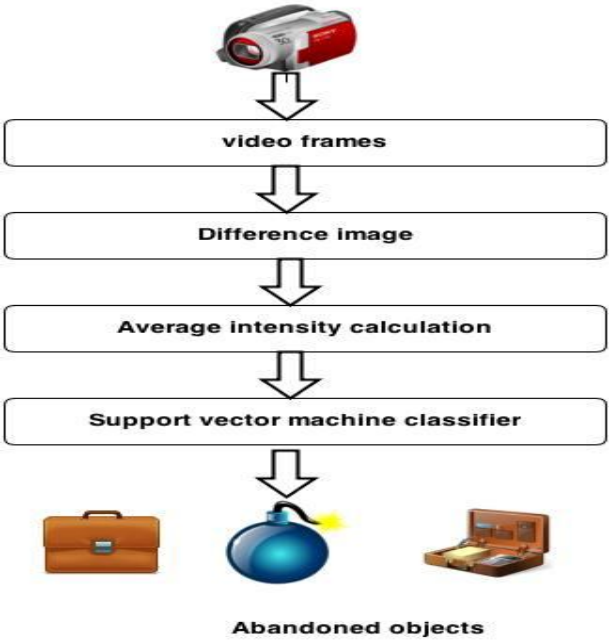
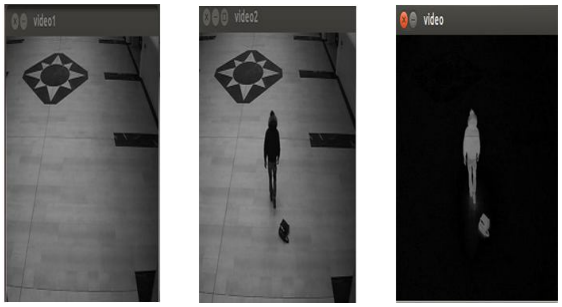


Figure. 9 Abandoned Object Detection

3.2.1 Difference Image

The difference image is defined as the residual image that results from subtracting the scenes at two consecutive time instants from each other. Difference images capture target motion in the scene and can be used for target tracking. Fig. 10 illustrates the basic idea behind difference images. Fig. 10(a), 10(b) show a scene with moving targets at two consecutive time instants. Fig. 10(c) shows the difference image resulting from subtracting the two image frames.



(a) (b) (c)

Figure. 10 (a). Image x_1 at time instance t_1 (b). Image x_2 at time instance t_2, and (c). Difference image.

4. EXPERIMENTAL RESULTS

Experiments were carried out on a PC with Intel IV cores 2.67GHz processor with C++ and the OpenCV library. The above approaches are implemented and experiments are performed on video clips freely available and PETS 2006 .The detection rate of proposed methods are shown in Table 1. The video data are processed at 20fps, and 320x240 resolution. We present some snapshots for occlusion detection and abandoned object detection in Fig.13, 14.

4.1 Occlusion detection

4.1.1 Occlusion degree and detection rate

The occlusion degree is calculated by the number of moving pedestrians and moving vehicles present in the frame with occlusion. For example if a video sequence contains 10 pedestrians, in which 5 pedestrians are occluded with each other, then the calculated occlusion degree is 50 %. More than 10 persons is considered as high, more than 5 as medium less than 5 as low occlusion degree. The occlusion rate is compared with the ground truth information.Fig.11 represents the graphical representation of occlusion detection.

4.1.2 Ground truth method

The ground-truth data allows us to do a quantitative comparison between occlusion detection method and the facts that are confirmed in an actual field. The ground truth method calculation is carried out manually.

Table 1. Detection rate of proposed method

Occlusion degree (%)	High	medium	low
Detection rate	80	90.3	91.2
Ground truth method	90	95	100

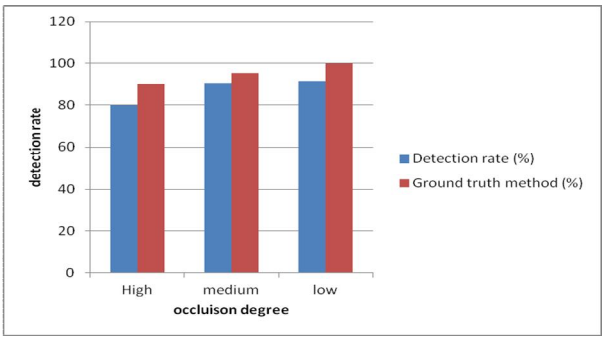


Figure. 11 Graphical reprsentation of our detection rate.

4.2 Abandoned Object Detection

4.2.1 Intensity Calculation of Difference Image

The intensity of the difference images are taken as an input to the SVM classifier. Fig. 12 shows the graphical representation of difference images intensity values of 346 difference images from the abandoned object dataset.

Table 2. Data set for abandoned object detection

S.NO	Dataset or video name	Video length
1	Abandoned object	59
2	Video 1	30
3	PETS 2006	60

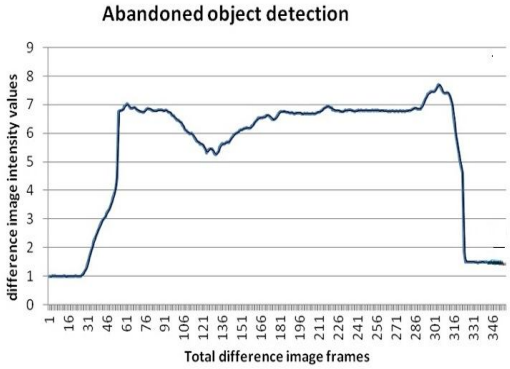


Figure. 12 Intensity values of the difference image in abandoned object detection

4.2.2 Results with SVM Classifier

The Support Vector Machine classifier is a binary classifier which looks for an optimal hyper plane as a decision function. Once trained on images containing some particular object, the SVM classifier can make decisions regarding the presence of an object, such as a human being. Classification results using quadratic discriminant analysis were found to be quite accurate. A training set of 60 cases, where 30 of them belong to one class - bag, and the rest belong to the second class - non-bag(people) is used. Using a test set of 50 cases, gave us an accuracy of 91%. Fig. 28 shows the SVM training and testing.

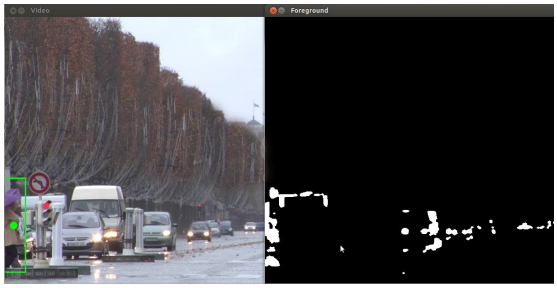


Figure. 13 (a). Occlusion detetion in frame 1 at video 4

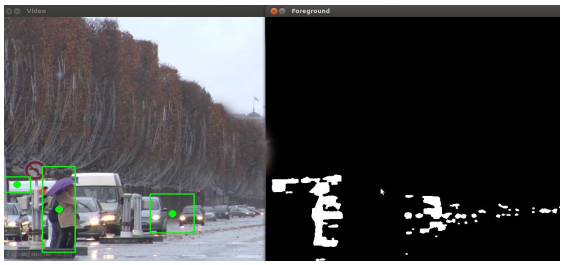
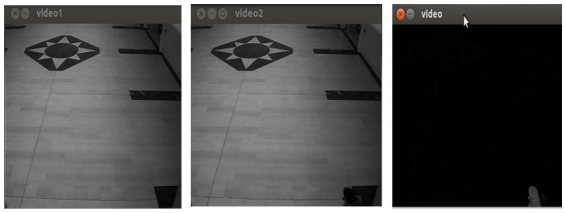
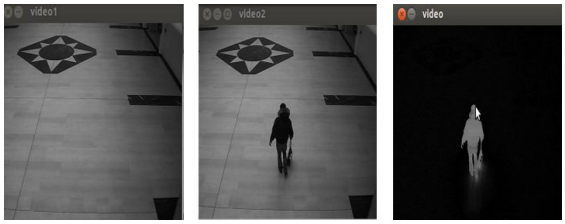


Figure. 13 (b). Occlusion detetion in frame 1 at video 10

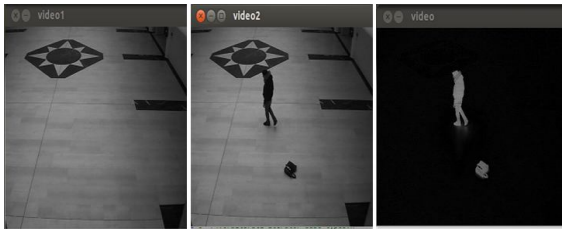
Figure. 13 Occlusion detection in random frame number.



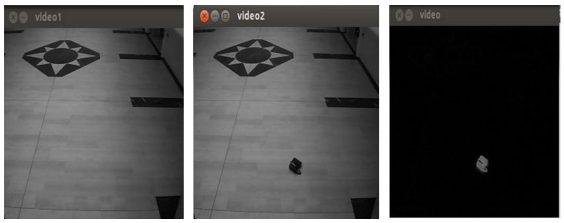
(a) Frame differencing done with first and 4th frame



(b) Frame differencing done with first and 12th frame



(c)Frame differencing done with first and 18th frame



(d) Frame differencing done with first and 26th frame.

Figure. 14 Abandoned object detection using frame differencing at frame 4, 12, 18, 26.

5. CONCLUSION

This research work will lay a stepping stone for the further developments of the automatic object tracking system in a secured area. A real-time multiple occluded objects detection and abandoned object detection system is presented. Experiments on complex indoor and outdoor environments show that the system can deal with difficult situations such as ghosts and illumination changes. Moreover, it can track multiple objects with long-duration and complete occlusion. While the system is highly computationally cost effective and

accurate. The testing results which are based on different scenarios have proved that our approach can be successfully applied in real world surveillance applications. Future work includes extending multiple camera view occlusion detection and abandoned object detection.

6. REFERENCES

[1] Pallavi S. Bangare, Nilesh J. Uke, Sunil L. Bangare, " An Approach for Detecting Abandoned Object from Real Time Video", International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 3, 2012.

[2] Geo Guo,Ting Tiang, Yizhou wang and Wen Gao, "Recovering missing Contours for occluded Object detection", IEEE Signal processing letters, Vol.19,No.8, Aug 2012.

[3] A.A.shafie,M.H Ali,Fadhlan Hafiz and Rosilizer M.Ali, "Smart video Surveillance System for Vehicle Detection and traffic Flow control", Journal of Engineering Science and technology, Vol. 6, No. 4, 2011.

[4] Tianyang Ma and Longin Jan Latecki, "From Partial Shape Matching through Local Deformation to Robust Global Shape Similarity for Object Detection", CPVR 2011.

[5] Manmohan Singh Rawat, Varun N.S, Vinotha S.R, "Occlusion Detection and Handling In Video Surveillance", 2011 International Conference on Future Information Technology, IPCSIT vol. 13 2011.

[6] Kung –Chung Lee,Anand Oka, Emmanual Pollakis and Lutz Lampe, " A Comparison between unscented Kalman filtering and Particle filtering for RSII-BasedTracking," Positioning Navigation and Communication (WPNC), 2010.

[7] Breitenstein, M.D., Reichlin, F., Leibe, B., Koller-Meier, E., Van Gool,L., " Robust tracking-by-detection using a detector confidence particle filter", IEEE International Conference on Computer Vision ,2009.

[8] Michael S. Darms, Paul E. Rybski, Christopher Baker, and Chris Urmson, "Obstacle Detection and Tracking for the Urban Challenge", IEEE Transaction on Intelligent Transportation systems,Vol. 10, No. 3, Sep 2009.

[9] Xiaoyu Wang, Tony X. Han, Shuicheng Yan, "An HOG-LBP Human Detector with Partial Occlusion Handling", Computer Vision, IEEE 12th International Conference, 2009.

[10] Junliang Xing, Haizhou Ai, Shihong Lao,"Multi-object tracking through occlusions by localtracklets filtering and global tracklets association with detection responses", Computer Vision and Pattern Recognition, 2009. CVPR 2009.

[11] Dollar, P., Wojek, C. ., Schiele, B. , Perona, P. ,"Pedestrian detection: A benchmark", Computer Vision and Pattern Recognition, . CVPR 2009.

[12] Pang, C.C.C. ; Lam, W.W.L. ; Yung, N.H.C. , "A Method for Vehicle Count in the Presence of Multiple-Vehicle Occlusions in Traffic Images", IEEE Transactions on Intelligent Transportation Systems, Sept. 2007.

[13] Junqiu WANG and Yasushi YAGI, "Integrating Color and Shape-Texture Features for Adaptive Real-time Object Tracking", IEEE Trans on. Image Processing, 2007.

[14] Valtteri Takala and Matti Pietik¨ainen, "Multi-Object Tracking Using Color, Texture and Motion", IEEE Trans on. Image Processing, 2007.

[15] Tao Yang, Stan Z.Li, Quan Pan, Jing Li, "Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes", Foundation of National Laboratory of Pattern Recognition and National Natural Science Foundation of China (60172037) , 2005.

[16] Clement Chun Cheong Pang, William Wai Leung Lam, Nelson Han Ching Yung, "A Novel method for Resolving vehicle occlusion in a monocular Traffic-Images Sequences", IEEE Transaction on Intelligent Transportation systems, Vol. 5, no. 3, sep 2004.

[17] Tao Zhao, Ram Nevatia "Tracking Multiple Humans in Crowded Environment", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)

[18] MichaelBramberger, RomanPflugfelder, BernhardRinner, Helmet Schwabach and Bernhard Strobl, "Intelligent Traffic Video Sensor : Architecture and Applications", IEEE 2003.

[19] P.Perez, C.Hue, J.Vermaak, and M.Gangnet "Color-Based Probabilistic Tracking", A. Heyden et al. (Eds.): ECCV 2002, LNCS 2350, pp. 661–675, 2002.

[20] Navneet Dalal and Bill Triggs, "Histograms of Oriented Gradients for Human Detection",IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.

[21] Asbjørn Berge, "INF 5300 Video Analysis Detecting motion", SINTEF.

[22] http://mateuszstankiewicz.eu/?p=189

[23] P. KaewTraKulPong, R. Bowden, "An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection", In Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01. Sept 2001.

[24] S. Guler, J. A. Silverstein, and I. H. Pushee, "Stationary Objects in Multiple Object Tracking", IEEE International Conference on Advanced Video and Signal-Based Surveillance, London , UK , Sep 2007.

[25] J. M. del Rincn, J. E. Herrero-Jaraba, J. R. Gomez, and C. Orrite Urunuela, "Automatic left luggage detection and tracking using multi cameras," in PETS, 2006, pp. 59–66.

[26] F. Porikli, Y. Ivanov, and T. Haga, Robust Abandoned Object detection Using Dual Foregrounds, EURASIP Journal on Advances in Signal Processing, Vol. 2008, Hindawi Publishing Corporation.

[27] Stauffer, C., Grimson, W.E.L.: Learning patterns of activity using real-time tracking. IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 22(8), 747–757 2000

[28] M.D.Beynon, D.J.Van Hook, M.Seibert, A.Peacock, D.Dudgeon "Detecting Abandoned Packages in a Multi-camera Video Surveillance System" IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'03) July 21 - 22, 2003 Miami, Florida P. 221.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 714 - 716, 2013, ISSN: 2319–8656

LOCATION BASED DETECTION OF REPLICATION ATTACKS AND COLLUDING ATTACKS

P.S.Nivedita Sai
Sri Sairam Engineering College
Chennai-44, India

T.P.Rani
Sri Sairam Engineering College
Chennai-44, India

Abstract -Wireless sensor networks gains its importance because of the critical applications in which it is involved like industrial automation, healthcare applications, military and surveillance. Among security attacks in wireless sensor networks we consider an active attack, NODE REPLICATION attack and COLLUDING attack. We use localized algorithms, ((ie) replication detection is done at the node level and eliminated without the intervention of the base station) to solve replication attacks and colluding attacks. Replication attacks are detected to using a unique key pair and cryptographic hash function. We propose to use XED and EED algorithm[1] (authenticates the node and tries to reduce the replication) , with this using the Event detected location , non-beacon node is used to find the location of a malicious node and by a simple threshold verification we identify malicious clusters.

Keywords: replication attacks ,collusion attacks, localized detection XED, EED

1. INTRODUCTION

Wireless Sensor Networks (WSNs) are used in various applications. They consist of many autonomous sensor nodes deployed in spatially distributed manner. They are used to sense various parameters like temperature, pressure etc. The network consists of small sensors and a unit which is used to store information also called data center. It consists of an antenna for transmission and a power source. Some typical examples are Industrial monitoring, Environment monitoring, Healthcare monitoring, Area monitoring, Passive Area location detection. These WSN's are more prone to attacks of different types as they are deployed under various conditions. This is because the attackers may intend to learn information from the WSNs or disable the functions of the WSNs. For example, on the battlefield, the enemies would hope to learn the private locations of soldiers by injecting wrong commands into the sensor network. It becomes important to ensure security of data transmitted , this security also will save considerate amount of battery power which will in increase the efficiency of the network. In this paper we have considered replication attack which is considered as a major compromise on the security. When a genuine node is compromised either by brutal force or by software attacks. This compromised node's id and key are copied into another node and randomly deployed in wireless scenario. This is replication attack , when this replicated nodes form a group and launch attacks against the benign nodes , this is collusion attacks. Collusion attacks results in attacks like selective forwarding , selective drop of packets, looping of data. There are many techniques which have been proposed for reducing this collusion attacks, some are deterministic (they use some abnormal pattern for detection) some are non deterministic . Centralized detection results in whole network synchronization and wastage of bandwidth hence here we make use of a localized detection algorthim.

2 .LITERATURE REVIEW

Many mechanisms have been proposed to overcome this replication and collusion attacks. The algorithms proposed in [1] it makes efficient usage of

key and hash pairs to authenticate users to detect replica but it doesn't consider the possibilities of collusion . A random walk model is used in [2], as nodes in a sensor network environment are randomly only deployed. Whereas in Witness collusion technique[3] uses three techniques , DIP,QP, WIP , the major shortcoming of these policies are they cannot detect collusion beforehand. Localized detection [4] uses omni-directional antennas which again emphasizes on the necessity of three beacons minimum .In RED model [5] the mechanism involved uses the mechanism of id obviousness and area obviousness but the major disadvantage is network wide synchronization required. Whereas in distributed detection [6] the topology information about the nodes is used but , all nodes stop working as soon as a replica is detected The detection protocols involving a central control have inherent limits such as a single point of failure.

To detect the node replicas in mobile sensor networks, two localized algorithms, XED and EDD, are proposed. The techniques developed in our solutions, challenge-and-response and with new counter-number with location based information, which are fundamentally different from the others.

3. PROPOSED SYSTEM

The idea behind XED is the basic key exchange mechanism where both the nodes initially during the setup phase will exchange a key , id pair and also a hash function value. These values are stored in a list or a hash table, every time they both encounter each other they will exchange these values and cross verify their authenticity.

For the generation of random numbers we use **[x^2 mod N]** [7] where

Let N {integers N|N } such that P, Q are equal length (|P| |Q|) are distinct primes =3 mod 4} be the set of parameter values.

For N € N, let Xn={x^2 mod N |x€ Zn*}

X=disjoint U_{NEN} Xn be the seed domain.

These random numbers are used in hash function which is generated using anyone of the cryptographic hash function family. These universal hash functions form a group and are stored together.

When a user needs to be authenticated anyone of the hash functions from the family of hash functions (n^2) is chosen and cuckoo hashing [8][9] procedure is used and the hashed values are stored in two tables following the code defined below.

procedure insert(*x*)
if lookup(*x*) then return
loop MaxLoop times
if *T*1[*h*1(*x*)] = *?* then *f T*1[*h*1(*x*)] *x*; return *g*
*x $ T*1[*h*1(*x*)]
if *T*2[*h*2(*x*)] = *?* then *f T*2[*h*2(*x*)] *x*; return *g*
*x $ T*2[*h*2(*x*)]
end loop
rehash(); insert(*x*)
end.

During the insertion process if all the positions in tables are filled then rehashing is done.

Time taken for both lookup and delete is O(n).

Advantages

Our algorithms possess the following advantages.

- Efficiency and Effectiveness: These algorithms are found to be more efficient then the other localized algorithms
- Network-Wide Revocation Avoidance: Since this is localized detect there is no need for all nodes to stop working as soon as a replica is found
- Time Synchronization Avoidance: There is no need for all nodes to operate in the same time slot for exchange of id's etc
- Security: Security level increased by a good amount
- Computational time: since we don't need to go through all the list the computational efficiency becomes O(n)

4. SIMULATION RESULTS

The preliminary stages of this work is network configuration and the Hash value is verified. We use random number generation and the cuckoo hashing technique. The network is deployed by using NS 2.34 and cygwin as an interface on Windows system. This process implementation is shown below.

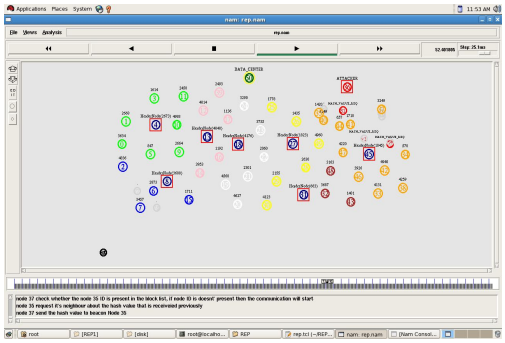


Figure 1: Node distribution

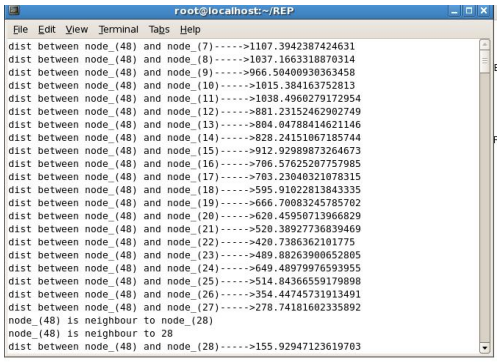


Figure 2: Trace file format

5. CONCLUSION

The first module of proposed system examined the WSNs configurations and clustering the WSN nodes. The sensor nodes clustering is done based their energy level because of entire WSNs mostly depend on power capacity so they can be communicated without communication break. A LEACH routing protocol is used to communicate with corresponding sensor nodes and routes to destination nodes are established. In addition the authenticity of the nodes are verified using XED algorithm

6. REFERENCES

[1] Chia-Mu Yu, Yao-Tung Tsou and Chun-Shien Lu," Localized Algorithms for Detection of Node Replication Attacks in Mobile Sensor Networks ,"IEEE transactions on information forensics and security, vol. 8, NO. 5, MAY 2013.

[2] Yingpei Zeng, Jiannong Cao, Senior Member, IEEE, Shigeng Zhang, Shanqing Guo and Li Xie ,"Random-Walk Based Approach to Detect CloneAttacks in Wireless Sensor Networks"In Proc.IEEE journal on selected areas in communications, vol. 28, NO. 5, JUNE 2010 677

[3] Amirali Salehi-Abari and Tony White ," On the Impact of Witness-Based Collusion in Agent Societies **",**in Proc. Of the International Conference on Principles of Practice in Multi-Agent Systems Pages 80-96.

[4] Zhuhong yOU, Max Q.-H. Meng, Huawei Liang, Shuai Li, Yangming Li, Wanming Chen ,Yajin Zhou, Shifu Miao, Kai Jiang and Qinglei Guo ," An ocalization Algorithm in Wireless Sensor Network Using a Mobile Beacon Node ," in Proc. of the 2007 International Conference on Information Acquisition July 9-11, 2007, Jeju City, Korea.

[5] Mauro Conti Di,Roberto Di Pietro,Luigi V. Mancini and Alessandro Mei, '' A Randomized, Efficient, and Distributed Protocol for the Detection of Node Replication Attacks in Wireless Sensor Networks,'' in Proc. Of MobiHoc'07, September 9-14, 2007, Montréal, Québec, Canada.

[6] Parno,A.Perrig,andV.Gligor,"Distributed detection of node replication attacks in sensor networks," in Proc. IEEE Symp. Security and Privacy (S&P), Oakland, CA, USA, 2005, pp. 49–63.

[7] L. Blum, M. Blum AND M. Shub" A simple unpredictable pseudo-random number generator" , in SIAM Journal on Computing ,Vol 15, issue 2, May 1986, page no-364-383.

[8] Rasmus Pagh and Flemming Friche Rodler,"Cuckoo hashing", Journal of Algorithms 51 (2004) 122–144

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 717 - 722, 2013

Gender Specification Using Touch less Fingerprint Recognition

Merlyn Francis
Fr.CRIT
Vashi, India

Oshin Koul
Fr.CRIT
Vashi, India

Priyanka Rokade
Fr.CRIT
Vashi, India

Abstract: Fingerprint recognition or fingerprint authentication refers to the automated method of verifying a match between two human fingerprints.Fingerprints are one of many forms of biometrics used to identify individuals and verify their identity. This article touches on two major classes of algorithms (minutia and pattern) and four sensor designs (optical, ultrasonic, passive capacitance, and active capacitance).Most of the current security and attendance systems are shifting towards automated biometric systems, the most popular biometrics being fingerprints. In Automated Fingerprint Identification Systems (AFIS), the fingerprint of an individual needs to be identified with that stored in the database. In this is a method which deals with fingerprint identification in the transform domain is considered and the main focus is on the reduction of the processing time. First, the mean of rows (or columns) of the fingerprint image is computed, this converts a two dimensional image signal into one dimension. The one dimensional Walsh transform of the row (or column) vector is generated and is distributed in a complex plane which is subjected to sectorization to generate the feature vector.The feature vector of a given test image is compared to those present in the database. The scores from row and column transform methods are fused using OR and MAX functions. The results with accuracy of more than 73% (for 16 sectors) and high computational speed show that the method can be used in fingerprint identification in application with requirements of less processing time.

Keywords: DWT, SVM, STFT, Gabor filter, SVD, Walsh transform

1. INTRODUCTION

For over a century, fingerprints have been one of the most highly used methods for human recognition. The determination and commitment of the fingerprint industry, government evaluation and needs, and organized standards bodies have led to the next generation of fingerprint recognition, which promises faster and higher quality acquisition devices to produce higher accuracy and more reliability. Automatic fingerprint recognition technology has now rapidly grown beyond forensic applications into civilian applications. In fact, fingerprint-based biometric systems are so popular that they have almost become the synonym for biometric systems.

Biometrics has three broad uses –

1. Verification, i.e. confirming another identifier such as a password, PIN or a photograph.
2. Identification, providing a discrete identifier (or identifiers) that are independent of what the individual knows/remembers (e.g. a password) or what the individual carries (e.g. an identity document or card).
3. Screening, enabling surveillance and sorting of groups of people (e.g. finding a person in a crowd or selecting travellers for detailed examination of passports).

Gender and Age information is important to provide investigative leads for finding unknown persons. Existing methods for gender classification have limited use for crime scene investigation because they depend on the availability of teeth, bones, or other identifiable body parts having physical features that allow gender and age estimation by conventional methods. Various methodologies has been used to identify the gender using different biometrics traits such as face, gait, iris, hand shape, speech and fingerprint. In this work, gender and age of a person is identified from the fingerprint using DWT and SVD.

2. EXISTING TECHNOLOGIES

Gender and age classification can be me made using the spatial parameters or frequency domain parameters or using the combination of both. Most of the findings are based on the spatial domain analysis and few were based on the frequency domain. Many studies were carried out for the human face gender classification by using frequency domain and various classifiers.

There are basically two methods of fingerprint feature extraction:

2.1. Discrete Wavelength Transform Based Fingerprint Feature extraction [3]

2.2. Singular Value Decomposition based Fingerprint Feature extraction [3]

2.1. Discrete Wavelet Transform based fingerprint feature extraction

Wavelets have been used frequently in image processing and used for feature extraction, de- noising, compression, face recognition, and image super-resolution. Two dimensional DWT decomposes an image into sub-bands that are localized in frequency and orientation. The decomposition of images into different frequency ranges permits the isolation of

the frequency components introduced by "intrinsic deformations" or "extrinsic factors" into certain sub-bands. This process results in isolating small changes in an image mainly in high frequency sub-band images. Hence, DWT is a suitable tool to be used for designing a classification system. The 2-D wavelet decomposition of an image is results in four decomposed sub-band images referred to as low–low (LL), low–high (LH), high–low (HL), and high–high (HH).Each of these sub-bands represents different image properties. Typically, most of the energy in images is in the low frequencies and hence decomposition is generally repeated on the LL sub band only (dyadic decomposition). For k level DWT, there are (3*k) + 1 sub-bands available. The energy of all the sub-band coefficients is used as feature vectors individually which is called as sub-band energy vector (E). The energy of each sub-band is calculated by using the equation (1).Where is the pixel value of the kth sub-band and R, C is width and height of the sub-band respectively.

$$E_k = \frac{1}{RC}\sum_{i=1}^{R}\sum_{j=1}^{C}|x_k(i,j)| \quad \text{.....(1)}$$

2.2. Singular Value Decomposition based fingerprint feature extraction

The Singular Value Decomposition (SVD) is an algebraic technique for factoring any rectangular matrix into the product of three other matrices. Mathematically and historically, it is closely related to Principal Components Analysis (PCA). In addition it provides insight into the geometric interpretation of PCA. As noted previously, the SVD has long been considered fundamental to the understanding of PCA. The SVD is the factorization of any k × p matrix into three matrices, each of which has important properties.

That is, any rectangular matrix A of k rows by p columns can be factored into U, S and V by using the equation (2).

$$A = U\,S\,V^T \quad \text{.....(2)}$$

Where

$$U = AA^T \quad \text{.....(3)}$$

$$V = A^TA \quad \text{.....(4)}$$

As the internal database contains images of size 260x300 pixels, the feature vector of SVD is of the size 1x260. The spatial feature extraction by using SVD is shown in Figure 1.

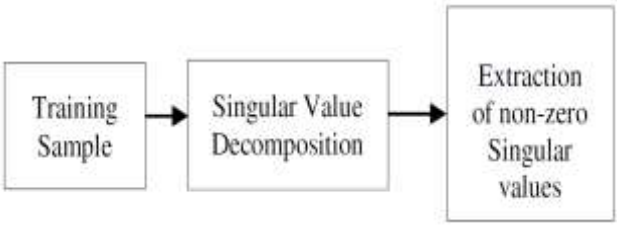


Figure 1. SVD based fingerprint feature extraction

3.TOUCH-LESS FINGERPRINT RECOGNITION

3.1. Overview

In this paper we will be discussing about the touch-less fingerprint recognition system. Most of the sensors available today use "touch" method since it is simple and little training is required. However, the touch-based electronic fingerprint scanner will lead to the weakening of durability if the device is used heavily. In addition, the pressure of the physical contacts will normally cause the touch-based fingerprint images to be degraded. On the other hand, images captured with touch-less devices are distortion free and present no deformation because these images are free from the pressure of contact. Touch-less fingerprint acquisition is a remote sensing technology to capture the ridge-valley pattern which provides essential information for recognition. Several approaches for touch-less fingerprint recognition system have been reported.

3.2. Algorithm

The block diagram of the proposed touch less fingerprint recognition system is shown below.

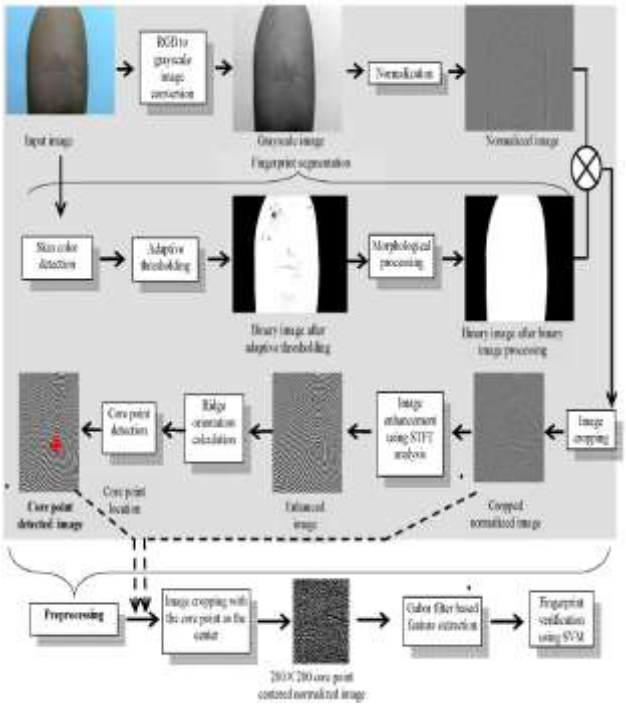


Figure 2. Block Diagram of Touchless Fingerprint Recognition System

3.3. Proposed System

The steps involved in implementation of this system are as follows:

1. Image acquisition
2. Image Preprocessing
3. Verification of results

3.3.1. Image acquisition:

Due to the lack of any standard database of images the fingerprint images are captured using digital camera. The digital camera used is Canon the PowerShot Pro1, with 8 mega pixels of effective resolution and a 7x optical zoom canon "L" series lens. The L-series lens' macro capabilities allow capturing high resolution images while "super macro mode" permits close focus to 1.2 inches enables to get clearer foreground pattern and blurred background pattern.

Table 1. Configuration of the camera.

ISO Speed	400
Image Size/ Resolution	640x480
Image Quality	Superfine
Super Macro	On
Colour Space	RGB
Drive Mode	Continuous shooting (Speed priority)

3.3.2 Image pre-processing:

The main sub- modules of the system are :

a. Pre-processing: The fingerprint images are pre-processed using proposed method which includes skin color detection normalization fingerprint segmentation image enhancement and core point detection. [1]

b. Gabor filter based feature extraction: The feature vectors are extracted by Gabor- filter from the images after pre-processing

c. Fingerprint verification using SVM: The fingerprint verification by using SVM classifier.

3.3.2.1. Pre-processing:

The first step in preprocessing is to convert the image from RGB format to grey scale [0-255]. The image is then normalized by changing the dynamic range of the pixel intensity values in order to reduce the damage caused by the illumination. The next step is segmentation done by skin color detection, adaptive thresholding and followed by the morphological processing. The fingerprint image thus obtained is multiplied with the binary mask from segmentation. The resultant image is cropped and enhanced by using Short Time Fourier Transform (STFT) analysis.
Finally, the core point is detected on the enhanced image.

3.3.2.2. Gabor filter based feature extraction:

It is desirable to obtain representations for fingerprints which are scale, translation, and rotation invariant. The present implementation of feature extraction assumes that the fingerprints are vertically oriented. In reality, the fingerprints in our database are not exactly vertically oriented; the fingerprints may be oriented up to away from the assumed vertical orientation.

The four main steps in our feature extraction algorithm are [2]:
1) Determine a reference point and region of interest for the fingerprint image
2) Tessellate the region of interest around the reference point
3) Filter the region of interest in eight different directions using a bank of Gabor filters (eight directions are required to completely capture the local ridge characteristics in a fingerprint while only four directions are required to capture the global configuration
4) Compute the average absolute deviation from the mean (AAD) of grey values in individual sectors in filtered images to define the feature vector or the Finger Code.

To meet these conflicting requirements of an accurate and reliable localization, we propose a new method of reference point determination based on multiple resolution analysis of the orientation fields. Our new method locates the reference point more precisely than the algorithm proposed by Hong and Jain.

Let us first define the orientation field, O, for a fingerprint image. The orientation field, O, is defined as a P×Q image, where O(i,j) represents the local ridge orientation at pixel (i,j). Local ridge orientation is usually specified for a block rather than at every pixel; an image is divided into a set of w×w non- overlapping blocks and a single orientation is defined for each block

A summary of our reference point location algorithm is presented below [2].

1) Estimate the orientation field O as described above using a window size of w×w.
2) Smooth the orientation field in a local neighborhood. Let the smoothed orientation field be represented as O'. In order to perform smoothing (low-pass filtering), the orientation image needs to be converted into a continuous vector field.
3) Compute an image containing only the sine component of O'.
4) Initialize A, a label image used to indicate the reference point.
5) For each pixel (i,j), integrate pixel intensities (sine component of the orientation field) in regions and assign the corresponding pixels in A, the value of their difference

$$A(i,j) = \sum_{R_I} \mathcal{E}(i,j) - \sum_{R_{II}} \mathcal{E}(i,j) \quad \text{....(5)}$$

The regions were determined by applying the reference point location algorithm. The geometry of regions is designed to capture the maximum curvature in concave ridges. Although this successfully detects the reference point in most of the cases, including double loops but is not very precise and consistent for the arch type fingerprints.
6) Find the maximum value in A and assign its coordinate to the core, i.e., the reference point.
7) For a fixed number of times, repeat steps 1–6 by using a window size of w'×w', where and w'<w and restrict the search for the reference point in step 6 in a local neighbourhood of the detected reference point.

3.3.2.3. Fingerprint verification using Singular Value Decomposition:

It is a binary classifier based on principle of structural risk minimization that maps an input sample to high dimensional

feature space. SVM [1] is adopted as it could optimally separate the two classes of genuine and imposters by constructing a hyper plane. The hyper plane is defined by

x * w + b = 0 **….(6)**

where,
w: is normal to the plane
b: is bias term.

The extension to non linear boundaries is achieved through transforming each data point to a higher dimensional space. Better seperability between two classes is achieved with proper transformation which is done by polynomial kernel or radial basis function (RBF) kernel.

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^n \quad \ldots(7)$$

Where n : order of the polynomial

$$K(x_i, x_j) = \exp\left[-\frac{1}{2}\left(\frac{\|x_i - x_j\|}{\sigma}\right)^2\right] \quad \ldots\ldots(8)$$

Where

σ :the width of the radial basis function.

Another technique for the above process is filterbank based matching.

3.3.2.4. Filterbank based matching:

Fingerprint matching is based on finding the Euclidean distance between the corresponding finger codes. The translation invariance in the finger code is established by the reference pointer achieved by cyclically rotating the features in the Finger Code itself. A single step cyclic rotation of the features in the Finger Code described by (21)–(23) corresponds to a feature vector which would be obtained if the image were rotated by 22.5°. A rotation by R steps corresponds to a R×22.5° rotation of the image. A positive rotation implies clockwise rotation while a negative rotation implies counterclockwise rotation. The Finger Code obtained after R steps of rotation is given by [2],

$$V_{i\theta}^{R} = V_{i'\theta'} \quad \textbf{….(9)}$$

i’ = (i + k + R) mod k + (i div k) ×k ….(10)

θ’ = (θ + 180° + 22.5° × R) mod 180° ….(11)

The process has been depicted in the following figure.

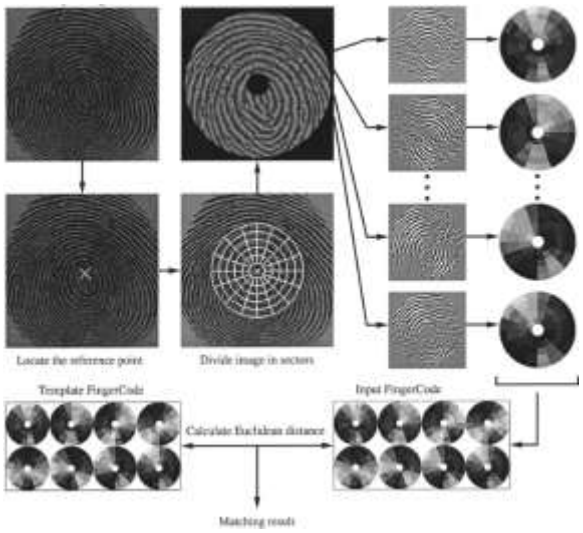


Figure 3. System diagram of fingerprint authentication system

4. FINGERPRINT GENDER CLASSIFICATION

4.1. Overview

The proposed system for gender classification is built based on the fusion of fingerprint features obtained by using DWT and SVD. This section describes two different stages named as learning stage and classification stage and the KNN classifier used for the gender classification [3].

4.2. Algorithm

4.2.1. Learning stage:

The feature vector V of size 1x260 obtained by SVD and the sub band energy vector E of size 1x19 obtained by DWT are fused to form the feature vector and used in the learning stage. The fusion of feature vector V and E is done by concatenation of features that are widely used for feature level fusion. The resulting feature vector is of the size 1x279 (1x260 +1x19).The learning stage is shown in Figure 4.

The learning algorithm is as follows:

[Input] all samples of fingerprint with known class (Gender).

[Output] the feature vector of all samples as database.

a) Decompose the fingerprint with 6 level decomposition of DWT.
b) Calculate the sub-band energy vector (E) using (1).
c) Calculate the Eigen vector (V) using (2).
d) Fuse the vectors E and V to form the feature vector for the particular fingerprint.
e) Insert this feature vector and the known class into the database.
f) Repeat the above steps for all the samples.

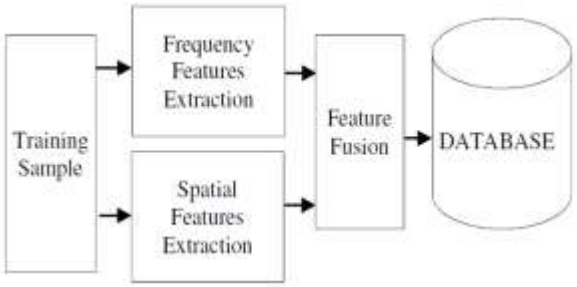


Figure 4. Learning Stage of the Proposed Gender Classification System

4.2.2. KNN Classifier:

In pattern recognition, the k-nearest neighbour algorithm (K-NN) is the generally used method for classifying objects based on closest training examples in the feature space. KNN is a type of instance-based learning where the function is only approximated locally and all computation is deferred until classification. In K-NN, an object is classified by a majority vote of its neighbours, with the object being assigned to the class most common amongst its k nearest neighbours (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbour. The neighbours are taken from a set of objects for which the correct classification is known. This can be thought of as the training set for the algorithm, though no explicit training step is required [3].

4.2.3. Classification Stage:

In the classification phase, the fused feature vector of the input fingerprint is compared with the feature vectors in the database by using the KNN classifier. The distance measure used in the classifier is 'Euclidean Distance'. The classification process is as follows.

***Classification algorithm* [3]:**

[Input] unknown fingerprint and the feature database

[Output] the class of the fingerprint to which this unknown fingerprint is assigned

a) Decompose the given unknown fingerprint with 6 level decomposition of DWT.
b) Calculate the sub-band energy vector (E) using (2).
c) Calculate the Eigen vector (V) using (1).
d) Fuse the vectors E and V to form the feature vector for the given unknown fingerprint.
e) Apply KNN classifier and find the class of the unknown fingerprint by using the database generated in the learning phase.

5. EXPERIMENTAL RESULTS

5.1. Gender classification using DWT only

The code is tested from 2nd level to 7th level and the success rate for the classification is identified. No appreciable results were obtained for the levels 2 to 4 and beyond the level 7, the results were not convincing. Significant success rate is obtained for the levels 5, 6 and 7. The sub-band energy vector for the level 5 is of the size 1x16 and these features are compared with templates stored in the database obtained during the learning stage. Similarly, the sub-band energy vectors are of the size 1x19 and 1x22 for the level 6 and 7 respectively. The results achieved by the 2-D DWT for the levels 5, 6 and 7 are listed in table 2 for each finger of the male and female [3].

Table 2. Gender Classification Rate For Different Levels Of DWT.

Finger No.	Level 5		Level 6		Level 7	
	Male	Female	Male	Female	Male	Female
1	86.99	91.67	87.67	94.32	87.67	90.63
2	89.04	86.46	89.86	90.88	89.04	83.33
3	90.41	81.25	92.79	85.88	91.10	86.46
4	93.15	79.17	95.46	79.68	89.73	78.13
5	91.78	73.96	94.92	75.92	93.15	72.92
6	93.84	72.92	92.8	73.87	93.15	71.88
7	89.73	77.08	93.17	80.69	91.78	78.13
8	91.78	83.33	92.57	84.28	91.10	79.17
9	89.73	86.46	90.73	89.78	93.15	87.50
10	86.99	85.42	86.76	93.58	87.67	87.50
Average	90.34	81.77	91.67	84.89	90.75	81.56

The overall classification rate for the level 5, 6 and 7 are 84.61%, 85.46% and 86.41% respectively. It is also observed that the success rate of right thumb finger (numbered as 6) of male are quite higher than the other fingers. Similarly, the success rate of left hand little finger (numbered as 1) of female are higher than the other fingers.

5.2. Verification results for touchless fingerprint recognition [1]

Verification experiments are conducted by using core point center normalized fingerprint images. After the preprocessing o\f the fingerprints the features derived from gabor filter are used for subsequent SVM verification.

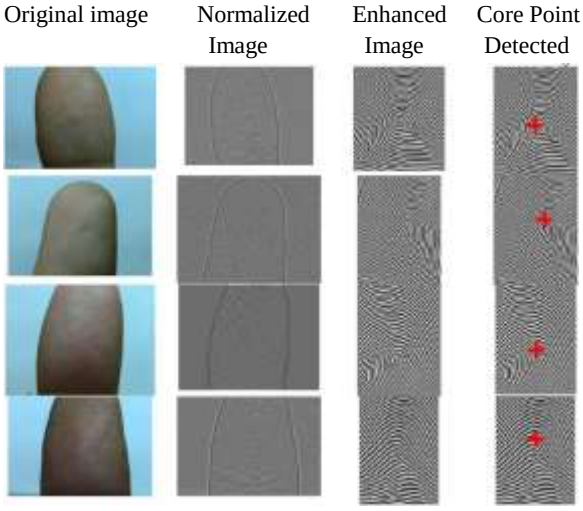


Figure 5. Results Of Proposed Preprocessing

Above are the results for successful core point detection.

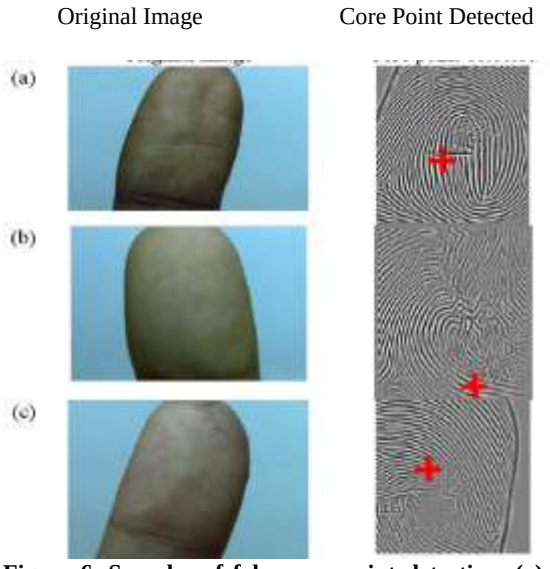


Figure 6. Samples of false core point detection. (a) Deep wrinkle image (b) motion blurriness image (c) defocus image.

There are two classes of errors in verification system
False Acceptance Rate (FAR). False Reject Rate (FRR).
System performance is determined by using Equal Error Rate (EER) i.e. FAR=FRR.
In filter based recognition the fingerprint images were collected of size 508x480 pixel. The age distribution was as follows:
1) Younger than 25 years: 46.5%;
2) Between the ages of 25 and 50: 51%;
3) Older than 50 years: 2.5%.
The images were taken in the order of right index, right middle, left index, left middle. These images were stored in the database MSU_DBI. The fingerprints were taken at the interval of 6 months. In practice, the ROC curve behaves in between these two extremes [2].

Table 3. FAR and FRR with different threshold values for MSU_DBI database.

Threshold Value	False Acceptance Rate (%)	False Reject Rate (%)
30	0.10	19.32
35	1.07	7.87
40	4.59	2.83

6. CONCLUSION

Average statistics along fingers for every pattern type was calculated together with different concordance and asymmetry properties for the corresponding fingers. The variation among females and males in the membership of the fingerprints to the different pattern types, and the average ridge count for fingers belonging to each pattern type, are very small, and thus are statistically insignificant. But there is no significant difference in the degree of asymmetry between males and females, and thus the asymmetry is not a good candidate for the classification process. The pattern type concordance between left and right corresponding fingers doesn't show significant statistical variations between females and males.

While the present approach tolerates small magnitudes of elastic distortion and local scaling (due to finger-pressure variations), it does not take care of significant nonlinear elastic distortion in the fingerprints. The inter-ridge densities in a fingerprint could be used to obtain a canonical representation to compensate for the large distortions due to shear and pressure variations caused by the contact of the finger with the sensing device.

The touch-less fingerprint recognition system is presented here where the experiment results signifies an improvement using the proposed algorithm in the segmentation, enhancement and core point detection for the fingerprint images captured by the digital camera. Moreover it has also presented an effective verification technique that employs SVM where feature vectors are extracted using the Gabor filter.

7. REFERENCES

1] "Touch-less Fingerprint Recognition System" - B.Y. Hiew, Andrew B.J. Teoh and Y.H.Pang, multimedia university, Malaysia. June 2007. Automatic Identification Advanced Technologies, 2007, IEEE Workshop, page no: 24-29.

2] "Filterbank Based Fingerprint Matching" - Anil K. Jain, Salil Prabhakar, Lin Hong, Sharath Pankanti. IEEE Transactions on Image Processing, Vol 9, No. 5, May 2000.

3] "Fingerprint Gender Classification using Wavelet Transform and Singular Value Decomposition Singular Value Decomposition" - Gnanasivam P and Dr. Muttan, IJCSI International Journal Computer Science Issues, Vol 9, Issue 2, No. 3, March 2012

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 723 - 725, 2013, ISSN: 2319–8656

Implementation Of Ontology Matching Using Protégé

Jayanth Kumar	Sujitha Reddy	Priya.M	Raju.J	Vanmathi.C
SITE	SITE	SITE	SITE	SITE
VIT University	VIT University	VIT University	VIT University	VIT University
Vellore, India	Vellore, India	Vellore, India	Vellore, India	Vellore, India

Abstract: Ontologies play an important role for interoperability between organizations and for semantic web because they aim at capturing domain knowledge. Different solutions exist to identify the common concepts and involve different sources using ontology. Matching is a key operation in semantic web since it takes inputs as ontologies; each consists of a set of discrete entities like tables, xml elements, classes, properties. In this paper, we developed two anthologies by which we matched those ontologies to get the similar classes, objects, instances by using Protégé. Protégé is a tool used for comparing ontologies, merging ontologies, data translation, query processing.

Keywords: Ontology,Ontology Matching,Merging Ontologies,Domain Knowledge

1. INTRODUCTION

An ontology is a broader knowledge model with a reasoning mechanism that facilitates knowledge on the semantic web [6]. Semantic web represents web content in a form that is easily machine process- able. It uses intelligent techniques to take advantage of these representations. Semantic web knowledge will be organized in conceptual spaces according to meaning. Semantic web technologies are Ontologies, Explicit Metadata, Logic and Inference and Agents [6]. The term ontology originates from a philosophy which means nature of existence

T.R Gruber's definition was redefined by R.Studer as "Ontology is an explicit and formal specification of a conceptualization" [3].

The knowledge representation languages are used to create sets of terms and axioms about the meanings of the terms as well as to specify classes, properties and relationships between classes and objects in a domain. Reasoning engines like Fact++, Hermi 1.2.8, DIG and Pellet 1.2.8 are used to infer the meanings and relationships between the classes, subclasses, and superclasses It will build a taxonomic structure [4] [5] and checks the model consistency.

Ontology for the semantic web is typically constructed in notational languages like RDF, RDFS and OWL.

RDF is Resource Description Framework and it is used for describing resources on the web. It is designed to understand and read by computers. This is written in XML and it is a part of W3C semantic web activity. RDF identifies things using web identifiers and describes the resources with properties, property values.

RDFS is a Resource Description Framework Schema. It does not provide actual application specific classes and properties. Classes in RDFS are like classes in object oriented programming, so it is easy for resources to define instances of classes, and subclasses of classes.

OWL is a web ontology language. It is built on top of RDF and is used for processing information on the web. It has been designed and interpreted by computers and not for people. Owl and RDF are one on the same thing but OWL is stronger language with the more interpretability than RDF. It comes with a larger vocabulary and a stronger syntax than RDF.

OWL has three types of languages

- OWL LITE
- OWL DL
- OWL FULL

Many applications such as data integration, e-commerce and semantic web services, which are exploiting internet content represented in a form that is easily processed by machine and enriched with metadata. Ontologies play an important role for interoperability between organizations and for semantic web because they aim at capturing domain knowledge in a generic way and provide a proper understanding of the domain [7] [8].

For developing the ontologies we need to go through these notational languages.

With the developed ontologies we can go with different types of operations like ontology matching, ontology merging, query processing and many more.

An ontology matching is like developing ontology and compares those two by which we will get similar classes, instances and properties of a class.

Ontology is considered here as taxonomy of concepts and the problems of matching are reduced to "for each concept node in one taxonomy, find the most similar node in the other taxonomy". Our concern discussed in this paper is to make corrections to eliminate the matching concepts.

The mapping is the most important step in Ontology alignment and it is described informally as where for each entity in one ontology we try to find a same entity in the second anthology, with the same meaning. The result of the mapping may be as simple as one-one correspondences between some concepts and they can be declarative concepts.

After some years people found tools for developing ontologies without using the notational If they are trying to develop ontologies automatically the code will be generated by using these tools, so need to go with the notational languages nowadays.

There are several tools which help in developing ontologies: [1] [2]

- Protégé
- Swoop
- Top Braid composer
- Oiled
- WebODE
- Ontolingua
- Internet business logic
- Onto track
- IHMC
- Cmap Ontology editor

- By using these tools we can define new concepts, relations and instances. These tools may contain additional features such as graphical browsing, search and constraint checking capabilities

2. RELATED WORK

a. Hacene Belhadefa [2011], detected and repaired the list of homologs concepts; which is based on the bidirectional comparison (Checking) between the two matched ontologies to filter this list of concepts. The measures used are classical just to show the reliability and reliability of their method, but they generalized it for supporting other advanced measures. Here an ontology is considered here as a taxonomy of concepts and the problem of matching is reduced to "for each concept node in one taxonomy, find the most similar node in the other taxonomy". Their main concern is to discuss in their paper is the reliability of the results and how to make corrections to eliminate the matching concepts when the degree is low [1].

b. Haridimos Kondylakis [et. al] presented a mapping language for information integration under a common knowledge representation model (LAV approach).They proposed a specific mapping annotation format that is capable to capture all those cases. They assume that the level of detail of this format is sufficient to produce complete mediation of data transformation algorithms without further input from the domain experts. They noted that their mappings will be used like in a Local as View mediation process and the mapping format should give enough specifications to an IT-expert to start building an integration algorithm without any help from the domain expert. Their plans for future work include the task of verification of their mapping mechanism [7]

c. M. Rahamatullah Khondoker,[et.al] Finded an appropriate tool to develop ontology is the first step towards ontology development in a lot of ontology development tools are available in the market, however, some are free and some are commercial. Ontology development is a complex and largely domain-oriented process that can be benefited from tool support. Ontology development tools are compared based on certain features such as modelling features/limitations, base language, web support and use, import/export format, graph view, consistency checks, multi-user support, merging, lexical support, and information extraction. Compares development tools based on user-experience is a scarce attention from the research community

3.ONTOLOGY MATCHING

Ontology alignment or ontology matching is the process of determining correspondences between concepts [3] [4]. A set of correspondences is also called as alignment. Ontology matching is the best solution for the heterogeneity problem because it finds correspondences between semantically related entities of the ontologies. These correspondences can be used for various tasks like ontology merging, query answering, and data translation and for navigation on the semantic web.The matching ontologies enable the knowledge and data expressed in the matched ontologies to interoperate.The alignment of two ontologies, Ontology1 and Ontology2 amounts to determining the correspondence between the various ontological entities by category. All methods of alignment determine the correspondences between ontological entities using measures of similarity. The measures of similarity or dissimilarity allow evaluating the similarity or the distance between two elements or (individuals).Subsequent Pages

4. IMPLEMENTATION DETAILS

Normally all the ontologies developed by using notational languages like XML, RDF, , OWL. By using these languages it is hard to implement or develop ontologies. To avoid them, tools are developed and one of them is Protégé which is good to develop ontologies as well as for matching ontologies to get the similar correspondences in a result. [5]

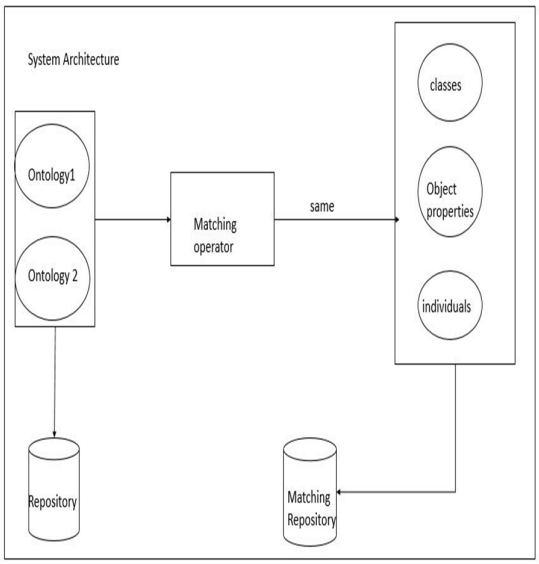


Fig 1: System Architetcure

The following steps take place in ontology matching:

A. Developing Ontologies

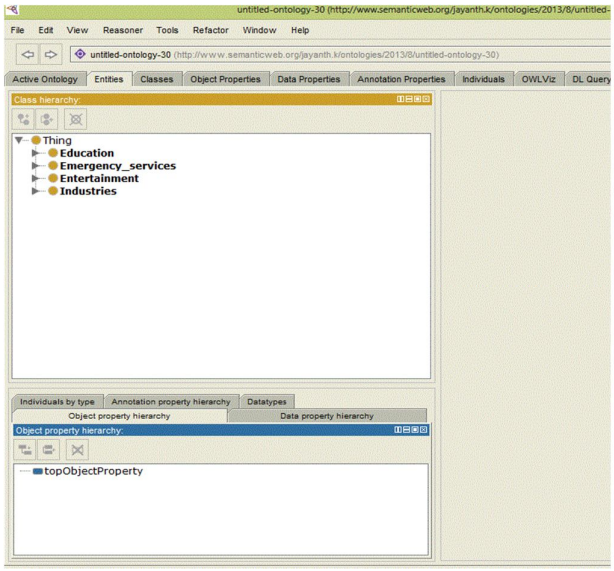


Fig 2: Developing Ontologies

B. Result of Ontology Matching

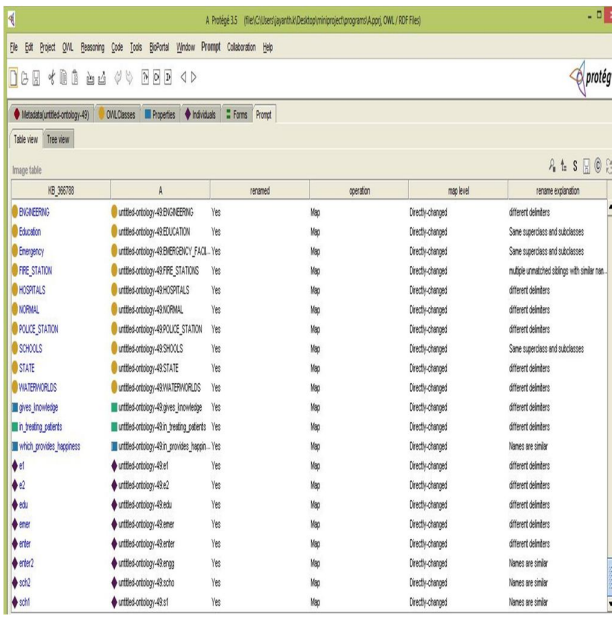


Fig 3: Matched Ontology

5. CONCLUSION

In this paper we developed ontologies using protégé and by comparing the developed ontologies we got the similar classes, objects and individuals and also we are able to search the queries by using the DL query it is possible when we start the reasoner. The developed ontologies are useful in the future for Merging two ontologies, query processing, Data translation, and for navigation of a web.

6. REFERENCES

[1] Hacene Belhadefa, "A new bidirectional method for ontologies matching", proceedings on 2011 International Conference on Power Electronics and Engineering Application

[2] M. Rahamatullah Khondoker, Paul Mueller, "Comparing Ontology Development Tools Based on an Online Survey", Proceedings of the World Congress on Engineering 2010 I WCE 2010, June 30 - July 2, 2010, London, U.K

[3] Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2004) "Ontological Engineering": with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web. 2nd printing, Springer-Verlag, London. ISBN 1-85233-551-3

[4] Jerome Euzenat, Jerome. Euzenat, Jerome Euzenat, Pavel Shvaiko "Ontology matching"Protégé (2000).

[5] Haridimos Kondylakis1, Martin Doerr1, Dimitri's Plexousakis1,"Mapping Language for Information Integration ", *Technical Report 385*, ICS-FORTH, December 2006

[6] Dario Salvi, Marco Picone, Mar´ıa Teresa Arredondo, Mar´ıa Fernanda Cabrera-Umpi´errez, ´ Angel Esteban, Sebastian Steger, and Tito Poli, "Merging Person-Specific Bio-Markers for Predicting Oral Cancer Recurrence through an Ontology" ,IEEE transactions on biomedical engineering aT tHE tIME oF 1, january 2013

[7] Review of ETL and Ontology/Schema Matching Techniques and Tool (2010), By: Geo scientific Data Integration,At:geodi.ucc.ie/files/deliverables/GeoDI_D25_Final.pdf

[8] Thomas R Gruber " A translation Approach to portable Ontology Specifications Knowledge Acquisition",5(2):199-220.1993

[9] Fonseca F., Davis C Câmara G."Bridging Ontologies and Conceptual Schemas Geographic Information Integration", Geoinformatica, Volume 7 Issue 4, 355-378. 2003.

[10] Kalyanpur, Aditya., Parsia, Bijan., Hendler, James "A Tool for Working with Web Ontologies" Proceedings of the International Journal on Semantic Web and Information Systems, Vol.1, No.1, Jan-Mar (2005)

[11] Gruber, T. 1993 "Toward principles for the design of ontologies used for knowledge sharing.", International Workshop on Formal Ontology, 43 (5) pp. 907-928.

[12] S. Steger and G. Sakas, "Image gradient based shape prior for the seg- mentation of not that spherical structures," in Proc.9th IEEE Int. Symp. Biomed. Image, May 2012, pp. 1252–1255.

[13] F.N. Noy and M.A. Musen, "PROMPT Algorithm and Tool for Automated Ontology Merging and Alignment," Proceedings on 17th Nat'l Conf. Artificial Intelligence (AAAI '00), pp. 450-455, 2000.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 726 - 730, 2013, ISSN: 2319–8656

Modeling and Analysis of Reliability in Grid using Petri Nets

Mahnoosh Vahebi
Department of Computer Science and Research Branch
Islamic Azad University,
Gilan. Iran

Mehdi Sadeghzadeh
Department of Computer Science and Research Branch
Islamic Azad University
Mahshahr, Iran

Reza Ebrahimi Atani
Department of Science and Research Branch,
Islamic Azad University,
Gilan. Iran

Abstract: Intelligent Grid computing is considered as an essential infrastructure in future life, and they are executed and implemented to save the energy, to reduce the costs and to increase reliability. This study has been carried out with the aim of identifying various samples of Grid networks and investigating the applied and functional fields. In this way, science and technology develops and expands through identifying their weaknesses and attempting to remove these weaknesses and to solve the problems. In addition, the purpose is to increase the reliability in grid networks for using the services. In fact, the purpose of this study is to provide and present a compact and intensive mathematical tool by using Petri nets in order to analyze reliability and certainty in Grid networks. In this study, reliability is investigated in Grid services, and through using Petri nets, a model is represented and proposed for computing as well as increasing reliability in Grid networks. In this research, a new model has been introduced for increasing reliability in Grid networks. In this model, by using series-parallel mode and also by adding redundancy technique to the system, reliability has been considerably improved.

Keywords: grid network- petri net- Reliability- cpn tools- Resource managment

1. INTRODUCTION

Nowadays, scientific issues and problems are very complicated and require great storage spaces and high computational power. Traditional techniques such as parallel and distributive computing are not appropriate methods for solving such problems. Therefore, Grid networks are proposed in late 90s, and nowadays, they are widely used in developed countries with the aim of preventing waste of energy and using them optimally.

Grid computing is considered as a software and hardware infrastructure, and allows us to access high computational capabilities in a certain, compatible, cheap and broad way. Grid computing is related to a set of heterogeneous resources in large scales such as personal computers, work stations of clusters and supercomputers. These resources vary in terms of computational capabilities and configuration, and they are organized according to different management policies. Grid computing environment is suitable and proper bed for solving the problems requiring long and difficult computing.

In software or a set of software use grid, some issues including information subscription, information dividing, its transferring by considering full security and management of information will be very difficult and important.

The most important issue and problem mentioned in this paper is resource subscription. In Grid, resource subscription is controlled by management system of control resources. The request sent by user to RMS is called task.RMS should divide the received task into sub-tasks, and should send them to resources available in the network. In order to reach the purpose of Grid and to use resources available in Grid environment, scheduling and distribution of sub-tasks among resources are performed by considering the quality of service. In the procedure of scheduling, sub-tasks should be distributed between resources in a way that maximum QOS is produced.

Service reliability in Grid environment should be studies as one of the most significant QOS criteria. Execution time is a duty and is considered as a factor for computing reliability in Grid environment. Always, it has been tried to perform a task in the time lesser than determined time. Usually, technique used in Grid environment to increase reliability is redundancy technique. After dividing tasks into sub-tasks, RMS does not assign each task to one resource; rather, it assigns it to several resources so that all of them process the related sub-task simultaneously. Also, if one of these resources is destroyed, other resources try to compensate it.

2. LITERATURE REVIEW

Azgami in [10] , has considered modeling of tasks distribution and reliability computing in Grid networks. In this paper, management system of resources has undertaken the task basis of Grid networks. This system receives the tasks from users, and divides them into sub-tasks. Then, it sends each sub-task to several available resources. Finally, after executing them by connected resources, outputs are collected, and user's request task is responded. Above mentioned operations are simulated by time colored Petri nets, and reliability is computed.

Zeng in [1] , has analyzed reliability in control center networks in smart Grid using stochastic Petri nets. In his article, the purpose is to present a compact and intensive mathematical tool by stochastic Petri nets (SPNs) to analyze reliability in control center network in smart Grid network.

In this paper, a method has been presented to repair the unit server that has been destroyed. Then, two supporting strategies in critical servers have been studies and investigated, and SPNs have been presented. Finally, above mentioned model combines with general model of the purpose of control center network.

In this paper, in control center, we have considered a network consisting of a control center and N sub-network. In control center, SCADA server, data base and application servers are

connected to local networks (LAN) supported by profile. One a section fails or is exposed to different attacks in control center, then it can be immediately repaired by user interface. In addition, supportive services are used in order to improve reliability.

Dastgheibi Fard, in[11]: has presented a new scheduler algorithm for grid computing. He has firstly introduced LMB algorithm. Then, weaknesses of this algorithm have been mentioned, and NLMB algorithm has been proposed and introduced. The purpose of this algorithm is to reduce removed deadlines, to increase sustainability of errors and to reduce return time. The proposed algorithm considers more parameters in scheduling in comparison with that latest presented algorithm, and reduces removed deadlines to a considerable degree. Also it increases sustainability of error.

3. PROPOSE ALGORITHM

In proposed model, following hypothesis have been taken into account:

1) In Grid environment, star topology has been used and applied, and RMS is connected to all resources. Communication lines are unique between each resource and RMS.
2) Users using Grid send their own requests to RMS. Then, RMS divides them into sub-tasks, and sends them to resources.
3) Sub-tasks do not depend on each other. As soon as resources receive sub-task data from RMS, they begin to execute it.
4) RMS is very quick and reliable, so the time of task processing by RMS can be ignored in comparison with the time of sub-task processing.
5) Grid environment reliability refers to the probability of presenting the results of a task in the time lesser than determined time, T(reliable).

In the proposed algorithm, redundancy technique is used to increase reliability. In this algorithm, one sub0task is allocated to several resources on the basis of redundancy technique. In previous algorithm, resources are provided for sub-tasks stochastically, while in proposed method, each sub-task considers the most appropriate resources on the basis of data required to execute each sub-task and its complex computing as well as bandwidth of resources and processing speed of each resource (see figure1 of the model).

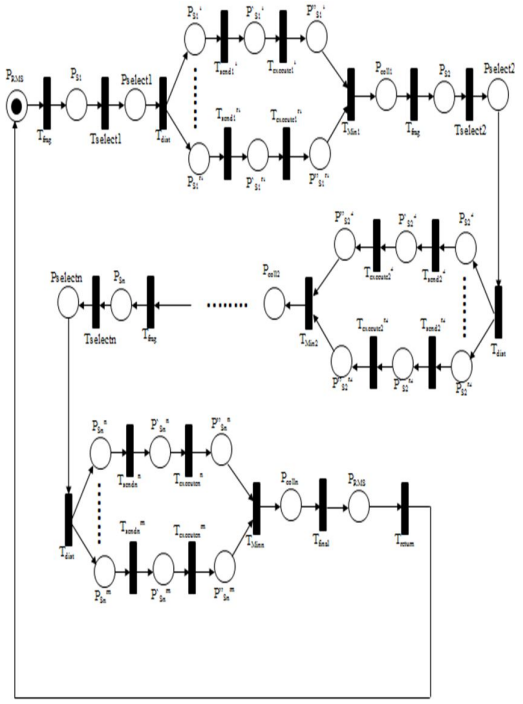


Figure 1. Proposed Model

Let's compute reliability as follows:

Data required executingS_i sub-task that should be sent from RMS toR_j resource is shown by rd_i. The results that should be sent from R_j resource to RMS are demonstrated by I_i. Data transferred and sent to execute S_i sub-task between RMS and R_j resource can be shown by a_i, and it is computed as follows:

$A_i=rd_i+I_i$

If bandwidth of communication lines between RMS and R_j resource is demonstrated by bw_j, then transferred time is computed as follows:

$T_{ij}=a_i/bw_j$

Transferred and sent time adds transferred time to tokens transferred from P^j location to P^i location. In addition, when tokens pass through transferring T(send), this issue should be considered that tokens should not face with failing during transferring data (q_{ij}). Therefore, when tokens pass through transferring, a pair of (T_{ij},q_{ij}) is added to them. If computing complexity of all entered tasks is shown by c, then computational complexity of each sub-task is computed according to the following equation:

$$\sum_{i=1}^{n} C_i = C$$

In this equation, n indicates the number of sub-tasks related to a specific task, and $c_i$is computational complexity of each s_i task. If processing speed of R_j is shown by P_{sj}, then processing time of s_i sub-task is obtained through the following equation:

$$T_{ij} = C_i / ps_j$$

As it is observed in figure, time transferring of T(execute) can add processing time of sub-task (T_{ij}) to time of data transferring (t_{ij}), and adds the obtained result to token. In

addition, the probability of lack of failing during sub-task (P_{ij}) processing is multiplied by the probability of lack of failing during data transferring (q_{ij}), and is added to token. Hence, tokens passing from transferring T(execute), it carries a pair of $(t_{ij}*T_{ij}),(p_{ij},q_{ij})$ that is an indication of successful execution time and its probability. When redundancy technique is used to increase reliability, each sub-task is processed by several processing resources in simulation. Therefore, T_{iRi} is an indication of minimum execution time of s_i sub-task in resources, and it is obtained by the following equation:

$$T_{i,Rri} = \min(t_{ij} + T_{ij}); \quad k \in R_{ri}$$

With regard to above mentioned equation as well as a pair of $(t_{ij}*T_{ij}),(p_{ij},q_{ij})$, the minimum execution time of s_i sub-task and probability of $(P_{i,Rri})$ have been placed over token after transferring through T(min) passage, and it is located in P(coll) place. In order to end a task, all sub-tasks should be executed. Hence, when tokens are placed over all P(coll) places, T(final) transferring begins, and it considers the greatest execution time of sub-task as execution time of task. It is obtained through the following equation:

$$T = \sum_{i=1}^{n} T_i$$

In this equation, n indicates the number of sub-tasks related to each sub-task. In other words, in order to execute a task, this task is firstly divided into sub-tasks. Afterwards, completion time is computed in each resource and for each sub-task. Minimum completion time is determined in each source and for each sub-task. After performing this process for all sub-tasks, minimum time is selected among the least completion time of resources for execution.

When tokens pass through T(max) transferring, and are placed in P(RMS2) location, the operations of task execution come to end. By beginning T(return), another task can enter the system.

The following equation is used to compute reliability:

$$R_{T^*} = \prod_{i=1}^{n} P_i * l(T < T^*)$$

In this equation, I indicates total number of t. p_i is the probability of executing a task in T_i time, and T^* is $T_{reliable}$ time. Also, l function is defined as follows:

$$l(true) = 1 \quad , \quad l(false) = 0$$

In addition, the following equation is used to compute p_i reliability:

$$P_i = 1 - \prod_{J=1}^{n} (1 - R_j)$$

4. CASE STUDY

Let's implement our proposed algorithm in a system, and measure reliability and response time according to mentioned methods.

In case studies, we try to present comprehensive and small samples. In this way, the samples are not very complicated. Also we try to present a small sample which is representative of real and large samples. For example, ATM system has been considered with a properC4ISR architectural structure because it is not very complex, and creating its architectural structure is possible. Then, the problem is explained, and an executable mode for architectural products (UML diagrams) is created. Afterwards, through using the proposed method and simulation of the model that can be used and executed in CPN tools, reliability and response time are computed. ATM interacts with user and bank. In this system, one of the main user operations is withdrawing money from the account. The following figure shows is withdrawing money from the account (see figure 2).

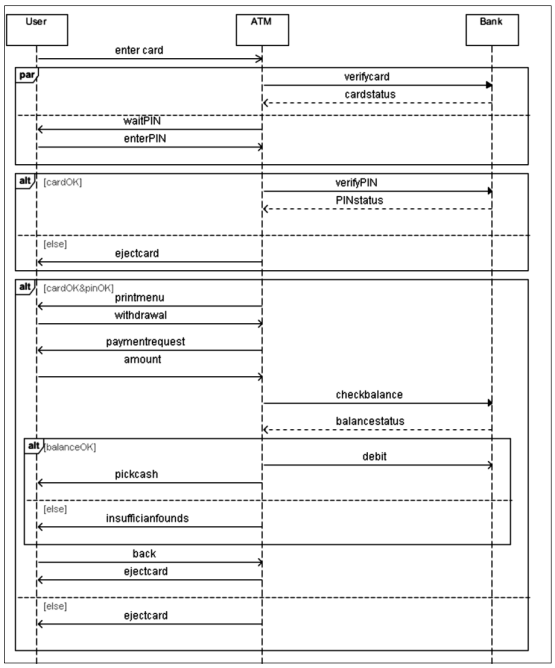


Figure 2. ATM system

4.1 Simulation

With regard to the model presented in chapter of proposed algorithm, ATM system shown in previous section is simulated with regard to explanations of proposed algorithm, primary characteristics of system should be firstly determined. These characteristics involve bandwidth of system, computational complexity for each job which has been entered the system, and data required for each job and work. After determining above mentioned issues and factors, system is simulated. ATM system has been implemented in CAPTOOLS simulator software as follows (see figure 3):

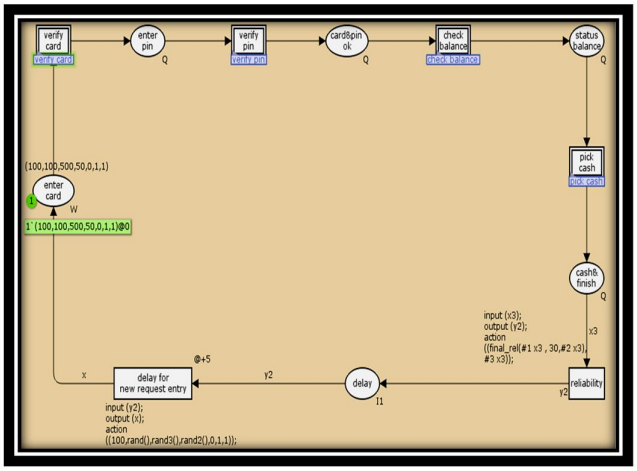


Figure 3.ATM simulation system by cpn tools

Let's consider subsystems of ATM system. Hence, there are four sub-systems as follows: card entry and investigating it, entering card password, selection of request type (deposit request), receiving money. In this paper, just one subsystem is explained in details. These subsystems have been implemented in CPNTOOLS simulator as follows (see figure 4):

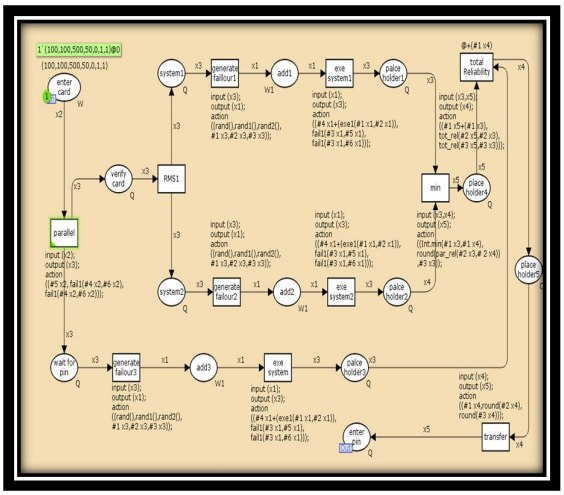


Figure 4. Subsystem of ATM System

After simulating the system, we compare the results obtained from system with parallel systems (Azgami model) and series system (normal mode). The following figure shows this comparison (see figure 5). According to diagram, it's obvious that our proposed model improves reliability to a considerable degree.

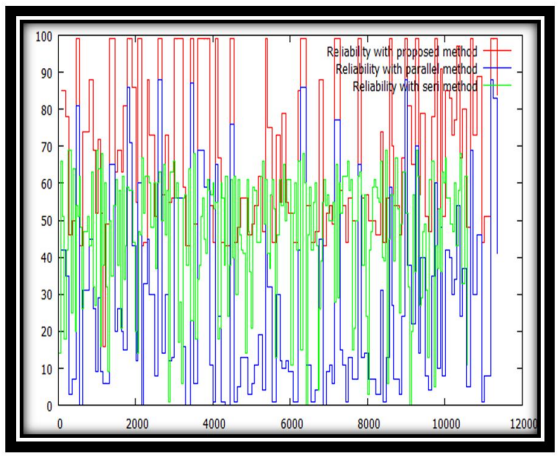


Figure 5. Simulation Diagram

5. CONCLUSION

In this paper, the proposed algorithm has considerably improved reliability through adding redundancy technique to system and using series-parallel modes. Since colored Petri nets involve powerful mathematical power in terms of evaluating software systems, we use VCPNTOOLS simulator to simulate the proposed model. Also, we compare our model with two other models (series and parallel models). The results obviously show that the proposed algorithm and model is preferred to other algorithms that have been proposed up to now.

6. SUGGESTIONS

Since no work is complete and there are certainly some faults in the work and job, the model presented in this paper has some faults. This issue can be considered as a background for performing great projects. With regard to proper application of Petri nets in improving systems, an optimum model has been proposed in this paper to increase reliability in Grid. Comparing this method with previous methods show improvement of efficiency and increasing accessibility.

In this paper, the following suggestions have been proposed:

1) Deleting some hypothesis of the problem and involving more real hypothesis and suggestions for problem solving.
2) Using above mentioned model and generalization of this method for other Grid topologies such as hierarchical and tree topology.
3) Considering data dependency among sub-tasks and expecting some sub-tasks to receive required data from sub-tasks.

7. REFERENCES

[1] Rongfei Zeng,Yixin Jiang,Chuang Lin, and Xuemin (Sherman) Shen "Dependability Analysis of Control Center Networks in Smart Grid using Stochastic Petri Nets" Department of Electrical and Computer Engineering Sept. 2012 (vol. 23 no. 9) pp. 1721-1730.

[2] Gholami Ehsan; Rahmani Amir Masoud; Habibizad navin Ahmad;" Using Monte Carlo Simulation in Grid Computing Systems for Reliability Estimation"; 2009 Eighth International Conference on Networks. PAGE: 380-384 E- ISBN 978-0-7695-3552-4.

[3] Tanenbaum As., Vansteen M.2007, Distributed Systems .Principles and paradigms.Prentice Hall,2007. iSBN 0-13-239227-5.

[4] Allan Mazur∗, Todd Metcalfe;" America's three electric grids: Are efficiency and reliability functions of grid size"; Electric Power Systems Research 89 (2012) 191–195. VOLUME 89 Elsevier.

[5] Mohammad Khanli Leyli, Etminan Far Maryam, Ghaffari Ali.," Reliable Job Scheduler using RFOH in Grid Computing", Journal of Emerging Trends in Computing and Information Sciences, ©2009-2010 CIS Journal. All rights reserved. VOL. 1, NO. 1, JULY 2010 ISSN 2079-8407.

[6] Parsa Saeed , Azadi Parand Fereshteh-.," Estimation of service reliability and performance in grid environment, Journal of King Saud University – Engineering Sciences (2012) 24, 151–157. July 2012.

[7] Y.S. Dai, M. Xie, K.L. Poh.," Reliability Analysis of Grid Computing Systems", Department of Industrial and System Engineering, National University of Singapore. ISBN: 0-306-48496-X.vol. 79, no. 1, pp. 103-112.

[8] Hiroyuki Mori, Hiroki Kakuta" Modified SPEA2 for Probabilistic Reliability Assessment in Smart Grids" , Procedia Computer Science 6 (2011)..

[9] Abawajy j.2004.placement of filereplicas in data environments:lecture Note in computer science (LNCS).springer,heidelborg.3038:66-73.

[10] Azgomi mohammad abdollahi,entezari maleki reza, esfand 87" Reliability modeling distribution of tasks and

calculate a star topology grid services using colored Petri Nets Time".

[11] Dastgheybifar Gholamhoseyn, Ansari Sara, Lotfi Somayeh ,1386 " A new scheduling algorithm for computational grid" Thirteenth Annual Conference of Computer Society of Iran

[11] Naseri a, Rahmani a , Jola, Esfand 1389 " New method for optimizing the time - cost of resource allocation in grid system, and implemented it using gravitational search algorithm" First National Conference on Soft Computing and Information Technology.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 731 - 737, 2013

Survey on Dynamic Resource Allocation Strategy in Cloud Computing Environment

N.Krishnaveni
Dept. of CSE
Erode Sengunthar Engineering College
Thudupathi, India

G.Sivakumar
Dept. of CSE
Erode Sengunthar Engineering College
Thudupathi, India

Abstract-Cloud computing becomes quite popular among cloud users by offering a variety of resources. This is an on demand service because it offers dynamic flexible resource allocation and guaranteed services in pay as-you-use manner to public. In this paper, we present the several dynamic resource allocation techniques and its performance. This paper provides detailed description of the dynamic resource allocation technique in cloud for cloud users and comparative study provides the clear detail about the different techniques.

Keywords: Cloud computing, Dynamic Resource Allocation, Cloud users, Resources, Data center, Virtual machine

1. INTRODUCTION

Cloud computing allows customers to scale up and down their resources based on needs. Cloud computing technology makes the resources as a single point of access to the client and cost is pay per usage. Cloud computing is a computing technology, where a pool of resources are connected in private and public networks and to provide these dynamically scalable infrastructure for application. Cloud computing is not application oriented and this is a service oriented. It offers the virtualized resources to the cloud users. Cloud computing provide dynamic provisioning and thus can allocate machines to store data and add or remove the machines according to the workload demands. Cloud computing platforms such as, those provided by Microsoft, Amazon, Google, IBM. Cloud computing is an environment for sharing resources without the knowledge of the infrastructure and can makes it possible to access the applications and its associated data from anywhere at any time.

Cloud environment provide the four types of cloud.

- Public cloud
- Private cloud
- Hybrid cloud
- Community cloud

Cloud computing offers three types of services

- Software as a service(Saas)
- Platform as a service(Paas)
- Infrastructure as a service(IaaS)

Virtualization technology

Cloud computing is based on the virtualization technology. Virtualization technology is used to allocate the data center resources dynamically based on the application demands.

Virtualization having two types,

- Para virtualization
- Full virtualization

Live migration

Virtual machine live migration technology makes it possible to mapping between the virtual machines (VMs) and the physical machines (PMs) while applications are running. Live migration increase the resource utilization and provide the better performance result.

2. RESOURCE ALLOCATION

In cloud computing, Resource allocation is the process of assigning available resources to the needed cloud applications. Cloud resources can be provisioned on demand in a fine-grained, multiplexed manner. In cloud the resource allocation is based on the infrastructure as a service (IaaS).In cloud platforms, resource allocation takes place at two levels:

- when an application is uploaded to the cloud, the load balancer assigns the requested instances to physical computers, to balance the computational load of multiple applications across physical computers
- When an application receives multiple incoming requests, these requests should be assigned to a specific application instance to balance the

computational load across a set of instances of the same application

Resource allocation techniques should satisfy the following criteria:

- Resource contention arises when two applications try to access the same resource at the same time
- Resource fragmentation arises when the resources are isolated. There would be enough resources but cannot allocate it to the needed application due to fragmentation.
- Scarcity of resources arises when there are limited resources and the demand for resources is high
- The multiple applications needed different types of resources such as cpu,memory,I/O devices and the technique should satisfy that request
- Over provisioning of resources arises when the application gets surplus resources than the demanded one

3. RESEARCH ISSUES IN DYNAMIC RESOURCE ALLOCATION TECHNIQUES

In this paper we have analyzed some of the dynamic resource allocation techniques in cloud environment

Dynamic Optimization of Multi-Attribute Resource Allocation in Self-Organizing Clouds

In Existing system generate the more messages for a single request. The proposed system using the SOC and it achieves the maximized resource utilization and it also delivers optimal execution efficacy.

SOC:

SOC connect a large number of desktop computers on the internet by P2P network. Each participating computer act as a resource provider and resource consumer.

SOC having two main issues:

- Locating a qualified node to satisfy a user task's resource demand with bounded delay
- To optimize a task's execution time by determining the optimal shares of the multi-attribute resources to allocate to the tasks with various QoS constraints, such as the expected execution time

Algorithm:

- Dynamic optimal proportional share
- Multi range query protocol

DOPS:

This algorithm used to redistribute available resources among running tasks dynamically, such that these tasks could use up the maximum capacity of each resource in a node, while each task's execution time can be further minimized.

Procedures:

Slice handler: It is activated to equally scale the amount of resources allocated to tasks.

Event handler: It is used for resource redistribution upon the events of task arrival and completion.

Multi Range Query Protocol:

This algorithm used to locate qualified nodes in the SOC environment; we design a fully-decentralized range query protocol, namely pointer-gossiping CAN (PG-CAN), DOPS to find the qualified resources with minimized contention among requesters based on task's demand. It is unique in that for each task, there is only one query message propagated in the network during the entire discovery.

Range query protocol proactively diffuses resource indexes over the network and randomly route query messages among nodes to locate qualified ones that satisfy tasks' minimal demands. To avoid possibly uneven load distribution and abrupt resource over-utilization caused by un-coordinated node selection process from autonomous Participants.

Dynamic Resource Allocation for Spot Markets in Clouds

As a demand of each VM type can fluctuate independently at run time, it becomes a problem to dynamically allocate data center resources to each spot market to maximize cloud provider's total revenue.

We present a solution to this problem that consists of 2 parts:

- Market analysis for forecasting the demand for each spot market
- A dynamic scheduling and consolidation mechanism that allocate resource to each spot market to maximize total revenue.

Cloud providers specify a fixed price for each type of VM offerings.

- When total demand is much lower than data center capacity, the data center becomes under-utilized, i.e., the cloud provider is to encourage customers to submit more requests.
- When total demand rises over the data center capacity, it is desirable for the cloud provider to motivate the customers to reduce their demand.

A promising solution to this problem is to use market economy to reshape the demand by dynamically adjusting the price of each VM type.

- When total demand is high, the mechanism raises the price to ensure resources are allocated to users who value them the most.
- When total demand is low, the mechanism lowers the prices and provides incentive for customers to increase their demand.

Dynamic resource allocation framework consists of the following components:

Market Analyzer:

- Analyze the market situation and forecast the future demand and supply level
- Predict the future demands use AR(auto regressive model)

Capacity planner:

Capacity planner decide the expected price of each VM

Different pricing schemes:
In the fixed pricing scheme, price of a VM type does not vary with the current supply and demand.
In the uniform pricing scheme, the price of a VM type is adjustable at run-time.
VM scheduler:

- Make online scheduling decision for revenue maximization
- Dynamic resource allocation policy has the multiple machine configuration and this will amplified when the demand pattern changes over the time.

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment

Cloud computing allows business customers to scale up and down their resource usage based on needs. In this paper, using virtualization technology to allocate data center resources dynamically based on application demands and support green computing by optimizing the number of servers in use
Technique:

1. Virtualization technology

2. Skewness

Goals:

Overload avoidance: The capacity of a PM should be sufficient to satisfy the resource needs of all VMs running on it.

Green computing: The number of PMs used should be minimized as long as they can still satisfy the needs of all VMs. Idle PMs can be turned off to save energy.

Virtualization technology: This technology used to allocate datacenter resources based on the application demands.

Skewness: This is used to measure the unevenness multidimensional resource utilization of a server. To minimizing skewness, we can combine different types of workloads.
Skewness can be measured based on,
Hot spot: If the utilization of any of its resources is above a hot threshold. This indicates that the server is overloaded and hence some VMs running on it should be migrated away.
Cold spot: If the utilizations of all its resources are below a cold threshold. This indicates that the server is mostly idle and a potential candidate to turn off to save energy.
Achieve the goals to make the following contributions,
1. Develop a resource allocation system that can avoid overload in the system
2. skewness to measure the uneven utilization of a server.
3. Design a load prediction algorithm that can capture the future resource usages of applications accurately without looking inside the VMs.
Load prediction algorithm:
Exponentially weighted moving average (EWMA): Predict the CPU load and we measure the load every minute and predict the load in the next minute.

Heterogeneity-Aware Resource Allocation and Scheduling in the Cloud

In cloud computing environment data analytics are the key applications and it should be account for the heterogeneity of environments and workloads. In cloud computing environment does not provide the fairness among jobs when multiple jobs share the cluster. In Proposed, Resource allocation on a data analytics system in the cloud to hold the heterogeneity of the underlying platforms and workloads and the architecture shows how to allocate resources to a data analytics cluster in the cloud.

Technique:

Data analytic cluster: Data analytics workloads have heterogeneous resource demands because some workloads may be CPU whereas others are I/O-intensive. In cloud various resource demands of data analytics workloads, we scale the cluster according to demands.

For scaling process the resource allocation strategy having two levels

(1) Divides machines into two pools - core nodes and accelerator nodes

(2) Dynamically adjusts the size of each pool to reduce cost or improve utilization.

Data analytic cloud contains the components are,
Core nodes: Host the data and computations
Accelerator nodes: This is added to the cluster temporarily when additional computing power needed.
Analytic engine:Runs the application both the pools
Cloud driver:

- This manages the nodes allocated to the analytic cloud and decides when to add/remove what type of nodes to/from which pool.
- The user submits the job to the cloud driver with the hints about the job.Cloud driver keeps the history of routinely processed query, this is used to estimate the submission query and update the hints provided.
- This is also monitor the storage system to estimate the incoming data rate. This will predict the resource requirements to process queries and to store data.
- Many productions query are submitted with tight deadlines, the cloud driver will add the nodes to the accelerator pool temporarily to handle the job rather than allocating more core nodes.

- When adding nodes, the cloud driver makes the decision on which resource container to use.

Priority Based Resource Allocation Model for Cloud Computing

Cloud computing is a model which enables on demand network access to a shared pool computing resources. A cloud environment consists of multiple customers requesting for resources in a dynamic environment with possible constraints. In existing system cloud computing, allocating the resource efficiently is a challenging job. The cloud does not show the Qos, SLA.

This paper proposed allocates resource with minimum wastage and provides maximum profit. The developed resource allocation algorithm is based on different parameters like time, cost, No of processor request etc.
Algorithm:

Priority algorithm:
Priority algorithm that mainly decides priority among different user request based on many parameters like cost of resource, time needed to access, task type, number of processors needed to run the job or task.

Resource Allocation Model:
In this model client send the request to the cloud server. The cloud service provider runs the task submitted by the client. The cloud admin decides the priority among the different users request.
Each request consists of different task and it have the different parameters such as ,
Time- computation time needed to complete the particular task,
Processor request- refers to number of processors needed to run the task. More the number of processor, faster will be the completion of task.
Importance- refers to how important the user to a cloud administrator (admin) that is whether the user is old customer to cloud or new customer.
Price- refers to cost charged by cloud admin to cloud users.

Survey on Resource Allocation Strategies in Cloud Computing

In cloud computing, the important problem is to manage the Qos and to maintain the SLA for cloud users that share cloud resources. In this paper proposed Dynamic resource allocation can be based on,
1. Topology Aware Resource Allocation (TARA)
2. Linear scheduling strategy
3.Dynamic resource allocation for parallel data processing

Topology Aware Resource Allocation (TARA):
This allocation scheme based on the Iaas cloud systems. IaaS systems are usually unaware of the hosted application's requirements and therefore allocate resources independently of its needs.
Overcome this problem this method using the prediction engine and genetic algorithm based search.
Prediction engine with lightweight simulator to estimate the performance of the given resource allocation.
Genetic algorithm based search technique that allows TARA to guide the prediction engine

Linear scheduling strategy:
Linear Scheduling performs tasks and resources scheduling respectively. Here, a server node is used to establish the IaaS cloud environment and KVM/Xen virtualization with LSTR scheduling to allocate resources which maximize the system throughput and resource utilization.

Dynamic resource allocation for parallel data processing:
This technique used to allocate and deallocate the resources from a cloud during job execution. Some of the particular task can be assigned to different types of virtual machines and these task are automatically instantiated and terminated during job execution.

Survey on Resource Allocation Strategies in Cloud Computing (2012)

In cloud computing, the important problem is to manage the Qos and to maintain the SLA for cloud users that share cloud resources.
In cloud computing there are many RAS techniques,

Execution time: Estimating the execution time for a job is a hard task for a user and errors are made very often. This paper proposed technique is matchmaking strategy and it is based on Any-Schedulability criteria for assigning jobs to resources in heterogeneous environment.

Policy: This is based on the two levels: security and processor.

Security policy proposed a decentralized user and virtualized resource management for IaaS by adding a new layer called domain in between the user and the virtualized resources. Based on role based access control (RBAC), virtualized resources are allocated to users through domain layer.

Processor policy for resource allocation means the job is allocates to the cluter,then the number of processor needed for the subsequent job allocation. The number of processors in each cluster is binary compatible. Job migration is required when load sharing activities occur.

Virtual machine:The dynamic availability of infrastructure resources and dynamic application demand, a virtual computation environment is able to automatically relocate itself across the infrastructure and scale its resources in cloud environment. This is also based on the load, cost, speed and the type of application.

Utility function: Dynamically manage VMs in IaaS by optimizing some objective function such as minimizing cost function, cost performance function and meeting QoS objectives. The objective function is defined as Utility

property which is selected based on measures of response time, number of QoS, targets met and profit.

Hardware Resource Dependency:Improve the hardware utilization, we propose the Multiple Job Optimization (MJO) scheduler. Jobs can be classified by hardware-resource dependency such as CPU, Network I/O, Disk I/O and memory bound. MJO scheduler can detect the type of jobs and parallel jobs of different categories. Based on the categories, resources are allocated.

Auction:In this method the cloud provider collects the user's proposals and determines the price. This not provides the more profit. This achieved by using market based resource allocation strategy. Dynamically adjust the resources in single VM according to various resource requirements of workloads.

SLA:RAS to focusing on SLA has driven user based QoS parameters to maximize the profit for SaaS providers. The mappings of customer requests in to infrastructure level parameters and policies that minimize the cost by optimizing the resource allocation within a VM.

A Dynamic Resource Allocation Methods for Parallel Data Processing

Nephele is the first data processing framework to explicitly exploit the dynamic and probably heterogeneous. In existing system the resource overload is high. The proposed system increases the efficacy of the scheduling algorithm for the real time cloud computing services. The algorithm utilizes the turnaround time utility efficiently by differentiating it into a gain function for a single task.

The algorithm assigns high priority for early completion task and less priority for abortions/deadlines.
Cloud computing performance can be improved by,

Associate each task with the time utility function (TUF).this is not important to measure the profit when completing a job in time but also account the penalty when a job is aborted or discarded.

In nephele architecture, the client submits the task job manager. Job manager allocate and deallocate VMs. VMS can be differentiated based on the instance type. For example,"m1.small" is a instance type means, it refers 1 cpu core, 1GB RAM, 128GB disk. The task manager receive tasks from the job manager at a time and decides how many and what type of instances job should be executed. The algorithm proves,

- Preemptive scheduling provides the maximum profit than the non-preemptive scheduling.
- Non-Preemptive scheduling provides the maximum penalty than the preemptive scheduling.

Efficient Idle Desktop Consolidation with Partial VM Migration

Idle desktop systems are frequently left powered, often because of applications that maintain network presence. Idle PC consumes up to 60% of its peak power desktop VM often large requiring gigabytes of memory. These VM creates bulk transfer and utilize server memory inefficiently. In existing technique using the ballooning method, this not ensures the quick resume and provides the strain to the network.

Proposed system: Using the partial VM migration technique. This migrates only the working set of an idle VM.it allows user applications to maintain the network presence while the desktop sleeps and to transfer the execution of an Idle VM and it fetches the VM's memory and disk state on –demand.

Partial migration leverages two insights:

- First, the working set of an idle VM is small, often more than an order of magnitude smaller than the total memory allocated to the VM.
- Second, rather than waiting until all state has been transferred to the server before going to sleep for long durations, the desktop can save energy by micro sleeping early and often, whenever the remote partial VM has no outstanding on-demand request for state.

Working set migration: when consolidating a VM from the desktop to the server, partial VM migration transfers memory state only as the VM requires for its execution.

State Access Traces

Mean Memory working set was only 165.63 MiB with standard deviation of 91.38MiB.

The mean size of disk access during idle times was 1.16 MiB with standard deviation of 5.75MiB.

Heuristic Based Resource Allocation Using Virtual Machine Migration: A Cloud Computing Perspective

Virtualization and VM migration capabilities enable the data center to consolidate their computing services and use minimal number of physical servers.
In previous works, the issue of SLA violation has not received thorough analysis. In this work, we devise an algorithm that will keep the migration time minimum as well as minimizing the number of migrations. This will play a major role in avoiding the performance degradation encountered by a migrating VM.

The main aim is to placing the virtual machine using the technique bin packing algorithm and gradient search technique. The heuristic based VM migration scenario is partitioned as follows:

- Determining when a physical server is considered to be overloaded requiring live migration of one or

more VMs from the physical server under consideration.

- Determining when a physical server is considered as being under loaded hence it becomes a good candidate for hosting VMs that are being migrated from overloaded physical servers.
- Selection of VMs that should be migrated from an overloaded physical server. VM selection policy (algorithm) has to be applied to carry out the selection process.
- Finding a new placement of the VMs selected for migration from the overload and physical servers and finding the best physical

This empirical study seeks to achieve the following goals:

- Carrying out the live migration of VMs in a manner that preserves free resources in order to prevent SLA violations
- Optimal utilization of resources
- Performing minimal number of migrations to the extent possible
- Efficient server consolidation through VM migrations

4. COMPARISON OF THE DYNAMIC RESOURCE ALLOCATION TECHNIQUE

TITLE	ADVANTAGE	PARAMETER RESULT
Dynamic Optimization of Multi-Attribute Resource Allocation in Self Organizing Cloud	Locating qualify nodes and optimize task execution time	Throughput Ratio: 60% improvement
Priority Based Resource Allocation Model for Cloud Computing	Resource wastage is minimized	Parameters: No.of users, Time to run, No.of processor, job type, User type
Dynamic Resource Allocation Using Virtual Machine for Cloud Computing Environment	Server overload is minimized	Migration of VM for resource Requirement
Survey on Resource Allocation Strategies in Cloud Computing(2013)	It should maintain the SLA and also manage the Qos	Strategies: Virtual machine,SLA,utility
Heterogeneity Aware Resource Allocation In Cloud	Provide the fairness among jobs when multiple jobs are submitted	The result is based on the Instance Type Ex:m1.small
Dynamic Resource Allocation for Parallel Data Processing in cloud	Overload is avoided	Gain utility: Preemptive>Non-Preemptive Penalty: Non Preemptive> preemptive
Efficient Idle Desktop Consolidation with Partial VM Migration	This migrates only working set of an idle VM	This can deliver the 85%to 104% of the energy saving compare full VM migration
Survey on Resource Allocation in Cloud Computing	This avoids the resource contention and scarcity of resources	Technique: Topology aware resource allocation
Heuristic Based Resource Allocation Using Virtual Machine Migration: A Cloud Computing Perspective	Less SLA violation and less performance degradation	Average: SLA violation Is reduced to 17.64 to 16.44
Dynamic Resource Allocation for Spot Market in Cloud	Total revenue is maximized	Income:15173.28 Loss:1083.63 NetIncome:14089.65

5. COMPARISON OF PERFORMANCE EVALUATION

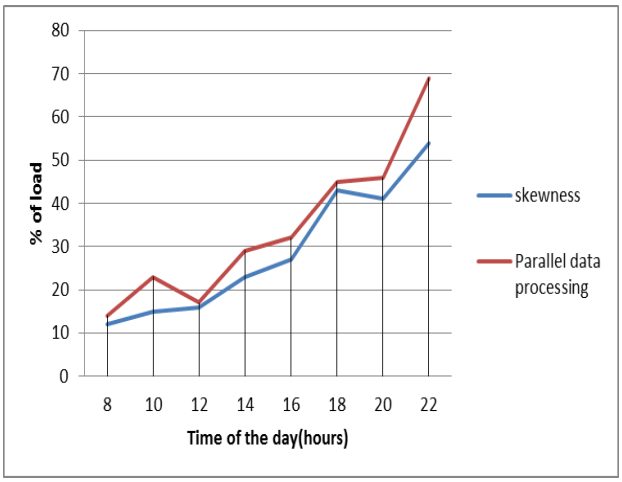


Fig.1 Overload of the server

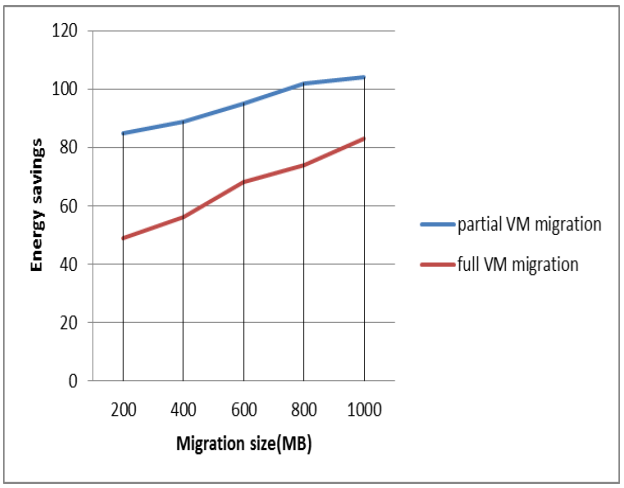


Fig.2 Energy saving based on VM migration

6. CONCLUSION

This paper addresses the theoretic study of various dynamic resource allocation techniques in cloud environment. The detail description of the techniques is summarized and also summarizes the advantages with parameters of the various techniques in cloud computing environment.

7. REFERENCES

[1] Ronak Patel, Sanjay Patel" Survey on Resource Allocation Strategies in Cloud Computing" International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 2, Feb- 2013

[2] K C Gouda, Radhika T V, Akshatha M," Priority based resource allocation model for Cloud computing" International Journal of Science, Engineering and Technology Research (IJSETR)Volume 2, Issue 1, January 2013.

[3] Gunho Leey, Byung-Gon Chunz, Randy H. Katzy," Heterogeneity-Aware Resource Allocation and Scheduling in the Cloud", IJERT-2012.

[4] Nilton Bilay, Eyal de Laray, Kaustubh Joshi_, H. Andr´es Lagar-Cavilla ,Matti Hiltunen_ and Mahadev Satyanarayananz," Efficient Idle Desktop Consolidation with Partial VM Migration", Journal of computer application-2012.

[5] Venkatesa Kumar, V. And S. Palaniswami," A Dynamic Resource Allocation Method for Parallel data processing in Cloud Computing", Journal of Computer Science 8 (5): 780-788, 2012.

[6] Sheng Di and Cho-Li Wang," Dynamic Optimization of Multi-Attribute Resource Allocation in Self-Organizing Clouds", IEEE Transactions on parallel and distributed systems, - 2013.

[7] V.Vinothina, Dr.R.Sridaran, Dr.padmavathiganapathi," A Survey on Resource Allocation Strategies in Cloud Computing "International Journal of Advanced Computer Science and Applications, Vol. 3, No.6, 2012.

[8] Qi Zhang, Eren G¨urses, Raouf Boutaba, Jin Xiao," Dynamic Resource Allocation for Spot Markets in Clouds", Journal of computer science-2012.

[9] Zhen Xiao, Senior Member, IEEE, Weijia Song, and Qi Chen," Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment", IEEE Transactions on parallel and distributed systems, vol. 24, no. 6, June 2013.

[10] Ts`epomofolo, R Suchithra," Heuristic Based Resource Allocation Using Virtual Machine Migration: A Cloud Computing Perspective", International Refereed Journal of Engineering and Science (IRJES) Volume 2, Issue 5(May 2013)

.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 738 - 742, 2013

Fault Injection Test Bed for Clock Violation

E. Kavitha
Dept. of ECE
MRE College
Secunderabad, AP, India

P.S. Indrani
Dept. of ECE
MRE College
Secunderabad, AP, India

M. J. C. Prasad
Dept. of ECE
MRE College
Secunderabad, AP, India

Abstract: In this paper, the International Data Encryption (IDEA) algorithm synthesis models will be used as test encryption algorithm. The Xilinx Digital clock manager component will be used for generation of clocks for different frequencies and phase shifts. The encryption output with faults introduced and without faults introduced is compared as a function of ratio of used clock frequency and maximum frequency of operation reported by synthesis tool. The clock generation, clock switching, interface adopter to IDEA core and UART interface will be realized and tested in FPGA hardware in integrated form. FPGA based test bed is realized for injecting faults through clock glitches, to result in setup and hold violations. The UART interface is realized on FPGA to provide PC based controlling for this fault injection. Xilinx chip scope tools will be used for verifying the output at various levels in FPGA hardware.

Keywords: International Data Encryption algorithm (IDEA), UART, FPGA, Digital clock Manager (DCM), PLL.

1. INTRODUCTION

To increase performance, a lot of cryptographic algorithms are implemented in hardware for that purpose FPGAs are frequently used for this purpose. Such implementations are however prone to various types of attacks intended to compromise their security. One possible way to perform such an attack is to inject transient faults affecting the normal circuit operation. One of the cryptographic algorithms most commonly used is the International Data Encryption algorithm (IDEA) is a block cipher [2]. The mentioned algorithm works on 64-bit plain text and cipher text block (at one time). For encryption, the 64-bit plain text is divided into four 16-bits sub-blocks. I denote these four blocks as X1 (16 bits), X2 (16 bits), X3 (16 bits) and X4 (16 bits). Each of these blocks will perform of operation 8 ROUNDS and one OUTPUT TRANSFORMATION phase.

A long term objective of our research is to develop an efficient method for protecting FPGA-based implementations of cryptographic algorithms through effective concurrent testing of various types of faults, including faults injected by the attackers [3]. An essential part of this research is to develop a method and tool for the evaluation of susceptibility of FPGA based circuits to fault injection attacks. In this paper, I present such a method and tool. It allows us to examine an FPGA-based circuit, in particular an implementation of a cryptographic algorithm, subjected to a fault injection attack based on clock glitching [7].

2. EXPERIMENTAL SETUP

The circuit under test (CUT) and the tester are both implemented on a low cost FPGA Spartan 3E development board. I used VHDL for defining custom components and Xilinx chip scope which is the system on-chip building tool, for creating standard library components and connections.

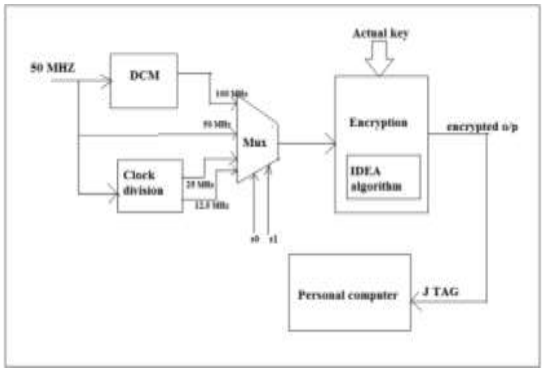


Figure 1: Experimental setup

A simplified diagram of experimental setup implemented in the device is shown in above Fig 1.

2.1 Analysis of IDEA

In IDEA algorithm, I taken input text of size 64 bits at a time and divide it in evenly; i.e., 64 bit plain text is divided into 4 sub-blocks, each of 16 bits in size [1]. The basic operations needed in the entire process for 8 rounds are

1. Multiplication modulo 2^**16** +1.
2. Addition modulo 2^**16**.
3. Bitwise XOR.

And, operations needed in the OUTPUT TRANSFORMATION phase –

1. Multiplication module 2^**16** +1.
2. Addition modulo 2^**16**.

All the above mentioned operations are performed on 16 bit sub-blocks. For simplicity of expressing the operations. Now, let us take a look on the key generation for the encryption process while using the 25-bit circular left shift operation on the original key, it produce other subsequent sub-keys, used in different rounds [2]. For instance, among the total no. of 52 keys- Sub-key Z1 is having first 16bits of the original key, sub-key Z2 is having the next 16 bits, and so on till sub-key Z6; i.e., for ROUND1, sub-keys Z1 to Z6 use first 16x6= 96 bits of the original cipher key. In the ROUND2, sub-key Z7 &Z8 take the rest of the bits (bits 97 to 128) of the original cipher key. Then we perform circular left shift (by 25bits) operation on the original key.

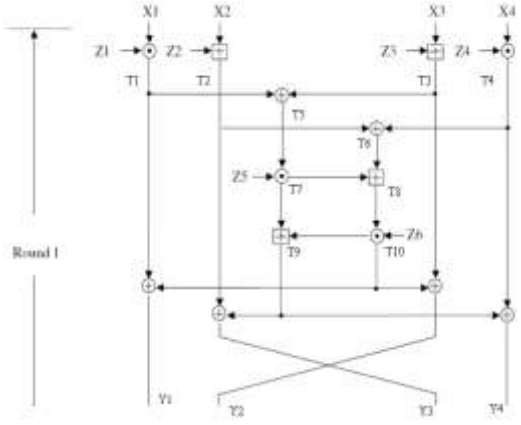


Figure 2: IDEA Encryption/Decryption sub key generation Architecture

As a result the 26th bit of the original key shifted to the first position and becomes the first bit (of the new shifted key) and the 25th bit of the original key is moves to the last position and becomes the 128th bit (after first shift). This process continues till ROUND8, and also in the OUTPUTPUT TRANSFORMATION phase; i.e., after the ROUND8, the key is again shifted left by 25 bits and the first 64 bits of the shifted key is taken for use, and used as sub-keys Z49 to Z52 in the OUTPUT TRANSFORMATIONMATION phase [2].

2.1.1 Output transformation stage

The final round of IDEA algorithm is also called output transformation stage. It only uses 4 sub-keys. The block diagram of final round is given below. The VHDL code for IDEA final round module is given.

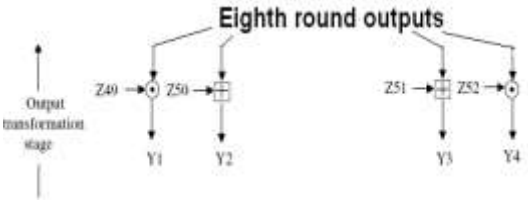


Figure 3: Block diagram of output transformation stage

The general IDEA architecture uses eight rounds with total 48 keys and final output transformation round with 4 sub-keys is implemented in VHDL using structural modeling style. The IDEA Decryption module also uses the same hardware, but the Decryption sub-keys are different. The encryption followed by decryption module is used for testing the complete IDEA algorithm with the following input.

Main Key: Z = (5a14 fb3e 021c 79e0 6081 46a0 117b ff03)
64-bit plaintext: X = (X1, X2, X3, X4) = (7fa9, 1c3, ffb3, df05)

The same text is used in simulating the other IDEA architectures.

2.2. Fault injection

The basic idea of our implementation of the fault injection based on clock glitching is to switch from a normal operation clock to a faster clock; so that one clock cycle is slightly shorter than CUT can handle [6].This idea is depicted in Figure 4. In order to generate single faults, the frequency of the faster clock has to be adjusted very precisely, more accurately than can be achieved using an on-chip PLL circuitry for clock generation or phase shifting. An external clock signal generated by Tektronix AWG 5002B Arbitrary Waveform Generator is used instead. The external clock is

going to feeds an internal PLL circuitry where it is divided by 4 to produce the slower clock. It also passes through unchanged to produce the faster (high speed) clock.

To switch clocks, I using a Clock Control Block, the dedicated clock management built-in component available in the device. The result of the operations on clock signals is shown in Figure 4, as "output clock". The last trace in Figure 4 is the real output clock registered by the 1 GHz Tektronix oscilloscope. The faster clock frequency is 150 MHz and it can be noticed that the signal is not distorted too much [7]. Moreover, additional measurements, made by the MXG-9810AVolcraft frequency counter show that the faster clock has the same frequency (with accuracy of 1 Hz) as the clock supplied to the FPGA by the waveform generator.

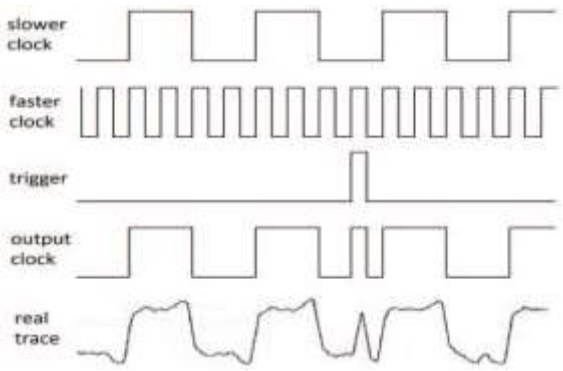


Figure 4: Clock glitch generation

3. TESTING PROCEDURE

Steps involved in project work

To minimize the test application time, the IDEA circuit is performed using the FPGA Spartan 3E, by running a dedicated application. Only the final results are sent through the host computer for display. Implementation of IDEA algorithm in VHDL coding. Implementation of sub key generation module from 128 bit key in VHDL and the basic Arithmetic and logic blocks in VHDL. The VHDL structural modeling of 9 rounds for IDEA encryption module is programmed.

VHDL structural modeling of 9 rounds for IDEA decryption module. Writing test bindles for individual components and also for top level modules and Simulation using modelsim. Verifying the modelsim outputs with expected results. Synthesizing the developed IDEA modules on Xilinx Spartan 3E FPGA using Xilinx ISE tool. Downloading the IDEA encryption and Decryption modules on Spartan 3E development board using the IMPACT tool. The Faults will be introduced and corresponding output results are observed and the Clock generator is going generate the clock and oscilloscope is going to generate the different clock frequencies [6-7]. The FPGA Spartan 3E kit is connected to host computer as shown in below figure.

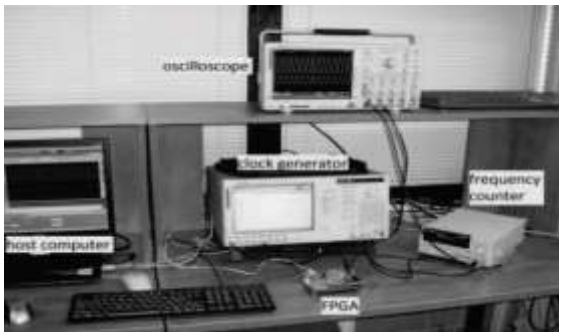


Figure 5: Measurement setup

4. OBSERVATIONS AND RESULTS

4.1. Chip scope resullts

IDEA algorithm has a maximum freqency of approximately 19MHZ frequency. So the clock freqency applied should be less than this freqency or approximately around this frequency. If we apply the freqency more than this freqency then the metastability condition takes place and the output will be corrupted depending on the applied input clock freqencies. The chipscope results of idea encryption algorithm with the corresponding different clock freqencies as shown in figure.

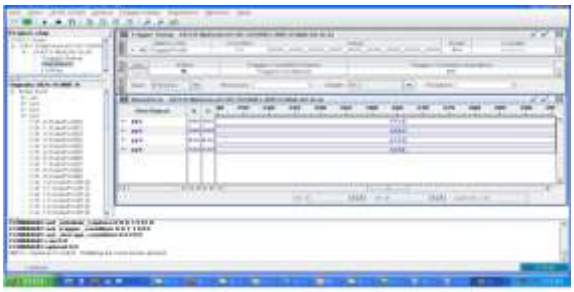


Figure 6: clock frequency of 12.5 MHZ frequency

The below chipscope results show how the idea encryption algorithm is behaving with the corresponding 100 MHZ clock frq.The chipscope results show how the idea encryption

algorithm is behaving with the corresponding 50 HZ clock freqency.

Figure 7: clock frequency of 50 MHZ frequency

4.2. Observed output results

The analyzer outputs are observed for different clock frequencies (12.5 MHZ, 25 MHZ, 50 MHZ, and 100 MHZ). The input is FFCC, 8BBD, E1D6, 849B which is generated by input generator.

S.NO	INPUT	FREQENCY	ANALYSER OUTPUT
1.	FFCC 8BBD E1D6 849B	12.5 MHZ	FFCC 8BBD E1D6 849B
2.	FFCC 8BBD E1D6 849B	25 MHZ	FFCC 8BBD E1D6 849B
3.	FFCC 8BBD E1D6 849B	50 MHZ	5B89,68SD,AFD6,etc.,, 94D6,3589,5E91,etc.,, B66F,FB62,08F9,etc.,, 636A,1E53,557C,etc.,,
4.	FFCC 8BBD E1D6 849B	100 MHZ	8802,7CC7,FE3F,C197,BA1B,etc.,, AF9C,4D0C,1310,7B4B,B56D,etc.,, BA9B,D913,BF1B,86E3,682B,etc.,, FD87,3AFE,CB54,4A59,504F,etc.,,

Table 1: output comparison table

5. CONCLUSION

The presented method and tool for injecting faults in an FPGA Spartan 3E circuit, based on Clock glitching and it has some unique features that allow us to thoroughly examine and analyze the impact of such faults on the operation of the circuit. The IDEA (International Data Encryption Algorithm) is a strong block-cipher text.

Though there are many arithmetic operations involved in the entire algorithm, only three different of operations are involved (as mentioned above). As the cipher key size is 128bits, in that respect IDEA is too strong (having taken care for weak keys).In particular, through recise adjustment of the frequency of an external clock Generator; we can control the number of faults occurring at the output of the circuit under Test.

6. FUTURE SCOPE

The presented solution is intended for injecting a single fault (single clock glitch) during an encryption operation. For more complex circuits the more complex experimental setup can be implemented, to allow dynamic configuration of fault injection conditions .In particular, the trigger unit could be redesigned and interfaced with the Avalon bus, so that it could be reconfigured by the software processor, depending on the testing scenario.

In order to decrease the testing time, the NIOS II processor can be clocked by an independent clock freq as faster than the external clock freq is divided by 4 and the solution requires some changes to the interface adaptor unit to account for a difference in clock frequencies for the NIOS II processor and the circuit under test. The proposed approach can be used not only in the case when the CUT is implemented in the same FPGA; Although it appears that only small changes need to be done to our experimental setup, no attempt has been made to verify this idea. It makes the algorithm more secure and less susceptible to cryptanalysis.

7. Acknowledgment

I E. Kavitha would like to thank P. S. Indrani Associate professor, who guided me through out to complete my work successfully. I would like to thank my HOD (ECE Dept.) Dr. M. J. C. Prasad for providing us constant support and providing us the resources needed.

8. REFERENCE

[1]. Blum M. and Gold wasser S., “An efficient probabilistic public-key encryption scheme which hides all partial information,” Advances in Cryptology-CRYPTO’84, Lecture notes in computer science (Springer-Verlag), pp.289-299, (1995).

[2]. Biryukov, Alex; Nakahara, Jorge Jr.; Preneel, Bart; Vandewalle, Joos, "New Weak-Key Classes of IDEA", Information and Communications Security, 4th International Conference, ICICS 2002.

[3]. William Stalling “Cryptography and Network Security”.

[4]. Bruce Schiener “Applied Cryptography “.

[5] Chang H.S., “International Data Encryption Algorithm” CS-627-1 Fall, 2004.

[6] K. Bousselam, G. Di Natale, M-L.Flottes, B. Rouzeyre, "Evaluation of concurrent error detection techniques on the Advanced EncryptionStandard", Proc. 16th IEEE On-Line Testing Symposium, 2010.

[7] J. Balasch, B. Gierlichs and I. Verbauwhede, "An In-depth and Black-box Characterization of the Effects of Clock Glitches on 8-bit MCUs",Proc. Workshop on Fault Diagnosis and Tolerance in Cryptography,2011.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 743 - 747, 2013, ISSN: 2319–8656

A New Design of Multiplier using Modified Booth Algorithm and Reversible Gate Logic

K.Nagarjun
Department of ECE
Vardhaman College of Engineering,
Hyderabad, India

S.Srinivas
Department of ECE
Vardhaman College of Engineering,
Hyderabad, India

Abstract: In this paper we propose a new concept for multiplication by using modified booth algorithm and reversible logic functions. Modified booth algorithm produces less delay compare to normal multiplication process. Modified booth algorithm reduces the number partial products which will reduces maximum delay count a the output. by combining modified booth algorithm with reversible gate logic it will produces further less delay compare to all other. In the past years reversible logic functions has developed as an important research area. Implementing reversible logic has the advantage of reducing the gate count, garbage outputs as well as constant inputs. Addition subtraction operations are realized using reversible DKG gate. This modified booth algorithm with reversible gate logic are synthesized and simulated by using Xilinx 13.2 ISE simulator.

Keywords: MBE, Reversible logic gate, DKG, reversible logic circuits. Carry save adder tree.

1. INTRODUCTION

The Continuous advances of the microelectronic technologies makes better use of input energy, to encode the data more efficiently, to transmit the information more faster and reliable, etc. In Particular, many of these technologies handle low power consumption to meet the requirements of various outboard applications. In these applications, a multiplier is a fundamental arithmetic unit and used in a great extent in circuits. The fastness of multiplication and addition arithmetic's decides the execution speed and performance of the total calculation. Because the multiplier needs the longest delay within the basic operational blocks in digital systems, the critical path is evaluated by multiplier in general. For high speed multiplication, the modified radix-4 booth's algorithm (MBA)[1] is generally used however this cannot completely solve the problem.

In general, the multiplier uses Booth's algorithm and array of full adders (FAs)[3], or Wallace tree rather than the array of FAs., i.e., this multiplier primarily consists of the three parts: Booth encoder[6], a tree to compact the partial products such as Wallace tree[1], and the final adder[1]. Because Wallace tree is to add the partial products. The most efficient way to gain the fastness of a multiplier is to cut down the number of the partial products as multiplication proceeds a series of additions for the partial products. To cut down the number of calculation steps for the partial products, MBA[1] algorithm has been employed mostly where Wallace tree has taken the role of increasing the fastness to add the partial products. To increase the fastness of the MBA algorithm, many parallel multiplication architectures have been explored and they have employed to various digital filtering calculations.

here In this design we use reversible logic gates[2] in the place of full adders to reduce the power and delay. A reversible logic circuit should have features like use minimum number of reversible gates[2], use minimum number of garbage outputs, use minimum constant inputs.

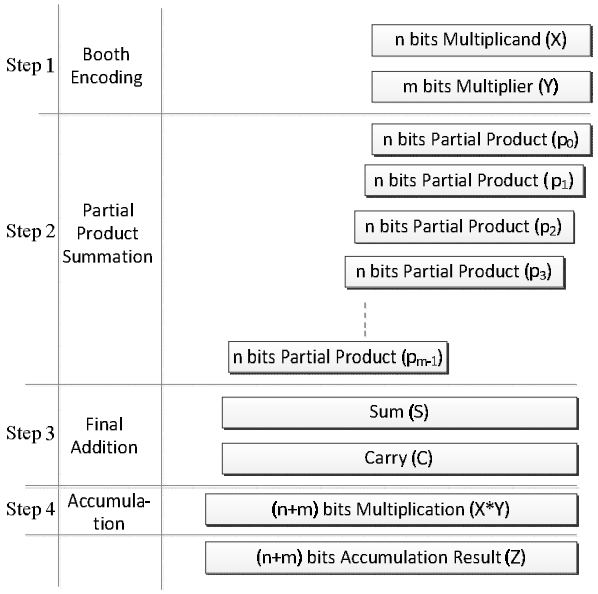


Fig 1: Basic arithmetic steps for multiplication and accumulation

In this paper, A New Design of Multiplier using Modified Booth Algorithm And Reversible Gate Logic are implemented.Section2 discusses overview of MAC.Section3 introduces reversible logic gate function Section4 covers the proposed MAC and architecture of CSA tree. Experimental results are showing the simulation results of proposed design.

2. OVERVIEW OF MAC

In this section, general basic MAC[1] operation is introduced. A multiplier can be divided into three functional steps. The first is radix-2 modified Booth encoding[1] in which a partial product is produced from the multiplicand and the multiplier. second is the adder array or partial product compression to

add all the partial products and change them into the form of sum and carry. The last step is the final addition in which the final multiplication result is produced by adding the sum and carry. If the process to accumulate the multiplied results is included, a MAC consists of four steps, as shown in Fig. 1, which shows the operational steps.

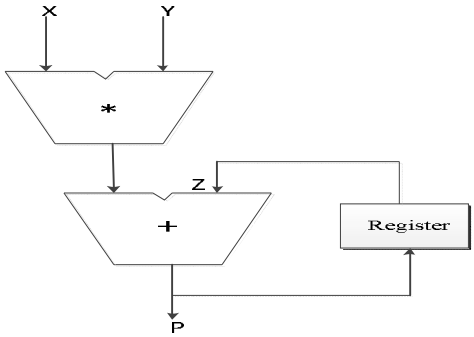


Fig 2: Hardware architecture of general MAC

A Basic hardware architecture of this MAC is shown in Fig. 2. It carries out the multiplication operation by multiplying the input multiplier X and the multiplicand Y. This result is added to the previous multiplication result Z as accumulation step. Fig. 2. Hardware architecture of the general MAC. The -bit 2's complement binary number can be expressed as

$$X = -2^{N-1}x_{N-1} + \sum_{i=0}^{N-2} x_i 2^i \quad x_i \in 0,1. \qquad \ldots. (1)$$

If the equation (1) is expressed in base-4 type redundant sign digit form to apply the radix-2 Booth's algorithm, it would be.

$$X = \sum_{i=0}^{N/2-1} d_i 4_i \qquad \ldots.. (2)$$

$$d_i = -2x_{2i+1} + x_{2i} + x_{2i-1.} \qquad \ldots.. (3)$$

If the equation (2) is used, multiplication can be expressed as

$$X \times Y = \sum_{i=0}^{N/2-1} d_i 2^{2i} Y \qquad \ldots..(4)$$

If these equations are used, the before mentioned multiplication accumulation results can be expressed as

$$P = X \times Y + Z = \sum_{i=0}^{N/2-1} d_i 2^i Y + \sum_{j=0}^{2N-1} z_i 2^i. \qquad \ldots\ldots.(5)$$

Each of two terms on the right side of (5) is evaluated independently and the final result is produced by adding the two results. The MAC architecture enforced by (5) is called the standard design . If the N-bit data are multiplied, the number of generated partial products is proportional to N . In order to add that partial products serially, the time of execution is also proportional to N . The fastest architecture of a multiplier, which uses radix-2 Booth encoding to generate partial products and a Wallace tree based on CSA as the adder array to add the partial products. If radix-2 Booth encoding is used, the number of partial products, i.e., the inputs to the Wallace tree, is reduced to half, resulting in the decrease in CSA tree step. In addition, the signed multiplication based on 2's complement numbers is also possible. Due to these reasons, most current used multipliers adopt the Booth encoding.

3.Reversible Logic DKG Gate

Reversible function is the main objective of the reversible logic theory. A 4* 4 reversible logic DKG[2] gate that can work uniquely as a reversible Full adder and a reversible Full subtractor is shown in Fig 3. It can be verified that input pattern representing to a particular output pattern can be uniquely determined. If the input A=0, the proposed gate acts as a reversible Full adder DKG gate, and if input A=1, then it acts as a reversible Full subtractor DKG gate. It has been proved that at least the two garbage outputs are required by the reversible full-adder circuit to make the output combinations unique.

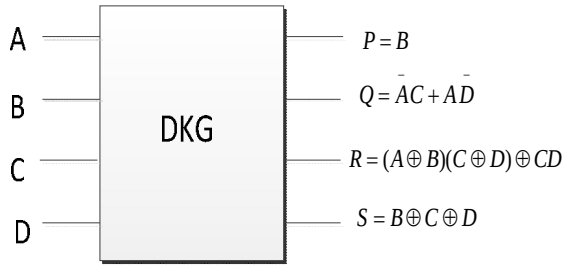


Fig 3: Reversible DKG gate

The binary full adder/subtractor is capable of handling the one bit of each input along with a carry in/borrow in generated as a carry out/ borrow from addition of previous lower order bit position. n binary full adders/subtractors are cascaded to add two binary numbers each of n bits.

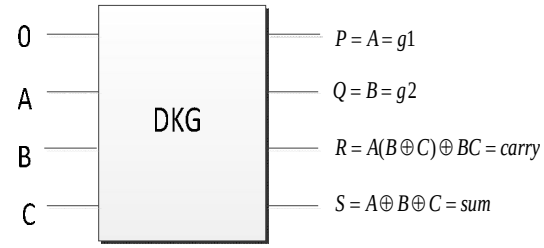


Fig 4: DKG gate implemented as Full adder

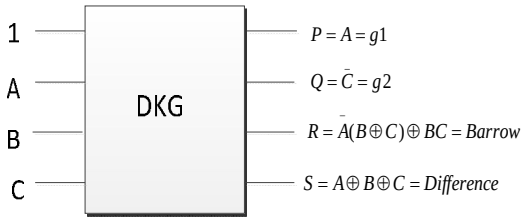


Fig 5: DKG gate implemented as Full subtractor

A Parallel adder/subtractor is an cascaded of full adders/subtractors and inputs are simultaneously applied. The

carry/borrow produced at a stage is propagated to the next stage. When the control input A=0, the circuit behaves as a parallel adder, generates a 4 bit sum and a carry out, as shown in Fig 4. If the control input A=1, the circuit behaves as a parallel subtractor, generates a 4 bit difference and a borrow out, as shown in Fig 5.

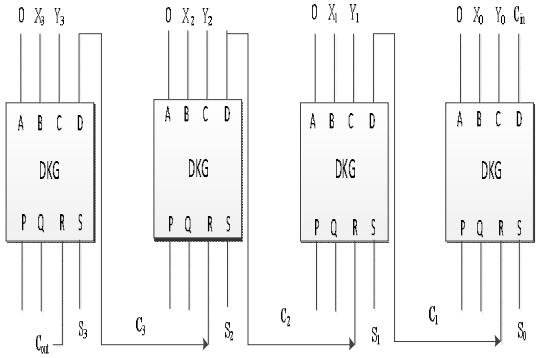


Fig 6: Reversible Parallel adder when A/S=0

4. PROPOSED ARCHITECTURE

In this section, By examining the various architectures, we have observed that delay can be improved by using higher radix MBA which reduces number of partial product rows that eventually reduces number of multiplication thereby improving speed. Thus we propose a new high speed and area efficient MAC architectures which will be an improvement over the existing Architecture by replacing Radix-2 with Radix-4 and Modified Booth Encoder in the Multiplication Stage. The output of multiplication and accumulation stages will be combined using hybrid reduction through CSA, CLA and HA which enhancing speed and efficiency. In addition, the Final Stage of CSA will include a New Reversible logic DKG gate was replaced a full adder. The new adder will be an improvement over the existing CSA. by replacing Reversible logic DKG[2] gate in CSA The design will be implemented using Verilog language and simulated using Xilinx ISE13.2 Simulator. We expect the proposed MAC will be useful in high performance signal processing system.

In Fig. 7 the proposed architecture were shown and it contain booth encoder and CSA & Accumulator tree and the tree contains half adder reversible logic DKG gate which will reduce the delay of the adder and a final adder to add the final sum and carry.

The architecture of the hybrid-type CSA that complies with the operation of the proposed MAC is shown in Fig.8, which performs 8X8-bit operation. It was formed based on the previous architectures. In Fig.8, S_i is to simplify the sign expansion and N_i is to compensate 1's complement number into 2's complement number. S[i] and C[i] and correspond to the i th bit of the feedback sum and carry. Z[i] is the i th bit of the sum of the lower bits for each partial product that were added in advance and Z[i]is the previous result. In addition corresponds to the i th bit of the I th partial product. Since the multiplier is for 8 bits, totally four partial products(P_0[7:0]~P_3[7:0]) are generated from the Booth encoder.

,

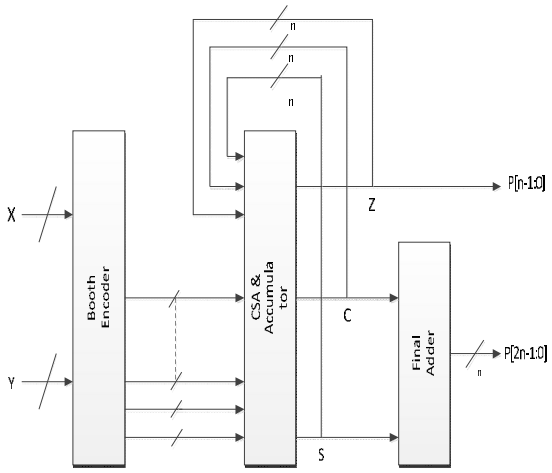


Fig 7: Hardware architecture of the proposed MAC

This CSA requires at least four rows of DKG full adders for the four partial products. Thus, totally five DKG full adder rows are necessary since one more level of rows are needed for accumulation. For an 8-bit MAC operation, the level of CSA is (n/2+1) . The white square in Fig. 8 represents an reversible DKG full adder and the gray square is a half adder (HA). The rectangular symbol with five inputs is a 2-bit CLA with a carry input.

The critical path in this CSA is determined by the 2-bit CLA. It is also possible to use FAs to implement the CSA without CLA. However, if the lower bits of the previously generated partial product are not processed in advance by the CLAs, the number of bits for the final adder will increase. When the entire multiplier or MAC is considered, it degrades the performance.

the characteristics of the proposed CSA architecture have been summarized and briefly compared with other architectures. For the number system, the proposed CSA uses 1's complement, but ours uses a modified CSA array without sign extension. The biggest difference between ours and the others is the type of values that is fed back for accumulation. Ours has the smallest number of inputs to the final adder.
In the carry save adder architecture more number of full adders were used and that will replaced by reversible DKG gate in the proposed system by this the overall performance was further increased the half adders in carry save adder were used as it is because no bigger difference in the normal half adder and reversible DKG half adder.

In addition, we compared the proposed architecture with that the previous booth multiplier. Because of the difficulties in comparing other factors, only delay is compared. The sizes of both MACs were 8X8 bits and implemented by using Xilinx 13.2 simulator. The delay of ours was 30.38ns while in previous it was 34.35ns,which means that ours improved about 4.2ns of the delay performance. This improvement is mainly due to the reversible logic DKG full adder. The architecture from the previous should include a normal full adders that will consumes maximum delay for that we used is very effective and gives maximum performance.

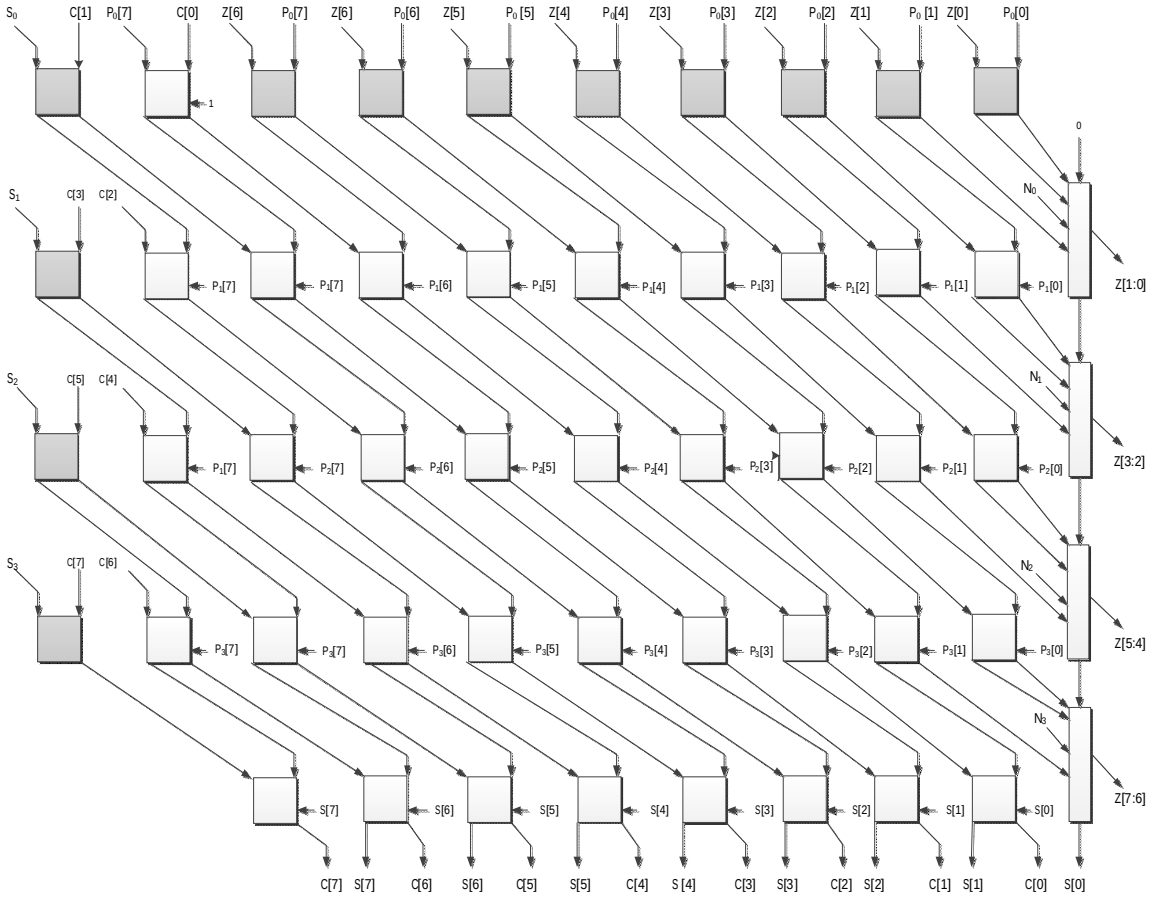


Fig 8: Hardware architecture of proposed CSA tree

5. EXPERIMENTAL RESULTS

The proposed architecture was simulated and synthesized by using Xilinx 13.2 tool and cadence virtuoso

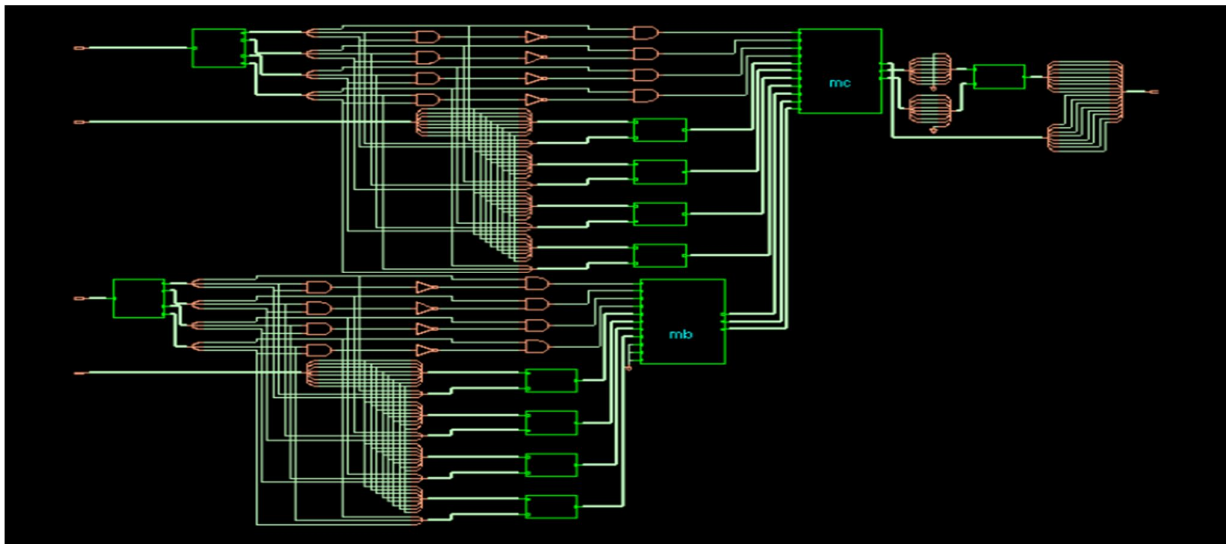


Fig 9: TOP Level Schematic of proposed MAC

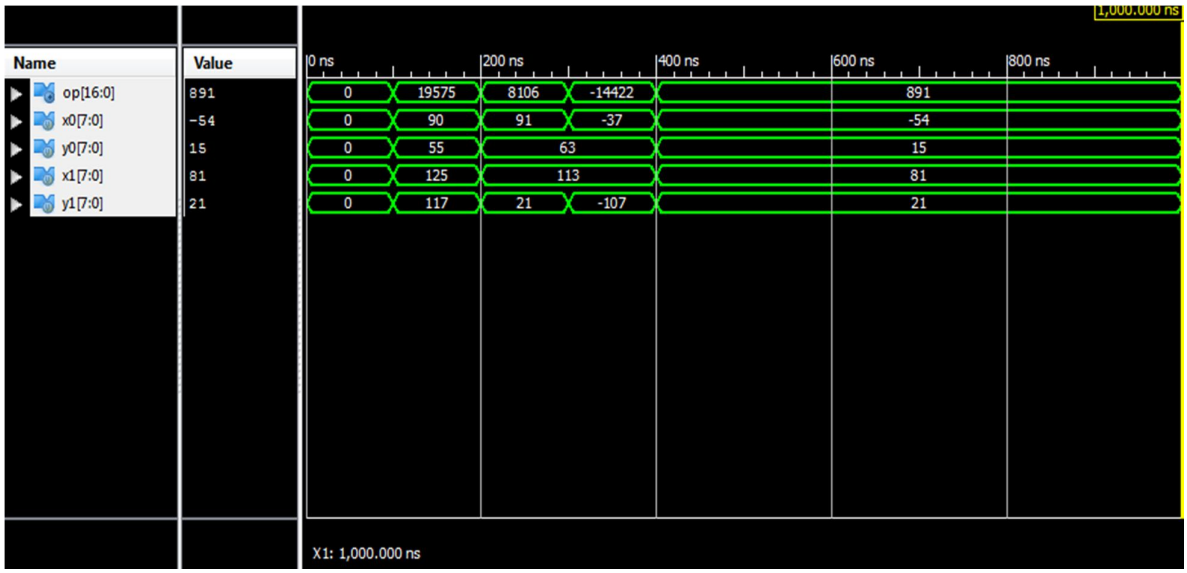


Fig 10: Output of Modified booth algorithm with reversible gate logic

Table 1.Comparison table of different parameters

Parameter	MBE [1]	MBE with reversible gate logic
No. of slices used	178 out of 4656 (3%)	162 out of 4656 (3%)
No. of LUT's	355 out of 9312(3%)	294 out of 9312 (3%)
Total Delay	34.535ns	30.38ns

6.CONCLUSION

In this paper, a new MAC architecture to perform the multiplication-accumulation, To efficiently process the digital signal processing and multimedia application this architecture was proposed. the overall MAC performance has been improved By eliminating the independent accumulation process that has the greatest delay and merging it to the compression process of the partial products, almost twice as much as in the previous work and by replacing full adder in the CSA with reversible logic gate further improves the performance. The proposed hardware was implemented and synthesized through Xilinx ISE 13.2 tool. Consequently, the proposed architecture can be used effectively where we requiring high throughput such as a real-time digital signal processing.

7.REFERENCES

[1] "A New VLSI Architecture of Parallel Multiplier–Accumulator Based on Radix-2 Modified Booth Algorithm "Young-Ho Seo, Member,IEEE,and Dong-Wook Kim, Member, IEEE

[2] "A Distinguish Between Reversible And Conventional Logic Gates" B.Raghu Kanth, B.Murali Krishna , M. Sridhar , V.G. Shanti Swaroop,

[3] O. L. MacSorley, "High speed arithmetic in binary computers," Proc.IRE, vol. 49, pp. 67–91, Jan. 1961.

[4] S. Weser and M. J. Flynn, Introduction to Arithmetic for Digital Systems Designers. New York: Holt, Rinehart and Winston, 1982.

[5] A.R. Omondi ,Computer Arithmetic Systems. Englewood Cliffs, NJ:Prentice-Hall, 1994.

[6] A.D.Booth,"A signed binary multiplication technique," Quart. J.Math., vol. IV, pp. 236–240, 1952.

[7] C. S. Wallace, "A suggestion for a fast multiplier," IEEE Trans. Electron Computer., vol. EC-13, no. 1, pp. 14–17, Feb. 1964.

[8] A.R. Cooper, "Parallel architecture modified booth multiplier", Proc. Inst. Electr Engg. Vol. 135, pp 1251-1988.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 748 - 751, 2013, ISSN: 2319–8656

An Efficient Anonymous Location Service for Geographic Ad Hoc Routing in MANET

Sathya.B
Maharaja Engineering college
Avinashi, TamilNadu,
India

Umarani.S
Maharaja Engineering college
Avinashi, TamilNadu
India

Archana.S.S
R.V.S College of Engg. & Tech
Coimbatore, TamilNadu
India

Rajeshwari.P
Maharaja Engineering college
Avinashi, TamilNadu
India

Sumathi.V.S
Maharaja Engineering college
Avinashi, TamilNadu
India

Abstract: Anonymous routing schemes in MANETs can be classified into on-demand or reactive routing methods, proactive routing methods and anonymous middleware routing method. A finer classification of reactive routing methods includes hop-by-hop encryption and redundant traffic routing which either generate high cost or cannot provide full anonymity protection to data sources, destinations, and routes. An Anonymous Location-based Efficient Routing protocol (ALERT) was used to offer high anonymity protection at a low cost. Like other anonymity routing algorithms, ALERT is not completely bulletproof to all attacks. To prevent the occurrence of stronger and active attackers, we propose a Secure Cryptographic Based Mix-Zones Routing Protocol (SCMIX).The idea for mix-zones is to prevent the adversary from accessing the content of messages, including the Node's signatures. All legitimate nodes within the mix-zone obtain a symmetric key and utilize this key to encrypt all their messages while within the zone.

Keywords: Anonymity, GPSR, Mix-Zone, Routing protocol, Zone Partition.

1. INTRODUCTION

A Mobile Ad Hoc Networks (MANET) is an autonomous system of mobile nodes. It consists of mobile platforms for example a router with multiple hosts and wireless communications devices. Herein simply referred to as 'nodes' which are free to move. It also may operate in isolation or may have gateways to and interface with fixed network. There are many important research questions in MANET. However, power efficiency is one of the most important issues. It is important to realize that issues such as QoS support, TCP performance, speed of routing repair process and others are secondary if nodes have a high probability of running out of energy resources.

Energy awareness in wireless ad hoc networks actually spans across several communication layers. Advances in battery technology are very slow compared to the results achieved in integrated circuit technology particularly in comparison to the rate of growth in communication speeds. Therefore, saving transmission power represents one of the most significant methods for long term wireless system performance.

2. LITERATURE SURVEY

An Anonymous Location-based Efficient Routing protocol (ALERT) [1] dynamically partitions the network field into zones and randomly chooses nodes in zones as intermediate relay nodes, which form a non-traceable anonymous route. In addition, it hides the data initiator/receiver among many initiators/receivers to strengthen source and destination anonymity protection. Thus, ALERT offers anonymity protection to sources, destinations, and routes. It also has strategies to effectively counter intersection and timing attacks. ALERT achieves better route anonymity protection and lower cost compared to other anonymous routing protocols. Also, ALERT achieves comparable routing efficiency to the GPSR geographical routing protocol.

ALARM: Anonymous Location Aided Routing in Suspicious MANETs [2] addresses a number of issues arising in suspicious location-based MANET settings by designing and analyzing a privacy-preserving and secure link-state based routing protocol (ALARM). ALARM uses node's current locations to securely disseminate and construct topology snapshots and forward data. With the aid of advanced cryptographic techniques (e.g., group signatures), ALARM provides both security and privacy features, including node authentication, data integrity, anonymity, and non-traceability (tracking-resistance). It also offers protection against passive and active insider and outsider attacks. It also offers resistance to certain insider attacks.

Anonymous Geo Forwarding in MANETs through Location Cloaking [3] addresses the problem of destination anonymity for applications in mobile ad hoc networks where geographic information is ready for use in both ad hoc routing and Internet services and proposes protocols that use the destination position

to generate a geographic area called an anonymity zone (AZ). A packet for a destination is delivered to all the nodes in the AZ, which make up the anonymity set. The size of the anonymity set may decrease because nodes are mobile, yet the corresponding anonymity set management is simple. We design techniques to further improve node anonymity and reduce communication overhead.

In [4], Vehicular Networks (VNs) seek to provide, among other applications, safer driving conditions. To do so, vehicles need to periodically broadcast safety messages providing precise position information to nearby vehicles. However, this frequent messaging (e.g., every 100 to 300ms per car) greatly facilitates the tracking of vehicles, as it suffices to eavesdrop the wireless medium. As a result, the driver's privacy is at stake. In order to mitigate this threat, while complying with the safety requirements of VNs, we suggest the creation of mix-zones at appropriate places of the VN. We propose to do so with the use of cryptography, and study analytically how the combination of mix-zones into mix-networks brings forth location privacy in VNs. Finally, we show by simulations that the proposed mix system is effective in various scenarios. Our results show that, although the unlinkability of individual mix-zones can be relatively low in some cases, the accumulated unlinkability of the mix-networks is generally very high.

[5] Safety critical applications for recently proposed vehicle to vehicle ad-hoc networks (VANETs) rely on a beacon signal, which poses a threat to privacy since it could allow a vehicle to be tracked. Mix-zones, where vehicles encrypt their transmissions and then change their identifiers, have been proposed as a solution to this problem. In this work, we describe a formal analysis of mix-zones. We model a mix-zone and propose a formal definition of privacy for such a zone. We give a set of necessary conditions for any mix-zone protocol to preserve privacy. We analyze, using the tool ProVerif, a particular proposal for key distribution in mix-zones, and the CMIX protocol.

3. EXISTING METHOD

3.1 ALERT Routing Algorithm

ALERT uses dynamic Hierarchical Zone Partition. It dynamically partitions a network field into zones and randomly chooses nodes in zone as intermediate relay nodes. This intermediate relay node forms non traceable anonymous route. It uses the GPSR algorithm to send the data to the relay node.

As shown in Figure. 1, the given area is vertically partitioned into two zones X1 and X2. We then horizontally partition zone X1 to Y1 and Y2. After that, we vertically partition zone Y2 into two zones. This type of zone partitioning consecutively splits the smallest zone in an alternating vertical and horizontal manner. This partition process is known as hierarchical zone partition.

ALERT uses the hierarchical zone partition. In each step, it randomly chooses a node in the partitioned zone as an intermediate relay node which is called data forwarder, thus dynamically generating an unpredictable routing path for a message. The zone with k nodes where D exists is called as the destination zone which is denoted as Z_D. k is used to control the degree of anonymity protection for the destination.

In ALERT, each and every data source or forwarder executes the hierarchical zone partition. It first checks whether itself and destination are in the same zone. If both are in same zone, it divides the zone alternatively in the horizontal and vertical directions. This process is repeated until itself and Z_D are not in the same zone. It then randomly chooses a position in the other zone called temporary destination (TD), and uses the GPSR routing algorithm to send the data to the node closest to TD. This node is defined as a random forwarder (RF). ALERT aims at achieving k-anonymity for destination node D, where k is a predefined integer. Thus, in the last step, the data are broadcasted to k nodes in Z_D, providing k-anonymity to the destination.

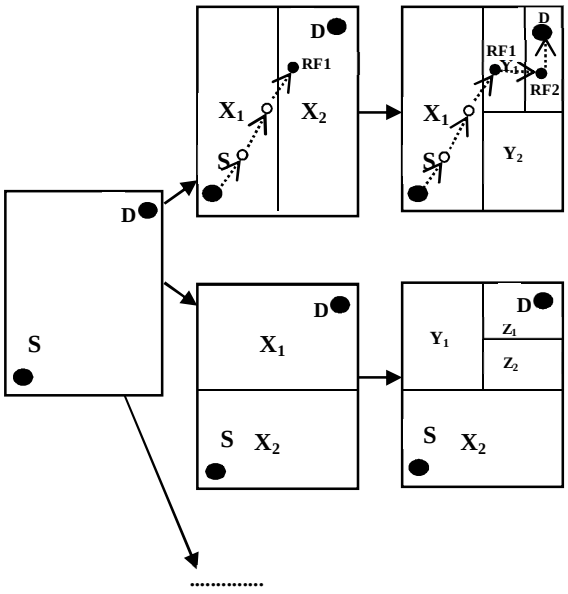


Figure. 1 Hierarchical Zone Partitions

3.2 Destination Zone Position

We use Z_D rather than D is to avoid exposure of D. Zone position refers to the upper left and bottom-right coordinates of a zone. There may be problem occur to find the position of Z_D, which is needed by each packet forwarder to check whether it is separated from the destination after a partition and whether it resides in Z_D. Let H denote the total number of partitions in order to produce Z_D. H is calculated by

$$H = log_2\left(\frac{\rho \cdot G}{k}\right)$$

where G is the size of the entire network area
k is the number of nodes in Z_D
ρ is the node density

3.3 Source Anonymity

ALERT aims to achieve the anonymity by restricting a node's view only to its neighbors and constructing the same initial and forwarded messages. This makes it difficult for an intruder to tell if a node is a source or a forwarding node. To strengthen the anonymity protection of the source nodes, we further propose a lightweight mechanism called "notify and go". The idea behind "notify and go" is a number of nodes send out packets at the same time as S in order to hide the source packet among many other packets.

"Notify and go" has two phases: "notify" and "go." In the first "notify" phase, S piggybacks its data transmission notification with periodical update packets to notify its neighbors that it will send out a packet. The packet includes two random back-off time periods, t and t0. In the "go" phase, S and its neighbors wait for a certain period of randomly chosen time €[t,t+t0] before sending out messages.

S's neighbors generate only several bytes of random data just in order to cover the traffic of the source. T should be a small value that does not affect the transmission latency. A long t0 may lead to a long transmission delay while a short t0 may result in interference due to many packets being sent out simultaneously. Thus, t0 should be long enough to minimize interference and balance out the delay between S and S's farthest neighbor in order to prevent any intruder from discriminating S.

3.4 Strategies against Attacks

ALERT has strategies to effectively counter intersection and timing attacks.

3.4.1 Timing Attacks

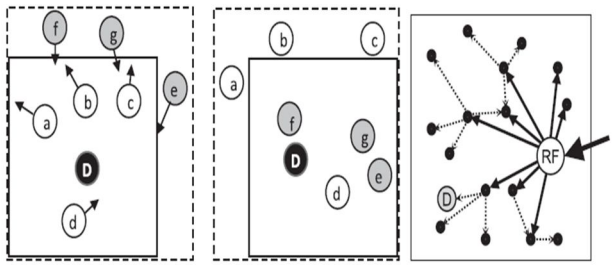
In timing attacks, an intruder can identify the packets transmitted between S and D through packet departure and arrival times. From this observation, the attacker can detect S and D. For example, two nodes A and B communicate with each other at an interval of 4 seconds. After a long observation time, the intruder finds that A's packet sending time and B's packet receiving time have a fixed five second difference such as (16:00:56, 16:01:00) and (12:08:33, 12:08:37). Then, the intruder would guess that A and B are communicating with each other. Avoiding the exhibition of interaction between communication nodes is a way to counter timing attacks.

In ALERT, the "notify and go" mechanism and the broadcasting in Z_D both put the interaction between S-D into two sets of nodes to obfuscate intruders. Also, the routing path between a given S-D and the communication delay (i.e., time stamp) change constantly. This again keeps an intruder from identifying the S and D.

3.4.2 Counter Intersection Attacks

In an intersection attack, an attacker may have the information about active users at a given time through repeated observations. The attacker with this information can determine the sources and destinations that communicate with each other. Though ALERT offers k-anonymity to D, an intersection attacker can still identify D from repeated observations of node movement and communication if D always stays in Z_D during a transmission session. This is because as long as D is conducting communication, the attacker can monitor the change of the members in the destination zone containing D. As time elapses and nodes move, all other members may move out of the destination zone except D. As a result, D is identified as the destination because it always appears in the destination zone.

Figure. 2. a). is the status of a Z_D after a packet is broadcasted to the zone. The arrows show the moving directions of nodes. We can see that nodes a, b, c, d, and D are in Z_D. Figure. 2. b) is the subsequent status of the zone the next time a packet is transmitted between the same S-D pair. This time, nodes d, e, f, g, and D are in Z_D. Since the intersection of the in-zone nodes in both figures includes d and D, D could be identified by the attacker. Therefore, the longer an attacker watches the process, the easier it is to identify the destination node.



a) b) c)

Figure. 2 Intersection Attack and Solution

To counter the intersection attack, ZAP dynamically enlarges the range of anonymous zones to broadcast the messages or minimizes communication session time. However, the former strategy increases the communication overhead, while the latter may not be suitable for long duration communication. Instead of adopting such a mitigating mechanism, we propose another strategy to resolve this problem. Note that the attacker can be puzzled and lose the cumulated observation by making it occasionally fail to observe D's reception of packets. Since packets are delivered to Z_D constantly in long-duration sessions rather than using direct local broadcasting in the zone, the last RF multicasts packet pkt1 to a partial set of nodes, say m nodes out of the total k nodes in the zone. The m nodes hold the packets until the arrival of the next packet pkt2. Upon the arrival of the next packet, the m nodes conduct one-hop broadcasting to enable other nodes in the zone to also receive the packet in order to hide D.

Figure. 2. c) shows the two-step process with the first step in solid arrows and the second step in dashed arrows. We can see that the first step reaches a number of nodes in the destination zone, but the destination is reached in the second step. Because the deliveries of pkt1 and pkt2 are mixed, an attacker observes that D is not in the recipient set of pkt1 though D receives pkt1 in the delivery time of pkt2. Therefore, the attacker would think that D is not the recipient of every packet in Z_D in the transmission session, thus foiling the intersection attack.

4. PROPOSED METHOD

Like other anonymity routing algorithms, ALERT is not completely bulletproof to all attacks. Also, ALERT cannot be applied to all network models. ALERT can be applied to Random Way Point model and Group Mobility Model. To

prevent the occurrence of stronger and active attackers, we propose a Secure Cryptographic Based Mix-Zones Routing Protocol (SCMIX).

An unobserved zone functions as a mix zone where the mobile nodes change pseudonym and mix with each other. Note that the Mobile nodes do not know where the mix zone is (this depends on where the adversary installs observation spots).

We propose to create mix-zones at predetermined locations and to force pseudonym changes to take place within those regions. Since the location of mix-zone is fixed, the adversary can identify them and thus could easily attempt to eavesdrop transmissions originating in the mix-zone area. The adversary observes the timing and the location of the entering and exiting node in order to derive a probability distribution over the possible mappings. To solve this problem the timing of events depends on the delay characteristics of the intersection structure. Likewise, the location of entering and exiting nodes depends on their trajectory in an intersection.

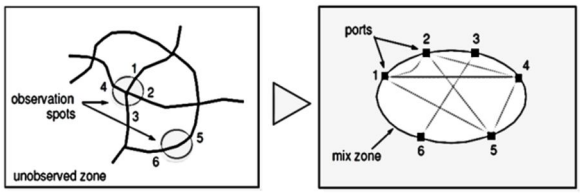


Figure. 3 Mix-Zone Concepts

The idea for mix-zones is to prevent the adversary from accessing the content of messages, including the Node's signatures. All legitimate nodes within the mix-zone obtain a symmetric key and utilize this key to encrypt all their messages while within the zone. We improve location privacy of mix-zones via extensions to the SCMIX protocol. SCMIX Protocol distributes Keys for encrypting beacon messages while in the mix-zone.

While the mobile node is inside of the cryptographic mix zone, all communication is encrypted and therefore an adversary cannot read-out useful information (including meta-information) from its messages. Nodes in the mix-zone forward the symmetric key to Mobile nodes that are in direct transmission range outside of the mix zone such that these nodes are also able to decrypt messages. Mobile Nodes then change pseudonyms while being inside of the mix-zone.

5. CONCLUSION

The concept of mix zone refers to a service restricted area where mobile users can change their pseudonyms so that the mapping between their old pseudonyms and new pseudonyms are not revealed. Since the location of mix-zones is fixed, the adversary can identify them and thus could easily attempt to eavesdrop transmissions originating in the mix-zone area. To solve this problem, we propose a Secure Cryptographic Based Mix-Zones Routing Protocol (SCMIX). In our future work, we intend to study how the frequency of the pseudonym change influences the level of privacy achieved**.**

6. REFERENCE

[1] Zhi Z. and Choong Y. K. (2005), 'Anonymizing Geographic Ad Hoc Routing for Preserving Location Privacy', Proc. Third Int'l Workshop Mobile Distributed Computing (ICDCSW).

[2] Defrawy K. E. and Tsudik G. (2007), 'ALARM: Anonymous Location- Aided Routing in Suspicious MANETs', Proc. IEEE Int'l Conf. Network Protocols (ICNP).

[3] Wu X., Liu J., Hong X. and Bertino E. (2008), 'Anonymous Geo- Forwarding in MANETs through Location Cloaking', IEEE Trans. Parallel and Distributed Systems, Vol. 19, No. 10, pp. 1297-1309.

[4] Freudiger J., Raya M., Felegyhazi M., Papadimitratos P. and Hubaux J. (2007), 'Mix-zones for location privacy in vehicular networks', Proc. First Int'l Workshop onWireless Networking for IntelligentTransportation Systems (WiN-ITS'07).

[5] Dahl M., Delaune S. and Steel G. (2010), 'Formal analysis of privacy for vehicular mix-zones', Proc. of the 15th European conference onResearch in computer security (ESORICS'10), pp. 55–70.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 752 - 758, 2013, ISSN: 2319–8656

Alert Routing In Wireless Video Sensor Networks for Monitoring Applications

Youssef Benabbassi
Bechar University,
Bechar, Algeria

Hafid Haffaf
Oran University,
Oran, Algeria

Congduc Pham
Pau University,
Pau, France

Abstract: Wireless sensor networks are appropriate monitoring tools used in surveillance applications. Organizations all over the world are realizing the effectiveness of protecting people, places and things with advanced video surveillance systems to increase safety. Wireless infrastructure allows them to deploy and extend video surveillance capability in virtually any indoor or outdoor environment. The network lifetime is directly related to the energy resources of the sensor nodes and can be extended by energy-aware protocols. The LIUPPA, France Laboratory has proposed a wireless video sensors model dedicated to intrusion detection surveillance application. Our aim is to extend this work by new algorithms that optimize the alert propagation message diffusion.

The main idea of our algorithm is to reduce the flooding of the message alert by propagating them only for the list of the sensors located in the field of view (FOV) of the sensor alerted. This reduces considerably the effects of implosion. In this paper, we describe the WVSN model proposed by the LIUPPA. Moreover, we explained and discussed the result of our algorithm under Omnetpp/Castalia Simulator.

Keywords: Wireless Sensor Networks; Routing; Surveillance Application, Wireless Video Sensor Networks; Energy Consumption.

1. INTRODUCTION

In the past, the wired sensor networks were used in the factory control and automation system. An exponential growth in wireless communication techniques has made it possible for wireless control and automation system [1].

Rapid advances in Micro-Electro-Mechanical Systems (MEMS) technology have made possible the existence of tiny sensor nodes equipped with sensing, communication and processing capabilities. When sprayed in an area, they are capable of forming a multi-hop wireless network known as Wireless Sensor Network (WSN)[2]. The availability of low-cost hardware is enabling the development of wireless multimedia sensor networks (WMSNs), i.e., networks of resource-constrained wireless devices that can retrieve multimedia content such as video and audio streams, still images, and scalar sensor data from the environment [3]. Wireless Video Sensor Networks (WVSNs) are a category of WSNs in which sensor nodes are equipped with a digital camera.

Wireless sensor networks are appropriate tools to monitor an area for surveillance. Video surveillance is being implemented on an unprecedented scale. Public safety agencies, municipalities, maritime ports, transit authorities and other government agencies depend on video surveillance to protect their constituents and property. For law enforcement, video surveillance is essential to deter crime, improve incident response time and provide forensic data. Video surveillance is also essential for organizations that have infrastructures spread over large areas, such as public transportation, university and corporate campuses, retail, construction, logistics, mining, oil and gas [4].

WVSNs suffer from the same problems as WSNs such as application dependency, energy constraints, high sensor density, limited bandwidth and dynamic topology. Terrance Boult and al. in [5] reviews significant issues, design constraints and accomplishments of work on multi-camera sensor networks combining multiple omni-directional imaging sensors, traditional stationary cameras and pan-tilt sensors. To analyze the surveillance performance of the network, the author in [6] proposes deployment quality measures and a trade-off between the number of sensors and the deployment quality have been discussed.

The performance of a surveillance wireless sensor network is generally measured with its detection capability within a monitored zone. Can Komar and al. provide a tool to the network to derive the expected detection performance and risk analysis framework for a given sensor network with realistic border surveillance scenario parameters [7]. Antonio-Javier and al. introduces that a video-surveillance is a solution to detect and identify intruders as well as to better take care of a process [8]. They propose a new platform called integrated wireless sensor network solution for precision agriculture. Congduc Pham and al. from the LIUPPA[19], France Laboratory; propose a multiple-level activity model that uses behavior functions to define application classes and allows for adaptive scheduling based on the criticality application of multiple cover sets per sensor node [9]. Our aim is to extend this application with an algorithm proposed to manage the alert propagated message diffusion. The main idea of our algorithm is to reduce the flooding of the message alert in order to reduce energy consuming.

The paper is organized as follows. Section 2 presents the related works in wireless routing protocols. Section 3 presents

the WVSN model developed at the LIUPPA France Laboratory which is the main stone of our approach. Section 4 presents our algorithm proposed to manage the diffusion of the message alert through the network taking to reduce considerably the effects of implosion. Section 5 presents the simulation results under the Omnetpp/Castalia Simulator. Finally, section 6 concludes the paper and points out open research problems.

2. ROUTING PROTOCOLS IN WSN

Routing protocols in WSN are topical issue in this domain and are covered by a rich literature researchs. Kemal Akaya and al. in [10] surveys recent routing protocols for sensor networks and presents a classification for various approaches pursued also each routing protocol is described and discussed. Ian F.Akyilidiz in [11] surveyed the state of the art in algorithms, protocols, and hardware for wireless multimedia sensor networks and discussed in detail the open research issues. Adamu M.Z. and al. in [12] re-simulate different routing protocols using a Matlab based simulator and give simulation results and performance metrics to serve as a benchmark for future comparison for the research community. Chee-Yee Chong and al. presents in [13] some recent research results in sensor network algorithms, including localized algorithms and directed diffusion, distributed tracking in wireless ad hoc networks, and distributed classification using local agents. K.Beydoun treated in [14] the problem of routing in wireless sensor networks and proposed a design of a hierarchical routing protocol for sensor networks based on network zones partitioning.

Routing allows information transport from source to destination through a network connection. The routing problem is to determine optimal path of the packets through the network with regard to some performance criteria such as energy consumption. The goal is to find the low investment cost which routes the nominal traffic and guarantees quality of service (QoS) [11].

In general, routing in sensor networks can be classified, according to structure of the network, in flat routing, and Hierarchical routing. In the flat routing, all nodes typically have the same roles and features. However, hierarchical routing is performed at several levels in the sense that the view of the network is reduced. Some nodes may play specific roles in the network to route information.

Following the method of creation and maintenance of roads in routing packets, routing protocols in wireless sensor networks can be classified into three categories: proactive, reactive and hybrid [12].

The roads in the proactive routing are calculated in advance. Each node maintains multiple routing tables by exchanging control packets between neighbors. The need to maintain and check the validity of tables routing permanently (including further information which will probably not used) is the main drawback of proactive protocols. By cons, they have the important advantage requires no delay before transmitting a packet since the road is already known. OLSR and FSR are two examples of proactive protocols.

Unlike proactive protocols, reactive protocols will calculate the route on request. However, the demand routing leads to a slow global process because of the search for roads. This type of protocol has the disadvantage of being very expensive in packet transmission in the determination of routes but has the advantage of not having to maintain information unused routing tables. AODV and DSR are two examples of reactive protocols.

The hybrid routing protocols or "mixed" combine two previous types of methods (proactive and reactive). The proactive protocol is applied in a limited scope around the source (limited number of neighbors), while the reactive protocol is applied beyond the scope (distant neighbors). This combination is achieved in order to exploit the advantages of each method and bypass their limitations. ZRP and CBRP are two examples of protocols hybrids [14].

In fact, numerous research investigations on the performance and benefit of routing protocols can improve throughput, increase reliability, reduce end to end delay, and mitigate network congestion.

3. WVSN MODEL

A wireless video sensors network consists of a set of autonomous nodes with a small onboard camera. In sensor networks scalar systems considering capacity omni-directional capture, two nodes are regarded as redundant if they are close to each other. In video sensor networks, cameras have a field of view and optionally, zoom capabilities. In this case two nodes can be redundant even if they are relatively distant from each other. Sometimes multiple views are desirable to resolve ambiguities; in other situations, remote nodes can provide more usable information depending on weather conditions, for instance.

At the LIUPPA laboratory P. Congduc & al develop a wireless video sensor networks for surveillance applications [15]. This model also addresses the management of energy, since the scarcity of this resource has a strong impact on coverage as it is not realistic that all nodes can be active simultaneously. The model proposes a based coverage approach for scheduling adaptively video nodes activity, taking into account energy considerations and objectives of the application. The proposed model is based on approach using several sets covering in order to manage redundancy fields of view of the cameras. It is based on a distributed algorithm that helps each node to organize its neighbors into non-disjoint sets, each of which being a cover set that overlaps its field of view. Then, based on the activity of its neighbors, a node decides to be active or not, without compromising its coverage area. To take into account the energy and the objectives of the application, the approach provides a multi-level model based on Bezier curves activity [15]. This approach defines two classes of application: high and low risk applications. In the first class, the application

does not need high frame capture rate, but in the second class, the majority of nodes capture at high frame capture speed to ensure better detection.

The surveillance zone of a video node *v* is usually represented by its field of vision. Video nodes are random positions and random directions. In [15,16] the authors considered a 2D model that defines the field of view by the 4-tuple *v*(P, R_s, V, α), where P represents the position, R_s the coverage radius, V the direction vector of the camera, and α the semi angle of view of the sensor video as illustrated in Fig 1

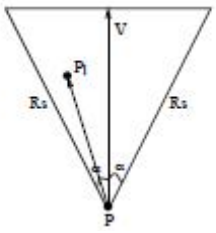


Fig 1: Video Sensing Model

First, we suppose that all nodes are identical, i.e they have the same coverage radius R_s and the same viewing angle. The approach is fully distributed where each node must provide its own coverage independently. For the sake of simplicity, a node *v* generally covers a triangular area with its field of vision, either by itself if it is being in an active state or by redundant nodes when it is in an inactive state. This approach allows nodes to find among their neighbors which of them cover their field of vision as illustrated in Fig 2.

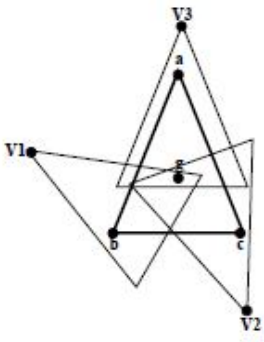


Fig 2: FoV coverage

It begins to calculate the cover set according to their cardinality, giving priority to those that contain fewer elements. In the case of two or more sets have the same cardinality; they will be classified according to their energy level. A video node *v* receives messages of activity of its neighbors and checks if any of its cover set is satisfied or not. If a set covering *v* is active, the node *v* is asleep and sends its decision to its neighbors. On the other hand, if no set covering *v* is satisfied that node *v* decides to stay active and also broadcasts its position to its neighbors.

In video sensor networks, autonomy of a sensor is directly related to the speed of image capture. The rate of capture can be synonymous with quality of monitoring and it depends on the criticality of the application. An application designed for intrusion detection should have a fairly high speed capture to avoid missing intrusion. In other applications, the capture rate can be greatly reduced. Thus, it is important to adjust the capture rate to allow a fair compromise between autonomy, criticality and coverage. The model proposes two approaches to regulating the speed of image capture. The first approach is called static control. This is a naive approach; it is to set a constant speed of capture throughout the life of the sensor according to the criticality of the application. Its major drawback is the rapid depletion of the network. The second approach is called dynamic approach which optimizes use of energy in the network. It varies the speed of capture of a node *v* as a function of the number of sensors they cover. In other words, the larger the area of a sensor is covered, the more it can afford to quickly capture.

4. ALERT ROUTING ALGORITHM

Our contribution is to extend the previous wireless video sensor networks model defined in section 3 by managing the alert diffusion. In this model, when a node detects an intrusion, it floods automatically the entire network by bypassing the message alerts to the whole network. This technique doesn't preserve energy and could create implosion effect. This effect is caused by duplicated messages sent to the same nodes. When an intrusion occurs, two nodes sensing the same region send alert packets to the same neighbors. The main idea of our algorithm is to reduce the flooding of the message alert by propagating them only for the list of the sensors located in the field of view (FoV) of the sensor alerted in such a way that only nodes able to retransmit the message in the wright direction are considered.

Figure 3 shows the operation of the noSelectiveFoV strategy. In this strategy when a node v detects an intrusion it spreads a message of warning to all its neighbors. Thereafter, any node receiving this message will carry an alert dissemination. Its major inconvenient is the risk of duplication of the same warning message.

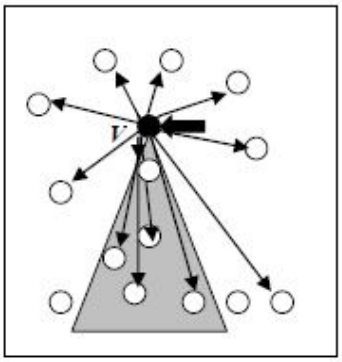


Fig 3: noSelectiveFov Alert propagation

What is more, it is disadvantageous in terms of energy consumption and obstructs the traffic in the network. In the contrary, in selectiveFoV technique, as shown in Figure 4, once the node v detects an intrusion it spreads a warning message to the nodes that are within its field of view. Thus, it greatly reduces the effect of this flood alert that arrives more quickly and then, increases the lifetime of the network.

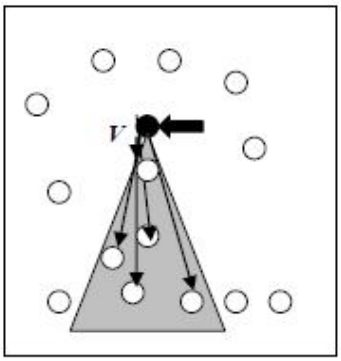


Fig 4: SelectiveFov Alert propagation

By varying the angle of view of the camera, a virtual field of vision is obtained as shown in Figure 5. The VirtualFoV strategy is based on this principle to expand the number of nodes to be alerted in the case of intrusion detection.

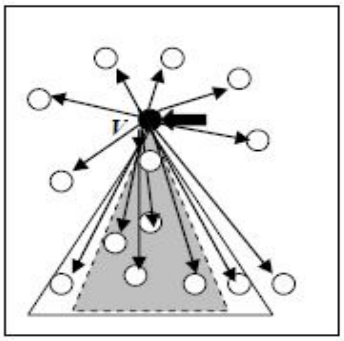


Fig 5: VirtualFov Alert propagation

So, our algorithm improves the phenomenon of diffusion in the network, and will obviously reduce energy consuming. When a node *v* detects an intrusion it looks in its table of neighbors the list of nodes *v`* that are in its field of view. If this node *v`* is in active mode will be alerted and if it was already alerted it switches reinforcement. We shall give now, the proposed algorithms.

Algorithm selectiveFOV:

Input: intrusion detected

Output: alert packet

Begin

List := ""

If node *V* detects intrusion **Then**

Begin

index := id_Sender

V := neighbors[index].p

For i:=0 to nbNeighbors **Do**

Begin

If (neighbors[i].p is_inside(*V*.fov) **Then**

Begin

If (neighbors[i].isActive=True) **Then**

Begin

Increase AlertCriticalitylevel

SetAlertCriticalityLevel

list := list + "#" + neighbors[i].id

End

Else

Begin

neighbors[i].status := Active

Endif

Endif

Add list to Alert packet

Send Alert packet to RoutingModule

Enddo

Endif

list_id := Packet.list

For each node *V'* receives Alert Packet **Do**

Begin

If *V'*.id is in list_id **Then**

Begin

If (*V'*.Alerted==False) **Then**

Alerted := True

Else

Begin

Increase AlertCriticalitylevel

SetAlertCriticalityLevel

Endif

Send Alert to neighbors

Endif

Enddo

End

"noSelectiveFOV" refers to the technique used by LIUPPA [19]. It is based on the dissemination of alert after intrusion detection. It is applied for all neighboring nodes, and may causes flooding of the network.

"selectiveFOV" method is our new technique which is based on sending the alert to the subset of neighboring nodes that are in the FOV of the node that detected the intrusion. This technique significantly reduces the propagating phenomena of sending alert and it will have a significant impact on energy consumption in the network.

"virtualFOV" is another variant of selectiveFOV where we vary the angle of view to obtain a virtual FOV.

We propose an algorithm to integrate and manage the three propagation techniques alerts in the previous model. If the severity level is at a low level implementation, we will propagate alerts based on selectiveFoV algorithm to neighbors which are in the field of view of the node that detected the intrusion. In the presence of intrusions case, there will be reinforcements alert based on virtualFoV algorithm. It will expand its field of vision by increasing the angle of view of its sensors. The alert level here is considered a medium level. It will subsequently alert a larger number of neighbors in the virtual field of vision of the node that detected the alarm. But once the alert goes to a higher level, the model uses the spread alert based on a flood represented here by the algorithm noSelectiveFoV

Algorithm scheduling Alert:

Begin

Level := Current_CriticalityLevel;

If Level = "Low" **Then**

Uses SelectiveFov algorithm for routing alert

Else

if Level = "Medium" **Then**

Uses VirtualFov algorithm for routing alert

Else

Uses NoselectiveFov algorithm for routing alert

Endif

Endif

End.

5. SIMULATION AND RESULTS

To evaluate the performance of our alert routing techniques we concluded a series of simulations based on the discrete event simulators Omnetpp/Castalia (http:/www.omnetpp.org) [17] and (http:/Castalia.research.nicta.com.au) [18] under the LIUPPA wireless video sensors network model.

The LIUPPA simulation model represents a randomly deployed wireless video sensor network in an 75mx75m field. Each sensor node is characterized by: its position (x,y), a depth of view for the camera, a line of sight for the camera and an angle of view (AoV). The sensor`s field of view (FoV) is then represented by a triangle as shown in fig 1. Depending on the number of nodes deployed in the field, the simulation will determine the cover sets for each sensor node and will compute a percentage of coverage for each cover set. Then each sensor will decide to be active or not, and will decrease its energy level according to its frame capture rate. The frame capture rate for each sensor is determined by the size of its cover sets. Also depending on the criticality level of the application, the capture rate for a given number of cover sets is varied when the initial cover set computation ends, energy node will decide to be active or not and the simulation model will determine which cover sets are active. The simulation ends when all sensor nodes have determined their cover sets and have computed for each cover set the percentage of coverage of the cover set [16]. This section aims to discuss the results of the simulation of our algorithm. Simulation was realized using Castalia environment consisting of 300 nodes. Application layer represents the videoSensorModel, Bypass Castalia in the routing layer and CSMA as Mac Layer.

These simulations are run ten times randomly generated by Castalia. The different figures below summarize our results. Fig.6 gives an idea of the number of the sent alert packets issued by each technique. We note that the technique selectiveFOV has significantly reduced the number of these packets.

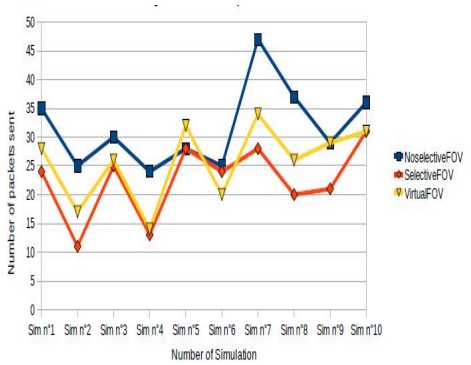


Fig 6. Number of Packet Sent

With selectiveFOV technique the number of received packets is reduced compared to the noSelectiveFOV as shown in fig 7. If we take the results of the simulation "sim n° 8", we find that the number of received packets is equal to 150 packets with the selectiveFOV technique against 350 packets with the noSelectiveFOV.

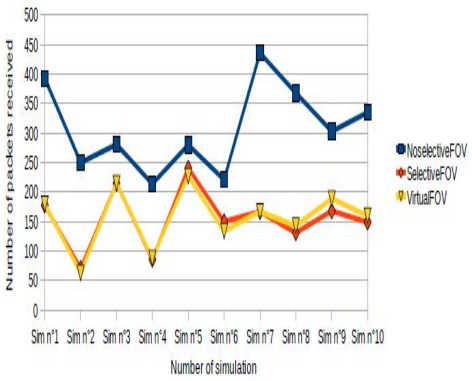


Fig 7. Number of Packet received

Fig 8 summarizes the number of detected intrusions in the network. In this case, the behavior of the three methods converges to the same values in the most phases of simulation.

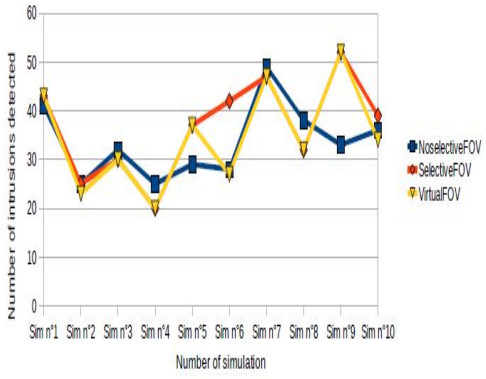


Fig 8. Number of intrusion detected

Fig 9 summarizes the difference in the energy consumption. The selectiveFOV technique gave a good preservation of the energy, and this for the various tests on the basis of the progressive variation of the running time.

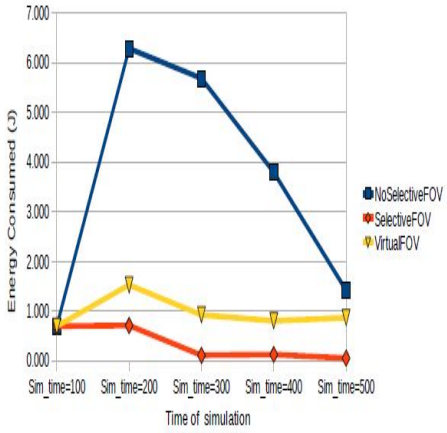


Fig 9. Energy Consumed

The simulation shows that our algorithm has given better results and then it allows good energy preservation.

6. CONCLUSION

This paper addresses the problem of routing alert for wireless video sensor networks deployed for critical surveillance applications. Increasing the network lifetime is of prime importance for mission-critical video sensor networks. Based on a coverage model that handles FoV concepts, we proposed a routing algorithm to manage alert propagation in the network based on a selective field of view to improve forward alerts propagation. It reduces in one hand, considerably the effects of implosion, and the number of alert packets in the network, and It takes in the other hand, into account the risk level in intrusion detection systems. Simulation results show the performance of our strategies in terms of percentage of send packets, received packets, intrusion detection and energy preservation. Future works will investigate in more details how to secure alert packets propagation in the network.

As well as security problem, the effect of the sensor failures on the Quality of Service (QoS) is an open research topic. Since sensor failures are common, fault tolerance of the network should be investigated because loss of individual sensor nodes or a group of sensor nodes should not hamper the task accomplishment of the network.

7. REFERENCES

[1] Pratikkumar,D. 2009. Localisation and Surveillance using Wireless Sensor Network and Pan/Tilt Camera. University Dayton, Ohio.

[2] Saaed, U. 2010. Energy and Distorsion Analysis of Video Compression Schemes in Multi-Hop Wireless Video Sensor Networks. University of Sciences and Technology Istanbul, Pakistan.

[3] Ian, F. and Kaushik, R. 2008. Wireless Multimedia Sensor Networks. Proceeding of the IEEE Vol.96, N°10.

[4] Aruba networks. 2010. Improving Coverage and Quality in Video Surveillance Applications.

[5] Terrance, E., Johnson, R., and Tracy, P. 2006. A decade of networked intelligent video surveillance. ACM Workshop on distributed camera network.

[6] Ertan, O. 1997. Deployment Quality Measures in Surveillance Wireless Sensor Networks. Thesis, Computer Engineering, Ege University.

[7] Can, C., Mehmet, Y., and Cem, E. 2012. Detection Quality of Border Surveillance Wireless Sensor Networks in The Existence of Trespassers favorite Paths. Computer Communication Journal 35 1185-1199.

[8] Antonio-Javier, G., Felipe, G., and Joan, G. 2011. Wireless Network deployment for Integrating Video Surveillance and Data Monitoring in Precision Agriculture over Distributed Corps. Computers and electronics in Agriculture Journal 75 288-303.

[9] Congduc, P., Adellah, M., and Rachid, S. 2011. Risk-based Adaptive Scheduling in Randomly Deployed Video Sensor Networks for Critical Surveillance Applications. Journal of Network and Computer.

[10] Akaya, K., Mohamed, Y. 2005. A Survey on Routing Protocols for Wireless Sensor Networks. Ad Hoc Networks 3 325-349.

[11] Ian, F.A., Tommaso, M., and Kaushik, R.C. 2007. A survey on Wireless Multimedia Sensor Networks. Computer Networks 51 921-960.

[12] Adamu, M.Z., Li-Minn, A., and Kah, P.S. 2012. Classical and Swarm Intelligence Based Routing Protocols for Wireless Sensor Networks: A survey and Comparison. Journal of Networks and Computer applications 35 1508-1536.

[13] Chee-Yee, C., and Srikanta, P.K. 2003. Sensor Networks: Evolutions, Opportunities, and Challenges. Proceeding of the IEEE, Vol.91, N°.8.

[14] Beydoun, K. 2009. Designing a Hierarchical Routing Protocol for Wireless Sensor Networks. PhD, University of Franche-Compte, France.

[15] Congduc, P., Abdellah, M., and Rachid, S. 2009. Scheduling Visual Sensor Nodes in Critical Surveillance Application. University of Pau, France.

[16] Congduc, P., and Abdellah, M. 2010. Performance Study of Multiple Cover-set Strategies for Mission-critical video Surveillance with wireless video Sensors. IEEE WIMBO.

[17] Omnetpp Network Simulation Framework. 2013. http//www.omnetpp.org.

[18] Castalia-NICTA. 2013. http://web.univ-pau.fr/cpham. 2013.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 759 - 762, 2013, ISSN: 2319–8656

A Hybrid approach for enhancing the capability of Spam Filter

Amandeep Kaur
Punjab Institute of Technology
Kapurthala, Punjab,
India

Malti Sarangal
Punjab Institute of Technology
Kapurthala, Punjab,
India

Abstract: In this paper, A hybrid approach for enhancing the capability of spam filter is proposed. This hybrid approach has defined the best features like Whitelisting, Blacklisting, Greylisting, Bayesian filtering, eXpurgate technology, Social closeness that are used to make sure that all fetched mails are checked against spam or not and redirected to user inbox or spam folder respectively**.**

Keywords: UBE; UCE; Bayesian; Blacklist; Whitelist; Greylist; eXpurgate technology.

1. INTRODUCTION

Electronic mail (email) has become a promising component for Internet users. The most common problem observed in maintaining email inbox is the incoming spam mails [7]. UCE or UBCE (Unsolicited Bulk Commercial Email) is the fastest and cheapest method of advertising the commercial websites [3]. UCE stands for Unsolicited Commercial E-mail where the word unsolicited means something unwanted, not requested or invited. These Spam mails are forcefully sent to many users at once. These mails fill the user inbox with thousands of unwanted mails that create the difficulty for user to differentiate between legitimate mails and spam mails. Many copies of the same mail are sent to many users at once. Spammers are the group of people who willingly spread spam in order to advertise their products, commercial websites across the internet [7]. Because sending spam mails through internet is the cheapest and fastest way of advertisement. The most common method of blocking the spam mails and let the only legitimate means useful traffic to pass through is called SPAM Filter. As the spammers enhanced their techniques the spam filters enhanced too to support the needs of user.

Adaptation to these new spam techniques is not observed in the most of spam filters available today. Also these filters do not deal with image spam and currently image spam is become a serious problem that spammers are using now a days. No doubt the commercial available email providers are using unique featured, efficient spam filters but they are not available online. So an individual user cannot use these filters for his machine. Some filters with excellent features come at higher price that a normal user can not afford [7]. So, this hybrid approach defines all unique features that will be available online at free of cost. It is combination of all the excellent features of the already available filters and also will remove the limitations of those filters. It has the ability to learn and adapt from the user's choices and establish a "Blacklist", "Whitelist" and "Greylist" of the messages and SPAM. It will be more effective and accurate in blocking the unwanted messages.

2. PROBLEMS CAUSED BY SPAM

2.1 Cost

Unwanted Spam mails costs a lot to the email providers. Spammers' costs are almost always borne by end users, because spammers often steal hardware and network resources. Spammers use networks of hijacked computers (botnets) as email clients [10]. Spam mails also wastes the network bandwidth by increasing the traffic over internet [2].

2.2 Time Constraint

- It wastes the precious time of the organizations as organizations spend a lot of time in identifying whether the incoming mail is useful for their organization or it is unwanted mail i.e. Spam mail before passing it into spam organization's email inbox. All this time is wasted, costs the company a lot more than the amount spent in initially sending out the spam mail by the spammers [7].
- It wastes the time of user also because he has to spend time in differentiating between the legitimate mails and useful mails. Employee time spent on checking, interacting and removing SPAM emails [2].
- Network administrator's time required to spend dealing with SPAM (scanning, cleaning) and/or associated problems on viruses and malicious applications [2].

2.3 Malware spreading/ Phishing

There are lot of problems that Spam mails create when spread across the internet. Many of the spam carry website links, that on clicking redirect to foreign sites that are harmful to the user's computer. It redirects the users control to phishing sites. Confidential (Personal) information of user is requested through the 'data fields' of such sites using which spammers obtain important personal information such as credit card information of the users [7].

2.4 Blank spam mails

Spam mails can be blank mails with blank body as well as no information even in the subject of the mail. Sender information is also made unavailable to the end recipient. Blank emails are sent by the spammers which enables them to differentiate between valid email address and invalid address under an email provider. Invalid addresses mails bounce back thus providing spammers with only valid email addresses to further send spam mails. Blank mails sometimes also spread malware which can harm the data in the user's computer. Trojans in the form of attachments are sent [7].

2.5 Forwarded mails

These forwarded mails are another problem causes spam. In some spam mails, spammers initialize a spam mail and send it to few users, in order to stop receiving similar spam mails further, the user is forced to forward the mail to some others in the mailing list. Hence, even if half the users forward the mail, the amount of spam created is immense and would require lot of cost to be removed off the internet [7].

2.6 Garbage/ Not legal data

Most of the spam mails prevalent are useless mails consisting of nothing that is meaningless to the user. Spam mails usually contain information about schemes and products that are not of much use to the individuals. Fraudulent schemes, solutions for situations, free advice, links to phishing websites etc. are sent through spam mails that are only contain the garbage material. Illegal content also spread across the Internet via spam mails. In certain countries, laws are implemented against display or spread of certain content. Spammers, against those laws, try to spread out content that is considered illegal through spam mails [7].

3. LITERATURE REVIEW

A solution proposed way back when Internet came into existence was to implement spam filters to avoid spam from filling email inbox to the brim. A SPAM filter is a set of instructions for determining the status of the received email. SPAM filters are used to prevent SPAM email passing through to the recipient. The challenge is how to design an effective SPAM filter that allows desired email mail to pass through while blocking the unwanted SPAM emails. The potential unwanted problem is that often a SPAM filter may identify a legitimate email as a SPAM, and block it (false positive), or identify SPAM email as legitimate email, and allow it to pass through (false negative). Of these two cases, implications on the false positive can be very serious as important legitimate emails may not reach the receiver. A means to quantify the effectiveness of a SPAM filter can be based on the percentage of SPAM emails being blocked, whitelist allowing legitimate emails to pass through to the recipients and blocks the mails that come from unknown sender. Listed below are three commonly used methods in SPAM filtering [2].

3.1 Blacklist Filter

Black list is effectively a list of emails that is not allowed to pass through. This can be based on the assumption that the email could contain a common word or phrase in the header, an IP address, or domain name. The use of a black list SPAM filter in isolation can result in false positive error. Assuming the word "results" is a keyword in the list, the following example will block both emails. If the email header is (your exam results), another email has (use our product for quick results), what is going to happen is the filter will block both emails. (False positive) [4].

3.2 Whitelist Filter

In this case, all the emails are treated as SPAM except the ones in the white list database. The database is built using a confirmation process by the recipient. The problem with this time consuming technique is that it causes unnecessary burden to the users [5].

3.3 Bayesian Filter (Content Focus)

This approach is an extension of text classification technology, which searches the textual content of an email and uses algorithms to identify SPAM email. The algorithms are able to classify the occurrence of certain words and phrases in terms of how and where they appear in the email. The challenge with content filtering is that SPAM emails sometimes contain images, which are difficult to interpret their contents [6].

4. EXISTING WORK

Every Incoming mail is parsed through these three filters step by step after that if mail is identified as Spam then passed to Spam folder or is allowed to enter into user's inbox.

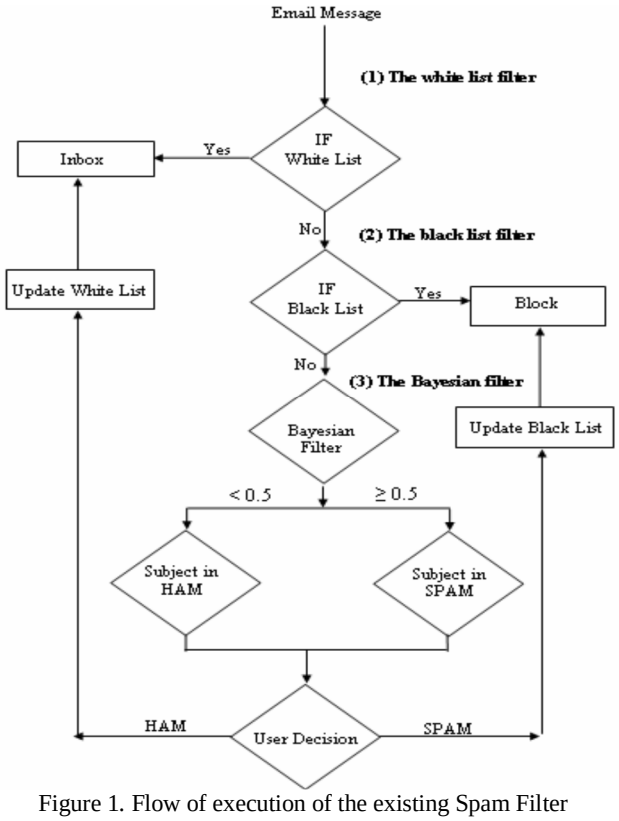


Figure 1. Flow of execution of the existing Spam Filter

Table 1: Comparison between different methods of SPAM filters

Spam Filter Methods	Block known SPAM	Block unknown emails	Self learning	Easy to use
Blacklist	✓			✓
Whitelist		✓		✓
Bayesian			✓	✓
Fingerprint	✓		✓	
Password		✓		✓
Challenge/ Response		✓		
Community Base	✓			✓
Mobile agent	✓			
Encryption and trust		✓		✓
Copyright tokens		✓		

This is the comparison of various filters methods that are used in detecting the Spam mails and and allows only legitimate mails to pass through into the user's inbox. Blacklist filter method is very efficient in blocking the known Spam mails and it is very easy to use. Whitelist filter is very much similar to blacklist filter but it blocks the unknown emails. Bayesian filter is a self learning filter as it automatically learns from new spam mail techniques. Fingerprint filter assign a fingerprint (distinct identifier) for spam message. It constructs the database for SPAM mails and prevent them from passing through. In Password filters passwords are required to be in the email to pass through the filter. But it blocks the new legitimate emails that does not have password yet. Challenge/Response blocks unapproved mail until response arrives and allows only legitimate senders to pass through after their response. But it blocks new legitimate mails and also annoy legitimate senders by asking for response with each message. Community Base filter method blocks mail based on community agreement means blocks a SPAM that a group decides to block but it does not block a new SPAM and also a one major drawback of this filter is that conflict may arise between the users while taking the decision about a particular mail is Spam or not. Encrytion and Trust Send mail with digital signature. Digital signature is very hard to fake and also used to sign and encrypt message that is sent out thus provide the security. But this technique is too complicated for the users and also cost and time wastage for small group of users. Mobile agent is a filter that works on remote system to perform the filtering on email server [2].

Table 2: Comparison between different approaches

Approaches	**Pers-onali-zed**	**Attack-resilient**		**User friendly**
		Imper-sonation	**Poison**	
Content based spam filters				
Static keyword	No	Yes	No	No
Machine-learning	No	Yes	No	No
			No	Yes
Collaborative	No	Yes		
Identity-based spam filters				
Black/white list	No	No	Yes	No
Social-interaction-based	No	No	Yes	Yes
Reputation	No	No	Yes	No
Social network aided content and identity based spam filter				
SOAP	Yes	Yes	Yes	Yes

5. PROPOSED WORK

No perfect SPAM filter has been found so far the following proposed approach [8] is aimed to enhance the capability of spam filter that can block SPAM emails and let legitimate emails to pass through using a combination of techniques including the use of the above approaches. It has defined the best features like Whitelisting , Blacklisting , Greylisting , Bayesian filtering, eXpurgate technology, Social closeness that are used to make sure that all fetched mails are checked against spam or not and redirected to user inbox or spam folder respectively. First of all user login with his details, user's credentials are checked if user is authorized, then before entering into user's inbox some techniques are used to check whether incoming mail is SPAM or not. If the incoming mail is identified as SPAM then it is redirected to SPAM folder otherwise incoming mail is allowed to enter into user's inbox. Firstly, eXpurgate technology is used for SPAM detection. With 2 step checking and adding extra header to incoming mail. The scope of false positive occurrence is decreased with help of eXpurgate technology. Then whitelist filter checks the incoming email against the white list. If the email address is found in the white list, then the filter will allow the message to pass through to the INBOX. If the sender mail id is not present in the white list, and if the mail sent by the sender is identified as spam then the mail id is added to another list called the Greylist. If another spam mail is sent by the same sender for the second time, the sender is then added into the blacklist thus blocking any further incoming mails into the inbox. Blacklist filter checks against the black list and blocks the known spam mails.

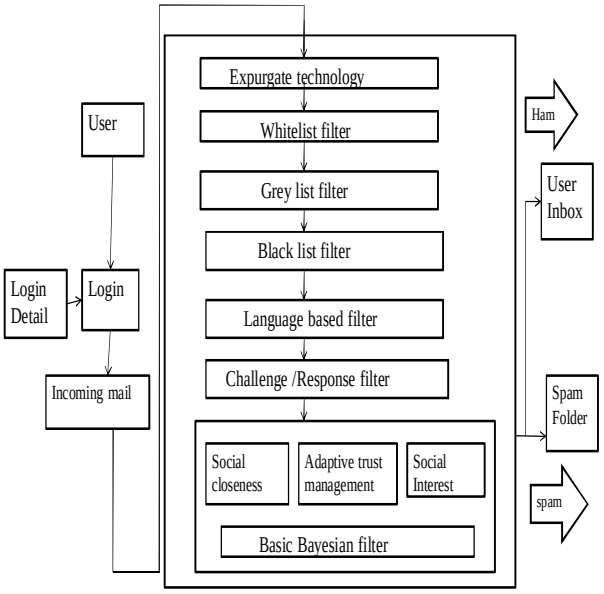


Figure 2. Proposed Approach

Language Based Filter is also used to remove out any incoming mail that is in any other language besides user's mailbox language preference. Challenge/Response filter sends an automated message that asks the sender to provide the return confirmation of his email address. If the filter has not recognized the incoming message as a white list or a black list, the Bayesian filter will be applied on <SUBJECT> field and the content <BODY> of the message. The filter scans through the message, and creates a probability of every word it knows about. Self learning Bayesian filter makes the approach more efficient and accurate. This approach integrates three new components to the Bayesian filter: (1) social closeness-based spam filtering (2) social interest-based spam filtering and (3) adaptive trust management. Based on the three social-based components, after parsing the keywords of an email, It adjusts the weights of the keywords. Then, it resorts to the Bayesian filter for spam evaluation. The weights are adjusted based on the closeness between the receiver and the sender, the receiver's (dis)interests, and the receiver's trust of the sender. If the closeness is high, the likelihood that the emails sent between them are spam is low, and then the weight is decreased otherwise weight is increased. Social closeness-based spam filtering helps filter to be resilient to poison attacks. Adaptive trust management helps filter to be resilient to impersonation attacks. Social interest-based spam filtering component contributes to the personalized feature. After processing the incoming mail through a all these filters, If the incoming mail is identified as SPAM then it is redirected into Spam folder otherwise the incoming mail is allowed to enter into the user inbox if it is identified as HAM means legitimate mail.

6. CONCLUSION

This paper provided the background problem caused by SPAM emails, and it also described the methodology of hybrid approach. This paper comprises a hybrid of the popular White List, Black List, Greylist, eXpurgate technology, Social closeness, Adaptive trust management, Social interest/disinterest and Bayesian Filters approaches that will effectively and accurately block the Spam mails and allow only legitimate mails to pass through to user's inbox based on the user's preferences. It is intelligent in the sense that it learns from the user's feedbacks and it is able to determine whether an incoming email message is a SPAM or not and also it adapts the new spam techniques.

7. REFERENCES

[1] Ze Li, Haiying Shen, "SOAP: A Social Network Aided Personalized and Effective Spam Filter to Clean Your E-mail Box",2011.

[2] Tarek Hassan, Peter Cole, Chun Che Fung, "An Intelligent SPAM filter – GetEmail5", 2006

[3] Khong,W.-K.(2001).The Law and Ecnomics of Junk Emails (SPAM).LAW AND ECNOMICS.Hamburg,University of Hamburg

[4] Moore D., Shannon C., Voelker G. M., and Savage S. "Internet Quarantine: Requirements for Containing Self-Propagating Code" IEEE INFOCOM April 2, 2003.

[5] Eric Allman "Features: Spam, Spam, Spam, Spam, Spam, the FTC, and Spam" Queue- Vol. 1 Issue 6, pages 62 - 69, September 2003

[6] Androutsopoulos I., Koutsias J., Chandrinos K. V., Paliouras G., and Spyropoulos C. D. "An Evaluation of Naïve Bayesian Anti-Spam Filtering" 11th European Conference on Machine Learning- Barcelona, Spain, pp 9-17, 2000

[7] Divya Tara Puvvula(2012) " SPAMKILLER: A TOOL FOR DETECTING AND FILTERING SPAM"

[8] Tarek Hassan, Peter Cole and Chun Che Fung "Towards Eradication of SPAM: A Study on Intelligent Adaptive SPAM Filters" Proceedings of the 5th PEECS Symposium, Perth, Western Australia, pp 203-206, September 2004

[9] C. Binzel and D. Fehr. How social distance affects trust and cooperation: experimental evidence from A slum. In *Proc. of ERF*, 2009

[10] Bindu V,Ciza Thomas "Spam War:Battling Ham against Spam"IEEE 2011

[11] "Anti-Spam Filtering Using Neural Networks and Baysian Classifiers" Proceedings of the 2007 IEEE International Symposium on Computational Intelligence in Robotics and Automation Jacksonville, FL, USA, June 20-23, 2007

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 763 - 770, 2013, ISSN: 2319–8656

Semantic Annotation: The Mainstay of Semantic Web

Thabet Slimani
Taif University
Taif, Saudia Arabia

Abstract:

Given that semantic Web realization is based on the critical mass of metadata accessibility and the representation of data with formal knowledge, it needs to generate metadata that is specific, easy to understand and well-defined. However, semantic annotation of the web documents is the successful way to make the Semantic Web vision a reality. This paper introduces the Semantic Web and its vision (stack layers) with regard to some concept definitions that helps the understanding of semantic annotation. Additionally, this paper introduces the semantic annotation categories, tools, domains and models.

Keywords: semantic annotation,

1. INTRODUCTION

Currently, despite the large amounts of documents and resources available online, semantic analysis is not enough supported by internet search engines (they typically match words syntactically). However, the requirement of massive metadata for the web content allows various Semantic Web applications to appear and gain broad approval. A typical Web application would provide and use new access methods based on the associated metadata. The Semantic Web was devised by Tim Berners-Lee as a network that includes a content semantically-enriched which contains links to explicit, formal semantics. A good number of the semantic content available has been generated automatically by the mean of wrapping or by using annotation services. However, Semantic Web success depends on the accomplishment of a great number of users creating and exploiting semantic content. This achievement requires tools that reduce the complexity of semantic technologies.

Various IE technologies are currently available allowing named entity recognition within the text, events, relations and scenarios in which they exist. Metadata assigned to a document can range from author reference of the document, to annotations of all the entities referred to in the text. To make this metadata readable by machines for effective structuring, discovery, automation, integration, and reuse is an important issue in semantic research. Based on the category of annotation, the automatic (versus manual) extraction of metadata approach is scalable, author-independent, and not expensive and enriches the web content of a specific user. At present, the technology available to provide automatic semantic annotation is not yet mature to achieve intuitive, scalable, and accurate model for generation and representation of such annotations.

This paper presents first a comprehensive introduction to Semantic Web (layers and content enrichment of web resources). Semantic annotation categories (manual, semi-automatic and automatic annotation) are presented in sections 3. Section 4 offers a brief explanation of the semantic annotation models and domains. Some existing annotation tools classified by text, images and ontologies are described in the section 5. Finally, section 6 gives a conclusion with perspectives for future work.

2. SEMANTIC WEB INTRODUCTION

The Semantic Web is a vision created by Tim Berners-Lee, the inventor of the WWW [1]. The success of the current WWW leads to a new challenge: A huge number of data is only human understandable; machine support is limited or absent. Berners-Lee suggests mechanisms to describe data in Semantic Web terms which will facilitate applications to exploit data (machine processable) in more ways and support the user in his task. The relevant pages and can thus improve both precision and recall. The definition of a Semantic Web structure is crucial. The structure has to be defined, and then has to be filled with life. To do this task, one should start with the easier tasks first. The following steps show the direction where the Semantic Web is heading:

- Provide a common syntax for machine understanding.
- Create common vocabularies.
- Support logical language
- Use the language for exchanging proofs.

The layer structure of the Semantic Web suggested by Tim Berners-Lee reflects the previous steps that follow the understanding that each step alone will already add value, so that the Semantic Web can be implanted in an incremental approach.

2.1 Layers of Semantic Web

The layers of Semantic Web suggested by Berners-Lee is a stack which shows how technologies that are standardized for the Semantic Web are organized to make the Semantic Web possible. This architecture is discussed in detail in [2] and [3], which also address recent research issues (Figure 1):

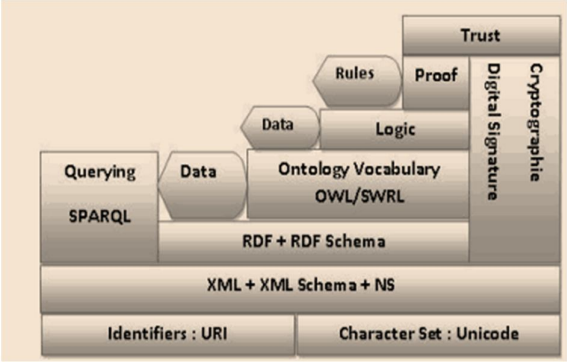


Figure. 1 The layers of Semantic Web.

The bottom layers contain technologies providing common syntax. Uniform Resource Identifier (URI) provides means for uniquely identifying semantic web resources (entities)[1] , while Unicode serves to represent and manipulate text in many languages useful for exchanging symbols. The Extensible Markup Language (XML) is a markup language that enables creation of documents composed of structured data, and XML Schema allows the definition of grammars for valid XML documents. Semantic web gives meaning (semantics) to structured data. XML documents can refer to different namespaces to make explicit the context (and therefore meaning) of different tags. XML Namespaces provide a way to use markups from more sources. Semantic Web aims to connect data together, which needs to refer more sources in one document.

The explained two layers are nowadays broadly accepted, and the number of XML documents is growing quickly. XML is the first step in the right direction, but it only formalizes the structure of a document and not its content. The Resource Description Framework (RDF)[2] is a framework for creating statements in a form denoted by triples. This form enables the representation of information about resources in the form of graph and can be seen as the first layer where information becomes machine understandable: According to the W3C recommendation[3], RDF "is a foundation for processing metadata; it provides interoperability between applications that exchange machine understandable information on the Web". The components of each RDF document consist of three types of entities: Resources (subjects and objects), properties (predicates/relations). Resources represent Web pages, parts or set of Web pages, or anything (real-world object) that can have a URI. Properties are specific attributes, or relations describing resources. The combination of a resource together with a property having a value for that resource forms a Statement (known as the subject, predicate and object). A value is either a literal, a resource, or other statement. A Statement which may be represented as a triple of the form (Subject, Property, Object) asserts that a resource recognized by the subject, has a property whose value is the recognized by the object (either another resource or a literal). Consequently, a property is a binary relationship between two resources or between a resource and a literal value. Figure 2 shows an example of RDF statements. Two of the Researcher of the "ASSW project" (i.e., their Web pages) are represented as resources 'URI-Mam' and 'URI-Ram'. On the lower right of Figure 2, the statement consists of the resource 'URI-Mam' and the property 'cooperates-with' with the value 'URI-Ram' (resource). The resource 'URI ASSW' has as value for the property 'title' the literal "Annotation System for Semantic Web".

RDF is basically a directed graph with labelled edges and partially labelled nodes. The definition of a simple modelling language on top of RDF is realized by the RDF Schema (RDFS)[4] which includes classes, IS-a relationships between classes and properties, and properties characterized by domain/range restrictions. RDF and RDF Schema are structured in XML syntax, but they do not use the tree semantics of XML. An extension of RDFS including more advanced constructs to describe semantics of RDF statements based on description logic is provided by Web Ontology Language (OWL)[4]. It allows states additional constraints, such as for example cardinality, value restrictions, or characteristics of properties such as transitivity. The ontology vocabulary denotes the next layer. Gruber [5], define an ontology as "an explicit formalization of a shared understanding of a conceptualization". Most of the definitions realized by different research communities share a certain understanding in common: That means, ontology is a formal model which explicitly represents the consensual knowledge of a domain. The domain entities are modelled through a set of concepts, a hierarchy on them, and relations between concepts. By instantiating these ontological concepts, concrete facts and information items which can be stored in the ontology. Most of these definitions also include axioms in some specific logic. The core of own ontology definition is presented in the following section.

At the layer of ontology vocabulary it is possible to query any RDF-based data (i.e., including statements involving RDFS and OWL) with the use of the latest RDF query language (SPARQL) [6]. According to Berners-Lee, the next layer is Logic. Nowadays, the integration between ontology and the logic levels is treated by the most researchers. This integration is encouraged by the ability of the most ontologies to allow for logical axioms. With the applicability of logical deduction, we can infer new knowledge from the information which is stated explicitly. For instance, the axiom given above allows to logically infer that the researcher addressed by 'URI-RAM' cooperates with the researcher addressed by 'URI-MAM'. The feasibility of the type of inference depends deeply on the logics chosen.

[1] Refers to a locatable URI, e.g., an http://www.w3schools.com/RDF address. It is often used as a synonym, although URLs are a subclass of URIs, see http://www.w3.org/Addressing
[2] http://www.w3.org/TR/REC-rdf-syntax/
[3] http://www.w3.org/TR/REC-rdf-syntax-grammar-20040210/
[4] RDF Vocabulary Description Language 1.0: RDF Schema. 2004 Available from: http://www.w3.org/TR/rdf-schema/.

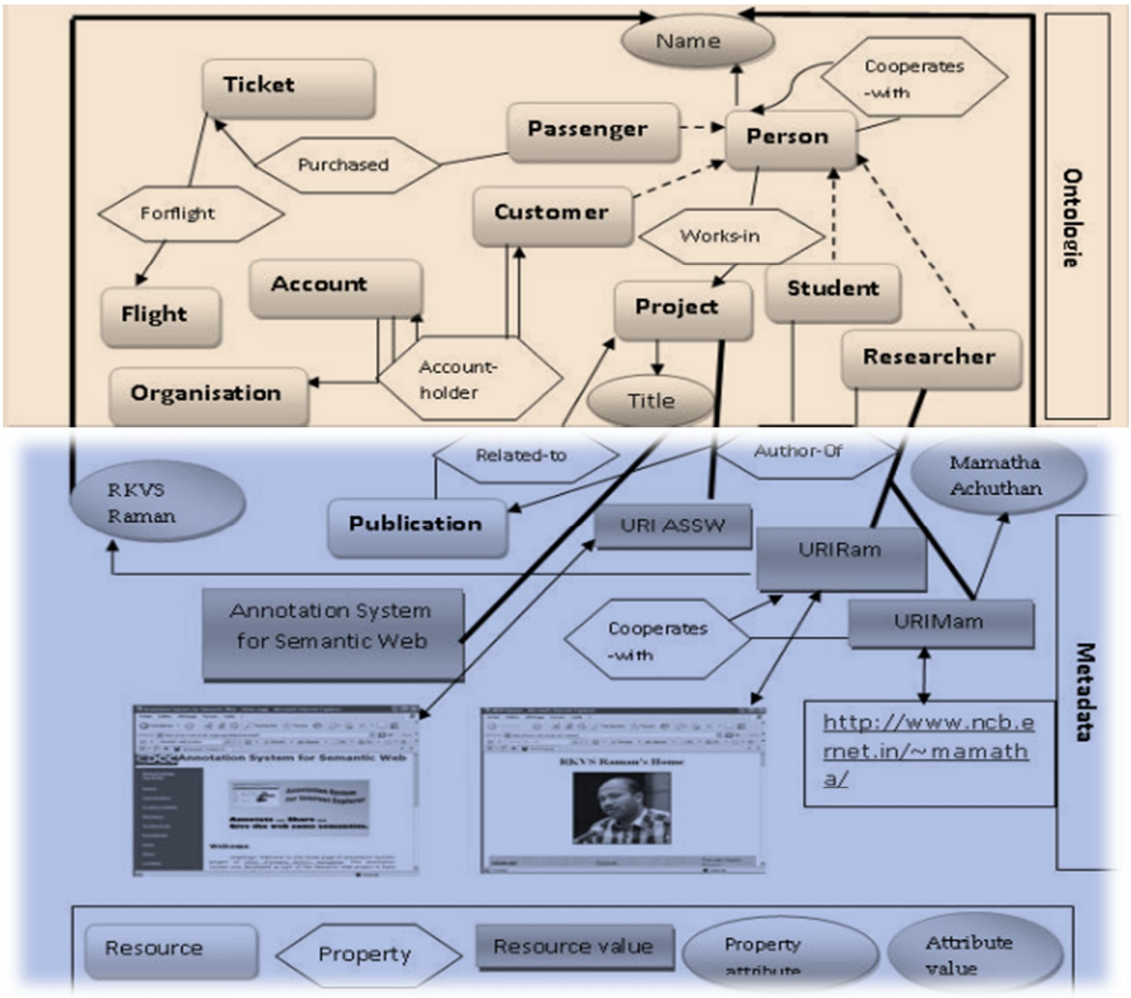


Figure. 2 The relation between resources in RDF graph and ontology.

The remaining layers are **Proof and trust**. The top layers contain technologies that are not yet standardized which require the ability to check the validity of the statements made in the (Semantic) Web, and Trust to derive statements will be supported by (a) verifying that the premises come from trusted sources and by (b) relying on formal logic during deriving new information. Consequently, the way it processes information will increase in the presence of validated statements. Then, the author must provide a proof which should be provable by a machine. At this level, it is not required that the machine of the reader finds the proof, but he needs only to check the proof given by the author. These two layers are rarely undertaken in current research.

3. SEMANTIC ANNOTATION CATEGORIES

Semantic annotation is the process that creates semantic labels of documents for the semantic Web, aiming to support advanced searching (based on concepts), reasoning about Web resources and the information visualization based on ontology. Additionally annotation is used to convert syntactic structures into knowledge structures. In other terms, semantic annotation consists to generate specific metadata and usage schema, enabling new information access methods and extending the existing ones.

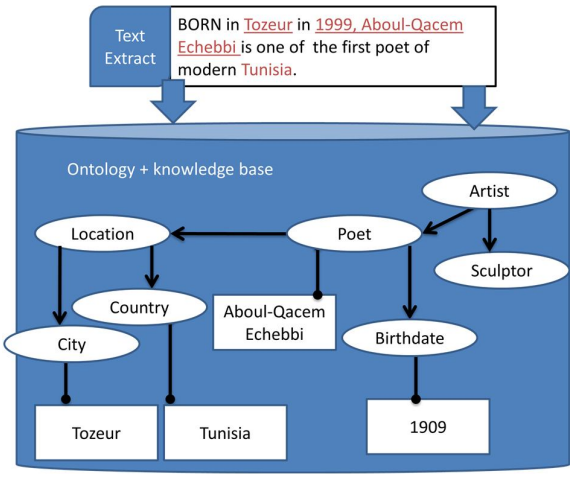


Figure. 3 The relation between resources in RDF graph and ontology.

In a nutshell, Semantic Annotation is about assigning to the entities in the text links to their semantic descriptions (as presented in Figure 3). Several types of annotation models are accessible in the literature and in the existing end-user applications. The models of semantic Web community have been abstracted by the Subject-Object-Predicate triple usefully for most of the annotation kind discussed here.

3.1 Manual Annotation

Manual annotation (MA) is a methodology that transforms the existing syntactic resources into interlinked knowledge structures by adding information to some level of document (word, phrase or paragraph) which constitutes metadata. The process of manual annotation is expensive, and regularly does not consider that multiple standpoints of a data source, involving multiple ontologies, can be useful to support the requirements of different users. Furthermore, MA is more easily feasible today, by means of authoring tools such as Semantic Word. MA is more precise compared to automatic annotation, but is very labor-intensive. As an example of manual annotation, in Protégé, a user can create an instance each time he wants to annotate by the selection of a piece of text in the loaded document and the class selection from the loaded ontology. As another example, in clinical documents, the instances that occur at different locations in the document actually refer to the same real world entity. Such annotations on equivalences can be valuable to deduce new knowledge and useful for medical care related applications [7]. The instance relationship annotation is another important type of annotation. Semantator [8] is an example of a tool that allows users to create a single relationship between two instances at a time. Semantator allows a given user to select two instances and add them to the relationship candidate list. After that, he can choose any object property from the loaded ontology and decide the subject of this new relationship. Another important type of annotation is instance relationship annotation allowing users to create a single relationship between two instances at a time.

3.2 Semi-automatic Semantic Annotation

The semi-automatic annotation process needs human intervention at some annotation level. This category of annotation systems differs in their architecture, methods and tools of information extraction, the manual work amount required to achieve annotation, performance, storage management and other features. GATE is an example of tools that performs semi-automatic annotation. The NCBO annotator [9] and cTAKES [10] are two other tools using semi-automatic annotation, in addition to Semantator. Similarly to that supported by NCBO annotator, cTAKES tool is used in Semantator for semi-automatic annotation. In a different manner to NCBO annotator, cTAKES is designed for clinical domain, uses NLP techniques and supports negation and time constraints. Currently, cTAKES allows annotation with the SNOMED CT and RxNorm dictionaries.

3.3 Automatic Semantic Annotation

Automated annotation of web documents is an important task in the Semantic Web effort. Semantic metadata created using automated annotation or tagging tools accompanied with best results are built on various machine learning algorithms which need training sets. Automated annotation tools can afford semantic metadata for semantic web and knowledge management [11]. Automatic semantic annotation can be realized on the base of automatic annotating algorithms: such as PANKOW (Pattern-based Annotation through Knowledge On the Web) and C-PANKOW (Context-driven and Pattern based Annotation through Knowledge on the Web) [12] for texts using Qtag[5] patterns and Google API. Additionally, Automatic semantic annotation can be based on statistical algorithms for image and video annotations. Nevertheless, current annotations based on automatic algorithms need to be improved and corrected. As an interesting tool for automatic semantic annotation, AeroDAML [13] loads a specific ontology and automatically produces the corresponding semantically marked up page which can then be checked by a human. KIM[6] [14], for example, is an automatic semantic annotation platform which uses information extraction based on GATE[7] [15] information extraction system, GATE with Annie[8] extension. SemTag [16] is the distinguished semantic annotation solution that operates within the distributed architecture and capability to process large scale data. To support annotation tasks, SemTag uses the Seeker [16] information retrieval platform and annotate web pages using Stanford TAP ontology [17]. Additionally, SemTag allows identifying but not creating new instances in the ontology. Ontea [18] is a platform for pattern based automated semantic annotation which allows pattern implementation over regular expressions, the implementation or integration of other methods based on patterns such as wrappers, solutions using structure of document, XPath, language patterns, e.g. C-PANKOW or GATE.

4. SEMANTIC ANNOTATION DOMAINS AND MODELS

4.1 Semantic Annotation Models

Semantic Annotation may be classified into four models: tags, attributes, relations and ontologies. Tags are located at the bottom level and correspond to the easiest form of annotation from the user point of view; while ontologies are at the top level and represent the hardest form of annotation from the user point of view.

- **Tags:** A tag annotation element is a keyword (word or sequence of characters without spaces) or a term assigned to a resource that, implicitly, describes a particular property of a resource. The annotation meaning planned by the annotator is not known by the computer and other users, only if the natural language used is unambiguous. The place names where picture are taken, the name of the person on a picture, or the topic of a news article are examples of tag.
- **Attributes:** An attribute annotation element is a pair of two elements: the name of the attribute and the value of the attribute. The name of the attribute defines the property of the annotated resource (e.g.,

[5] http://www.english.bham.ac.uk/staff/omason/software/qtag.html
[6] http://www.ontotext.com/kim/semanticannotation.html
[7] http://gate.ac.uk/
[8] http://gate.ac.uk/ie/annie.html

"Country", "birthdate") and the attribute value specifies the corresponding value (e.g., "Tunisia", "1909").

- **Relations:** a relation annotation element is a pair of two components: the relation name and the related resource. The annotated resource is related with relation by the relation name. In other words, the model of relation annotation is an extension of the model of attribute annotation to the domain of resources, allowing the user to interconnect these resources. For example, a citation referencing another paper in given scientific paper is an example of the annotation of relation defining a link between these documents.
- **Ontologies:** The ontology model describes the metadata that align a resource or a part of it with some of its properties and characteristics description according to a formal conceptual model (ontology). As defined by Studer et al., "an ontology is an explicit specification of a (shared) conceptualization" [5]. Ontologies are useful for domain knowledge capturing (in a generic manner) and the specification of a commonly granted understanding of a domain (that may be reused and shared within communities or applications). The design of ontologies may be realized with the following elements: concepts notion, concepts instances, concepts and instances properties, restrictions on these properties, relations between concepts and relations between instances. The user that use ontology annotation model is able to describe and connect existing resources by the resources structuring (concepts or as instances) and by the definition of the restrictions that hold between relations and properties.

4.2 Annotation Domains

Annotation domains were classified to document annotations, semantic wikis, semantic blogs, and tagging. A short introduction to each annotation domain with the specification of the associated role is presented in the following sections:

- **Document annotations:** Annotations of documents is the attachment of comments, notes, explanations, or other types of external remarks to a Web document or to a selected part of a document. Annotation of document can be realized manually (performed by user(s)), semi-automatic by automatic suggestions, or completely automatic.
- **Semantic Wikis:** Wiki can be defines as an environment for collaborative hypertext authoring that allows people collecting, describing, and authoring information in a collaborative manner. As a promising tools, semantic Wikis allow users to make formal descriptions of resources (wiki pages) enabling metadata insertions through semantic annotations and link relations between those resources. They need ontologies as conceptual models for their content organization. Annotations are needed to refer to an ontological model which defines concepts and properties associated to pieces of wiki contents.
- **Semantic Blogs:** A blog is a web site or online journal including comments, reflections and a lot of hyperlinks provided by the writer, but presented in reverse chronological order. However, the success of blogging can be reinforced when it accompanied with machine-readable content (annotation) which is beneficial of blog content consumers. Most commonly, an annotation in blogs is a statement about a post. For example, while a blog solution allows the classification of posts with simple categories or topics such as "research", "teaching"; we can say that blog posts are annotated with these categories. The process that transform blogs from simple online record to full participants in an information sharing network exploiting the metadata richness is called semantic blogging.

- **Tagging:** Organizing electronic content in a collaborative form by marking content with descriptive terms (tagging by keywords or tags) is a common way for future navigation, filtering or search. The tags express some undetermined relation between the resource and whatever the term refers to. As an example, del.icio.us[9], Technorati[10] or Flickr[11] are three tagging systems allowing users to associate one or more tags to a web resource.

5. ANNOTATION TOOLS

Annotation metadata can have several formats (textual, ontological, image or multimedia). The following paragraphs give a detailed description for each tool format.

5.1 Ontology-based annotation tools

Based on the criteria of tools that capture the requirement of providing explicit formal meaning to annotations, the following tool was selected:

- **Kim Plugin** [19] (Sirma Inc.): **KIM** platform is a part of the SWAN (Semantic Web ANnotator) project. It is a fully automatic and unsupervised tool for semantic annotation which works with its own meta-ontology.
- **Melita** [20] (University of Sheffield): Melita is a tool developed to define and develop an ontology-based annotation services. It is a semi-automatic annotation tool based on the Amilcare Information Extraction Engine [21].
- **Ont-O-Mat** [22] (AIFB): Is an implementation of the S-CREAM, a framework that supports both manual and interactive semi-automatic annotation of texts. Ont-O-Mat uses an automated data extraction technique from Amilcare (an adaptive Information Extraction) system designed to support active annotation of documents.
- **MnM** [23] (KMI): In a similar manner to Melita and Ont-O-Mat, MnM tool provides both automated and semi-automated support, based on the Amilcare system that support ontologies formalized in RDF.
- **C-Pankow** [24] (AIFB)[12]: C-Pankow is a fully automatic and unsupervised tool for semantic

[9] https://delicious.com/
[10] http://technorati.com/
[11] http://www.flickr.com/
[12] Smore 5.0 was not evaluated due to lack of support of annotation metadata.

annotation that support ontologies formalized in RDF.

5.2 Image Annotation Tools

Semantic annotation of images necessitates multimedia ontologies. Several vocabularies can be exploited (Dublin Core, FOAF), but they don't provide suitable models to describe sufficient multimedia content for sophisticated applications. The following section gives some example of image annotation tools:

- **ALIPR** [25] is a real-time automatic image tagging engine system fully automatic and high speed annotation for online pictures. It annotates images based on content.
- **GIAM** (Generalized image annotation methods) [26] [27][28] are designed to be used across a large number of images but needs a high intra-category clustering with adequate intercategory separation. Because the search space grows, categorical separation becomes challenging with GIAM.
- **SIAM** (Specialized image annotation methods): Conversely to GIAM, specialized annotators frequently perform well within their domains [29][30], but need a-priori assumptions about the data which, for a general image set are incorrect.
- **SpiritTagger** [31] : Is a system of image annotation invented to explore knowledge extraction through mining of millions of global photographs referenced with a geographical coordinate.
- **SML** (supervised multiclass labeling) [32]**:** it also 1) produces a natural ordering of semantic labels at annotation time, and 2) eliminates the need to compute a "nonclass" model for each of the semantic concepts of interest.
- **CRM (**Continuous-space Relevance Model**) [33]:** CRM is an image annotation and retrieval tool based on probabilistic model. It is designed to reduce an image to a real-valued feature vectors set, and subsequently model the joint probability of observing feature vectors by means of potential annotation words.
- **MBRM** (Multiple Bernoulli Relevance Models)[33]: MBRM is based on the CRM model presented above and uses a multiple-Bernoulli distribution for modeling image annotations over CRM.

5.3 Text Annotation Tools

There is a good number of textual information that would be more useful if it were annotated for the Semantic Web, but the nature of the data makes it difficult to do so. As examples of this type of data are the EBay posts texts, internet classifieds like Craig's list, bulletin boards such as Bidding for Travel, or even the text summary below the hyperlinks returned after querying Google. As example of tools:

- **Amaya** [34]: Amaya is an annotation tool allowing user to make annotations in the same tool they use for browsing and for editing text by mark-up Web documents in XML or HTML. It is a good example of a single point of access environment. Amaya manual annotation of web pages is allowed, but requires the features to support automatic annotation. All Amaya annotations may be realized by **Annozilla**[13] browser aiming to make readable in the shadow Amaya developments and the Mozilla browser.
- **AktiveDoc** [35]: AktiveDoc is a client-server application integrated in a Web Based KM system. It allows annotation of documents at three levels: free text statements, on-demand document enrichment and ontology based content annotation. AktiveDoc provides Semi-automatic annotation of content based on Amilcare. AktiveDoc is able to provide automatic suggestions about relevant content, given its design for knowledge reuse. ActiveDoc functionality is extended to free text, in addition to filling forms functionality and other pre-determined structures.
- **Magpie** [36]: Magpie is a real-time annotation of web resources that relates text strings to ontology concepts of the user's choice. With Magpie, an appropriate web service can be linked to highlighted strings. Even though, the annotation of documents is automatic, Magpie has the disadvantage to produce manually the lexicons specific parts of text strings subjects for each ontology. In Magpie, the work on automating lexicon generation is in progress.
- **Thresher** [37]**:** Is a similar system to Magpie (similarity in the use of wrappers that generates RDF on the fly as users browse deep web resources) that lets non-technical users teaching their browsers semantic web content extraction from HTML. Additionally, Thresher allows the user to access semantic services for recognizing objects. Since Thresher is part of the Haystack semantic browser [38], users can also do the personalization of the ontology that use.
- **WiCKOffice [39]**: WiCKOffice provides annotation for word processor files, similar to OntoOffice[14](A commercial annotation system for Microsoft Office applications available from Ontoprise). It helps in filling forms using data extracted from knowledge bases and demonstrates the effectiveness of writing within a knowledge aware environment to support possibilities (automatic assistance for form).

6. CONCLUSION

This paper has presented a number of issues related to the representation and the usage of the semantic annotation. Firstly, it describes some concepts of the Semantic Web, including metadata annotation which helps to make Semantic Web vision a reality. Additionally, this paper has presented three classes semantic annotation with some recognized systems (manual, semi-automatic and automatic). The domains, tools and the models of semantic annotations are also described. In future we plan to evaluate the described approaches of semantic annotation based in several criteria and different domains.

[13] Annozilla annotator (http://annozilla.mozdev.org/index.html accessed on 3 August 2004).

[14] OntoOffice tutorial (http://www.ontoprise.de/documents/tutorial ontooffice. pdf accessed on 30 November 2004).

7. REFERENCES

[1] Berners-Lee, T., Hendler, J., and Lassila, O. 2001. The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. In *Scientific American*.

[2] Patel-Schneider, P. and Fensel, D. 2002. Layering the semantic web: Problems and directions. In First International Semantic Web Conference (ISWC2002), Sardinia, *Italy*, 16–29.

[3] Patel-Schneider, P. and Simeon, J. 2002. Building the semantic web on xml. In *First International Semantic Web Conference (ISWC2002), Sardinia, Italy*, 147–161. Springer-Verlag.

[4] Dean, M., Connolly, D., Harmelen, F., Hendler, J., Horrocks, I., L. McGuinness, D., F. Patel-Schneider, P., and Stein, L. 2002. Owl web ontology language 1.0 reference. In *W3C Working Draft. http ://www.w3.org/TR/owlref/*.

[5] Gruber, T. 1993. Toward principles for the design of ontologies used for knowledge sharing. *International Journal Human-Computer Studies*, volume 43, number 5-6:pages 907–928.

[6] Prud'hommeaux, E. and Seaborne, A. 2005. *SPARQL:Query Language for RDF*.

[7] Tao, C., Solbrig, H.R., Sharma, D.K, Wei, W.Q, Savova, G.K, Chute, C.G. 2010. Time-oriented question answering from clinical narratives sing semantic-web techniques, Proceedings of the 9th international semantic web conference on the semantic web, November 07-11, Shanghai, China.

[8] Song, D., Chute, C. G., & Tao, C. 2011. Semantator: a semi-automatic semantic annotation tool for clinical narratives. In *10th International SemanticWeb Conference (ISWC2011)*.

[9] Noy, N. F., Shah, N. H., Whetzel, P. L., Dai, B., Dorf, M., Gri_th, N., Jonquet, C., Rubin, D. L., Storey, M.-A., Chute, C. G., and Musen, M. A. 2009. BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res, 37(Web Server issue), 170-173.

[10] Savova, G.K., Masanz, J.J., Ogren, P.V., Zheng, J., Sohn, S., Kipper-Schuler, K.C., Chute, C.G. 2010. Mayo clinical text analysis and knowledge extraction system (ctakes): architecture, component evaluation and applications. Journal of the American Medical Informatics Association 17(5), 507-513.

[11] Uren, V. et al. 2005. Semantic Annotation for Knowledge Management: Requirements and a Survey of the State of the Art. Journal of Web Semantics: Science, Services and Agents on the WWW, Vol. 4, 2005, No. 1, 14–28.

[12] P. Cimiano, G. Ladwig, and S. Staab, Gimme'. 2005. The context: context-driven automatic semantic annotation with c-pankow., in WWW, 332–341.

[13] Kogut, P. , Holmes, W. 2001. AeroDAML: Applying Information Extraction to Generate DAML Annotations from Web. In the Proceeding of First International Conference on Knowledge Capture (K-CAP 2001). Workshop on Knowledge Markup and Semantic Annotation.

[14] Kiryakov, A., Popov, B., Terziev, I., Manov, D., Ognyanoff, D. 2005. Semantic Annotation, Indexing, and Retrieval. Elseviers Journal of Web Semantics, Vol. 2, No. 1.

[15] Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V. 2002. GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL '02), Philadelphia.

[16] Dill, S., Eiron, N. et al.2003. A Case for Automated Large-Scale Semantic Annotation. Journal of Web Semantics.

[17] Guha, R., McCool, R. Tap: Towards a Web of Data. http://tap.stanford.edu/.

[18] Laclavık, M., Seleng, M., Hluchy, M.C.L. Ontea. 2009. Platform for Pattern Based Automated Semantic Annotation. Computing and Informatics, Vol. 28, 1000–1024, V 2009-May-13.

[19] Popov, B. et al. 2003. Kim - semantic annotation platform, in ISWC, 834–849.

[20] Ciravegna, F. et al. 2002. User-system cooperation in document annotation based on information extraction. In EKAW, 122–137.

[21] Ciravegna, F. et al. 2002. Amilcare: adaptive information extraction for document annotation. In SIGIR, 367–368, 2002.

[22] Handschuh, S., Staab, S. and Ciravegna, F. 2002. S-cream - semi-automatic creation of metadata. In EKAW, 358–372.

[23] Vargas-Vera, M. et al. 2002. Mnm: Ontology driven semi-automatic and automatic support for semantic markup. In EKAW, 379–391.

[24] Cimiano, P., Ladwig, G., and Staab, S. 2005. Gimme' The context: context-driven automatic semantic annotation with c-pankow., in WWW, pp. 332–341, 2005.

[25] Li, J. and Wang, J.Z. 2006. Real-time Computerized Annotation of Pictures. In the Proceedings *of the ACM Multimedia Conference*, 911-920, ACM, Santa Barbara, CA, October 2006.

[26] Wang, J. Z.; Li, J.; and Wiederhold, G. 2001. SIMPLIcity:Semantics-sensitive integrated matching for picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence 23:947–963.

[27] Zhang, Q., Goldman, S. A.,Yu,W., and Fritts, J. E. 2002. Content based image retrieval using multiple-instance learning. In International Workship on Machine Learning, 682–689.

[28] Chen, Y., and Wang, J. Z. 2002. A region-based fuzzy feature matching approach to content-based image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence 24:1252–1267.

[29] Szummer, M., and Picard, R. 1998. Indoor-outdoor image classification. In IEEE International Workshop on Content-Based Access of Images and Video Databases, 42–51.

[30] Vailaya, A., Jain, A., and Zhang, H. 1998. On image classification: City images vs. landscapes. Pattern Recognition 31:1921–1935.

[31] Moxley, E. Kleban, J. Manjunath, B. S. 2008. Spirittagger: a geo-aware tag suggestion tool mined from flickr, Proceedings of the 1st ACM international

conference on Multimedia information retrieval, October 30-31, Vancouver, British Columbia, Canada.

[32] Carneiro, G., Chan, A. B. Moreno, P. J. and Vasconcelos, V. 2006. "Supervised Learning of Semantic Classes for Image Annotation and Retrieval," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 29(3), 394-410.

[33] S. Feng, S., Manmatha, R. and Lavrenko, V. 2004. Multiple Bernoulli Relevance Models for Image and Video Annotation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*.

[34] Quint, V., Vatton, I. 2004. An Introduction to Amaya, W3C NOTE 20-February-1997, 1997. (http://www.w3.org/TR/NOTE-amaya-970220.html accessed on 28 July 2004).

[35] Lanfranchi, V., Ciravegna, F., Petrelli. D. 2005. Semantic Web-based document: editing and browsing in AktiveDoc, in the Proceedings of the 2nd European Semantic Web Conference, May 29–June 1, Heraklion, Greece.

[36] Dzbor, M. Motta, v, Domingue, J. 2004. Opening up magpie via semantic services, in: Proceedings of the 3rd International Semantic Web Conference, November 2004, Hiroshima, Japan.

[37] Hogue, A. Karger, v. 2005. Thresher: automating the unwrapping of semantic content from the World Wide Web, in: Proceedings of the 14th International World Wide Web Conference (WWW2005), May 10–14, Chiba, Japan, 86–95.

[38] Huynh, D., Kerger, D., Quan, D. 2002. Haystack: a platform for creating, organizing and visualizing information using RDF, in: Proceedings of the 11th International World Wide Web Conference (WWW2002), Hawaii, USA.

[39] Carr, L. Miles-Board, T., Woukeu, A., Wills, G., Hall, W. 2004. The case for explicit knowledge in documents, in: Proceedings of the ACM Symposium on Document Engineering (DocEng '04), Oct 28–30, Milwaukee, Wisconsin, USA, pp. 90–98.

International Journal of Computer Applications Technology and Research
Volume 2– Issue 6, 771 - 774, 2013, ISSN: 2319–8656

A Survey on Enhancing the Efficiency of Various Web Structure Mining Algorithms

Arun Kumar Singh
Department Of Computer Science
IIMT Engineering College
Meerut, India

Avinav Pathak
Department Of Computer Science
IIMT Engineering College
Meerut, India

Dheeraj Sharma
Department Of Computer Science
IIMT Engineering College
Meerut, India

Abstract - With the fast pace in internet technology, users get easily confused in large hyper text structure. Providing the relevant information to user is primary aim of the website owner. In order to achieve this goal, they use the concept of web mining. Web mining is used to categorize users and pages by analyzing the users' behavior, the content of the pages, and the order of the URLs that tend to be accessed in order [1]. Web structure mining executes very important role in this approach. It's defined as the process of analyzing the structure of hyperlink using graph theory. There are many proposed algorithms for web structure mining such as PageRank Algorithm, HITS, Weighted page rank Algorithm, Topic Sensitive Pagerank Algorithm (TSPR), weighted page content rank Algorithm (WPCR) etc. In this paper, we have described the outline of all the algorithms and identify their strengths and limitations with a broader survey and description.

Keywords: HITS, Page rank algorithm, TSPR, Web mining weighted pagerank

I. INTRODUCTION

Web mining [2] is an application of data mining [9]. As we know data mining is the process of extracting the useful information from large amount of data present in any organization. Web mining is defined as the process of extracting relevant information from World Wide Web data. Two different approaches are taken to define the web mining. 1) "Process-centric view" which defines web mining as sequence of task. 2) "Data-centric view" which defines web mining in terms of the types of web data that is being used in the mining process. Mainly web mining is divided into three types: 1) Web Content Mining 2) Web usage mining 3) Web Structure Mining. Web content mining is the process of extracting the useful information from web document. To carry out this task there are two main methods like agent based approach and database approach. Web usage mining is the process of mining the relevant information from web history. Web usage mining process can be divided into three stages: 1) Pre-processing 2) Pattern discovery 3) Pattern analysis. In preprocessing stage, data is cleaned and partitioned into set of user's transaction that represents the activity of each user during the visiting of different sites. In Pattern discovery stage database, machine learning and statistical operations are performed to obtain the hidden patterns that reflect the behavior of user. Pattern analysis: In this, discovered patterns are further processed, filtered and analyses in user model that can be used in model that can be used as input to applications such as visualization tools and report generation tools. Web structure mining is very complex task. It is the process of analyzing the hyperlink and extract relevant information from it. It is also used to mine the structure of a document, analyze the structure of page to describe the HTML or XML usage. The goal of the Web Structure Mining is to generate the structural summary about the Web site and Web page. Web Structure mining will further categorize the Web pages and generate the information like the similarity and relationship between different Web sites. This type of mining can be either performed at document level that is referred to as intra-page [11] or at hyperlink level that referred to as inter-page mining. Due to the significance of this mining technique, there have been several algorithms proposed to solve this. In this paper we will describe and analyze web structure mining algorithms and identify their strengths and limitations. The rest of this paper is organized as follows. In Section 2 Process of web structure mining is described and several proposed algorithms are introduced. Section 3 gives the detailed of these algorithms is given. In Section 4 we provide a Comparison of the different web structure mining algorithms. In Section 5 we conclude.

2. PROCESS OF WEB MINING

Web structure mining is also known as "Link Analysis" function. It is a past area of research but with the increasing interest in Web mining, the research of structure analysis had increased and these efforts had resulted in a newly emerging research area also called Link Mining. The Web contains a multiple variety of objects with almost no unifying structure, with differences in the style and content much larger than in traditional collections of text

documents. Link mining is divided into four parts and is shown in following figure:

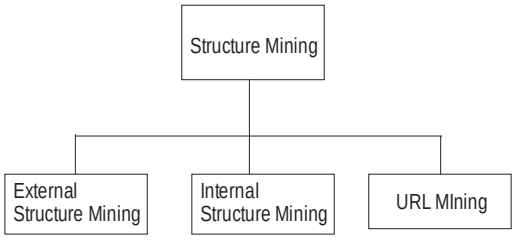


Fig 1.0

The objects in the WWW are web pages, and links are in-, out- and co-citation i.e. two pages that are both linked to the same page. There are some possible tasks [2] of link mining which are applicable in Web structure mining and are described as follows: **1. Link-based Classification: -** is the most recent upgrade of a classic data mining task to linked Domains. The task is to focus on the prediction of the category of a web page, based on words that occur on the page, links between pages, anchor text, html tags and other possible attributes found on the web page. **2. Link-based Cluster Analysis.** The aim in cluster analysis is to find naturally occurring sub-classes. The data is segmented into groups, where similar objects are grouped together, and dissimilar objects are grouped into different groups. Different than the previous task, link-based cluster analysis is unsupervised and can be used to discover hidden patterns from data. **3. Link Type.** There are a wide range of tasks concerning the prediction of the existence of links, such as predicting the type of link between two entities, or predicting the purpose of a link. **4. Link Strength.** Links could be associated with weights. **5. Link Cardinality.** The main task here is to predict the number of links between objects.

There are some uses of web structure mining like it is:

1. Used to rank the user's query
2. Deciding what page will be added to the collection
3. page categorization
4. finding related pages
5. finding duplicated web site and also to find out similarity between them

3. PROPOSED ALGORITHMS

A. Pagerank algorithm

Pagerank algorithm is link analysis algorithm[8] that was discovered by Larry page. This algorithm is used by Google internet search engine. In this algorithm numerical weight is assigned to each element of hyperlink set of document such as World Wide Web, with the purpose of measuring the relative importance of that particular set in that hyperlink. Pagerank is a probability distribution algorithm used to represent the person's randomly clicking on links will arrive at any particular page. A probability is expressed as a numeric value between 0 and 1. That numerical value is defined as damping factor. It is represented as d and usually its value set to be 0.85. Also C (A) is the number of link going out of that particular page and is known as backlink. Pagerank of any page is evaluated by: PR (A) = (1-d) + d (PR (T1)/C (T1) + ... + PR (Tn)/C (Tn)) (1) Where PR (A) is pagerank of particular web page A D is damping factor. PR (T1) is page link with main page PR (A) C is outlink. Page Rank can be calculated using a simple iterative algorithm, and corresponds to the principal Eigen vector of the normalized link matrix of the web. Page Rank algorithm needs a few hours to calculate the rank of millions of pages.

B. HITS

This algorithm [9] was given by Kleinberg in 1997. According to this algorithm first step is to collect the root set. That root set hits from the search engine. Then the next step is to construct the base set that includes the entire page that points to that root set. The size should be in between 1000-5000. Third step is to construct the focused graph that includes graph structure of the base set. It deletes the intrinsic link, (the link between the same domains). Then it iteratively computes the hub and authority scores. In HITS concept, he identifies two kinds of pages from the Web hyperlink structure: authorities (pages with good sources of content) and hubs (pages with good sources of links). For a given query, HITS will find authorities and hubs. According to him, A good hub is a page that points to many good authorities; a Good authority is a page that is pointed to by many good hubs". Although HITS provides good search results for a wide range of queries, HITS did not work well in all cases due to the following three reasons: 1 Mutually reinforced relationships between hosts. Sometimes a set of documents on one host point to a single document on a second host, or sometimes a single document on one host point to a set of document on a second host. 2. Automatically generated links. Web document generated by tools often have links that were inserted by the tool. 3. Non-relevant nodes. Sometimes pages point to other pages with no relevance to the query topic.

C. Weighted pagerank algorithms [6]:

Wenpu Xing and Ali Ghorbani proposed a Weighted Pagerank algorithm which is an extension of the Pagerank algorithm. This algorithm assigns a larger rank values to the more important pages rather than Dividing the rank value of page evenly among its outgoing linked pages, each outgoing link gets a value proportional to its importance. In this algorithm weight is assigned to both backlink and forward link. Incoming link is defined as number of link points to that particular page and outgoing link is defined as number of links goes out. This algorithm is more efficient than pagerank algorithm because it uses two parameters i.e. backlink and forward link. The popularity from the number of in links and out links is recorded as Win and Wout respectively. Win (v, u) is the weight of link (v, u) calculated based on the number of in links of page u and the number of in links of all reference pages of page v.

D. Weighted page content rank algorithm:

Weighted Page Content Rank Algorithm (WPCR) [7] is a proposed page ranking algorithm which is used to give a sorted order to the web pages returned by a search engine in response to a user query. WPCR is a numerical value based on which the web pages are given an order. This algorithm employs web structure mining as well as web content mining techniques. Web structure mining is used to calculate the importance of the page and web content mining is used to find how much relevant a page is? Importance here means the popularity of the page i.e. how many pages are pointing to or are referred by this particular page. It can be calculated based on the number of in links and out links of the page. Relevancy means matching of the page with the fired query. If a page is maximally matched to the query, that becomes more relevant. This algorithm is better than the pagerank as well as weighted pagerank algorithm because its complexity is less than both the algorithm and is < (Ologn).

E. Topic Sensitive PageRank Algorithm [5]:

In this algorithm, separate scores are evaluated, multiple important scores for each page under several topics that form a composite Pagerank score for those pages matching the query. At query time, the similarity of the query is compared to each of these vectors or topics; and subsequently, instead of using a single global ranking vector, the linear combination of the topic-sensitive vectors is weighed using the similarity of the query to the topics. This method yields a very accurate set of results relevant to the context of the particular query. For each web document query sensitive importance score. The results are ranked according to this composite score. It provides a better scalable approach for search rankings using Link analysis. For each Web page, compute an importance score per topic. At query time, these importance scores are combined based on the topics of the query and associated context to form a composite Page rank score for those pages matching the query. This score can be used in conjunction with other scoring schemes to produce a final rank for the result pages with respect to the query. This algorithm will improve the order of web pages in the result list so that user may get the relevant pages easily.

4. COMPARISON OF DIFFERENT ALGORITHMS

By analysing the literature review of significant web page ranking algorithms, it is concluded that each algorithm has some relative strengths and limitations. A tabular summary is given below in table 1.0, which summarizes the techniques, Advantages and limitations of some of important web page ranking algorithms:

5. CONCLUSION

This paper described various proposed web structure mining algorithms like pagerank algorithm, weighted pagerank algorithm, weighted content pagerank algorithm (WCPR), HITS etc. We examined their strengths and limitations and provide comparison among them. Hence, we can say that this paper may be used as a reference by researchers when deciding which algorithm is suitable. The present status defines that the algorithms have worked efficiently well and can be improved. Slight change in the parameters may fulfill the present day conditions but it has to checked that the goal is not disturbed.

6. REFERENCES

[1] R. Kosala and H. Blockeel, “Web Mining Research: A survey”, In ACM SIGKDD Explorations,

[2] Q. Lu and L. Getoor. Link-based classification. In *Proceedings of ICML-03*, 2003.

[3] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. *Technical report Stanford Digital Libraries SIDL-WP-1999-0120*, 1999.

[4] S. Pal, V. Talwar, and P. Mitra. Web mining in soft computing framework: Relevance, state of the art and future directions *IEEE Trans. Neural Networks*, 13(5):1163–1177 2002.

[5] Mishra Shesh Narayan et al (2012), “An Effective algorithm for web mining based on Topic sensitive pagerank algorithm”, International journal of computer science and software engineering.

[6] Wenpu Xing and Ali Ghorbani, “Weighted Pagerank Algorithm”, In proceedings of the 2rd Annual Conference on Communication Networks & Services Research, PP. 305-314, 2004.

[7] Bhatia Tamanna (2011), “Link analysis algorithm for web mining”, International journal of computer science and software engineering.(0976-8491).

[8] Munibalaji T and , Balamurugan C (2012), “Analysis of Link Algorithms for Web Mining”, International Journal of Engineering and Innovative Technology (IJEIT) Volume 1, Issue 2,

[9] Mukhopadhyay et al.(2006),” A Syntactic Classification based Web Page Ranking Algorithm”, 6th International Workshop on MSPT Proceedings.

[10] http://www.csbdu.in/econtent/Data%20Mining%20&%20 Warehousing/Unit%20III.pdf

[11] http://www.ieee.org.ar/downloads/Srivastava-tut-pres.pdf

Table 1 Comparison of Algorithms

Algorithm	Page Rank	HITS	Weighted Page Rank	Weighted Content Page Rank	Topic Sensitive Page Rank
Main Technique	Web Structure Mining	Web Structure Mining	Web Structure Mining and Content Mining	Web Structure Mining and Content Mining	Web Structure Mining
I/O Parameters	Backlink	Content ,Backlink, Forward Link	Content ,Backlink, Forward Link	Content ,Backlink, Forward Link	Content ,Backlink, Forward Link
Working	This algorithm computes the score of pages at the time of indexing of pages.	It computes the hubs and authoroties of relevant pages.	Weight of web pages is calculated on the basis of input and output links and on weight basis relevance of page is decided.	It gives weight to web links based on 3 attributes: relative position on page,tag where link is contained,length of anchor text	It computes the score of page according to the importance of content available on page.
Efficiency	Very Less	Moderate	Average	Average	Good
Search Engine	Google	Clever	Google	Google	Google
Significance	High, Backlinks are considered	Moderate , Hubs and Authorities are utilized.	HIgh, The pages are sorted according to relevance.	High	High , Score according to importance is calculated
Drawbacks	Results come at the time of indexing and not at query time.	Topic Drift and Efficiency problem.	Relevancy is ignored.	Relevant position is not effective resulting in illogical position of pages.	Only used for text , images are not.
Complexity	O(logn)	<O(logn)	<O(logn)	<O(logn)	<O(logn)
Quality of Result	Medium	Less than Page Rank	Higher than page rank	Higher	Higher than all.