IJACSA

WHERE WISDOM SHARES

International Journal of Advanced Computer Science and Applications

Volume 3 Issue 6

June 2012

ISSN 2156-5570(Online)
ISSN 2158-107X(Print)

SAI

www.ijacsa.thesai.org

IJACSA

WHERE WISDOM SHARES

INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS

THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org

Editorial Preface

From the Desk of Managing Editor…

IJACSA seems to have a cult following and was a humungous success during 2011. We at The Science and Information Organization are pleased to present the June 2012 Issue of IJACSA.

While it took the radio 38 years and the television a short 13 years, it took the World Wide Web only 4 years to reach 50 million users. This shows the richness of the pace at which the computer science moves. As 2012 progresses, we seem to be set for the rapid and intricate ramifications of new technology advancements.

With this issue we wish to reach out to a much larger number with an expectation that more and more researchers get interested in our mission of sharing wisdom. The Organization is committed to introduce to the research audience exactly what they are looking for and that is unique and novel. Guided by this mission, we continuously look for ways to collaborate with other educational institutions worldwide.

Well, as Steve Jobs once said, Innovation has nothing to do with how many R&D dollars you have, it's about the people you have. At IJACSA we believe in spreading the subject knowledge with effectiveness in all classes of audience. Nevertheless, the promise of increased engagement requires that we consider how this might be accomplished, delivering up-to-date and authoritative coverage of advanced computer science and applications.

Throughout our archives, new ideas and technologies have been welcomed, carefully critiqued, and discarded or accepted by qualified reviewers and associate editors. Our efforts to improve the quality of the articles published and expand their reach to the interested audience will continue, and these efforts will require critical minds and careful consideration to assess the quality, relevance, and readability of individual articles.

To summarise, the journal has offered its readership thought provoking theoretical, philosophical, and empirical ideas from some of the finest minds worldwide. We thank all our readers for their continued support and goodwill for IJACSA. We will keep you posted on updates about the new programmes launched in collaboration.

We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJACSA provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

We regularly conduct surveys and receive extensive feedback which we take very seriously. We beseech valuable suggestions of all our readers for improving our publication.

Thank you for Sharing Wisdom!

Managing Editor
IJACSA
Volume 3 Issue 6 June 2012
ISSN 2156-5570 (Online)
ISSN 2158-107X (Print)

Editorial Board

Reviewer Board Members

CONTENTS

A Study of FR Video Quality Assessment of Real Time Video Stream

Xinchen Zhang
School of Physical Science and Technology
Central China Normal University
Wuhan, China

Lihau Wu, Yuan Fang, Hao Jiang
School of Electronic Informaion
Wuhan Unicersity
Wuhan, China

***Abstract*—To assess the real-time transmission video's quality, this paper persents a approach which used FR video quality assessment (VQA) model to satisfy the objective and subjective measurement requirement. If we want to get the reference video in the measuring terminal and to make a assessment, there are two problems which are how to certain the reference video frame and how to make the objective score close to the subject assessment. We present in this paper a novel method of computing the order number of the video frame in the test point. In order to establish the relationship between the objective distortion and the subjective score, we used the "best-fit" regressed curve model and the BP neural network to describe prediction formula. This work is the mainly aim to get the high accurency assessment results with the human subjective feeling. So we select huge video sources for testing and training. The experimental results show that the proposed approach is suit to assess the video quality using FR model and the converted subjective score is available.**

Keywords- video quailty assessment (VQA); full reference; objective performance; subjective score; BP neural network.

I. INTRODUCTION

Video quality assessment (VQA) plays improtant role in various video communication applications, while the demand for video applications into everybody's life are rapidly growing. In these applications, the digital video is always processed through many stages before it display in the receiving terminal. The display video must be degraded in the end device, but there is an improtant question that the video has been more or less degraded. VQA provides the procedure to measure the quality of the video. The quality of the video can be described by the distortion to the original video or the feeling with human visual system (HVS). Then there are two mainly methods: the subjective score method and the objective estimation method. The subjective methods for video quality are consistent with actual conditions, but cost too much time. The subjective methods are not fit to real-time video application. So the objective video qualtiy assessment methods have been widely used, such as signal-to-noise ratio (PSNR), which needs the original video as the Full Reference (FR). But in some real-time video applications, the reference videos could not be obtained for the measure system or not get the reference video frame synchronization. Some existing researches proposed novel No-Reference (NR) objective methods to solve the problem about the absence of the reference video. However, it is difficult to design an objective NR VQA because of the limited understanding of the HVS.

The published papers about VQA can be classied into: (1) objective FR methods, (2) objective NR methods and (3) from objective to predict subjective scores motheds. These methods can be explained as follows:

Ref. [1] proposed a structural distortion measurement method for image based on FR, which is different from traditional error sensitiveity. The proposed method used multiscale structural similarity index (MS-SSIM) algorithm, which correlate with human visual preception significantly. Ref. [2] provides us the visual information fidelity index (VIF) and Ref. [3] explains video quality metric (VQM), which also use full reference video to measure the video quality. Ref. [4]-[6] give us some new algorithms based on SSIM to product different "structure" to estimate the video quality.

Ref. [7] provides an estimate of the mean square error distiortion at the marcoblock level without original content, duing working at the receiver sider for H.264 codec and considering the action of error concealment techniques. Ref. [8] introduced NR video quality assessment measure method based on coding quantization, packet loss and error progation and temporal effects of HVS.

Ref. [8] generally describes the objective and subjective VQA method, and emphatically introduces the subjective video database which made for wireless video applications and included 160 distorted videos. Ref. [9] gives us a standard test video database with the subjective sroces and the definition of the objective algorithms to calculate the video distortion, which was produced by VQEG born from a need to bring together experts in subjective video quality assessment and objective quality measurement in 1997.

After above introduciton, it is obvious that the objective measurement method is suitbale for real-time video quality measurement and the FR methods can estimate the distortion more accurately by using original video content. Although the NR methods got the achieved progess, the objective FR methods are still more reliable in practical industry applications. However, because we can not obtain the original video and synchronize the original video frame in the receiving terminal easily showing in Fig. 1, the FR methods were only used in encoding efficiency and distortion research generally. Then how to synchronize the frame of the decoding end to the original video is an important question to promote the objective FR methods for real-time video quality measurement. VQEG [9] reported and provided us that the PSNR, one of the objective FR performace, and the subjective scores of the test video database, also recommended the "best-fit" regressed curve to get the best evaluation relationship between the objective performances and the subjective scores. But in the practical real-time video application, such as video surveillance, video conference, and so on, the video contents are similar, or are belong to the same type, which is not like the test video database include a wide variety of the video contents. In real-time vdieo applications in a special scene, is there a better algorthm to evaluate the subjective scores from the objective performance which can be automatically calculated by computer. The main contribution of this paper is to propose a novel approach to synchronize the frame order between the sending and receiving end by adding the special marcoblock mark and study in using BP neural network to train and establish the maping of the objective performance and the subjective scores in oder to esitmate the VQA more better in real-time video applicaiton.

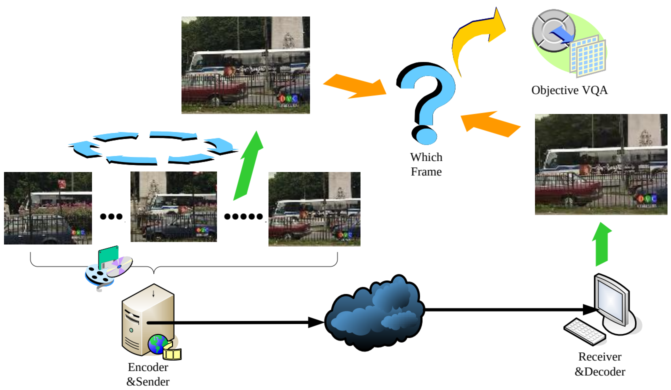


Figure 1. Which frame is the reference in the sender to the reveiver

The rest of the paper is organized as follows. Section II descirbes the novel approch how to add the special mark in the video before the encoder and to compute the frame order number to obtain the objecive preformance based on FR. Section III discusses the BP neural network principle and the training and evaluation procedure of the quality esitmaiton, which estblished the correlation with human visual results. In seciton IV, the experimental results are compared and discussed. Finally, the paper makes a conclusion of this paper.

II. Marking Method

In this section, we propose a novel method to add special mark in some macroblock and not destroy the video content. Using this adding mark, the frame's order number can be calculated in the receiver. Fisrtly, the marking method is introduced. Secondly, the judgement criterion of the parsing procedure is described. Finally, the processing results and the feasibility are discussed.

A. *Adding special mark*

In the past, the number of the video sequence can be marked by the semantic method in different video compression standard. But this method can only provide us the frame index, which clear the order to find the prediction relationship. In real-time video communication system, we also cannot confirm which frame into the encoder is the displaying one in the reveiving end. Then we cannot determine the reference frame to calculate the objective performance. In this paper, the proposed method is based on adding some additional information in the video content, which can be easily parsed to contain the reference frame order number.

The novel method to mark the order contains two procedures as Fig. 2 shows: adding the special mark to the original video before encoding, parsing video sequence order number after decoding. The details of the first procedure is described as follows.

- Define the order number of the original video sequence (Format YUV420) frame in binary system.
- Select first 4 marcoblocks (16x16) of the luminance component of the original video, (maybe select more marcoblocks accoding to how many frame).
- Conbine every 2-bit of the binary of the frame number, and replace the 2-bit value by 4 possible values complying with the rule. The rule is (00) to 0, (01) to 85, (10) to 170, (11) to 255.
- Using these values to fill the selected marcoblocks' luminance component, then the added-mark video sequence was produced.

For example, one video sequence contains 256 frames. Then we can select the first 4 marcoblocks to fill 4 possible values (4^4=256) to represent the 256 frames' number. If we are dealing with the number 60 frame, it's number can be notated by (00111100) in binary. Then we can deduce that (0, 255, 255, 0) are the values to fill the first 4 marcoblocks accoding to the rule.

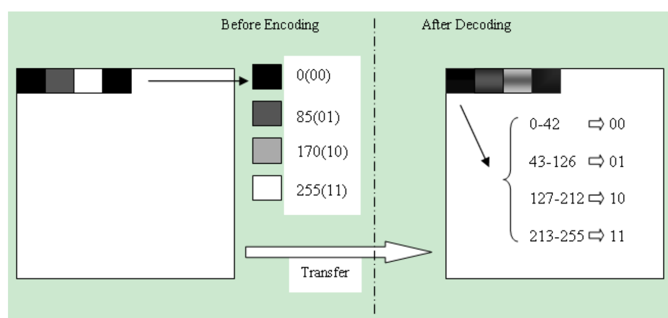


Figure 2. The procedures of the adding mark and parsing number

B. *Parse the frame number*

Through the real-time video communication system, the receiving video can be decoded from the payload video stream. If the decoder video format is YUV 4:2:0, we can observe the luminance component of the decoded video. By parsing the first 4 marcoblocks' values, we can calculate the average of the pixel value in one marcoblock and get the number of the decoded frame in the encoding end. The analysis criterion and the processing is as follows.

- Calculating the average pixel value of the marcoblock. We define A to represent the average, then A= Σ

$x_{ij}/256$, $i=0,1,2,\cdots,15$, $j=0,1,2,\cdots,15$. x_{ij} means the value of the pixel which is in the 16x16 marcoblock.

- After calculating, there are four average values which can be represented by (A_1, A_2, A_3, A_4). Judging A_j ($j=1,2,3,4$), it belongs to which renge. There are four renges devided in the uniform. The judgement is as the following formula (1):

$$A_j \in \begin{matrix} [0,42], then & (00) \\ [43,126], then & (01) \\ [127,212], then & (10) \\ [213,255], then & (11) \end{matrix} \quad (1)$$

- Finally, (A_1, A_2, A_3, A_4) is converted to 8 bit binary which is the order number of the frame in original. It is easy to transfer it to decimal digits.

Additional explanation is that the pixel values of the marcoblock maybe be changed after coding, transfering and decoding. However the features of the marcoblock an not changed more, and the average value will not exceed the boundary of the renge.

C. Discussion

In order to prove the availability of the novel method, we input the video adding the marks to the encoder with extreme QP and transmission parameters setting, it also can promote correct the order of the video frame.

For this purpose, the test video sequence is named BUS, which is CIF format (352x288, YUV 4:2:0, 30fps, 150 frames) and download form JVT test video. Firstly, because this sequence has 150 frames which is not exceed 255, we only select the first 4 marcoblocks to mark the frame number. They are on the top-right of the picture. According to the order number of the frame, the different values are filled to the each marcoblock's luminance pixels. Fig. 3 shows the number 39 frame had been added the mark marcoblocks. The marked marcoblock's luminance value are (0, 170, 85, 255), which means (00100111) in binary.

Figure 3. The added-mark video, frame number is 39.

In order to test the limit case in the worse, we select the H.264 encoder's QP as 48. In this situation, the picture quality is very bad with very low bit-rate. Fig. 4 shows the encoded video after adding the mark of the first 4 marcoblocks. The test result prove that the high compression distortion cannot change the feature of the adding mark, also means the coding error cannot effect the mark's judgement. And the wide boundary of the judgement range can guarantee the correctness of the method. The right frame number will be obtained in the reveiving end, which helps to seek the objective distortion of the decoded video and the reference video.

Another effection is the transmission error. This error maybe loss the first 4 blocks or change the features of the mark. If this happened, that means the channel cannot transmit the correct data. The first 4 marcoblock is following the picture header information. If they are not correct, the other data are very likely incorrect. Then the video quality becomes very low and the system are not suit video communication. So if the added mark exceed the range, we cannot find the right number of the frame, it represents the video system is abnormal. Error judged number brings the objective performance to bad, this just explains the poor video quality.

Figure 4. The encoded video with QP 48 after adding mark

III. ESTABLISH THE SUBJECIVE AND OBJECTIVE ESTIMATION

In order to research the relationship between the subjective score and the objective distortion, the BP neural network training method and the "berst-fit" regressed curve model will be introduced in this section. Using those methods, it is easy to convert the objective results based on full reference model to the subjective score as human visual evaluation. The subjective score is regarded as the final VQA result.

A. BP neural network

Back-Propagation (BP) neural network (NN) is a common method of training artificial neural networks so as to minimize the objective function. BPNN is a good idea for the situation, while there are a large amount of input/output data is available and we do not sure the realtionship which is complex and variable problem. The past researches can provide a large amount of the training video and the subjective scores as output data. Because the subjective human visual is not a certain question, which depend the human experience and mood. Recently, some neural-based approaches to video quality evaulation have been proposed [10]-[12].

BP neural network can learn and store a lot of input - output mode mapping, without prior mathematical equations that describe the mapping relationship. The learning rule is to use the steepest descent method, back-propagation to adjust the network weights and thresholds, so that the squared error of the network is minimum. BP neural network model topology including input layer, hidden layer and output layer. BP network is a multilayer feed-forward neural network. The neuron transfer function is S-shaped function, the output of the continuous quantity between 0 to 1, it can be any nonlinear mapping from input to output $y = f(x)$. The mapping function $y = f(x)$ is estimated during a traning phase, where the network learns to associate input vectors to output vectors. The main advantageof the NN method is its ability to process nonlinear problems of VQA [12].

We design the BPNN include 2 part researches, one is the structure of BPNN and another is the training rule of studying of the network connection point weights variation. Fig. 5 shows a basic BP neural model, which has R inputs, each input through an appropriate weight and the next layer is connected to the network output can be expressed as:

$$a = f(w \times p + b) \tag{2}$$

In (2), f is the transfer function to describe the relationship between inputs and outputs.

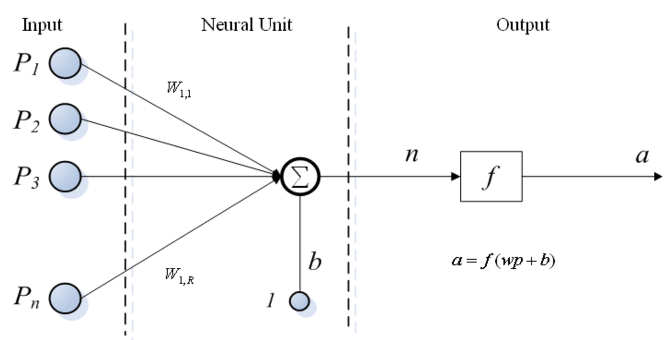


Figure 5. The BP neural model

B. Training and prediction method

Evaluation model is to train the neural network on fact. Before training it is need to construct network architecture. It requires four input conditions were: maximum and minimum of R-dimensional input samples constitute dimensional matrix, the number of neurons in each layer, the transfer function of the layers of neurons, and the training function. Now we build a two-level neural network. The first level contains 15 neurons n_1, the transfer function is *tansig*(). The second level is single neuron n_2, its transfer function named *traingd*() is linear model for training. Fig. 6 shows the two-level structure of the BP neural network, in the figure it only darws one-dimensional network structure of the input samples. $IW^{1,1}$ is the input layer to hidden layer weights. $LW^{2.1}$ is the hidden layer to output layer weights. b^1 is the adjustment paremeter of the input layer to hidden layer. b^2 is the adjustment parameter of the hidden layer to output layer.

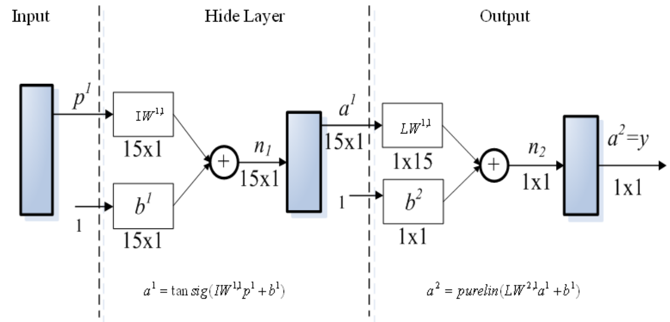


Figure 6. The structure of the two BP neural network for training

In training, another important thing is the input training data. For the aim of the training, we use MS_SSIM and MOVIE which are the objective quality performace parameters as the input vector *p*.

C. The "best-fit" regressed curve model

In statistics, nonlinear regression is a form of regression analysis in which observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. The data are fitted by a method of successive approximations. Curve fitting is the process of constructing a curve, or mathematical function, that has the best fit to a series of data points, possibly subject to constraints.

The "best-fit" regressed curve model is recommended by VQEG [13] and has been widely used in evaluating the performance of algorithms. We used a 4-parameter logistic function, constrained to be monotonic to transform the objective score to the subjective one, as in (3),.

$$Q_j^{'} = \beta_2 + \frac{\beta_1 - \beta_2}{1 + e^{-(Q_j - \beta_3 / |\beta_4|)}} \tag{3}$$

In the following experiments, we use function *nlinfit*() in Matlab softare to find the best values for fixing the parameter β. In (3), the VQA value of the objective esitmation method is represented by $Q_j, j = 1,2,3...150$, and $Q_j^{'}, j = 1,2,3...150$ is the VQA score of the subjective estimation method form human.

IV. Experimental Results

For the experiment, a total 150 video sequences are generated from 10 reference video sequences (named "Blue Sky", "River Bed", "Pedestrian area", "Tractor", "Sunflower", "Rush hour", "Station", "Park run", "Shields", "Mobile and Calendar") download from VQEG [13]-[14] and LIVE database [15]. The uncompressed RAW sequences are YUV 4:2:0 format with a resolution of 768x432. One original video are subject to H.264 video coding with different bit-rate (from 200kbps ot 5Mbps) and different packet loss rate (3%、5%、10%、20%) to simulate the network transmission loss, different bit error (0.5%-10%) as the stream in the wireless channel, MPEG-2 video coding with different bit-rate (from 700kbps to 4Mbps). Each RAW sequence changes into 15 video sequences through different processing, which include different main coding distortion and transmission distortion to improve the measurement experiment stability. These 10 RAW vdieo sequences are described as Table I.

TABLE I. THE RAW REFERENCE VIDEO SEQUENCES

Name	*Play Time*	*Frames*	*Description*
Blue Sky	8.68s	217	Circle move, tree
River Bed	10s	250	Camera still, river
Pedestrian area	10s	250	Camera still, street
Tractor	10s	250	Horizontal move
Sunflower	10s	250	Camera still, flower
Rush hour	10s	250	Camera still, Traffic
Station	10s	250	Camera still, tarin
Park run	10s	500	Horizontal move
Shields	10s	500	Horizontal move
Mobile&Calendar	10s	500	Horizontal move

Then we calculate the objective performance using MS-SSIM and MOIVE. Paper [16] compared the evaluation performance and computational complexity of the various objective algorithms. The PSNR is only used as an indicator of quality. MS-SSIM seems to perform the best amongs the algorithms. It is easy to implement and real-time estimates for available. The multiscale SSIM index [17] corrects the viewing-distance dependence of SS-SSIM and accounts for the multiscale nature of both natural images and human visual system. The MS-SSIM index performs better (relative to human opinion) than the SS-SSIM index on images. The computation formula of MS-SSIM is as (4)

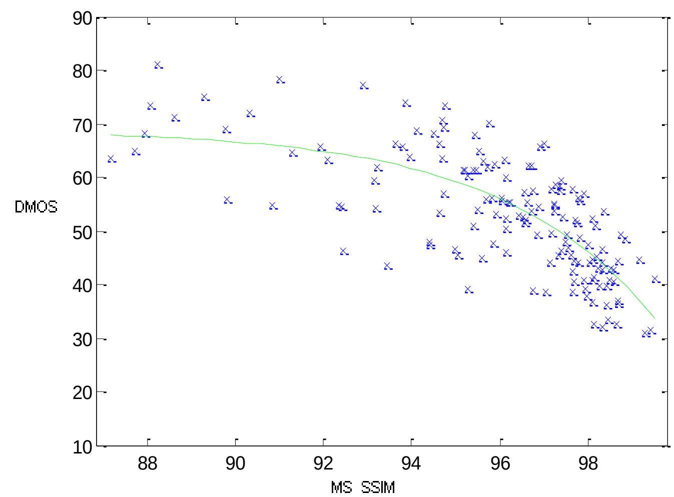
$$MS\_SSIM(x,y)=[l_M(x,y)]^{\alpha_M}\cdot\prod_{j=1}^{M}[c_j(x,y)]^{\beta_j}[s_j(x,y)]^{\gamma_j} \qquad (4)$$

Motion-based Video Integrity Evaluation (MOIVE) [18] intergrates both spatial and temporal aspects of distortion assessment. MOIVE explicitly uses motion information from the reference video and evaluates the quality of the test video along the motion trajectories of the reference video. MOVIE contains five main processing steps. The first composition named Linear Decomposition is to decompose the reference and test videos into multiple spatio-temporal bandpass channels using a family of Gabor filters. The second composition is to measure spatial distortions of the reference and test videos, named Spatial MOIVE. The third composition is Motion Estimation, which computes from the reference video sequence in the form of optical flow fields. The fourth composition is Temporal MOVIE which uses the spatio-temporal Gabor decompositions of the reference and test video sequences, and the optical flow field computed from the reference video using the outputs of the Gabor filters to esitmate the temporal video quality. The last step is to compute the MOVIE Index, as (5).

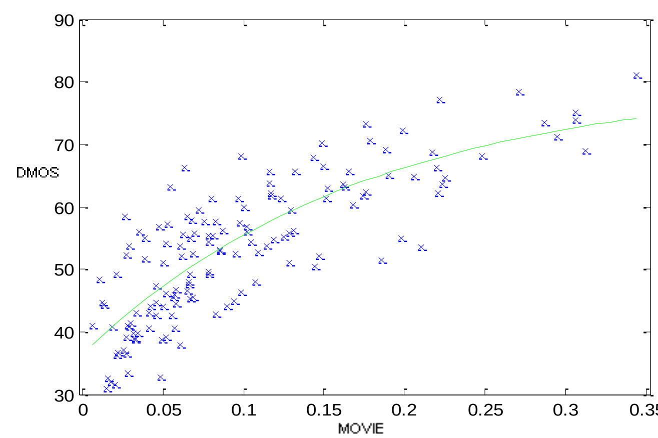
$$\text{MOVIE}=\text{Spatial MOVIE}\times\text{Temporal MOVIE} \qquad (5)$$

In (5), Paper [18] said $\text{Spatial MOVIE}=\frac{1}{\tau}\sum_{j=1}^{\tau}FQ_S(t_j)$, $FQ_S(t_j)=\frac{\sigma_{Q_S(x,y,t_j)}}{\mu_{Q_S(x,y,t_j)}}$, $\text{Temporal MOVIE}=\sqrt{\frac{1}{\tau}\sum_{j=1}^{\tau}FQ_T(t_j)}$, $FQ_T(t_j)=\frac{\sigma_{Q_T(x,y,t_j)}}{\mu_{Q_T(x,y,t_j)}}$.

The experimental results of the "best-fit" regressed curve model is shown as Fig. 7. The scatter plots of the MS-SSIM versus DMOS is with the best parameters $\beta_1=-1563.646$, $\beta_2=68.847$, $\beta_3=112.587$, $\beta_4=3.427$. The scatter plots of the MOIVE versus DMOS is with the best parameters $\beta_1=80.273$, $\beta_2=-1798.651$, $\beta_3=-0.645$, $\beta_4=0.175$.



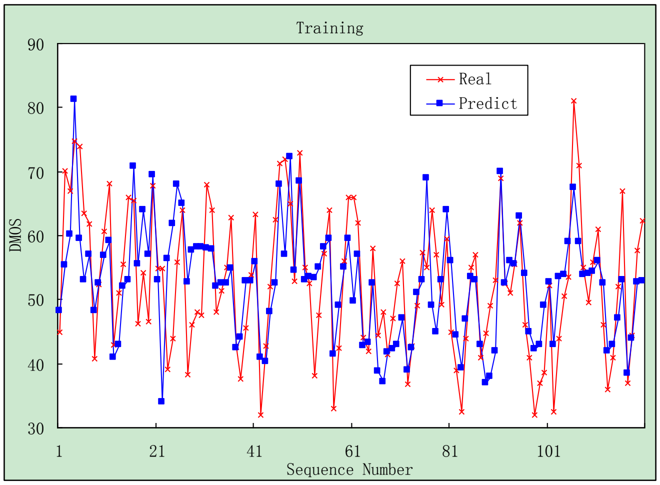
(a) The scatter plots of the MS-SSIM versus DMOS



(b) The scatter plots of the MOVIE versus DMOS

Figure 7. The "best-fit" regressed curve results

We use 120 video sequences as the training inputs for BP neural network, and the rest of 30 sequences as the test data. As mentioned earlier, the training is for study of fixing the network parameters or relationship, and the testing uses this network to estimate the subjective scores.



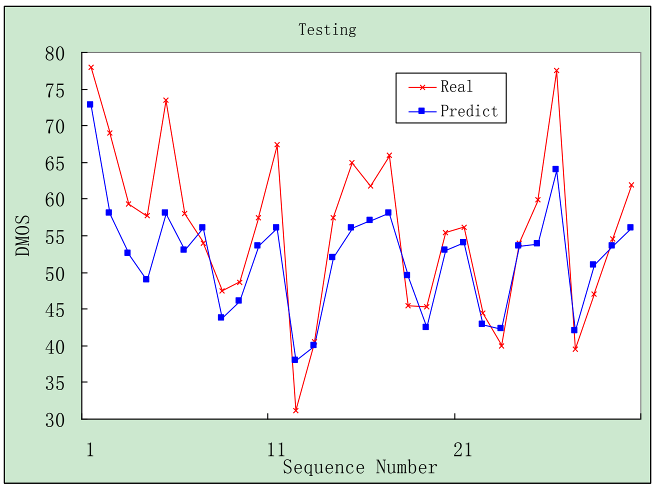
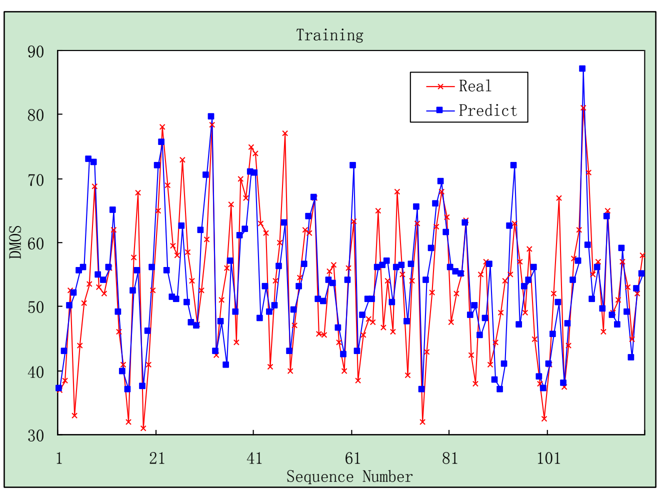


Figure 8. BP Neural network experimental results (MS-SSIM VS. DMOS)

Fig. 8 gives the experimental results about MS-SSIM versus DMOS. The two polylines are the subjective values by real test and the predictive values from BPNN computation.



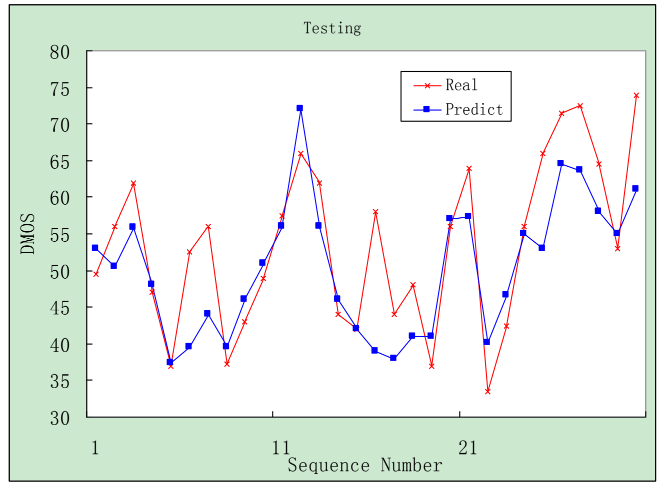


Figure 9. BP Neural network experimental results (MOIVE VS. DMOS)

Fig. 9 gives the experimental results about MOIVE versus DMOS.

For performance evaluation suggested by the VQEG, we used three metrics to evaluate the performance of the proposed method. The root-mean-squared error (RMSE), the Pearson correlation coefficient (PCC), the spearman rank ordered correlation coefficient (SROCC) were computed between the fitted objective data and the corresponding subjective data. The performance comparison of those algorithms is as Table II shows. From the results, we concern that the BPNN is close to the best-fit curve performance, and MOIVE algorithm is slightly better than MS_SSIM.

TABLE II. PERFORMANCE OF

	Name	***PCC***	***SROCC***	***RMSE***
Best Fit Curve	***MS_SSIM***	0.7387	0.7321	7.3982
	MOVIE	0.8116	0.7890	6.519
BPNN	***MS_SSIM***	0.7008	0.6771	7.0774
	MOVIE	0.7882	0.7678	6.9719

V. CONCLUSIONS

This paper has proposed a novel method to calculate the order number of the reference video in receiving end by adding special mark in marcoblocks. The method had been used in video conference system VQA and web stream VQA to solve which frame can be used as the reference in FR model. The method can resist to high coding error from the video encoder, while it can obtain the right frame number in the receiving terminal after the video decoded. Then this paper designs a two-level BP neural network structure to fix the relationship between the objective performance and the subjective DMOS score. We use a large of video sequences to train the BPNN and to fix the parameters of each level. The experimental results shows that it's performace is close to the "best-fit" regressed curve. In future, it is an interesting study on using BPNN to sovle the NR VQA. We will also care the research on wireless video quality assessment application.

ACKNOWLEDGMENT

This work was supported by self-determined research funds of CCNU from the colleges' basic research and operation of MOE.

REFERENCES

[1] Z. Wang, E.P. Simoncelli, and A.C. Bovik, "Multiscale structural similarity for image quality assessment,"in Proc.IEEE Asilomar Conf. Signals Syst. Comput., pp.1398–1402, Nov. 2003.

[2] H.R. Sheikh and A.C. Bovik, "Image information and visual quality," IEEE Trans. Image Process., vol.15, no.2, pp.430–444,Feb. 2006.

[3] M.H. Pinson and S.Wolf, "A new standardized method for objectively measuring video quality," IEEE Trans. Broadcasting, vol.50, no.3, pp.312–313,Sep.2004.

[4] Z.Wang and Q.Li, "Video quality assessment using a statistical model of human visual speed perception," J. Opt. Soc. Am., vol.24, no.12, pp.B61–B69, Dec.2007.

[5] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," IEEE Signal Process. Lett., vol.13, no.4, pp.600–612, Apr.2004.

[6] A .K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," IEEE J. Sel.Topics Signal Process., vol.3, no.2, pp.193–201, Apr.2009.

[7] M. Naccari, M. Taglisasacchi and S. Tubaro, "No-reference video quality monitoring for H.264/AVC coded video", IEEE Trans. on Multimeida, Vol. 11, No. 5, pp:932-946, August, 2009.

[8] F.Z. Yang, S.Wan, Q.P. Xie, H.R. wu, "No-Reference Quality Assessment for Networked Video via Primary Analysis of BitStream," IEEE transactions circuits and system for video technology, vol. 20, No.11, 2010, pp:1544–1554.

[9] Video Quality Experts Group (VQEG), Final Report from the Video Quality Experts Group on the Validation of Objective Models of video quality assessment phase II, August 2005, http://www.its.bldrdoc.gov/vqeg/project-pages/frtv-phase-ii/frtv-phase-ii.aspx.

[10] P. Gastaldo, S. Rovetta and R. Zunino, "Objective quality assessment of MPEG-2 video stream by using CBP neural networks," IEEE Transl. on Neural Networks, vol. 13, No. 4, pp. 939–947, 2002.

[11] A. Bouzerdoum, A. Havstad, A. Beghdadi, "Image quality assessment using a neural network approach", Proceedings of the Fourth IEEE International Symposium on Signal Processing and Information Technology, December 2004, pp. 330-333.

[12] J. Choe, K. Lee, C. Lee, "No-reference video quality measurement using neural networks", 16th International Conference on Digital Signal Processing, July 2009, pp. 1197-1200.

[13] Video Quality Experts Group (VQEG), "Final Report from the VQEG on the validation of Objective Quality Metrics for Video Quality Assessment", www.vqeg.org, March 2000.

[14] Video information processing lab. YUV test sequences. Available: http://vip.cs.nctu.edu.tw/resource_seq.html.

[15] A.K. Moorthy, A.C. Bovik. Live Wireless Video Quality Assessment Database [Online]. Available: http://live.ece.utexas.edu/research/quality/live_wireless_video.html.

[16] A.K. Moorthy, K. Seshadrinathan, R. Soundararajan, etc. Wireless video quality assessment: a study of subjective scores and objective algorithms, IEEE Trans. on circuits and systems for video technology, vol. 20(4), april (2010) 587-599

[17] Z. Wang, E.P. Simoncelli, and A.C. Bovik, "Multiscale structural similarity for image quality assessment," in Proc. IEEE Asilomar Conf. Signals Syst. Comput., Nov.2003, pp.1398–1402.

[18] K. Seshadrinathan and A. C. Bovik, "Motion Tuned Spatio-temporal Quality Assessment of Natural Videos", IEEE Transactions on Image Processing, vol. 19, no. 2, pp. 335-350, Feb. 2010.

AUTHORS PROFILE

Xinchen Zhang was born in Wuhan, China in 1977. He graduated from Wuhan University with a B.Sc. Degree in Eelectronic and Information System and M.Eng. Degree in Circuit and System in 1999 and 2001. He earned a Ph.D in Photograghic Measurence and Remote Sense in 2006. Now he is Associate Professor, Pyscial Science Technology School, Central China Normal University. Dr. Zhang is a registered member of IEEE. He is awarded China Scholarship Council Sponsorship to visit Lioms, ISIMA, UBP, Clermont-ferrand, France in 2011-2012.

Lihua Wu was born in Hubei, China in 1987. He graduated from Wuhan University with B.Eng. Degree in Comm. Engineering in 2009. He is a Ph.D student in School of Electronic Information, Wuhan University.

Yuan Fang was born in Hubei, China in 1988. She graduated from CCNU with B.Eng. Degree in Information Technology in 2010. Now She is a Master student in School of Electronic Information, Wuhan University.

Hao Jiang was born in Wuhan, China in 1976. He graduated from Wuhan University with a B.Eng. Degree in Eelectronic Engneering in 1998 and M.Eng. Degree in Communication Engneering in 2001. He earned a Ph.D in Communication Engneering in 2004. Now he is Associate Professor, School of Electronic Information, Wuhan University.…

Effective Listings of Function Stop words for Twitter

Murphy Choy
School of Information System
Singapore Management University
Singapore

***Abstract*— Many words in documents recur very frequently but are essentially meaningless as they are used to join words together in a sentence. It is commonly understood that stop words do not contribute to the context or content of textual documents. Due to their high frequency of occurrence, their presence in text mining presents an obstacle to the understanding of the content in the documents. To eliminate the bias effects, most text mining software or approaches make use of stop words list to identify and remove those words. However, the development of such top words list is difficult and inconsistent between textual sources. This problem is further aggravated by sources such as Twitter which are highly repetitive or similar in nature. In this paper, we will be examining the original work using term frequency, inverse document frequency and term adjacency for developing a stop words list for the Twitter data source. We propose a new technique using combinatorial values as an alternative measure to effectively list out stop words.**

Keywords- Stop words; Text mining; RAKE; ELFS; Twitter.

I. Introduction

Text mining comprises of a series of tasks that includes selection of approach, parameter setting and the creation of a stop word list [14][31]. The creation of a stop word list is often viewed as an essential component of the text mining which requires manual labor and investigations to produce. Stop words lists are rarely investigated and validated compared to the results of the mining process or mining algorithm. The lack of research into stop words list creation resulted in extensive use of pre-existing stop word lists which might not be suitable given the differences in the context of the textual sources. Research in the area has identified the weaknesses of standardized stop words list [3][4][23].

With the spread of social media platforms and adoption of such technologies in business and daily life, social media platforms have become one of the most important forms of communication for internet users and companies. Some companies are using Facebook and Twitter system to provide real time interaction with their customers. These social media platforms are beneficial to companies building consumer brand equity [12]. The platforms also act as low cost effective measures to manage complex relations between companies and consumers. The nature of social media also promotes open and transparent resolution of disputes and allows for greater visibility of the disputes to the senior management. Social Media has also proven to be very effective in communicating news such as the occurrence of earthquakes [25][9] and political office election [21][28].

The enormous amount of textual information from Twitter and social media requires extensive amount of data preparation and analysis to reap any benefits. There are many approaches to analyze the data. However, due to the nature and assumptions of the techniques as well as the huge amount of data collected, the data quality has to be of a very high level of quality in order to be effective [5][13][27]. To improve the quality of textual data, many authors have proposed different techniques to extract an effective stop word list for a particular corpus [22][29]. In the next section, we will focus on the common approaches to the development of stop words list.

II. Current Approaches

A stop words list refers a set of terms or words that have no inherent useful information. Stop words create problems in identification of key concepts and words from textual sources when they are not removed due to their overwhelming presence both in terms of frequency as well as occurrence in textual sources. Several authors [30][24][17] have argued for the removal of stop words which make the selection of the useful terms more efficient and reduce the complexity of the term structure. The current literature divides the stop words into explicit stop words and implicit stop words.

The common approach is to manually assemble a stop words list from a list of words. This approach is used by several authors [10] and has proven to be generally applicable to a variety of situation [17]. Even though the generic stop words lists generally achieved high accuracies and robust in nature, customized stop words lists occasionally outperforms especially in technical areas. These customized stop words lists were developed based on the entropy lists or unions of the standard stop lists with entropy lists mixed in [23]. Other authors held the opinion that any words that appear too rarely or were longer than a certain length should be removed [16].

There have been other attempts to use a variety of frequency measures such as term frequency, document frequency or inverse document frequency [15][18]. Each of these measures has proven to be effective in extracting the most common words that appear in the documents. The combination of term frequency with inverse document frequency (TF-IDF) measure was widely quoted by text books and papers [15][29] as the most popular implicit approach for creating a stop words list. There were also attempts in using Entropy approach to calculate the probability of a word being a stop word. [32] In non Anglophone languages, there have successes in using weight Chi Square method in classifying stop words. [33] In Rose et. Al. (2010), the authors proposed a new measure called the adjacency measure to establish whether a particular word is a stop word or a content word. In the next section, we will examine the algorithm described by Rose et. Al.

III. Rapid Automatic Keyword Extraction Stop word list

In the paper "Automatic keyword extraction from individual documents" by Rose et. Al., the authors describe a process to determine the usefulness of that word in describing the contents. Every word is identified and the word co-occurrences are calculated with a score is calculated for each word. Several scoring techniques based on the degree and frequencies of words were evaluated in the paper. In the paper, Adjacency frequency is defined as the number of times the word occurred adjacent to keywords. Keyword frequency is defined as the number of times the word occurred within keywords. The authors noted that selection by term frequency will increase the likelihood of content-bearing words to be added to the stop words list for a specialized topic that result in removal of critical information words. Rose et. Al. describes the adjacency algorithm as 'intuitive' for words that are adjacent to keywords are less likely to be useful than those that are in it. The authors subsequently tested the algorithm using several standardized documents and found the algorithm to be very effective.

However, there are several issues with the use of the adjacency measure.

1) Adjacency measure first assumes the presence of a keyword in which we can use to determine words that are adjacent. This results in the technique being usable only in the case where keywords are specified. In most textual sources, keywords are not available. In the case of Twitter, while you can use query keywords, it may not be useful for general trend extraction from tweets.

2) Adjacent words might be descriptive words which cannot be found within the keywords. In this case, the measure punishes these words.

3) Adjacency measures assumes multiple keywords in order for the between keywords to be found. This is an unlikely situation given that keywords are likely to single words. This makes it very difficult to be applied to Twitter or documents where the keywords are single words.

Given the restrictive nature of the RAKE stop words list generator, it is very difficult to apply the algorithm to a wide spectrum of text mining problems. In the next section, we will extend on the ideas given in Rose et. al. (2010) and present an effective algorithm in listing functional stop words using the combinatorial counts as measure of information value.

IV. Effective Listings of Functional Stop words using combinatorial counts

The authors noted that while the adjacency-within factor cannot be easily computed, the combinatorial factor can be computed easily. The combinatorial factor is defined as the number of unique word combination that can be found in the collection of tweets given a start word. The mathematical form is expressed below.

$$\mathrm{TCF} = \sum_{i=1}^{n} f\left(w_{p,n}, w_{p+1,n}\right) \qquad (1)$$

Where n is the number of tweets, p is the position of the word and w_p is the word in the position p. The function f is the indicator function with the following behavior.

$$f = \begin{cases} 1, \text{where} w_p = w \\ 0, \text{where} w_p \neq w \end{cases} \qquad (2)$$

Where w is the word that is being investigated.

The measure is computationally simple and implementable in a variety of programming languages natively. The combinatorial nature of the measure may not be intuitive. Any words can be linked by a number of words in a language to form meaning combinations. Words designed to convey a precise meaning needs to be linked up in a particular combination for the correct meaning to be conveyed. However, words which are commonly used as bridges in sentences will naturally accumulate a large number of combinations in any collection of documents or tweets. If the collection contains a strong theme or event, the words related will have smaller combinations of words. Theoretically, if there are certain words which are important, the number of combinations should only be one. For example, in any discussion about Linear Algebra, many of the technical terms used will naturally have little variations such as 'Linear Models', 'Complement Set'. This is in contrast to words such as 'in the' and 'that is'.

This measure is an alternative approach to the classical techniques of term-frequency and inverse-document frequency. This approach measures the information value of the word not through the conventional Kullback – Leibler framework but through the combinatorial nature of words. As opposed to measuring the information value of words to establish the stop words, the technique focuses on the extreme number of combinations that most non-meaningful words display to establish stop words. Moreover, the use of combinations allow us to naturally manage both words with high and low occurring frequency which presents a problem for the classical framework of TF*IDF without using transformation.

V. Experimental Setup

To validate the prowess of the measure, we conducted experiments with several techniques commonly used in development of stop word list. For all the experiments conducted, we have selected 9 3-days periods containing tweets with the key word search of 'Earthquake'. Each of this period starts 24 hours before the beginning of an earthquake and last till 48 hours after the occurrence of the earthquake. The reason for selecting 9 different periods and earthquakes is to ensure that the experiments will be as unbiased as possible. The use of query based tweets is to ensure that we have some form of central themes which provides some kind of comparison for the words which are not useful or meaningful. This two conditions enable us to assess the overall performance for the techniques tested effectively and unbiased.

The control factor for this experiment is the Fox's and Manu's stop word list. The choice of having two stop word lists is to double validate the techniques as both stop word lists are commonly used for text mining purposes. At the same time, both stop word lists have different words which can be useful as a further comparison between the efficacies of the techniques. All the words found in both stop word lists are determined to be stop words in the tweets through human examinations of the tweets using random samples of 1000 unique tweets from each period. For the classical techniques such as term frequency and inverse document frequency, we varied the cutoff thresholds before determining the optimal threshold by calculating the precision of the generated list with the stop list for different range of values. In total, we generated about 10 lists per technique.

Once we have generated the lists, we then compare the list across the different levels of threshold in increasing level of liberty in allowing the word to be considered stop word. Both precision and recall are calculated together with F-measure by comparing the list with the control stop word lists. The technique which consistently outperformed the other techniques will be considered to be the most effective stop word list generator.

VI. Results and Analysis

Using the experimental approach described above, we have generated the various stop words lists and compared their performance at detecting stop words which are listed in the Fox's and Manu's list. In the following sections, we will first compare the various measures and their performance with the Manu's list which is the smaller of the two lists. After the initial comparison, we will then further compare the results using the Fox's list for a second level of validation. The results are plotted with the F-Measures and the threshold levels.

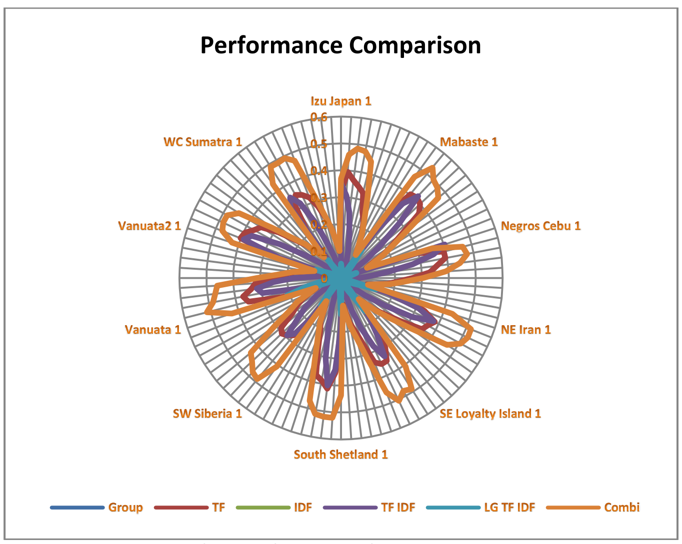


Chart 1: Comparison of the performance of the measures with the Manu's List using Radar Chart

From the chart 1, we can see that the combination technique outperforms most of the other techniques by a fair margin. With the exception of a few initial threshold, where TF*IDF or Log (TF)*IDF variant performs better, the new proposed approach is distinctly better than the other techniques. This superior performance could be attributed to the smaller list of stop words generated by combination approach compared to the other techniques. This effect is further compounded by the small list of stop words in the Manu instance. Many of the words included in the new stop word lists include new words which could be stop words in the context of the Twitter contents.

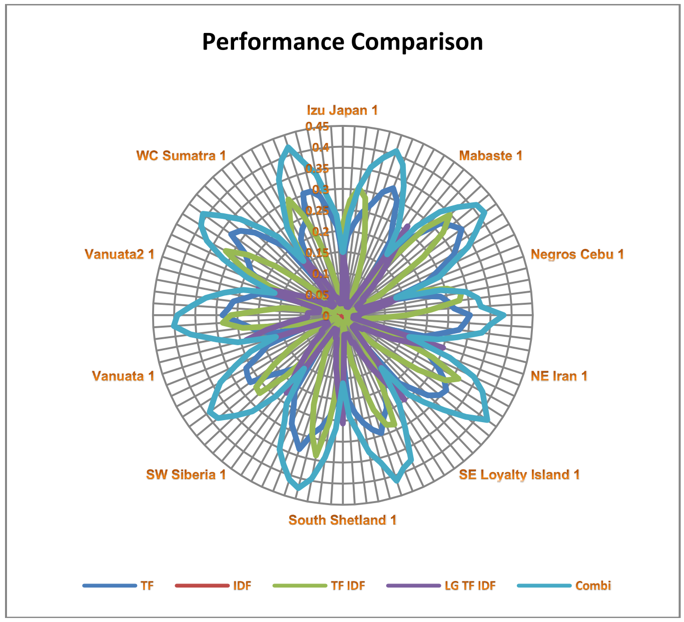


Chart 2: Comparison of the performance of the measures with the Fox's List using Radar Chart

From the chart 2, we can see that the combination technique outperforms most of the other techniques by a fair margin. However, the technique is not as strong as some of the other techniques in the initial threshold levels in some cases as evident in the breaks in the lines of the radar charts. The drop in performance could be attributed to the larger list of stop words covered by Fox's list which is almost three times the size of Manu's list. At the same time, as mentioned earlier, the stop word list generated by the combination technique is also smaller than its TF*IDF and variant counterparts. However, the combination technique still outperforms the other techniques beyond the initial threshold which indicates its superior performance on the overall.

VII. Conclusion

In this paper, we proposed a new method for automatically generating a stop word list for a given collection of tweets. The approach is based on the combinatorial nature of the words in speeches.

We investigated the effectiveness and robustness of the approach by testing it against 9 collections of tweets from different periods. The approach is also compared with the existing approaches using TD*IDF and variants. The results indicated that the new approach is comparable to existing approaches if not better in certain cases.

The direct nature of the combinatorial approach is not normalized and additional research is needed to produce the normalized measure. Other newer approaches such as page-rank approach will also require more research to understand the effectiveness. Future research will also need to investigate the scenario of three or more combinations of words to determine whether they are stop words.

VIII. ACKNOWLEDGEMENT

The author will like to thank the reviewers for their guidance and suggestions. The author will also like to express gratitude to the staffs of School of Information System, Singapore Management University for their guidance and support.

REFERENCES

[1] Busemann, S., Schmeier, S. and Arens, R. G. 2000. Message Classification in the Call Center, Proceedings of the Sixth conference on Applied Natural Language Processing, 158–165.

[2] Blake, C. 2010. Text Mining, ARIST, Vol 45

[3] Chakrabarti, S., Dom, B., Agrawal, R. and Raghavan, P. 1997. Using Taxonomy, Discriminants, and Signatures for Navigating in Text Databases, Proceedings of the 23rd International Conference on Very Large Databases, 446–455.

[4] Chakrabarti, S., Dom, B., Agrawal, R. and Raghavan, P. 1998. Scalabe Feature Selection, Classification and Signature Generation for Organizing Large Text Databases into Hierarchical Topic Taxonomies, The VLDB Journal, Springer-Verlag, 7, 163–178.

[5] Cooley, R., Mobasher, B., and Srivastava, J. 1999. Data Preparation for Mining World Wide Web Browsing Patterns, Knowledge Information System, 1-27.

[6] Corbin, J. and Strauss, A. 1990. Grounded Theory Research: Procedures, Canons, and Evaluative Criteria, Qualitative Sociology, 13(1), 3-21.

[7] Corney, M., de Vel, O., Anderson, A., and Mohay, G. 2002. Gender-preferential Text Mining of E-mail Discourse, The 18th annual Computer Security Applications Conference (ACSAC2002).

[8] de Vel, O., Corney, M. and Mohay, G. 2001. Mining E-Mail Content for Author Identification Forensics, SIGMOD Record, ACM Press, 30(4), 55–64.

[9] P. Earle. 2010, Earthquake Twitter. Nature Geoscience, 3:221.

[10] C. Fox. 1992, Lexical analysis and stoplists. In: Information Retrieval - Data Structures & Algorithms, p. 102-130. Prentice- Hall.

[11] Glaser, B., and Strauss, A. 1967. The Discovery of Grounded Theory, Chicago: Aldine Publishing Company.

[12] Jothi,P. et al. 2011. "Analysis of social networking sites: A study on effective communication strategy in developing brand communication", Journal of Media and Communication Studies Vol. 3(7), pp. 234-242, July 2011

[13] Jung, W. 2004. An Investigation of the Impact of Data Quality on Decision Performance, Proceedings of the 2004 International Symposium on Information and Communication Technology (ISICT '04), 166–171.

[14] Keogh, E., Lonardi, S. and Ratanamahatana, C. A. 2004. Towards Parameter-Free Data Mining, Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 206-215.

[15] Konchady, Manu. 2007, Text Mining Application programming. Charles River Media Publishing.

[16] Koprinska, I, Poon, J., Clark, J. and Chan, J. 2007. Learning to Classify Email, Information Science, 177, 2167–2187.

[17] Manco, G., Masciari, E., Ruffolo, M. and Tagarelli, A. 2002. Towards An Adaptive Mail Classifier, Proceedings of Italian Association for Artificial Intelligence Workshop.

[18] Mannings, D and Schuetze, H. 1999, Foundation of Statistical Natural Language Processing, MIT Press.

[19] Marwick, A. D. 2001. Knowledge Management Technology, IBM Systems Journal.

[20] Moreale, E. and Watt, S. 2002. Organisational Information Management and Knowledge Discovery in Email within Mailing Lists, In H. Yin et al. (Eds.), Intelligent Data Engineering and Automated Learning, Lecture Notes in Computer Science, 2412/2002, 217-224, Berlin / Heidelberg:Springer-Verlag.

[21] Mungui-Pippidi, Alina and Munteanu, Igor. "Moldova's 'Twitter Revolution.'" Journal of Democracy 20/3 (July 2009): 136-142.

[22] Salton, G. 1971. The SMART Retrieval System—Experiments in Automatic Document Processing, Upper Saddle River, NJ, USA: Prentice-Hall, Inc..

[23] Silva, C and Ribeiro, B. 2003. The Importance of Stop Word Removal on Recall Values in Text Categorization, Proceedings of the International Joint Conference on Neural Networks, 3, 1661-1666.

[24] Sinka, M. P., and Come D. W. 2003. Evolving Better Stoplists for Document Clustering and Web Intelligence, Proceedings of the 3rd Hybrid Intelligent Systems Conference, Australia, IOS Press.

[25] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors. In WWW '10: Proc. of the 19th international Conf. on World wide web, pages 851–860, New York, NY, USA, 2010. ACM.

[26] Tang, J., Li, H., Cao, Y. and Tang, Z. 2005. Email Data Cleaning, Proceedings of the eleventh ACM SIGKDD international conference on Knowledge Discovery in Data Mining, Chicago, Illinois, USA, 2005, 489–498.

[27] Tayi, G. K. and Ballou, D. P. 1998. Examining Data Quality, Communications of the ACM, ACM Press, 41(2), 54–57.

[28] Tumasjan, A.; Sprenger, T. O.; Sandner, P. G.; and Welpe, I. M. 2010. Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In Proc. 4Th Intl. AAAI Conf. on Weblogs and Social Media (ICWSM).

[29] Rose, S., D. Engel, N. Cramer, and W. Cowley. 2010. Automatic keyword extraction from individual documents. In M. W. Berry and J. Kogan (Eds.), Text Mining: Applications and Theory. John Wiley and Sons, Ltd.

[30] Van Rijsbergen, C. J. 1979. Information Retrieval, Newton, MA: Butterworth- Heinemann.

[31] Xu, R. and Wunsch, D. 2005. Survey of Clustering Algorithms, IEEE Transactions on Neural Networks, 16(3), 645-678.

[32] Z. Yao, and C. Ze-wen, "Research on the construction and filter method of stop-word list in text Preprocessing", Fourth International Conference on Intelligent Computation Technology and Automation, 2011.

[33] F. Zou, F. L. Wang, X. Deng, S. Han, and L. S. Wang, "Automatic Construction of Chinese Stop Word List", Proceedings of the 5th WSEAS International Conference on Applied Computer Science, Hangzhou, China, April 16-18, 2006 (pp1010-1015).

AUTHORS PROFILE

Murphy Choy is an instructor with School of Information System at Singapore Management University. He received his MSc finance from University College Dublin, Ireland. He has published papers in risk management, text analytics and operation research. His research interest is in the application of data mining and operation research to real life problem.

A Comparative Study of White Box, Black Box and Grey Box Testing Techniques

Mohd. Ehmer Khan
Department of Computer Science,
Singhania University,
Jhunjhunu, Rajasthan, India

Farmeena Khan
Department of Computer Science,
EIILM University,
Jorethang, Sikkim, India

***Abstract*—Software testing is the process to uncover requirement, design and coding errors in the program. It is used to identify the correctness, completeness, security and quality of software products against a specification. Software testing is the process used to measure the quality of developed computer software. It exhibits all mistakes, errors and flaws in the developed software. There are many approaches to software testing, but effective testing of complex product is essentially a process of investigation, not merely a matter of creating and following route procedure. It is not possible to find out all the errors in the program. This fundamental problem in testing thus throws an open question, as to what would be the strategy we should adopt for testing. In our paper, we have described and compared the three most prevalent and commonly used software testing techniques for detecting errors, they are: white box testing, black box testing and grey box testing.**

Keywords- Black Box; Grey Box; White Box.

I. INTRODUCTION

Software testing identifies defects, flaws or errors in the application code that must be fixed. We can also define software testing as a process of accessing the functionality and correctness of a software through analysis. The main purpose of testing can be quality assurance, reliability estimation, validation and verification. Software testing is a fundamental component of software quality assurance and represents a review of specification, design and coding. The main objective of software testing is to affirm the quality of software system by systematically testing the software in carefully controlled circumstances, another objective is to identify the completeness and correctness of the software, and finally it uncovers undiscovered errors. [1] [2]

The three most important techniques that are used for finding errors are: [1]

1) White Box Testing Technique: It is the detailed investigation of internal logic and structure of the code. In white box testing it is necessary for a tester to have full knowledge of source code.

2) Black Box Testing Technique: It is a technique of testing without having any knowledge of the internal working of the application. It only examines the fundamental aspects of the system and has no or little relevance with the internal logical structure of the system.

3) Grey Box Testing Technique: White box + Black box = Grey box, it is a technique to test the application with limited knowledge of the internal working of an application and also has the knowledge of fundamental aspects of the system.

II. WHITE BOX TESTING TECHNIQUE

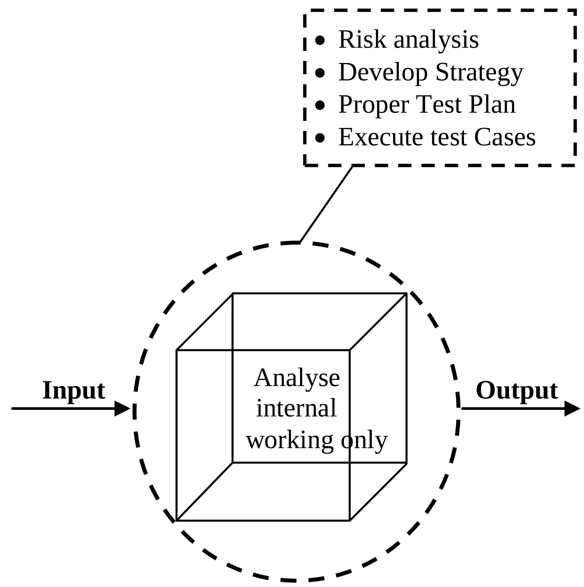


Figure 1. Represent white box testing

White box testing is a test case design method that uses the control structure of the procedural design to derive test cases. White box testing can uncover implementation errors such as poor key management by analyzing internal workings and structure of a piece of software. White box testing is applicable at integration, unit and system levels of the software testing process. In white box testing the tester needs to have a look inside the source code and find out which unit of code is behaving inappropriately. [3]

Some of the advantages and disadvantages of white box testing technique are listed below: [3] [4]

Advantages

- It reveals error in hidden code by removing extra lines of code.
- Side effects are beneficial.

- Maximum coverage is attained during test scenario writing.

Disadvantages

- It is very expensive as it requires a skilled tester to perform it.
- Many paths will remain untested as it is very difficult to look into every nook and corner to find out hidden errors.
- Some of the codes omitted in the code could be missed out.

Some of the synonyms of white box testing are glass box testing, clear box testing, open box testing, transparent box testing, structural testing, logic driven testing and design based testing.

Some important types of white box testing techniques are briefly described below: [3]

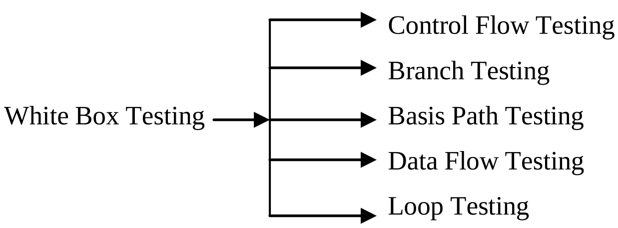


Figure 2. Represent different forms of white box testing techniques

1) Control Flow Testing: It is a structural testing strategy that uses the program control flow as a model control flow and favours more but simpler paths over fewer but complicated path.

2) Branch Testing: Branch testing has the objective to test every option (true or false) on every control statement which also includes compound decision.

3) Basis Path Testing: Basis path testing allows the test case designer to produce a logical complexity measure of procedural design and then uses this measure as an approach for outlining a basic set of execution paths.

4) Data Flow Testing: In this type of testing the control flow graph is annoted with the information about how the program variables are define and used.

5) Loop Testing: It exclusively focuses on the validity of loop construct.

III. BLACK BOX TESTING TECHNIQUE

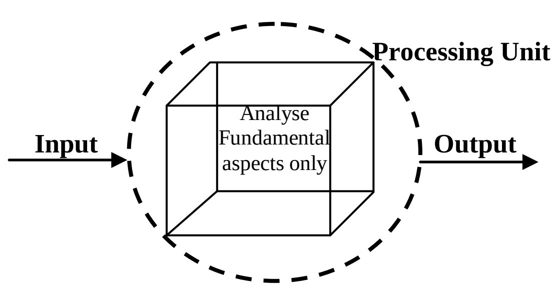


Figure 3. Represent black box testing

Black box testing treats the software as a “Black Box” – without any knowledge of internal working and it only examines the fundamental aspects of the system. While performing black box test, a tester must know the system architecture and will not have access to the source code. [5]

Some of the advantages and disadvantages of black box testing technique are listed below: [4] [5]

Advantages

- Efficient for large code segment.
- Tester perception is very simple.
- Users perspective are clearly separated from developers perspective (programmer and tester are independent of each other).
- Quicker test case development.

Disadvantages

- Only a selected number of test scenarios are actually performed. As a result, there is only limited coverage.
- Without clear specification test cases are difficult to design.
- Inefficient testing.

Some of the synonyms of black box testing technique are opaque testing, functional testing, close box testing, and behavioural testing.

Some important types of black box testing techniques are briefly described below: [5]

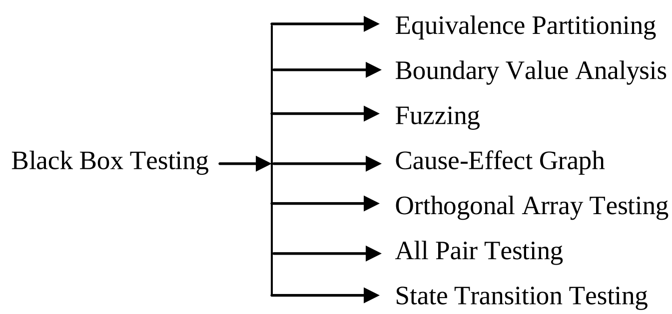


Figure 4. Represent differnet forms of black box testing techniques

1) Equivalence Partitioning: It can reduce the number of test cases, as it divides the input data of a software unit into partition of data from which test cases can be derived.

2) Boundary Value Analysis: It focuses more on testing at boundaries, or where the extreme boundary values are chosen. It includes minimum, maximum, just inside/outside boundaries, error values and typical values.

3) Fuzzing: Fuzz testing is used for finding implementation bugs, using malformed/semi-malformed data injection in an automated or semi-automated session.

4) Cause-Effect Graph: It is a testing technique, in which testing begins by creating a graph and establishing the relation between the effect and its causes. Identity, negation, logic OR and logic AND are the four basic symbols which expresses the interdependency between cause and effect.

5) Orthogonal Array Testing: OAT can be applied to problems in which the input domain is relatively small, but too large to accommodate exhaustive testing.

6) All Pair Testing: In all pair testing technique, test cases are designs to execute all possible discrete combinations of each pair of input parameters. Its main objective is to have a set of test cases that covers all the pairs.

7) State Transition Testing: This type of testing is useful for testing state machine and also for navigation of graphical user interface.

IV. GREY BOX TESTING TECHNIQUE

Grey box testing technique will increase the testing coverage by allowing us to focus on all the layers of any complex system through the combination of all existing white box and black box testing.

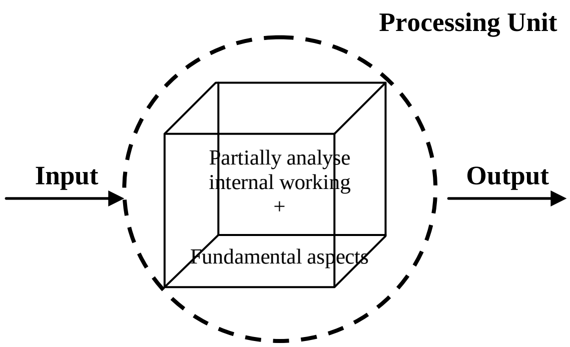


Figure 5. Represent grey box testing

In grey box testing the tester must have knowledge of internal data structures and algorithm, for the purpose of designing test cases. Examples of grey box testing technique are: [6]

- Architectural Model
- Unified Modeling language (UML)
- State Model (Finite State Machine)

In grey box testing the codes of two modules are studied (white box testing method) for the design of test cases and actual test are performed in the interfaces exposed (black box testing method).

Some of the advantages of grey box testing technique are listed below: [4] [6]

- Grey box testing provides combined benefits of white box and black box testing techniques.
- In grey box testing, the tester relies on interface definition and functional specification rather than source code.
- In grey box testing, the tester can design excellent test scenarios.
- The test is done from the user's point of view rather than designer's point of view.
- Create an intelligent test authoring.
- Unbiased testing.

Some of the disadvantages of grey box testing technique are listed below:

- Test coverage is limited as the access to source code is not available.
- It is difficult to associate defect identification in distributed applications.
- Many program paths remain untested.
- If the software designer has already run a test case, the tests can be redundant.

The other name of grey box testing is translucent testing. Different forms of grey box testing techniques are briefly described below: [6]

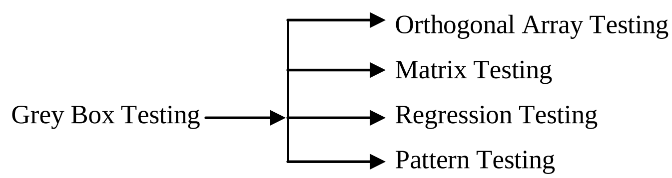


Figure 6. Represent different forms of grey box testing techniques

1) Orthogonal Array Testing: This type of testing use as subset of all possible combinations.

2) Matrix Testing: In matrix testing the status report of the project is stated.

3) Regression Testing: If new changes are made in software, regression testing implies running of test cases.

4) Pattern Testing: Pattern testing verifies the good application for its architecture and design.

V. FUTURE OF SOFTWARE TESTING

With the changing trends in the software industry, software testing too changes. The existing new technologies like Service Oriented Architecture (SOA), wireless technologies, mobile services etc. has opened new path to testing. Some of the changes which we will see in the industry over the next few years are listed below: [8]

- Testers will provide light weight models that developers can run against their codes.
- Early review and modeling will exposes many ambiguous bugs.
- As in the future developer's code is full of testability hooks, errors will be more detectable.
- Static analyser (detection tools) will come in main stream.
- Useful matrices such as spec coverage, model coverage and code coverage drives the projects.
- Combinatorial tools will allow testers to prioritize their testing.
- The testers will provide visible and value added services throughout the software development process.
- Tester can develop test harnesses stubs and drivers written in and interacting with a variety of programmatic languages.

- Tomorrows' tester will be professionally more educated, examine and accredited professional.

TABLE I. COMPARISON BETWEEN THREE FORMS OF TESTING TECHNIQUES [6] [7]

S. No.	Black Box Testing	Grey Box Testing	White Box Testing
1.	Analyses fundamental aspects only i.e. no proved edge of internal working	Partial knowledge of internal working	Full knowledge of internal working
2.	Granularity is low	Granularity is medium	Granularity is high
3.	Performed by end users and also by tester and developers (user acceptance testing)	Performed by end users and also by tester and developers (user acceptance testing)	It is performed by developers and testers
4.	Testing is based on external exceptions – internal behaviour of the program is ignored	Test design is based on high level database diagrams, data flow diagrams, internal states, knowledge of algorithm and architecture	Internal are fully known
5.	It is least exhaustive and time consuming	It is somewhere in between	Potentially most exhaustive and time consuming
6.	It can test only by trial and error method	Data domains and internal boundaries can be tested and over flow, if known	Test better: data domains and internal boundaries
7.	Not suited for algorithm testing	Not suited for algorithm testing	It is suited for algorithm testing (suited for all)

In the near future we will see a shift towards new techniques and testing transformed business operations, the way people interact with the systems and information it provides, and therefore mitigating the risk and increasing the benefits of business change.

VI. CONCLUSION

We can define software testing as an activity aimed at evaluating an attribute, or capability of a program to determine, that it meets its required specification. Software testing can provide an independent view of the software to allow the business to appreciate and understand the risk of software implementation.

To carry out software testing in a more effective manner, in our paper we have described and compared three main software testing techniques.

REFERENCES

[1] Mohd. Ehmer Khan, "Different Forms of Software Testing Techniques for Finding Errors," IJCSI, Vol. 7, Issue 3, No 1, pp 11-16, May 2010

[2] Software Testing Methodologies by Azhar available at http://azhar-paperpresentation.blogspot.com/2010/04/software-testing-methodologies.html

[3] Mohd. Ehmer Khan, "Different Approaches to White Box testing Technique for Finding Errors," IJSEIA, Vol. 5, No. 3, pp 1-13, July 2011

[4] Software Testing Methods available at http://www.tutorialspoint.com/software_testing/testing_methods.htm

[5] Mohd. Ehmer Khan, "Different Approaches to Black Box Testing Technique for Finding Errors," IJSEA, Vol. 2, No. 4, pp 31-40, October 2011

[6] Grey Box Testing from Wikipedia available at http://en.wikipedia.org/wiki/Gray_box_testing

[7] Comparison among Black Box Testing, Gray Box Testing & White Box Testing available at http://www.softwaretestinggenius.com/articalDetails?qry=404

[8] Predicting the Future of Testing by Harry Robinson available at http://www.stickyminds.com/sitewide.asp?Function=edetail&ObjectType=COL&ObjectId=6887

AUTHORS PROFILE

Mohd. Ehmer Khan

I completed my B.Sc in 1997 and M.C.A. in 2001 from Aligarh Muslim University, Aligarh, India, and pursuing Ph.D (Computer Science) from Singhania University, Jhunjhunu, India. I have worked as a lecturer at Aligarh College Engineering & Management, Aligarh, India from 1999 to 2003. From 2003 to 2005 worked as a lecturer at Institute of Foreign Trade & Management, Moradabad, India. From 2006 to present working as a lecturer in the Department of Information Technology, Al Musanna College of Technology, Ministry of Manpower, Sultanate of Oman. I am recipient of PG Merit Scholarship in MCA. My research area is software engineering with special interest in driving and monitoring program executions to find bugs, using various software testing techniques.

Farmeena Khan

I completed my BCA in 2003 from Dr. B. R. Ambedkar University, Agra, India and MBA in 2011 from Janardan Rai Nagar Rajasthan Vidyapeeth University, Rajasthan, India, and pursuing M.Sc (Computer Science) from EIILM University, Jorethang, Sikkim, India.

A Study On OFDM In Mobile Ad Hoc Network

Malik Nasereldin Ahmed
Computer Science & Information system
University Technology Malaysia
Johor Bahru - Malaysia

Abdul Hanan Abdullah
Computer Science & Information system
University Technology Malaysia
Johor Bahru - Malaysia

Satria Mandala
Computer Science & Information system
The state of Islamic University of
MALIKI–Indonesia

***Abstract*— Orthogonal Frequency Division Multiplexing (OFDM) is the physical layer in emerging wireless local area networks that are also being targeted for ad hoc networking. OFDM can be also exploited in ad hoc networks to improve the energy performance of mobile devices. It is important in wireless networks because it can be used adaptively in a dynamically changing channel. This study gives a detailed view about OFDM and how it is useful to increase the bandwidth. This paper also gives an idea about how OFDM can be a promising technology for high capacity wireless communication.**

Keywords- Ad hoc; OFDM; MANET.

I. INTRODUCTION

A. Ad hoc Networks

Ad hoc networks are a new paradigm of wireless communication for mobile hosts where node mobility causes frequent changes in topology. Mobile nodes that are within each other's radio range communicate directly via wireless links called peer to peer or single hop ad hoc network. If destination mobile node is out of range of the source mobile node, other nodes in the range of the source and destination acts as router to transmit packets between source and destination and this is called multi – hop ad hoc networks. The figures 1 and 2 show single and multi-hop ad hoc network. Ad hoc networks are self-configurable and autonomous systems consisting of routers and hosts, which are able to support mobility and organize themselves arbitrarily. Moreover, the ad hoc network can be either constructed or destructed quickly and autonomously without any administrative server or infrastructure. If the nodes of ad hoc networks are mobile and with wireless communication to maintain the connectivity, it is known as mobile ad hoc network (MANET) and require an extremely flexible technology for establishing communications in situations which demand a fully decentralized network without any fixed base stations, such as battle fields, military applications and other emergency and disaster situations.

There are some challenges in mobile environments like limitations of the wireless network, packet loss due to transmission errors, variable capacity links, frequent disconnections / partitions, limited communication bandwidth and broadcast nature of the communications. In addition, limitation imposed by mobility dynamically changing topologies / routers lacks of mobility awareness by system / applications. Limitations of the mobile computer such as short battery lifetime and limited capacities will create more problems for the transmission.

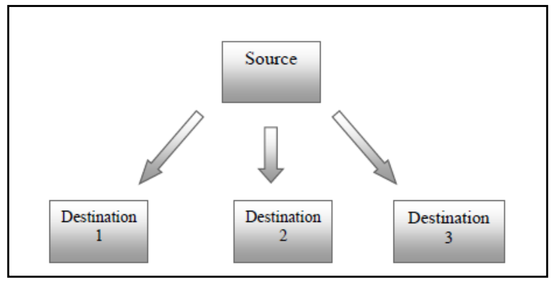


Figure 1. Single hop ad hoc network

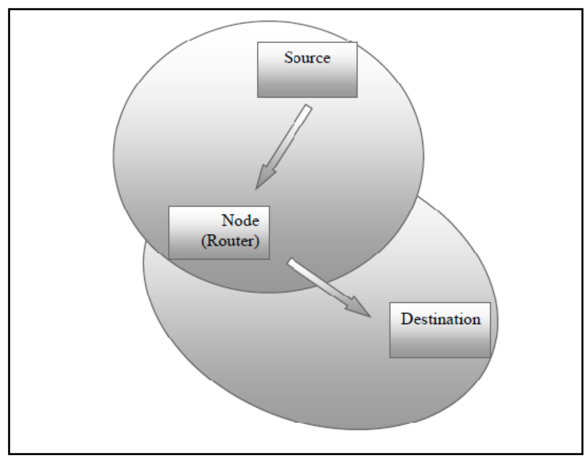


Figure 2. Multi hop ad hoc network

Ad hoc network has some problems include routing, broadcasting, multicasting, geo casting, clustering, area monitoring, data and service access. Topology control problems include neighbor discovery, topology construction and maintenance, activity scheduling, position discovery, partition detection, location updates, and network management. In addition, there are two fundamental problems network information capacity and minimum energy per bit.

Multiplexing or Muxing, is the process where multiple analog or digital signals combined in one signal over shared medium. For example, it is possible to transfer many phone calls in one signal at the same time. At the receiver side, the combined signals must be separate into the original signal. This process is called Demultiplexing.

Multiplexing techniques may be divided into many different types and they are space division multiplexing

(SDM), frequency division multiplexing (FDM), time division multiplexing (TDM) and code division multiplexing (CDM).

A. OFDM

OFDM is a modulation technique as well as multiplexing technique as it is divide a single high data rate stream into a number of lower rate streams that are data transmitted simultaneous over some narrow sub channel. OFDM technique offering better use of ad hoc network like inter-symbol interference (ISI), OFDM avoids this problem by sending many low speed transmissions simultaneously and increase the network throughput.[1][2][3]

In OFDM, the sub carrier frequencies are chosen so that the sub carriers are orthogonal to each other. This greatly simplifies the design of both the transmitter and the receiver. Unlike conventional Frequency Division Multiplexing (FDM), a separate filter for each sub channel is not required. The orthogonality allows high spectral efficiency with a total symbol rate near the Nyquist rate for the equivalent baseband signal. Almost the whole available frequency band can be utilized. [4]

In conventional OFDM systems most of the approaches to combat ICI are towards using frequency synchronization and interference cancellation. They are usually very complex and sometimes there is loss in bandwidth efficiency OFDM system create high data rate with long symbol duration by combining low data rate. It reduces equalization complexity by implementing with Inverse Fast Fourier Transform (IFFT) at the transmitter and Fast Fourier Transform (FFT) at the receiver that converts the wide band signal affected by frequency selective fading into N narrowband flat fading signals.[8][9] The beneficial since OFDM enables support of more antennas and large band widths since it simplifies equalization dramatically in MIMO system. Hence, the available bandwidth is utilized very efficiently in OFDM systems. The figure 3 shows the physical model of the design [10][11][12] .

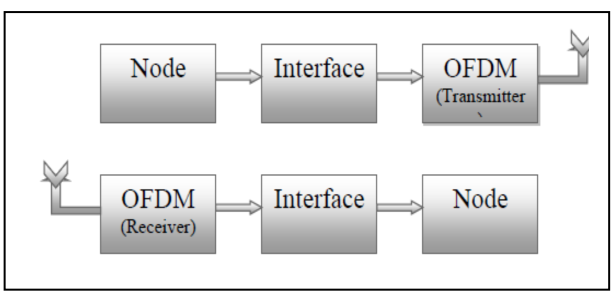


Figure 3. Physical Model

II. REVIEW OF THE LITERATURE

Dandan wang and Uf tureli[7] proposed a new transceiver architecture with MIMO – OFDM in the physical layer and an adaptive multiple antennas receiver initiated busy tone medium access to improve the performance of ad hoc networks. They have presented both theoretical and numerical analysis of the throughput and delay.

Gavin yeng et.al.[6] presented a detailed OFDM and channel modeling on the performance evaluation of higher layer protocols. The result of the model given by the author, show that device and wireless channel can impact packet delivery ratios and even point out a deficiency of the auto rate fall back protocol.

Swati chowduri et.al.[5] proposed a model which implements a MIMO- OFDM for MIMO based mobile ad hoc network and compared its bit error rate with different modulation technique. They have described different approaches for channel estimation and data detection. They have evaluated the performance of OFDM system using phase shift keying and quadrature amplitude modulation with OFDM using generalized gamma fading distribution.

Kamol kaemarungsi and prashant Krishnamurthy[4] studies shows that adaptive OFDM can be also exploited in ad hoc networks to improve the energy performance of mobile devices. They have evaluated the improvement in performance of adaptive OFDM over non – adaptive OFDM in ad hoc networks using simulations. Their study shows that how much of energy can be preserved by employing adaptive OFDM on the physical layer and how the channel information gain from the loading algorithm will help improve the energy preservation of ad hoc wireless network.

In this paper we have done a simulation to show that the OFDM got the characteristics to improve the bandwidth in the MANET. Moreover the study reveals the advantages of OFDM over other multiplexing technique.

III. SIMULATION AND RESULTS

OFDM is a physical layer encoding technology for transmitting signals through the RF. This method breaks one high – speed data carrier into several lower speed carriers. These are in turn transmitted in parallel across that particular RF spectrum. The 802.11a subcommittee has elected to use the transmission technique for their standard in the 5 GHz unlicensed national information infrastructure bands. OFDM also used in the 2.4 GHz ISM band as the physical layer standard for the 802.11g standard. The simulation design contains three parts. The figure 3 shows the detailed view of the design.

Node: It is the receiving end; it may be a laptop, desktop or a mobile. In the MATLAB 'Signal generator' selected to act as a node and it generates continuous square pulse at 1 amplitude and 1Hz frequency.

Interface: Discrete input signal at 0.5 intervals.

OFDM: This block contains three units. P / S to convert signals from parallel to serial; Inverse Fast Fourier Transform (IFFT) is useful for OFDM because it generates samples of a waveform with frequency components satisfying orthogonally conditions. The following Fourier transform shows the in-phase and quadrature components.

$$s(t) = \sum_{n=0}^{N-1} [a(n)\cos(\omega_n t) + b(n)\sin(\omega_n t)] \quad (1)$$

a(n) – is a real sequence representing the in-phase component.

b(n) – is a real sequence representing the quadrature component.

Fourier Fast Transform (FFT) : A FFT is an efficient algorithm to compute the Discrete Fourier Transform (DFT) and its inverse. There are many distinct FFT algorithms involving a wide range of mathematics from simple complex – number arithmetic to group theory and number theory.

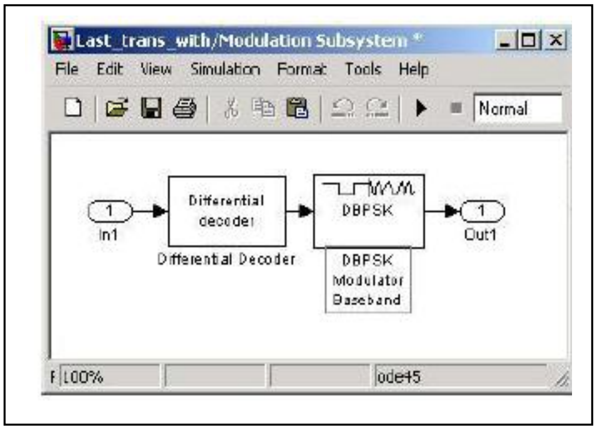


Figure 4. Modulation and Decoder Subsystem

A. System Transmitter

Transmission system generates digital signals and it will be quantized and enter into the OFDM system. After that the signal sends through radio wave. In the simulation, we have proved that OFDM increases bandwidth. Bandwidth describes the signal media carrier and the extent of use. This can be known by discovering the signal in the frequency domain and measuring the distance between the f 1and f2 at -3dB. The following figure5 shows the OFDM transmitter and figure 6 shows the normal increased bandwidth.

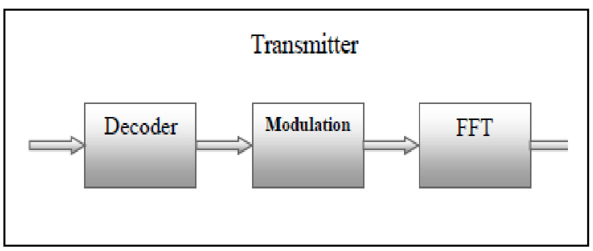


Figure. 5 OFDM Transmitter

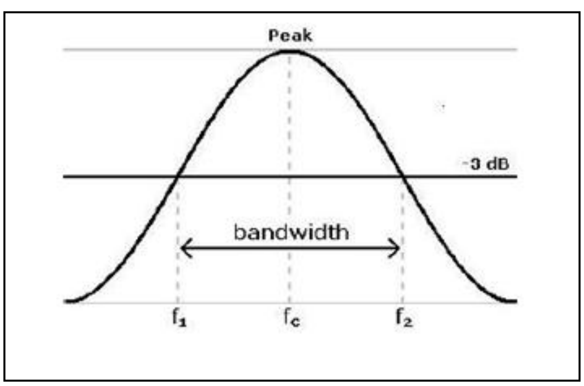


Figure 6. Normal Bandwidth

We have made this simulation to prove that the OFDM increases the bandwidth. First part of our simulation is transmitter without OFDM. The following figures show the transmitter without OFDM and output of the system.

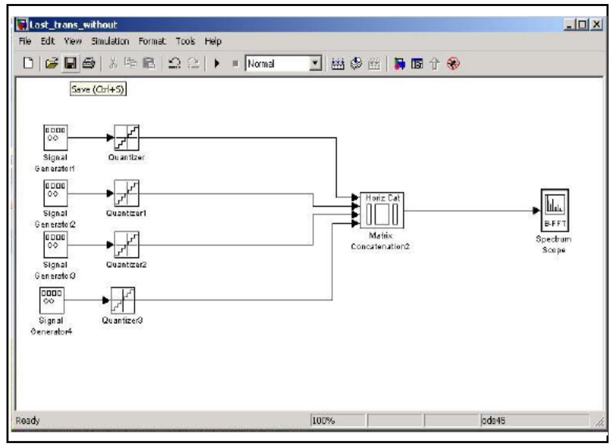


Figure 7. Transmitter without OFDM

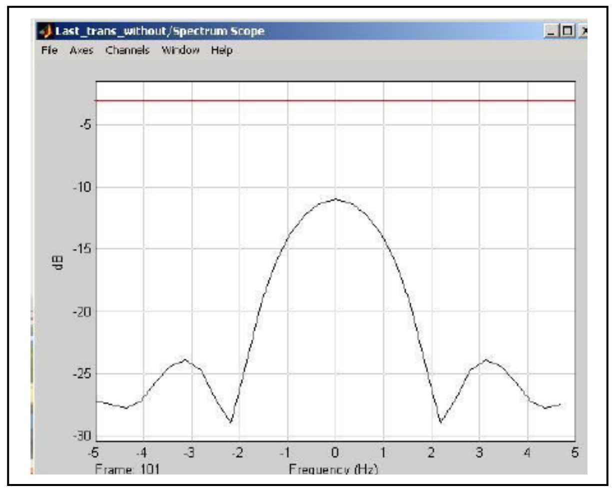


Figure 8. Output of the Transmitter without OFDM

The straight red line in the figure represents the -3dB. It is clear that bandwidth did not reach the maximum level. The same process with OFDM is shown below.

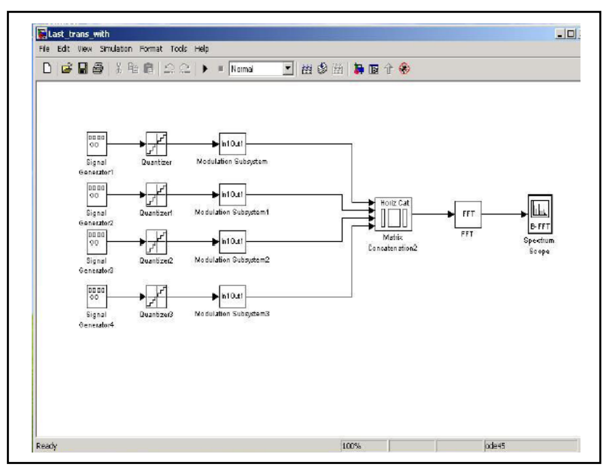


Figure 9. Transmitter with OFDM

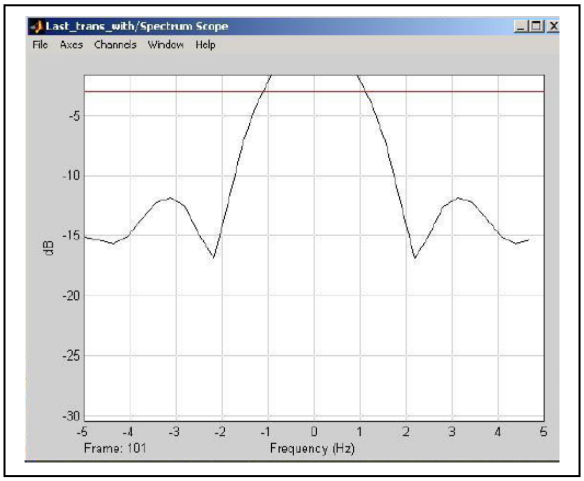


Figure 10. Output of the Transmitter with OFDM

The output shows clearly that the OFDM increases the bandwidth.

IV. Conclusion

This paper has presented on the effects of detailed OFDM and its advantages over other multiplexing techniques. The results show that the OFDM acts as an accelerator to increase the bandwidth.

Adaptive OFDM system has optimum throughput performance for low and high mobility conditions. OFDM is a promising technology suitable for high capacity wireless communications and it is computationally efficient by using FFT techniques to implement the modulation and demodulation functions.

Reference

[1] Shan Chu, Xin Wang, " Opportunistic and Co operative spatial multiplexing in MIMO Ad-hoc networks:, IEEE / ACM Trans. On Networking. Vol.18, no.5, Oct 2010.

[2] B.Sathish kumar, K.R.Shankar Kumar, R.Radhakrishnan, " An efficient inter carrier interference cancellation schemes for OFDM systems", IJCSIS Vol.6, No.3, 2009.

[3] D.K.Sharma, A.Mishra and Rajiv Savena, "Analog & Digital Modulation Technique: An overview", International Journal of Computer Science and Communication technologies, Vol 3, No.1, July 2010.

[4] Kamol Kaemarungsi, Prashant Krishnamurthy," On the use of Adaptive OFDM to preserve energy in Ad-hoc wireless networks", Proceedings of 13th MPRG, Page(s): 1 -4.

[5] Swati Chowdhuri, Dr.Arun kumar Mondal, Dr.P.K.Baneerjee," Implementation of MIMO – OFDM system in mobile Ad-Hoc networks", International journal of Engineering Research and Applicaition:, Vol.1, Iss.3, Page(s): 753 – 757.

[6] Gavin Yeung, Mineo Takai, Rajive Bagrodia, Alireza Mehrnia, Babak Daneshrad,"Detailed OFDM modeling in network simulation of mobile Ad-Hoc networks.

[7] Dandan Wang, UF Tureli, " Joint MIMO – OFDM and MAC design for broadband multi hop Ad Hoc networks", DRAFT, Nov. 1, 2005.

[8] P.H.Moose, " A technique for orthogonal frequency division multiplexing frequency offset correction", IEEE transactions on communications Vol 42, No.10, 1994.

[9] H.Yang, " A Road to future broadband wireless access: MIMO – OFDM based Air Interface", Communications Magazine, Vol.43, No.1, Jan 2005, Page(s): 53 – 60.

[10] R.Rom and M.Sidi, " Multiple Access Protocols Performance and Analysis", Springer Verilog 1990.

[11] IEEE standard for wireless LAN Medium Access Control (MAC) and physical layer specifications, Nov. 1997, P802.11.

[12] T.Tang, M.Park, R.W.Heath and S.M. Nettles, " A joint MIMO – OFDM transceiver and MAC design for mobile Ad hoc networking, International workshop on wireless ad – hoc networks 2004.

[13] Q.Zhao and L.Tang, "A multiqueue service room MAC protocol for wireless networks with multipacket reception", IEEE / ACM trans. Networking Vol.11, No.1, Feb 2003 Page(s):125 – 137.

[14] Charles.E.Perkins, Elizabeth.M.Royer, "Ad-hoc on demand distance vector routing", in proceedings of 2nd IEEE workshop on mobile computing systems and application. Page(s): 90 – 100, 1999.

[15] Gavin Holland, Nitin Vaidya, Paramvir Bahl, " A Rate Adaptive MAC protocol for multi hop wireless networks", in proceedings of mobicom 2001, July 2001.

[16] Mineo Takai, Jay Martin, Rajive Bagrodia, "Effects of wireless physical layer modeling in mobile Ad-hoc network", in proceedings of mobihoc 2001, page(s): 87 – 94, Oct 2001.

Evaluation of Spectral Efficiency, System Capacity And Interference Effects on CDMA Communication System

Ifeagwu E.N.[1], Ekeh j[2], Ohaneme C.O.[1], Okezie C.C. [1]
[1]Department of Electronics and Computer Engineering, Nnamdi Azikiwe University, Awka
[2]Department of Electrical/ Electronic Engineering, Enugu State University of Technology,

***Abstract*— Wireless communication technology have been developed based on exploring new mobile communications frequency bands, reasonable use of frequency resources and minimization, portability and multifunction's of mobile stations. The technology of wireless communications with duplex transmission is one of the fastest expanding in the world today. A very effective solution in achieving network performance in terms of system capacity and spectral efficiency with respect to the CDMA wireless network is presented in this paper. The effect of how interference grows as the number of users increases is analyzed. Equally, this paper presents the major problems in CDMA as the multiple access interference which arises due to the deviation of the spreading codes from orthogonality.**

Keywords- Bandwidth; Capacity; Interference; Pseudo-noise.

I. INTRODUCTION

Wireless communication technologies have basically been developed based on exploring new mobile communication frequency bands, reasonable use of frequency resources and minimization, portability and multi-functions of mobile stations. The concept of cellularization sprung up as the demand for mobile communication increased [1].

It seeks to make an efficient use of available channels by employing low-power transmitters so that within certain distances, the same frequency could be reused without dcisrupting the integrity of the conversation taking place on a channel.

Cellular services are now being used every day by millions of people worldwide. The major objective of a cellular system is to be able to handle as many calls as possible in a given bandwidth with the specified blocking probability (reliability) [2]. An efficient use of the available spectrum is achieved by finding ways to allow multiple users to share the available spectrum simultaneously. The Code Division Multiple Access (CDMA) was designed to achieve this feet. CDMA is a form of multiplexing and a method of multiple access that divides up a radio channel not by time nor frequency, but instead by using different pseudo-random code sequence for each user such that the transmitting end modulates the signals that it sends using the pseudo-random codes, and the receiving end detects the corresponding signals by demodulating the mixed signals using the same pseudo-random code [3]. The pseudo-random code is a periodic binary sequence with a noise-like waveform. It is also called pseudo-noise codes and is used to differentiate subscribers. The codes are shared by both the mobile station (MS) and the base station (BS). As a crude analogy to the CDMA scheme, imagine a large room (channel) containing many people wishing to communicate amongst each other. Possible mechanisms for avoiding confusion include: having only one person speak at a time (time division) [4], having people speak at different pitches (frequency division), or in different directions (spatial division). The CDMA approach is to have them speak different languages to each other. Groups of people speaking the same language can understand each other, but not any of the other groups. Similarly, in CDMA, each group of users is given a shared code. There are many codes occupying the same channel, but only the users associated with a particular code can understand each other. We have several modes of CDMA. They are;

1) *Direct Spectrum CDMA (DS-CDMA).*
2) *Frequency Hopping CDMA (FH-CDMA).*
3) *Multi-carrier CDMA (MC-CDMA)*
4) *IS-95 CDMA*

Code Division Multiple Access (CDMA) relies on the use of spread spectrum techniques to achieve multiple communication channels in a designated segment of the electromagnetic spectrum. With CDMA, each user's narrowband signal is modulated by a high rate special code (pseudo-random binary sequence). This causes the spreading of the bandwidth of the user's signal resulting in a wideband signal. A large number of CDMA users share the same frequency spectrum. In the receivers, the signals are correlated by the appropriate pseudo-random code which de-spreads the spectrum. The other users' signals whose codes do not match are not de-spread and therefore only appear as noise and represent a self-interference generated by the system. The signal-to-interference ratio for CDMA is determined by the ratio of the desired signal power to the total interference power from all the other users.

Capacity of cellular systems is of major concern to designers due to its economic value. For any multiuser communication system, the measure of its economic usefulness is not the maximum number of users which can be serviced at one time, but rather the peak load that can be supported with a given quality and with availability of service [5]. The radio channel capacity of a CDMA cellular system is defined as the number of subscribers that can access a particular base station of the system simultaneously. The capacity of CDMA is therefore only limited to the amount of interference that can be tolerated from other users [2].

II. CDMA Capacity Derivation

Using a unique pseudo-random noise (PN) code, a CDMA transmitting station spreads the signal in a bandwidth wider than actually needed. Each authorized receiving station must have the identical PN code to retrieve the information. Other channels may operate simultaneously within the same frequency spectrum as long as different, orthogonal codes are used. The long PN code is called a ***chip***. The chip is modulated by the information data stream. The ratio of the chip rate Rc to the information data rate Rb is called the processing gain G:

$$G = \frac{R_c}{R_b} \qquad (1)$$

Because the CDMA carrier bandwidth,

BT= R_c;

The processing gain, G can also be expressed as

$$G = \frac{B_T}{R_b} \qquad (2)$$

Then, if Eb is the average bit energy of any channel's signal, the power spectral density of the transmit signal (energy per chip) is:

$$\frac{E_b R_b}{R_c} = \frac{E_b}{G} \qquad (3)$$

In typical CDMA technology, when speech pauses, the transmit signal is switched off to save power. The voice activity state of channel n is denoted by $\alpha_n \in \{0;1\}$ where {0;1} representing both the on and off states. The expected value is $E\{\alpha_n\} = \alpha$. Instead of switching off the transmit power, real systems transmit at a low power and data rate during voice inactivity. Therefore α can also be considered as the average transmit power divided by the transmit power during talk spurts. Let us define $\bar{I}_i$ to be the mean value of the power spectral density of the interference noise caused by other channels. Because there are N channels in total, there are N - 1 channels other than the channel under consideration. Based on (3),

$$\bar{I} = \alpha(N-1)E_b / G \qquad (4)$$

If N_O stands for thermal noise of the system, then the mean total noise power spectral density I_{tot} is the sum of the interference noise and the thermal noise.

$$I_{tot} = \bar{I} + N_o = \alpha(N-1)E_b / G + N_O \qquad (5)$$

We can then derive the number of total CDMA channels N from (5) to be

$$N = 1 + \frac{G}{\alpha}\left[\left(\frac{E_b}{I_{tot}}\right)^{-1} - \left(\frac{E_b}{N_O}\right)^{-1}\right] \qquad (6)$$

From [7], Claude Shannon's work relates the amount of information carried, channel bandwidth, Signal-to-noise ratio, and detection error probability; it shows the theoretical upper limit attainable.

Shannon's capacity equation is given as

$$C = B_w \log2\,[1+S/N] \qquad (7)$$

Where B_w= Bandwidth in Hertz, C= Channel capacity in bits/second, S/N= Signal-to-noise ratio.

Capacity is determined by the balance between the required SNR for each user, and the spread spectrum processing gain. The figure of merit of a well-designed digital receiver is the dimensionless signal-to-noise ratio (SNR).

$$S/R = \frac{S}{(N-1)S + \eta} \qquad (8)$$

where, S: power of the received signal from each user, N: number of users in the cell, η: Background noise. The SIR could be increased (interference reduced) by controlling the terms in the denominator of (8).

Eb/No (Energy-per-bit-to-Noise density ratio) could be expressed by:

$$\frac{E_b}{N_o} = \frac{S/R}{(N-1)S/W + \eta/W} \qquad (9)$$

where, E_b: Energy per bit seen at the receiver,

W: spreading bandwidth,

R: information bit rate,

N_o: Noise power spectral density.

Following from (9), the maximum number of users in the cell (capacity) could be expressed as:

$$N = 1 + \frac{W/R}{E_b/N_o} - \frac{\eta}{S} \qquad (10)$$

Where W/R is the spread bandwidth-to-data rate ratio, sometimes called processing gain. This shows that the number of users depends on the SNR required by the BS for adequate performance. The SNR requirement could be translated into an interference level threshold for an SIR admission policy.

For a multi-cell system, the base station suffers from inter-cell interference in addition to the intra-cell interference. The effect of this interference on the SNR seen at the BS under consideration (BS_o) could be observed by simply adding an inter-cell interference term to (9).

$$\frac{E_b}{N_o} = \frac{S/_R}{(N-1)^S/_W + {}^I/_W + {}^\eta/_W} \quad (11)$$

Where, I: the total inter-cell interference seen by BS_o receiver from users in neighbouring cells who are not controlled by BS_o.

III. Interference

The type of interference experienced in CDMA communication systems is the Multiple Access Interference or the Multi-user Interference. Multiple Access Interference (MAI) is a factor, which limits the capacity and performance of CDMA system. In contrast to FDMA and TDMA techniques which are frequency bandwidth limited, all users transmit in the same frequency band and are distinguished at the receiver by the user specific spreading code. All other signals are not de-spread because they use different codes. These signals appear as interference to the desired user. As the number of users increase, the signal to interference ratio (SIR) decrease until the resulting performance is no longer acceptable. Thus, this multi–user interference must be reduced to achieve higher capacities.

There are several ways of improving CDMA communication system capacity. They are:

Receiver beamforming,voice activation technology, power control, multiuser detection, using rake receivers and soft handoff.

A simple equation for the uplink capacity U of a single CDMA cell is given by:

$$U = 1 + WG/(E_b/N_O) - (\sigma 2/G) \quad (12)$$

Where the value of E_b/N_0 represents that required for adequate link performance. The scalar **σ2** is the background noise power and S is the received signal power for each user. Finally G is the ratio of the antenna gain for the desired user to that of interfering user in that cell. The value of G depends on the beam pattern for each user, but will roughly be proportional to the array size M.

As a result, antenna arrays can improve the capacity in two ways:

- Increasing the antenna gain G and hence the array M. This reduces the average level of interference from each user in the cell, permitting a capacity increase.

- Reducing the required, E_b/N_0 antenna array can provide increased space diversity at the base station, which can permit the receiver to operate at lower power signal. This increases the tolerance of the receiver to multiple access interference.

IV. Spectral Efficiency

Direct-Sequence Spread-Spectrum code-division multiple access (CDMA) has well-known desirable features: dynamic channel sharing, robustness to channel impairments, graceful degradation, ease of cellular planning, etc. These advantages result from the assignment of "signature waveforms" with large time–bandwidth products to every potential user of the system. Each signature can be viewed as a vector in an N-dimensional signal space, where is N the spreading gain or number of chips per symbol [12]. The presence of multipath fading leads to the loss of orthogonality between signals transmitted by the same base station, which results in additional self-interference. The existence of this self-interference renders the problem of computing the spectral efficiency even more difficult, as the SINR (Signal-to-interference plus noise ratio) becomes now a ratio of non-independent random variables. Consider a downlink CDMA multi-cellular system in a frequency-selective Rayleigh fading channel and let the composite CDMA downlink signal comprises M_1 mutually orthogonal signals, which are destined to M_1 different mobiles in the same cell. Though these signals appear orthogonal at the transmitter side, however, the orthogonality is lost at the receiver side after being propagated over a multipath fading environment, resulting in intra-cell interference (self-interference) [13]. This interference adds to the noise and the inter-cell interference from neighbouring cells.

The fundamental figure of merit is the *spectral efficiency* C, defined as the total number of bits per chip that can be transmitted arbitrarily reliably. Since the bandwidth of the CDMA system is (roughly) equal to the reciprocal of the chip duration, the spectral efficiency can be viewed as the bits per second per hertz (bits/s/Hz) supported by the system. Note that if the code rates (bits per symbol) employed by each individual user are identical and denoted by R then the spectral efficiency is equal to the product.

$$C = \frac{K}{v} R \quad (13)$$

Where K is the number of users and v is the spreading factor.

In a system where no spreading is imposed, the encoders are able to control the symbols modulating each chip independently. Therefore, the Cover-Wyner capacity region of the conventional Gaussian multi-access channel applies to this case and the spectral efficiency in the absence of spreading is given by:

$$C = \frac{1}{2}\log(1 + \frac{K}{v} SNR) \quad (14)$$

Once the spectral efficiency is determined, it is possible to obtain the minimum bandwidth necessary to transmit a predetermined information rate or the maximum information rate that can be supported by a given bandwidth. In order to compare different systems (with possibly different spreading gains and data rates), the spectral efficiency must be given as a function of $\frac{E_b}{N_o}$. If the spectral efficiency of the system reaches the optimum level, then SNR can be substituted by:

$$SNR = \frac{2N}{K}\frac{E_b}{N_o} C \quad (15)$$

Despite their overlap in time and frequency, the users can be completely separated at the receiver by means of a matched filter front-end provided the signature waveforms are mutually orthogonal. In that case, single-user error-control coding and decoding is sufficient. Moreover, channel distortion (such as multipath) and out-of-cell interference are common impairments that destroy the orthogonality of signature waveforms. Optimal spectral efficiency in non-orthogonal CDMA requires joint processing and decoding of users.

In our analysis of spectral efficiency we consider, in addition to optimal decoding (performed by a bank of matched filters), some popular linear multiuser detector front-ends:

• Single-user matched filter,

• Decorrelator.

V. Simulation and Results

From (5),

$$I_{tot} = I_i + N_o = \alpha(N-1)^{E_b}/_{G} + N_o$$

While writing the codes;

α is taken as 'a'= 3/8 (voice activity factor), Eb=2dB (1.9953) is taken as 'E' (Average bit energy of the channel's signal), N (Number of channels) which varies from 5 to 24, No (Noise power spectral density) whose values are generated by the MATLAB function randn(1,20), G=240 (Processing gain). The generated table is shown in table 1.

A plot of I against K from Program1 produces Figure 1.

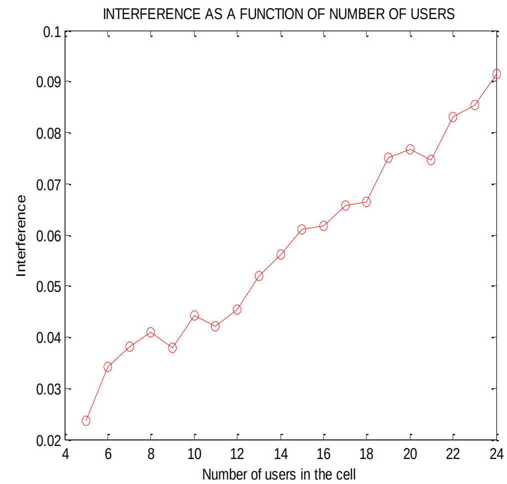


Fig 1: Graph of interference against Number of users

This shows that as the Number of users increase, the generated interference also increases. This necessitates the need for power control during transmission from MSs (mobile stations) to reduce the rate of growth of interference on the channel as users increase.

From (14); C $=\frac{1}{2}\log(1+\frac{K}{N}SNR)$

During plotting, K/N=B =0.1(Number of users per chip),C is taken as C (Spectral efficiency in the absence of spreading), SNR (Signal-to-noise-ratio), which ranges from 1 to10.The generated table is shown in Table 2.

Table 1 : Interference and number of channels used by users

A	G	E	No	N	I
3/8	240	1.9953	0.0112	5	0.0237
3/8	240	1.9953	0.0185	6	0.0341
3/8	240	1.9953	0.0195	7	0.0382
3/8	240	1.9953	0.0192	8	0.0410
3/8	240	1.9953	0.0129	9	0.0378
3/8	240	1.9953	0.0162	10	0.0443
3/8	240	1.9953	0.0109	11	0.0421
3/8	240	1.9953	0.0111	12	0.0453
3/8	240	1.9953	0.0145	13	0.0519
3/8	240	1.9953	0.0157	14	0.0562
3/8	240	1.9953	0.0175	15	0.0611
3/8	240	1.9953	0.0151	16	0.0619
3/8	240	1.9953	0.0159	17	0.0657
3/8	240	1.9953	0.0135	18	0.0665
3/8	240	1.9953	0.0189	19	0.0750

TABLE 2: SPECTRAL EFFICIENCY AND SNR

CONSTANT	B	SNR	C
0.5	0.1	1	0.021
0.5	0.1	2	0.040
0.5	0.1	3	0.057
0.5	0.1	4	0.088
0.5	0.1	5	0.102
0.5	0.1	6	0.115
0.5	0.1	7	0.128
0.5	0.1	8	0.139
0.5	0.1	9	0.151
0.5	0.1	10	0.161

A plot of C against SNR from Program 2 produces Figure. 2.

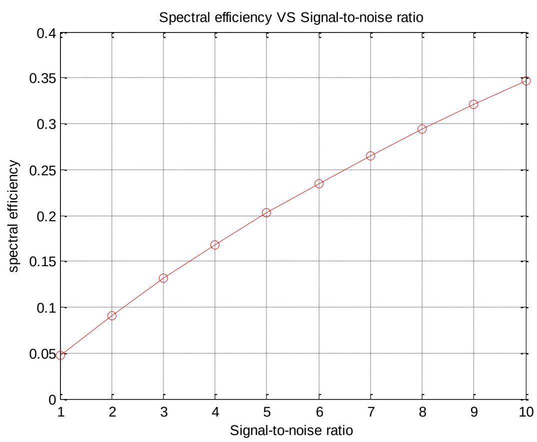


Fig. 2: Graph of Spectral efficiency as a function of Signal-to-noise ratio

It can be seen that at the receiver, as the Signal-to-noise ratio increases, the Spectral efficiency increases as well. In CDMA, as the number of users increase, multi-user interference increases causing the Signal-to-noise ratio at the receiver to decrease.

For acceptable quality of service a threshold Signal-to-noise ratio is set so that when exceeded, more users would not be accommodated. Also, interference reduction schemes are used to counter the effects of interference on a received signal to improve the efficient use of the allocated spectrum by adding more users until the threshold is reached.

From (15),

SNR= $\frac{2N}{K}\frac{E_b}{N_o}C$

It follows that C= 0.5B$\frac{SNR}{A}$

During plotting, B= K/N (Number of users per chip) which is the variable term;

A= Eb/No= 3 (Noise density ratio);

SNR= 100dB= 10

(Signal to noise ratio); C (Spectral efficiency). The generated table is;

TABLE 3: SPECTRAL EFFICIENCY AND NUMBER OF USERS PER CHIP

CONSTANT	B	SNR	A	C
0.5	1	10	100	0.05
0.5	2	10	100	0.10
0.5	3	10	100	0.15
0.5	4	10	100	0.20
0.5	5	10	100	0.25
0.5	6	10	100	0.30
0.5	7	10	100	0.35
0.5	8	10	100	0.40
0.5	9	10	100	0.45
0.5	10	10	100	0.50
0.5	11	10	100	0.55
0.5	12	10	100	0.60
0.5	13	10	100	0.65
0.5	14	10	100	0.70

From Program3, a plot of C against B produces Figure 3.

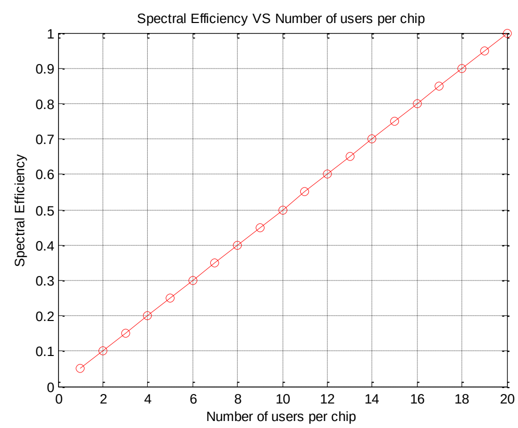


Fig. 3: Graph of Spectral efficiency as a function of Number of users per chip

It shows that as the Number of users increase, the Spectral efficiency increases as well. The implication of this is that within the threshold Signal-to-noise ratio already set, the more the number of users a channel carries, the better the utilization of the channel.

From (17) and (18),

$$\mathrm{C}^{\mathrm{sumf}}(\beta, \mathrm{SNR}) = \beta \log_2\left(1 + \frac{\mathrm{SNR}}{1+\mathrm{SNR}\beta}\right)$$

$$\mathrm{C}^{\mathrm{deco}}(\beta, \mathrm{SNR}) = \beta \log_2\left(1 + \mathrm{SNR}(1-\beta)\right), \qquad \beta \leq 1.$$

Where β = B (Number of users per chip), $\mathrm{C}^{\mathrm{deco}}$ = G (Decorrelator), $\mathrm{C}^{\mathrm{sumf}}$ = F Single-user matched filter). The generated table is:

TABLE 4: SPECTRAL EFFICIENCY OF SINGLE USERS

B	SNR	G	F
0.1	1	0.093	0.093
0.1	2	0.149	0.142
0.1	3	0.189	0.173
0.1	4	0.220	0.195
0.1	5	0.246	0.212
0.1	6	0.268	0.225
0.1	7	0.287	0.236
0.1	8	0.304	0.244
0.1	9	0.319	0.252
0.1	10	0.332	0.258

From Program4, plots of G against SNR and F against SNR produce the graphs in Figure 4.

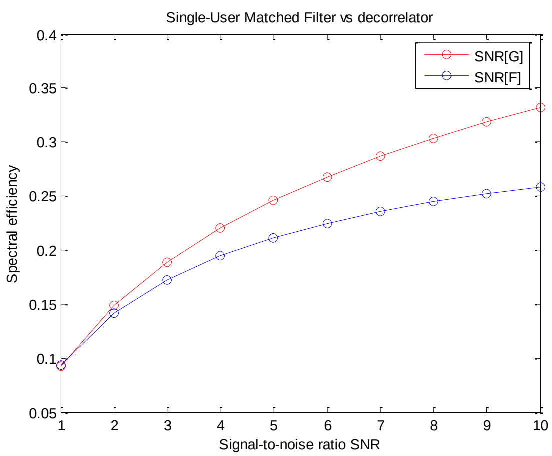


Fig. 4: Spectral efficiencies of Single-User

It shows the Spectral efficiencies of Single-user matched filter and Decorrelator when used at the receiver. Evidently, when faced with the same SNR, the Decorrelator performance is better than that of the Single-user matched filter. Hence, for effective bandwidth utilization and improved capacity, Decorrelators are preferred to Single-user matched filter.

VI. CONCLUSION

In this paper, through our simulations, we have analysed how interference grows as the number of users increase. We have also shown how the summation of interference and noise vary the capacity and spectral efficiency of a CDMA communication system. We have discussed the major problem in CDMA as the multiple Access interference (MAI) which arises due the deviation of the spreading codes from perfect orthogonality. Finally, the capacity of CDMA is interference limited .the obvious way to increase capacity of the CDMA is to reduce the level of interference through the methods described in this project.

REFERENCES

[1] Abdullah Aburomeh, (2009) “Capacity Analysis of Cellular CDMA Systems” (ECE 559 Project)

[2] A. M. Viterbi, A. J. Viterbi (2008) “Erlang Capacity of a power Controoled CDMA system”

[3] Darren D. Chang and Olivier L. de Weck, 2005‘*Basic capacity calculation methods and benchmarking for MF-TDMA and MF-CDMA communication satellites’* .INTERNATIONAL JOURNAL OF SATELLITE COMMUNICATIONS AND NETWORKING; 23:153–171.

[4] K. S. Gilhousen et al., 1991 “*On the capacity of a cellular CDMA system*,” *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303–312,

[5] Tony Dean, Phil Fleming, Alexander Stolyar, 1998 ‘*ESTIMATES OF MULTICARRIER CDMA SYSTEM CAPACITY’*. Proceedings of the Winter Simulation Conference.

[6] L. Nithyanandan, P. Dananjayan, ‘*Analytical BER computation for interference cancellation receiver in MC-CDMA system’*. Dept. Of Electronics and communication Engineering, Pondicherry Engineering College, Pondicherry-605014, India.

[7] J. C. Liberti and T. S. Rappaport,1999 ‘*Smart Antennas for Wireless Communications’: IS-95 and Third Generation CDMA Applications*. Englewood Cliffs, NJ: Prentice-Hall

[8] Aylin Yener, Roy D. Yates, Sennur Ulukus, 2001 ‘ *Interference Management for CDMA Systems Through Power Control, Multiuser Detection, and Beamforming*’. IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 49, NO. 7

[9] Jin Yu, Yu-Dong Yao, and Jinyun Zhang-2004 “IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 53, NO. 5, pp 1423.

[10] Kamal Shahtalebi, Gholam Reza Bakhshi, Hamidreza Saligheh Rad, 2007 ‘*Interference Cancelation in Coherent CDMA Systems Using Parallel Iterative Algorithms*’

[11] Robert Akl, Manju Hegde, Alex Chandra,1998 ‘*CCAP: CDMA Capacity Allocation and Planning’*. Washington University, Missouri, USA

[12] Sergio Verd´u, and Shlomo Shamai (Shitz), 2009 ‘*Spectral Efficiency of CDMA with Random Spreading*’. IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 45, NO. 2.

[13] Khairi Ashour Hamdi, and Pongsatorn Sedtheetorn, 2008‘*On the Spectral Efficiency of CDMA Downlink Cellular Mobile Networks’*. IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 7, NO. 11

[14] Shlomo Shamai (Shitz), and Sergio Verdú,2001 ‘*The Impact of Frequency-Flat Fading on the Spectral Efficiency of CDMA’*. IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 47, NO. 4.

Modeling of Traffic Accident Reporting System through UML Using GIS

Dr. Gufran Ahmad Ansari
Assistant Professor
Department of Computer Science
College of Computer, Qassim University
Al-Qassim, Kingdom of Saudi Arabia (KSA)

Dr. M. Al-shabi
Assistant Professor
Department of Information Technology
College of Computer, Qassim University
Al-Qassim, Kingdom of Saudi Arabia (KSA)

***Abstract*— Nowadays vehicles are increasing day by day in town and cities roads. It is a well-known problem to manage traffics on the roads in towns and cities. A lot of accident occurs on the road due to careless driving and Technical faults in vehicles. The main problem of traffic authorities is to manage the traffic on the road for the smooth functioning of vehicles that can reduce the accident and violation on the road. There is a tremendous demand from traffic authorities to develop a system that can helps to avoid the accident and keep the accident report data and also maintain the accident report data. The main objective of this paper is to model a Traffic Accident Reporting System (TARS) through UML using GIS to solve the above problem. Authors are also proposed the sequence and activity diagram for the above proposed model.**

Keywords- UML; GIS; TARS; Sequence diagram; Activity Diagram.

I. INTRODUCTION

Models are playing very important role to understand real time problems. A model gives an overall idea about the actual problem in a very simple and clear way [1]. The Object Management Group introduced the Unified Modeling Language (UML) for the software designers to develop useful, efficient, effective designs and quality model system for the industry peoples [2, 3].The Unified Modeling Language (UML) is a modeling language that covers a large range of different application domains and which is used to design a scientific and research problems [4]. UML model is an accepting a view of actual real world problem and explain in the form of pictures and notations [5]. UML have nine standard diagrams for graphic representation of a system which is represent the different points of view of the system and that are classes, interaction- sequence, objects, interaction-communication, state, use , activities, components and display[6].Some of the important domain oriented UML models are designed and shown in [7,8,9].Geography always plays the important role in human's life. A geographic information system (GIS) is a kind of system which is used to capture, designed, store, manipulate, analyze and manage all types of geographically referenced data [10,11]. The geographic knowledge is applied to human routine tasks such as unfamiliar with the city or searching the exact street or station etc. [12]. Recently there are some important research papers about explaining the GIS system in a very effective and efficient way is given in [13, 14]. These are the some paper that explaining and prevention about the accident in a very simple way [15, 16].

II. UML CLASS MODEL FOR TRAFFIC ACCIDENT REPORTING SYSTEM

A Traffic Accident Reporting System (TARS) has been designed with the use of UML concepts and which is shown in Fig1. UML class diagram demonstrate the structure of the system by depicting classes, attributes and relationship. The complete Traffic Accident Reporting System TARS have been designed with attributes and functions. The different properties have been used like association, aggregation, inheritances etc in the form of sub classes and shown in the UML class model. In a UML class diagram Viewers class multiple associations with TARS and TARS has a single association with Viewers class. Similarly TARS class also has a single association with GIS class and multiple associations with Insurance_Com and Traffic_Officer classes. GIS class has a single association with TARS class and multiple associations with Vehicle_detail class also GIS class has a multiple association with Street class and Street class also has a single association with GIS class. Vehicle_datail class multiple associations with GIS class and also multiple associations with Report class. Report class has a multiple associations with Vehicle_Detail, Traffic_Officer and Insurance_Com class also the Traffic_Officer class and Insurance_Com class multiple associations with Report class. TARS also have a multiple association with Traffice_Officer class and Traffic_Officer class also has the multiple associations with TARS class.

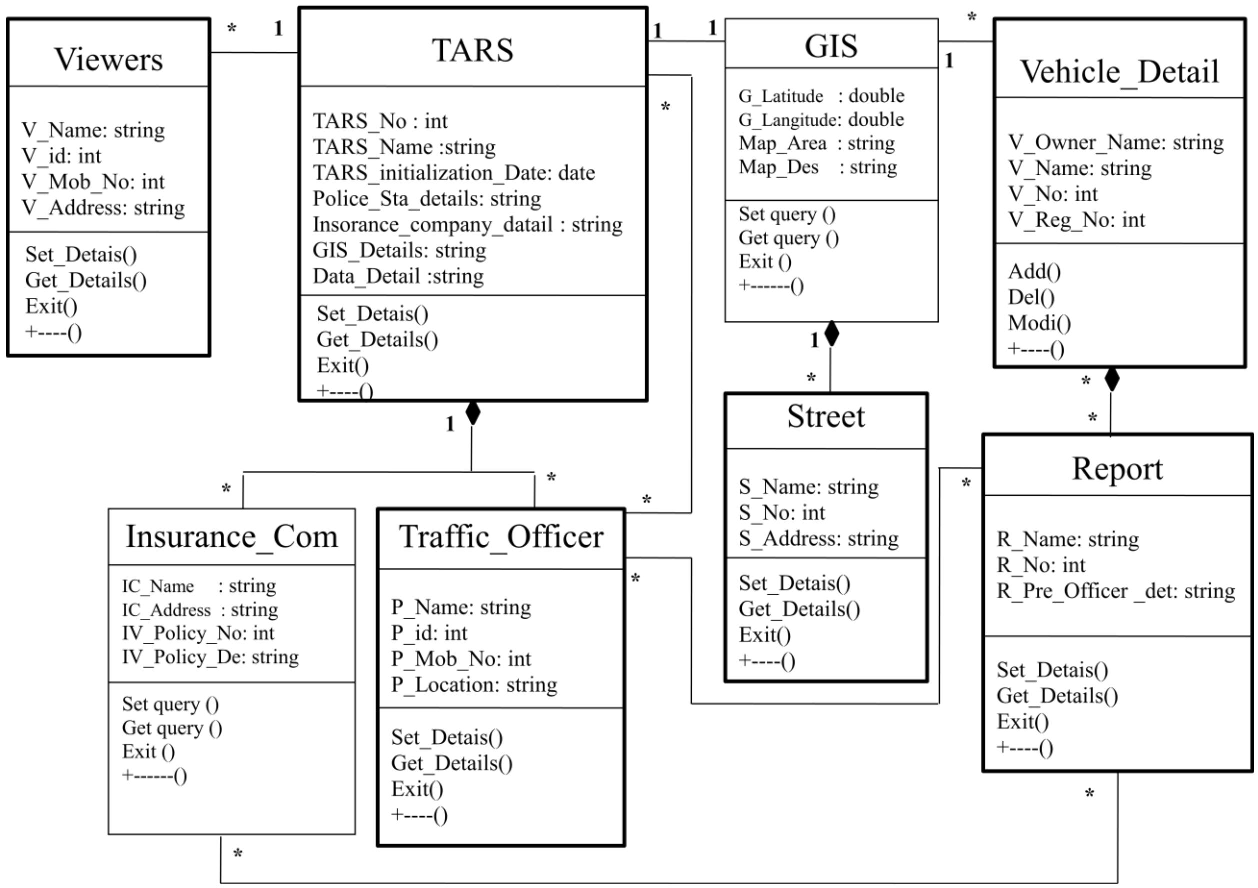


Figure 1. UML Class Diagram Traffic Accident Reporting System

III. UML Activity Diagram For Traffic Accident Reporting System

An activity diagram is a kind of flowchart that shows the flow of control step-by-step [17]. The activity diagram shows the various activities one by one with the moving for both controlled and uncontrolled activities. The UML activity diagram of the above model is shown in Fig.2. The activity diagram represents the complete process of Traffic Accident Reporting System. According to the activity diagram viewers see the accident and inform to the TARS with mail or through the phone even viewers can go directly traffic police and inform about the accident. When the TARS get the information about accident the system will connect GIS system through internet and find the exact location of the accident. After finding the exact location of the accident the TARS search the nearest police station and inform the responsible person also the same time TARS send the information to the insurance company. When the police officer gets the information about the accident he reaches the proper location prepares the report about the accident and send to the TARS. After that clear the traffic and go back to the other task. Similarly insurance officer reaches the accident place and prepares report and sends to the insurance company and go for other task.

IV. UML Sequence Diagram For Traffic Accident Reporting System

The sequence diagram is representing the interactions between objects. The sequence diagram passes the message from top to bottom. The sequence diagram of above TARS UML model is given in Fig.3 This sequence diagram of TARS have five important object which are shown on the top of the diagram in the form of rectangles boxes with their class names. The five main objects are Accident Effected party, TARS, GIS, Officer and Insurance Company. The communication between two objects is shows by an arrow and the message of that arrow. According to the sequence diagram is given in Fig. 3. The Accident Effected party inform to the Traffic Accident Reporting System (TARS) through the mail or phone. The TARS connect to GIS system and find exact location of the accident. After finding the exact location, system searches the nearest traffic police station and informs and assigns the task to the traffic police officer. Similarly at the same time also inform to insurance company. The traffic police officer and insurance company officer reach the accident place and prepare the report. The insurance company officer sends the report to the company. The traffic police officer prepare the report and send to TARS also the officer clear the traffic after that go to the other task.

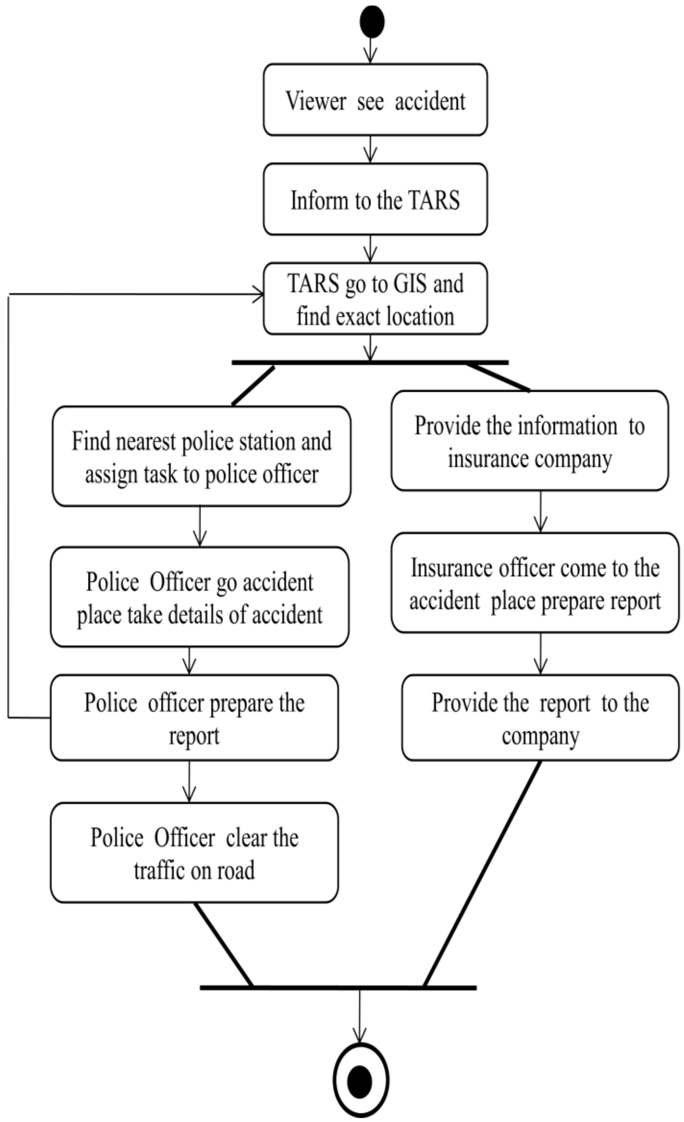


Figure 2. UML Activity Diagram for Traffic Accident Reporting System

V. A Case And Experimental Study

Saudi Arabia has the highest Road Accident Death Toll in the world and averages of 17 Saudi Arabian residents are died on the country's road each day. A report by the Kingdom's General Directorate of Traffic has revealed [18, 19]. Let us consider the accidental data of Saudi Arabia for evaluating the TARS model.

The following table shows statistics from the General Directorate of Traffic accident and traffic violation recorded in the Kingdom of Saudi Arabia during the years 2000-2008 [20]. The bar chart is giving for the TARS that shown in Fig. 4. The blue color bars showing year and the brown color is shows the number of injuries happened in a year.

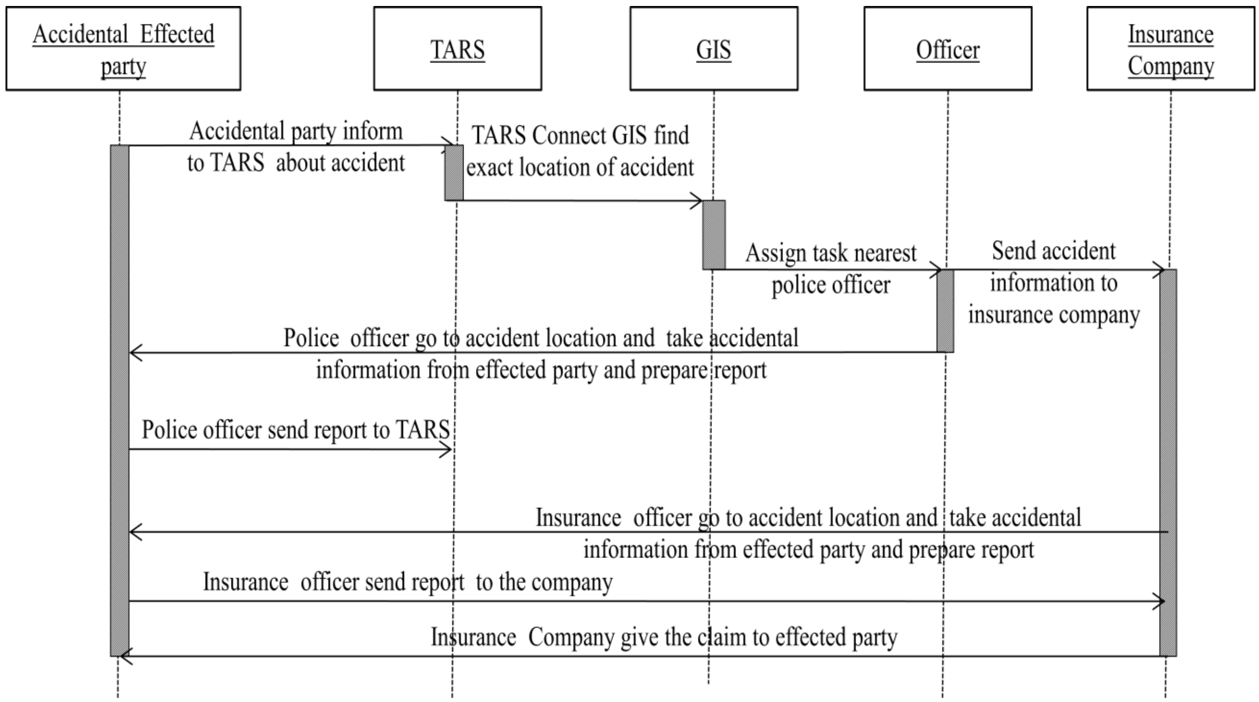


Figure 3. UML Sequence Diagram for Traffic Accident Reporting System

TABLE I. TRAFFIC ACCIDENTS AND TRAFFIC VIOLATIONS IN SAUDI ARABIA

Year	No. of Injuries	No. of Deaths	No. of Accidents
2000	28,998	4,419	280,401
2001	28,379	3,913	305,649
2002	28,372	4,161	223,816
2003	30,439	4,293	261,872
2004	34,811	5,168	293,281
2005	34,441	5,982	296,051
2006	35,884	5,883	283,648
2007	36,025	6,358	435,264
2008	36,489	6,458	485,931
Total	**293,838**	**46,635**	**2,865,913**

Source: General Directorate of Traffic Accident and Traffic Violation KSA [20].

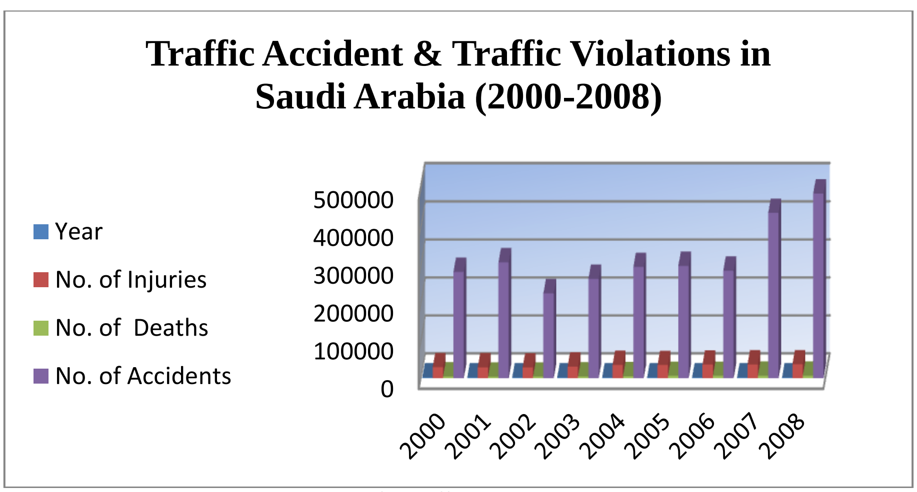


Figure 4. Bar chart for Traffic Accident Reporting System

The green bar is showing the no of death happen in each year and last Byzantium bar is showing the no of accident happened in each year from 2000 to 2008.

VI. Conclusion And Future Work

From the above study of work it is accomplished that the UML is a powerful modeling language to solve scientific and research problems. In this paper a complete modeling of Traffic Accident Reporting System has been done through the UML and its results are shown in the form of bar chart graph. This model is a simple and has a reusability property also model can easily enhance, modify and updated according to the need of data. This basic work can be expended in the field of data mining using UML and expert system.

References

[1] G. Booch, J. Rumbaugh, and I, Jacobson "The Unified Modelling Language User guide", Addison-Wesley, Reading, MA, 1999.

[2] OMG Unified Modeling Language Specification, Available online via http://www.omg.org, 2001.

[3] Kovacevic, S. UML and User Interface Design, inUML'98. Mulhouse – France, year, 1998.

[4] Roff, T., UML: A Beginner's Guide, Tata McGraw-Hill Edition, Fifth Reprint, 2006.

[5] Craig Larman: "Applying UML and Pattern", Prentice Hall, 1998.

[6] Booch, G., Rumbaugh, J. & Jacobson, I. The UML reference manual (2nd ed.). Boston: Addison- Wesley, 2004.

[7] V. Saxena and Ansari, G.A., Ajay Pratap **"**Enhancing Security through UML", **"**International Journal of Computer Sciences, Software Engineering and Electrical Communication Engineering" Vol. 2(1), pp 31-36 June 2011.

[8] Ansari, G.A., "A Domain Oriented Modeling of Indian Education System through UML" the Icfai Journal of Systems Management (IJSM) (ICFAI Press India), Vol. VIII; No.3; August, 2010.

[9] V. Saxena and Ansari, G.A., "UML Modeling & Protection of Domain Based System" International Journal of Computer Science and Network Security (IJCSNS), South Korea, Vol. 8, No. 7, pp. 338-344, July 2008.

[10] Geographic Information Systems as an Integrating Technology: Context, Concepts, and Definitions" ESRI, 9 June 2011.

[11] Amdahl, G. Disaster Response: GIS for Public Safety, ESRI Press: Redlands, California, 2001.

[12] De Smith, M.G., Goodchild, M.F. & Longley, P.A., Geospatial Analysis: A Comprehensive Guide to Priciples, Techniques, and Software *Tools.*Winchelsea Press: Leicester, 2007.

[13] Fonseca, Frederico, Sheth, Amit "The Geospatial Semantic Web" UCGIS, White Paper, 2002.

[14] Longley, P.A, Goodchild, M.F, Maguire, D.J and Rhind, D.W "Geographic Information Systems and Sciences, Chichester: Wiley, 2nd edition, 2005.

[15] A.S. Al-Ghamdi, Z. Nemeth and R.Rogness. "Forecasting Traffic Accidents in Saudi Arabia by Using a Time Series Model" Presented at the 72nd Annual Meeting of TRB Conference, Washington, D.C, 1993.

[16] S.E. Asogwa, The Use of the Police for Limited Road Accident Data Collected in Developing Countries. Accident Analysis and Prevention, Vol. 14, No.3, pp. 203-208 1982.

[17] Ericcson. Maria. "Activity Diagrams: What They Are and How to Use Them" http://www.ibm.com/developerworks/rational/library/2802.html Accessed on dated 1-1-2012.

[18] Public Health (2000) 114, 37–39. http://www.publichealthjrnl.com/article/S0033-3506%2800%2900306-1/abstract 2000

[19] The General Directorate of Traffic, Ministry of Interior, Riyadh. The Publications of Road Accident Statistics, for the years 1974-1994.

[20] Yassar Hauas "Workshop for Setting Reginioal and National Road Traffic Casuality Reduction Targets in the ESCWA Abu Dhabi, UAE 16-17 June 2009.

Author's Profile

Dr. Gufran Ahamd Ansari received his Bachelor degree (B.Sc. Computer Science) from Shia P.G. College, Lucknow, India in 1997, MCA from DR. B.R. Ambedkar University Agra in 2002 and Ph.D (Computer Science) from Babasaheb Bhimrao Ambedkar (A Central) University, Lucknow, U.P. India in 2009. He is currently working as an Assistant Professor in Department of Information Technology, College of Computer Qassim University Saudi Arabia. He has produced several outstanding publications in National & International Journal on various research problems related to the domain based UML modeling, Security, Testing and Designing.

Dr. M. Al-Shabi received his Bachelor degree (B.Sc. Computer Science) from Technology University at Iraq (1997). , Post graduate Master (M. Sc. Computer Science from Putra Malaysia University at 2002), and PhD (Computer Science) from Putra Malaysia University, Malaysia (2006). He is currently an assistant professor of College of Computer at Qassim University. Kingdom of Saudi Arabia. Prior to joining Qassim University he worked in the Faculty of computer at Sana'a University, Yemen. His research interests include: wireless security, cryptography, UML, Stenography Multistage interconnection network, parallel computing and Apply Mathematic.

Towards Multi Label Text Classification through Label Propagation

Shweta C. Dharmadhikari
PICT
Pune, India

Maya Ingle
DAVV
Indore, India

Parag Kulkarni
EkLaT Solutions
Pune, India

***Abstract*— Classifying text data has been an active area of research for a long time. Text document is multifaceted object and often inherently ambiguous by nature. Multi-label learning deals with such ambiguous object. Classification of such ambiguous text objects often makes task of classifier difficult while assigning relevant classes to input document. Traditional single label and multi class text classification paradigms cannot efficiently classify such multifaceted text corpus. Through our paper we are proposing a novel label propagation approach based on semi supervised learning for Multi Label Text Classification. Our proposed approach models the relationship between class labels and also effectively represents input text documents. We are using semi supervised learning technique for effective utilization of labeled and unlabeled data for classification. Our proposed approach promises better classification accuracy and handling of complexity and elaborated on the basis of standard datasets such as Enron, Slashdot and Bibtex.**

Keywords- Label propagation; semi-supervised learning; multi-label text classification.

I. INTRODUCTION

The area of text classification is getting more popular among the researchers. The major objective of text classification system is to organize the available text documents systematically into their respective categories [7]. This categorization of text documents facilititates ease of storage, searching, retrieval of relevant text documents or its contents for the needy applications. Three different paradigm exists under text classification and they are single label(Binary) , multiclass and multi label. Under single label a new text document belongs to exactly one of two given classes, in multi-class case a new text document belongs to just one class of a set of m classes and under multi label text classification scheme each document may belong to several classes simultaneously [3]. In real practice many approaches are exists and proposed for binary case and multi class case even though in many applications text documents are inherently multi label in nature. Eg. In the process of classification of online news article the news stories about the scams in the commonwealth games in india can belong to classes like sports, politics, country-india etc.

Multilabel text classification problem refers to the scenario in which a text document can be assigned to more than one classes simultaneously during the process of classification.. It has attracted significant attention from lot of researchers for playing crucial role in many applications such as web page classification, classification of news articles, information retrieval etc. Generally supervised methods from machine learning are mainly used for realization of multi label text classification. But as it needs labeled data for classification all the time, semi supervised methods are used now a day in multi label text classifier. Many approaches are preferred to implement multi label text classifier. Through our paper we are proposing label propagation approach for multi label text classifier, it uses existing label information for identifying labels of unlabeled documents. We are representing input text document corpus in the form of graph to exploit the ambiguity among different text documents. The ambiguity is represented in the form of similarity measures as a weighted edge between text documents. With the setting of semi supervised learning we have focused on not only graph construction but also sparsification and weighting of graph to improve classifiers accuracy. We apply the proposed framework on standard dataset such as Enron, Bibtex and slashdot.

The rest of the paper is organized as below. Section 2 describes literature related to semi supervised learning methods for multi label text classification system; Section 3 highlights mathematical modeling of our approach. Section 4 describes our proposed label propagation approach for building multi label text classifier followed by experiments and results in Section 5, followed by a conclusion in the last section.

II. RELATED WORK

Multilabel text classifier can be realized by using supervised, unsupervised and semi supervised methods of machine learning. In supervised methods only labeled text data is needed for training. Unsupervised methods relies heavily on only unlabeled text documents; whereas semi supervised methods can effectively use unlabeled data in addition to the labeled data[1][2]. The traditional approach towards multi-label learning either decomposes the classification task into multiple independent binary classification tasks or identifies rank to find relevant set of classes. But these methods do not exploit relationship among class labels. Few popular existing methods are binary relevance method, label power set method, pruned sets method, C4.5, Adaboost.MH & Adaboost.MR, ML-kNN, Classifier chains method etc[20]. But all these are lacking the capability of handling unlabeled data ie these are based on principle of supervised learning.

While designing a multi label text classifier the major objective is not only to identify the set of classes belonging to given new text documents but also to identify most relevant out of them to improve accuracy of overall classification process. Graph based approaches are known for their effective exploration of document representation and semi supervised methods explores both labeled and unlabeled data for classification that's why accuracy of multi label text classifier can be improved by using graph based representation of input documents in conjunction with label propagation approach of semi supervised learning[16][17].

Table 1 summarizes few existing well-known representative methods for multi label text classifier based on semi supervised learning, few uses only graph based framework and few uses both.

TABLE 1: STATISTICS OF POPULAR ALGORITHMS FOR MLTC BASED ON SEMI SUPERVISED LEARNING AND GRAPH BASED REPRESENTATION

Algorithm and Year of proposal	Working Theme	Datasets used for experimentation	Merits	Demerits
Multi-label classification by Constrained Non-Negative Matrix Factorization [2006] [8]	Optimization of class labels assignment by using similarity measures and non negative matrix factorization.	ESTA	Powerful representation of input documents using NMF and also works for large scale datasets	Parameter selection is crucial.
Graph-based SSL with multi-label [2008][9]	Exploits correlation among labels along with labels consistency over graph.	Video files: TECVID 2006.	Effective utilization of unlabeled data.	Can not applicable to text data , more effective on video data.
Semi supervised multi-label learning by solving a Sylvester Eq [2010][10]	Graph construction for input documents and class labels.	Reuters	Improved accuracy	May get slower on convergence.
Semi-Supervised Non negative Matrix Factorization [2009][11] .	Performs joint factorization of data and labels and uses multiplicative updates performs classification.	20-news, CSTR, k1a,k1b,WebKB4, Reuters	Able to extract more discriminative features	High computational complexity.

In preprocessing stage graph based approaches can effectively represents relationship between labeled and unlabeled documents by identifying structural and semantical relationship between them for more relevant classification ; and during training phase semi supervised methods can propagate labels of labeled documents to unlabeled documents based on some energy function or regularizer. Our proposed work is based on the same strategy.

III. MATHEMATICAL MODEL OF PROPOSED SYSTEM

In this section we are introducing some notions related with text classification. We are firstly representing the input document corpus in the form of graph.The process of graph construction deals with conversion of input text document corpus , X to graph G ie X → G , where X represents input text document corpus x1,x2,..,xn wherein each text document instance x_i in turn represented as m-dimensional feature vector. And G represents overall graph structure as G=(V,E) where V = set of vertices corresponding to document instance x_i ; E represents set of weighted edges between pair of vertices where associated edge weight corresponds to similarity between two documents. Generally weight matrix W is computed to identify the similarity between pair of text documents. Various similarity measures such as cosine, Jacobi or kernel functions K(.) like RBF kernel , Gaussian kernel can be used for this purpose.

Now we are defining our graph based multi label text classifier system S as follows:

S = { X, Y, T, $\hat{y}$, h}; where X represents entire input text document corpus = {x1,x2,..,xn}. Out of these |L| numbers of documents are labeled and remaining are unlabeled.Y represents set of possible labels = {Y1,Y2,…,Yn}. T represents multilabel training set of classifier of the form {(x1,Y1), (x2,Y2),….., (xn,Yn)} where xi $\in X$ is a single document instance and Yi $\subseteq$ Y is the label set associated with xi . $\hat{y}$ represents set of estimated labels = {$\hat{Y}$l , $\hat{Y}$u}. The goal of the system is to learn a function h i.e.

h : X → 2y from T which predicts set of labels for unlabeled documents i.e. xl+1 ..xn

With this graph based setting, we are using semi supervised learning to propagate labels on the graph from labeled nodes to unlabeled nodes and compare the estimated labels $\hat{y}$ with the true labels.

IV. PROPOSED APPROACH

We are mainly using theme of smoothness assumption of semi supervised learning to propagate the labels of labeled documents to unlabeled documents. Smoothness assumption of semi supervised learning states that "if two input points x1,x2 are in a high-density region are close to each other then so should be the corresponding outputs y1,y2". Thus based on this we mainly emphasized on exploiting relationships between input text documents in the form of graph and relationship between the class labels in the form of correlation matrix. The purpose behind this is to reduce classification errors and assignment of more relevant class labels to new test document instance.

During classifiers training phase we are computing similarity between input documents to identify whether they are in high density or low density region. We evaluated relationships between documents by using cosine similarity measure and represented it in the form of weighted matrix, W as :

$$Wij = \cos \quad \frac{X1 \,.\, X2}{|X1|.|X2|}$$

Where X1and X2 are two text documents represented in the feature space. Large cosine value indicates similarity and small value indicates that documents are dissimilar. After that we performed graph sparcification by representing it in the form of diagonal matrix in order to reduce consideration of redundant data.While identifying relationships between class labels we computed correlation matrix C mxm where m is no. of class labels using RBF kernel. Each class is represented in the form of vector space whose elements are said to be 1 when corresponding text document belongs to the class under consideration. Then in testing phase, in order to provide relevant label set to unlabeled document we computed energy function E to measure smoothness of label propagation. This energy function measures difference between weight matrix W and dot product of sparcified diagonal matrix with correlation matrix.

$$\mathbf{E = \sum Wij \ - \ D^{-1}Cij}$$

The labels are propagated based on minimum value of Energy function. It indicates that if two text documents are similar to each other than the assigned class labels to them are also likely to be closer to each other. In other words two documents sharing highly similar input pattern are likely to be in high density region and thereby the classes assigned to them are likely to be related and propogated to those documents which in turn resides in same high density region.

After this label propagation phase, we obtained labels of all unlabeled document instances. We computed accuracy to verify correct assignment of label sets. The corresponding results are given in table [III]. We once again ensured the working by applying all this document and label set to existing classifier chains method which is supervised in nature. We used decision tree(J48 in WEKA),SVM (SMO & libSVM) separately as base classifiers and computed the results. The corresponding results are given in table [IV].

The summary of our proposed label propagation approach is given as:

Input - T: The multi label training set {(x1,Y1), (x2,Y2),….., (xn,Yn)}.

z: The test document instance such that z $\in$ X

Output – The predicted label set for z.

Process: Compute the edge weight matrix W as $Wij = \text{arc}\cos \frac{X1\,.\,X2}{|X1|.|X2|}$ and assign W_{ii}=0

- Sparcify the graph by computing diagonal degree matrix D as $D_{ii}=\sum_j W_{ij}$
- Compute the label correlation matrix C $_{mxm}$ using RBF kernel method
- Initialize $\hat{Y}^{(0)}$ to the set of ($Y_1, Y_2, \ldots, Y_l$,0,0,……..,0)
- Iterate till convergence to $\hat{Y}^{(\infty)}$
  1. $E = \sum W_{ij} - D^{-1}C_{ij}$
  2. $\hat{Y}^{(t+1)} = E$
  3. $\hat{Y}^{(t+1)}{}_{l} = Y_l$
- Label point z by the sign of $\hat{Y}^{(\infty)}{}_i$

V. EXPERIMENTS AND RESULTS

In this section, in order to evaluate our approach we conducted experiments on three text based datasets namely Enron , Slashdot , Bibtex and measured accuracy of overall classification process. Table II summarizes the statistics of datasets that we used in our experiments.

TABLE II: STATISTICS OF DATASETS

Dataset	No. of document instances	No. of Labels	Attributes
Slashdot	3782	22	500
Enron	1702	53	1001
Bibtex	7395	159	1836

Enron dataset contains email messages. It is a subset of about 1700 labeled email messages[21]. BibTeX data set contains metadata for the bibtex items like the title of the paper, the authors, etc. Slashdot dataset contains article titles and partial blurbs mined from Slashdot.org [22].

We used accuracy measure proposed by Godbole and Sarawagi in [13] . It symmetrically measures how close y_i is to Zi ie estimated labels and true labels. It is the ratio of the size of the union and intersection of the predicted and actual label sets, taken for each example and averaged over the number of examples. The formula used by them to compute accuracy is as follows:

$$Accuracy = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{Y_i \cap Z_i}{Y_i \cup Z_i}\right]$$

We also computed precision , recall and F-measure values , the formula used to compute them is as follows:

F-Measure = 2.0 x precision x recall
precision + recall

$$\textbf{F-Measure} = \frac{1}{N}\sum_{i=1}^{N}\frac{2[Y_i \cap Z_i]}{[Z_i]-[Y_i]}$$

We evaluated our approach under a WEKA-based [23] framework running under Java JDK 1.6 with the libraries of MEKA and Mulan [21][22]. Jblas library for performing matrix operations while computing weights on graph edges. Experiments ran on 64 bit machines with 2.6 GHz of clock speed, allowing up to 4 GB RAM per iteration. Ensemble iterations are set to 10 for EPS. Evaluation is done in the form of 5 × 2 fold cross validation on each dataset. We first measured the accuracy, precision, Recall and after label propagation phase is over. Table III enlists accuracy measured for each dataset.

TABLE III: RESULTS AFTER LABEL PROPAGATION PHASE

Evaluation Criterion	Enron	Slashdot	Bibtex
Accuracy	90	89	92
Precision	50	49	48
Recall	49	47	46
F-measure	50	47	47

After label propagation phase, we obtained labels of all unlabeled documents. Thus we get entire labeled dataset as a result now. We applied this labeled set to Ensemble of classifier chains method which is supervised in nature[24] and measured accuracy ,precision, recall on three different base classifiers of decision tree(J48 in WEKA) , and two variations of support vector machine (SMO in WEKA , libSVM).We also measured overall testing and building time required for this process. The Ensemble of classifier chains method (ECC) is proven and one of the efficient supervised multi label text classification technique, we verified our entire final labeled dataset by giving input to it. The results are enlisted in table IV

TABLE IV: RESULT AFTER USING SUPERVISED MULTI LABEL CLASSIFIER

Datasets	Slashdot			Enron			Bibtex		
Parameters	SMO	libSVM	J48	SMO	libSVM	J48	SMO	libSVM	J48
Test Time	70	28.9	29	69	28.9	29	68	29	29
Build Time	173	2610	2546	173	2609	2546	172	2610	2546
Accuracy	0.5	0.51	0.31	0.5	0.52	0.31	0.4	0.51	0.31
Precision	0.9	0.49	0.44	0.9	0.51	0.44	0.9	0.49	0.44
Recall	0.4	0.56	0.53	0.4	0.56	0.53	0.4	0.56	0.53

VI. CONCLUSION AND FUTURE WORK

We have proposed a novel label propagation based approach for multi label classifier. It works in conjunction with semi supervised learning setting by considering smoothness assumptions of data points and labels. The approach is evaluated using small scale datasets (Enron, Slashdot) as well as large scale dataset (Bibtex). It is also verified against traditional supervised method. Our approach shows significant improvement in accuracy by incorporating unlabeled data along with labeled training data. But significant amount of computational time is required to calculate similarity among documents as well as class labels.The input text corpus is well exploited as a graph however, in the future the use of feature extraction methods like NMF with Latent Semantic indexing may provide more stable results.

REFERENCES

[1] J. Zhu. Semi-supervised learning Literature Survey. Computer Science Technical Report TR 1530 , University of Wisconsin – Madison , 2005.

[2] Olivier Chapelle , Bernhard Schfolkopf , Alexander Zien. Semi-Supervised Learning2006 , 03-08 , MIT Press.

[3] G. Tsoumakas, I. Katakis. Multi-label classification: An overview. International Journal of Data Warehousing and Mining, 3(3):1-13, 2007.

[4] A. Santos , A. Canuto, and A. Neto, "A comparative analysis of classification methods to multi-label tasks in different application domains", International journal of computer Information systems and Industrial Management Applications". ISSN: 2150-7988 volume 3(2011), pp. 218-227.

[5] R.Cerri, R.R. Silva , and A.C. Carvalho , "Comparing Methods for multilabel classification of proteins using machine learning techniques",BSB 2009, LNCS 5676,109-120,2009.

[6] G. Tsoumakas , G. Kalliris , and I. Vlahavas, " Effective and efficient multilabel classification in domains with large number of labels", Proc. Of the ECML/PKDD 2008 workshop on Mining Multidimensional Data (MMD' 08)(2008) 30-44.

[7] Nigam, K., McCallum, A. K., Thrun, S., & Mitchell, T. M. (2000). Text classification from labeled and unlabeled documents using EM. Machine Learning,39, 103–134.

[8] Y. Liu, R. Jin, L. Yang. Semi-supervised Multi-label Learning by Constrained Non-Negative Matrix Factorization .In: AAAI, 2006.

[9] Z. Zha, T. Mie, Z. Wang, X. Hua. Graph-Based Semi-Supervised Learning with Multi-label. In ICME. page 1321-1324, 2008.

[10] G. Chen, Y. Song, C. Zhang. Semi-supervised Multi-label Learning by Solving a Sylvester Equation. In SDM, 2008.

[11] Semi-supervised Nonnegative Matrix factorization. IEEE. January 2011.

[12] Qu Wei , Yang, Junping, Wang. Semi-supervised Multi- label Learning Algorithm using dependency among labels. In IPCSIT vol. 3 2011.

[13] S. Godbole and S. Sarawagi , "Discriminative methods for multi-labeled classification", 8th Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2004.

[14] R. Angelova , G. Weikum . "Graph based text classification : Learn from your neighbours". In SIGIR'06 , ACM , 1-59593-369-7/06/0008".

[15] T.Jebara , Wang and chang , "Graph construction and b-matching for semi supervised learning". In proceedings of ICML-2009.

[16] Thomas, Ilias & Nello. " scalable corpus annotation by graph construction and label propogation". In proceedings of ICPRAM, 25-34, 2012.

[17] P. Talukdar , F. Pereira. " Experimentation in graph based semi supervised learning methods for class instance acquisition". In the proceedings of 48th Annual meet of ACL. 1473-1481.2010.

[18] X. Dai, B. Tian , J. Zhou , J. Chen. "Incorporating LSI into spectral graph transducer for text classification" . In the proceedings of AAAI. 2008.

[19] S.C. .Dharmadhikari , Maya Ingle , parag Kulkarni .Analysis of semi supervised methods towards multi-label text classification. IJCA , Vol. 42, pp. 15-20 ISBN :973-93-80866-84-5.

[20] S.C. .Dharmadhikari , Maya Ingle , parag Kulkarni .A comparative analysis of supervised multi-label text classification methods. IJERA , Vol. 1, Issue 4 , pp. 1952-1961 ISSN : 2248-9622.

[21] http://mulan.sourceforge.net/datasets.html

[22] http://MEKA.sourceforge.net

[23]www.cs.waikato.ac.nz/ml/weka/

[24] J. Read, B. Pfahringer, G. Homes , and E.Frank , "Classifier chains for multi-label classification"., Proc. Of European Conference on Machine Learning and knowledge discovery in Databases, LNAI 5782(254-269),2009

Segment, Track, Extract, Recognize and Convert Sign Language Videos to Voice/Text

P.V.V.Kishore
Dept. of ECE, Research Scholar
Andhra University College of Engineering
Andhra Pradesh, INDIA

P.Rajesh Kumar
Dept. of ECE, Associate Professor
Andhra University College of Engineering
Andhra Pradesh, INDIA.

Abstract— **This paper summarizes various algorithms used to design a sign language recognition system. Sign language is the language used by deaf people to communicate among themselves and with normal people. We designed a real time sign language recognition system that can recognize gestures of sign language from videos under complex backgrounds. Segmenting and tracking of non-rigid hands and head of the signer in sign language videos is achieved by using active contour models. Active contour energy minimization is done using signers hand and head skin colour, texture, boundary and shape information. Classification of signs is done by an artificial neural network using error back propagation algorithm. Each sign in the video is converted into a voice and text command. The system has been implemented successfully for 351 signs of Indian Sign Language under different possible video environments. The recognition rates are calculated for different video environments.**

Keywords- Sign Language; Multiple Video object Segmentations and Trackings; Active Contours; skin colour, texture, boundary and shape; Feature Vector; Artificial Neural Network.

I. INTRODUCTION

Sign language is a vision based language of hearing impaired people, which involves the use of hands, the face and body. Sign language recognition system works on five essential parameters; hand shapes, hands movement, hand and head orientation, hand and head location and facial expressions. Out of these five parameters the foremost fundamental requirement is hand shapes. The second most important parameters are hand movements, hand and head orientation and their location in the video frame. These parameters can be incorporated into a sign language recognition system to improve the performance of the system by segmenting and tracking hands and head of the signer in the video. Segmented hand shapes and their locations are used to design a feature vector. The created feature vector is used to train the neural network. In the last decade extensive research had been done to understand hand gestures from the movement of human hands [1,2]. Compared to tracking techniques to rigid objects [3], tracking non-rigid objects such as bare human hands are a very complex and challenging task.

Going back to early days of sign language, Stokoe et.al [4] showed that signs were made of basic articulatory units. Initially, they were referred as cheremes. At present, they are being referred as phonemes, in resemblance with words in spoken language.

The major difficulty in sign language recognition, compared to speech recognition, is to recognize [5] different communication attributes of a signer, such as hands and head movement, facial expressions and body poses simultaneously. All these attributes are considered as a good recognition system.

The second major problem, faced by sign language recognition system designers, is tracking the signer in the clutter of other information available in the video. This is addressed by many researchers as signing space [6]. A sign language space can be created with entities such as humans or objects stored in it around a 3D body centered space of the signer. The entities are executed at a certain location and later referenced by pointing to the space. Another major challenge faced by researchers is to define a model for spatial information, containing the entities, created during the sign language dialogue.

Additional difficulties arise in the form of background in which signer is located. Most of the methods, which have been developed so far, are used simple backgrounds [7,8,9,10] in controlled set-up such as simple backgrounds, special hardware like data gloves[11,12,13], restricted sets of actions, restricted number of signers, resulting different problems in sign language feature extraction.

Our research is directed towards recognizing gestures of Indian Sign language [33] under real time conditions such as varied lighting, different backgrounds and make the system independent of the signer. For segmentation and tracking, we have employed active contour models, which are capable of segmenting and tracking non rigid hand shapes and head movements. Features are extracted from segments of hand and head shapes including tracking information, in the form of hand locations from each video frame. This feature vector is used to train the artificial neural network, the network outputs and determine the voice command relevant to the trained sign.

Active contours, popularly known in the research community as 'snakes', is an active research area with applications to image and video segmentation predominantly to locate object boundaries. They are also used for video object tracking applications. Active contours come under the category of model based segmentation and tracking methods, which are giving good results in the last few years [14,15,16]. The active contours were first introduced by Terzopoulos [17,18]. The basic idea, behind active contours, is to start with a curve anywhere in the image and move the curve in such a way that it sticks to the boundaries of the objects in the image. Thus it separates the background of the image from its objects. The original "snakes" algorithm was prone to topological disturbances and is exceedingly susceptible to initial conditions. However, with invention of level sets, [19] topological changes in the objects are automatically handled.

For tracking non-rigid moving objects, the most popular model based technique, is active contours [20]. Active contours can bend themselves based on the characteristics of the objects in the image frame. Previously created active contour models were only capable of tracking an object in a video sequence with a static background [21]. Jehan-Besson et.al [22] tracked objects in successive frames by matching the object intensity histogram using level sets. Though, the tracking error was in minimum level, it increased the object experiences intensity transformation due to noise or illumination changes. Almost all active contours methods discussed above were suffered from the problems related to cluttered backgrounds, lacking in texture information and occlusions from other objects in a video sequence. These problems can cause the performance of tracking non-rigid objects to decrease drastically.

Segmentation and tracking results are fused together to form a feature matrix, which is unique from the methods proposed for sign language feature extraction in [23, 24,25]. Finally, an artificial intelligence approach is proposed to recognize gestures of sign language as in [26,27,28,29]. Hidden Markov Models (HMM) were used extensively for classification for sign language [30, 31, 32].

The rest of the paper is organized as follows: sect. 2 introduces the proposed system for recognizing gestures of Indian Sign Language (INSLR), Sect.3 presents proposed segmentation and tracking using active contours, Sect.4 gives the creation of feature matrix, Sect.5 pattern classification using neural network, Sect.6 discusses experiments under various test conditions, and the last Sect.7, briefly concludes with future scope of our research.

II. Proposed System For Gesture Recognition

The proposed system has four processing stages namely, hand and head segmentation, hand and head tracking, feature extraction and gesture pattern classification. Fig. 1 shows the block process flow diagram of the proposed method.

From the Fig. 1 we can comprehend the overall working process of the system. Video segmentation stage divides the entire frames of images of the signers into hands and head segments of the signer. Hand and head tracking module gives the location of each hand and head in each video frame. Shape features, extracted from segmentation and location information from tracking, are fused into a feature matrix for each sign and are optimized before saving to the database. This process is repeated for all the available signs using a video of a signer. The sign language database can be restructured to add new signs into the database.

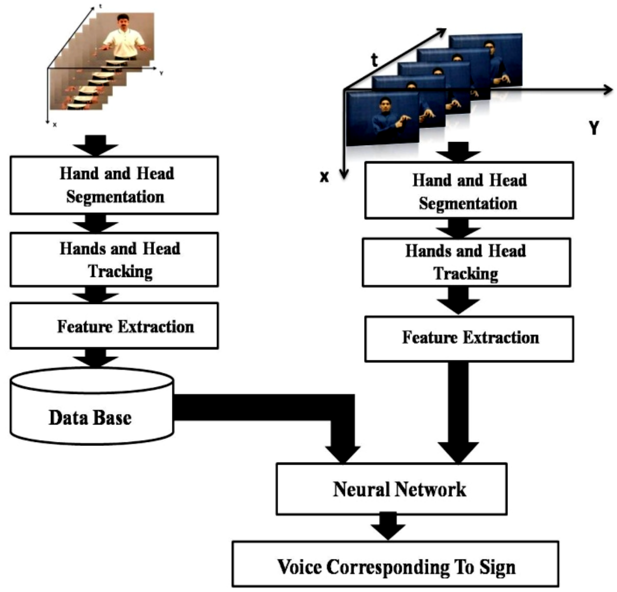


Figure 1. Sign Language Recognition System Architecture

Testing is done by giving a sign video sequence that is not initially trained and does not have a feature vector in the database. Unlike American Sign Language or British sign language, Indian sign language does not have a standard database that is available for use. Hence, we have created our own video database of Indian signs in collaboration with Indian Deaf Society [33] and shanthi ashram school for deaf children. We have created 480 signs of alphabets, numbers, words and sentences by multiple signers of Indian sings and list is growing. A total of twenty different signers volunteered under different conditions with a total number of 4800 gesture videos for a total of 480 signs. To test the robustness of our system we also used sign language videos from YouTube, which are under different video backgrounds and in real environments taking the total video database count to 6800. Neural network is trained using the feature vector from the database and tested using the feature vector from the input of the system. The output of the neural network results in a voice and text message to the corresponding input sign.

III. Segmentation and Tracking

A. Active Contours-Level Sets

James A. Sethian and Stanley Osher [34] presented boundaries of $\mho(x)$ implicitly and model their propagation using appropriate partial differential equations. The boundary is given by level sets of a function $\phi(x)$. In level sets method, the interface boundary is characterized by a zero level set function $\phi(x) = 0, where\ \phi: \mathbb{R}^{\mathbf{2}} \rightarrow \mathbb{R}$. $\mho$ is defined for all values of x,

$$\mho = \{\phi(x) = 0, x \in \mathbb{R}^2\} \qquad (1)$$

The sign of $\phi(x)$ defines whether x is inside the contour $\mho$ or external to it. The sets $\mho^{Int} = \{x, \phi(x) \leq 0\}$ and $\mho^{Ext} = \{x, \phi(x) > 0\}$. The level set evolves based on the curvature κ of the image objects and assuming the curve moving towards the outward normal $\vec{\boldsymbol{n}}$ defined in terms of parameter ϕ as

$$\boldsymbol{\kappa} = \nabla.\left[\frac{\nabla\phi}{|\nabla\phi|}\right] \;\; and \;\; \vec{\boldsymbol{n}} = \frac{\nabla\phi}{|\nabla\phi|} \tag{2}$$

Usually the curve $\mho$ evolution is a time dependent process and the time dependent level set function $\phi: \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ as $\mho(t) = \{\phi(x,t) = 0, x \in \mathbb{R}^2\}$. One way to solve is to approximate spatial derivatives of motion and update the position of the curve over time. This method of solving the level sets is prone to unsteadiness due to erroneousness detection of position of the curve.

A different approach was proposed from the theory of level sets in [34]. Start with a zero level set $\phi(x) = 0$ of higher dimension function and entrench the object curvature. Initializing the level set function ϕ at $t = 0$, we have

$$\phi(x, t = 0) = \pm d \tag{3}$$

Where $'d'$ is signed distance function (sdf) from x to the curvature of the image object. If d is a positive value, x is outside the object boundary. if d is a negative value, x is inside the object boundary. The goal is to construct an equation for evolution of $\phi(x,t)$ to embrace the object boundaries from zero level set $\phi(x) = 0$. we can propagate the zero level set $\phi(x) = 0$, by solving a convection equation containing the velocity field v, which propagates all the level sets as

$$\phi_t + v.\nabla\phi = 0 \tag{4}$$

The motion is in normal velocity of the curve which is given by eq.2 as $v = F\vec{\boldsymbol{n}} = F\frac{\nabla\phi}{|\nabla\phi|}$. Inserting in eq.4 we get a level set equation of the form

$$\phi_t + F|\nabla\phi| = 0 \tag{5}$$

Eq.5 is a type of Hamilton-Jacobi equation. The speed term F is dependent of object curvature $\boldsymbol{\kappa} = \nabla.\left[\frac{\nabla\phi}{|\nabla\phi|}\right]$, which can be formulated as

$$F(\boldsymbol{\kappa}) = F_0 + F_1(F) \tag{6}$$

Eq.6 drives the contour to level out with the high curvature regions together with a diffusion term.

B. Hands and Head Segmentation and Tracking

This section presents the video object segmentation, tracking of hands and head simultaneously from a range of video backgrounds under different lighting conditions with diverse signers. A video sequence is defined as a sequence of image frames $\mathrm{I(x, y, t)}: \mathrm{D} \to \mathrm{R}$ where the images change over time. Alternatively, a succession of image frames can be represented as $I_{(n)}$ where $0 \leq n \leq \infty$. The basic principle behind our proposed segmentation and tracking technique is to localize the segment and track one or more moving objects of the n^{th} frame from the cues available from previous segmented and tracked objects in frames $I_{(1)}, I_{(2)}, \dots\dots\dots.. I_{(N)}$ such that subsequent contours $\Gamma^1, \Gamma^2 \dots\dots\dots \Gamma^N$ are available. The sign videos are composed of many moving objects along with the hands and head of the signer. Signer's hands and head are considered as image foreground denoted by $I^f_{(n)}$ and rest of the objects are as the image background $I^b_{(n)}$, for the image $\mathrm{I_{(n)}}$ in the video sequence. Foreground contour of the hands and head might be denoted by $\Gamma^f_{(n)}$. Our proposed video segmentation algorithm segments hands and head of signers using colour, texture, boundary and shape information about the signer given precedent understanding of hand and head shapes from $\boldsymbol{I^f_{(n-1)}}$ and $\boldsymbol{I^b_{(n-1)}}$..

C. Colour and and Texture Extraction Modules

Colour plays a vital role in segmentation of complex images. There are various colour models, but the RGB colour model is most commonly used for video acquisition. Processing on a RGB colour frame increases the size of feature vector and thereby making the tracking process lethargic. Instead of working with gray scale images which store intensity information about each pixel, we used each of the three colour planes separately to extract each colour feature vector. This allows us to work with only one plane at a time depending on the background colour level. We choose manually the colour plane which highlights the human object from a background of clutter. Once a colour plane is identified, texture features are calculated using coorelogram of pixel neighbourhood [35,36]. Texture is an irregular distribution of pixel intensities in an image. Allam.et.al [37] established that co-occurrence matrix (CM's) produce better texture classification results than other methods. Gray Co-occurrence matrix (GLCM) presented by Haralick.et.al [38] is most effectively used algorithm for texture feature extraction for image segmentation.

Let us consider a color plane of our original RGB video. The R color plane is now considered as $M \times N$ R coded 2D image. The element of co-occurrence matrix $C_{d,\theta}$ defines the joint probability of a pixel x_i of R color intensity $\boldsymbol{r_i}$ at a distance d and orientation θ to another pixel x_j at R color intensity $\boldsymbol{r_j}$.

$$C_{d,\theta}(r_i, r_j) = Pr\{I(z_1) = r_i \wedge I(z_2) = r_j : |z_1 - z_2|_\theta = d\} \tag{7}$$

where $|z_1 - z_2|_\theta$ gives the distance between pixels. For each co-occurrence matrix, we calculate four statistical properties: contrast (C), correlation (CO), energy (EN) and homogeneity (H) defined as follows

$$CO = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)}{\sigma_i\sigma_j} C_{d,\theta}(r_i, r_j) \tag{8}$$

$$C = \sum_{i,j} |i-j|^2 \, C_{d,\theta}(r_i, r_j) \tag{9}$$

$$EN = \sum_{i,j} \left(C_{d,\theta}(r_i, r_j)\right)^2 \tag{10}$$

$$H = \sum_{i,j} \frac{C_{d,\theta}(r_i, r_j)}{1 + |i-j|} \tag{11}$$

The logic, in which the above parameters are used as texture feature, is described as follows; the contrast represents inertia and variance of the texture the correlation term gives correlation between different elements of GLCM, CO is high for more complex textures. From eq.41 μ_i and μ_j mean values along i and j directions.σ_i and σ_j represent variances, energy term describes the consistency of the texture, homogeneity is taken as a measure of coarsenesses of the texture. We used four different orientations $\theta = \{0,45,90,135\}$ and two distance measures $d = (1,-1)$ for calculation of GLCM matrix.

Finally, a feature vector $f^{vect}(\boldsymbol{x})$ is produced which is a combination of any one or all of the color planes and texture vector. Thus $\boldsymbol{f^{vect}}(\boldsymbol{x}) = \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}) \ldots\ldots\ldots. f_n(\boldsymbol{x})\}$ the feature vector contains color and texture values of each pixel in the image. This is a five dimension feature vector containing the first vector for any of the three color planes and the next four vectors for texture. We can also use all three color planes to represent color, and then the feature vector becomes a seven dimension feature vector.

Most of the image sequences contain many classes of colour and texture. Hence, we classify them as background and foreground pixels using Mixture of Gaussians (GMM) [39] clustering algorithm.

Given n- dimensional feature vector $\mathrm{f^{vect}(x)}$ and the GMM algorithm classify this n-vector into $k-$categories. Thus, we model the object and background in an image frame using two mixture of Gaussian probability density functions (pdf): M^{ho} and M^b. The probability of finding a value of feature vector in the reference frame i.e. the first frame $I^{(1)}$is given by

$$p_{in}\left(\frac{f^{vect}(x)}{O_{hh}}\right) = \sum_{i=1}^{C_{obj}} w_{i,t} * p\left(\frac{f^{vect}(x)}{O_{hi}}\right) \quad for\ x \in Object \tag{12}$$

$$p_{out}\left(\frac{f^{vect}(x)}{B_{bk}}\right) = \sum_{i=1}^{C_{bck}} w_{j,t} * p\left(\frac{f^{vect}(x)}{b_{bj}}\right) \quad for\ x \in Background \tag{13}$$

where $p_{in}(f^{vect}(x)/O_{hh})$ and $p_{out}(f^{vect}(x)/B_{bk})$ represent mixture probabilities of human object and background. O_{hh}represents object hand and head portions. B_{bk} represents video background. $p(f^{vect}(x)/O_{hi})$ and $p\big(f^{vect}(x)/b_{bj}\big)$terms represent probabilities of multivariate guassians of object and background. C_{obj} and C_{bck} denote the number of clusters in the object and background mixtures. Weight terms $w_{i,t}$ and $w_{j,t}$ gives the estimate of the mixing parameters of two mixture models C_{obj}andC_{bck}. Maximum Likelihood Estimation (MLE) [40] algorithm is used to calculate the parameters $O_{hh} = \big(O_{hi}, w_{i,t}\big) \forall\ 1 \le i \le C_{obj}$ and $B_{bk} = \big(b_{bj}, w_{j,t}\big) \forall\ 1 \le j \le C_{bck}$. The MLE consists of estimating the unknown parameters of two mixtures by minimizing the likelihood functions

$$\lambda^0 = \prod_{x \in \boldsymbol{\mathcal{I}}_{\boldsymbol{f}}^{(\boldsymbol{1})}} p_{in}(f^{vect}(x)/O_{hh}) \tag{14}$$

$$\lambda^b = \prod_{x \in \boldsymbol{\mathcal{I}}_{\boldsymbol{b}}^{(\boldsymbol{1})}} p_{out}(f^{vect}(x)/B_{bk}) \tag{15}$$

The estimation problem comes down to solve the following equations

$$O_{hh} = arg \max_{O_{hh}}[\lambda^0] \tag{16}$$

$$\overline{B_{bk}} = arg \max_{B_{bk}}[\lambda^b] \tag{17}$$

We will assume, at this point, is that objects in the video sequence pretty much remain same when compared to background, that varies due to camera movement or changes in background scenes. This can be taken care by periodically updating the background clusters with some threshold if the changes in consecutive background frame cross the specified threshold.

For a pixel x in any frame $I_{(n)}$ where $n > 1$, the above procedure is used to calculate the object mixture probability $\overline{p_{in}(x)}$ and background mixture probability $\overline{p_{out}(x)}$ on a neighbourhood of pixels which is eight in this paper. To track the objects in the current frame $I_{(n)}$, the initial contour position in previous frame $I_{(n-1)}$ is moved towards new object boundaries in current frame $I_{(n)}$ by minimizing the following energy functional of the level set

$$E^{CT}(I^f) = \iint_{\mathcal{I}^f} D\big(p_{in}, \overline{p_{in}(x)}\big)dx + \iint_{\mathcal{I}^b} D\big(p_{out}, \overline{p_{out}(x)}\big)\, dx \tag{18}$$

Where p_{in} and p_{out} are object and background PDF's in the reference frame $I_{(1)}$. $\overline{p_{in}(x)}$ and $\overline{p_{out}(x)}$ are object and background PDF's in the current frame $I_{(n)}$. $D\big(p_{in}, \overline{p_{in}(x)}\big)$and $D\big(p_{out}, \overline{p_{out}(x)}\big)$ represent KL-distance symmetric which is computed between two PDF's by the following equation

$$D\big(p_{in}, \overline{p_{in}(x)}\big) = \frac{1}{2}\left[\int_z p_{in}(z) * \log(p_{in}(z)/\overline{p_{in}(z)})dz + \int_z \overline{p_{in}(z)} * \log(p_{in}(z)/\overline{p_{in}(z)})dz\right] \tag{19}$$

Where the integrals are calculated over domain z. The energy function in eq.18 is used to track the object boundaries in $I_{(n)}$by calculating the local statistics of the pixels within the object separating them from background. We can implement this by assigning pixel x to object in the current frame when $D\big(p_{in}, \overline{p_{in}(x)}\big) > D\big(p_{out}, \overline{p_{out}(x)}\big)$ and to the background otherwise.

D. Object Boundary Module

In the earlier module the focus was on extracting region information with the objective of minimizing the object contour energy function, which segments the objects of interest in the video sequence. But poor lightning can impact image region information in a big way. Hence, we extracted boundary edge map of the image objects which only depends on image derivatives. The way out would be to couple the region information in the form of colour and texture features to boundary edge map to create a good tracking of image objects.

We define the boundary B^o as pixels that are present in edge map of the object. The boundary pixels can be calculated by using gradient operator on the image. To align the initial contour Γ^0 from previous frame to the objects in the current frame to pixels on the boundary we propose to minimize the following energy function

$$E^B(I^f) = \int_{arc(Length\ og\ Obj)} g\big(\mathcal{B}^O(x)\big)dx \tag{20}$$

Where $arc(Length\ of\ obj)$ is the length of the object boundary. The function g is an edge detection function. The boundary energy reaches to a minimum when the initial contour aligns itself with the boundary of the objects in the image. The minimization of energy in eq.20 also results in a smooth curve during evolution of the contour [41].

E. Shape Information Module

Before Even with colour, texture and boundary values of pixels in the image, the greatest challenge comes when object pixels and background pixels share the same colour and texture information. This happens because we are trying to track non- rigid objects that are hands and head of the signer along with finger positions and orientations, which change frequently in sign video. The problem influences the propagation of contour and results in meagre tracking of non-rigid objects in video sequences. The active contour can be influenced by giving information regarding the shape of the object computed from the previous frames. The method followed in [42] is used to construct the influence of shape of non-rigid objects in the image sequence. As for the first frame $I_{(1)}$, where prior shape information is not available, we just used the region and boundary information for segmentation. The segmented objects in frame one are used to initialize contours in the next frame for tracking. For $I_{(n)} \forall\, n = 1$ the tracking of $I_{(n)}$ is given by the level set contour $\Gamma^{(n)}$ which minimizes the energy function

$$E^S = \min_T \int_{\Gamma^{int}} \phi_0(T^{-1}x)dx \tag{21}$$

where the minimum is calculated over Euclidian Similarity Transformations $T: \mathbb{R}^2 \to \mathbb{R}^2$ which is a combination of translational and rotational parameters. Minimizing over groups of transformations to achieve rigid object interactions was proposed by chan and zhu [33]. We propose to use a non-rigid shape influence term in this paper. Now let us recollect Γ, which indicates the active contour and Γ^0 be the active contour for the shape from the first frame. Let $\phi: \mathbb{R}^2 \to \mathbb{R} = \phi(x)$ be a level set distance function associated with contour Γ, and $\phi_0: \mathbb{R}^2 \to \mathbb{R} = \phi_0(x)$ is a level set function with contour Γ^0 from first image frame $I_{(1)}$. Let x be a pixel in the image space R fixed, $\phi(x) = \phi(\mho; x)$ is actually a function of contour Γ. The initial contour Γ aligns itself with the object contour Γ^0 in the first frame that is the initial contour for the next frame in the video sequence coming from the contour in the previous frame. Hence the shape interaction term proposed in this paper has the from

$$E^S(I^f) = \int_{\Gamma^{int}} \phi_0(x)dx \tag{22}$$

Thus by applying shape energy to the level set we can effectively track hands and head in sign videos and we could differentiate between object contour modifications due to motion and shape changes.

F. Level Set Energy Function For segmentation and Tracking Combining Colour, Texture, boundary and Shape Modules

By integrating the energy functions from colour, texture, boundary and shape modules we formulate the following combined energy functional of the active contour as

$$E_{int}(I^f) = \alpha E_{CT}(I^f) + \beta E^O(I^f) + \gamma E^S(I^f) \tag{23}$$

where α, β, γ are weighting parameters that provide stability to contribution from different energy terms. All terms are positive real numbers. The minimization of the energy function is done with the help of Euler-Lagrange equations and realized using level set functions. The resultant level set formulation is

$$\frac{d\phi^n(x,t)}{dt} = \big((\alpha R_{CT}(\phi^n) + \beta R^b(\phi^n) + \gamma R^S(\phi^n)\big)\|\nabla\phi^n\| \tag{24}$$

where

$$R_{CT}(\phi^n) = -D\big(p_{in}, \overline{p_{in}(x)}\big) + D(p_{out}, \overline{p_{out}(x)}) \tag{25}$$

$$R^b(\phi^n) = g(B^O(x)) + \nabla . g(B^O(x))\left[\frac{\nabla\phi}{|\nabla\phi|}\right] \tag{26}$$

$$R^S(\phi^n) = \phi_0(x) \tag{27}$$

The numerical implementation of eq.24 is computed using narrowband algorithm [43]. The algorithm approximates all the derivatives in eq.24 using finite differences. The level set function is reinitialized when the zero level set clutches the boundary of the object in the image frame.

IV. Feature Vector

After Successful segmentation and tracking the feature extraction, a feature vector is being created, that is stored in the database at training stage as templates or can be used as inputs to pattern classifiers such as Hidden Markov Models (HMM), Neural Networks, Fuzzy Inference Systems or Decision Trees, which have limited memory. Hence, it becomes important to design a feature matrix that optimizes the performance of the classifier.

The feature matrix f^{Mat} derived from a video sequence is a fusion of hand and head segments representing shapes in each frame along with their location in the frame. The shape information for an n[th] frame is presented as ϕ^n, which is a binary matrix equal to frame size. Tracking active contour results in location of hands and head contours in each frame and giving their location $(x^{(n)}, y^{(n)})$ values.

Temporal dimension is mostly neglected during feature extraction in video processing. In other words, video is considered as an extension of images. In addition to 2D nature of image, the temporal dimension is managed using a technique called temporal pooling [44]. Here, the extracted features are temporally pooled into one feature value for the whole video sequence as illustrated in fig. 2. Temporal pooling is largely done by averaging. Other simple statistical techniques are employed such as standard deviation, median or minimum/maximum.

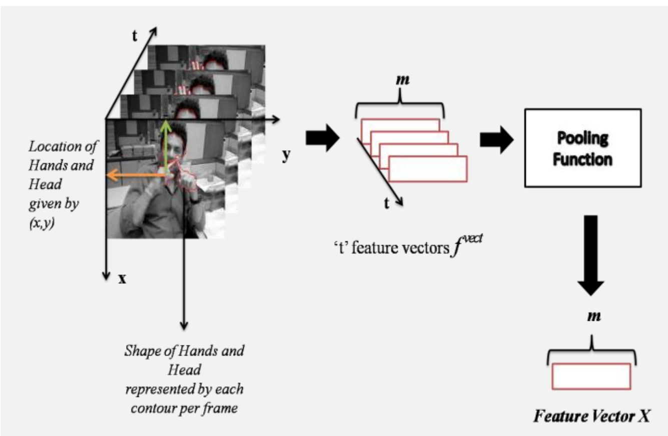


Figure 2. Feature Vector Design using Temporal Pooling

For an n[th] frame in the video sequence, the first row of feature vector f^{vect} consists of pixel values with the shape of active contour i.e. the pixels representing segmented head and hand shapes. The second and third row consists of $(x^{(n)}, y^{(n)})$ location information of those pixels that are segmented. The feature vector for each frame is a three dimensional vector f^{vect} representing shape and location information about each segmented hand and head contours. For an entire video in a sign language recognition system, each f^{vect} of three dimension becomes f^{Mat} feature matrix stacked in the form shown in fig. 2. A row in the feature matrix represents 3D f^{vect} for a particular frame and it is represented with time't'.

Temporal pooling is engaged to reduce the dimensionality of the feature vector for a particular video sequence. Averaging on each row of the f^{Mat} with respect to number of pixels in a frame leave us with a new reduced feature vector f^{Nvect}, which uniquely represents a particular video sequence. Finally each row in the new feature matrix f^{NMat} consists of feature vector representing each sign video sequence, stored in the database as templates.

V. PATTERN RECOGNITION-NEURAL NETWORK

For handling large data such as in our system it is customary to use solutions that are fast enough without losing its accuracy. One of the few systems that can handle our large data matrices is Feed-Forward artificial neural network (ANN). Artificial neural network (ANN) is extensively used in detection of cancer [45], classifications [46], face recognition [47] and finger print recognition [48] ,to name a few. The feature matrix f^{NMat} is given as input for training the feed-forward neural network shown in fig. 3. Generally, a Feed-Forward neural network is a combination of three layers of neurons: input layer, hidden layer and output layer. The neurons in these layers are activated by using a nonlinear sigmoid activation function. Let $x_{i,j}(itr)$ where $1 \le i \le N$ *and* $1 \le j \le M$, be the input to the neural network derived from feature matrix f^{NMat}. Where M and N denote the number of columns and rows of f^{NMat}. *itr* is the number of iterations called Epochs in neural network terminology. The neural network outputs are denoted by $y_{i,j}(itr)$ where $1 \le i \le N$ *and* $1 \le j \le M$.

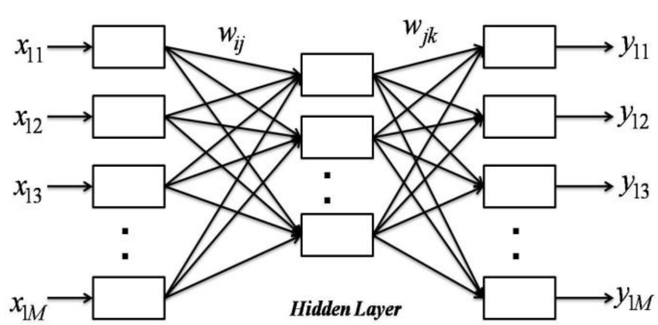


Figure 3. Neural Network Architecture

The back propagation algorithm follows the following steps in determining the output as discussed in [49]. The output of j[th] unit in the first hidden layer is calculated from the values in the hidden layer as

$$y_j(itr) = f\left(\sum_{i=1}^{M} x_i(itr) w_{ij}(itr) - \Theta_j\right) \tag{28}$$

Where M is number of neurons in the hidden layer, w_{ij} is the weight vector connecting i[th] unit in between hidden layer and input layer and j[th] unit in between hidden layer and output layer. Θ is the threshold value of j[th] unit in hidden layer. $f(\bullet)$ is the activation function, which we choose as sigmoid function, defined as

$$f(\bullet) = \frac{1}{1 + e^{-x_{ij}(itr)}} \tag{29}$$

The output layer output is calculated as

$$y_k(itr) = f\left(\sum_{j=1}^{M} x_{jk}(itr) w_{jk}(itr) - \Theta_j\right) \tag{30}$$

where M is the number of neurons in the j^{th} layer for k^{th} neuron.

During back propagation process the error gradient from the output layer is calculated from the relation

$$\Delta_k(itr) = y_k(itr)[1 - y_k(itr)]e_k(itr) \tag{31}$$

where $e_k(itr)$ is the error in the output layer given as

$$e_k(itr) = y_{Tk}(itr) - y_k(itr) \tag{32}$$

where $y_{Tk}(itr)$ is the desired or targeted output of the system. Adjustment of the weights of the neurons is done iteratively using the equation

$$w_{ij}(itr+1) = w_{ij}(itr) + \mu y_j(itr) + \Delta_k(itr) \tag{33}$$

Where μ is adaptive learning rate of the neural network during training stage. After $(k-1)^{th}$ iteration and the k^{th} feed-forward propagation, let the output vector be $y_{ij} = [y_{1j}, y_{2j}, y_{3j}, ..., y_{ij}]^T$ and the target vector $y_{Tk} = [y_{1k}, y_{2k}, y_{3k}, ..., y_{Tj}]^T$. The expression for adaptive learning rate is

$$\mu = \frac{\alpha}{e^{(-\frac{1}{j}\sum_{i=1}^{M}(y_{ij}-y_{Tk})^2)}} \tag{34}$$

where the value of α is. $0 \le \alpha \le 1$ The learning rate is controlled by error obtained during iterations.

The error gradient for the hidden layer is calculated using the relation

$$\Delta_j(itr) = y_j(itr)[1 - y_j(itr)]\sum_{k=1}^{i}\Delta_k(itr)w_{jk}(itr) \tag{35}$$

The above discussed procedure is run iteratively to reach the expected output.

VI. EXPERIMENTS AND RESULTS

The first frame $I_{(1)}$ is segmented by calculating the feature vector consisting of single colour plane (R or G or B) and texture information along with boundary edge map. Then the proposed method is applied to the remaining frames of the video sequence. The segmentation and tracking result of the previous frames is used as a mask or initial contour for the current frame along with location of the mask and so on. The energy minimization function in eq.23 is employed to process the level sets and to produce an optimal contour for segmentation and tracking for each frame. In all video sequences initial contour is cropped manually that can be placed near to the object of interest.

In the first experiment we started with a video sequence which is shot in a lab using web cam with dark background and with an additional constraint that signer should also wear a dark shirt. This video sequence is part of the database we have created for sign language recognition project. Our Sign language database consists of 351 signs with 8 different signers. Figure 4 (frame size 320×480) shows our segmentation and tracking algorithm with values of $\alpha = 0.3, \beta = 0.5$ and $\gamma = 0.2$. The object and background clusters are made of three and two clusters.

The experiments are performed in R colour plane. As such we can do it any colour plane or with colour videos. The problem with full colour sequences pertaining to sign language is that the sign language videos contain large sequence of frames with lot of information to be extracted. Fig. 4(a) shows four frames from a video sequence of the signer performing a sign related to English alphabet 'X'. This simple sign contains 181 frames. Fig. 4(b) shows the results obtained from our proposed method. The inclusion of prior shape term in the level sets segments the right finger shape in spite of being blocked by the finger from left hand. This segmentation and tracking result will help in good sign recognition. Fig. 4(b) and 4(c) shows the effectiveness of the proposed algorithm against the Chan-Vese (CV) model in [50].

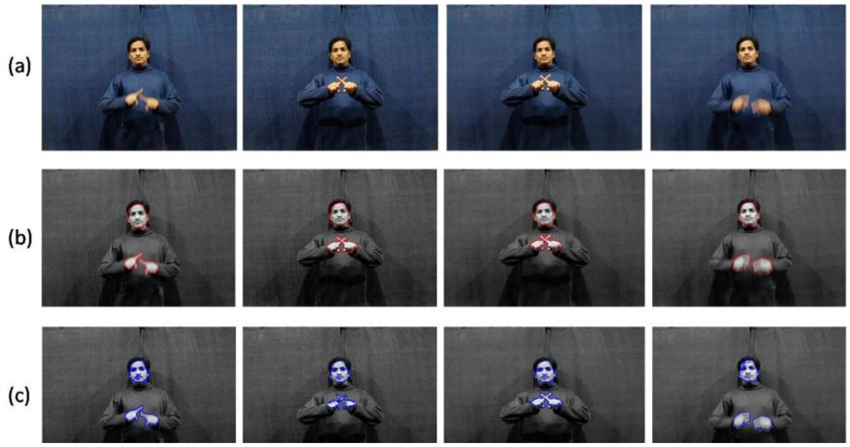


Figure 4. Experiment one showing our proposed segmentation algorithm on sign language videos under laboratory conditions. Frames 10, 24, 59,80 are shown. Row (a) shows four original frames. Row (b) shows the results from proposed algorithm and Row (c) Shows results of CV algorithm in [50].

The results of segmentation and tracking are fused into a feature vector and a database is created using 351 signs with 4 different signers. The size of feature matrix f^{NMat} for the neural network is 1404×80. The Neural network can classify 351 signs. Hence the target vector of neural network is 351×80.

The neural network, used, is feed forward and back propagation network, which has 1404 instances of input and 351 instances of output neurons and 652 neurons in the hidden layer as shown in fig. 5. The weights and bias for each neuron are initialized randomly and network is ready for training. The training process requires a set of examples of proper network behavior, network inputs and target outputs. During training the weights and biases of the network are iteratively adjusted to minimize the network performance function, which in case of feed forward networks, is mean square error (MSE).

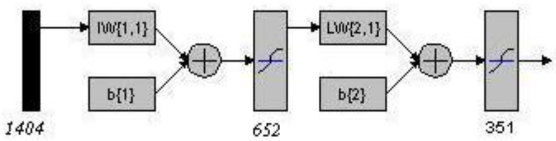


Figure 5. Neural Netwok used for testing sign videos with simple backgrounds.

The neural network shown in fig. 5 is trained for simple backgrounds using four sets of sign videos for each gesture consisting of different signers. The epochs versus error graph for training the network is shown in fig. 6.

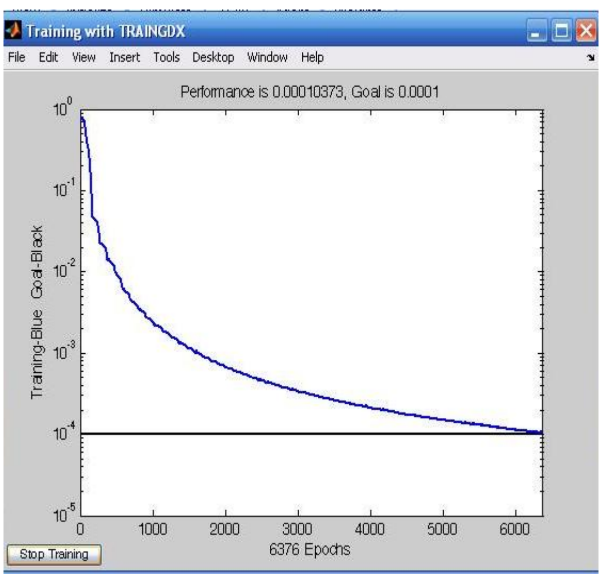


Figure 6. Neural Netwok Training graph for simple backgrounds with four samples per gesture.

The recognition rate is defined as the ratio of the number of correctly classified signs to the total number of signs:

$$\text{Recognition Rate}(\%) = \frac{\text{Perfectly Classified Signs}}{\text{Total Signs Used}} \times 100 \quad (36)$$

TABLE I. RESULTS OF TESTING USING GESTURES WITH SIMPLE BACKGROUNDS

Data	Total Number of Samples	Correctly Recognized samples	Performance Rate of the Neural Network (%)
Number of input neurons : 1404 Number of hidden neurons: 652 Number of output neurons: 351 Activation Function: sigmoid Learning rate: 0.05 Momentum Factor: 0.9 Error Tolerance: 0.0001 Number of training samples used: **1404** Number of testing samples: **2106**			
Training With 4 Samples/Gesture			
Training	**1404**	**1404**	100%
Testing	**2106**	**2009**	95.39%
Total	3510	122	97.23%

The recognition rate is defined as the ratio of the number of correctly classified signs to the total number of signs:

$$\text{Recognition Rate}(\%) = \frac{\text{Perfectly Classified Signs}}{\text{Total Signs Used}} \times 100 \quad (36)$$

Table. I gives the results of testing the proposed network. For simple backgrounds the neural network is classifying signs with a recognition rate of 97.23%.

We observe occlusions of hands and head very frequently in sign language videos. Most sign language recognition systems insist that the signer should face the camera directly to avoid occlusions of hands largely. This problem is solved using our proposed level set method. It is shown in fig. 7. We initialized contour for only right hand of the signer in the first frame.

With the left hand coming in the path of right hand as can be observed from the original sequence in figure 7(a), it's difficult to segment and track the shape of right hand. But the results in figure 7(b) show the segmentation and tracking of right hand only with occlusions from left hand. This breakthrough will help in designing sign language systems with utmost robustness.

In the next experiment, the supremacy of our proposed technique is shown, where the video sequence contains fast moving objects in contrast to hand and head movements. The video sequence is shot on an Indian road and in the natural environment. Fig. 8 shows the original sequence in column (a) along with the results of CV model [50] in column (b) and our method in column (c).

Table. I give the results of testing the proposed network. For simple backgrounds the neural network is classifying signs with a recognition rate of 97.23%.

Figure 7. Showing the influence of prior shape knowledge. Here only the occluded right hand is segmented and tracked. Column (a) shows original image sequence and column (b) the segmentation and tracking result.

Observation of third row exposes the disadvantage associated with CV method without prior shape term in energy level set for tracking. In this video (see Fig. 8) the background object suddenly appears in the frame to which the method in [50] technique provides much resistance and the final segmentation and tracking result include the suddenly appeared object. But providing prior shape information along with object colour, texture and boundary edge map proves the strength of our method. Also, we get the unwanted tracks in the form of leaves of trees in the background for left hand of the signer in row three for the method used in [50], which is not an issue with our method.

Using the sequences in fig.7 and fig.8, which represents a more real time scenario for testing our proposed system, we created a data base of around 110 signs. These videos are taken under different background clutters like in fig. 7 and fig. 8. We have chosen two samples of each sign for training the network and two signs of similar gesture for testing the network. The proposed network is shown in fig. 8 containing 220 input neurons, 181 hidden neurons and 110 output neurons. The video sequences, used here, contain continuous gestures with cluttered backgrounds, such as offices, on road, restaurants, and videos from [33].

Figure 8. Frames of a sign video sequence on a Indian road and under natural environment. Column (a) is original sequence of frames 39, 54,79 and 99. Column (b) method used in [50] and Column (c) our proposed tracking method.

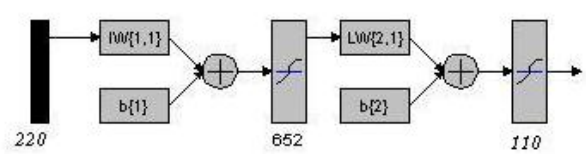


Figure 9. Neural Netwok Architecture for cluttered backgrounds with two samples per gesture.

The neural network is trained with a feature matrix extracted from sign videos with complex backgrounds. The training graph for the network in fig.9 is shown in fig.10. We observe that even though the input has two sets of signs for each gesture, it took more number of iterations to train due to occlusions and cluttered background.

Table. II gives the results of testing the proposed network. For complex and variable backgrounds, the neural network is classifying signs with a recognition rate, of which is slightly less than the recognition rate for simple backgrounds i.e. 95.68%.

In the third and final experiments, the proposed neural network is trained using a combination of two signs from simple backgrounds and two signs from complex backgrounds for each sign. A total of 200 signs are trained four from each gesture. The neural network employed for this task consists of 200 input neurons, 652 hidden neurons and 50 output neurons to classify 50 signs and convert them in to audio and text. The network architecture for this job is shown in fig. 10.

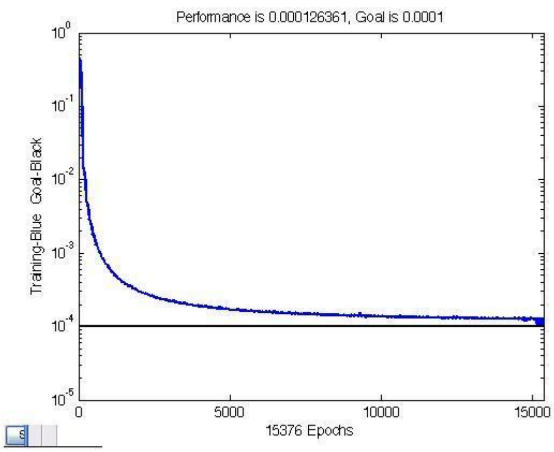


Figure 10. Neural Netwok Training graph for cluttered backgrounds with two samples per gesture.

TABLE II. RESULTS OF TESTING, USING GESTURES WITH CLUTTERED BACKGROUNDS

Number of input neurons : 220 Number of hidden neurons: 652 Number of output neurons: 110 Activation Function: sigmoid Learning rate: 0.05 Momentum Factor: 0.9 Error Tolerance: 0.0001 Number of training samples used: **220** Number of testing samples: **220**			
Data	Total Number of Samples	Correctly Recognized samples	Performance Rate of the Neural Network (%)
Training With 2 Sample/Gesture			
Training	220	220	100%
Testing	220	201	91.5%
Total	440	421	95.68%

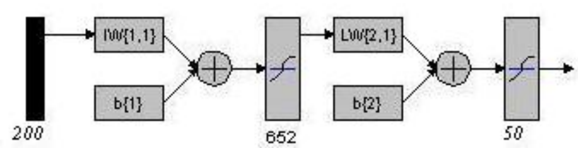


Figure 11. Neural Netwok Architecture for cluttered backgrounds with two samples per gesture along with two samples from simple backgrounds.

The neural network is trained with a feature matrix, combining the feature matrix from simple backgrounds and feature matrix from complex backgrounds. The training graph produced for this feature matrix is shown in fig. 12.

Table. III gives the results of testing the proposed network. For complex backgrounds and simple backgrounds combined the resultant recognition rate, and has slightly decreased due to shape occlusions for longer periods of time. This can be avoided by using regularly updating the shape module after losing shape information during longer occlusion period. For all our experiments, if the occlusion period is more than 10 frames, we have to reinitialize the contour. The recognition rate for our third experiment is 85%.

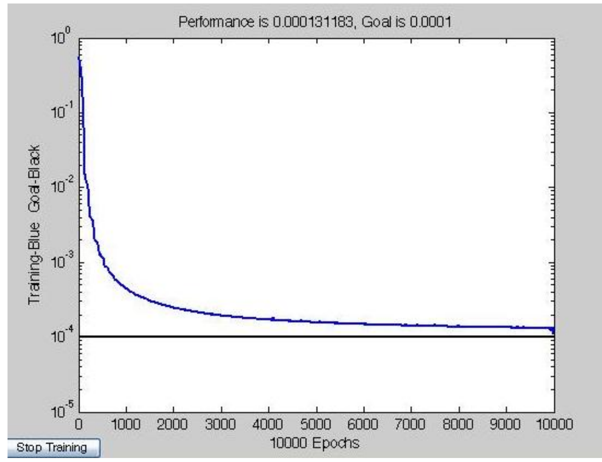


Figure 12. Neural Netwok Training graph for cluttered as well as simple backgrounds with each two samples per gesture.

TABLE III. RESULTS OF TESTING USING GESTURES WITH SIMPLE CLUTTERED BACKGROUNDS

Number of input neurons : 200 Number of hidden neurons: 652 Number of output neurons: 50 Activation Function: sigmoid Learning rate: 0.05 Momentum Factor: 0.9 Error Tolerance: 0.0001 Number of training samples used: **200** Number of testing samples: **500**			
Data	Total Number of Samples	Correctly Recognized samples	Performance Rate of the Neural Network (%)
Training With 4 Sample/Gesture			
Training	200	190	95%
Testing	200	150	75%
Total	400	340	85%

VII. CONCLUSIONS

This paper carries us slightly closer to building sign language recognition system that performs well under natural backgrounds. The proposed method combines effectively the colour, texture, boundary and prior shape information to produce an effective video segmentation and tracking of sign language videos under various harsh environments such as cluttered backgrounds, poor lighting, fast moving objects and occlusions. The colour and texture information is extracted statistically by creating a feature vector and classifying each pixel in the frame to object and background pixel. Boundary information is provided by divergence operator along with the curvature of the object under consideration. Including shape information from the previous frame it is done a whole lot of difference to the level set minimization to segment correctly and track effectively the occulted hand from other hand and also head some times. We have effectively demonstrated by experimentation of the proposed method by applying it to video sequences under various conditions. Nevertheless, there are challenges which are to be addressed to apply this method to continuous sign language recognition systems to carry out the recognition in real time. The recognition rate of the overall sign language recognition system is 93% which is a good score for our proposed real time system.

The segmentation experiments are run on an Intel core i3 2.5GHz processor with 3GB RAM using MATLAB software. The average running time was 3 frames per second with around 40 level set iterations per frame. The resolution of the video is kept to minimum along with colour information which in this paper was restricted to only one colour plane. This is to reduce the computation time which otherwise increases for higher resolutions and full colour processing. The speed of the algorithm can be increased by using fast numerical schemes for determining the level sets as given in [51].

REFERENCES

[1] Imagawa I et. al, 'Recognition of local features for camera-based sign language recognition system', Proc. ICPR'00, pp849-853, 2000.

[2] P.V.V.Kishore, P.Rajesh Kumar, E.Kiran Kumar & S.R.C.KIshore . "Video Audio Interface for Recognizing Gestures of Indian Sign Language" *International Journal of Image Processing (IJIP), CSC Journals,* Vol. 5, No.4, pp479-503, Sep 2011

[3] Yang M, Ahuja N, 'Recognizing hand gesture using motion trajectories',CVPR'99, pp892-897, 1999.

[4] W. Stokoe, D. Casterline, and C. Croneberg. A Dictionary of American Sign Language on Linguistic Principles. *Gallaudet College Press, Washington D.C.*, USA, 1965.

[5] U. R. Wrobel. Referenz in eb¨ardensprachen: Raum und Person. Forschungsberichte des Instituts f¨ur Phonetik und Sprachliche Kommunikation der Universit ¨at M¨unchen, 37:25–50, 2001.

[6] C. Vogler and D. Metaxas. A framework for recognizing the simultaneous aspects of american sign language. Computer Vision & Image Understanding, 81(3):358–384, March 2001.

[7] Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, and Nitin V Pujari," Finger Detection for Sign Language Recognition", Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2009), vol. I, Hong Kong, March 2009.

[8] J. Mackie ,B. McCane, "Finger Detection with Decision Trees", University of Otago, Department of Computer Science.

[9] Claudia N¨olker, Helge Ritter, "Detection of Fingertips in Human Hand Movement Sequences", Gesture and Sign Language in Human-Computer Interaction, Proceedings of the International Gesture Workshop, Published by Springer, 1997, pp. 209-218.

[10] E. Stergiopoulou, N. Papamarkos, "Hand gesture recognition using a neural network shape fitting technique", Engineering Applications of Artificial Intelligence (ELSEVIER), vol. 22,pp.1141–1158, 2009, doi: 10.1016/j.engappai.2009.03.008.

[11] Mehdi, S.A.; Khan, Y.N. "Sign language recognition using sensor gloves**"** Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02. Volume 5, pp18-22, Nov. 2002.

[12] W.Kadous. Recognition of Australian Sign language using instrumented glove. Master Thesis. University of New South Wales. October 1995.

[13] Tan Tian Swee, A.K. Ariff, Sh-Hussain. Salleh, Siew Kean Seng, and Leong Seng Huat, "*Wireless Data Gloves Malay Sign Language Recognition System"* Proceedings of the international conference on Information and computing Systems, 2007, ICICS'07,pp56-59.

[14] Collins, R.T., Liu, Y., Leordeanu, M. (2005) "*Online selection of discriminative tracking features"*. IEEE Trans. Pattern Anal. Mach. Intell. Vol.27 No.10 , pp1631–1643.

[15] Ginmo Chung and Luminita Vese. (2003) "Image segmentation using a multilayer level set approach.Technical Report" 03-53, UCLA.

[16] R. Malladi, J.A. Sethian, and B.C. Vemuri. (1995) *"Shape modeling with front propagation: A level set approach*". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.17 No. 2,pp158-175.

[17] D. Terzopoulos and K. Fleischer (1988) "*Deformable models"*. The Visual Computer, Vol.4 No.6, pp306-331.

[18] D. Terzopoulos, J. Platt, A. Barr, and K. Fleischer.(1987) "*Elastically deformable models*" In Comp.Graphics Proc., pages 205-214. ACM Press/ ACM SIGGRAPH.

[19] Osher, S., Sethian, J. (1988) Fronts propagating with curvature-dependant speed: algorithms based on Hammilton-Jacobi formulations.J. Comput. Phys. Vol.79 No.1, pp12–49.

[20] L. Brkthes, P. Menezes, E Lerasle and J. Hayet "Face tracking and hand gesture recognition for human-robot interaction" Pmcwdingr of the 2004 IEEE InternatIona1 Confennee on RobotIco 6 Automation New Orleans. April 2004.

[21] Mansouri.A.R, Region tracking via level set pdes without motion computation. IEEE Trans. Pattern Anal. Mach. Intell. 2002, Vol.24, no.7, 947–961.

[22] Jehan-Besson, S, Barlaud, M.,Aubert,G.: Detection and tracking of moving objects using a new level set based method. Proceedings of IEEE International Conference on Pattern Recognition, Barcelona, Spain, 3–8 September, 2000, pp. 7112–7117.

[23] Newby, G. (1993). Gesture recognition using statistical similarity. In Proceedings of virtual reality and persons with disabilities.

[24] Birk, H., Moeslund, T. B., & Madsen, C. B. (1997). Real-time recognition of hand alphabet gestures using principal component analysis. In Proceedings of the 10th Scandinavian conference on image analysis.

[25] Rangarajan, K., & Shah, M. (1991). Establishing motion correspondence. CVGIP: Image Understanding, 54, 56–73.

[26] Dorner, B., & Hagen, E. (1994). Towards an American sign language interface. Artificial Intelligence Review, 8(2–3), 235–253.

[27] Fels, S. (1994). Glove-TalkII: mapping hand gestures to speech using neural networks—an approach to building adaptive interfaces. PhD thesis, Computer Science Department, University of Toronto.

[28] Fukushima, K. (1989). Analysis of the process of visual pattern recognition by the neocognitron. Neural Networks, 2, 413–420.

[29] Haykin, S. (1999). Neural networks: a comprehensive foundation. Prentice Hall.

[30] Charniak, E. (1993). Statistical language learning. Cambridge: MIT Press.

[31] Starner, T. (1995). Visual recognition of American sign language using hidden Markov models. Master's thesis, Massachusetts Institute of Technology.

[32] Seymore, K., McCallum, A., & Rosenfeld, R. (1999). Learning hidden Markov model structure for information extraction. AAAI 99 workshop on machine learning for information extraction.

[33] www.deafsigns.org , Indian sign Language Dictionary.

[34] M.Kass, A Witkin, D Terzopoulos ,"Snakes:active Contour Models", *Int. J. of Computer Vision,* 1987,pp 321-331.

[35] R. Malladi, J.A. Sethian, and B.C. Vemuri. *"Shape modeling with front propagation: A level set approach*". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.17 No. 2, 1995, pp158-175.

[36] Tuceryan, M., Jain, A, Texture analysis. In: Chen, C.H., Pau, L.F., Wang, P.S.P. (eds) The Handbook of Pattern Recognition and Computer Vision, 2nd edn. chap. 2.1, 207–248.World Scientific Publishing, Singapore.

[37] Allili, M.S., Ziou, D.,(2006), *"Automatic colour-texture image segmentation by using active contours"*. In: Proceedings of 1st IEEE International Workshop on Intelligent Computing in Pattern Analysis/Synthesis, Xi'an, China, 26–27, LNCS 4153, pp. 495–504.

[38] S. Allam, M. Adel, P. Refregier, "Fast algorithm for texture discrimination by use of a separable orthonormal decomposition of the co-occurrence matrix," Applied Optics, vol.36, 1997, pp.8313–8321.

[39] R. M. Haralick, K. Shangmugam, I. Dinstein, (1973) *"Textural Feature for Image Classification,"* IEEE Trans on Systems, Man, Cybernetics, 3(6), pp.610—621.

[40] McLachlan, G., Peel, D, Finite Mixture Models, Wiley Series in Probability and Statistics (2000).

[41] A Dempster, N. Laird, and D. Rubin. "Maximum likelihood from incomplete data via the EMalgorithm," *Journal of the Royal Statistical Society*, vol. 39, 1977, 1-38.

[42] Allili, M.S., Ziou, D.: Automatic colour-texture image segmentation by using active contours. In: Proceedings of 1st IEEE International Workshop on Intelligent Computing in Pattern Analysis/Synthesis, Xi'an, China, 26–27, August 2006, LNCS 4153, pp. 495–504.

[43] Cremers, D., Soatto, S.: A pseudo-distance for shape priors in level set segmentation. In: Proceedings of 2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision, Nice, France, 13–16 October 2003, pp169–176.

[44] K.Seshadrinnathan and A.Bovik, "Motioin Tuned Spatio-Temporal quality assesment of quality videos" IEEE Trans on Image Processing, vol.19, No.2,pp335-350,Feb.2010.

[45] Braik, M., Sheta, A., "A New Approach for Potentially Breast Cancer Detection Using Extracted Features and Artificial Neural Networks", Journal of Intelligent Computing Vol. 2, No. 2, June 2011, pp.54-71.

[46] Honjun Lu., Rudy S., huan lui., 1996, Effective data mining Using Neural Networks, IEEE Transactions On Knowledge and Data Engineering, Vol. 8, No. 6, , pp. 957-961.

[47] Radha, V., Nallammal., N, October 19-21, 2011, "Neural Network Based Face Recognition Using RBFN Classifier" In Proceedings of the World Congress on Engineering and Computer Science, Vol. I, San Francisco, USA.

[48] Darwis, Y. and, Sheta, A., 2008, " Minutiae Extraction for Fingerprint Recognition" In Proceedings of the Fifth IEEE International Multi-Conference on System, Signal and Devices (SSD'08), Amman, Jordan.

[49] Michael, N., 2005 Artificial Intelligence: A Guide to Intelligent Systems, Pearson Education Limited.

[50] T. Chan and L. Vese(2001) "*Active contours without edges*". IEEE Transations on Image Processing,Vol. 10 No.2,pp266-277.

[51] Weickert, J., K ¨ uhne, G.: Fast methods for implicit active contour models. In: Osher, S., Paragios, N.: (eds) Geometric Level Set Methods in Imaging, Vision and Graphics, chap. 3, pp. 44–57. Springer, Heidelberg.

AUTHORS PROFILE

P.V.V.Kishore (SMIEEE'07) received his M.Tech degree in electronics from Cochin University of science and technology in the year 2003, and currently pursuing PhD at Andhra University College of engineering in Department of ECE from 2008. He is working as research scholar at the Andhra university ECE department. He received B.Tech degree in electronics and communications engineering from JNTU, Hyd. in 2000. His research interests are digital signal and image processing, computational intelligence, human computer interaction, human object interactions. He is currently a student member of IEEE.

Dr. P.Rajesh Kumar (MIEEE'00, FIETE'09) received his Ph.D degree from Andhra University College of Engineering for his thesis on Radar Signal Processing in 2007. He is currently working as associate professor at Dept. of ECE, Andhra University College of engineering, Visakhapatnam, Andhra Pradesh. He is also Assistant Principal of Andhra University college of Engineering, Visakhapatnam, Andhra Pradesh. He as produced numerous research papers in national and international journals and conferences. He has guided various research projects. His research interests are digital signal and image processing, computational intelligence, human computer interaction, radar signal processing.

RFID: The Big Player in the Libraries of the Future

Priya. S[2]
Part Time Ph.D. Scholar
School of Management Studies
Anna University of Technology, Coimbatore

Dr. Venkatesh. J[1]
School of Management Studies
Anna University of Technology, Coimbatore
Coimbatore, India

Thenmozhi.S[4]
Part Time Ph.D. Scholar in Computer Applications
Kongu Engineering College
Erode, India

Aarthy.C[3]
Full Time Ph.D. Scholar, School of Management Studies
Anna University of Technology, Coimbatore
Coimbatore, India

Balasubramanie.P[5]
Department of Computer Science and Engineering,
Kongu Engineering College
Erode, India

***Abstract*— This paper suggest the idea of developing an automatic Radio Frequency Identification (RFID)-based system for library search and to observe the current literature to define whether current technology and developed RFID-based systems are supportive for building that. To fully understand its key points, implementations, risks, and benefits, the fundamentals of radio frequency are to be recognized and revised. The paper recognizes key benefits and risks of radio frequency identification through analysis of library literature. To make library systems efficient and successfully operational RFID solutions can be used to reduce operating costs through reducing the labor costs, improving automation, refining tracking and tracing, and inhibiting the loss of resources under any conditions. The projected automated RFID-based system is a unique idea by itself. Due to the fact that enhanced organization of books and resources becomes possible, resources are not lost, stealing is avoided, and customers are served on time and correctly.**

Keywords- Readers; Libraries; Solicitations; Automation; Tracing.

I. INTRODUCTION

Imagine a library where each book has its own place on a particular shelf. Customers like to take a book off the shelf, look it through and then put it away and take another one and do the same till the right book is found. There are some people who place back the book in the right place but many people either leave the books in some corner of the library or return them to the wrong places. This later situation is hard to identify and can make librarian's frightening. Until books find their rightful homes, with limited number of staff that a library has, days will pass and in some cases more time is required. When a client needs a book for research or even relaxation and needs it instantly needed then tracking down the missing book starts - the librarian gets called in, and on occasion organization hears about the case as well. In this case, tracking and locating is not easy and finding the book somewhere in a very large public The Electronic Library or university library is not a simple job. Do labeling book and materials in the library help such situations? Curran and Porter have projected and defined a library model that utilizes Radio Frequency Identification (RFID) to augment and speed up the current customer book search and identification procedures. The hardware used in the design and implementation of the patterns are a laptop to host the server, a router to create the wireless network, a PDA to host the solicitations, RFID tags and an RFID reader to carry out the RFID message.

The user can search for a book on the shelf by entering the book information in any of the search norms text boxes and pressing the Search button. The system has to search the catalogue to return the appropriate book. The functionality and benefits accessible by the RFID systems match the needs and areas of enhancement for libraries. The development and evaluation of the library application has proven that RFID can be successfully integrated into library systems. 2 per cent of libraries in the USA use RFID technology and 8 per cent global. All RFID retailers in the library market bid a product with anti-collision (the ability to read several tags concurrently). The actual speed at which this can be achieved, and total number of tags that can be read does vary substantially

II. RFID: DYNAMIC AND INERT TAGS

Radio frequency credentials are a term used for technologies employing radio waves for identifying discrete items automatically. The most common way is storing a serial number recognizing a product and related information on a chip attached to a probe. RFID is used much related to bar codes. It is intended to track items without demanding a line of sight. To read a bar code its lines had to stay in sight of the scanner to recognize product correctly. Radio Frequency Identification is a site determination technology that has been receiving a lot of profitable attention in recent times, particularly in the areas of asset tracking and supply chain management. RFID is not a new spectacle. It has been around for years. It was used primarily for proximity access control. Thereafter, it was evolved-to-be-used in supply chain tracking, toll barrier control, and even protecting vehicles. There are four types of tags in industry: Inert; Dynamic; Semi Inert; Semi Dynamic. In a market study conducted in August 2010, it was found that Dynamic RFID systems have a sequence of valuable features including:

1) *Improved reliability because of high enactment;*
2) *Improved security/access control including stealing decline;*
3) *The skill to link tags together in software for custodianship; the ability to automate identification and location by removing social intervention;*
4) *Improved data reliability because of correctness and consistency;*
5) *Better read accuracy and longer read choices; and*
6) *Increased data transmission rate.*

Inert tags get their energy from a remote RFID reader. A Dynamic tag uses a set for both the chip and the transmission of data on the antenna. Semi-Inert tags use a small onboard set to power the chip. Semi-Dynamic tags use the battery for operating the antenna but the chip depends on the Radio Frequency (RF) energy from the reader. The lives of Dynamic tags are limited while of Inert tags are infinite. Dynamic tags are heftier than the inert tags and more costly as well. Table I identifies variances between Dynamic and Inert RFID tags.

III. RFID Frequencies and Overheads:

RFID is a method for sending and getting data without any links occurs between the interrogators and tags using electromagnetic waves. RFID tags can hold more data than data carrier systems such as bar code system. RFID systems work at a number of diverse frequencies including 125 KHz, 13.56 MHz, 2.45 GHz and 5.8 GHz and for UHF 860-950 MHz Low frequency tags work along 120 KHz-140 KHz frequencies while high frequency tags work along the 13.56 MHz radio frequencies. Ultra high frequency (UHF) tags work along the 850-900MHz range. Low rate tags are less costly and use less power related to other kinds of tags, though high and ultra-high tags have better ranges and transfer data faster. However, these two types of tags use more power and are classier. Table I condenses the areas of solicitation of various chips for different frequencies. Knowing that manufacturers target specific industries once begin to produce a tag and try to expand to other areas as well it is important to know the producers and users at the same time.

Table I: Difference Between Dynamic And Inert Rfid Tags

Disputes	**Dynamic RFID Tags**	**Inert RFID Tags**
Power Bases	In-house	Energy transmitted from the reader
Power Constancy	Unceasing	When it's in the field of reader
Communication Array	Long Range	Short range
Data Storage	Large	Small
Signal Power From Tag to Reader	High	Low
Signal Power From Reader to Tag	Low	High
Active Life	5-10 yr.'s	Infinite
Memory	2Mb	up to 16Kb

The reader conveys an electromagnetic field that "wakes-up" the tag and provides the power essential for the tag to operate. RFID readers usually cost around $1,000-2,000. The tag cost can be broken down into the following components: chip; inlay/element with antenna; assembly; and licensing. Chips cost is about $0.45 to $0.55 while inlay cost ranges from $0.04 to $0.30 and assembly from 0.04 to 0.06. In contrast with the price of one chip being estimated to reach $0.12 now is still very high. There exist few ways in helping to decrease costs significantly: utilization of a worldwide RFID chip that can be used for many applications; capable of handling multiple solicitations; reducing the cost of packaging antenna to the chip; and instinctive handling versus manual.

IV. RFID Usage in Libraries:

There are opinions on the side of the use of RFID in the libraries as well as not using it. Arguments presented in favor of RFID application in libraries are that: libraries use inert tags instead of Dynamic tags; tags used in libraries have a very short read range; the data stored on the library RFID tag does not disclose any important information; and the real threat comes from liabilities in the library's database. There are others who have anxieties about the use of RFID in the libraries of the future. Concerns about RFIDs are referred as a noise. Librarians should keep in mind that since the many of the integrated library systems and RFID vendors have a entrusted interest in endorsing the use of RFID's for amplified sales that will assist their company's profitability, so it is natural that they will minimize the security distresses. Some RFID vendors are retorting to the security criticisms and are now offering a next generation of RFID labels that reportedly have enriched security features than those labels previously available. Management of an organization must regulate whether adopting RFID will be useful to its firm or one of its companies or its processing units or not. Performance metrics should be aligned with the firm's objectives and reflect such topics as improved cash-to-cash cycles, leaner portfolios, reduced stock outs, and more precise data.

V. RFID Risks in Library

To implement RFID, gaining management vow is a big challenge. Here, management looks into the Return on Investment (ROI) to measure RFID investment before commits to its implementation. A challenge that companies face with is the high cost of implementation. To justify the acceptance of RFID technology into business, cost-benefit analysis is a must. The key risks aspects relate to the library of the future having RFID-based system with the identified competences are discussed in the following.

Initial Cost:

The cost is one of the major factors manipulating acceptance of RFID, although the production costs of RFID have condensed and Strange Technology has cut the tag price to less than $US0.20. At present, the costs of RFID adoption comprise the major speculation in hardware, application software, middleware, and tags, and the cost of assimilating the RFID-based system with the legacy systems, of consultancy fees, and of employee training. Therefore, the cost of RFID tags may endure to present a major hurdle for RFID deployment. Another cost of RFID implementation for many companies is the major investment in large scale IT infrastructure. Other costs for RFID adoption may be substantial, including the purchase of initial hardware/software, integrating RF-enabled technology into dissemination and warehousing activities and prevailing management systems, and additional maintenance costs for application renovations, readers and software, and employee training. A library with over 200,000 collections must have adequate funds to invest in the purchase of tags alone although, the costs for the reader, marginal equipment and application software will be even more of the problem. With the economic crisis in most cities and countries around the world it is most likely that a great number of public libraries cannot own an independent and full RFID-based system soon.

A. Skilled RFID Workers:

A survey conducted by the Computing Technology Industry Association revealed that 80 per cent of the responding companies said that there are not ample numbers of skilled RFID workers. About two-thirds of respondents pointed out that training their employees to become expert in RFID is the biggest challenge they faced in order to prosper in the RFID market. Hence, gaining management obligation in implementing RFID will be a big issue considering this difficult hurdle.

B. Access Rate:

There are many factors that can impact the read/write efficiency of RFID. Some of those are: metal, mist, distance, and improper positioning of antennas. A point, well justified, is that when the distances among the tags are very close, interference between them may be made or invalid access may occur. For example, if one client is taking books to initiate the loan process under self-check in/out equipment (RFID reader), and another patron is standing too close to the first patron, the reader doing the self-check in/out may sense the tags of books which are held by the wrong patron. Libraries that highlight patron services might not be able to sustain such check in/out errors. It is observed that even with extremely strict requirements for baggage identification in airports the identification rate is still not at 100 per cent.

C. Patron Privacy and Issues:

A big anxiety is the offensive of privacy due to the use of RFID that has been a major issue fueling the hostility from consumer protection organizations. Privacy activists are concerned about tracking customers while other researchers talk about the RFID standards and privacy. Some customers will be opposed to any RFID related system due to the alleged privacy issues surrounding it.

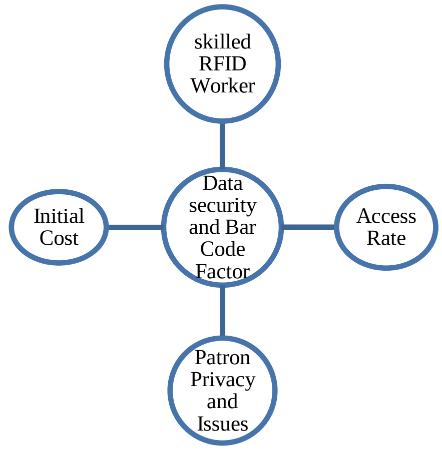


Fig1. Risk Factors in Library

All patron events such as reading, surfing, and action behavior will be detected by readers which are installed in libraries. But library book tags will not contain any customer information and most executions will be using Inert read only tags that can only be read from small distances. There may be some customers that reject to accept the new RFID-based system at first.

Some customers might be unaware or reluctant to use any new technology for not putting their own or family security and privacy at risk. Therefore it may become essential to have someone at the library to have free lectures for them or give away leaflets on the topic to customers discussing the issues. However, it must be said that RFID patron cards open many opportunities for offering a wealth of information about individuals. Security based on insignificance, and exclusive protocols, are not enough. This may become an even greater problem if there is conjunction in the protocols and technology that will allow all these systems to readily talk to one another.

D. Data Security:

The biggest issue that must be taken into attention when a new technology is addressed and employed is the level of the security that it may provide or needs in order to keep executive data at the safe level. Hence, organization data security policies must be inspected to ensure that security of customer data are not compromised at any price.

E. Barcode Factor:

Barcodes, very popular around the world, for almost any product, are also employed by both library systems, and book publishers. Therefore, the fame of barcodes is an issue since almost all book retailers around the world are set up with the bar coding system. Between 5-10 trillion bar codes are printed every year, and around 5billion bar code scans occur every day. Bar codes are low-cost, standardized, and, in some cases, are already achieving a satisfactory performance level. Although the popularity of bar codes is not believed to be a restrictive, their popularity has not helped spread larger-scale RFID arrangement. Taking these key issues into thought, we will notice barcode factor is a big player in the spread of the RFID.

VI. RFID Benefits in Library

RFID-based systems include much remuneration that exceeds their disadvantages and risks. In the following sections some of these benefits are briefly discussed.

A. Information Management:

Barcodes, book cards, and magnetic strips can all be combined into one RFID tag. With a suitable RFID tag that has memory for recording information and then delivering it to the system, bibliographical records and movement status can be stored on that. This system is capable of locating the location of material in the library when it is essential to be located. Videotapes and diskettes are unable to use magnetic strips to impose entrance guard because demagnetization will destroy the data on the material. Because RFID tags do not use demagnetization to modify data, they can use tags to manage magnetic materials the same as the books.

B. Circulation:

With the use of RFID-based system effective operation of library and circulation begins. It brings the opportunity of not scanning barcodes one by one at all. Having RFID readers by the librarians' sides they can check in/out books in a short time and constantly. The confirmation of materials and controlling the entrance guard for not taking out materials by someone without checking is easy and highly manageable. Library cards will include RFID tags. Readers will detect and fetch information from library cards when customers enter a library, and it will be spread to a backend system process. After that, the front desk shows lent materials, overdue books, reserve materials, and other circulation status on the monitor about this patron. Librarians depend on these messages to deliver service.

C. Portfolio:

Batch processing can also apply to libraries to perform inventory or shelf reading. Take hand-held readers to sweep shelves, for instance - readers can instantly detect all of the collection within this range, including anomalous situations such as books put on the wrong shelf. Libraries can use RFID to replace barcodes and obtain several advantages: reduce lines at the front desk; decrease repeatable tasks; increase interaction with customers; extend internal security; lower the cost of deploying and managing collections; secure collections, checking and accepting automation; and raise the efficacy of inventory and arrangement.

D. Aid in Searching and Positioning:

The application of RFID in industrial circles provides material flow management similar to library circulation. A library automation system can only query about check in/out situations. But, it is not probable for that to determine where it is if it is not at the right place in the shelf. When, as a part of the RFID-based system, a reader is installed on each gate in the library, once a client takes a book or other item and enters another room, the system will sense who took it and where it was left. Then, this information will be passed to the automation system to record the position of the material.

E. Data Precision and Reliability:

The effective arrangement of RFID has a potential to quickly provide accurate and reliable data that surpasses the bar coding or manual competences available today. This can have a major impact, mainly in busy libraries such as university libraries and public libraries in populated areas.

F. Application Statistics for Serials:

There is not a proper and precise method for calculating the reading rates of magazines that are placed on the journal racks in the library. Often, many journals are not used at all and some are used far more than anticipated. General library systems cannot record when periodicals are used in open shelves and read in the libraries. Usually, a patron's response to the intended questionnaires is not very precise. However, utilizing the detection scope of RFID, it is possible to determine the location of such materials from the periodical rack and the readers. If readers remain unnoticed on one tag for a while, this shows that the magazine was taken off the shelf and is being read.

VII. The Big Picture RFID-Based System

The functionality and aids offered by the RFID systems match the needs of library systems and the enhancements to be made in other situations. The development and assessment of the library applications proved that RFID could effectively be integrated into library systems. However, no single solution can maximize the value of RFID but rather a combination of technology needs to be taken into deliberation to make the system functional, useful, and convenient. To obtain a necessary level of visibility, tags need to be read at many points in the system. The library of the not too distant future will be very diverse with what we have today. It will work with a combination of computer hardware/software, RFID technology, robots, conveyor belts, computer experts, and few librarians. Already many of these technologies and systems are in place in some libraries. This is a big picture of the RFID-based system for large libraries where a blend of technologies is used for wisely managing situations better and more proficiently. RFID is used to do one or more of the following important tasks, as they are necessary:

1) *read labels instead of bar codes;*
2) *define the location of materials in the library;*
3) *trace of the materials; finding whether journals and newspapers are at their place or in use;*
4) *identify the client in the library and serve them if it is necessary;*
5) *manage the inventory of books and materials; quick identification of lost materials on an hourly/daily basis;*
6) *Online data collection and consumption;*
7) *prevent from losing its materials through theft or other forms of possible loss;*
8) *use robots to take books off the shelf for patron or librarian; and*
9) *use conveyor belt to send book down the line from the shelf to patron or circulation desk.*

A good material-handling system as it is used in large depositories can be extremely helpful in the research and university libraries. Since the use of material handling systems can be quite posh for large libraries it is possible to set-up such systems in a portion of the library having large demands on the hourly and daily basis. Using RFID-based systems, it is possible to determine where the substantial is currently located and then by the help of robots the materials can be brought to the conveyor belt and routed easily to the patron or circulation desk. This is quite possible with the aid of the current technology if the working system at the library is an RFID-based structure.

VIII. MANAGEMENT EMPLOYMENT:

RFID is a flexible technology and has many Dynamic features that can be unified into different systems. It can exclusively identify any object on which a tag is attached. The tag can be read in any direction. The reader does not require line of sight for each tag and can be read through most materials. However, the type of RFID-based system described here is fully valid in libraries that serve blind and partially restricted customers.

The capability of the RFID-based system defined previously allows customers to manage a portion of their own needs without the help of librarian or assistance. However, having a robot as a part of the system would help to find materials for groups of customers that might be the only users of such libraries (for example, in a school for blind or partially disabled students). RFID-based systems can be integrated into existing library systems to recover the efficiency of the main procedures carried out in any library and increase the quality of services to be provided. The ability of RFID to exclusively identify every item is very suitable for libraries. An RFID-based library system would bring with itself many properties as are listed in the following: promptness the finding of books; improve the stock control of the library; capability to track down erroneous books; increase the security; bring ease to library system; aides disabled students, e.g. blind/partially blind; ability to self-checkout and return books without the need for human interference; improves library workflow; increase staff efficiency; enhance customer service; and improve process efficacy.

IX. CONCLUSION:

More establishments from manufacturers to government agencies, retailers and providers are announcing RFID skills into their supply chains, for asset tracking and on time management, and for the safety and supervisory purposes. However, as companies explore these noteworthy advantages through pilot programs the effects of RFID technologies on the firm network must be reflected. The types of operations that can be done by RFID technology and the benefits presented by that match the needs of libraries and the areas of improvement that management have in mind. RFID technology can help in restructuring major library procedures such as evaluation and book searches. Library employees as well as customers share the same convenience and ease of operations.

However, customers can face risks in libraries with RFID systems unless the rules protecting customers are changed and stronger one replaced them. Currently, only about 12 percent of libraries world-wide are using RFID, but this amount will soon rapidly increases as libraries understand the benefits and convenience of incorporating RFID into their processes. As already noted, the library of the future is going to be diverse from what we have today. It will be partly made conceivable with the aid of already existing technologies plus RFID-based systems. As a extension of the ideas outlined in this paper, the authors are working with a group of specialists from various fields to make such a prototype RFID-based system with just such additional technological skills a reality for libraries.

REFERENCES

[1] Boulard, G. (2005), "RFID: promise or perils?" State Legislatures Magazine, December, pp. 22-24.

[2] Butters, A. (2007), "RFID systems, standards and privacy within libraries", The Electronic Library, Vol. 25, pp. 430-439.

[3] Chao, C.-C., Yang, J.M. and Jen, W. (2007), "Determining technology trends and forecasts of RFID by a historical review and bibliometric analysis from 1991 to 2005", Technovation, Vol. 27 No. 5, pp. 268-279.

[4] Chow, H.K.H., Choy, K.L. and Lee, W.B. (2007), "A dynamic logistics process knowledge-based system - an RFID multi-agent approach", Knowledge Based Systems, Vol. 20, pp. 357-372.

[5] Curran, K. and Porter, M. (2007), "A primer on radio frequency identification for libraries", Library Hi Tech, Vol. 25 No. 4, pp. 595-611.

[6] IDTechEx (2006a), "RFID forecasts, players and opportunities 2006-2016", IDTechEx market research, URL.

[7] IDTechEx (2006b), "RFID retail case studies", IDTechEx Case Studies, (URL/.infrastructure, December 2000).

[8] Lee, D. and Park, J. (2008), "RFID-based traceability in the supply chain", Industrial Management & Data Systems, Vol. 108 No. 6, pp. 713-725.

[9] Potter, B. (2005), "RFID: misunderstood or untrustworthy?" Network Security, April, pp. 17-18.

[10] Smith, A. (2005), "Exploring radio frequency identification technology and its impact on business", Information Management & Computer Security, Vol. 13 No. 1, pp. 16-28.

[11] Spekman, R. and Sweeny, P. (2006), "RFID: from concept to implementation", International Journal of Physical Distribution & Logistics Management, Vol. 36 No. 10, pp. 736-754.

[12] Tien, L. (2005), "T08... RFID and privacy: potential and perils", program presented at: Michigan Library Association Annual Conference, Grand Rapids, MI, 25-28 October.

[13] Wray, G. (2006), "Turning into radio frequency identification (RFID)", Cognotes, Midwinter Meeting Highlights, p. 10

[14] Wyld, D. (2006), "RFID 101: the next big thing for management", Management Research News, Vol. 29 No. 4, pp. 154-173.

[15] Yu, S.-C. (2007), "RFID Implementation and Benefits in Libraries", the Electronic Library, Vol. 25 No. 1, pp. 54-64.

AUTHOR'S PROFILE:

PROFILE OF Dr .J. Venkatesh Ph.D., (MAIN AUTHOR):

Prof. Dr. Venkatesh, J received a MBA degree in 1998 and a PhD in Management in 2008, both from the University of Bharathiar, and second PhD in International Business from the University of Alagappa, Tamilnadu, India in 2011. He is working as an Associate Professor at Anna University of Technology Coimbatore, Tamilnadu, India, specialized in the field of Systems, Production, Finance and Marketing. He published many papers/articles on management and systems oriented empirical, research and conceptual papers/articles in leading referred journals and magazines. His present research focuses on mobile ad-hoc networks and its application to management. He has published two books in the area of ERP with reputed publication agency.

PROFILE OF Ms. C.Aarthy (CO-AUTHOR):

Ms. C.Aarthy is a full time research Scholar (Management studies) in Anna University of Technology, Coimbatore- Tamil Nadu. Her research focuses on enabling system oriented global business processes. She is a MBA (International Business Management) degree holder in 2011 from Anna University of Technology, Coimbatore and received Engineering (Computer Science and Engineering) Degree from Avinashilingam University for Women. She has published many articles on IT and management in leading national and international journals

PROFILE OF S.Thenmozhi (CO-AUTHOR):

Ms. S.Thenmozhi is a part time Ph.D research Scholar in Computer Applications at Kongu Engineering College Perundurai, Tamil Nadu. She is specialized in digital image processing, intranet, extranet, supply chain management and ERP. She is working as assistant professor at Ganamani Institute of Technology, and she has published many articles on IT and management in leading national and international journals.

PROFILE OF Dr.P.Balasubramanie (CO-AUTHOR):

Dr.P.Balasubramanie is working as a professor in the department of Computer Science and Engineering at Kongu Engineering College, Perundurai, Tamil Nadu. He is specialized in Image Processing, Networks and Cloud Computing. He is a dynamic and vibrant academician and researcher.

ComEx Miner: Expert Mining in Virtual Communities

Akshi Kumar
Dept. of Computer Engineering
Delhi Technological University
Delhi, India

Nazia Ahmad
Dept. of Computer Engineering
Delhi Technological University
Delhi, India

***Abstract*— The utilization of Web 2.0 as a platform to comprehend the arduous task of expert identification is an upcoming trend. An open problem is to assess the level of expertise objectively in the web 2.0 communities formed. We propose the "ComEx Miner System" that realizes Expert Mining in Virtual Communities, as a solution for this by quantifying the degree of agreement between the sentiment of blog and respective comments received and finally ranking the blogs with the intention to mine the expert, the one with the highest rank score. In the proposed paradigm, it is the conformity & proximity of sentimental orientation of community member's blog & comments received on it, which is used to rank the blogs and mine the expert on the basis of the blog ranks evaluated. The effectiveness of the paradigm is demonstrated giving a partial view of the phenomenon. The initial results show that it is a motivating technique.**

Keywords- expert; web 2.0; virtual community; sentiment analysis.

I. INTRODUCTION

Expert identification is an intricate task because experts and their expertise are rare, expensive, unevenly disseminated, hard to qualify, continuously varying, unstable in level, and often culturally isolated and oversubscribed. The expert seekers behavior further complicates this, as they typically have improperly articulated requirements, are ignorant of expert's performance history, and are not well equipped to differentiate between a good and a bad expert.

Web 2.0 [1] is an evolution from passive viewing of information to interactive creation of user generated data by the collaboration of users on the Web. The proliferation of Web-enabled devices, including desktops, laptops, tablets, and mobile phones, enables people to communicate, participate and collaborate with each other in various Web communities, viz., forums, social networks, blogs. Thus, evidently the Internet now forms the basis for the constitution of virtual communities. According to the definition of Howard Rheingold in [2], virtual communities are social aggregations that emerge from the Net when enough people carry on public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace.

Thus, the expected alliance of these active areas of research, namely, Expert Identification & Web 2.0, fills the gaps that exist in the diversified Web. In response to the identified need to better exploit the knowledge capital accumulated on the Web 2.0 as a place for people to seek and share expertise, the operative challenge is to mine experts in virtual communities. Expert identification in virtual communities is noteworthy for the following reasons [3]. Firstly, virtual communities are knowledge pools where members communicate, participate and collaborate to gather knowledge. Intuitively, we tend to have more confidence on an expert's text. Secondly, virtual communities allow interaction of novices with experts, which otherwise in real world is tedious and expensive.

Instigated by the challenge to find experts in the virtual communities, we propose a **Com**munity **Ex**pert Mining system called the ComEx Miner system, where, firstly we build an interest similarity group, an online community which is a virtual space where people who are interested in a specific topic gather and discuss in depth a variety of sub-topics related to the topic using blogs. We further propose to mine the sentiment of the each group member's blog along with the sentiment of their respective comments. This is based on the intuition that the blogger and the commenter talk about the same topic or product, treated as feature for opinion orientation identification and if the blog's sentiment about a topic/product matches with the commenter's sentiment about the topic/product this implies that blogger's knowledge about the topic/ product is acceptable as people agree to what has been talked about in the blog. This degree of acceptance matching would then help to rank the blog and mine the expert with highest blog rank.

The main components of the ComEx Miner are:

Interest Mining Module: This module puts forward an algorithm for Interest Group construction by uncovering shared interest relationships between people, based on their blog document entries. The key point of constructing this Collaborative Interest Group is the calculations of interest similarity relations and application of the K-means clustering technique to cluster researchers with similar interests into the same group.

- ***Expert Mining Module:*** The ranking of member's blog within the built group is done on the basis of the score obtained by conjoining the blog & average comment orientation. This helps to identify the expert, the one with the highest ranking blog. The module is further divided into the following sub-modules:

- ***Sentiment Mining Module:*** The goal of this module is to perform sentiment analysis of the group member's blogs and the comments received on the respective blog. It gives the strength of the blogs and strength of their respective comments.
- ***Blog Ranking:*** Once the blog strength and comment strength has been determined; this module ranks the blogs by calculating the blog score, a metric which combines the respective blog's strength to the average comment strength. The blogs are then ranked as per the blog score and the expert is identified as the one with the highest rank.

The paper is organized into 4 sections. Section 2 highlights the related and background work pertinent to the research carried. Section 3 illustrates the proposed ComEx Miner System expounding the methodology used to mine the expert from a virtual community, followed by section 4 which demonstrates the results and analysis of proposed paradigm with the help of sample data. Finally, the conclusion lists out the key contributions of the research work presented.

II. Related Work

We seek *guidance* from people who are familiar with the choices we face, who have been helpful in the past, whose perspectives we value, or who are recognized experts [4]. Expert finding addresses the task of identifying the right person with the appropriate skills and knowledge [5]. There are various approaches related to expert identification & expertise search available in literature. Bogers et al. [6] used two methods: content based expert finding using academic papers and expert finding using social citation network between the documents and authors for finding experts. Breslin et al. [7] introduced a concept of re-using and linking of existing vocabularies in the semantic web, which can be used to link people based on their common interest. They described that a framework made by the combination of popular ontologies FOAF, SIOC, SKOS could allow one to locate an expert in a particular field of interest. Metze et al. [8] proposed a system to provide exchange of information by determining experts who can answer a given question. They provided a prototype expert finding system which enables individual within a large organization to search for an expert in certain area. Schall and Dustdar [9] addressed the problem of expertise mining based on performed interactions between people. Their approach comprised of two steps: Firstly, of offline analysis of human interaction considering tagged interaction links. Secondly, composition of ranking scores based on performance. Huh et al. [10] presented a grid enabled framework of expertise search (GREFES) engine, which uses online communities as sources for expert on various topics. They also suggested an open data structure SNML (Social Network Markup Language) for sharing community data. Smirnova and Balog [11] have argued that in real world, the notion of best expert depends on the individual performing the search. They proposed a user oriented model that incorporates user-dependent factor. It is based on the assumption that the user's preferences for an expert is balanced between the time needed to contact the expert and the knowledge value gained after .Li et al.[12] describe a method for finding expert through rules and taxonomies. They have proposed a combination of RDF FOAF facts and RuleML FOAF rules.

Punnarut and Sriharee [13] have introduced a method for finding expertise research using data mining and skill classification ontology. Zhang et al. [14] utilize an online community to find the people who may have expertise for answering a particular question. They analyze the experts by considering interactions of the people in questioning and answering the questions. Tang et al. [15] propose an expertise search system that analyses information from a web community. They use ontology to determine the correlation between information collected from different sources.

In the research presented in the paper, we intend to mine the experts in an online community which is a virtual space where people who are interested in a specific topic gather and discuss in depth a variety of sub-topics related to the topic using blogs. The conformity & proximity of sentimentally orientation of community member's blog & comments received is then used to rank the blogs and mine the expert on the basis of the blog ranks evaluated. The next section furnishes the details of the proposed paradigm.

III. The proposed ComEx Miner System

In general, an expert is someone who possesses a high level of knowledge in a particular area. This entails that experts are reliable sources of relevant resources and information. An open problem thus arises to assess the level of expertise objectively. We propose the "ComEx Miner System" that realizes Expert mining in virtual communities, as a solution for this by quantifying the degree of agreement between the sentiment of blog and respective comments and finally ranking the blogs with the intention to mine the expert, the one with the highest rank score.

Figure 1 shows the architectural overview of ComEx Miner System proposed in this research.

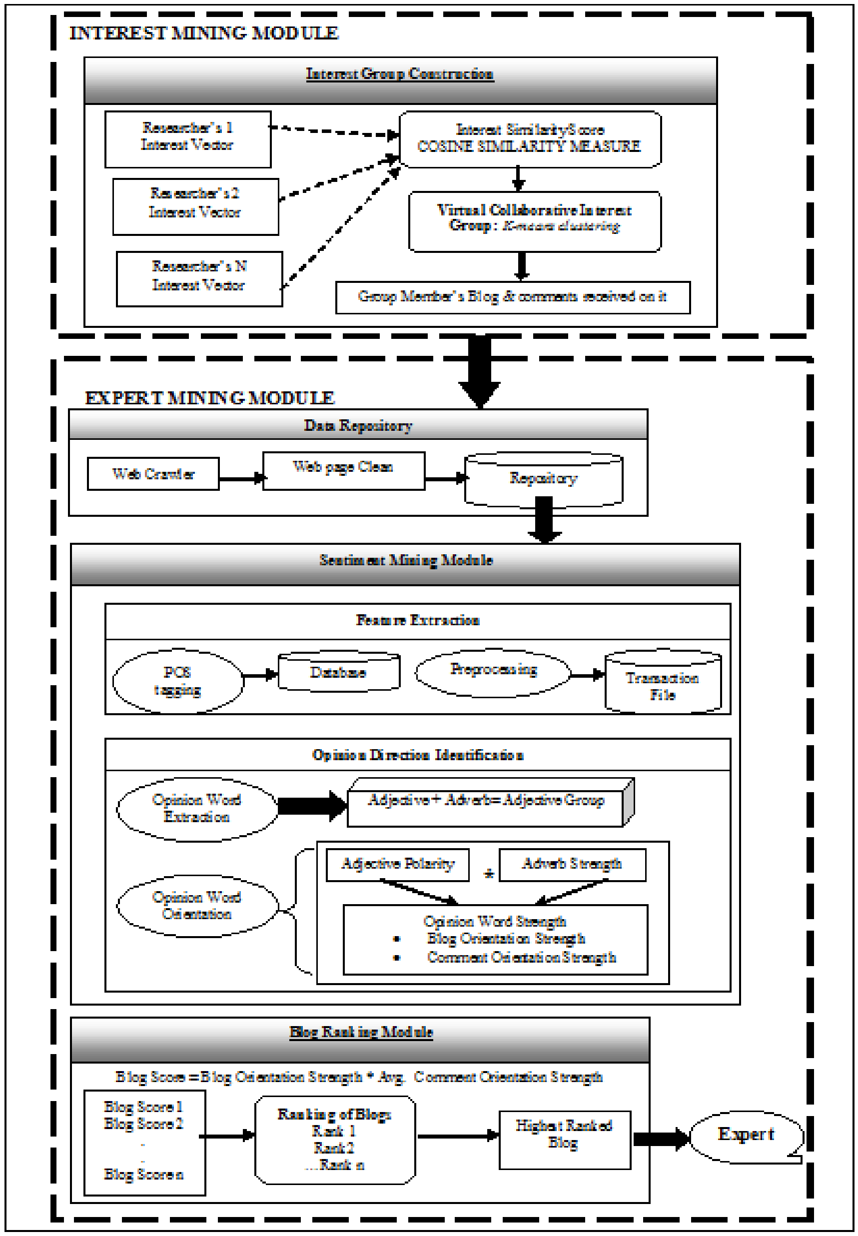


Figure 1. System Architecture of the ComEx Miner System

The following sub-sections expound the details of the ComEx Miner:

A. Interest Mining Module

In this module, we focus on the problem of discovering people who have particular interests. The Interest Group construction algorithm is based on interest similarity, which can cluster researchers with similar interests into the same group and facilitate collaborative work.

The following sub-sections expound the details of the Collaborative Interest Group construction [4]:

1) Interest Vector: Each researcher writes blog entries according to his or her interest. The interest vector of the researcher, V_i, is represented as a bag-of-words with frequently used words being assigned high weights. The interest vector is calculated by the equation described below:

and;

$$V_i = (s_{i1}, s_{i2}, s_{i3}, \dots) \tag{1}$$

And

$$s_{ik} = ef_i(w_k) \times \log\left(\frac{N_u}{uf(w_k)}\right) \tag{2}$$

where s_{ik} means the strength of interest in word w_k; $ef_i(w_k)$ means the number of entries containing w_k in researchers i's site; $uf(w_k)$ means the number of researchers who use w_k; and N_u means the number of researchers.

INTEREST MINING MODULE

Input: Researchers' Blog which contains their research papers

Output: Construction of Collaborative Interest Group

Steps:

1. **Interest Vector**
   - We calculate the interest vector V_i for each researcher i as follows :-

     for each researcher i
       for each frequently-used word w_k in his blog
     {find the values of entry-frequency $ef(w_k)$ & user-frequency $uf(w_k)$ calculate the strength of interest in word w_k (product of ef & log of inverse uf)}
       endfor
     endfor

2. **Interest Similarity Score**
   - We calculate the interest similarity score R_{ij} between researchers i and j using the cosine similarity of V_i and V_j

3. **Collaborative Interest Group Construction**
   We construct the collaborative interest group by using the technique of K-means clustering algorithm. It consists of two basic steps as follows:
   - We find the total number of clusters, denoted by K with the help of researcher groups so formed.
   - And then we assign points to the closest centroid by taking the proximity measure as the distance between two researchers.

2) Interest Similarity Score: A similarity score represents how similar the interests of a pair of researchers are. If researcher i and j have similar interests, their interest vectors should be similar. Thus, we calculate the similarity score between them, R_{ij}, using the cosine similarity of V_i and V_j as described below.

$$R_{ij} = \frac{V_i \times V_j}{|V_i||V_j|} \tag{3}$$

All elements of V_i and V_j are positive and thus the range of R_{ij} is 0 to 1.

3) Collaborative Interest Group Construction: Construction of an interest group is done to cluster the researchers with similar interests into the same group and facilitate collaborative work. Collaborative Interest Group Construction is done by using the technique of K-means clustering algorithm [16] where K is a user-specified parameter and it refers to the total number of clusters required.

Each point is then assigned to the closest centroid, and each collection of points assigned to a centroid is a cluster. The centroid of each cluster is then updated based on the points assigned to the cluster. We keep repeating this procedure again and again and update steps until no point changes clusters, or equivalently, until the centroids remain the same.

Basic K-means algorithm
1: Select K points as initial centroids.
2: **repeat**
3: Form K clusters by assigning each point to its closest centroid.
4: Re-compute the centroid of each cluster.
5: **until** centroids do not change.

a) Finding total number of clusters, denoted by K:

The value of K is found out by first forming the researcher groups. Total number of researcher groups formed is equal to the total number of researchers and researchers belonging to a particular group can carry out the co-operative work among themselves. Each group will have its respective threshold value which will decide the membership of a particular researcher in that group. T_i denotes the threshold for group i and is found out by averaging all the similarity scores corresponding to researcher i.

Membership criteria:

$$\begin{array}{l}\text{If } R_{ij} >= Ti\text{, then researcher j belongs to group i}\\ \text{else, researcher j belongs to some other group}\end{array} \tag{4}$$

Now, once all the researcher groups have been formed, then the value of K is equivalent to the minimum number of groups required to cover all the data points.

b) Assigning Points to the Closest Centroid: To assign a point to the closest centroid, we need a proximity measure that quantifies the notion of 'closest' for the specific data under consideration. We use the proximity measure as the distance between any two researchers, denoted by d_{ij} and is given as:

$$d_{ij} = 1 - R_{ij} \tag{5}$$

where d_{ij} denotes the distance between researchers i and j R_{ij} denotes the similarity score between researchers i and j.

c) Centroids and Objective Functions: The next step is to re-compute the centroid of each cluster, since the centroid can vary, depending on the proximity measure for the data and the goal of clustering.

Once the virtual collaborative interest similarity group is put together, the next step is to identify the expert from this group. To realize this task, the sentiment of each group member's blog along with the sentiment of their respective comments is analyzed for opinion strengths. As mentioned previously the degree of acceptance matching would then help to rank the blog and mine the expert with highest blog rank.

B. Expert Mining Module

The expert mining module is divided into three sub-modules; namely, the Data Repository module which collects the web pages from the member's blog & comments, cleans them and then stores them in the repository, the Sentiment Mining Engine that receives these cleaned web pages from the repository and then provides orientation strengths of blogs & respective comments by extracting opinion features and opinion words and the Blog ranking module which finally

EXPERT MINING MODULE

Input: Member's blog and comments on each blog

Output: Expert, one with the highest ranking blog

Steps:

1. **Data Repository**
   - Web Crawler: Crawls the member's blog and respective comments, collects them as web pages.
   - Web page Cleaning: Remove HTML tags
   - Stores in the "Repository"

2. **Sentiment Mining Module**
   - Feature Extraction: POS Tagging ; Preprocessing
   - Opinion Direction Identification:
     - ➢ Opinion Words Extraction
     - ➢ Opinion Words Orientation
       - ✓ Adjective Polarity
       - ✓ Adverb Strength
       - ✓ Opinion Word Strength

Blog Orientation Strength & Comment Orientation Strength

Blog Ranking Module

- Blog Score = Blog Orientation Strength* Avg. Comment Orientation Strength
- Rank the blogs by their Blog Score .

ranks the blogs on the basis of combined orientation strength of blog & comments to mine the expert as the one with the highest ranking blog.

The details of each of these sub-modules are given in the sections below.

1) *Data Repository:* This sub-module deals with collecting the web pages and storing them in the repository. Firstly the web crawler periodically crawls the member's blog and respective comments to collect them as web pages. Thereafter, these pages are cleaned up to remove the HTML tags and then are organized properly to be stored in the " Repository".

2) *The Sentiment Mining Module:* This sub-module deals with providing the actual orientation strenghts of both member's blog & comments received on it. The Sentiment Mining Module receives the web pages from the repository, i.e., if there are k members in a group then k blogs and their respective n comments will be processed to finally calculate the opinion strengths using the following three steps:-

a) Feature Extraction: This is most basic and crucial step for providing orientation strength by identifying those features that the bloggers & commenters have expressed their opinion on. Such features are known as Opinion Features. We make use of both the data mining and NLP Techniques to perform the task of feature extraction. We extract the opinion features with the help of POS Tagging and Preprocessing techniques.

- *POS Tagging (Part of Speech Tagging)*

  POS Tagging is done to find out the features of the product that have been written about. As we know, features are usually noun or noun phrases in the review sentences. Therefore, we use NL Processor linguistic Parser [17] to parse each text, to split texts into sentences and to produce POS Tag for each word (whether the word is a noun, verb, adjective etc.) NL Processor generates XML output and deals only with explicit features, which are the features that occur explicitly as nouns or noun phrases. Each sentence is then saved in a Database along with the POS Tag information of each word in the sentence.

- *Pre-Processing*

  In this sub-step, a transaction file is created which consists of pre-processed noun/noun-phrases of the sentences in the database. Here pre-processing includes the deletion of stop words, stemming and fuzzy matching.

b) Opinion Direction Identification: In this step, we find out the opinion direction using the opinion features extracted in the previous step. To find the opinion direction, we will first extract the opinion words in the text and then find out their orientation strengths. It includes the following sub-steps:

- *Opinion Words Extraction*

  In this sub-step, we extract the opinion words from the text given by the member's in their respective blog & by the commenter's in their comments on that blog. Opinion words are the words that people use to express their opinion (either positive, negative or neutral) on the features extracted in the previous steps. In our work, we are considering the opinion words as the combination of the adjectives along with their adverbs. We have called them collectively as an <u>Adjective-Group</u> (AG). Although, we can compute the sentiment of a certain texts based on the semantic orientation of the adjectives, but including adverbs is imperative. This is primarily because there are some adverbs in linguistics (such as "not") which are very essential to be taken into consideration as they would completely change the meaning of the adjective which may otherwise have conveyed a positive or a negative orientation.

  For example;

  One user says, "***This is a good book***" and;

  Other says, "***This is not a good book***"

Here, if we had not considered the adverb "not", then both the sentences would have given positive review. On the contrary, first sentence gives the positive review and the second sentence gives the negative review. Further, the strength of the sentiment cannot be measured by merely considering adjectives alone as the opinion words. In other words, an adjective cannot alone convey the intensity of the sentiment with respect to the document in question. Therefore, we take into consideration the adverb strength which modify the adjective; in turn modifying the sentiment strength.

Adverb strength helps in assessing whether a document gives a *perfect* positive opinion, *strong* positive opinion, a *slight* positive opinion or a *less* positive opinion.

For example;
One user says, ***"This is a very good book"*** and ;
Other says, ***"This is a good book"***

The Algorithm used for extraction of Opinion Words is given below:

> *For each sentence in the review database*
> If (it contains a product feature, *extract all the Adjective-Group* i.e. adjectives and their adverbs as opinion words)
> *For each feature in the sentence*
> The nearby adjective and adverb is recorded as its effective opinion (which modifies the noun / noun phrase which is a product feature)

- *Opinion Words Orientation*
  In this sub-step, we find out the orientation strength of the opinion word. As our opinion word consists of adjective + adverb, therefore to find out the orientation of the opinion word, we first find out the polarity of the adjective in the opinion word and then identify the strength of its corresponding adverb in the opinion word which modifies the adjective. Finally, the product of the adjective polarity and the adverb strength gives us the strength (orientation) of the opinion word. The details for finding adjective polarity, calculating adverb strength and deducing the final opinion word strength are as follows:

a) Adjective Polarity

Here, we will identify the semantic orientation for each of the adjective. As we know, words that have a desirable state (e.g. good, great) have a positive orientation, while words that have an undesirable state (e.g. bad, nasty) have a negative orientation. In general, adjectives share the same orientations as their synonym and opposite orientations as their antonyms. Using this idea, we propose a simple and effective method by making use of the adjective synonym set & antonym set in WordNet [18] to predict the semantic orientation of adjectives. Thus, our method is to use a set of seed adjectives whose orientations we know, & then grow this set by searching in the WordNet. The complete procedure for predicting adjective polarity is given below: Procedure "*determine_ polarity"* takes the target adjective whose orientation needs to be determined and the adjective seed list as the inputs.

> 1. Procedure **determine_polarity** (target_adjective w_i , adjective_ seedlist)
> 2. begin
> 3. if (w_i has synonym s in adjective_ seedlist)
> 4. { w_i's orientation= *s*'s orientation;
> 5. add w_i with orientation to adjective_ seedlist ; }
> 6. else if (w_i has antonym a in adjective_ seedlist)
> 7. { w_i's orientation = opposite orientation of *a*'s orientation;
> 8. add w_i with orientation to adjective_ seedlist; }
> 9. end

Note:

1) *For those adjectives that Word Net cannot recognize, they are discarded as they may not be valid words.*
2) *For those that we cannot find orientations, they will also be removed from the opinion words list and the user will be notified for attention.*
3) *If the user feels that the word is an opinion word and knows its sentiment, he/she can update the seed list.*
4) *For the case that the synonyms/antonyms of an adjective have different known semantic orientations, we use the first found orientation as the orientation for the given adjective.*

b) Adverb Strength

We collect all the adverbs which are used to modify the adjectives from English lexicon. Based on the different emotional intensity expressed by the adverb, we mark the negative adverbs with a negative score and other positive adverbs with different score in different sentiment level. The score is ranging from -1 to +1 and a higher score expresses a stronger sentiment. For example, we consider that the adverb "**extremely**" has higher strength than "**more**" does, but lower than that of "**most**". Consequently, "**most**" is marked with 0.9, "**extremely**" with +0.7, and "**more**" with +0.3. Negative adverbs, such as "not", "never", "hardly", "seldom", are marked with a negative score accordingly.

c) Opinion Word Strength

It is calculated by the product of adjective polarity i.e. $P(adj_i)$ and the adverb strength i.e. $S(adv_i)$ and is given by the following formula:

$$S(OW_i) = P(adj_i) \bullet S(adv_i) \quad (6)$$

where, $S(OW_i)$ represents the sentiment of i^{th} opinion word , $P(adj_i)$ represents the polarity of i^{th} adjective and $S(adv_i)$ represents the strength of i^{th} adverb. The value of $P(adj_i)$ is either -1 or +1 and the value of $S(adv_i)$ ranges from -1 to +1.

Therefore, the strength of each opinion word i.e., $S(OW_i)$ will also lie in the range of -1 to +1.

Note:

Sometimes, there is no adverb in the opinion word, so the S(adv) is set as a default value 0.5 . When there is no adjective in the opinion word, then the P(adj) is set as +1.

d) Blog & Comment Orientation Strength: After extracting all the opinion words from the blog and finding their respective strength, the overall strength of a Blog B is calculated by averaging the strength of opinion words as shown below:

$$S(B) = \frac{1}{|OW(B)|} * \sum_{i=1}^{|OW(B)|} S(OW_i) \quad (7)$$

Entry \ Researcher	i	j	k	n	m
1.	W_1, W_{16}, W_3, W_2, W_{17}, W_9, W_{24}, W_{25}	W_{14}, W_8, W_6, W_7, W_{17}, W_{21}, W_{25}	W_{11}, W_7, W_2, W_9, W_{19}, W_{21}, W_{25}	W_{13}, W_{13}, W_{10}, W_{14}, W_{21}, W_{22}	W_{10}, W_{15}, W_2, W_{21}, W_{23}, W_{24}
2.	W_4, W_2, W_3, W_{14}, W_{11}, W_{18}, W_{21}, W_{23}	W_1, W_{16}, W_{11}, W_7, W_{18}, W_{17}, W_6, W_{23}	W_{14}, W_{10}, W_4, W_9, W_{19}, W_{20}	W_{11}, W_{13}, W_6, W_5, W_{20}, W_{21}, W_{22}, W_{25}	W_{14}, W_{16}, W_9, W_8, W_{18}, W_{23}, W_{24}
3.	W_1, W_2, W_6, W_{13}, W_{20}	W_7, W_3, W_{18}, W_8, W_{17}, W_{24}	W_9, W_{19}, W_{11}, W_{10}, W_{17}, W_{23}	W_{13}, W_{14}, W_{18}, W_{12}, W_{20}, W_{22}	W_{15}, W_{19}, W_1, W_{16}, W_{20}, W_{23}, W_{24}
4.	W_1, W_2, W_4, W_8, W_{15}, W_{10}	W_6, W_6, W_7, W_{17}, W_{22}	W_{12}, W_9, W_{19}, W_{16}, W_{24}	W_{17}, W_{13}, W_2, W_{20}, W_{21}, W_{22}	W_{11}, W_{17}, W_6, W_{15}, W_{24}, W_{25}
5.	W_1, W_2, W_5, W_3, W_{19}	W_7, W_{18}, W_{15}, W_2, W_{18}, W_6, W_{17}, W_1	W_{19}, W_9, W_{17}, W_{10}, W_{10}	W_{18}, W_7, W_{13}, W_{13}, W_{20}, W_{23} W_{24}	W_3, W_{13}, W_{22}, W_{23} W_{24}, W_{25}

TABLE I. SAMPLE BLOG ENTRIES OF 5 RESEARCHERS [3]

where; |OW(B)| denotes the size of the set of opinion words extracted from the blog and $S(OW_i)$ denotes the sentiment strength of i^{th} opinion word. As the overall strength of the blog is calculated by averaging the strength of the opinion words, therefore the strength of the review i.e. S(B) will also lie in the range of -1 to +1; where, S(B) = -1 indicates a strong negative opinion, S(B) = +1 indicates a strong positive opinion and S(B) = 0 indicates a neutral opinion.

Similar to blog orientation, comment orientation, S (C), of each comment received on a particular blog is determined. Once the orientation of every comment is known, the average comment orientation, Avg. S (C), is calculated (dividing the total comment orientation by the no. of comments).

3) Blog Ranking Module: This module takes as input the blog orientation strength and the average comment orientation strength for each member to compute the blog score.

Blog Score = Blog Orientation Strength* Avg. Comment Orientation Strength

The member's blogs are are then ranked on the basis of this computed blog score. Finally, the expert is identified as the one with the highest ranking blog score.

IV. ILLUSTRATION

To clearly illustrate the use and effectiveness of the proposed system, a case study is presented to describe a typical scenario and examine the result of each module of the approach.

A. *Interest Mining Module:* To demonstrate the Interest mining module we directly take the sample data calculations of interest vector and interest similarity from [3], where there are 5 researchers viz. i, j, k, n & m. Therefore, $N_u = 5$ and there are 5 entries in each of the researcher's blog site. The following table I shows the blog entries of each of the Researcher i, j, k, n & m.

The key point of constructing this Collaborative Interest Group is the calculations of interest similarity relations and application of the K-means clustering technique to cluster researchers with similar interests into the same group.

Interest Vector calculations: *We have the interest vector corresponding to each of the researcher i, j, k, n & m represented as V_i, V_j, V_k, V_n, V_m. The vectors using equation (2) is shown below:*

For Researcher i:

$$V_i = (0.8874, 0.4846, 1.1938, 1.3979, 0.6989)$$

For Researcher j:

$$Vj = (0.8874, 1.9897, 0.7959, 0.4845, 0.6655)$$

For Researcher k:

$$V_k = (1.9897, 0.6655, 0.1938, 0.6988, 1.9897)$$

For Researcher n:

V_n = (1.9897, 0.1938, 0.8874, 0.2907, 0.8874)

For Researcher m:

V_m = (1.1938, 0.4436, 0.3876, 0.4845, 0.1938)

Interest Similarity Score calculations: *The calculated values of Similarity Score between each of the 2 researchers:*

R_{ij} = 0.7063; R_{ik} = 0.7110; R_{in} = 0.7502; R_{im} = 0.8064;

R_{jk} = 0.6688; R_{jn} = 0.6132 ; R_{jm} = 0.7424;

R_{kn} = 0.8786; R_{km} = 0.8140; R_{nm} = 0.9169

As all the elements of both the vectors taken at a time to calculate the similarity score are positive, thus the range of similarity score is between0 to 1.

This indicates that:

The value of 1 means that the 2 researchers have exactly similar interests and;

The value of 0 means that the 2 researchers do not have any similar interests at all.

Therefore, we can say that:

The researchers n & m have almost similar interests (as R_{nm}= 0.9169, approx 1)

The researchers k & n have similar interests to a very great extent (as R_{kn} = 0.8786)

The researchers "k & m" and "i & m" have quite a lot similar interests (as R_{km} = 0.8140 and R_{im} = 0.8064)

- The researchers "j & k" and "j & n" have quite less similar interests (as R_{jk} = 0.6688 and R_{jn} = 0.6132)

a) Collaborative Interest Group Construction

We construct the collaborative interest group by using the technique of K-means clustering algorithm with the help of two basic steps. We first construct the researcher groups by finding the membership of each of the researcher using the formula defined in equation (4). This step would give us the total number of clusters required, denoted by K. And then we assign points to the closest centroid by taking the proximity measure as the distance between two researchers using the formula defined in equation (5).

CONSTRUCTION OF RESEARCHER GROUPS

1) Membership for group i

Step 1: Calculate the threshold for this group i.e. T_i

$$T_i = \frac{1}{5}\,[R_{ii} + R_{ij} + R_{ik} + R_{in} + R_{im}]$$
$$= \frac{1}{5}\,[1 + 0.7063 + 0.7110 + 0.7502 + 0.8064]$$
$$= 0.79478$$

Step 2: Deciding the members for group i

As we can see, $R_{ii} > T_i$ and $R_{im} > T_i$, therefore Researcher i and Researcher m belong to group i.

We find Membership for group j, group k, group n, group m in a similar way and the following Researcher Groups are formed with their respective members:

Group i	Group j	Group k
Researcher i and m	Researcher j	Researcher k and n

Group n	Group m
Researcher k, n and m	Researcher n and m

CONSTRUCTION OF CLUSTERS

1) Total number of clusters

Now as we know total number of clusters i.e. K is equivalent to the minimum number of groups required to cover all the data points. Therefore, K=3. In other words, we can say that there are total three number of clusters required with the centroid as i, j, and n respectively.

1st cluster	2nd cluster	3rd cluster
i	j	n

2) Assigning points to the closest Centroid

In this step we assign points (researcher m and k) to the closest centroid by taking the proximity measure as the distance between two researchers. Therefore using the formula defined in equation (5), we calculate the distance of these two researchers with each of the above researchers:

d_{ki} = 0.289;d_{kj} = 0.3312; d_{kn} = 0.1214

Since d_{kn} is minimum, therefore researcher k belongs to the 3rd cluster with centroid as n.

d_{mi} = 0.1936; d_{mj} = 0.2576; d_{mn} = 0.0831

Similarly, Since d_{mn} is minimum, therefore researcher m also belongs to the 3rd cluster with centroid as n.

So, after the first iteration we have the following clusters:

1st cluster	2nd cluster	3rd cluster
i	j	n

Now, the 2nd iteration begins. We recompute the centroid of the 3rd cluster.

Distance between each of the two researchers is as follows:

d_{ij} = 0.2937; d_{in} = 0.2498; d_{jn} = 0.3868;

d_{km} = 0.186; d_{ki} = 0.289; d_{kj} = 0.3312;

d_{kn} = 0.1214; d_{mi} =0.1936; d_{mj} = 0.2576; d_{mn} = 0.0831

Assuming n to be the centroid:

S1= d_{nm} + d_{nk} = 0.1214 + 0.0831= 0.2045

Assuming m to be the centroid:
$S2= d_{mk} + d_{mn} = 0.186 + 0.0831= 0.2691$

Assuming k to be the centroid:
$S3= d_{km} + d_{kn} = 0.186 + 0.1214= 0.3074$

Since S1 is minimum, therefore n remains the centroid.

B. *Expert Mining module:*

As described in section III, the expert mining module is divided into three sub-modules, namely the data respository; sentiment mining & blog ranking modules, here we demonstrate them & examine their effectiveness. We consider a group with 4 members and analyze their blogs & cooments received on them. Our final task is to determine the expert from this group of 4 members.

BLOG 1:-

> The colours are boring. The headlights are not very strong and rear seats are less comfortable. There's hardly any boot space. The ride is not too bad, but there is a little stiffness and it crashes over sharp bumps. Ground clearance is very poor and is unstable at high speeds above 100km/h.

Comments:

1) Yes there's very little boot space.
2) I agree.
3) There are only 3 colours available.
4) I think the ride is quite good.
5) Instability at high speeds is a major drawback.
6) According to me, the seats are very comfortable.
7) Ground clearance is a very common issue in India!

The Sentiment mining module works in the following manner:

Feature Extraction: Each of the sentences along with their POS tag information is saved in the Repository. Sample XML for the blog 1 about car above:

<S>
<ART><WC = 'the'> *the*</W></ART>
<N><WC = 'colours'> colours </W></N>
<V><WC = 'are'> are </W></V>
<A><WC='boring'>boring</W></A>
</S>

<S>
<ART><WC = 'the'> *the*</W></ART>
<N><WC = 'headlights'> *headlights* </W></N>
<V><WC = 'are'> *are*</W></V>
<AG><WC='not'>*not*</W><WC='very'>*very*</W><WC= 'strong'> *strong*</W></AG>
<CONJ><WC='and'>*and*</W></CONJ>
<N><WC = 'rear-seats'> *rear- seats*</W></N>
<V><WC = 'are'> *are*</W></v>
<AG><WC='less'>*less*</W><WC='comfortable'>*comfortable*</W>
</S>

<S>
<P><WC = 'There'> *there*</W></P>
<V><WC = 'is'> *is*</W></V>
<AG><WC = 'hardly'> *hardly*</W><WC = 'any'> *any*</W></AG>
<N><WC = 'boot space'> boot space</W></N>
</S>

<S>
<ART><WC = 'the'> *the*</W></ART>
<N><WC = 'ride'> ride </W></N>
<V><WC = 'is'> *is*</W></V>
<AG><WC = 'not'> *not*</W><WC = 'too'> *too*</W><WC = 'bad'> *bad*</W></AG>
<WC = ','>, </W>
<CONJ ><WC = 'but'> *but*</W></CONJ>
<P><WC = 'There'> *there*</W></P>
<V><WC = 'is'> *is*</W></V>
<ART><WC = 'a'> *a*</W></ART>
<A><WC = 'little'> *little*</W></AG>
<N><WC = 'stiffness'> *stiffness*</W></N>
<CONJ><WC = 'and'> *and*</W></CONJ>
<N><WC = 'it'> *it*</W></N>
<V><WC = 'crashes'> *crashes*</W></V>
<P><WC = 'over'> *over*</W></P>
<A><WC = 'sharp'> *sharp*</W></A>
<N><WC = 'bumps'> *bumps*</W></N><WC='.'>.</W>
</S>

<S>
<N><WC = 'Ground clearance'>Ground clearance</W></N>
<V><WC = 'is'> *is*</W></V>
<AG><WC='very'>*vary*</W><WC='poor'>*poor*</W>
<CONJ><WC = 'and'> *and*</W></CONJ>
<V><WC = 'is'> *is*</W></V>
<A><WC='unstable'>unstable</W></A>
<P><WC = 'at'> *at*</W></P>
<N><WC = 'high speeds'> *high speeds*</W></N>
<P><WC = 'above'> *above*</W></P>
<N><WC = '100km/h'> 100km/h</W></N>
</S>

To determine the opinion word orientation, we establish the Adjective Polarity & the Adverb Strength: For adjective polarity, we use a set of seed adjectives whose orientations we know, & then grow this set by searching in the WordNet. We consider the following initial Adjective Seed-List, shown in Table II (with positive & negative orientations):-

TABLE II. SEED LIST OF ADJECTIVES

Positive Orientation	**Negative Orientation**
Great	Sharp
Blend	Dirty
Amazing	Sick
Compact	Unfortunate
Affordable	Bad
Reasonable	Boring
Excellent	Nasty
Big	Wrong
Fast	Poor
Comfortable	Awful
Strong	Scary
Beautiful	Dull
Impressive	Inferior
Good	Unstable
Exciting	Jerky
Stiff	Noisy
Variety	Common
Smooth	Okay okay

High	Bulky
Value-for-money	Low
Spacious	Drawback
Effective	
Major	
Attractive	
Stylish	
Streamlined	
Maneuverable	
Better	
Value for money	

We manually mark the strengths of a few frequently used adverbs with values ranging from -1 to +1 based on our intuitions. We consider the most frequently used adverbs (for our illustration) along with their strength as below in table III:-

TABLE III. ADVERB STRENGTHS

Adverb	**Strength**
Complete	+1
Most	0.9
Extremely	0.8
Absolutely	0.7
Too	0.7
Very	0.6
Indeed	0.6
More	0.4
Much	0.3
Reasonably	0.2
Any	0.1
Quite	-0.2
Pretty	-0.3
Little	-0.4
Less	-0.6
Not	-0.8
Never	-0.9

Opinion Strength Calculations: The strength of each opinion word is given by the formula defined in equation (7)

Opinion Words (for blog):

1. boring -1 * +0.5 = -0.5
2. not very strong -0.8 * +0.6 * +1 = -0.48
3. less comfortable -0.6 * +1 = -0.6
4. hardly any -1 * +0.1 = -0.1
5. not too bad -0.8 * +0.7 * -1 = +0.56
6. little stiff -0.4 * -1 = +0.4
7. sharp -1 * +0.5 = -0.5
8. unstable -1 * +0.5 = -0.5

Total Blog Orientation Strength = S (B_1) =
(- 0.5 - 0.48 - 0.6 - 0.1 + 0.56 + 0.4 - 0.5 - 0.5) / 8 = - **0.215**

Opinion Words (for comments):

1. very little +0.6 * -0.4 = -0.24
2. quite good -0.2 * +1 = -0.2
3. major drawback -1 * +1 * +0.5 = -0.5
4. very comfortable +0.6 * +1 = +0.6
5. very common +0.6 * -1 = -0.6

Average Comment Orientation Strength = Avg. S (C_1) =
(- 0.24 - 0.2 - 0.5 + 0.6 - 0.6) / 5 = -0.188/5 = **-0.0376**

Blog Score$_1$ = S (B_1) *Avg. S (C_1) = **-0.215** * **-0.0376 = +0.008**

BLOG 2:-

> The drive is reasonably smooth but gets jerky at higher speeds. Only manual transmission is available and that too is a little poor. The diesel model has a very noisy engine even for a new car. There is a very good variety of colours and a reasonably high mileage. All in all, it's value for money and a good buy.

Comments:

1) Jerky drive.
2) Mileage is good.
3) Its an okay okay buy.
4) Colour choices are good.
5) Engine is a little noisy.
6) Transmission is good.
7) I found it to be a smooth car.

Now we calculate the S (B_2) & Avg. S(C_2) to compute blog score $_2$

Opinion Words (for blog):

1. reasonably smooth 1* 0.2= 0.2
2. jerky -1 * 0.5=-0.5
3. little poor -1*-0.4= 0.4
4. very noisy -1*0.6= -0.6
5. very good 1*0.6= 0.6
6. reasonably high 1*0.2=0.2
7. good 1*0.5= 0.5
8. value for money 1*0.5=0.5

Total Blog Orientation Strength = B (S_2) =
(0.2+ (-0.5) + 0.4 + (-0.6) + 0.5 +0.2+ 0.5 + 0.5) / 8= +1.2/8 **= +0.15**

Opinion Words (for comments):

1. Jerky -1 *0.5= -0.5
2. Good 1* 0.5= 0.5
3. Okay okay -1* 0.5= -0.5
4. Good 1* 0.5= 0.5
5. Little noisy -0.4* -0.6= 0.24
6. Good 1* 0.5= 0.5
7. Smooth 1*0.5= 0.5

Average Comment Orientation Strength = Avg. S (C_2) =
(-0.5 +0.5 + (-0.5) + 0.5 + 0.24 +0.5 + 0.5) / 7= (+1.24)/7 = **+0.1771**

Blog Score$_2$ = S (B_2)* Avg. S(C_2) = **0.15 * 0.1771= + 0.02656**

BLOG 3:-

> This car is a complete blend of great power and style, with exciting features. It has very good fuel efficiency and engine is pretty impressive too. It's very spacious for its size and the drive is absolutely smooth. It has got beautiful interiors and the compact dimensions make it an excellent traffic warrior.

Comments:

1) Good review!
2) Interiors are indeed attractive.
3) Engine is a little noisy guys.
4) Car is quite stylish!
5) The car is spacious but bulky too!

Now we calculate the S (B_3) & Avg. S (C_3) to compute blog score $_3$

Opinion Words (for blog):

1. complete blend +1 * +1 * +0.5 = +0.5
2. great +1 * +0.5 = +0.5
3. exciting +1 * +0.5 = +0.5
4. very good +0.4 * +1 = +0.4
5. pretty impressive -0.3 * +1 = -0.3
6. very spacious +0.4 * +1 = +0.4
7. absolutely smooth +0.7 * +1 = +0.7
8. beautiful +1 * +0.5 = +0.5
9. compact +1 * +0.5 = +0.5
10. excellent +1 * +0.5 = +0.5

Total Blog Orientation Strength = B (S_3) =

(+ 0.5 + 0.5 + 0.5 + 0.4 - 0.3 + 0.4 + 0.7 + 0.5 + 0.5 + 0.5) / 10 = **+0.42**

Opinion Words (for comments):

1. good +1 * +0.5 = +0.5
2. quite stylish -0.2 * +1 = -0.2
3. indeed attractive +0.6 * +1 = +0.6
4. little noisy -0.4* -0.6= +0.24
5. spacious +1 * +0.5 = +0.5
6. bulky -1 * +0.5 = -0.5

Average Comment Orientation Strength = Avg. S (C_3) =

(+ 0.5 - 0.2 + 0.6 + 0.24 + 0.5 - 0.5) / 5 = (+1.14)/5 = **+0.228**

Blog Score$_3$ = S (B_3)* Avg. S (C_3) = +0.42 * +0.228 = +**0.09576**

BLOG 4:-

The size of this car is never big and this makes its price pretty reasonable and affordable. It does not demand any maintenance and its performance and safety are also amazing. Not much of car service is required. The cooling is very effective and this car is not very smooth on hilly terrains.

Comments:

1) Yes, very low maintenance required.
2) The ride is not very smooth.
3) Affordable price .Just right for middle class families.
4) Streamlined shape.
5) Cooling is not good especially for Delhi summers.
6) Its a good buy.
7) Better cars are available in the market.

Now we calculate the S (B_4) & Avg. S (C_4) to compute blog score$_4$

Opinion Words (for Blog):

1. never big 1* -0.9= -0.9
2. pretty reasonable 1 *-0.3=-0.3
3. Affordable 1*0.5=0.5
4. any 0.5*-0.1= -0.05
5. amazing 1*0.5=0.5
6. not much 0.5*0.3*-0.8=-0.12
7. very effective 1* 0.6= 0.6
8. not very smooth 1* -0.8* 0.6= -0.48

Total Blog Orientation Strength = B (S_4) =
-0.9 + (-0.3) +0.5 + (-0.05) +0.5 + (-0.12) +0.6+ (-0.48) / 8 = (-0.25) / 8 = **-0.03125**

Opinion Words (for Comments)

1. very low -1*0.6= -0.6
2. not very smooth 1*-0.8*0.6=-0.48
3. Affordable 1*0.5=0.5
4. Streamlined 1*0.5=0.5
5. not good 1*-0.8=-0.8
6. good 1*0.5=0.5
7. Better 1*0.5=0.5

Average Comment Orientation Strength = Avg. S (C_4) =
(-0.6 + (-0.48) +0.5 +0.5 + (-0.8) + 0.5 +0.5) / 7 = (0.12) / 7 = **+0.0171**

Blog Score$_4$ = S (B_4)* Avg. S (C_4) = -0.03125* 0.171= **-0.0005**

TABLE IV. BLOG RANKING

Blog	***Blog Score***	***Blog Rank***
Blog1	+0.008	3
Blog 2	+0.02656	2
Blog 3	+0.09576	1
Blog 4	-0.0005	4

Thus, comparing all the blog strengths, according to our approach, the highest blog score is for blog 3 and therefore the Expert is blogger 3!

Limitations:

1) It covers comments only written in English.
2) No abbreviations or acronyms can be accounted for.
3) It does not cover interrogative sentences.
4) A negative adjective and a negative adverb convert into a positive opinion word.
5) A positive adjective and a negative adverb also convert into a negative opinion word.
6) The method has no way of detecting and dealing with emoticons.

V. Conclusion

We proposed a novel ComEx Miner System for mining experts in virtual communities. This work is exploratory in nature and the prototype evaluated is a preliminary prototype. The major contributions of this research are:

i. Constructing a collaborative interest group known as the virtual community which will cluster researchers with similar interests in a same group and thereby facilitate collaborative work.

ii. Accessing the expertise from the virtual community using sentiment analysis of each group member's blog & comments received on it. Their combined orientation strength determined the blog score which enabled to rank the blogs and identify the expert as the one with the highest blog rank.

The practice result proves that this algorithm has the characteristics of highly effective group arranging and identifying expert. This study is just one step in this direction. Due to the complex nature of framework, it is impossible to consider and incorporate all the factors that could have an impact on the effectiveness and efficiency of this system. . More research needs to be done in order to validate or invalidate these findings, using larger samples.

References

[1] L. Colazzo, A. Molinari and N. Villa., "Collaboration vs. Participation: the Role of Virtual Communities in a Web 2.0 world". International Conference on Education Technology and Computer, 2009, pp.321-325.

[2] H.Rheingold, "The Virtual Community: Homesteading on the Electronic Frontier", revised edition. The MIT Press, 2000.

[3] Kumar, A. & Bhatia, MPS., "Community Expert based Recommendation for solving First Rater Problem". International Journal of Computer Applications (IJCA), Vol. 37, No.10, 7-13, January 2012, Foundation of Computer Science, USA.

[4] Kumar, A. & Jain, A., "An Algorithmic Framework for Collaborative Interest Group Construction". Recent Trends in Networks and Communications, LNCS-CCIS, Springer, Volume 90, Part 3, 2010, pp.500-508.

[5] Balog, K., Bogers, T., Azzopardi, L., de Rijke, M., van den Bosch, A.: "Broad expertise retrieval in sparse data environments". In: SIGIR 2007, pp. 551–558.

[6] Bogers, T., Kox, K., van den Bosch, A. "Using Citation Analysis for Finding Experts in Workgroups". In: Proc. DIR. (2008)

[7] John G. Breslin, Uldis Bojars, Boanerges Aleman-Meza, Harold Boley, Malgorzata Mochol, Lyndon J. B. Nixon, Axel Polleres, and Anna V. Zhdanova. "Finding experts using Internet-based discussions in online communities and associated social networks". First International ExpertFinder Workshop, Berlin, Germany, January 16, 2007, 2007.

[8] F. Metze, C. Bauckhage, T. Alpcan, K. Dobbrott, and C. Clemens, "A community-based expert finding system," in IEEE Int. Conf. on Semantic Computing (ISCS 2007), Irvine, CA, September 2007.

[9] Schall, D. & Dustdar, S. "Dynamic Context-Sensitive PageRank for Expertise Mining". SocInfo 2010, pp. 160-175

[10] Huh, E.-N., Lee, P. and Newby, G. "A Framework of Online Community based Expertise Information Retrieval on Grid". OGF Document Series GFD-I.164, January 2010.

[11] E. Smirnova and K. Balog. "A User-oriented Model for Expert Finding". In: 33rd European Conference on Information Retrieval (ECIR 2011), LNCS 6611,2011,pp. 580-592.

[12] Jie L"i, Harold Boley, Virendrakumar C. Bhavsar, and Jing Mei. Expert finding for eCollaboration using FOAF with RuleML rules". Montreal Conference on eTechnologies (MCTECH), 2006.

[13] Punnarut, R. & Sriharee, G. "A Researcher Expertise Search System using Ontology-Based Data Mining". Proceedings of Seventh Asia-Pacific Conference on Conceptual Modelling, APCCM'10, 2010, Volume 110, pp. 71-78.

[14] Zhang, J., Ackerman, M.S., Adamic, L. "Expertise Networks in Online Communities: Structure and Algorithms". The 16th international conference on World Wide Web, 2007, PP.7-16.

[15] Tang, J., Zhang, J., Zhang, D., Yao, L. and Zhu, C. "ArnetMiner: An Expertise Oriented Search System for Web Community. Semantic Web Challenge". In Proceedings of the 6th International Conference of Semantic Web (ISWC'2007).

[16] Pang-Ning Tan, Michael Steinbach and Vipin Kumar. "Introduction to Data Mining". Pearson Education.

[17] POS Tagger: http://www.infogistics.com/textanalysis.html

[18] WordNet: http://wordnet.princeton.edu/

Authors Profile

Akshi Kumar is a PhD in Computer Engineering from University of Delhi. She has received her MTech (Master of Technology) and BE (Bachelor of Engineering) degrees in Computer Engineering. She is currently working as a University Assistant Professor in Dept. of Computer Engineering at the Delhi Technological University, Delhi, India. She is editorial review board member for '*The International Journal of Computational Intelligence and Information Security'*, Australia, ISSN: 1837-7823; *'International Journal of Computer Science and Information Security'*, USA, ISSN: 1947-5500; '*Inter-disciplinary Journal of Information, Knowledge & Management'*, published by the Informing Science Institute, USA. (ISSN Print 1555-1229, Online 1555-1237) and '*Webology*', ISSN 1735-188X. She is a life member of Indian Society for Technical Education (ISTE), India, a member of International Association of Computer Science and Information Technology (IACSIT), Singapore, a member of International Association of Engineers (IAENG), Hong Kong, a member of IAENG Society of Computer Science, Hong Kong and a member of Internet Computing Community (ICC), AIRCC. She has many publications to her credit in various journals with high impact factor and international conferences. Her current research interests are in the area of Web Search & Mining, Intelligent Information Retrieval, Web 2.0 & Web Engineering.

Nazia Ahmad is doing M.Tech (Master of Technology) in Computer Technology & Application from Delhi Technological University, Delhi, India and has done her B.Tech (with Distinction) also in Computer Science and Engineering. She is currently working as an Assistant Professor in Dept. of Computer Engineering at the Delhi Technological University, Delhi, India.

Identification of the Strong Factor that Influence the Improvement of the Primary Education using the Evidential Reasoning

Mohammad Salah Uddin Chowdury
Department of Computer Science & Engineering
BGC Trust University Bangladesh
Chittagong, Bangladesh

Smita Sarker
Department of Computer Science & Engineering
BGC Trust University Bangladesh
Chittagong, Bangladesh

***Abstract*— Primary education is one of the important social sectors in the developing countries. It plays a gigantic role in promoting the social development of the concerned countries. More specifically, not the only education but the quality primary education which depends on such factors will contribute a lot for its smooth acceleration. In primary education, if the quality will maintain then automatically all the problems in primary education i.e. enrollment, completion, drop-out and so on will be robotic and gradually solved. The improvement primary education is influenced by both multiple qualitative and quantitative factors. This paper presents evidential reasoning(ER) approach to find out significant factors that are aggregated in assessment of performance of primary education. A case study of Dhaka, Chittagong, Rajshahi and Khulna districts in Bangladesh is provided to illustrate the implementation process of the ER approach for finding strong factors of primary education. For that reason, firstly we assess the performance of four districts then we determine the weakness and strength of specific factors of particular district. In this paper we also show the relation of lowest or best performing districts to its specific factors.**

Keywords- Assessment; evidential reasoning; district-wise primary education; factors; key performance indicator (KPI); multiple attribute decision analysis (MADA); uncertainty; utility interval.

I. INTRODUCTION

A Primary education in Bangladesh that is the largest unitary authorities in the world plays an important role in education system. In order to improve primary education Bangladesh government has undertaken a lot of necessary steps. The Primary and Mass Education Division (PMED) prepared in 1997 a comprehensive Primary Education Development Program (PEDP) which aimed at enhancement of education planning and management capacity, increasing equitable access to primary schooling and improvement of the quality of primary education through its several projects [10]. In the World Education Forum held at Dakar, Senegal in April 2000, the government of Bangladesh has committed to achieve of Education for All goals and every citizen by the year 2015[12]. Now 75% of total schools are controlled by the government and around 83% of the total children enrolled in the primary level educational institution go to these schools [10], [12]. The quality of primary education is evaluated in terms of the classroom climate, teaching style, classroom management, and understanding the subject matters during the lesson in the class [13]. PEDP-II defines 14 key monitoring indicators, the Key Performance Indicators (KPI) and the Primary School Quality Level (PSQL) indicators which act as the basis for the sector programme performance report, setting expectations that will instill purpose to ongoing monitoring and evaluation activities for the benefit of the planning process [1], [2], [10].

A number of survey reports were published to reflect the performance of primary education of Bangladesh based on some statistical measurement. According to case study of two district of Tamil Nadu it is observed that they assess the primary education with taking few major factors such as completion, repetition, and dropout rates. After analyzing these factors they identify the weaken areas which contribute to the lack of acceptable quality schools and focus the learning environment, school governance and management issues [11].But this assessment procedure do not follow systematic computational methods. For this reason the result is reflected by unwanted uncertainties. At that case, Evidential Reasoning approach is very effective which enable both qualitative and quantitative measurement under the multiple attributes decision analysis [3], [4], [5], 6].

In this paper we select four focus districts as our problem areas where two significant factors of KPI such as enrollment and outcomes are selected for performance measuring factors. The main objective of this paper is to find out the strong factor that influence the performance of primary education best district using ER approach by aggregating basic attributes of these two factors. Finally we show the ranking of district wise primary education as well as their factors.

We organize the research activities as follows. In section 2, we explain the ER approach for ADPE outlined and illustrated by subsequent sub-sections 2.1, and 2.2. The experimental result is outline by section 3.Finaly we concluding our remarks at section 4 in which we show the outcomes of evaluation with the discussion of suggestion of future work.

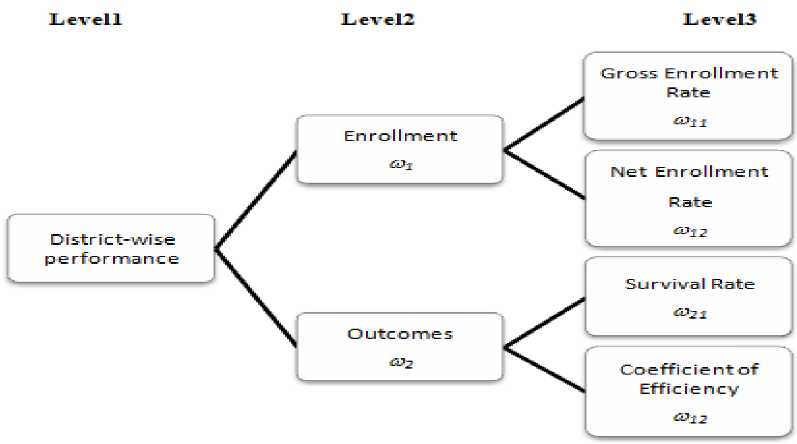


Figure 1. Evaluation hierarchy of the district-wise primary education

TABLE1 1. District wise Comparison of Enrollment and Outcomes (2005-2007)

General attributes	Basic attributes	District level Primary Education (2005,2006,2007)			
		Dhaka	Chittagong	Rajshahi	Khulna
Enrollment	Gross Enrollment Rate(GER)	87.9%,91.4%,91.2%	85.7%,95%,95.3%	91.2%,92.2%,93.4%	100.4%,97.3%98.1%
	Net Enrollment Rate(NER)	82.7%,87.5%,86.0%	84.3%,88.8%,89.6%	90.1%,86.5%,88.5%	92.9%,91.3%,92.4%,
Outcomes	Survival Rate	69.5%,61.6%,47.2%	58.1%,63.6%,62.6%	53.3%,57.7%,59.9%	64.4%,60.3%,63.2%
	Coefficient of efficiency	75.2%,67.7%,58.0%	61.5%,66.6%,64.8%	59%,65.8%,66.7%	66.8%,66.1%,65.8%

2. The Evidential Reasoning Approach for Identification strong factors in the Primary Education Assessment

2.1. Identification of Assessment Factors and Evaluation Grades

We apply the evidential reasoning approach to analyze the performance of four main districts wise primary education including *Dhaka*, *Chittagong*, *Rajshahi*, and *Khulna*. Here only qualitative performance attributes are considered for demonstrating purpose. The major performance attributes are considered as *enrollment* and *outcomes.* For facilitating the assessment these attributes are further classified basic factors such as *gross enrollment, net enrollment, survivable rate* and *coefficient of efficiency* which are shown in the figure 1.

According to School Survey Report 2007 we draw the scenario of primary education of four districts as the following table [1]. Considering the data of three years of the following table 1 we define the evaluation scale as follows:

Gross Enrollment Rate and Net Enrollment Rate

Indifferent (I):80-85% Average (A): 85-90%, Good (G): 90-95, Excellent (E): 95-100%

Survival Rate

Indifferent (I):45-55%, Average (A):55-60%, Good (G):60-65%, Excellent (E): 65-70%

Coefficient of efficiency

Indifferent (I):45-55%, Average (A):55-65%, Good (G):65-75, Excellent (E): -75-85%

In the following table 2 we summarize the whole assessment problem where I, A, G, and E indicate the evaluation grades Indifferent, Average, Good and Excellent respectively, and the number in a bracket denotes the degree of a belief to which an attribute is assessed to a grade. Now we consider the primary education of Dhaka district where we use the grades as defined before and represent the following distribution as follows[3] , [4], [5],[6],[7],[8]:

S(Gross Enrollment Rate)={ (average, 0.2) , (good, 0.8) } (1a)
S(Net Enrollment Rate)={ (indifferent, 0.2) , (average, 0.8) } (1b)
S(Survival Rate)={ (average, 09) } (1c)
S(coefficient of efficiency)={ (average, 0.2) , (good, 0.6) } (1d)

Due to uncertainty some incomplete assessment are shown in the table 2.For example:

TABLE 2
Decision matrix for district wise primary level education assessment

General attribute	Basic attributes	Districts types			
		Dhaka (a_1)	Chittagong (a_2)	Rajshahi (a_3)	Khulna (a_4)
Enrollme -nt (e_1)	gross enrollment rate(e_{11})	A(.2) G(.8)	A(.1) G(.9)	G(1)	E(1)
	net enrollment rate(e_{12})	I(.2) A(.8)	I(.1) A(.9)	A(.2) G(.8)	G(1)
Out-comes (e_2)	survival Rate(e_{21})	A(.9)	G(.8) A(.2)	I(.1) A(.9)	G(1)
	coefficient of efficiency (e_{22})	A(.2) G(.6)	A(.9) G(.1)	A(.4) G(.6)	G(.8)

The gross *enrollment rate* of Dhaka is complete because the sum of belief is 1 but the *coefficient of efficiency* is incomplete because .8<1.

2.2 Computational steps of aggregating assessment

Firstly we show the total calculation for assessment of enrollment of Dhaka primary education .The enrollment (e_1) is assessed by two basic attributes: gross enrollment rate (e_{11}) and net enrollment rate(e_{12}).

From (1a) and (1b), we have

$\beta_{1,1} = 0$, $\beta_{1,2} = 0.2$, $\beta_{1,3} = 0.8$, $\beta_{1,4} = 0$

$\beta_{2,1} = 0.2$, $\beta_{2,2} = 0.8$, $\beta_{2,3} = 0$, $\beta_{2,4} = 0$

On the basis of importance on the performance of primary education suppose the hypothetical weights for two attributes are: ω_{11}=0.55 and ω_{12}=0.45.

We get the basic and combined probability masses ($m_{n,i}$) by using following recursive equations [4], [5], [6], [7], [8]:

$m_{n,i} = \omega_{n,i}\beta_{n,i}$ for i=1, 2…N

$m_{1,1} = 0;$ $m_{2,1} = 0.11;$ $m_{3,1} = 0.44;$ $m_{4,1} = 0;$

$\bar{m}_{H,1} = 0.45 \quad \tilde{m}_{H,1} = 0$

$m_{1,2} = 0.09;$ $m_{2,2} = 0.36;$ $m_{3,2} = 0;$ $m_{4,2} = 0;$

$\bar{m}_{H,2} = 0.55 \; ; \tilde{m}_{H,2} = 0$

$$K_{I(2)} = \left[1 - \sum_{t=1}^{4}\sum_{\substack{j=1 \\ j\neq t}}^{4} m_{t,I(1)} m_{j,2}\right]^{-1}$$

$$= [1-(0+..+0+.0099+0.0396+0.1584+0+..+0)]^{-1}$$

$$= [1-.2079]^{-1} = 1.2625$$

And $m_{H,i} = \bar{m}_{H,i} + \tilde{m}_{H,i}$ (i=1,2) now we have

$m_{1,I(2)}=K_{I(2)}(m_{1,1},m_{1,2}+m_{1,1},m_{H,2}+m_{1,2}m_{H,1})=1.2625(0+0+0.09*0.45)$ $=0.0511$

$m_{2,I(2)} = K_{I(2)}(m_{2,1},m_{2,2}+ m_{2,1},m_{H,2}+ m_{2,2}\, m_{H,1})$
$=1.2625(0.11*0.36+00.11*0.55+0.36*0.45) =0.3309$

$m_{3,I(2)} =K_{I(2)}(m_{3,1},m_{3,2}+ m_{3,1},m_{H,2}+ m_{3,2}\, m_{H,1})$
$=1.2625(0+0+0.44*0.45) =0.3055$

$m_{4,I(2)} = K_{I(2)}(m_{4,1},m_{4,2}+ m_{4,1},m_{H,2}+ m_{4,2}\, m_{H,1})=0$

$$\bar{m}_{H,I(2)} = K_{I(2)}\left[\bar{m}_{H,I(1)}\bar{m}_{H,2}\right] = 0.3124$$

$$\tilde{m}_{H,I(2)} = K_{I(2)}\left[\tilde{m}_{H,I(1)}\tilde{m}_{H,2} + \bar{m}_{H,I(1)}\tilde{m}_{H,2} + \tilde{m}_{H,I(1)}\bar{m}_{H,2}\right] = 0$$

Now the combined degrees of belief are calculated by using equation as follows [4], [5], [6], [7], [8]:

$$\beta_1 = \frac{m_{1,I(2)}}{1-\bar{m}_{H,I(2)}} = \frac{0.0511}{1-0.3124} = 0.0743$$

$$\beta_2 = \frac{m_{2,I(2)}}{1-\bar{m}_{H,I(2)}} = \frac{0.3309}{1-0.3124} = 0.4812$$

$$\beta_3 = \frac{m_{3,I(2)}}{1-\bar{m}_{H,I(2)}} = \frac{0.3055}{1-0.3124} = 0.4443$$

$$\beta_4 = \frac{m_{4,I(2)}}{1-\bar{m}_{H,I(2)}} = 0$$

$$\beta_H = \frac{\tilde{m}_{H,I(2)}}{1-\bar{m}_{H,I(2)}} = 0$$

Then the enrollment rate of primary education of Dhaka district is assessed by

S (enrollment rate)

= {(indifferent, 0.0743), (average, 0.4812), (good, 0.4443)} (2a)

After repeating above procedures we assess the outcomes attribute as follows:

S (outcomes)

={(average, 0.7094), (good, 0.1751)} (2b)

Finally the performance of Dhaka district primary education is assessed by enrollment (e_1) and outcomes (e_2) as shown in table 3.

TABLE 3
Assigned weights, beliefs and calculated probability masses for level 2 attributes of Dhaka

Attributes	Weight	Belief				Probability Mass						
	ω_i	$\beta_{1,i}$	$\beta_{2,i}$	$\beta_{3,i}$	$\beta_{4,i}$	$m_{1,i}$	$m_{2,i}$	$m_{3,i}$	$m_{4,i}$	$m_{H,i}$	$\bar{m}_{H,i}$	$\tilde{m}_{H,i}$
enrollment (e_1)	.5	.0743	.4821	.4443	0	.0372	.2412	.2222	0	.5	.5	0
Outcomes (e_2)	.5	0	.7094	.1751	0	0	.3547	.0876	0	.5577	.5	0.0577

By using above similar actions we find the assessment degree of for Dhaka district as follows:

S(Dhaka) = { (indifferent, 0.0328 *), (average, 0.6272) , (good, 0.2954) }* (3a)

Similarly we can generate the overall assessment of primary education of other three districts such as Chittagong, Rajshahi and Khulna:

S(Chittagong) = { (indifferent, 0.0153*), (average, 0.4381) , (good, 0.5467) }* (3b)

S(Rajshahi) = { (indifferent, 0.0271*), (average, 0.4034) , (good, 0.5695) }* (3c)

S(Khulna) = { ((good, 0.7281), (excellent, 0.2613) } (3d)

According to the statements 3a-3d we show the distributed district-wise performance of our given four districts in the following figure 3.

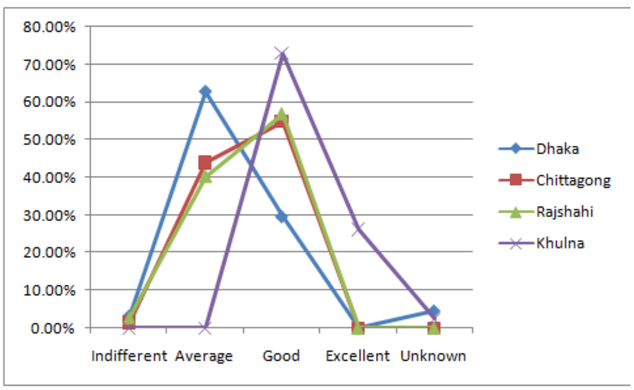


Figure 2. Distributed Assessment of District-wise performance

II. Experimental Results and Analysis

To precisely rank the four districts, their utilities need to be estimated. To do so, the utilities of the four individual evaluation grades need to be estimated first. The above partial rankings of alternatives could be used to formulate regression models for estimating the utilities of grades [4],[5].[6],[7],[8]. The maximum, minimum, and the average expected utility on the particular districts are given on the table 4.

TABLE 4
Utilities on district wise performance

	U_{min}	U_{max}	U_{avg}	Rank
Dhaka	0.4049	0.4548	0.4321	4
Chittagong	0.5109	0.5109	0.5109	3
Rajshahi	0.5147	0.5147	0.5147	2
Khulna	0.7491	0.7730	0.7611	1

Now from above analysis it clear that the performance ranking of four district-wise primary education is stated as follows.

Khulna>Rajshahi>Chittagong>Dhaka

Now we identify the strong entity of best performing district Khulna with considering the table 5.Firstly we compute the utility of these factors then we rank the factors on basis on their utility value.

Using the following formula we find out the maximum, minimum and average utilities of particular factor.

$$u_{\max}(y) = \sum_{n=1}^{N-1} \beta_n u(H_n) + (\beta_N + \beta_H) u(H_N) \qquad (4a)$$

$$u_{\min}(y) = (\beta_1 + \beta_H) u(H_1) + \sum_{n=2}^{N} \beta_n u(H_n) \qquad (4b)$$

$$u_{avg}(y) = \frac{u_{\max}(y) - u_{\min}(y)}{2}. \qquad (4c)$$

TABLE 5. Utility and Rank of the Factors

	U_{min}	U_{max}	U_{avg}	Rank
Enrollment	0.8677	0.8677	0.8677	1
Outcomes	0.6418	0.6839	0.6629	2

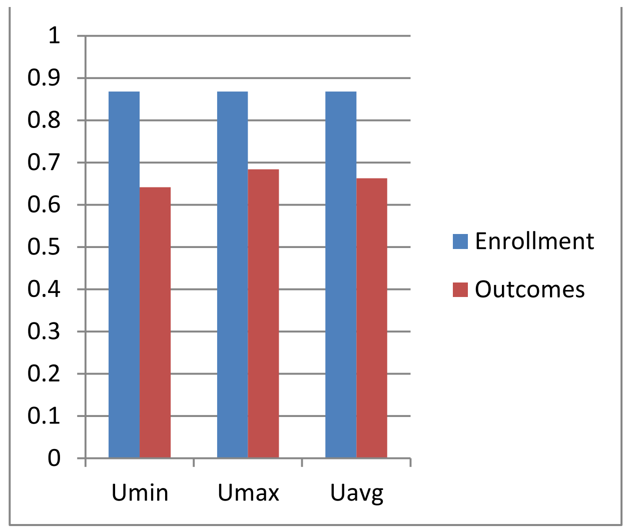


Figure 3. Utility interval of the factors

Considering the equations we find the results which shown the table 5. Hence we the rank of the factor is as follow.

Enrollment>Outcomes

III. Conclusion

Primary Education requires many qualitative and quantitative attributes in their assessment which need strong methodologies in order to find out their significant impact. In this paper we apply evidential reasoning approach that is appropriate for such attributes assessment under uncertainties. Because of major role in the assessment of primary education we focus our attention on two attributes i.e. enrollment and outcomes which are successfully aggregated by ER approach. The uncertainties among various attribute of particular district are properly handled by our suggested approach. On the above study we summarize that the performance of Khulna district is highest among all other districts. In the case of particular district assessment process the relative importance of each attribute is also measured by identifying the strengths and weakness of it's on each district. In this paper we identify that the enrollment f the strong factor of Khulna district primary education. For that reason it is clear that if we increase the enrollment rate particular district then the performance of the primary education of this district is stronger. We present the result of an individual district in the form of interval from minimum utility to maximum utility in a systematic and effective way. When a set of necessary steps to increase the performance of such weaken factors of weaken districts then the expected quality of primary education is achieved.

ACKNOWLEDGMENT

In this work we are grateful to Prof. Dr.Mohammad Shahadat Hossain, Department of Computer Science & Engineering, University of Chittagong, Bangladesh for his sound idea on Evolutionary Computing specially Evidential Reasoning.

REFERENCES

[1] School Survey Report of Second Primary Education Development Programme (PEDP-II), ,Directorate of Primary Education ,Ministry of Primary and Mass Education, Government of the Peoples' Republic of Bangladesh, pp.1-131 ,2007(Published 2008).

[2] Bangladesh Primary Education Annual Sector Performance Report , Government of the People's Republic of Bangladesh, Directorate of Primary Education, Second Primary Education Development Programme, pp. 1-72,May 2009.

[3] J. B. Yang and M. G. Singh, "An Evidential Reasoning Approach for Multiple Attribute Decision Making with Uncertainty," *IEEE Trans. Syst.,Man, Cybern.*, vol. 24, no. 1, pp. 1–4, 1994.

[4] J. B. Yang and D. L. Xu, "On the Evidential Reasoning Algorithm for Multiple Attribute Decision Analysis with Uncertainty," *IEEE Trans. Syst., Man, Cybern. A*, vol. 32, pp. 289–304, May 2002.

[5] J. B. Yang and D. L. Xu, "Nonlinear Information Aggregation via Evidential Reasoning in Multiattribute Decision Analysis Under Uncertainty," *IEEE Trans. Syst., Man, Cybern. A*, vol. 32, no. 4, pp. 376–393, May 2002.

[6] J. B. Yang, "Rule and Utility Based Evidential Reasoning Approach for Multiple Attribute Decision Analysis Under Uncertainty," *Eur. J. Oper.Res.*, vol. 131, no. 1, pp. 31–61, 2001.

[7] D. L. Xu, "Assessment of Nuclear Waste Repository Options Using the Er Approach," *Int. J. of I T & DM* vol. 8, no. 3, pp. 581–607, 2009.

[8] P. Gustafsson, R. Lagerström, P. Närman, and M. Simonsson , "The Ics Dempster-Shafer how to ," unpublished

[9] Y.Wang, J. B. Yang, and D. L. Xu, "Environmental Impact Assessment using the Evidential Reasoning Approach," *Eur. J. Oper.Res.*, vol. 174, pp. 1885–1913, 2005.

[10] Baseline Report of Second Primary Education Development Programme (PEDP-II), Government of the People's Republic of Bangladesh, Directorate of Primary Education, Second Primary Education Development Programme, pp. 1-138, June 2006

[11] S. Grover,N. H. Singh, and N. Stromquist, "The Quality of Primary Education" *A Case Study of Madurai and Villupuram Districts in Tamil Nadu,India*, pp. 1-77,April 2002.

[12] A F M Fazle Rabbi, "Primary Education in Bangladesh," *Viability of Achieving Millennium Development Goals*, pp. 1-86, 2007.

[13] M. Nurul I. Shekh, "The State of Primary Educational Quality in Bangladesh: An Evaluation" unpublished

Automated Biometric Voice-Based Access Control in Automatic Teller Machine (ATM)

Yekini N.A.
Department of Computer Technology
Yaba College of Technology, Lagos Nigeria.

Itegboje A.O.
PHD Candidate, SMED Southern university,
Baton Rouge, LA US.

Oyeyinka I.K.
Department of Computer Technology
Yaba College of Technology, Lagos Nigeria

Akinwole A.K.
Department of Computer Technology
Yaba College of Technology, Lagos Nigeria.

***Abstract*— An automatic teller machine requires a user to pass an identity test before any transaction can be granted. The current method available for access control in ATM is based on smartcard. Efforts were made to conduct an interview with structured questions among the ATM users and the result proofed that a lot of problems was associated with ATM smartcard for access control. Among the problems are; it is very difficult to prevent another person from attaining and using a legitimate persons card, also conventional smartcard can be lost, duplicated, stolen or impersonated with accuracy. To address the problems, the paper proposed the use of biometric voice-based access control system in automatic teller machine. In the proposed system, access will be authorized simply by means of an enroll user speaking into a microphone attached to the automatic teller machine. There are 2 phases in implementation of the proposed system: first training phase, second testing or operational phase as discussed in section 4 of this paper.**

Keywords- Automatic Teller Machine (ATM), Biometric, Microphone, Voiced-Based Access Control, Smartcard Access Control, Voiced-Based Verification System

I. INTRODUCTION

The biometric recognition systems, used to identify person on the basis of physical or behavioral characteristics (voice, fingerprints, face, iris, etc.), have gained in popularity during recent years especially in forensic work and law enforcement applications [1]. Automatic Teller Machine was invented to address the following issues in banking system: Long queue in banking hall, Quick access to fund withdrawal, banking at any time, Improvement in the quality of banking services to customers.

Safety of bank customer fund in banking has always been a concern since ATM was introduced. Access control for automatic teller machine represent an important tool for protecting hank customers fund and guarantee that the authentic owner of the ATM card [smartcard] is the one using it for transaction [9]. The most important authentication method for ATM is based *on* smartcard [Njemanze, P.C. 2007). It is very difficult to prevent another person from attaining and using a legitimate person's card.

The conventional smartcard can be lost, duplicated, stolen, forgotten or impersonated with accuracy [7]. This conventional security procedure in ATM cannot guarantee the required security for ATM.

An intelligent voiced-based access control system, which is biometric in nature, will enable automatic verification of identity by electronic assessment of one or more behavior and/or physiological characteristics of a person. Recently biometric methods used for personal authentication utilize such features as face, voice, hand shape, finger print and Iris [4]. In other to overcome the problems of smartcard access control in ATM. This paper proposed an intelligent voice-based access control system which is a biometric technique that offers an ability to provide positive verification of identity from individual voice characteristics to access automatic teller machine. The use of voice as a biometric characteristic offers advantages such as: it is well accepted by the uses, can be recorded by regular microphones, the hardware costs are reduced, etc. The paper discusses; ATM and it model network, drawback in smartcard based access control for ATM based on survey of 1000 users of ATMs, proposed voiced based access control and conclusion that compare the advantages of the system over current: available technology.

II. AUTOMATIC TELLER MACHINE MODEL NETWORK

Automatic teller machine is online with bank, each transaction will be authorized by the bank on demands and it uses real-time online processing technique which directly updated the account from which transaction takes place. Figure I, gives the ATM model network.

The ATM model in figure 1 work as follow; Bank customer inserts the smartcard (smartcard) in the ATM machine. The machine then request for a personal identification Number PIN if the supply PIN is correct, access will be authorized and transaction will continue, the customer then enter the amount to withdrawal, and if the customer has enough money in the account then the amount will he paid.

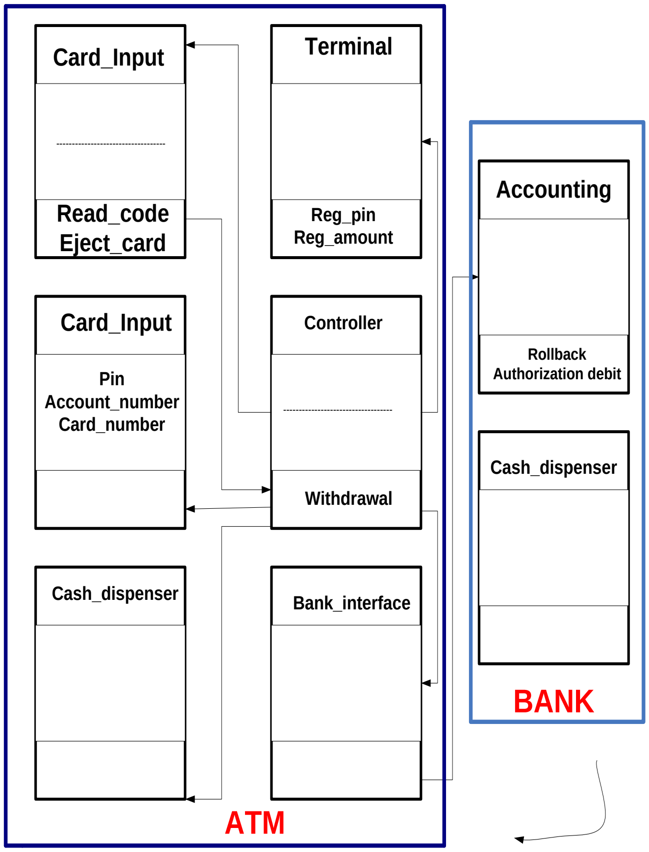


Figure l: ATM Model Network [8].

The whole work is being monitored by the controller class. In principle this is not necessary, but for working with a secure model the controller class is needed as a dispatcher of actions and it would have a log file with the trace of every transaction earned out by ATM.

The class card_input has the methods for reading the code of the client's card and for ejecting the card from machine. It interacts through the controller with the class terminal, where the methods reg_PIN and reg_amount are defined.

In other to verify whether the PIN of a particular users is correct or not, the class card will have the information of the cardholder i.e. card_number, PIN, and Account number. The controller will interact with the bank the bank using the information of the card holder in order to get the authorization to pay (or not) request amount.The bank_interface will send the request to the accounting class, which belongs to the bank package, in other to call the debit method of the accounting class [5].

The accounting class has the methods of rollback, authorization and debit which directly interact with the accounting class. Rollback is for rollback a transaction in case anything is wrong and should leave the account and the teller machine in the original state; authorization will authorize or not an operation and debit will extract the requested amount of money from the account in case the operation is authorized [8].

III. RESEARCH FRAMEWORK AND METHODOLOGY

Research framework and methodology is based on the survey that covered a sample of one thousand ATM users in Lagos state. The choice of the location is based on the fact that Lagos state is the economy nerve center of Nigeria and it has more branches of the banks and ATM location compare to any other state in Nigeria. The following questionnaire was used to get information that prompts us to proposed the voiced based access control

QUESTIONNAIRE

Set A

S/N	Question	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
1.	Banking would have been better if ATM was never invented	817	155	5	10	13
2.	Withdrawal of money from bank using ATM is faster compared to normal banking	836	34	0	156	714
3.	Withdrawal of money from ATM is more secure	619	181	9	101	100
4.	There is need for better security access for ATM that will guarantee only one person to a card.	0	9	11	99	881
5.	Biometric Technique as in face recognition, fingerprint, voice recognition etc for access control in ATM would provide better security	0	11	9	109	871

Set B

		YES	NO
6.	Has there been a time where your card was rejected by ATM card because you enter a wrong PIN	618	382
7.	Have you ever misplaced your ATM card before?	719	281
8.	Have you been a victim of ATM fraud before	699	301
	If yes to question no (8) please give brief detail	Detail vary	

The result was later converted to column charts as shown figure 2a and b.

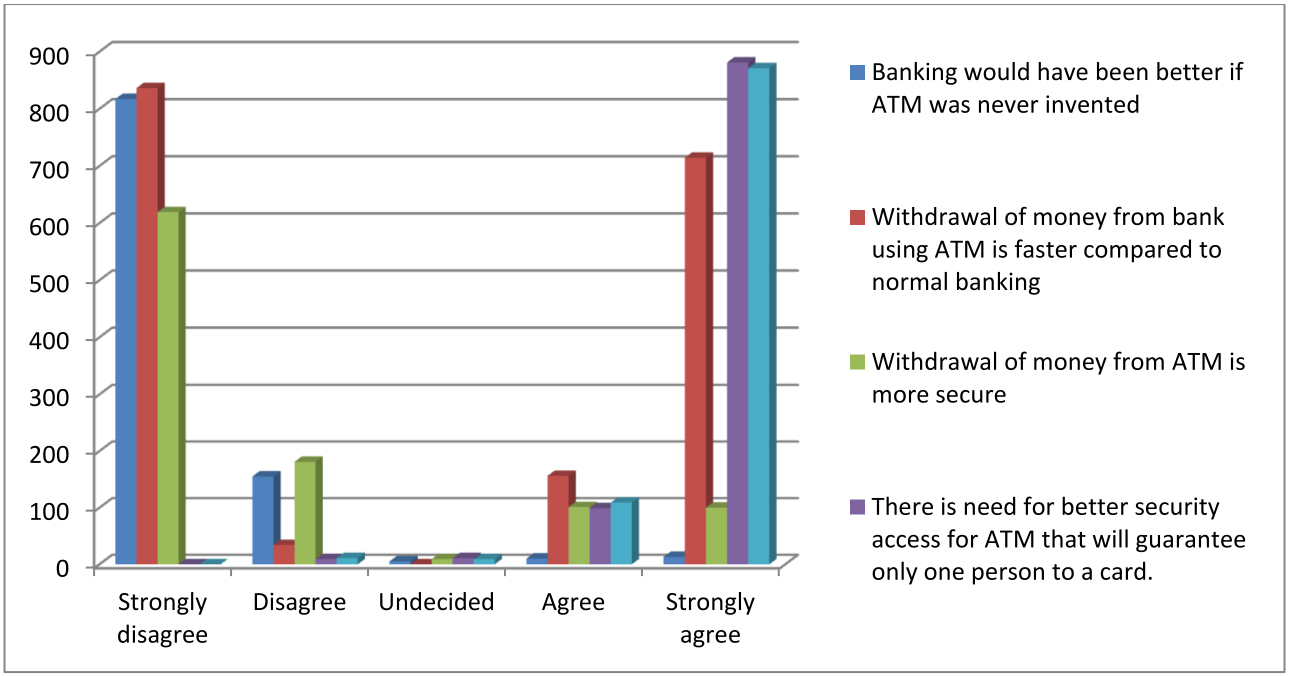


Figure 2a: Column chart for question set A

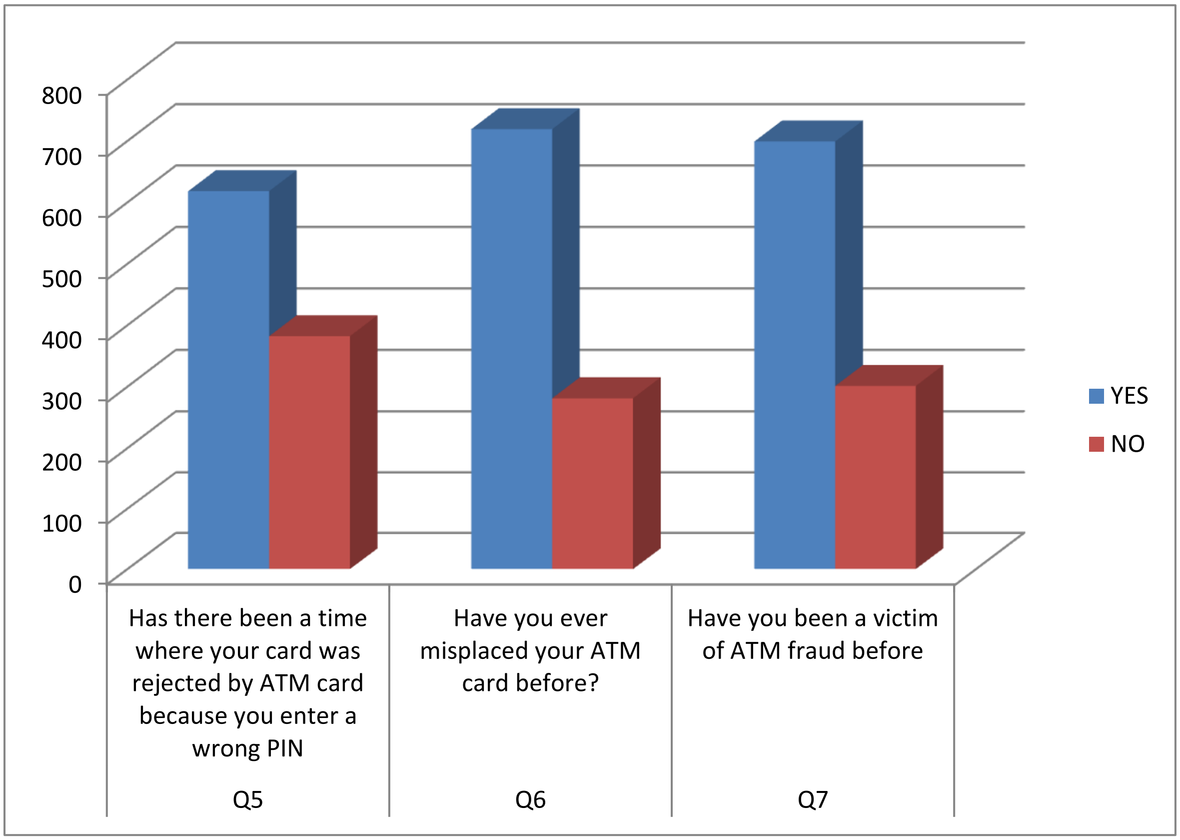


Figure 2b: Column Chart for Question set B

The result obtained from the questionnaire shows that, there is need for better security approaches to ATM access control. The result is analyses as follows:

Question 1: 81.7%, the response to the question implied that invention of ATM is a welcome innovation in banking sector.

Question	Strongly Disagree	%	Disagree	%	Undecided	%	Agree	%	Strongly Agree	%
Q1	817	81.7	155	15.5	5	0.5	10	1	13	13
Q2	336	83. 6	34	3.4	0	0	156	I5.6	714	71.4
Q3	619	61.9	181	18.1	9	0.9	101	10.1	100	10
Q4	0		9	0.9	11	1.1	99	99	881	88.1
Q5	0		11	1.1	9	0.9	109	109	871	87.1

Question 2: the majority of the respondent prefers using ATM because it enables quick access to withdrawal of money.

Question 3: the majority of the respondent strongly agreed that fund withdrawal through ATM is not secure compared to using face-face with cashier.

Question 4: larger percentage of the respondents strongly agreed that there is need for better security that will guarantee one user with one

Question 5: biometric approach to ATM access control would provide better security in ATM.

Question 6: majority of people interview has forgotten their password before and as a result they were unable to use their card to withdraw money from their account.

Question 7: larger percentage of people Intel-viewed has misplaced their ATM card before and as a result they were unable to use their card to withdraw money from their account.

Question 8: 69.9% of the people interviewed were once a victim of ATM fraud. Based on the detail given the approaches use in defraud the users varied and it was not discussed in this write up. The only common thing is that their card was used by another person to withdraw money from their account.

Based on the result analysis, it was discovered that Smart card access control has the following drawbacks;

1) *Magnetic stripe or chip distortion: information require for the card to function is stored on the chip or magnetic stripe. There is possibility of magnetic surface being distorted as result of continuous use and improper handling. This distortion can leads to damage and rendered the card useless.*
2) *Misplacement of the card: theirs is possibility of the card being misplaced and as a result rendered the card useless.*
3) *Stolen or theft: this card could be stolen by another person even with the password. There have been a case of burglar's forcefully collected ATM card and password from the legitimate owner and even follow such person to the ATM location to confirm that the PIN number given to them is correct.*
4) *Card fraud: recently, there has been reported case of card fraud. Various methods were use by fraudster in perpetuating this crime; among others are: for a low tech form of fraud, the easiest is to simply steal an ATM card. A later variant is of this is to trap the card inside ATMs card reader [3]. Advance fraud in ATM involve the installation of a magnetic card reader over the real ATMs card slot and use of a wireless surveillance camera or a modified digital camera to observe the user PIN. Card data is then cloned out on a second card and the criminal attempt a standard cash withdrawal [6]. Consequent to the identified drawbacks. The authors of this paper proposed the design of the intelligent voice-based access control in automatic teller machine.*

IV. PROPOSED INTELLIGENT VOICE-BASED SYSTEM

Proposed system description, figure 2 shows the schematic diagram of proposed intelligent voiced-based access control for ATMs. The proposed system basically consists of 3 main components:

1) *Voice sensor*
2) *Speaker verification system and*
3) *ATM access control*

A low cost microphone commonly used in computer system is used as voice sensor to record the ATM user voice. The recorded voice is then sent to the voiced-based verification system which will verity the authenticity of the user based on his/her voice.

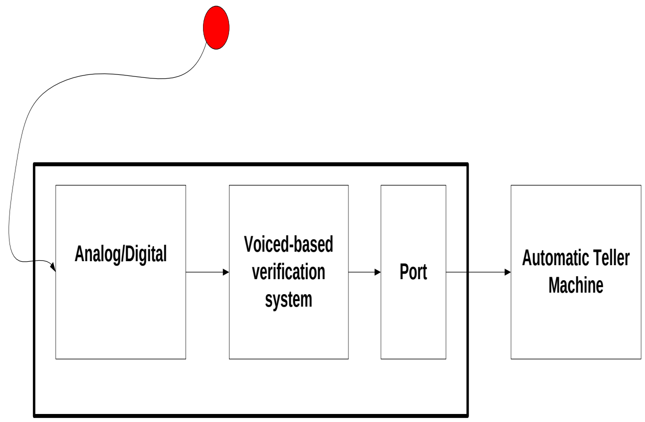


Figure 2: Proposed Intelligent Voice-Based ATM Access Control

Voice base verification, will enable a decision signal which will accept or reject the access that will be sent through the parallel port of the system. The intelligent voice-based access control will generally make possible decision as;

4) *Authorized person (ATM user) accepted*
5) *Unauthorized person (ATM user] rejected*

The qualities to measure rate of access control accuracy to reject the authorized person is called false rejection (FRR), and that to measure rate of access control to accept the unauthorized person is called false acceptance rate [FAR].

Mathematically, both rates are expressed as percentage using the following simple calculation [2].

NFR and NFA are the numbers of false rejections and false acceptance respectively, while NFA and NIA are the number of authorized person attempts. To generate high security of the ATM access control system it is expected that the proposed system have both low FRR and low FAR.

A. *Voiced-based verification System*

The use of voice for biometric measurement becomes more popular due to the following reasons; natural signal generation, convenient to process or distributed, and applicable for remote access. There are 2 kinds of voiced-based recognition or speaker recognition (Campbell J.P. 1997].

1) *Speaker identification*
2) *Speaker recognition*

In speaker verification system, the system decodes that a user is who claims to be. On other hand speaker identification decides the person among a group of person. Speaker recognition is further divided into 2 categories which are text dependent arid text independent.

Tex dependent speaker recognition recognizes the phrase that spoken, whereas in text identification the speaker can alter any word. Thy most appropriate method for voice-based ATM access control is based on concept of speaker verification, since the objective in the access control is accept or reject a user to gain access into ATM. Figure 3a and b. show the basic structure of the proposed system. There are 2 phases in the proposed system.

1. Training enrolment as shown in figure 2a, the authorized persons are registered and their voices are recorded. The recorded voices are then extracted. Features extracted from the recorded voices are used to develop models of authorized persons.

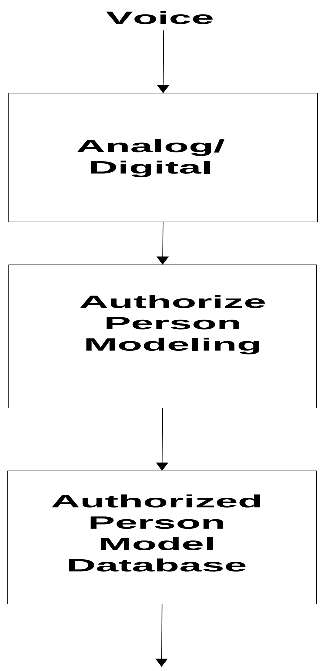


Figure 3a: Training phase

2. Testing or operational phase as shown in figure 3b, in this phase a person who want to access the ATM is required to enter the claimed identity and his/her voice. The entered voice is processed and compared with the claimed person model to verify his/her claim. At this point system decides whether the feature extracted from the given voice matches with the model of claimed person. Threshold is set in order to give a definite answer of access acceptance or rejection. When degree of similarity between a given voice and model is greater than threshold the system will accept the access, otherwise the system will reject the person to access the ATM.

V. Merits of Intelligent Voice-Based Access Control

Precession of classification: voice-based system has low false acceptance rate (FAR] and false rejection rate FRR compared to smartcard, and even other biometric method like fingerprint.

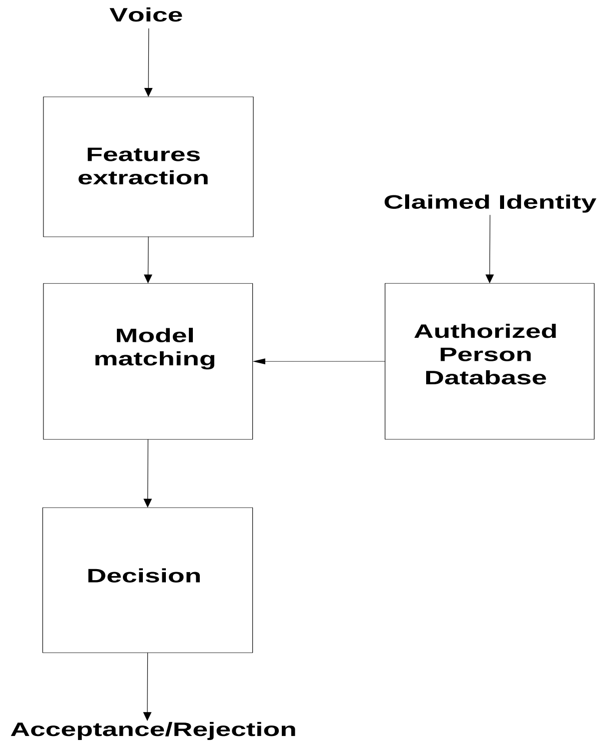


Figure 3b: Testing (operational) Phase

2. **Reliability**: as for other biometric, a test has shown that, spoofing method can be used to bypass finger print check, facial recognition can be broken in by showing a video of someone face or a still image of person. Voice recognition is reliable in the sense that no two people have the same voice signal.

3. **No misplacement**: smartcard can be misplaced or forgotten but a user voice is part of his body, as a result it can be forgotten or misplaced

4. **Economy**: the use of voice recognition will save the banks cost of producing smartcard.

5. **Fraud**: use of this method will reduce fraud discussed earlier.

VI. Conclusion

This paper presents a conceptual framework for use of intelligent voice-based access control in ATMs which is biometric approach. Research has shown that ATM user have encountered several problems in the past which include; chip distortion, card misplacement.

Card fraud, etc. these entire problems are associated with using smartcard access control in ATM. To overcome these problems it is advisable that government should partner with banking sector to implement the use of biometric technique "intelligent voice-based access control" in ATMs, as this will eliminate completely the problems associated with smartcard access control.

REFERENCES

[1] Atkins, W., 2001: A testing for face recognition technology. Biometric Technology Today, vol 147, pp. 195-197.

[2] Campbell, J, P., 1997. Speaker recognition; a tutorial. In Procc. IEEE, pp; 1437-1462.

[3] Dade, L. A. et al. 1998. Human brain function encoding and recognition: Anal of the New York Academy of Sciences, 355, 572 - 574

[4] Kung, S. Y., M W, Mack, and S. H. Lin, 2004. Biometric authentication machine learning approach. Prentice Hall

[5] Njemanze, P. C. 2007. Cerebral lateralization for processing Laterality, 12, 31 -49.

[6] Schoon G. A. A. and deBurn J. C. 1994, Forensic science international, pill.

[7] Wahyudi et al, 2007. Speaker recognition identifies people by their voices. In Proc. Of conference on security in computer application (2007).

[8] Yekini, N. A., and Lawal, 0. N. 2000. 1CT for accountants and Bankers: Hasfem Publication,

[9] Zhang, D, *d.,* 2000. Automated biometric technologies and systems. Kluwer academic Publisher.

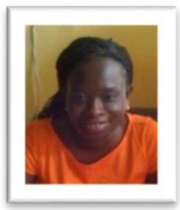
REFERENCES

Yekini Nureni Asafe., majors in Data Communication & Networking, Computer Application, and Computer Architecture. He obtained NCE in Physic from Lagos state college of education, B.Sc. Degree in Electronic and Computer Engineering from Lagos State University and Master Degree in Computer Science from University of Lagos, Nigeria.

He is a member of International association of Engineers (**IAENG**) and Nigeria Computer Society (**NSC**), International association of Computer Science and Information Technology (**IACSIT**). He currently lectures Data Communication and Networking, Assembly Language and Computer architecture in Yaba College of Technology, Lagos, Nigeria.

Itegboje A.O., obtained her B.sc degree in mathematics education, PGD in computer science and MSc in computer science from university of Lagos, former; HOD computer department, Director of MIS center, Yaba College of Technology. Lagos Nigeria. She is currently a PHD research candidate, SMED Southern university, Baton Rouge, LA US.

Oyeyinka I.K., a chief lecturer. Major in network congestion and control. He obtained his B.sc degree in mathematics and PGD in computer science from University of Lagos, and PhD Computer science from university of Abeokuta. He is a chief Lecturer in the department of Computer Technology, and currently the Director Center for Information Technology management center of Yaba College of Technology, Nigeria. He's a member of Computer Professional Council of Nigeria [CPN] and Nigeria Computer Society [NCS].



Akinwole A.K., a principal technologist. Major in Mobile Computing and ADhoc Network Performance. She obtained her academic Higher Diploma degree (HND) from Yaba College of Technology, PGD in Computer Science from Namidi Azikwe University and MTech in Computer Science from Ladoke Akintola University (LAUTECH) Ogbomosho. She is currently and academic staff in the department of Computer Technology, Yaba College of Technology. Lagos Nigeria.

Resolution Enhancement by Incorporating Segmentation-based Optical Flow Estimation

Osama A. Omer
Department of Electrical Engineering
South Valley University
Aswan, Egypt

***Abstract*— In this paper, the problem of recovering a high-resolution frame from a sequence of low-resolution frames is considered. High-resolution reconstruction process highly depends on image registration step. Typical resolution enhancement techniques use global motion estimation technique. However, in general, video frames cannot be related through global motion due to the arbitrary individual pixel movement between frame pairs. To overcome this problem, we propose to employ segmentation-based optical flow estimation technique for motion estimation with a modified model for frame alignment. To do that, we incorporate the segmentation with the optical flow estimation in two-stage optical flow estimation. In the first stage, a reference image is segmented into homogeneous regions. In the second stage, the optical flow is estimated for each region rather than pixels or blocks. Then, the frame alignment is accomplished by optimizing the cost function that consists of L_1-norm of the difference between the interpolated low-resolution (LR) frames and the simulated LR frames. The experimental results demonstrate that using segmentation-based optical flow estimation in motion estimation step with the modified alignment model works better than other motion models such as affine, and conventional optical flow motion models.**

Keywords- Optical flow; image segmentation; Horn-Schunck; super resolution; resolution enhancement.

I. INTRODUCTION

Multi-frame super resolution (SR) is the process of producing a resolution-enhanced frame from multiple low-resolution (LR) frames with sub-pixel shift. SR received much attention in computer vision and image processing communities over the past three decades [1-21] (for review see [6]). SR process typically includes three steps: (i) image registration (motion estimation), (ii) the alignment of LR frames on the high-resolution (HR) grid, and (iii) image restoration.

SR methods can be categorized, based on the domain in which the process is done, into time domain [1-5, 7-8, 10-21] and frequency domain [9]. In another way, they can be categorized, based on the incorporated motion model, into global motion model (including translation [3] and affine [2] models) and local motion model (including optical flow [7, 12] and block matching [13]). Furthermore, they can be categorized, based on the alignment process, into non-uniform interpolation [8], deterministic and stochastic regularization [10], and projection onto convex sets (POCS) [11].

Most of the existing super-resolution algorithms [2-3, 9-11] cannot cope with local motion, because they assume that motion model can be globally parameterized. To overcome the problems of registration error in locally moving parts, three techniques appeared in the literature. The first is to use different global (or local) weights for different registration error levels [14,16]. The second is to use local motion (or multi-motion) estimation to improve the accuracy of registration in the locally moving parts [7, 12, 15, 19]. The third is a combination of the previous two techniques [17, 18]. Among these techniques, the first technique is widely used in SR for its simplicity. The idea of using different weights for different registration is based on rejecting pixels or even whole frames that have high registration error. On the other hand, the main idea behind using local motion estimation techniques is to incorporate information from different frames as much as possible [12, 15, 19]. The main problem of multi-motion estimation [19] is the complexity, since it requires estimation of motion for each moving object in all frames, which is complex and not always accurate because of the fact that the motion of different object affects the motion of other objects. Also, using conventional optical flow estimation for motion estimation [7, 12] is sensitive to noise. In addition, using block-matching results in blocking artefacts due to dividing frames into blocks and incorrectly assuming that all pixels within each block have the same motion vector. Moreover, using combination of local motion estimation and weighting technique adds more complexity to the algorithm as proposed in [17,18]. In addition, algorithm proposed in [15] requires high computational time, since it perform region matching using full search.

On the other hand, region segmentation has been employed for image SR in [14, 15, 20, 21]. In [14], segmentation is employed in a region rejection based on the registration inaccuracies. While in [15], region is employed in a region matching in the registration step. In [20], a region-based super-resolution algorithm is proposed in which different filters are used according to the type of region. In this method the segmentation information is not fully used where it is used only to classify regions into homogeneous and inhomogeneous regions. In [21], the image is segmented into background and different objects and each of these are super-resolved separately using a traditional technique and then the super-resolved regions are merged to construct the HR image. This algorithm is very complex since it requires segmentation of moving objects and registration of each object separately.

In order to overcome these difficulties, we propose a modified gradient-based optical flow estimation algorithm based on the Horn-Schunck algorithm, so that tries to overcome the problems of conventional optical flow estimation by the use of image segmentation.

Based on the assumption that pixels in a homogeneous arbitrary shaped region have the same motion then motion discontinuities go along with discontinuities in the intensity image. In addition, using segmentation reduces the susceptibility to noise. Furthermore, we propose to incorporate the modified gradient-based optical flow estimation algorithm for estimating motion of each region in image registration.

Our claim in this work is that the enhancement of the motion estimation results leads to enhancement in the SR results. Therefore, we establish a framework of multi-frame SR method that utilizes segmentation-based optical flow estimation.

The method in the proposed framework is summarized as follows. First, the frames are interpolated to make optical flow estimation in sub-pixel accuracy, and then the interpolated reference frame is segmented into arbitrary shaped regions using watershed transform [22].

The locally moving segments are then motion compensated using search. The next step is image alignment, where we fuse the motion compensated frames to produce one blurred HR frame by using the L_1-norm in the frame alignment step. The so generated HR frame is de-blurred by using a regularization-based restoration method.

II. Problem Describtion

The problem of multi-frame SR is to estimate a high-resolution frame out of observed successive LR frames. Assume that N LR frames of the same scene denoted by Y^k ($1 \leq k \leq N$), each containing M^2 pixels, are observed, and they are generated from the HR frame denoted by X, containing L^2 pixels, where $L \geq M$. The observation of N LR frames are modelled as follows:

$$Y^k = \mathbf{DHF}^k X + V^k \tag{1}$$

where $\mathbf{F}^k$, $\mathbf{H}$ and $\mathbf{D}$ are the motion, blurring, and down-sampling operators, respectively. The size of $\mathbf{F}^k$, $\mathbf{H}$ and $\mathbf{D}$ are $L^2 \times L^2$, $L^2 \times L^2$ and $M^2 \times L^2$, respectively. X is the unknown HR frame, Y^k is the k-th observed LR frame, and V^k is an additive random noise for the k-th frame with the same size as Y^k.

Throughout the paper, we assume that $\mathbf{D}$ and $\mathbf{H}$ are known and the additive noise is Gaussian. The problem here in this paper is to estimate the motion for each frame, $\mathbf{F}^k$, and to find the original image, X.

The assumption of known and constant down-sampling operator is a practical assumption, since in resolution enhancement applications, the enhancement factor is determined by the user and then the down-sampling operator is known. Also, even if in practical applications, the blurring operator is not known, it can be estimated by any of the blur estimation algorithms for example see [27].

III. Modified Horn-Schunck Optical Flow Estimation

In the conventional Horn-Schunck optical flow estimation algorithm, an equation that relates the changes in image brightness at a point to the motion of the brightness pattern has been derived. Assuming that the image brightness at the point (x,y) in the image plane at time t be denoted by E(x, y, t). The brightness of a particular point in the pattern is constant, so that

$$\frac{dE}{dt} = 0 \tag{2}$$

Then, by using chain rule for differentiation, it becomes

$$E_x u + E_y v + E_t = 0 \tag{3}$$

Where $u = \frac{\partial x}{\partial t}$, $v = \frac{\partial y}{\partial t}$, $E_x = \frac{\partial E}{\partial x}$, $y = \frac{\partial E}{\partial y}$, $E_t = \frac{\partial E}{\partial t}$. To allow smoothness of the optical flow, the smoothness measure term is added in the minimization process. The Laplacian of the x- and y-components of the flow are used as the smoothness measure function. So that the optical flow estimation problem is described as the minimization of

$$\varepsilon^2 = \varepsilon_b^2 + \alpha^2 \varepsilon_c^2,$$

Where

$$\varepsilon_b = E_x u + E_y v + E_t,$$

$$\varepsilon_c^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2,$$

α is the weighting parameter that weights the importance between ε_b and ε_c. This problem is solved iteratively as follow

$$u^{(n+1)} = u^{(n)} - E_x\left[E_x u(^{(n)} + E_y v^{(n)} + E_t\right] / [\alpha^2 + E_x^2 + E_y^2] \tag{4}$$

And

$$v^{(n+1)} = v^{(n)} - E_y\left[E_x u^{(n)} + E_y v^{(n)} + E_t\right] / [\alpha^2 + E_x^2 + E_y^2] \tag{5}$$

Where n is the iteration number.

The main modification of the Horn-Schunck algorithm is represented by the incorporation of the segmentation information in the optical flow estimation, which done in two stages. The two stages are described as follows.

A. Image Segmentation

Without loss of generality, we used the watershed transform for image segmentation in this work. The watershed transform is a region segmentation approach, in which the image is supposed to take high gradient values in the neighbourhood of edges and low gradient values for interior pixels. The segmentation can be obtained by removing some of weakest edges, which will create a number of lakes by grouping all the pixels that lie below a certain threshold. This can reduce the influence of noise and reduce the over-segmentation problem. It is then determined for each pixel in which direction the rain would flow if it would fall on the topographic activity surface. The segmentation process is done for the reference frame only to reduce the computational complexity.

The main steps of the watershed segmentation algorithm are summarized as follows (see [22] for details):

1) Define the floating point activity as $|\nabla f| = \sqrt{f_x^2 + f_y^2}$, where ($f_x$; f_y) is image gradient,
2) Assign a label for each pixel position,
3) For each pixel position find the weak edges, in which the floating point activity value is less than certain threshold (μ) (the thresholding is done for the floating point activity in the positions ep; swp; sp and sep as defined in Fig. 1),
4) Remove these weak edges by merging the corresponding labels,
5) For further merge regions and remove the effect of noise do; for each pixel position find the direction of rain-falling among the eight directions (direction of rain-falling is defined as the direction which has the smallest difference between the floating point activity of the central position (cp) and the other position (nwp; np; nep;wp; ep; swp; sp or sep),
6) Finally, the pixels that have the same label belong to the same region.

An example for segmented image is shown in Fig. 2. This figure shows the suitability of the watershed segmentation algorithm.

nwp	np	nep
wp	cp	ep
swp	sp	sep

Figure 1: A diagram showing central pixel and neighboring pixels.

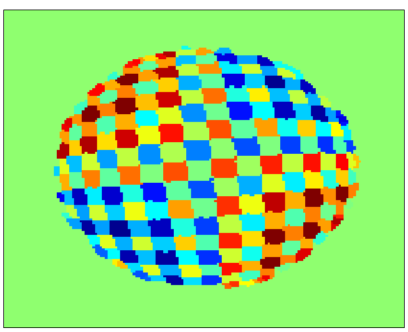


Figure 2: An example for segmented image form rotation sequence.

B. Segmentation-based Horn-Schunck Algorithm

As Fig. 2 shows an example for the segmented image where each colour represents a homogeneous region, our assumption is that regions' discontinuity is the motion discontinuity. So that, rather than estimating optical flow for each pixel as in [23,24,25] or estimation a single optical flow for each block by assuming that all pixels in a certain block [26] have the same motion, we assume that motion of all pixels in each homogeneous region is the same. Therefore, we suggest modifying Eqs. (3) and (4) to be

$$u_{\Re}^{(n+1)} = u_{\Re}^{(n)} - median_{x,y}\{E_x[E_x u^{(n)} + E_y v^{(n)} + E_t]/[\alpha^2 + E_x^2 + E_y^2]\} \quad \forall\, x,y \in \Re \tag{6}$$

And

$$v_{\Re}^{(n+1)} = v_{\Re}^{(n)} - median_{x,y}\{E_y[E_x u^{(n)} + E_y v^{(n)} + E_t]/[\alpha^2 + E_x^2 + E_y^2]\} \quad \forall\, x,y \in \Re \tag{7}$$

Where ($u_{\Re}^{(n+1)}$, $v_{\Re}^{(n+1)}$) is the motion vector for all pixels in the region $\Re$, $median_{x,y}\{.\}$ is the median value evaluation function. The median value is estimated done for all values of x, y in region $\Re$. Here we suggested using median function for its robustness against noise.

IV. Resolution Enhancement Based on Modified Optical flow Estimation

In this section, we illustrate the method for optimizing a cost function that consists of the error between the simulated LR frames and the interpolated observed LR frames.

The proposed cost function incorporates the effect of local motion.

A. Cost Function

We start with the resolution enhancement problem which can be described as an optimization problem as follows:

$$J_1[X] = \sum_{k=1}^{N} \left\| \mathbf{HF}^k X - \tilde{Y}^k \right\|_1 \tag{8}$$

where $||.||_1$ is the L_1-norm which describes the cost function measuring error and $\tilde{Y}^k$ is defined as the upsampled and interpolated frame from the observed LR frame Y^k:

$$\tilde{Y}^k = \boldsymbol{\mathcal{J}}(\mathbf{D}^T Y^k)$$

where $\boldsymbol{\mathcal{J}}(.)$ is the interpolation operator, which is defined on the missed pixels positions only.

On the other hand, the general traditional cost function directly comes from (8) is [2,4]

$$J_2[X] = \sum_{k=1}^{N} \left\| \mathbf{DHF}^k X - Y^k \right\|_1 \tag{9}$$

This cost function is so ill-posed that we need to add a regularization term in practice. Indeed, $J_1[X]$ and $J_2[X]$ are related and $J_1[X]$ includes ``less'' ill-posedness than $J_2[X]$ does as proved in [15]. The cost function $J_1[X]$ is modified as follows

$$J_3[X] = \sum_{k=1}^{N} \left\| \mathbf{HF}^k X - \tilde{Y}^k \right\|_1 + \lambda \left\| \mathbf{C} X \right\|_1 \tag{10}$$

Where $\mathbf{C}$ is a matrix that represents a Laplacian high pass operator and λ is the regularization parameter. The regularization term in (10) is added to solve the ill-posedness of the inverse problem. The regularization term incorporates the smoothness property of the HR frame.

B. Two Steps-Based Minimization of $J_3[X]$

The reconstruction of the HR image (X) can be divided into two independent steps, namely, the fusion and restoration steps. The fusion problem is described as

$$\hat{Z} = arg \min_{Z} \sum_{k=1}^{N} \left\| F_k Z - \tilde{Y}_k \right\|_1 \quad (11)$$

This can be obtained by median filtering as:

$$\hat{Z}(i,j) = \mathrm{median}_k (F_k^T \tilde{Y}_k)(i,j) \quad (12)$$

i.e. all the registered frames are fused by using median operator. Then, the problem is reduced to a restoration problem as follows:

$$\hat{Z} = \mathbf{H}\,\hat{X} \quad (13)$$

where $\hat{X}$ is the estimated version of X (the HR frame). $\hat{X}$ can be obtained from $\hat{Z}$ by using a restoration step as follows.

$$\hat{X} = \underset{X}{\mathrm{argmin}} \left[\left\| \mathbf{H}X - \hat{Z} \right\|_1 + \lambda \left\| \mathbf{C}X \right\|_1 \right] \quad (14)$$

The steepest decent solution to the minimization problem in (14) is:

$$\hat{X}^{(n+1)} = \hat{X}^{(n)} - \beta \Big\{ \mathbf{H}^T \mathrm{sign}(\mathbf{H}\hat{X}^{(n)} - \hat{Z}) + \lambda\, \mathbf{C}^T \mathrm{sign}(\mathbf{C}X^{(n)}) \Big\} \quad (15)$$

where β is a scalar representing the step size in the direction of the gradient.

V. Simulation Results

A. Data Sets

Two different video sequences including Table Tennis (352×240 SIF format) and Football (352×240 SIF format) are used to evaluate the proposed algorithm. For these sequences, the YUV components are available. Moreover, we assume that the coloured sequences are already demosaicked or captured by three CCD sensors. In addition to test the proposed modification for the Horn-Schunck optical flow estimation algorithm, we test the optical flow estimation with two famous sequences in the field of optical flow estimation, namely, the rotation and div sequences.

B. Experiment Setup

In the simulation, we assumed that the available sequences are HR sequences then we generate LR sequences from these sequences by applying blurring, down-sampling operators and adding noise to the HR sequences. Then we used different SR algorithm to reverse these operations.

The LR frames were generated from the original HR video sequences according to the model as in (1), where the frames were blurred by Gaussian operator (5×5) with the variance equal to 1, down-sampled by a decimation factor equals 2 in the horizontal and vertical directions, and distorted by an additive white Gaussian noise with 30 dB signal-to-noise ratio.

To demonstrate the efficiency of the proposed modification of the Horn-Shunck optical flow estimation algorithm, simulation results are presented for two different image sequences, with different motion types including rotation and div motions, in comparison with the traditional Horn-Schunck [23], and phase-based [25] optical flow estimation algorithms.

C. Optical flow Analysis

Optical flow estimation is one of the main contributions of this work. The efficiency of the proposed modification, which incorporates segmentation with the gradient-based optical flow estimation, is demonstrated as follows.

Fig. 3 shows the results of image sequence that includes rotation motion. Even if phase-based algorithm [25] estimates approximate motion for moving parts, it fails to estimate correct motion at the boundaries as shown in Fig. 3a, where parts of the background are estimated to have high motion.

Fig. 3b shows the results of conventional Horn-Schunck algorithm [23], from this figure, it fails to estimate correct motion at boundaries as shown by estimating motion in the background parts. The main problem of the Horn-Schunck algorithm is treated and solved with the proposed modification s shown in Fig. 3c.

Another example that demonstrates the efficiency of the proposed modification is shown in Fig. 4. This sequence contains div motion. The resulting motion vectors by using phase-based algorithm is scattered and inconsistent for the same region which shows the failure of this algorithm in this sequence as shown in Fig. 4a. Horn-Schunck suffers from the error motion vectors near the boundaries and smooth areas as shown in Fig. 4b. However, the proposed modification has solved this problem as shown in Fig. 4c.

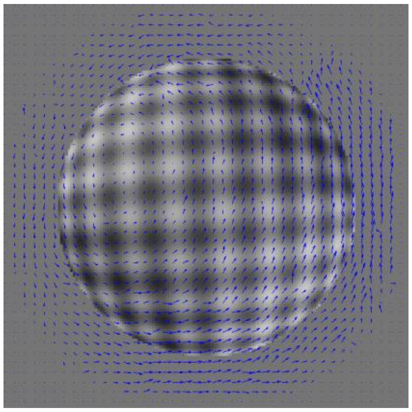
D. Resolution Enhancement Results

Fig. 5 shows the results of SR reconstruction for “Table Tennis” sequence. The segmented regions and the bicubic-interpolated frame are shown in Figs. 5a and 5b, respectively. In this figure, the estimated HR frames by incorporating different motion estimation algorithms are shown. From this figure, we can see that as a global motion model, affine motion is not suitable for sequences that contain locally moving objects. This is clear in the disappearance of the ball, as shown in Fig. 5c.

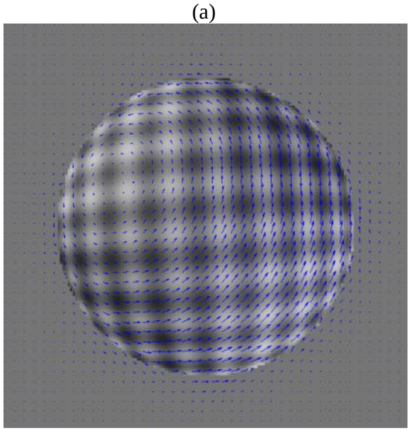
In addition, although conventional optical flow can estimate the flow of each pixel, is very sensitive to noise which results in a noisy edge as shown in Fig. 5d, where the ball is deformed due to the noise effect on the flow estimation process.

On the other hand, segmentation-based optical flow overcomes the problems of the conventional optical flow by assuming a constant motion vector in an arbitrary shaped region (as shown in Fig. 5e).

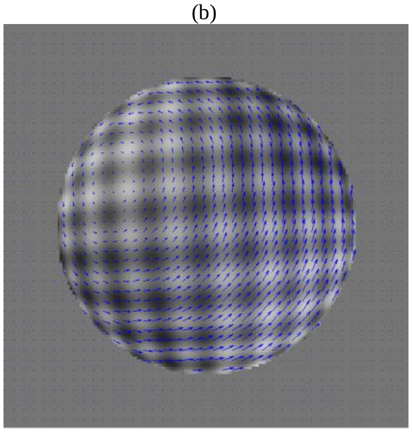
As another example, the results of Football sequence are shown in Fig. 6. In this figure, the failure of the affine motion is clearer since this sequence contains multi-objects and faster motion as shown in Fig. 6c. In case of using optical flow, the effect of noise is clear at edges, which are not sharp as shown in Fig. 6d. On the other hand, incorporating segmentation-based optical flow estimation preserves sharp edges as shown in Figs. 6e.



(a)

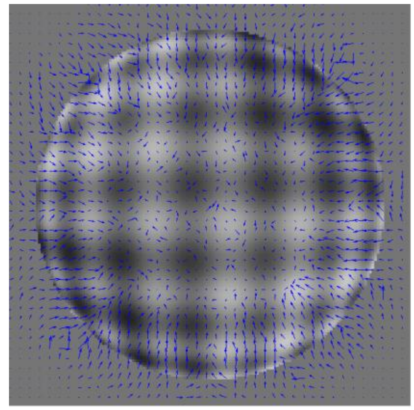


(b)

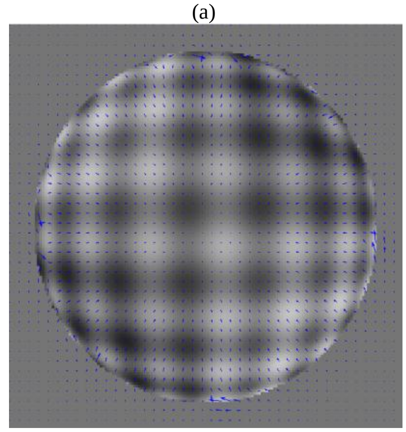


(c)

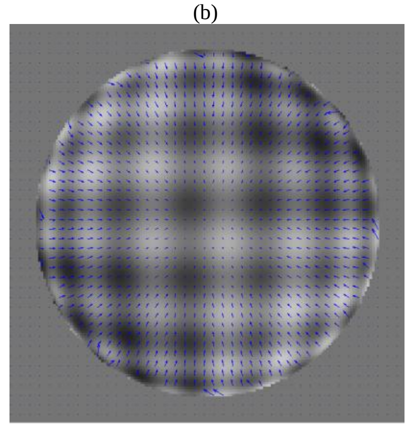
Figure 3: Image sequence including rotation motion, image with optical flow using; (a) phase-based algorithm [25], (b) conventional Horn-Schunck algorithm [23], and (c) the proposed segmentation -based Horn-Schunck algorithm.



(a)

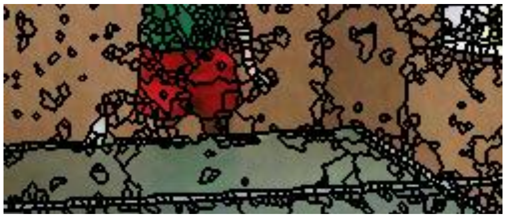


(b)

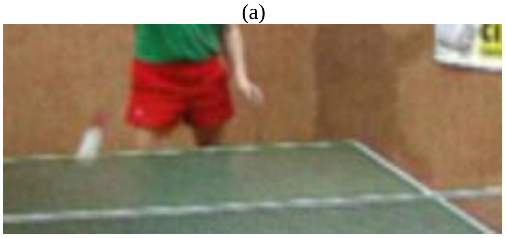


(c)

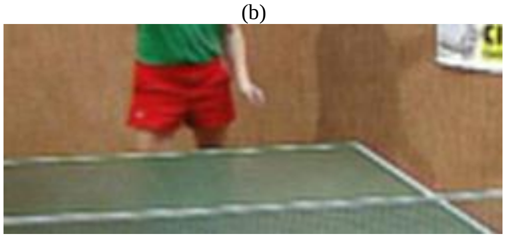
Figure 4: Image sequence including div motion, image with optical flow using; (a) phase-based algorithm [25], (b) conventional Horn-Schunck algorithm [23], and (c) segmentation-based Horn-Schunck algorithm.



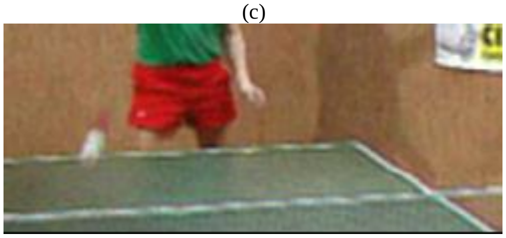
(a)



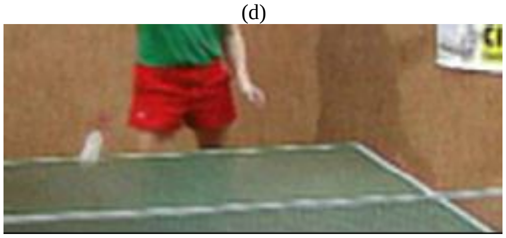
(b)



(c)



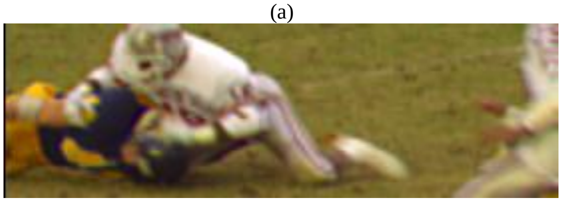
(d)



(e)

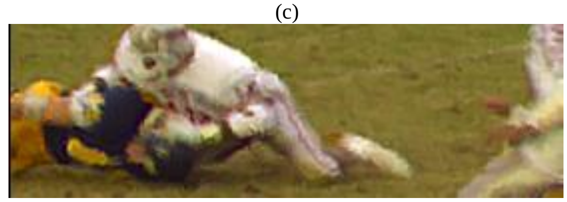
Figure 5, Table tennis sequence, (a) segmented regions, (b) bicubic interpolation, (c) estimated HR frame by incorporating affine motion estimation, (d) estimated HR frame incorporating conventional optical flow estimation, and (e) estimated HR frame by incorporating modified optical flow estimation.

(a)

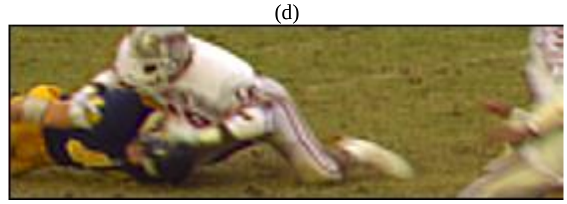


(b)

(c)



(d)



(e)

Figure 6, Football sequence, (a) segmented regions, (b) bicubic interpolation, (c) estimated HR frame by incorporating affine motion estimation, (d) estimated HR frame incorporating conventional optical flow estimation, and (e) estimated HR frame by incorporating modified optical flow estimation.

VI. CONCLUSION

This paper consists of two main contributions. First, we proposed a modification for the Horn-Schunck optical flow estimation algorithm to overcome problems of conventional gradient-based optical flow estimation algorithms, which are the handling of un-textured regions and the estimation of correct flow vectors near motion discontinuities. We assumed that motion of all pixels in each homogeneous region is the same. **Second**, we incorporated the proposed segmentation-based optical flow estimation in the resolution enhancement technique. The proposed algorithm overcomes some of the super-resolution problem for video sequences such as the non-rigidity problem where the assumption of general local motion estimates a different motion vector for different parts. The proposed algorithm gave promising results for low resolution sequences with slow/fast motion.

REFERENCES

[1] Z. Jiang, T.T. Wong and H. Bao, "Practical super-resolution from dynamic video sequences," in Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR 2003), Madison, Wisconsin, USA, June 2003.

[2] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Dynamic de-mosaicing and color super-resolution of video sequences," Proceedings of the SPIE conference on image reconstruction from incomplete data III, Vol. 5562, October 2004.

[3] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure transnational motion and common space invariant blur," IEEE Trans. on Image Processing, Vol.10, no.8, pp.1187-93, August 2001.

[4] S. Farsiu, M. Elad, P. Milanfar, "Multiframe demosaicing and super-Resolution of color images," IEEE Trans. On Image Processing, Vol. 15, No. 1, January 2006.

[5] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multi-frame super-resolution," IEEE Transactions on Image Processing, vol. 13, no. 10, pp. 1327-1344 , October 2004.

[6] S. Cheol Park, M. Kyu Park, “Super-resolution image reconstruction: a technical overview,” IEEE Signal Processing Magazine, May 2003.
[7] W. Zhao, H. Sawhney, "Is super-resolution with optical flow feasible?" Proc. ECCV, vol.1, pp.599-613, 2002.
[8] J. J. Clark, M. R. Palmer, and P. D. Lawrence, “A transformation method for the reconstruction of functions from non-uniformly spaced samples,” IEEE Trans. Acoustics Speech, Signal Processing, vol.33, no.4, pp. 1151-1165, 1985.
[9] S. P. Kim, N. K. Bose, and H. M. Valenzuela, “Recursive reconstruction of high resolution image from noisy under-sampled multi-frames,’ IEEE Trans. Acoustics Speech, Signal Processing, vol. 38, pp. 1013-1027, June 1990.
[10] M. C. Hong, M. G. Kang, and A. Katsaggelos, “An iterative weighted regularized algorithm for improving the resolution of video sequences,” Proc. Int. Conf. Image Processing, vol. 2, pp. 474-477, 1997.
[11] H. Stark and P. Oskoui, “High resolution image reconstruction from lower-resolution image sequences and space varying image restoration,” J. Opt. Soc. Am. A, vol. 6, pp. 1715-1726, 1989.
[12] S. Baker and T. Kanade, “Super resolution optical flow,” Tech. Rep. CMU-RI-TR-99-36, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, October 1999.
[13] D. Barreto, L.D. Alvarez, and J. Abad, “Motion estimation techniques in super-resolution image reconstruction a performance evaluation,” in Virtual Observatory: Plate Content Digitalization, Archive Mining and Image Sequence Processing, edited by Heron Press, vol. I, 254-268, April 2005.
[14] O. A. Omer and T. Tanaka, “Region-based weighted-norm with adaptive regularization for resolution enhancement,” Digital Signal Processing, doi:10.1016/j.dsp.2011.02.005.
[15] O. A. Omer and T. Tanaka, “Extraction of high-Resolution frame from low-resolution video sequence using region-based motion estimation,” IEICE Trans. Fundamentals, Vol. E93-A, No.4, April 2010.
[16] H. He, L. P. Kondi, “An image super-resolution algorithm for different error levels per frame,” IEEE Trans. Image Process., vol. 15, no. 3, pp. 592-603, March 2006.
[17] R. R. Schultz, R. L. Stevenson, “Extraction of high-resolution frames from video sequences,” IEEE Trans. Image Process., vol. 5, no. 6, pp. 996-1011, June 1996.
[18] Hardie, K.J. Barnard, J.G. Bognar, E. Armstrong, and E.A. Watson, “High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system,” Opt. Eng., vol. 37, no. 1, pp. 247-260, 1998.
[19] A. J. Patti, M. I. Sezan, and A. M. Tekalp, “Robust methods for high-quality stills from interlaced video in the presence of dominant motion,” IEEE Trans. Circuits and Systems for Video Technology, vol. 7, no. 2, pp. 328-342, April 1997.
[20] B. Choi, S.D. Kim and J.B. Ra, “Region-based super-resolution using adaptive diffusion regularization,” Opt. Eng., vol. 47, no. 2, pp. 027006-1-027006-11, February 2008.
[21] A. van Eekeren, K. Schutte, J. Dijk, D.J.J. de Lange and L.J. van Vliet, “Super-resolution on moving objects and background,” in Proc. Int. Conf. Image Process., October 2006, pp. 2709-2712.
[22] P. De Smet and D. De Vleschauwer, “Performance and scalability of highly optimized rain-falling watershed algorithm,” Proc. Int. Conf. on Imaging Science, Systems and Technology, CISST98, pp. 266-273, Las Vegas, NV, USA, July 1998.
[23] B. Horn and B. Schunck, “Determining optical flow,” Artificial Intelligence, 17:185–203, 1981.
[24] B. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” In International Joint Conference on Artificial Intelligence, pages 674–679, 1981.
[25] T. Gautama, and M.M. Van Hulle, “A phase-based approach to the estimation of the optical flow field using spatial filtering,” IEEE Trans. Neural Networks, Vol. 13, No. 5, pp. 1127-1136, 2002.
[26] T. Senst, V. Eiselein, R. H. Evangelio and T. Sikora, “Robust modified L2 local optical flow estimation and feature tracking,” Workshop on Applications of Computer Visions, PP. 685 – 690, 2011.
[27] Y. Yang, N. P. Galatsanos, and H. Stark, .Projection based blind de-convolution,. Optical Society of America, vol. 11, no. 9, pp. 2401- 2409, September 1994.

AUTHORS PROFILE

Osama A. Omer received his B.Eng. and M.Eng. degrees in electrical engineering from South Valley University, Aswan, Egypt, in 2000 and 2004 respectively. He received his Ph.D. degree from Tokyo University of Agriculture and Technology in 2009. He is now assistant professor at Aswan University. He spent the summer internship 2008 in Nokia/Tokyo research center. His interests include image/video super resolution and image/video compression.

Monte Carlo Ray Tracing Simulation of Polarization Characteristics of Sea Water Which Contains Spherical and Non-Spherical Particles of Suspended Solid and Phytoplankton

Kohei Arai 1
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Yasunori Terayama 2
Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*— Simulation method of sea water which contains spherical and non-spherical particles of suspended solid and phytoplankton based on Monte Carlo Ray Tracing: MCRT is proposed for identifying non-spherical species of phytoplankton. From the simulation results, it is found that the proposed MCRT model is validated. Also some possibility of identification of spherical and non-spherical shapes of particles which are contained in sea water is shown. Meanwhile, simulations with the different shape of particles, Prolate and Oblate show that Degree of Polarization: DP depends on shapes. Therefore, non-spherical shape of phytoplankton can be identified with polarization characteristics measurements of the ocean.**

Keywords- Monte Carlo Ray Tracing; Phytoplankton; Polarization characteristics.

I. INTRODUCTION

Red tide detection is getting more important. Due to some reason including global warming, red tide occurs more frequently and widely all over the coastal zone in the world. Harmful Algal Blooms: HABs [1] are getting more severe situation. There are two types of red tide, spherical and non-spherical shapes of phytoplankton[2] induced red tide. There is specific red tide phytoplankton to the coastal zones. For instance, *Chattonella Antiqua* [3], *Chattonella Marina* [4] is dominant phytoplankton in the Ariake Sea[5] area, in Kyushu Island, Japan. Almost every summer, *Chattonella* phytoplankton appears in the Ariake Sea area. The shapes of these phytoplankton are different each other. Namely, *Chattonella Antiqua* has non-spherical shape while *Chattonella Marina*, *Chattonella Globosa* [6] has relatively spherical shape. There is a strong demand on discrimination of spherical and non-spherical shapes of phytoplankton because it provides information on poison containing phytoplankton in some cases. It is possible to discriminate spherical and non-spherical shapes of phytoplankton by using polarization characteristics [1],[2]. Degree of polarization: DP [3] of the spherical shape of particles is almost zero while that of non-spherical shape of particles is greater than zero depending upon its spherical degree. Therefore, it is also possible to identify non-spherical shapes of phytoplankton through polarization measurements of sea surface with polarization radiometers or polarization cameras[7].

Simulation study on discrimination of spherical and non-spherical particles can be done with Monte Carlo Ray Tracing: MCRT method [4]-[15]. MCRT simulation model for the atmosphere is proposed [4] while MCRT model for aerosol particle is proposed [5]-[7]. Meanwhile, MCRT for layered clouds is discussed [8] while MCRT models for representation of forest parameters are proposed [9]-[13]. In particular, forest Leaf Area Index: LAI model based on MCRT is proposed [14] while Bi-Directional Reflectance Distribution Function: BRDF model with MCRT is proposed for tea trees [15].

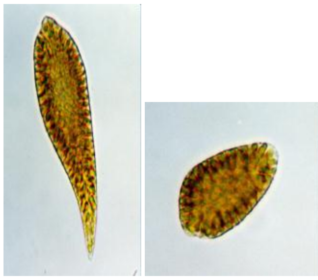
In this paper, simulation method is proposed then simulation results which show a validity of the method is described followed by some field experimental data which is also support a validity of the method. Then conclusion and some discussions are described finally.

II. PROPOSED METHODS

A. Research Background

The shapes of *Chattonella Antiqua*, *Chattonella Marina* are shown in Fig.1. *Chattonella Antiqua* has non-spherical shape while *Chattonella Marina* has relatively spherical shape as shown in Fig.1 clearly.

[1] http://en.wikipedia.org/wiki/Algal_bloom
[2] http://en.wikipedia.org/wiki/Phytoplankton
[3] http://www.minc.ne.jp/suishi/page022.html
[4] http://www.minc.ne.jp/suishi/page009.html
[5] Kohei Arai, Red Tides -Combines satellite and ground based detection-, SPIE Newsroom, 10/1117/2, 1201012, 003267, 2011.
[6] http://www.marinespecies.org/hab/aphia.php?p=taxdetails&id=246589
[7] http://www.sciencedirect.com/science/article/pii/026288569594383B



(a) *Chattonella Antiqua* (b)*Chattonella Marina*

Figure 1 The shapes of *Chattonella Antiqua*, and *Chattonella Marina*

Polarization characteristics for both spherical and non-spherical phytoplankton are different obviously. Therefore, it might be possible to identify non-spherical phytoplankton through polarization measurement for sea surface with polarization radiometer or polarization camera. There are influences due to ocean waves, polarization on the sea surface, etc. on polarization measurement data. In order to avoid the influence due to ocean wave, the pixels of which normal direction of the sea surface is directed to zenith are extracted from the polarization camera images. Meanwhile, the influence due to polarization effect at the sea surface can be avoided by setting the looking angle at 53 degree (Brewster angle [3] of sea surface). Thus the polarization characteristics of the sea surface in concern are measured and estimated.

B. Monte Carlo Ray Tracing: MCRT

In order to show a validity of the proposed method for non-spherical phytoplankton, MCRT simulation study and field experimental study is conducted. MCRT allows simulation of polarization characteristics of sea surface with designated parameters of the atmospheric conditions and sea surface and sea water conditions. Illustrative view of MCRT is shown in Fig.2.

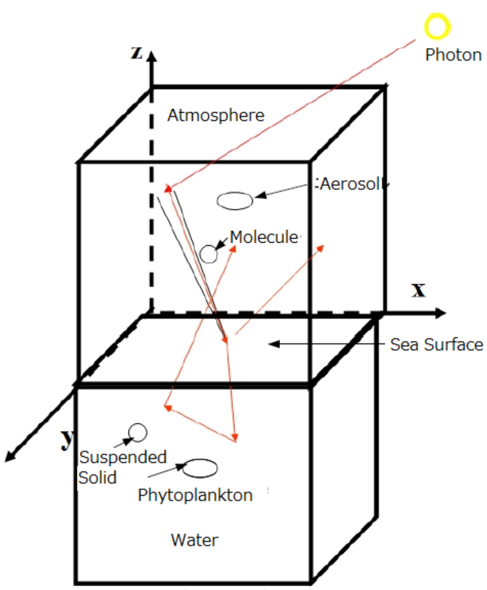
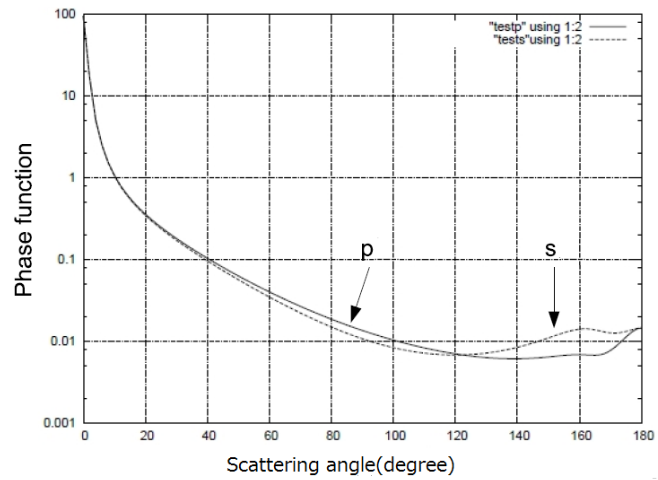
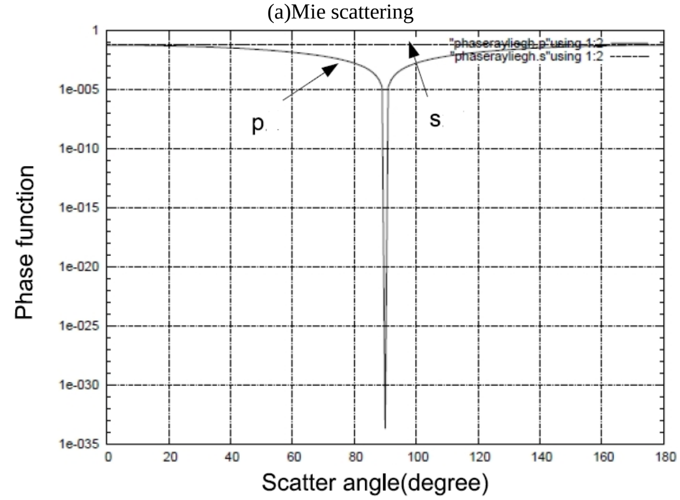


Figure 2 Illustrative view of MCRT for the atmosphere and sea water

Photon from the sun is input from the top of the atmosphere (the top of the simulation cell). Travel length of the photon is calculated with optical depth of the atmospheric molecule and that of aerosol. There are two components in the atmosphere; molecule and aerosol particles while three are also two components, water and particles; suspended solid and phytoplankton in the ocean. When the photon meets molecule or aerosol (the meeting probability with molecule and aerosol depends on their optical depth), then the photon scattered in accordance with scattering properties of molecule and aerosol. The scattering property is called as phase function. In the visible to near infrared wavelength region, the scattering by molecule is followed by Rayleigh scattering law [3] while that by aerosol is followed by Mie scattering law [3]. Example of phase function of Mie scattering is shown in Fig.3 (a) while that of Rayleigh scattering is shown in Fig.3 (b).



(a)Mie scattering

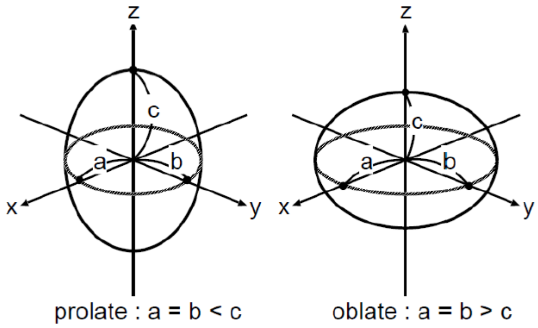


(b)Rayleigh scattering

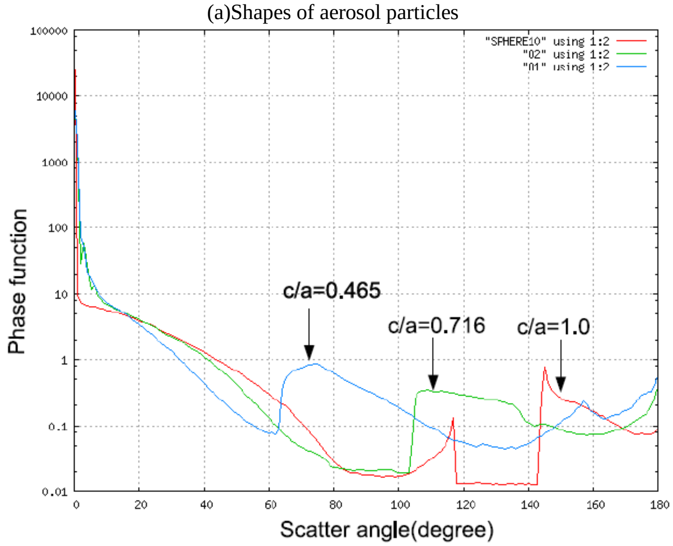
Figure 3 Phase functions for Mie and Rayleigh scattering

In the figure, scattering angle is defined as the angle between incidence and reflected angle from the particle. These phase functions can be calculated with Mie Code in the MODerate resolution atmospheric TRANsmission; MODTRAN[8]. In particular, Mie phase function depends on refractive index, size distribution, and shape of the aerosol particle in concern. For instance, phase function of prolate shape of particle differs from that of oblate shape of particle as shown in Fig.4.

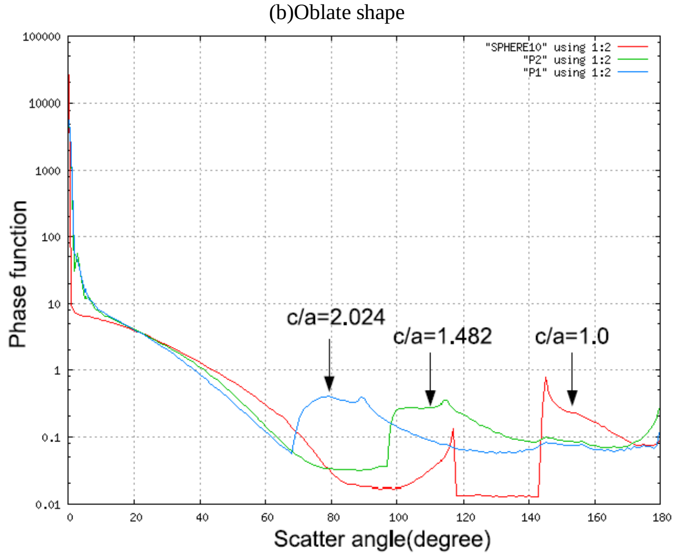
[8] http://modtran.org/



(a)Shapes of aerosol particles



(b)Oblate shape



(c)Prolate shape

Figure 4 Two types of shape of aerosol particle and their phase functions as a function of the ratio of long radius and short radius.

The shapes of suspended solid and phytoplankton are assumed to be oblate or prolate of elliptical shape of which c/a ratio and orientation angle are to be random. In the MCRT simulation, these are aligned with the certain average interval (density of such materials) with the different c/a ratio and orientation together with refractive index of suspended solid and phytoplankton[9].

Reflection, transpiration, and refraction of the photon at the sea surface are followed by Fresnell law [3] and Snell law [3] as shown in Fig.5.

Thus reflectance and transparency is calculated for p and s polarization separately together with total reflectance and transparency, Root Sum Square of p and s components as shown in Fig.6. Then the following Degree of Polarization: DP is calculated.

$$DP = (p-s)/(p+s) \tag{1}$$

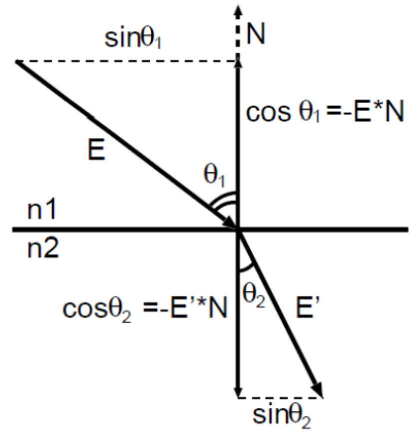


Figure 5 Reflection, transpiration, and refraction of the photon at the sea surface

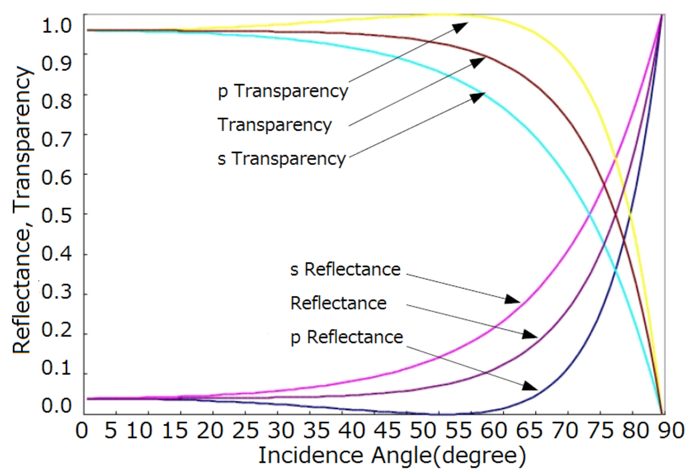


Figure 6 Reflectance and transparency is calculated for p and s polarization separately together with total reflectance and transparency, Root Sum Square of p and s components

III. Experiments

A. *Overall Characteristics DP of Sea Surface at the Brewster Angle of Observations*

Fig.7 shows an example of MCRT simulation result of DP calculation at around Brewster angle of observation angle as a function of scattering angle. At the Brewster angle, p reflectance is zero while s reflectance is around 0.14 as shown

[9] In this simulation, both refractive index of phytoplankton and suspended solid are assumed to be same.

in Fig.6. Therefore, theoretical DP reaches at 1.0 and gradually decreases in accordance with increasing of scattering angle. Meanwhile, theoretical DP is increased with increasing of scattering angle ranges from zero to Brewster angle.

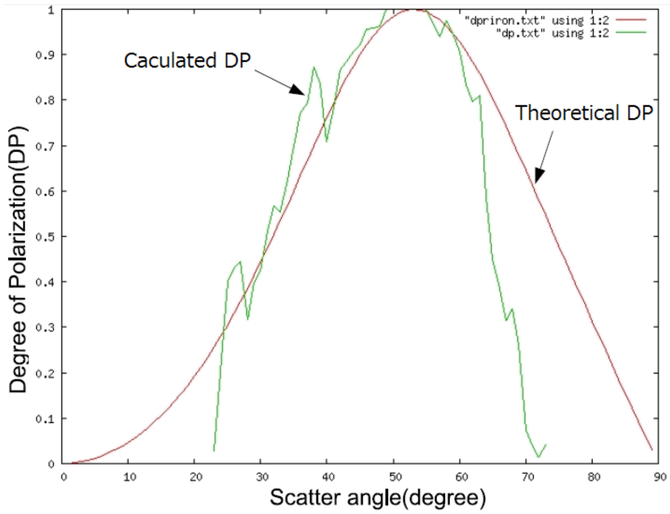


Figure 7 Example of MCRT simulation result of DP calculation at around Brewster angle of observation angle as a function of scattering angle

On the other hand, calculated DP (MCRT simulation result) shows the same trend with theoretical DP, the details are different though. MCRT simulation result of calculated DP goes up and down everywhere. One of the causes for these variations is that the number of photons is not enough. One of the conclusions from the MCRT simulation study is that observation angle has to be set at Brewster angle. Thus the polarization characteristics of the sea surface can be observed at the best observation condition.

B. Reflectance Calculated Based on Monte Carlo Ray Tracing: MCRT

At the observation angle of 53 degree of Brewster angle, p and s polarization reflectance are estimated with MCRT simulation. Table 1 shows the results.

TABLE I. P AND S REFLECTANCE FOR THE DIFFERENT OPTICAL DEPTH

Optical Depth	P Reflectance	S Reflectance
0.01	0.07425	0.7641
3.0	0.1218	0.8798
5.0	0.1509	0.9106

As shown in Fig.6, s reflectance is always greater than p reflectance. The result shows the same trend. In the simulation, optical depth of the particles (suspended solid and phytoplankton) is changed from 0.01, 3.0, and 0.5. On the other hand, reflectance of the sea surface is set at 0.1. This implies that the transparency of sea water is set at 0.9. Thus DP can be calculated through p and s reflectance (using the difference of the number of photons for p and polarization of MCRT simulation results).

Table 2 shows total reflectance of particles containing sea water and sea water only as well as the deviations of reflectance and DP from the spherical shape of particle and from the reflectance for the sea water only. In the simulation, optical depth of the particles of suspended solid and phytoplankton is set at 3.0. Shapes of the particles of P1, P2 are set at 1.95 and 4.62 of c/a ratio while those of O1, O2 are set at 0.46 and 0.24, respectively.

Particle size of these suspended solid and phytoplankton are same as spherical water. Therefore, equivalent scattering cross section of sea water only and of particle containing sea water are same.

TABLE II. TOTAL REFLECTANCE OF PARTICLES CONTAINING SEA WATER AND SEA WATER ONLY AS WELL AS THE DEVIATIONS OF REFLECTANCE AND DP FROM THE SPHERICAL SHAPE OF PARTICLE AND FROM THE REFLECTANCE FOR THE SEA WATER ONLY WHEN THE PARTICLE SIZE IS NOT CHANGED DEPENDING UPON C/A RATIO.

Shape	Reflectance	Deviation of reflectance (%)	Deviation of DP
O1	0.5	19.75	0.0136
O2	0.5	19.74	0.0127
Sphere	0.499	19.51	-
Sea Water	0.418	-	-
P2	0.5	19.74	0.0132
P1	0.5	19.75	0.0134

Reflectance deviation of spherical shape of particle from the reflectance of sea water only is 19.51% while that of non-spherical particles is around 19.75%. It is said that the reflectance of particles containing sea water is much different from that of sea water only while shape dependency of the reflectance is comparatively small, much smaller than that of the reflectance difference between sea water and particle containing sea water.

Meanwhile, the deviation of DP from the spherical shape of particle is varied depending upon c/a ratio in some extents. As is aforementioned, the particle orientations are random and are varied by time being. There are a variety of sizes for particles. Deviations of DP for the same size of different shapes and orientation angles of particles range from 0.0127 to 0.0136. Therefore, not so large influence due to the different shapes and orientation angles is suspected.

If the particle size is varied by the shape of particle, then the simulation result is remarkably changed from that of the case that the particle size is same even for the different shape. Table 3 shows the MCRT simulation results for that situation.

Due to the fact that scattering cross section is varied by the c/a factor, reflectance of the sea surface is changed with c/a factor as well as DP. The deviation of DP is totally depends on c/a ratio in this case. Because the averaged particle interval is same, cross section is dependent to c/a ratio.

TABLE III. TOTAL REFLECTANCE OF PARTICLES CONTAINING SEA WATER AND SEA WATER ONLY AS WELL AS THE DEVIATIONS OF REFLECTANCE AND DP FROM THE SPHERICAL SHAPE OF PARTICLE AND FROM THE REFLECTANCE FOR THE SEA WATER ONLY WHEN THE PARTICLE SIZE IS CHANGED DEPENDING UPON C/A RATIO.

Shape	Reflectance	Deviation of reflectance (%)	Deviation of DP
O1	0.577	57.70	0.0167
O2	0.575	57.51	0.00379
Sphere	0.169	16.87	-
Sea Water	0.1	-	-
P2	0.136	13.58	0.00199
P1	0.127	12.66	0.00298

IV. Conclusion

Monte Carlo Ray Tracing: MCRT based simulation method of sea water which contains spherical and non-spherical particles of suspended solid and phytoplankton is proposed. Using the proposed method, identifying non-spherical species of phytoplankton is available. From the simulation results, it is found that the proposed MCRT model is valid. Also some possibility of identification of spherical and non-spherical shapes of particles which are contained in sea water is shown. Meanwhile, simulations with the different shape of particles, Prolate and Oblate show that Degree of Polarization: DP depends on shapes. Therefore, non-spherical shape of phytoplankton can be identified with polarization characteristics measurements of the ocean.

Acknowledgment

The authors thank laboratory staff and members for their discussions and valuable comments as well as suggestions through MCRT simulation study. In particular, the author thanks Miss.Yoko Takamatsu for her great effort to create phase function of suspended solid and phytoplankton particles.

References

[1] Kohei Arai and Yasunori Terayama, A method for red tide detection and discrimination of red tide type(spherical and non-spherical shapes of red tide) through polarization measurements of sea surface, International Journal of Applied Science, 2, 3, 62-70, 2011

[2] Kohei Arai and Yasunori Terayama, Satellite and Ground based red tide detection method and system by means of peak shift of remote sensing reflectance, International Journal of Applied Science, 2, 3, 53-61, 2011

[3] Kohei Arai, Lecture Note on Remote Sensing, Morikita-Shuppan Publishing Incorporation, 2004.

[4] Tsutomu Takashima, Kazuhiko Masuda and Kohei Arai, Measurements of skylight polarization with surface reflectance over Railroad Valley, Nevada, Papers in Meteorological Geophysics, Vol.50, No.3, pp.113-124, 1999

[5] Kohei Arai, Influence of non-spherical shape of sea salt aerosol particle on the top of the atmosphere radiance, Journal of Japanese Remote Sensing Society, Vol.20, No.5, pp.32-39, (2001).

[6] Kohei Arai and Yoko Takamatsu, Phase function of relatively large size of sea salt aerosol particle containing bubbles, Journal of Japanese Remote Sensing Society, Vol.22, No.1, pp.40-49, 2002.

[7] Kohei Arai, Phase function of relatively large aerosol particles containing bubbles by means of a Ray Tracing, Advances in Space Research, Vol.29, No.19, 1813-1818, 2002.

[8] Kohei Arai, Adjacency effect of layered clouds estimated with Monte-Carlo simulation, Advances in Space Research, Vol.29, No.19, 1807-1812, 2002.

[9] Kohei Arai and Yoko Takamatsu, Influence of spume aerosol particle on the top of the atmosphere radiance, Journal of Japanese Remote Sensing Society, Vol.23, No.4, pp.355-363, 2003.

[10] Shigeki Kuzuoka and Kohei Arai, Forest parameter estimation based on Monte Carlo Ray Tracing, Journal of Japanese Remote Sensing Society, Vol.22, No.4, pp.394-404, 2003.

[11] Kohei Arai and Yaliu Ding, Reflection model taking into account multiple reflection between trees, Journal of Japan Society of Photogrammetry and Remore Sensing, 45, 6, 56-66.2006

[12] Yaliu Ding and Kohei Arai, Forest parameter estimation based on MCRT based radiative transfer model, Journal of Japanese Remote Sensing Society, 27, 2, 141-152, 2007

[13] Yaliu Ding, Kohei Arai, Forest parameter estimation, by means of Monte-Carlo simulations with experimental consideration of estimation of multiple reflections among trees depending on forest parameters, Advances in Space Research, 43, 3, 438-447, 2009.

[14] Kohei Arai and Yui Nishimura, Polarization model for discrimination of broad and needle shaped leaves and estimation of LAI using polarization measurements, Advances in Space Research, 44, 510-516, 2009.

[15] Kohei Arai, Monte Carlo ray tracing simulation for bi-directional reflectance distribution function and grow index of tealeaves estimation, International Journal of Research and Reviews on Computer Science, 2, 6, 1313-1318, 2011.

Authors Profile

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science, and Technology of the University of Tokyo from 1974 to 1978 also was with National Space Development Agency of Japan (current JAXA) from 1979 to 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post-Doctoral Fellow of National Science and Engineering Research Council of Canada. He was appointed professor at Department of Information Science, Saga University in 1990. He was appointed councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was also appointed councilor of Saga University from 2002 and 2003 followed by an executive councilor of the Remote Sensing Society of Japan for 2003 to 2005. He is an adjunct professor of University of Arizona, USA since 1998. He also was appointed vice chairman of the Commission "A" of ICSU/COSPAR in 2008. He wrote 30 books and published 332 journal papers.

Power Analysis Attacks on ECC: A Major Security Threat

Hilal Houssain, Mohamad Badra
LIMOS Laboratory
CNRS
France

Turki F. Al-Somani, Senior Member, IEEE
Computer Engineering Department
Umm Al-Qura University
Makkah, Saudi Arabia

***Abstract*— Wireless sensor networks (WSNs) are largely deployed in different sectors and applications, and Elliptic Curve Cryptography (ECC) is proven to be the most feasible PKC for WSN security. ECC is believed to provide same level of security such as RSA with a much shorter key length, and thus they seem to be ideal for applications with small resources such a sensor network, smartcard, RFID, etc. However, like any other cryptographic primitive, ECC implementations are vulnerable to Power Analysis Attacks (PAAs) that may reveal the secret keys by exploiting leaked power consumption from running cryptographic devices (e.g. smart cards, mobile phones etc.). In this paper, we present a comprehensive study of major PAAs and its countermeasures on ECC cryptosystems. In addition, this paper describes critical concerns to be considered in designing PAAs on ECC particular for WSNs, and illustrates the need to conduct, in the near future, intensive researches for the development of these specific PAAs.**

Keywords- Wireless Sensor Networks (WSNs); Elliptic curve cryptosystems (ECC); Side-channel attacks (SCA); Scalar multiplication.

I. INTRODUCTION

Wireless sensor networks (WSNs) [1] [2] are ad hoc networks comprised of a large number of low-cost, low-power, and multi-functional sensor nodes and one or more base stations sensors that collaboratively monitor physical and environmental conditions. There a wide range of applications for WSN, such as health monitoring, industrial control, environment observation, as well as office and even military operations. In most of these scenarios, critical information is processed and frequently exchanged among sensor nodes through insecure wireless channels. It is therefore crucial to add security measures to WSNs for protecting its data against using Public Key Cryptography (PKC) that is shown to be feasible in WSNs (e.g., [3] [4] [5]) by using Elliptic Curve Cryptography (ECC).

Side-Channel Attacks (SCA), introduced by Paul Kocher in 1999 [6] [7], exploit leaked side-channel information, such as power consumption, electromagnetic emanations and running time etc., from running cryptographic devices such as WSNs, smart cards, mobile phones, RFIDs etc., to reveal the secret keys. Although recently Kanthakumar et al. [8] in 2008, discussed the security of WSNs against SCAs; But so far, however, no serious research effort has focused on the SCAs on WSNs, and in specific with the presence of ECC cryptosystems.

In this paper, we present a comprehensive study of major PAAs and its countermeasures on ECC cryptosystems, taking into consideration the WSNs resource constraints. The rest of the paper is organized as follows. In section 2, we present a background on WSNs. Section 3 presents ECC and its use in various fields. PAAs and its countermeasures on ECC are then reviewed in sections 4 and 5. Discussion on the findings of the paper study is presented in section 6. Section 7 concludes the presented study.

II. WIRELESS SENSOR NETWORKS

WSNs [1] [2] comprise mainly of a large number of small sensor nodes with limited resources and are based around a battery powered microcontroller. Wireless sensors are equipped with a radio transceiver and a set of transducers through which they acquire data about the surrounding environment. WSNs form an ad-hoc multi-hop network, where nodes communicate with each other and with one or more sink nodes that interact with the outside world. Sensors in the WSN can receive commands via the sink to execute tasks such as data collection, processing and transfer. The number of nodes participating in a sensor network is mainly defined by several requirements such as the network connectivity and coverage, and the size of the area of interest.

There exist a large number of different applications for WSN: examples are health monitoring, industrial control, environment observation, as well as office and even military applications. For example, in the health monitoring applications, WSN can be used to remotely monitor physiological parameters, such as heartbeat or blood pressure of patients, and send a trigger alert to the concerned doctor according to a predefined threshold. In addition, sensor nodes may be deployed in several forms: at random, or installed at deliberately chosen spots.

III. ELLIPTIC CURVE CRYPTOGRAPHY

Here we present a brief introduction to elliptic curves. Let $GF(2^m)$ be a finite field of characteristic two. A nonsupersingular elliptic curve E over $GF(2^m)$ is defined to be the set of solutions (x, y) ϵ $GF(2^m)$ X $GF(2^m)$ to the equation,

$$y^2 + xy = x^3 + ax^2 + b, \qquad (6)$$

Where a and b ϵ $GF(2^m)$, $b \neq 0$, together with the point at infinity denoted by O. It is well known that E forms a commutative finite group, with O as the group identity, under the addition operation known as the tangent and chord method. Explicit rational formulas for the addition rule involve several arithmetic operations (adding, squaring, multiplication and inversion) in the underlying finite field. In affine coordinates, the elliptic group operation is given by the following:

Let $P = (x_1, y_1)$ ϵ E; then $-P = (x_1, x_1 + y_1)$.

For all P ϵ E, O + P = P + O = P. If $Q = (x_2, y_2)$ ϵ E and $Q \neq -P$, then $P + Q = (x_3, y_3)$,

Where

$$x_3 = (\frac{y_1 + y_2}{x_1 + x_2})^2 + \frac{y_1 + y_2}{x_1 + x_2} + x_1 + x_2 + a \qquad (7)$$

$$y_3 = (\frac{y_1 + y_2}{x_1 + x_2}) \cdot (x_1 + x_3) + x_3 + y_1 \qquad (8)$$

if $P \neq Q$ and,

$$x_3 = x^2{}_1 + \frac{b}{x^2{}_1} \qquad (9)$$

$$y_3 = x^2{}_1 + (x_1 + \frac{y_1}{x_1})x_3 + x_3 \qquad (10)$$

if P = Q.

Computing P + Q is called elliptic curve point addition (PADD) if $P \neq Q$ and is called elliptic curve point doubling (PDBL) if P = Q. Point subtraction is a useful operation in some algorithms. This operation can be performed with the PADD or PDBL formulas using the additive inverse of the point to be subtracted. For example, the point subtraction P − Q can be computed using the PADD operation where: P − Q = P + (−Q). The additive inverse of a point P = (x, y) is the point (x, x + y) for curves defined over the $GF(2^m)$fields.

Scalar multiplication is the basic operation for ECC, called ECSM (Elliptic Curve Scalar Multiplication). ECSM in the group of points of an elliptic curve is the analogous of exponentiation in the multiplicative group of integers modulo a fixed integer m. Computing kP can be done with the straightforward double-and-add method [9], as described in Algorithm 1 (Left to Right approach), and Algorithm 2 (Right to Left approach) based on the binary expression of $k = (k_{m-1},\ldots,k_0)$ where k_{m-1} is the most significant bit of k. However, several Elliptic Curve Scalar Multiplication (ECSM) methods have been proposed in the literature. A good survey is presented in [9].

Algorithm 1: The double-and-add method.

(Left to Right)

Inputs: *P*: Base Point, *k*: Secret key.

Output: *kP*.

1: Q ← P

2: for i = m − 2 downto 0 do

3: Q ← 2Q

4: if $k_i = 1$ then Q ← Q + P

Return Q.

Projective coordinate systems define points over the projective plane as triplets (*X*, *Y*, *Z*). Projective coordinate systems are used to eliminate the need for performing inversion. For elliptic curve defined over $GF(2^m)$, many different forms of formulas are found [10] for PADD and doubling. The projective coordinate system (*Pr*), so called homogeneous coordinate system, takes the form $(x, y) = (X/Z, Y/Z)$, while the Jacobian coordinate system takes the form $(x, y) = (X/Z^2, Y/Z^3)$ and the Lopez-Dahab coordinate system takes the form $(x, y) = (X/Z, Y/Z^2)$. The Mixed coordinate system, on the other hand, adds two points where one is given in a certain coordinate system while the other is given in another coordinate system. The coordinate system of the resulting point may be in a third coordinate system [10].

IV. POWER ANALYSIS ATTACKS (PAA) ON ECC

In 1996, Paul Kocher introduced the power analysis procedure; then, in 1999 he introduced the PAAs. These attacks have become a major threat against tamper resistant devices [6]. PAA [6] [7] allow adversaries to obtain the secret key in a cryptographic device, or partial information on it, by observing the power consumption traces. This is a serious threat especially to mobile devices such as WSNs, smart cards, mobile phones, RFIDs etc. Thus, implementers need algorithms that are not only efficient, but also PAA-resistant.

Two main PAA techniques are the Simple PAA (SPAA) and Differential PAA (DPAA).

A. *Simple Power Analysis Attack (SPAA)*

The main idea of the SPAA [7] is to get the secret d using the side-channel leakage information obtained through observing the power consumption from a single measurement trace. For instance, as ECSM is the basic operation for ECC, and the most straightforward algorithm for point multiplication on an elliptic curve is the double-and-add algorithm (See Algorithm 1 in Section III), where a PDBL is executed for each bit of the scalar and a PADD is executed only if the scalar bit is equal to one. If the power consumption trace pattern of PDBL is different from that of PADD, the side-channel leakage of the implementation reveals the presence of the PADD and thus the value of the scalar bits and attackers can easily retrieve the secret key from a single side-channel trace.

B. Differential Power Analysis Attack (DPAA)

In DPAA [7], the adversary makes use of the obvious variations in the power consumption that are caused by multiple data and operation computations, and use statistical techniques to pry the secret information. This attack uses a two round technique: data collection and data processing. A DPAA on ECSM is described in [11].

More advanced DPAA techniques applicable to elliptic curve cryptosystems, such as refined power analysis (RPA) [12], zero power analysis (ZPA) [13], and doubling attacks [14] were introduced.

a) RPA (also called Goubin-type DPA) [12] attack directs its attention to the existence of a point P0 on the elliptic curve E(K) such that one of the coordinates is 0 in K and P0≠O. RPA could deduce the next bit of the scalar by computing power consumption of chosen message and some chosen points on the elliptic curve.

b) ZPA attack [13] is an extension of RPA attack. This attack is based on the observation that that even if a point had no zero-value coordinate; the auxiliary register might take on a zero-value. Thus with this attack, all points with zero power consumption are noticeable.

c) Doubling attack (DA) [14] attack is based on the two queries; one is on some input P and the other one is on 2P. The DA can detect when the same operation is done twice, i.e., exploits the similar (PDBL) operations for computing dP and d(2P). There are two types of DA, normal and relative DA (relative doubling attack proposed by Yen et al. [15]), where the relative DA uses a totally different approach to derive the key bit in which the relationship between two adjacent key bits can be obtained as either $d_i = d_{i-1}$ *or* $d_i \neq d_{i-1}$.

In addition, Template Attack [16] is very similar to DPAA (Two rounds technique: Template building and matching), but requires access to a fully controllable device. In Template building phases (also called profiling phase), the attacker constructs a precise model of the wanted signal source, including a characterization of the noise. The matching phase comprises the actual attack. Another attack is the Carry-based Attack (CBA) [17]; it does not attack the ECSM itself but its countermeasures. The CBA depends on the carry propagation occurring when long-integer additions are performed as repeated sub-word additions. Moreover, an advanced statistical technique such as Principal Component Analysis (PCA) [18] can be used by an attacker to perform PCA transformation on randomly switched PADD and PDBL (as in ECSM using Montgomery ladder) and identify the key bit.

V. Countermeasures Of Paa On Ecc

Since 1996, many research efforts [19] [11] [20] [21] [22] [23] [24] [25] [26] [27] have been made to secure ECC method implementations, in special the ECSM, against PAAs. The Major challenge is to avoid additional computational cost, and to develop relatively fast cryptosystems without compromising security, due to the nature of WSNs as constrained devices.

A. Countermeasures of SPAA on ECC

There are different strategies to resist SPAA attacks. These strategies share the same objective, which is to render the power consumption traces that are caused by the data and operation computations during an ECSM independent from the secret key.

SPAA attacks can be prevented by using one of the following methods:

1) Making the group operations indistinguishable (by processing of bits "0" and "1" of multiplier indistinguishable by inserting extra point operations). As an example, the double-and-add-always algorithm, introduced in [11] (As shown in Algorithm 2), and Montgomery ladder [20] (as shown in Algorithm 3) ensures that the sequence of operations appear as an PADD followed by a PDBL regularly.

Double-and-add-always algorithm [11] is highly regular, and it requires no pre-computation or prior recoding. This algorithm requires n PDBL and n PADDs regardless of the value of the scalar multiplicand, and two temporary registers are needed to store the results of each iteration.

As for the Montgomery ladder [20], the execution time of the ECSM is inherently unrelated to the Hamming weight of the secret scalar, and this algorithm avoids the usage of dummy instructions. Montgomery ladder [20] resists the normal DA. However, it is attacked by the relative DA proposed by Yen et al. [15]. Moreover, recent studies have shown that processing the bits of multiplicand from left-to-right, as Montgomery ladder does, are vulnerable to certain attacks [14].

Algorithm 2: Double-and-add-always.

Inputs: *P*: Base Point, *k*: Secret key.

Output: *kP*.

1: R[0]←O.

2: for i = l − 1 downto 0 do

3: R[0] ← 2R[0], R[1] ← R[0] + P.

4: R[0] ← R[ki].

5: end for

Return R[0].

Algorithm 3: Montgomery powering ladder.

Inputs: *P*: Base Point, *k*: Secret key.

Output: *kP*.

Scalar Multiplication (*kP*):

1: R[0] ← P, R[1] ← 2P

2: for *i* from *l-2* downto 0 do

3: R[¬ki]←R[0]+R[1], R[ki]←2R[ki].

4: end for

Output R[0]

In addition, the authors in [23] proposed secure (same security level as double-and-add-always method [11] and the Montgomery method [20]) and efficient ECSM method (See algorithm 4) by partitioning the bit string of the scalar in half (Key splitting into half) and extracting the common substring from the two parts based on propositional logic operations. The computations for common substring are thus saved, where the computational cost is approximately (k/2)A+kD.

ALGORITHM 4: ECSM BASED PROPOSITIONAL LOGIC OPERATIONS [23].

Inputs: $B_2=(d_2^{k/2}\dots\ d_2^{e}\ \dots.d_2^{1})_2$, $B_1=(d_1^{k/2}\dots\ d_1^{e}\ \dots.d_1^{1})_2$, P

Output: *dP*.

1: Q[0]=Q[1]=Q[2]=Q[3]=O;

2: For e=1 to k/2 do /* scan B1 and B2 from LSB to MSB */

3: $Q[2d_2^e + d_1^e] = Q[2d_2^e + d_1^e] + P$; /* ADD */

4: P = 2P; /* DBL */

5: Q[1] = Q[1] + Q[3]; Q[2] = Q[2] + Q[3];

6: For e=1 to k/2 do

7: Q[2] = 2Q[2]; /* DBL */

8: Q[1] = Q[2] + Q[1];

Return Q[1].

2) Using of unified formulae for PADD and PDBL through inserting extra field operations [19] [21] [22] [24] [25] [26] [28] [29] [30], by rewriting the PADD and PDBL formulas so that their implementation provides always the same shape and duration during the ECSM.

An arithmetic was proposed in [19] and refined in [26] together with the use of Edwards coordinates for ECC as proposed by Bernstein and Lange in 2007 [31] uses the same formula to compute PADD and PDBL. In addition, Hesse [21] and Jacobi form [22] elliptic curves achieve the in-distinguishability by using the same formula for both an PADD and PDBL. Moreover, a method proposed by Moller [25] performs ECSM with fixed pattern of PADD and PDBL, employing a randomized initialization stage to achieve resistance against PAAs. The same way, Liadet and Smart [24] have proposed to reduce information leakage by using a special point representation in some elliptic curves pertaining to a particular category, such that a single formula can be used for PADD and PDBL operations.

3) Rewriting sequence of operations as sequences of side-channel atomic blocks that are indistinguishable for SPAA [28]. The idea is to insert extra field operations and then divide each process into atomic blocks so that it can be expressed as the repetition of instruction blocks which appear equivalent (same power trace shape and duration) by SCA. The atomic pattern proposed in [28] is composed of the following field operations: a multiplication, two additions and a negation. This choice relies on the observation that during the execution of PADD and PDBL, no more than two additions and one negation are required between two multiplications.

To reduce the cost of atomic pattern of [28], Longa proposed in his PhD thesis [29] two atomic patterns in the context of Jacobian coordinates. In [29] Longa expresses mixed affine-Jacobian PADD formula as 6 atomic patterns and fast PDBL formula as 4 atomic patterns. It allows performing an efficient left-to-right ESCM using fast PDBL and mixed affine-Jacobian addition protected with atomic patterns. In addition, the authors in [30] address the problem of protecting ECSM implementations against PAA by proposing a new atomic pattern. They maximize the use of squarings to replace multiplications and minimize the use of field additions and negations since they induce a non-negligible penalty.

B. Countermeasures of DPAA on ECC

Same as in SPAA, there are different approaches and techniques [32] [33] [11] [34] [35] [12] used to resist DPAA attacks. In general, the traditional and straightforward approach is by randomizing the intermediate data, thereby rendering the calculation of the hypothetical leakage values rather impossible.

Coron [11] suggested three countermeasures to protect against DPAA attacks:

1) Blinding the scalar by adding a multiple of #E. For any random number r and k' = k+r#E, we have k'P = kP since (r#E)P = O.

2) Blinding the point P, such that kP becomes k(P +R). The known value S = kR is subtracted at the end of the computation. Blinding the point P makes RPA/ZPA more difficult.

In [14], the authors conclude that blinding the point P is vulnerable to DA since the point which blinds P is also doubled at each execution. Thereafter, in [32], the authors proposed a modification on the Coron's [11] point blinding technique to defend the DA. The modified technique in [32] is secure against DPAA.

Randomizing the homogeneous projective coordinates (X,Y,Z) with a random $\lambda \neq 0$ to (λX, λY, λZ). The random variable λ can be updated in every execution or after each PADD or PDBL, which will makes the collection of typical templates more difficult for an attacker.

Although randomizing projective coordinates is an effective countermeasure against DPAA, it fails to resist the RPA as zero is not effectively randomized. Furthermore, if the device outputs the point in projective coordinates, a final randomization must be performed; otherwise [36] shows how to learn parts of the secret value.

Similar to Coron [11], Ciet and Joye [35] also suggested several similar randomization methods.

3) Random scalar splitting: k = k1 + k2 or k = [k/r]r +(k mod r) for a random r.

Random scalar splitting can resist DPAA attacks since it has a random scalar for each execution. In addition, it helps preventing RPA/ZPA if it is used together with Blinding the point P technique [37] [12] [38].

4) Randomized EC isomorphism.

5) Randomized field isomorphism.

In the same context, Joye and Tymen [34] proposed to execute the ECSM on an isomorphic curve and to change the intermediate representations for each execution of a complete ECSM.

In [33], the authors presented a PAA resistant ECSM algorithm, based on building a sequence of bit-strings representing the scalar k, characterized by the fact that all bit-strings are different from zero; this property will ensure a uniform computation behavior for the algorithm, and thus will make it secure against PAA attacks.

Window-method [39]: It is an improved ECSM algorithm (sometimes referred to as m-ary method). For a window width of w, some multiples of the point P up to (2w-1)P are precomputed and stored and the scalar k is processed w bits at a time. K is recoded to the radix 2w.k can be recorded in a way so that the average density of the nonzero digits in the recoding is 1/(w+£), where $0 \leq £ \leq 2$ depends on the algorithm.

VI. Discussion

The main focus of this study is in highlighting on the PAAs on ECC as a major security threat in the context of WSNs. In a point of fact, none of the proposed countermeasures against PAAs on ECC, which are suggested in literatures, have considered the case of WSNs.

Given the resource constraints of WSN nodes, designing countermeasure methods against PAAs seems a non-trivial problem, and it should be a matter of tradeoff between the available resources on WSN node and performance. Thus, some critical concerns need to be taken into consideration while designing such countermeasures:

1) Not include any dummy operations (limited battery life time), and

2) Not limited to particular family of curves, and thus can be implemented in any NIST standardized curves.

3) Immunity against DPAAs may be carefully designed by combining several data randomization countermeasures and selectively change the ordering of these countermeasures with a time short enough to avoid a successful DPAA.

4) Template attacks are serious security threats on WSN nodes especially that the template building is simple and fast.

In addition, as shown in Figure 1, different attacks could be thwarted by one or more countermeasures. For example, Random Projective Coordinate prevents three powerful attacks (DPA, DA, and Template attack). However, it is worthy to emphasis on the fact that find a countermeasure against all know attacks is extremely costly, especially in the context of constrained devices like WSN.

VII. Conclusions and Future Work

Taking into consideration the resource constraints of WSN nodes, its deployment in open environments make these nodes highly exposed to PAAs. This paper represents a comprehensive study of major PAAs on ECC. The contributions of this paper are as follows: First, we present a review of the major PAAs and its countermeasures on ECC. Second, we make a graphical presentation for the relation between PAAs on ECC and its countermeasures. In addition, this paper discussed the critical concerns to be considered in designing PAAs on ECC particular for WSNs. Those, this paper should trigger the need for intensive researches to be conducted in the near future on the PAAs on ECC in WSNs nodes, especially that ECC is considered as the most feasible PKC for WSN security.

Although attacks like PAAs in WSN are normally carried out in situations where the adversary can control the target device [40], SPAA together with Template Attacks are still considered serious security threats, and thus a robust a cost-effect security solutions should be implementation to thwart these attacks.

Acknowledgment

The authors would like to acknowledge the support of Umm Al-Qura University, Makkah, Saudi Arabia and the support of LIMOS, CNRS, University Blaise Pascal, Clermont-Ferrand II, France.

References

[1] D. Estrin, R. Govindan, J. Heidemann and S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," in Mobile Computing and Networking (MobiCom'99), Seattle, WA USA, (1999).

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless Sensor Networks: a survey," Computer Networks, vol. 38, no. 4, pp. 393-422, 15 March 2002.

[3] D. Malan, M. Welsh and M. Smith, "A public-key infrastructure for key distribution in TinyOS based on elliptic curve cryptography," in Proc. of the 1st IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON '04), pp. 71–80, Santa Clara, Calif, USA, 2004.

[4] H. Houssain, M. Badra and T. F. Al Somani, "Hardware Implementations of Elliptic Curve Cryptography in Wireless Sensor Networks," in Proc. 6th International Conf. on Internet Technology and Secured Transactions (ICITST 2011), Abu Dhabi, UAE, pp. 1-6, Dec 2011.

[5] N. Gura, A. Patel, A. S. Wander, H. Eberle and S. Chang Shantz, "Comparing elliptic curve cryptography and RSA on 8-bit CPUs," in Cryptographic Hardware and Embedded Systems — CHES 2004, vol.

3156 of Lecture Notes in Computer Science, pp. 119–132, Springer Verlag, 2004.

[6] P. Kocher, "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems," in Advances in Cryptology, Proc. CRYPTO '96, N. Koblitz, ed., pp. 104-113, 1996.

[7] P. Kocher, J. Jaffe and B. Jun, "Differential power analysis," in Proc. Adv. Cryptology – CRYPTO'99, Santa Barbara, CA, 1999, vol. 1666, pp. 388–397.

[8] K. Pongaliur, Z. Abraham, A. X. Liu, L. Xiao and L. Kempel, "Securing Sensor Nodes Against Side Channel Attacks," in Proceedings of the 2008 11th IEEE High Assurance Systems Engineering Symposium (HASE '08), IEEE Computer Society, Washington, DC, USA, 353-361, 2008.

[9] D. Hankerson, A. Menezes and S. Vanstone, Guide to Elliptic Curve Cryptography, Springer-Verlag, 2004.

[10] H. Cohen, G. Frey, R. Avanzi, C. Doche, T. Lange, K. Nguyen and F. Vercauteren, "Handbook of Elliptic and Hyperelliptic Curve Cryptography. Discrete Mathematics and Its Applications," Vol. 34, Chapman and Hall, CRC, USA, 2005, ISBN: 9781584885184.

[11] J. S. Coron, "Resistance against differential power analysis for elliptic curve cryptosystems," in Cryptographic Hardware and Embedded Systems – CHES 1999, Worcester, MA: Springer, 1999, vol. 1717, pp. 292–302.

[12] L. Goubin, "A refined power-analysis attack on elliptic curve cryptosystems," in Proceedings of PKC 2003, LNCS 2567, pp. 199-211. Springer Berlin / Heidelberg, 2003.

[13] T. Akishita and T. Takagi, "Zero-value register attack on elliptic curve cryptosystem," IEICE Transactions, 88-A(1):132–139, 2005..

[14] P. Fouque and F. Valette, "The doubling attack– why upwards is better than downwards," in Proc. CHES'03, 2003, vol. 2779, pp. 269–280.

[15] S. M. Yen, L. C. Ko, S. J. Moon and J. C. Ha, "Relative doubling attack against montgomery ladder," in Proc. ICISC'05, 2006, vol. 3935, pp. 117–128.

[16] S. Chari, J. R. Rao and P. Rohatgi, "Template Attacks," in Cryptographic Hardware and Embedded Systems, CHES, ser. LNCS, vol. 2523, 2002, pp. 13–28..

[17] P. Fouque, D. R´eal, F. Valette and M. Drissi, "The Carry Leakage on the Randomized Exponent Countermeasure," in Cryptographic Hardware and Embedded Systems - CHES, ser. LNCS, vol. 5154. Springer, 2008, pp. 198–213..

[18] L. Batina, J. Hogenboom, N. Mentens, J. Moelans and J. Vliegen, "Side-channel evaluation of FPGA implementations of binary Edwards curves," in in International Conference on Electronics, Circuits and Systems 2010, pp. 1255-1258, Athens, Greece, Dec. 12-15, 2010.

[19] E. Brier and M. Joye, "Weierstraß elliptic curves and side-channel attacks," in David Naccache and Pascal Paillier (Eds.), Public Key Cryptography, vol. 2274 of Lecture Notes in Computer Science, pp. 335–345. Springer, Berlin / Heidelberg, 2002.

[20] P. Montgomery, "Speeding up the Pollard and elliptic curve methods of factorization," Mathematics of Computation, vol. 48, no. 177, pp. 243–264, 1987..

[21] M. Joye and J. Quisquater, "Hessian elliptic curves and side-channel attacks," Cryptographic Hardware and Embedded Systems CHES 2001, LNCS 2162, Springer-Verlag, pp.402–410, 2001.

[22] O. Billet and M. Joye, "The Jacobi model of an elliptic curve and side-channel analysis," Applied Algebra, Algebraic Algorithms and Error-Correcting Codes 2003, LNCS 2643, Springer- Verlag, pp.34–42, 2003..

[23] W. Keke, L. Huiun, Z. Dingju and Y. Fengqi, "Efficient Solution to Secure ECC Against Side-channel Attacks," 2011 20 (CJE-3): 471-475.

[24] L. P.Y. and S. NP, "Preventing SPA/DPA in ECC systems using the Jacobi form," in Workshop on Cryptographic Hardware and Embedded Systems (CHES 2001), Paris, France, 2001, vol. 2162, pp. 391–401.

[25] M. B., "Parallelizable elliptic curve point multiplication method with resistance against side-channel attacks," in Int. Conf. on Information Security (ISC 2002), Sao Paulo, Brazil, 2002, vol. 2433, pp. 402–413.

[26] É. Brier, I. Déchène and M. Joye, "Unified PADDition formulæ for elliptic curve cryptosystems," In Embedded Cryptographic Hardware: Methodologies & Architectures., Nova Science Publishers, 2004..

[27] T. F. Al-Somani and A. A. Amin, "High Performance Elliptic Curve Scalar Multiplication with Resistance against Power Analysis Attacks," Journal of Applied Sciences, Volume 8 (24), 2008, pp. 4587-4594..

[28] B. Chevallier-Mames, M. Ciet and M. Joye, "Low cost solutions for preventing simple side-channel analysis: Side channel atomicity," IEEE Trans. Computers, 53(6):760–768, 2004..

[29] P. Longa, Accelerating the Scalar Multiplication on Elliptic Curve Cryptosystems over Prime Fields., PhD thesis, School of Information Technology and Engineering, University of Ottawa, 2007.

[30] C. Giraud and V. Verneuil, "Atomicity Improvement for Elliptic Curve Scalar Multiplication," CARDIS 2010: 80441.

[31] D. Bernstein and T. Lange, "Faster Addition and Doubling on Elliptic Curves," Advances in Cryptology - ASIACRYPT, K. Kurosawa (ed.), vol. 4833 of LNCS, pp. 29-50, Springer, 2007.

[32] S. Ghosh, D. Mukhopadhyay and D. R. Chowdhury, "Petrel: Power and Timing Attack Resistant Elliptic Curve Scalar Multiplier Based on Programmable GF(p) Arithmetic Unit," IEEE Trans. on Circuits and Systems 58-I(8) , : 1798-1812 (2011).

[33] M. Hedabou, P. Pinel and L. Bénéteau, "A comb method to render ECC resistant against Side Channel Attacks," IACR Cryptology ePrint Archive 2004: 342, 2004 .

[34] M. Joye and C. Tymen, "Protections against differential analysis for elliptic curve cryptography," In: [cKKNP01] Cryptographic Hardware and Embedded Systems – CHES 2001, Lecture Notes in Computer Science, Vol. 2162, pp. 377.

[35] M. Ciet and M. Joye, "(Virtually) Free Randomization Techniques for Elliptic Curve Cryptography," in Information and Communications Security (ICICS2006), LNCS 2836, Springer, 2003, pp. 348–359..

[36] D. Naccache, N. P. Smart and J. Stern, "Projective Coordinates Leak," In: Advances in Cryptology - EuroCrypt 2004, Lecture Notes in Computer Science, Vol. 3027, pp. 257–267. Springer, Berlin / Heidelberg, 2004.

[37] T. Akishita and T. Takagi, "Zero-Value Point Attacks on Elliptic Curve Cryptosystem," vol. 2851, pp. 218–233, 2003..

[38] J. Ha, J. Park, S. Moon and S. Yen, "Provably Secure Countermeasure Resistant to Several Types of Power Attack for ECC," in Information Security Applications (WISA), vol. 4867. Springer, 2007, pp. 333–344.

[39] E. K. Reddy, "Elliptic Curve Cryptosystems and Side-channel Attacks," International Journal of Network Security, Vol.12, No.3, PP.151-158, May 2011.

[40] G. d. Meulenaer and F.-X. Standaert, "Stealthy Compromise of Wireless Sensor Nodes with Power Analysis Attacks," MOBILIGHT 2010: 229-242.

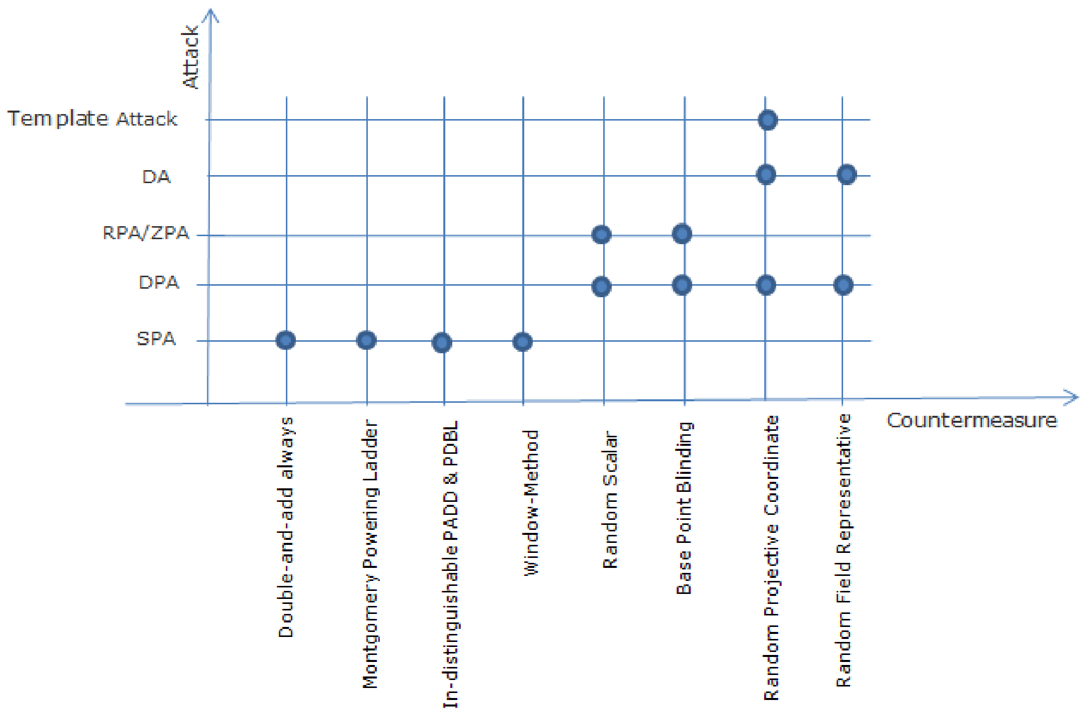


Figure 1 - PAAs vs. Countermeasures

A Survey on Resource Allocation Strategies in Cloud Computing

V.Vinothina, Sr.Lecturer,
Dept. of Computer Science,
Garden City College,
Bangalore, Karnataka, India.

Dr.R.Sridaran, Dean
Faculty of Computer Applications,
Marwadi Education Foundation's
Group of Institutions,
Rajkot, Gujarat, India.

Dr.PadmavathiGanapathi
Professor and Head,
Dept. of Computer Science,
Avinashilingam Institute of
Home Science and Higher
Education for Women,
Coimbatore, Tamil Nadu.
India.

***Abstract*— Cloud computing has become a new age technology that has got huge potentials in enterprises and markets. Clouds can make it possible to access applications and associated data from anywhere. Companies are able to rent resources from cloud for storage and other computational purposes so that their infrastructure cost can be reduced significantly. Further they can make use of company-wide access to applications, based on pay-as-you-go model. Hence there is no need for getting licenses for individual products. However one of the major pitfalls in cloud computing is related to optimizing the resources being allocated. Because of the uniqueness of the model, resource allocation is performed with the objective of minimizing the costs associated with it. The other challenges of resource allocation are meeting customer demands and application requirements. In this paper, various resource allocation strategies and their challenges are discussed in detail. It is believed that this paper would benefit both cloud users and researchers in overcoming the challenges faced.**

Keywords- Cloud Computing; Cloud Services; Resource Allocation; Infrastructure.

I. INTRODUCTION

Cloud computing emerges as a new computing paradigm which aims to provide reliable, customized and QoS (Quality of Service) guaranteed computing dynamic environments for end-users [22].Distributed processing, parallel processing and grid computing together emerged as cloud computing. The basic principle of cloud computing is that user data is not stored locally but is stored in the data center of internet. The companies which provide cloud computing service could manage and maintain the operation of these data centers. The users can access the stored data at any time by using Application Programming Interface (API) provided by cloud providers through any terminal equipment connected to the internet.

Not only are storage services provided but also hardware and software services are available to the general public and business markets. The services provided by service providers can be everything, from the infrastructure, platform or software resources. Each such service is respectively called Infrastructure as a Service (IaaS), Platform as a Service (PaaS) or Software as a Service (SaaS) [45].

There are numerous advantages of cloud computing, the most basic ones being lower costs, re-provisioning of resources and remote accessibility. Cloud computing lowers cost by avoiding the capital expenditure by the company in renting the physical infrastructure from a third party provider. Due to the flexible nature of cloud computing, we can quickly access more resources from cloud providers when we need to expand our business. The remote accessibility enables us to access the cloud services from anywhere at any time. To gain the maximum degree of the above mentioned benefits, the services offered in terms of resources should be allocated optimally to the applications running in the cloud. The following section discusses the significance of resource allocation.

A. Significance of Resource Allocation

In cloud computing, Resource Allocation (RA) is the process of assigning available resources to the needed cloud applications over the internet. Resource allocation starves services if the allocation is not managed precisely. Resource provisioning solves that problem by allowing the service providers to manage the resources for each individual module.

Resource Allocation Strategy (RAS) is all about integrating cloud provider activities for utilizing and allocating scarce resources within the limit of cloud environment so as to meet the needs of the cloud application. It requires the type and amount of resources needed by each application in order to complete a user job. The order and time of allocation of resources are also an input for an optimal RAS. An optimal RAS should avoid the following criteria as follows:

a) ***Resource contention*** *situation arises when two applications try to access the same resource at the same time.*

b) ***Scarcity of resources*** *arises when there are limited resources.*

c) ***Resource fragmentation*** *situation arises when the resources are isolated. [There will be enough resources but not able to allocate to the needed application.]*

d) ***Over-provisioning*** *of resources arises when the application gets surplus resources than the demanded one.*

*e) **Under-provisioning** of resources occurs when the application is assigned with fewer numbers of resources than the demand.*

Resource users' (cloud users) estimates of resource demands to complete a job before the estimated time may lead to an over-provisioning of resources. Resource providers' allocation of resources may lead to an under-provisioning of resources. To overcome the above mentioned discrepancies, inputs needed from both cloud providers and users for a RAS as shown in table I. From the cloud user's angle, the application requirement and Service Level Agreement (SLA) are major inputs to RAS. The offerings, resource status and available resources are the inputs required from the other side to manage and allocate resources to host applications [25] by RAS. The outcome of any optimal RAS must satisfy the parameters such as throughput, latency and response time. Even though cloud provides reliable resources, it also poses a crucial problem in allocating and managing resources dynamically across the applications.

TABLE I. INPUT PARAMETERS

Parameter	Provider	Customer
Provider Offerings	√	-
Resource Status	√	-
Available Resources	√	-
Application Requirements	-	√
Agreed Contract Between Customer and provider	√	√

From the perspective of a cloud provider, predicting the dynamic nature of users, user demands, and application demands are impractical. For the cloud users, the job should be completed on time with minimal cost. Hence due to limited resources, resource heterogeneity, locality restrictions, environmental necessities and dynamic nature of resource demand, we need an efficient resource allocation system that suits cloud environments.

Cloud resources consist of physical and virtual resources. The physical resources are shared across multiple compute requests through virtualization and provisioning [23]. The request for virtualized resources is described through a set of parameters detailing the processing, memory and disk needs which is depicted in Fig.1. Provisioning satisfies the request by mapping virtualized resources to physical ones. The hardware and software resources are allocated to the cloud applications on-demand basis. For scalable computing, Virtual Machines are rented.

The complexity of finding an optimum resource allocation is exponential in huge systems like big clusters, data centers or Grids. Since resource demand and supply can be dynamic and uncertain, various strategies for resource allocation are proposed. This paper puts forth various resource allocation strategies deployed in cloud environments.

The rest of the paper is organized as follows: In section II, a few work related to this topic is presented. Various resource allocation strategies and their impacts in cloud environments are discussed in section III. In section IV, some of the advantages and limitations of resource allocation in cloud are addressed. Finally the conclusion of the paper is given as section V.

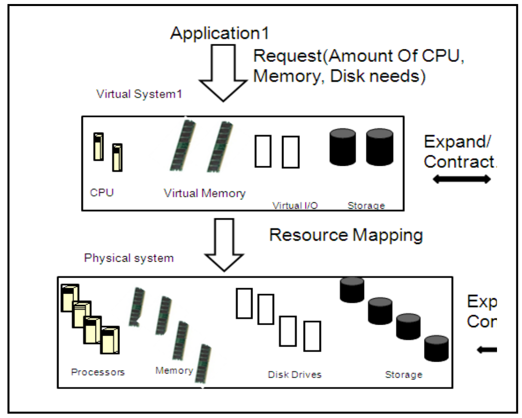


Figure1. Mapping of virtual to physical resources

II. RELATED WORK

Very little literature is available on this survey paper in cloud computing paradigm. Shikharesh et al. in paper [30] describes the resource allocation challenges in clouds from the fundamental point of resource management. The paper has not addressed any specific resource allocation strategy.

Patricia et al. [25], investigates the uncertainties that increase difficulty in scheduling and matchmaking by considering some examples of recent research.

It is evident that the paper which analyzes various resource allocation strategies is not available so far. The proposed literature focuses on resource allocation strategies and its impacts on cloud users and cloud providers. It is believed that this survey would greatly benefit the cloud users and researchers.

III. RESOURCE ALLOCATION STRATEGIES (RAS) AT A GLANCE

The input parameters to RAS and the way of resource allocation vary based on the services, infrastructure and the nature of applications which demand resources. The schematic diagram in Fig.2 depicts the classification of Resource Allocation Strategies (RAS) proposed in cloud paradigm. The following section discusses the RAS employed in cloud.

A. Execution Time

Different kinds of resource allocation mechanisms are proposed in cloud. In the work by Jiani at.al [15], actual task execution time and preemptable scheduling is considered for resource allocation. It overcomes the problem of resource contention and increases resource utilization by using different modes of renting computing capacities. But estimating the execution time for a job is a hard task for a user and errors are made very often [30]. But the VM model considered in [15] is heterogeneous and proposed for IaaS.

Using the above-mentioned strategy, a resource allocation strategy for distributed environment is proposed by Jose et al. [16]. Proposed matchmaking (assign a resource to a job) strategy in [16] is based on Any-Schedulability criteria for assigning jobs to opaque resources in heterogeneous environment. This work does not use detailed knowledge of the scheduling policies used at resources and subjected to AR's (Advance Reservation).

B. Policy

Since centralized user and resource management lacks in scalable management of users, resources and organization-level security policy [6], Dongwan et al. [6] has 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.

One of the resource allocation challenges of resource fragmentation in multi-cluster environment is controlled by the work given by Kuo-Chan et al. [20], which used the most-fit processor policy for resource allocation. The most-fit policy allocates a job to the cluster, which produces a leftover processor distribution, leading to the most number of immediate subsequent job allocations.

It requires a complex searching process, involving simulated allocation activities, to determine the target cluster. The clusters are assumed to be homogeneous and geographically distributed. The number of processors in each cluster is binary compatible. Job migration is required when load sharing activities occur.

Experimental results shows that the most-fit policy has higher time complexities but the time overheads are negligible compared to the system long time operation. This policy is practical to use in a real system.

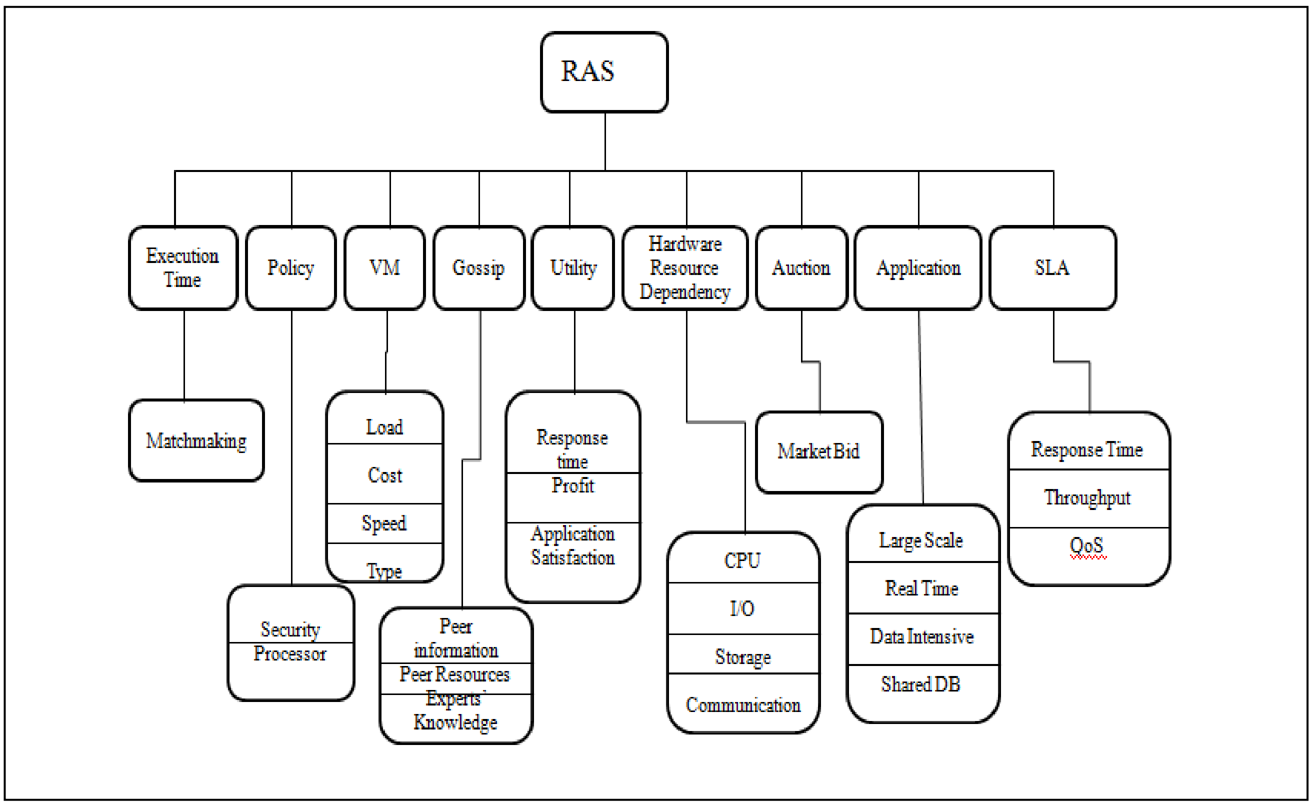


Figure2. Resource Allocation Strategies in Cloud Computing

C. Virtual Machine (VM)

A system which can automatically scale its infrastructure resources is designed in [24]. The system composed of a virtual network of virtual machines capable of live migration across multi- domain physical infrastructure. By using 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. But the above work considers only the non-preemptable scheduling policy.

Several researchers have developed efficient resource allocations for real time tasks on multiprocessor system. But the studies, scheduled tasks on fixed number of processors. Hence it is lacks in scalability feature of cloud computing [18]. Recent studies on allocating cloud VMs for real time tasks [36], [31], [17] focus on different aspects like infrastructures to enable real-time tasks on VMs and selection of VMs for power management in the data center. But the work by Karthik et al. [18], have allocated the resources based on the speed and cost of different VMs in IaaS. It differs from other related works, by allowing the user to select VMs and reduces cost for the user.

Users can set up and boot the required resources and they have to pay only for the required resources [3]. It is implemented by enabling the users to dynamically add and/or delete one or more instances of the resources on the basis of VM load and the conditions specified by the user. The above mentioned RAS on IaaS differs from RAS on SaaS in cloud because SaaS delivers only the application to the cloud user over the internet.

Zhen Kong et al. have discussed mechanism design to allocate virtualized resources among selfish VMs in a non-cooperative cloud environment in [44]. By non-cooperative means, VMs care essentially about their own benefits without any consideration for others. They have utilized stochastic approximation approach to model and analyze QoS performance under various virtual resource allocations. The proposed stochastic resource allocation and management approaches enforced the VMs to report their types truthfully and the virtual resources can be allocated efficiently. The proposed method is very complex and it is not implemented in a practical virtualization cloud system with real workload.

D. Gossip

Cloud environment differs in terms of clusters, servers, nodes, their locality reference and capacity. The problem of resource management for a large-scale cloud environment (ranging to above 100,000 servers) is addressed in [28] and general Gossip protocol is proposed for fair allocation of CPU resources to clients.

A gossip-based protocol for resource allocation in large-scale cloud environments is proposed in [9]. It performs a key function within distributed middleware architecture for large clouds. In the thesis, the system is modeled as a dynamic set of nodes that represents the machines of cloud environment. Each node has a specific CPU capacity and memory capacity. The protocol implements a distributed scheme that allocates cloud resources to a set of applications that have time-dependent memory demands and it dynamically maximizes a global cloud utility function. The simulation results show that the protocol produces optimal allocation when memory demand is smaller than the available memory in the cloud and the quality of the allocation does not change with the number of applications and the number of machines. But this work requires additional functionalities to make resource allocation scheme is robust to machine failure which spans several clusters and datacenters.

But in the work by Paul et al. [26] cloud resources are being allocated by obtaining resources from remote nodes when there is a change in user demand and has addressed three different policies to avoid over-provisioning and under-provisioning of resources. Recent research on sky computing focuses on bridging multiple cloud providers using the resources as a single entity which would allow elastic site for leveraging resources from multiple cloud providers [19]. Related work is proposed in [24] but it is considered only for preemptable tasks. Yang et al. [43] have proposed a profile based approach for scaling the applications automatically by capturing the experts' knowledge of scaling application servers as a profile. This approach greatly improves the system performance and resource utilization. Utility based RAS is also proposed for PaaS in [12].

In paper [8], Gossip based co-operative VM management with VM allocation and cost management is introduced. By this method, the organizations can cooperate to share the available resources to reduce the cost. Here the cloud environments of public and private clouds are considered. They have formulated an optimization model to obtain the optimal virtual machine allocation. Network game approach is adopted for the cooperative formation of organizations so that none of the organizations wants to deviate. This system does not consider the dynamic co-operative formation of organizations. Related work is discussed in [2] that use desktop cloud for better usage of computing resources due to the increase in average system utilization. The implication for a desktop cloud is that individual resource reallocation decisions using desktop consolidation and decision based on aggregate behavior of the system.

E. Utility Function

There are many proposals that 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 etc.

There are few works [4], [38] that dynamically allocate CPU resources to meet QoS objectives by first allocating requests to high priority applications. The authors of the papers do not try to maximize the objectives. Hence the authors' Dorian et al. proposed Utility (profit) based resource allocation for VMs which use live VM migration (one physical machine to other) as a resource allocation mechanism [7]. This controls the cost-performance trade-off by changing VM utilities or node costs. This work mainly focuses on scaling CPU resources in IaaS. A few works [1],[32] that use live migration as a resource provisioning mechanism but all of them use policy based heuristic algorithm to live migrate VM which is difficult in the presence of conflicting goals.

For multitier cloud computing systems (heterogeneous servers), resource allocation based on response time as a measure of utility function is proposed by considering CPU, memory and communication resources in [10]. HadiGoudarzi et al. characterized the servers based on their capacity of processing powers, memory usage and communication bandwidth.

For each tier, requests of the application are distributed among some of the available servers. Each available server is assigned to exactly one of these applications tiers i.e. server can only serve the requests on that specified server. Each client request is dispatched to the server using queuing theory and this system meets the requirement of SLA such as response time and utility function based on its response time. It follows the heuristics called force-directed resource management for resource consolidation. But this system is acceptable only as long as the client behaviors remain stationary.

But the work proposed in [13] considers the utility function as a measure of application satisfaction for specific resource allocation (CPU, RAM). The system of data center with single cluster is considered in [13] that support heterogeneous applications and workloads including both enterprise online applications and CPU-intensive applications. The utility goal is computed by Local Decision Module (LDM) by taking current work load of the system. The LDMs interact with Global Decision Module (GDM) and that is the decision making entity within the autonomic control loop. This system relies on a two-tier architecture and resource arbitration process that can be controlled through each application's weight and other factors.

F. Hardware Resource Dependency

In paper [35], to improve the hardware utilization, Multiple Job Optimization (MJO) scheduler is proposed. Jobs could be classified by hardware-resource dependency such as CPU-bound, Network I/O-bound, Disk I/O bound and memory bound. MJO scheduler can detect the type of jobs and parallel jobs of different categories. Based on the categories, resources are allocated. This system focuses only on CPU and I/O resource.

Eucalyptus, Open Nebula and Nimbus are typical open source frame works for resource virtualization management [39]. The common feature of these frameworks is to allocate virtual resources based on the available physical resources, expecting to form a virtualization resource pool decoupled with physical infrastructure. Because of the complexity of virtualization technology, all these frameworks cannot support all the application modes. The system called Vega LingCloud proposed in paper [39] supports both virtual and physical resources leasing from a single point to support heterogeneous application modes on shared infrastructure.

Cloud infrastructure refers to the physical and organizational structure needed for the operation of cloud. Many recent researches address the resource allocation strategies for different cloud environment. Xiaoying Wang et al. have discussed adaptive resource co-allocation approach based on CPU consumption amount in [25]. The stepwise resource co-allocation is done in three phases. The first phase determines the co-allocation scheme by considering the CPU consumption amount for each physical machine (PM). The second phase determines whether to put applications on PM or not by using simulated annealing algorithm which tries to perturb the configuration solution by randomly changing one element. During phase 3, the exact CPU share that each VM occupies is determined and it is optimized by the gradient climbing approach. This system mainly focuses on CPU and memory resources for co-allocation and does not considered the dynamic nature of resource request.

HadiGoudarzi et al. in paper [11] proposed a RAS by categorizing the cluster in the system based on the number and type of computing, data storage and communication resources that they control. All of these resources are allocated within each server. The disk resource is allocated based on the constant need of the clients and other kind of resources in the servers and clusters are allocated using Generalized Processor Sharing (GPS). This system performs distributed decision making to reduce the decision time by parallelizing the solution and used greedy algorithm to find the best initial solution. The solution could be improved by changing resource allocation. But this system cannot handle large changes in the parameters which are used for finding the solution.

G. Auction

Cloud resource allocation by auction mechanism is addressed by Wei-Yu Lin et al. in [37]. The proposed mechanism is based on sealed-bid auction. The cloud service provider collects all the users' bids and determines the price. The resource is distributed to the first k^{th} highest bidders under the price of the $(k+1)^{th}$ highest bid. This system simplifies the cloud service provider decision rule and the clear cut allocation rule by reducing the resource problem into ordering problem. But this mechanism does not ensure profit maximization due to its truth telling property under constraints.

The aim of resource allocation strategy is to maximize the profits of both the customer agent and the resource agent in a large datacenter by balancing the demand and supply in the market. It is achieved by using market based resource allocation strategy in which equilibrium theory is introduced (RSA-M) [41]. RSA-M determines the number of fractions used by one VM and can be adjusted dynamically according to the varied resource requirement of the workloads. One type of resource is delegated to publish the resource's price by resource agent and the resource delegated by the customer agent participates in the market system to obtain the maximum benefit for the consumer. Market Economy Mechanism is responsible for balancing the resource supply and demand in the market system.

H. Application

Resource Allocation strategies are proposed based on the nature of the applications in [33] [34]. In the work by Tram et al. [33], Virtual infrastructure allocation strategies are designed for workflow based applications where resources are allocated based on the workflow representation of the application. For work flow based applications, the application logic can be interpreted and exploited to produce an execution schedule estimate. This helps the user to estimate the exact amount of resources that will be consumed for each run of the application. Four strategies such as Naive, FIFO, Optimized and services group optimization are designed to allocate resources and schedule computing tasks.

Real time application which collects and analyzes real time data from external service or applications has a deadline for completing the task. This kind of application has a light weight web interface and resource intensive back end [34]. To enable dynamic allocation of cloud resources for back-end mashups, a prototype system is implemented and evaluated for both static and adaptive allocation with a test bed cloud to allocate resources to the application. The system also accommodates new requests despite a-priori undefined resource utilization requirements. This prototype works by monitoring the CPU usage of each virtual machine and adaptively invoking additional virtual machines as required by the system.

David Irwin et al. [5] have suggested the integration of high bandwidth radar sensor networks with computational and storage resources in the cloud to design end-to-end data intensive cloud systems. Their work provides a platform that supports a research on broad range of heterogeneous resources and overcomes the challenges of coordinated provisioning between sensors networks, network providers and cloud computing providers. Inclusion of nontraditional resources like Steerable sensors and cameras and stitching mechanisms to bind the resources are the requirement of this project. Resource allocation strategy plays significant role in this project.

Database replicas allocation strategy is designed in [27]. In that work, the resource allocation module divides the resource (CPU, Memory and DB replicas) allocation problem in two levels. The first level optimally splits the resources among the clients whereas the database replicas are expandable (dynamic) in the second level, based on the learned predictive model. It achieves optimal resource allocation in a dynamic and intelligent fashion.

I. SLA

In cloud, the works related to the SaaS providers considering SLA are still in their infancy. Therefore in order to achieve the SaaS providers' objective, various RAS specific to SaaS in cloud has been proposed. With the emergence of SaaS, applications have started moving away from pc based to web delivered-hosted services. Most of the RAS for SaaS focused towards customer benefits. Popovivi et al. [14] have mainly considered QoS parameters on the resource provider's side such as price and offered load.

Moreover Lee at.al [42] have addressed the problem of profit driven service request scheduling in cloud computing by considering the objectives of both parties such as service providers and consumers. But the author Linlin Wu et al. [21] have contributed to RAS by focusing on SLA 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 are also proposed in [21].

Managing the computing resources for SaaS processes is challenging for SaaS providers [29]. Therefore a framework for resource management for SaaS providers to efficiently control the service levels of their users is contributed by Richard et al. [29]. It can also scale SaaS provider application under various dynamic user arrivals/departures. All the above mentioned mainly focus on SaaS providers' benefits and significantly reduce resource waste and SLO violations.

IV. ADVANTAGES AND LIMITATIONS

There are many benefits in resource allocation while using cloud computing irrespective of size of the organization and business markets. But there are some limitations as well, since it is an evolving technology. Let's have a comparative look at the advantages and limitations of resource allocation in cloud.

A. Advantages:

1) The biggest benefit of resource allocation is that user neither has to install software nor hardware to access the applications, to develop the application and to host the application over the internet.

2) The next major benefit is that there is no limitation of place and medium. We can reach our applications and data anywhere in the world, on any system.

3) The user does not need to expend on hardware and software systems.

4) Cloud providers can share their resources over the internet during resource scarcity.

B. Limitations

1) Since users rent resources from remote servers for their purpose, they don't have control over their resources.

2) Migration problem occurs, when the users wants to switch to some other provider for the better storage of their data. It's not easy to transfer huge data from one provider to the other.

3) In public cloud, the clients' data can be susceptible to hacking or phishing attacks. Since the servers on cloud are interconnected, it is easy for malware to spread.

4) Peripheral devices like printers or scanners might not work with cloud. Many of them require software to be installed locally. Networked peripherals have lesser problems.

5) More and deeper knowledge is required for allocating and managing resources in cloud, since all knowledge about the working of the cloud mainly depends upon the cloud service provider.

In Appendix A, various resource allocations strategies and their impact are listed.

V. CONCLUSION

Cloud computing technology is increasingly being used in enterprises and business markets. In cloud paradigm, an effective resource allocation strategy is required for achieving user satisfaction and maximizing the profit for cloud service providers. This paper summarizes the classification of RAS and its impacts in cloud system. Some of the strategies discussed above mainly focus on CPU, memory resources but are lacking in some factors. Hence this survey paper will hopefully motivate future researchers to come up with smarter and secured optimal resource allocation algorithms and framework to strengthen the cloud computing paradigm.

REFERENCES

[1] A.Singh ,M.Korupolu and D.Mohapatra. Server-storage virtualization:Integration and Load balancing in data centers. In Proc.2008 ACM/IEEE conference on supercomputing (SC'08) pages 1-12, IEEE Press 2008.

[2] AndrzejKochut et al. : Desktop Workload Study with Implications for Desktop Cloud Resource Optimization,978-1-4244-6534-7/10 2010 IEEE.

[3] Atsuo Inomata, TaikiMorikawa, Minoru Ikebe, Sk.Md. MizanurRahman: Proposal and Evaluation of Dynamin Resource Allocation Method Based on the Load Of VMs on IaaS(IEEE,2010),978-1-4244-8704-2/11.
[4] D. Gmach, J.RoliaandL.cherkasova, Satisfying service level objectives in a self-managing resource pool. In Proc. Third IEEE international conference on self-adaptive and self organizing system.(SASO'09) pages 243-253.IEEE Press 2009 .
[5] David Irwin, PrashantShenoy, Emmanuel Cecchet and Michael Zink:Resource Management in Data-Intensive Clouds: Opportunities and Challenges .This work is supported in part by NSF under grant number CNS-0834243.
[6] Dongwan Shin and HakanAkkan :Domain- based virtualized resource management in cloud computing.
[7] Dorian Minarolli and Bernd Freisleben: Uitlity –based Resource Allocations for virtual machines in cloud computing(IEEE,2011),pp.410-417.
[8] DusitNiyato, Zhu Kun and Ping Wang : Cooperative Virtual Machine Management for Multi-Organization Cloud Computing Environment.
[9] FetahiWuhib and Rolf Stadler : Distributed monitoring and resource management for Large cloud environments(IEEE,2011),pp.970-975.
[10] HadiGoudaezi and MassoudPedram : Multidimensional SLA-based Resource Allocation for Multi-tier Cloud Computing Systems IEEE 4th International conference on Cloud computing 2011,pp.324-331.
[11] HadiGoudarzi and MassoudPedram: Maximizing Profit in Cloud Computing System Via Resource Allocation: IEEE 31st International Conference on Distributed Computing Systems Workshops 2011: pp,1-6.
[12] Hien et al. ,'Automatic virtual resource management for service hosting platforms, cloud'09,pp 1-8.
[13] Hien Nguyen et al.: SLA-aware Virtual Resource Management for Cloud Infrastructures: IEEE Ninth International Conference on Computer and Information Technology 2009, pp.357-362.
[14] I.Popovici et al,"Profitable services in an uncertain world". In proceedings of the conference on supercomputing CSC2005.
[15] Jiyani et al.: Adaptive resource allocation for preemptable jobs in cloud systems (IEEE, 2010), pp.31-36.
[16] Jose Orlando Melendez &shikhareshMajumdar: Matchmaking with Limited knowledge of Resources on Clouds and Grids.
[17] K.H Kim et al. Power-aware provisioning of cloud resources for real time services. In international workshop on Middlleware for grids and clouds and e-science, pages 1-6, 2009.
[18] Karthik Kumar et al.: Resource Allocation for real time tasks using cloud computing (IEEE, 2011), pp.
[19] Keahey et al., "sky Computing",Intenet computing, IEEE,vol 13,no.5,pp43-51,sept-Oct2009.
[20] Kuo-Chan Huang &Kuan-Po Lai: Processor Allocation policies for Reducing Resource fragmentation in Multi cluster Grid and Cloud Environments (IEEE, 2010), pp.971-976.
[21] Linlin Wu, Saurabh Kumar Garg and Raj kumarBuyya: SLA –based Resource Allocation for SaaS Provides in Cloud Computing Environments (IEEE, 2011), pp.195-204 .
[22] Lizhewang,JieTao,Kunze M.,Castellanos,A.C,Kramer,D.,Karl,w, "High Performance Computing and Communications",IEEE International Conference HPCC,2008,pp.825-830.
[23] M.SuhailRehman,MajdF.Sakr : Initial Findings for provisioning Variation in Cloud Computing(IEEE,2010),pp.473-479 .
[24] P.Ruth,J.Rhee, D.Xu, R.Kennell and S.Goasguen, "Autonomic Adaptation of virtual computational environments in a multi-domain infrastructure", IEEE International conference on Autonomic Computing, 2006,pp.5-14.
[25] Patricia Takako Endo et al. :Resource allocation for distributed cloud :Concept and Research challenges(IEEE,2011),pp.42-46 .
[26] Paul Marshall, Kate Keahey& Tim Freeman: Elastic Site(IEEE,2010),pp.43-52.
[27] PenchengXiong,Yun Chi, Shenghuo Zhu, Hyun Jin Moon, CaltonPu&HakanHacigumus: Intelligent Management Of Virtualized Resources for Database Systems in Cloud Environment(IEEE,2011),pp.87-98.
[28] RerngvitYanggratoke, FetahiWuhib and Rolf Stadler: Gossip-based resource allocation for green computing in Large Clouds: 7th International conference on network and service management, Paris, France, 24-28 October, 2011.
[29] Richard T.B.Ma,Dah Ming Chiu and John C.S.Lui, Vishal Misra and Dan Rubenstein:On Resource Management for Cloud users :a Generalized Kelly Mechanism Approach.
[30] ShikhareshMajumdar: Resource Management on cloud : Handling uncertainties in Parameters and Policies (CSI communicatons,2011,edn)pp.16-19.
[31] Shuo Liu Gang Quan ShangpingRen On –Line scheduling of real time services for cloud computing. In world congress on services, pages 459-464, 2010.
[32] T.Wood et al. Black Box and gray box strategies for virtual machine migration. In Proc 4th USENIX Symposium on Networked Systems Design and Implementation (NSDI 07),pages 229-242.
[33] Tram Truong Huu& John Montagnat: Virtual Resource Allocations distribution on a cloud infrastructure (IEEE, 2010), pp.612-617.
[34] WaheedIqbal, Matthew N.Dailey, Imran Ali and Paul Janecek& David Carrera: Adaptive Resource Allocation for Back-end Mashup Applications on a heterogeneous private cloud.
[35] Weisong Hu et al. : Multiple Job Optimization in MapReduce for Heterogeneous Workloads : IEEE Sixth International Conference on Semantics, Knowledge and Grids 2010,pp.135-140.
[36] Wei-Tek Tsai Qihong Shao Xin Sun Elston, J. Service-oriented cloud computing. In world congress on services, pages 473-478, 2010.
[37] Wei-Yu Lin et al. : Dynamic Auction Mechanism for Cloud Resource Allocation: 2010 IEEE/ACM 10th International Conference on Cluster, Cloud and Grid Computing, pp.591-592.
[38] X.Zhu et al. Integrated capacity and workload management for the next generation data center. In proc.5th international conference on Automatic computing(ICAC'08),pages 172-181,IEEE Press 2008.
[39] XiaoyiLu,Jian Lin, Li Zha and ZhiweiXu: Vega Ling Cloud: A Resource Single Leasing Point System to Support Heterogenous Application Modes on Shared Infrastructure(IEEE,2011),pp.99-106.
[40] Xiaoying Wang et al. : Design and Implementation Of Adaptive Resource Co-allocation Approaches for Cloud Service Environments : IEEE 3rd International Conference on Advanced Computer Theory and Engineering 2010,V2,pp,484-488.
[41] Xindong YOU, Xianghua XU, Jian Wan, Dongjin YU:RAS-M :Resource Allocation Strategy based on Market Mechanism in Cloud Computing(IEEE,2009),pp.256-263.
[42] Y.C Lee et.al,"Project driven service request scheduling in clouds". In proceedings of the international symposium on cluster & Grid Computing.(CC Grid 2010),Melbourne, Australia.
[43] Yang wt.al A profile based approach to Just in time scalability for cloud applications, IEEE international conference on cloud computing ,2009,pp 9-16.
[44] Zhen Kong et.al : Mechanism Design for Stochastic Virtual Resource Allocation in Non-Cooperative Cloud Systems: 2011 IEEE 4th International Conference on Cloud Computing :pp,614-621.
[45] ZhixiongChen,JongP.Yoon,"International Conference on P2P, Parallel,Grid,Cloud and Internet Computing",2010 IEEE:pp 250-257

AUTHORS PROFILE

1. **Ms. Vinothina V., M.Sc. (Computer Science), M.B.A. (Systems), M.Phil.,**is presently a Senior Lecturer, Department of Computer Science, **Garden City College**, Bangalore. She has done M.Sc. Computer Science from Madurai Kamarai University, Tamil Nadu. She has completed M.B.A. with specialization in Systems from Alagappa University and M.Phil. inImage Processing from Bharathidasan University, Tamil Nadu. She has presented 4 papers in the National conference, attended many workshops and Webinars. Her areas of interest are networking and Web Technologies. She is currently pursuing Ph.D in Bharathiyar University. She may be reached at rvvino@yahoo.com.

2. Dr. Sridaran has done his post graduation in Computer Applications and Management. He has been awarded the Ph.D in Computer Applications in 2010. Having started his career as an Entrepreneur, he has offered his consultancy services to various service sectors. He has also designed and delivered various training programs in the areas of IT & Management. He has published 14 research papers in leading Journals and Conferences and presently guiding four research scholars. He has got 17 years of academic experience and served in leading educational institutions at different capacities. He is currently the Dean, Faculty of Computer Applications, Marwadi Education Foundation's Group of Institutions, Rajkot, Gujarat. He may be reached at sridaran.rajagopal@gmail.com

3. Dr. PadmavathiGanapathi is the Professor and Head of the Department ofComputer Science, Avinashilingam University. Her area of research includes communication network, security and real time and multimediasystems. she has 170 publications at her credit.she is a life member of many professional organizations like, CSI, ISTE, ICSA, AACE, WSEAS,UWA etc. she has executed funded projects worth 2 crore from variousfunding agencies like AICTE, UGC, DRDO-NRB, DRDO-ARMREB.sheiscurrently guiding research scholars at M.Phil and Ph.D levels. She may be reached at ganapathi.padmavathi@gmail.com.

APPENDIX A

S.No	Resource Allocation Strategy	Impacts
1	Based on the estimated execution time of job .(Advanced Reservation, Best effort and immediate mode)	Estimation may not be accurate. If job could not finish its execution in estimated tine, it will affect the execution of other jobs.
2	Matchmaking strategy based on Any-Schedulability criteria.	Strategy mainly depends upon the user estimated job execution time of a job.
3	Based on role based security policy.	Follows decentralized resource allocation.
4	Most Fit Processor Policy.	Requires complex searching process and practical to use in real system.
5	Based on cost and speed of VM.	Allows the user to select VM.
6	Based on the load conditions specified by the user.	Instances of resources can be added or removed.
7	Based on gossip protocol (resources allocated by getting information for other local nodes)	It used decentralized algorithm to compute resource allocation and this prototype is not acceptable for heterogeneous cloud environment.
8	Utility function as a measure of profit based on live VM migration.	Focused on scaling CPU resources in IaaS.
9	Based on the utility function as a measure of price.	Allocate resources only in the lowest level of cloud computing and considered only CPU resource.
10.	Utility function as a measure of response time.	Lacks in handling dynamic client requests.
11	Based on utility function as a measure of application satisfaction.	Relies on two-tier architecture.
12	Based on the CPU usage of VM, active user requests are served. Adaptively new VM spawns, when the CPU usage reaches some critical point.(VR)	There is a limitation in the number of concurrent user monitor and the prototype is not capable of scaling down as the number of active user decreases.
13	Based on hardware resource dependency.	Considered only CPU and I/O resource.
14	Auction mechanism.	Not ensure profit maximization
15	Based on online resource demand predication.	Prediction may not be accurate and leads to over provisioning or under provisioning.
16	Based on workflow representation of the application.	The application logic can be interpreted and exploited to produce an execution schedule estimate. Again estimation may not be accurate.
17	Based on the machine learning technique to precisely make decisions on resources.	This prototype reduces the total SLA cost and allocate resources considering the both the request rates and also the weights.
18	Simulated annealing algorithm.	Lacks in handling dynamic resource request.
19	Based on constant needs of client and GPS.	Solution can be improved by changing the resource allocation and lacks in handling the large changes in parameters.
20	Stochastic approximation approach.	Very complex in nature.
21	Network game theory approach.	Lack in dynamic cooperative organization formation

Cluster-based Communication Protocol for Load-Balancing in Wireless Sensor Networks

Mohammed A. Merzoug and Abdallah Boukerram
Department of Computer Science
University Ferhat Abbas of Setif
Setif 19000, Algeria

***Abstract*—One of the main problems in wireless sensor networks is information collection. How the sensor nodes can send efficiently the sensed information to the sink since their number is very large and their resources are limited? Clustering provides a logical view, much more effective than the physical view, which extends the lifetime of the network and achieves the scalability objective. Although the assumption that all sensor nodes can communicate directly with each other or with the sink is not always valid because of the limited transmission range of the sensor nodes and the remoteness of the sink. In this paper, we introduce a cluster-based communication protocol that uses a multi-hop communication mode between the cluster-heads. The protocol aims to reduce and evenly distribute the energy consumption among the sensor nodes, and cover a large area of interest. The simulations show that the proposed protocol is efficient in terms of energy consumption, maximization of the network lifetime, data delivery to the sink and scalability.**

Keywords- wireless sensor networks; clustering; load-balancing.

I. Introduction

A wireless sensor network is composed of a large number of sensor nodes and one or more sink nodes (base stations). The sensor nodes are deployed inside the area of interest to collect useful information from the surrounding environment and report it to a base station located generally at the extremity of the area of interest. For example, the sensor nodes can monitor and report certain events like the movement of objects. The role of the base station is to gather the information sent by the sensor nodes and send it back to the user (control node), and eventually send queries to the sensor nodes. Generally, the base station is much more powerful in terms of resources than the sensor nodes.

A sensor node is a small device that includes four basic components: a data acquisition unit, a processing unit, a wireless communication unit and an energy unit. The sensor node is equipped with low-power batteries suitable for its small size, which limits the ability of the sensor node in terms of processing, storage and transmission. In most cases, the sensor nodes are disposable and should last until their energy runs out. Thus, the energy is the most precious resource in a wireless sensor network. The conservation of energy and the maintenance of the wireless sensor network as long as possible are important challenges.

When the sensor nodes communicate directly with the base station, the sensor nodes located farther away from the base station will have a higher energy load due to the long range communication. When the sensor nodes use a multi-hop communication to reach the base station, the sensor nodes located close to the base station will have a higher energy load because they relay the packets of other nodes [3].

The clustering-based communication mode is considered as the most suitable communication mode for the wireless sensor networks. Clustering consists in selecting a set of cluster-heads from the set of sensor nodes and then regrouping the remaining sensor nodes around the cluster-heads. The cluster-members send the data to the cluster-head that sends it back to the base station. Clustering gives better results, it reduces and balances the energy consumption and improves the lifetime and scalability of the wireless sensor network. Clustering is often used with a data aggregation technique. Thus, the number of sent messages and transmission ranges can be reduced.

When the sensor nodes make the decision to become cluster-heads based on a limited view, the formed clusters will not be effective. In the proposed protocol, the sensor nodes base on their local parameters and the parameters of their neighbors in order to have a global view and ensure a good distribution of the cluster-heads.

Most of the clustering-based protocols use a single-hop communication to send data from the cluster-heads to the base station. In fact, they assume that all sensor nodes can communicate directly with each other or with the base station [3] [4] [12]. This becomes impossible when the size of the area of interest increases. The proposed protocol uses a multi-hop communication between the cluster-heads to conserve energy and cover a large area of interest. To reduce the amount of information to be sent to the base station, we integrated data aggregation. Moreover, the rotation of cluster-heads and the use of the low-power sleep mode by the sensor nodes that do not participate in routing allow to balance the load and reduce energy consumption significantly.

The remainder of this paper is organized as follows. Section 2 provides an overview of related work. In Section 3, we define the system model. In Section 4, we present our protocol. Section 5 provides details of the proposed protocol. Section 6 presents the simulation setup and results. Section 7 concludes the paper and outlines future work.

II. Related Work

LEACH [3] is one of the first cluster-based communication protocols proposed for the wireless sensor networks. LEACH is executed in two phases: the setup phase and the steady-state phase. In the first phase, the cluster-heads are selected and the clusters are formed. In the second phase, the data transmission to the base station will take place. No negotiation is needed to determine the cluster-heads. Each sensor node makes the decision to become cluster-head or not independently of the other nodes in the network based on a predetermined value of the optimal percentage of cluster-heads for the network and the number of rounds in which the node has not been a cluster-head. A sensor node *n* chooses a random value between 0 and 1. If this value is less than a threshold *T(n)*, the node becomes a cluster-head for the current round. The threshold *T(n)* is calculated as follows:

$$T(n) = \begin{cases} \dfrac{P}{1 - P * \left(r \bmod \dfrac{1}{P} \right)} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where P is the desired percentage of cluster-heads ($P = 0.05$), r is the current round, and G is the set of sensor nodes that have not been cluster-heads in the last $1/P$ rounds.

The cluster-heads inform all the sensor nodes in the network of their election by broadcasting an advertisement message. Each sensor node chooses the closest cluster-head that requires the minimum communication energy, based on the signal strength of the advertisement message. After this step, the sensor nodes send join-request messages to the chosen cluster-heads.

Once the sensor nodes are organized into clusters, each cluster-head creates a TDMA schedule and broadcasts it to its members. Each node transmits its data to the cluster-head during its time slot specified in the TDMA table and turns off its radio waiting for the next transmission slot, which minimizes the consumed energy. On the other hand, the cluster-heads keep their radio on in order to receive data from cluster-members. Once a cluster-head has received the data from all members, it performs data aggregation and then transmits the aggregated data directly to the base station. Since the base station is generally far away, this transmission is a high-energy transmission. LEACH uses random rotation of the role of cluster-head after each round to evenly distribute the energy load among the sensor nodes.

LEACH-C [4] is a variant of LEACH that uses a centralized clustering algorithm. In the setup phase of each round, the base station receives the coordinates of the sensor nodes and their remaining energy. Based on this information and using the optimization method of simulated annealing, the base station selects the cluster-heads and forms the clusters. The steady-state phase is similar to that of LEACH.

In LEACH-F [4], the base station uses the same clustering algorithm used in LEACH-C. The clusters remain fixed and the sensor nodes in the same cluster rotate the role of cluster-head. Each sensor node can become a cluster-head. This method does not allow adding new nodes to the network.

III. System Model

There are several possible models for wireless sensor networks. Our protocol is designed for the wireless sensor networks where the sensor nodes are homogeneous (same initial energy, same processing and data storage capacities) and energy-constrained. Once deployed randomly in the area of interest (Fig. 1), the sensor nodes are left unattended. It is impossible to recharge or change the batteries in case of energy

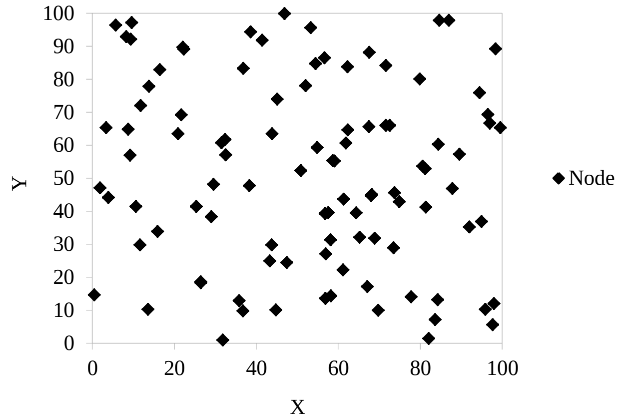


Figure 1. 100 sensor nodes randomly deployed in an area of 100 × 100 meters.

depletion. Therefore, the energy conservation and the maximization of the lifetime of each sensor node are required. The sensor nodes ignore their geographic location (no location system such as GPS). Our protocol is designed for surveillance applications, also called time-driven applications. The sensor nodes collect data from the surrounding environment at a fixed rate and send it to a single base station characterized by its unlimited capacity (energy, processing and storage). The base station is located at the extremity of the area of interest, away from the sensor nodes. All nodes including the base station are fixed (very common configuration for the applications of wireless sensor networks). Each node has multiple transmission power levels and controls the transmission power. Direct communication between all the nodes including the base station is not possible. The sensor nodes play two different roles: sensor node and cluster-head. The sensor node monitors its environment and transmits the collected data to its cluster-head. The cluster-head aggregates the collected data and forwards it to the base station. The cluster-head also participates in routing the data of other clusters.

A. Radio Energy Consumption Model

The energy consumed by the radio represents the largest portion of the energy consumed by a sensor node and the receive/listen mode represents often the main source of consumption. Therefore, the model of the energy consumed by the radio in different modes (transmit, receive/listen and sleep) affects the evaluation results and change the advantages of different protocols. Different radio energy consumption models were proposed [3] [4] [15]. Heinzelman et al. [3] proposed a simple model where the radio consumes $E_{elec}=50\ nJ/bit$ to run the transmitter or receiver circuitry and $\epsilon_{amp}=100\ pJ/bit/m^2$ for the transmit amplifier. Thus, to transmit a *k-bit* message a distance *d*, the radio consumes:

$$E_{TX}(k, d) = E_{elec} * k + \epsilon_{amp} * k * d^2 \quad (2)$$

To receive this message, the radio consumes:

$$E_{RX}(k) = E_{elec} * k \quad (3)$$

Castalia Simulator [1] provides realistic wireless channel and radio models. The energy is consumed according to the state of the radio. For example, the radio consumes the same energy in the receive mode and the listen mode in other words if it is receiving packets or just listening. Castalia provides three radio models: CC2420, CC1000 and BAN Radio. CC2420 and CC1000 define the real radios of the same name. BAN Radio describes the narrowband radio proposed in the IEEE 802.15 Task Group 6 documents.

In our work, we opted for the CC2420 radio. This radio is widely used in wireless sensor networks. CC2420 radio consumes *62mW* in the receive/listen mode. In the sleep mode, the radio consumes *1.4mW* (a single sleep level is defined). For the transmit mode, there are eight possible transmission power levels (*dBm*) with different power consumption (*mW*) (Table 2). In Table 2, the first column lists the different transmission levels and the second column shows how much energy the radio is spending when transmitting in a power level. For example, if the radio transmits with *0dBm*, it consumes *57.42mW*. These numbers that Castalia Simulator is using for the CC2420 radio are straight out of its datasheet. From Table 1 and 2, we notice that the receive/listen mode and the transmit mode consume a lot of energy and it is better to transmit than receive or remain in the listen mode. Therefore, the communication protocols must minimize not only the transmission distances but also the transmission and reception operations by keeping the sensor nodes in the sleep mode as often as possible.

Table 3 and 4 specify respectively the delays (in *msec*) and the power consumed (in *mW*) to switch between the three main modes of the CC2420 radio. For the transition delays, the radio takes *194µsec* to change from the sleep mode to either receive/listen or transmit modes, *10µsec* to change between transmit and receive modes and *50µsec* to enter the sleep mode. The no transitions are represented by "-" instead of a number. For the consumed power, the transitions to the sleep mode consume *1.4mW* while any other transitions take *62mW*.

TABLE I. POWER CONSUMED BY THE CC2420 RADIO IN THE RECEIVE/LISTEN AND SLEEP MODES

Mode	Power consumed (mW)
Receive/Listen	62
Sleep	1.4

TABLE II. POWER CONSUMED BY THE CC2420 RADIO IN THE TRANSMIT MODE

Transmission power level (dBm)	Power consumed (mW)
0	57.42
-1	55.18
-3	50.69
-5	46.20
-7	42.24
-10	36.30
-15	32.67
-25	29.04

TABLE III. MODE SWITCHING DELAYS
(DELAY TO SWITCH FROM COLUMN MODE TO ROW MODE, IN MSEC)

	Receive/Listen	Transmit	Sleep
Receive/Listen	-	0.01	0.194
Transmit	0.01	-	0.194
Sleep	0.05	0.05	-

TABLE IV. MODE SWITCHING POWERS
(POWER TO SWITCH FROM COLUMN MODE TO ROW MODE, IN MW)

	Receive/Listen	Transmit	Sleep
Receive/Listen	-	62	62
Transmit	62	-	62
Sleep	1.4	1.4	-

B. *Transmission power adjustment*

The sensor nodes have several transmission power levels. In order to minimize the amount of energy required for transmission, the sensor nodes can adjust their transmission power before communicating with each other or with the base station. For example in the case of a clustering-based communication protocol, the sensor nodes adjust their transmission power before communicating with the cluster-head. The cluster-heads can also adjust their transmission power to communicate with each other or communicate directly with the base station. This adjustment allows achieving a significant energy gain.

In order to determine the appropriate transmission power for the communication from a sensor node *A* to another sensor node *B*, the node *A* must first receive a packet from the node *B*. The sensor node *A* bases on the RSSI (Received Signal Strength Indicator) of the received packet, calculated by its radio. The node *A* needs other information as the sensitivity of the radio of the node *B* and the transmission power used to transmit the packet. Suppose that the sensor nodes are equipped with a CC2420 radio and the sensor node *B* has transmitted the packet with a transmission power of *0dBm*. The sensitivity of the CC2420 radio is *-95dBm*. If the RSSI of the packet is *-75dBm*, the sensor node *A* knows that there is a 20dB "link budget". In other words it can reduce the transmission power of the node *B* by *20dB* or less and remains above the radio sensitivity. The adjusted transmission power is calculated as follows [1]:

$$P_A > P_{TX} - (RSSI - S) \quad (4)$$

where

P_A: adjusted transmission power.
P_{TX}: transmission power of the transmitter node.

S: sensitivity of the radio of the transmitter node (in *dBm*).
RSSI: RSSI of the received packet.

This is true if the links are completely symmetric. In case of asymmetry, for example, both directions of a path between two sensor nodes vary by *4dB*, the possible variation has to be accounted in the formula:

$$P_A > P_{TX} + V - (RSSI - S) \quad (5)$$

where

V: possible variation.

IV. COMMUNICATION PROTOCOL

An optimal communication protocol for wireless sensor networks collects the data at the base station from the distributed sensor nodes while consuming their energy slowly and evenly. In other words, the sensor nodes die at about the same time avoiding the rapid and unbalanced depletion of energy. Cluster-based communication protocols are very effective. These protocols conserve the energy using different techniques: reducing the distance sensor nodes need to transmit their data, turning off the radio of the sensor nodes, aggregating data, rotating the role of cluster-head, etc.

The proposed protocol aims to reduce and balance the energy consumption. In fact, this protocol is designed with two main objectives: maximization of the network lifetime and scalability. The protocol operation is divided into rounds. Each round consists of three main phases: initialization phase, clustering phase and data transmission phase (Fig. 2). During the initialization phase, the sensor nodes exchange information and organize themselves into layers according to their distance from the base station (hop count). Our protocol is based on the clustering algorithm of Qin and Zimmermann [7]. In fact, the performance of any cluster-based communication protocol depends mainly on the clustering algorithm. The position of cluster-heads and the number of nodes in each cluster are very important factors. The clustering phase starts with the selection of cluster-heads and clusters formation. Each sensor node executes a distributed clustering algorithm in order to become cluster-head or join a cluster based on the information exchanged during the initialization phase. To reduce the intra-cluster collisions, each cluster-head creates a TDMA schedule. During the transmission phase, data collected by the cluster-members are sent to the cluster-head according to this schedule. The cluster-head aggregates the data and sends it to the upper layer. The data is then sent from cluster-head to cluster-head until reach the base station.

Clustering or re-clustering is not an end in itself. Despite the advantages of the rotation of cluster-heads, the frequent change is undesirable because it imposes a considerable overhead. In order to remedy this problem and minimize the consumed energy during initialization and clustering, the transmission phase is long compared to the other two phases.

V. PROTOCOL DETAILS

A. Initialization Phase

This phase assigns to each sensor node a level according to its distance with the base station and a weight of eligibility to become cluster-head according to its remaining energy. The base station starts the process by sending a setup message with a zero level (Fig. 3). Initially, only the sensor nodes that are within the reach of the base station can receive this message.

Then, these sensor nodes of a 1-hop explore the next-hop sensor nodes. The process continues until there will be no next-hop nodes. When a sensor node receives multiple setup messages, it chooses the one with the smallest level. Fig. 3 shows the structure of the network after the initialization phase.

While building the topology of the network and assigning levels, each sensor node collects the remaining energy of its neighbors regardless of the level of the neighbor. Based on its remaining energy and the remaining energy of its neighbors, each sensor node determines its suitability and the suitability of its neighbors to become cluster-head and assigns the votes. The vote given by a sensor node S_i to itself is:

$$V(S_i) = \frac{E_i}{\sum_{S_k \in N_i} E_k} \tag{6}$$

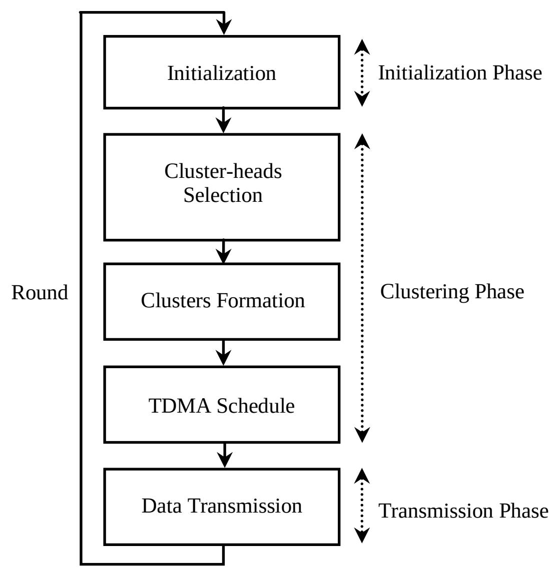


Figure 2. Proposed protocol operation

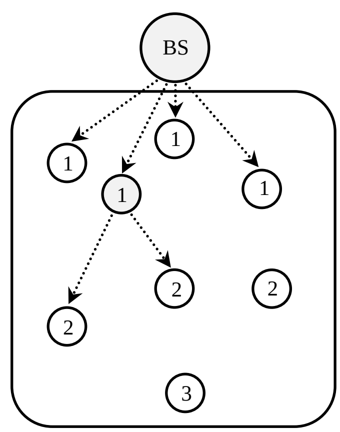


Figure 3. Initialization phase. Determination of levels and exchange of information. The base station (BS) broadcasts a message to the 1-hop nodes. A node explores the next-hop.

where E_i is the remaining energy of the sensor node S_i and N_i is the set of neighbors of the sensor node S_i. The vote given by a sensor node S_i to a neighboring sensor node S_j is:

$$V(S_i, S_j) = \frac{E_j}{\sum_{S_k \in N_i} E_k} \tag{7}$$

Each sensor node broadcasts the votes attributed to its neighbors. After receiving the votes, each sensor node determines its weight as follows:

$$W(S_i) = V(S_i) + \sum_{S_k \in N_i} V(S_k, S_i) \quad (8)$$

Each sensor node broadcasts its weight to its neighbors. During the clustering phase, the sensor nodes use these weights to elect themselves cluster-heads or remain regular sensor nodes. The advantage of this technique of vote lies in the fact that the importance of a sensor node must be determined by the node itself and all its neighbors instead of using the local properties alone. More votes a sensor node accumulates, the more important it is in the entire network.

B. Clustering Phase

In this phase, the sensor nodes elected to be cluster-heads form the clusters by broadcasting advertisement messages and receiving join-request messages. The clustering phase is divided into three sub-phases: cluster-heads selection, clusters formation and TDMA schedule.

Taking into account the energy during the selection of cluster-heads is of major importance. Since the cluster-heads consume more energy than the regular sensor nodes, they should be selected from the set of nodes with high remaining energy. In the cluster-heads selection sub-phase, each sensor node decides to become cluster-head or not based on the weights of eligibility calculated and broadcasted in the initialization phase. The sensor node compares its weight with the weights of its neighbors. If its weight is the highest, it becomes cluster-head for the current round and broadcasts an advertisement message. The cluster-heads use the CSMA/CA MAC protocol and the same transmission power to transmit their advertisement. During this phase, the non-cluster-heads nodes must keep their radio on (listen/receive mode) to hear the advertisement messages. A sensor node can receive the advertisement of several cluster-heads. In this case, the sensor node must decide to which cluster it will belong in this round. The decision must balance the load of the cluster-heads because the main objective of our protocol is to maximize the lifetime of the network. The sensor nodes choose to join the cluster whose cluster-head has the minimum degree (number of neighbors). To do this, each cluster-head calculates its fitness value and broadcasts it with the advertisement message. The fitness of a sensor node S_i is calculated as follows:

$$\text{fitness}(S_i) = \frac{1}{D(S_i)} \quad (9)$$

where $D(S_i)$ is the degree of the sensor node S_i (number of neighbors).

The sensor nodes receive the advertisement messages and choose the cluster-head with the maximum fitness. After choosing a cluster-head, the sensor node adjusts its transmission power according to the RSSI of the advertisement message and the transmission power used by the cluster-head. This adjustment saves the energy considerably. Then, the sensor node sends a join-request message to the cluster-head using the CSMA/CA MAC protocol. The cluster-heads must keep their radio in the listen/receive mode during this phase in order to hear the join-request messages. It is possible that some sensor nodes receive advertisement messages from cluster-heads that belong to different layers. We did not impose restrictions preventing nodes to join the cluster-heads of other layers. The cluster-heads can also receive the advertisement of other cluster-heads. In this case, the cluster-head selects the cluster-head that belongs to an upper layer (lower level) and has the maximum fitness as its parent-cluster-head. The cluster-heads do not send join-request messages to other cluster-heads.

After receiving the join-request messages, each cluster-head creates a TDMA schedule and broadcasts it to the cluster-members. The TDMA schedule assigns to each node in the cluster a time slot during which it can transmit. Thus, the sensor nodes can enter the sleep mode and conserve energy. The use of the TDMA approach also ensures the absence of collisions in the cluster.

After the clustering phase, it is possible that some sensor nodes do not belong to any cluster. These nodes turn off their radio and wait the next round in order to conserve energy. During the transmission phase, the sensor nodes transmit the collected data to the cluster-heads and the cluster-heads transmit the aggregated data to the base station.

C. Data Transmission Phase

This phase is divided into frames and each frame is divided into slots (Fig. 4). The number of slots in a frame corresponds to the number of nodes in the cluster including the cluster-head. Each sensor node sends its data to the cluster-head once per frame during its allocated transmission slot. During the slots of the other cluster-members, the node turns off its radio to minimize energy dissipation. Instead of transmitting all the received data to the base station, the cluster-head checks and combines the content of the data. This aggregation reduces the traffic load, since much less data needs to be transmitted to the base station.

The cluster-heads use the CSMA/CA MAC protocol to transmit the aggregated data to the base station. Each cluster-head sends the data to its parent-cluster-head. If a cluster-head does not have a parent-cluster-head then it broadcasts the data message. Any other cluster-head who receives this message sends it to its parent-cluster-head or broadcasts it back if it does not have a parent.

The cluster-heads consume more energy than the regular nodes because they have more responsibilities. In order to evenly distribute the energy consumption among all the nodes in the network, after each round our protocol repeats the process from the initialization phase. The re-clustering prevents the assignment of the role of cluster-head to some sensor nodes and therefore the rapid depletion of their energy.

VI. Simulation and Discussion

In this section, we analyze the performance of our protocol, simulated using OMNeT++/Castalia [1][2], against LEACH [3][4]. The comparison is made in terms of the following metrics:

- Energy consumption: this metric shows the energy consumed during the network operation.
- Network lifetime: this metric shows the number of sensor nodes that die during the network operation.

- Received data messages: this metric shows the number of data messages successfully delivered to the base station.

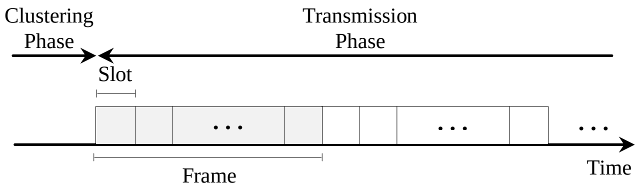


Figure 4. Data transmission phase

We also test the scalability of the proposed solution.

A. *Simulation Setup*

The simulations were executed in a square area of 100 × 100 meters, with 100 sensor nodes randomly deployed (Fig. 1). The base station is located at position 50, 150. Each sensor node has an initial energy of 100 J. The radio used is CC2420. The size of the data packets is fixed at 2000 bytes and the size of the control packets at 25 bytes. Table 5 summarizes the parameters used in our simulations.

B. *Simulation Results*

Since the cluster-heads consume more energy, our protocol assigns this role to sensor nodes with high remaining energy. After each round, these cluster-heads will be replaced with other sensor nodes with more remaining energy. LEACH does not take into account the remaining energy of sensor nodes during the selection of cluster-heads. The choice is made randomly and all sensor nodes in the network play that role periodically. Moreover, the cluster-heads communicate directly with the base station using the highest transmission power, which requires a high energy. Our protocol uses short distance transmissions to reduce energy consumption. Fig. 5 and Fig. 6 show that LEACH consumes more energy.

To analyze the network lifetime, we have chosen the following three definitions: the time until the first sensor node dies, the time until half of the sensor nodes die, and the time until the last sensor node dies. Since more than one sensor node is required to perform the clustering, the latter definition represents the lifetime of the network when 80% of the sensor nodes die.

In the case where the cluster-heads are not in the center of the clusters, some sensor nodes will consume more energy than others. LEACH does not guarantee a good distribution of cluster-heads because the selection is done randomly without considering the network parameters.

TABLE V. SIMULATION PARAMETERS

Parameter	Value
Number of sensor nodes	100
Dimensions of the area of interest	100 × 100 meters
Deployment of the sensor nodes	random
Location of the base station	x = 50, y = 150
Initial energy of the sensor nodes	100 J
Radio	CC2420
Data rate	250 Kbps
Data packets size	2000 bytes
Control packets size	25 bytes

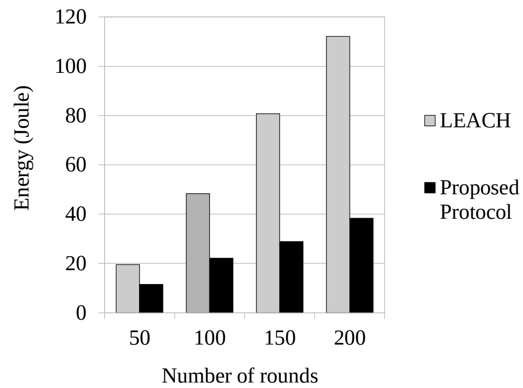


Figure 5. Energy consumption

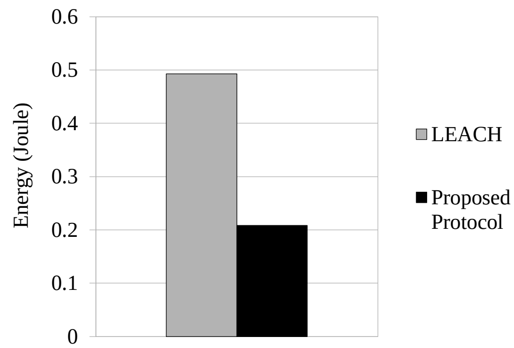


Figure 6. Average energy consumed per round

To form the clusters, the sensor nodes listen to the advertisement messages and choose the cluster-head with the strongest signal (largest RSSI). In the case of ties, a random cluster-head is chosen. Using this technique with the random selection of cluster-heads, a sensor node with low remaining energy can be cluster-head and have a large number of members. This reduces the lifetime of that node. When the network becomes dense, this situation is more likely to occur. Fig. 7 shows that the first node dies earlier in LEACH.

The main objective of our approach is to extend the network lifetime. This objective is achieved through: short distance transmissions, good distribution of cluster-heads, and load balancing among the cluster-heads. The sensor nodes communicate only with their close neighbors to save energy. In order to have a global view and ensure a good distribution of cluster-heads, the sensor nodes base on their parameters and the parameters of their neighbors to elect themselves cluster-heads. Knowing that the energy consumed by a cluster-head depends essentially on the number of cluster-members, it is very important to balance the load of the cluster--heads. The use of the fitness value of the cluster-heads tries to balance the clusters size. A sensor node tries to join the cluster-head with the minimum load. Fig. 8 and Fig. 9 show that the sensor nodes remain alive for longer time with the proposed protocol.

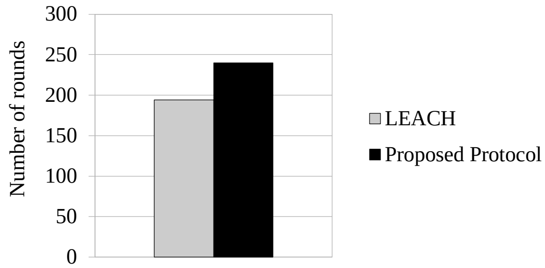


Figure 7. Network lifetime when the first sensor node dies

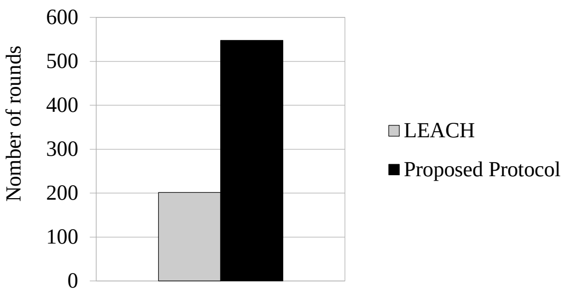


Figure 8. Network lifetime when 50% of sensor nodes die

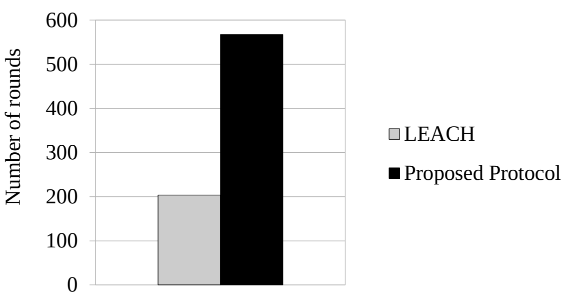


Figure 9. Network lifetime when 80% of sensor nodes die

Scalability is one of the most important features. In a large network, the communication protocol should not be affected and the control traffic should be negligible compared to the data traffic. Because a large number of sensor nodes generates a lot of transmissions. Sending and receiving messages consume energy. Fig. 10 shows that our protocol outperforms LEACH and generates a longer lifetime when the density of nodes increases. In both protocols, the sensor nodes send data to the cluster-heads over a single-hop. To reduce the amount of data to be transmitted to the base station, the cluster-heads aggregate the data received from the sensor nodes. In LEACH, the data is sent directly to the base station while in our protocol, it is sent from cluster-head to cluster-head until the base station. The transmissions over short distances conserve energy and avoid collisions. Fig. 11 shows the number of data messages received by the base station during the network operation. Since the rate of dead sensor nodes in the proposed protocol is less, more data messages will be sent to the base station.

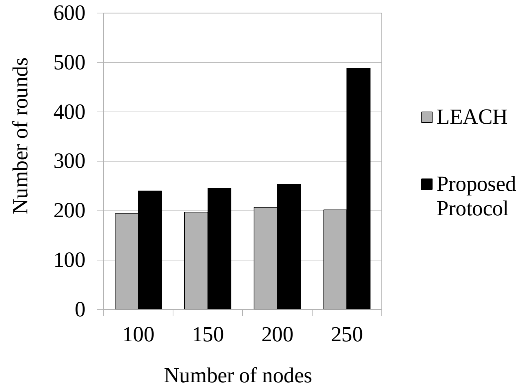


Figure 10. Network lifetime as a function of number of nodes when the first node dies

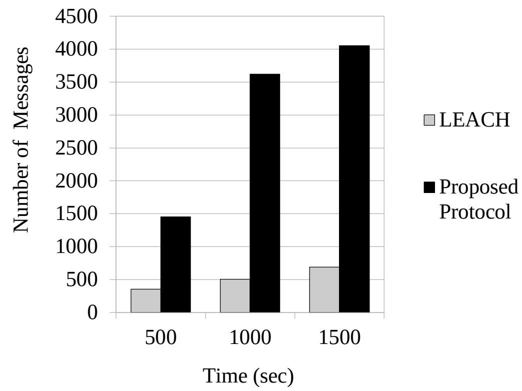


Figure 11. Number of data messages received by the base station

VII. Conclusion and Future Work

In this paper, we introduced a clustering-based communication protocol, which aims to evenly distribute the energy consumption among the sensor nodes and extend the lifetime of wireless sensor networks. Simulation results show that the proposed protocol is better than LEACH in terms of energy consumption, maximization of the network lifetime, data delivery to the base station and scalability. Our algorithm improves the formation of clusters by incorporating the information of the neighboring nodes.

However, like many other distributed clustering algorithms, it does not guarantee an optimal formation of clusters. Moreover, the problem of load-balancing remains. When the sensor nodes use the single-hop communication, some sensor nodes will have a high energy load due to the long-distance communication.

When the sensor nodes use the multi-hop communication, some sensor nodes will have a high energy load because they relay packets of other nodes. Our future work is to improve the protocol and find a much more suitable communication mode for wireless sensor networks.

REFERENCES

[1] Castalia : Simulator for wireless sensor networks. http://castalia.npc.nicta.com.au/

[2] OMNeT++ : Simulation environment. http://www.omnetpp.org/

[3] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In Proccedings of the 33rd Hawaii International Conference on System Sciences, 2000.

[4] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan. An application-specific protocol architecture for wireless microsensor networks. IEEE Transactions on Wireless Communications, 1(4) :660–670, October 2002.

[5] S. Lindsey and C. S. Raghavendra. PEGASIS : power-efficient gathering in sensor information systems. In IEEE Aerospace Conference, volume 3, pages 1125–1130, 2002.

[6] O. Younis and S. Fahmy. HEED : A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Transactions on Mobile Computing, 3(4) :366–379, October-December 2004.

[7] M. Qin and R. Zimmermann. VCA : An energy-efficient voting-based clustering algorithm for sensor networks. Journal of Universal Computer Science, 12(1), 2007.

[8] N. Israr and I. Awan. Multihop clustering algorithm for load balancing in wireless sensor networks. Internat. J. Simul. Systems Sci. Technol., 8(1) :13–25, 2007.

[9] N. Kim, J. Heo, H. S. Kim, and W. H. Kwon. Reconfiguration of clusterheads for load balancing in wireless sensor network. Computer Communications, 31(1) :153–159, January 2008.

[10] C. P. Low, C. Fang, and Y. H. Ang J. M. Ng. Efficient load-balanced clustering algorithms for wireless sensor networks. Computer Communications, 31 :750–759, March 2008.

[11] S. Soni and N. Chand. Energy efficient multi-level clustering to prolong the lifetime of wireless sensor networks. Journal of Computing, 2 :158–165, May 2010.

[12] F. Bajaber and I. Awan. Adaptive decentralized re-clustering protocol for wireless sensor networks. Journal of Computer and System Sciences, 77(2) :282–292, March 2011.

[13] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks : a survey. Computer Networks, 38 :393–422, 2002.

[14] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella. Energy conservation in wireless sensor networks : a survey. Ad Hoc Networks, 7(3) :537–568, May 2009.

[15] M. Holland, T. Wang, B. Tavli, A. Seyedi, and W. Heinzelman. Optimizing physical layer parameters for wireless sensor networks. ACM Transactions on Sensor Networks, 7(4), Nov. 2011.

[16] I. Demirkol, C. Ersoy, and F. Alagöz. MAC protocols for wireless sensor networks : a survey. IEEE Communications, 44(4) :115–121, April 2006.

[17] K. Akkaya and M. Younis. A survey on routing protocols for wireless sensor networks. Ad Hoc Networks, 3(3) :325–349, 2005.

[18] A. A. Abbasi and M. Younis. A survey on clustering algorithms for wireless sensor networks. Computer Communications, 30 :2826–2841, 2007.

Semantic Searching and Ranking of Documents using Hybrid Learning System and WordNet

Pooja Arora
Research Scholar, Banasthali Vidyapith
Assistant Professor, MCA Department AKGEC
Ghaziabad, India

Prof. Om Vikas
Senior Member, IEEE
India

***Abstract*— Semantic searching seeks to improve search accuracy of the search engine by understanding searcher's intent and the contextual meaning of the terms present in the query to retrieve more relevant results. To find out the semantic similarity between the query terms, WordNet is used as the underlying reference database. Various approaches of Learning to Rank are compared. A new hybrid learning system is introduced which combines learning using Neural Network and Support Vector Machine. As the size of the training set highly affects the performance of the Neural Network, we have used Support Vector Machine to reduce the size of the data set by extracting support vectors that are critical for the learning. The data set containing support vectors is then used for learning a ranking function using Neural Network. The proposed system is compared with RankNet. The experimental results demonstrated very promising performance improvements. For experiments, we have used English-Hindi parallel corpus, Gyannidhi from CDAC. F-measure and Average Interpolated Precision are used for evaluation.**

Keywords- Learning to Rank; English-Hindi Parallel Corpus; Hybrid Learning; Support Vector Machine (SVM); Neural Network (NN); Semantic Searching; WordNet; Search Engine.

I. INTRODUCTION

Information retrieval is the method of searching documents, and information within documents and metadata about documents in databases and on the World Wide Web. The main idea is to locate terms that the user specify in his query. Many documents contain the desired semantic information, even though they do not contain the user specified query terms. For retrieving those documents semantic searching is required. So, semantic similarity of the terms must also be considered while calculating the score of the particular document. For that, firstly the user query is expanded by replacing all the query terms by their synonyms and then searching is performed according to the changed query and score of all the retrieved relevant documents is calculated using the ranking function. We have used WordNet for query expansion. Finally, all the retrieved documents are ranked according to their relevance score.

First we discuss learning to rank and compare its various approaches. Then we have proposed our approach of learning a ranking function using Support Vector Machine (SVM) and Neural Network (NN). [3]

II. LEARNING TO RANK

The problem of ranking has recently gained much attention in information retrieval. The task of Learning to Rank has emerged as an active and growing area of research both in information retrieval and machine learning. The goal is to design and apply methods to automatically learn a function from training data, such that the function can sort objects (e.g. documents) according to their degrees of relevance, preference, or importance as defined in a specific application. In information retrieval, Learning to Rank is used to generate an effective ranking function that is used to rank documents according to their relevance score.

The learning process, formalized as follows, consists of two steps: **training and test**. Given a query collection Q = { $q_1,\ldots\ldots,q_m$ } and a document collection , D = { $d_1, \ldots\ldots, d_n$ } , the training corpus is created as a set of query-document pairs , each (q_i , d_j) ϵ Q × D, upon which a relevance judgment indicating the relationship between q_i and d_j is assigned by a labeler . The relevance judgment can be a score e.q. sim(q_i , d_j) specifying the degree of relevance between q_i and d_j .[18]

The inputs to the learning algorithm comprise training instances, their feature vectors and the corresponding relevance judgments. The output is a ranking function R_f, where $R_f(q_i,d_j)$ is supposed to give the true relevance score for q_i and d_j. During the training process, the learning algorithm attempts to learn a ranking function such that a performance measure (e.g. Mean Average Interpolated Precision (MAIP), F-measure, etc.) with respect to the output relevance judgment can be optimized.

In the test phase, the learned ranking function is applied to determine the relevance between each document d_i in D and a new query q_{m+1}. The learning is greatly affected by various factors such as performance measure used for evaluation, the form of training instance, etc.

Various Approaches of Learning to Rank

Learning to Rank approach is classified into Pointwise approach, Pairwise approach and Listwise approach depending upon the type of instance used for learning. Their comparison is shown in Table 1.

Pointwise Approach: The Pointwise approach solves the problem of ranking by means of regression or classification on single documents. It takes features of a single document with respect to query as input : $\varphi(q_k,d_i^k)$. In the training phase, learner learns to classify each instance of the document as relevant or irrelevant. In test phase, model assigns a unique score to each instance according to its relevance to queries. After operation, it produces ordered categories as output: $\{R_f(q_k,d_1^k)$, $R_f(q_k,d_2^k)$ . . ., $R_f(q_k,d_n^k)\}$ where n is the number of documents retrieved and $R_f(q_k,d_n^k)$ is the score of the n^{th} document. Various algorithms were proposed using this approach. For instance, Crammer & Singer propose a ranker Prank based on the Perceptron which maps a feature vector x to the real with a learned weight vector *w* such that the output of the mapping function is just *w*.x. Prank regards a query/document pair as an instance for the input, and each instance is corresponding to a rank level.[9] Harrington has proposed a simple but very effective extension of Prank, which approximates finding the Bayes point by averaging over Prank models.[6] RankProp is also a neural net ranking model proposed by Caruana.[2] RankProp alternates between two phases: an MSE regression on the current target values, and an adjustment of the target values themselves to reflect the current ranking given by the net.

Pairwise Approach: The Pairwise approach transforms ranking to classification on document pairs. It takes document pairs as input: $\{\varphi(q_k,d_i^k),\varphi(q_k,d_j^k)\}$ such that one document is more relevant than another. On these instances system performs pairwise preference learning. The label pair $\{label_i^k, label_j^k\}$ is used to indicate the order of i^{th} and j^{th} document. $label_i^k < label_j^k$ means i^{th} document is more relevant than j^{th} document. After operation, model gives a binary value to an indicator variable y_{ij} of +1 or -1, depending upon the order of the documents in the instance pair. $y_{ij} = +1$ if $R_f(q_k,d_i^k) \geq R_f(q_k,d_j^k)$ and $y_{ij} = -1$ if $R_f(q_k,d_i^k) < R_f(q_k,d_j^k)$. Many algorithms were trained using this approach. For instance, Herbrich cast the problem of learning to rank as ordinal regression – learning the mapping of an input vector to a member of an ordered set of numerical ranks. [7]

They model ranks as intervals on the real line, and consider loss functions that depend on pairs of examples and their target ranks. RankBoost is another ranking algorithm that is trained on pairs. [21] In this algorithm, results are given using decision stumps as the weak learners. It attempts to solve preference learning problem directly, rather than solving an ordinal regression problem.

Dekel has provided a very general framework for ranking using directed graphs, where an arc from A to B means that A is to be ranked higher than B.[4] Joachims proposed RankSVM algorithm, which uses Support Vector machine for optimizing search performance using click-through data. It aims to minimize the number of discordant pairs, which is similar to RankBoost, and to maximize the margin of pair.[8] RankNet is another algorithm that employs relative entropy as a loss function and gradient descent as an algorithm to train a neural network model for document retrieval.[1]

Listwise Approach: In Listwise approach, there is no classification of instances or instance pairs. It tackles the ranking problem directly by optimizing the ordering of the whole list. It treats the list of documents associated with the same query as learning instance to obtain rank and query level information. It takes document collection with respect to query as input : { $\varphi(q_k,d_1^k),\varphi(q_k,d_2^k)$, . . , $\varphi(q_k,d_n^k)$} and produces permutation of these documents as output : Π_n^k where $\varphi(q_k,d_i^k)$ is the feature vector of the i^{th} document w.r.t. k^{th} query. ListNet was one of the first listwise method.

It this, the listwise loss function is defined as cross entropy between two parameterized probability distributions of permutations; one is obtained from the predicted result and the other is from the ground truth.[22] RankCosine was another method. In this, the listwise loss function is defined on the basis of cosine similarity between two score vectors from the predicted result and the ground truth.[16] ListMLE is another listwise method that employ the likelihood loss as the surrogate loss function. They maximize the sum of the likelihood function with respect to all the training queries.[5]

TABLE I. COMPARISON OF DIFFERENT APPROACHES

	Pointwise Approach	**Pairwise Approach**	**Listwise Approach**
Number of instances	Equal to number of training elements	Equal to half of the training elements	Takes whole list as a training instance
Implementation Complexity	O(n)	$O(n^2)$	More complex
Training Time	More	Less	Less
Characteristic	More suitable to ordinal regression	More suitable to learning to rank	More suitable to learning to rank
Technique	Transforms ranking to regression, classification or ordinal regression	Transforms ranking to pairwise classification	Simply represents learning to rank problem
Document Dependence	No dependence between training documents is considered	Document dependence is considered	More dependence between documents is considered
Existing Theories	Easy to use existing theories and algorithms	Easy to use existing theories and algorithms	New theory needed
Flexibility	Flexible to ensure people expected precision	More flexible than Pointwise approach	Less flexible

III. OUR APPROACH

Machine learning provides wide range of algorithms for learning a ranking function. Some of the algorithms are more suitable than others. Two potential learning techniques used are Neural Network (NN) and Support Vector Machine (SVM). NNs are partly inspired on biological learning system and are the most effective learning methods currently known. SVM is based on margin maximization and is the most elegant of all kernel-learning methods. Each has its own advantages and disadvantages. The comparison of both the learning approaches is mentioned in Table 2. In our approach, we have used SVM for pruning to documents that are more critical for learning a ranking function and thus reduced the size of the learning data set. And then NN is used to train the system.

TABLE II. COMPARISON OF SVM AND NN LEARNING

SVM	**NN**
Deterministic Algorithm	Non-Deterministic Algorithm
SVM takes into account learning examples as well as structural behavior. It achieves better generalization due to structural risk management.	Use data empirical risk minimization which stops training once learning error is within a specified margin. This leads to non-optimal model & the solution is often plagued by local minimum problem.
An objective function is convex, so that unlike in the cases of many NN models, any local minimum of an SVM model is also a global minimum.	Suffers from local convergence problem
SVM is comparatively faster than NN. SVM solution is sparse; it only involves the support vectors.	NN suffer from long learning times, which become worse as the volume of data grows

A. Architecture of Hybrid Learning System

Our hybrid learning system works on user feedback as shown in figure 1. Initially, we gave query to the Information Retrieval System which is then internally expanded by the synonyms using WordNet. Initially, term frequency is used to calculate the score of the document and ranked accordingly. Then we select relevant documents out of the result set returned and generate training set for learning the ranking function. This training set is then used by SVM to extract support vectors which are critical for learning. These support vectors are then used to learn a ranking function using back-propagation algorithm using NN. Then the whole result set is ranked again according to the new ranking function. This ranking is compared with the previous ranking using Kendall's Rank correlation coefficient. If the correlation coefficient value is less than the threshold value then the new training set is generated based on user feedback and the whole process is repeated again. We have taken threshold value as 0.6.

B. Implementation

Outliers or meaningless vectors are identified by SVM and can therefore be easily eliminated. Support vectors that are critical in classification are extracted from SVM. Then these support vectors along with their corresponding actual output labels are used to train NN to perform final ranking.

In this way NN training examples are pruned using SVM so that the training examples that are closest to decision boundary are left for final training of the learner.

C. Pruning to Documents Using Support Vector Machine (SVM)

SVM can be characterized as an algorithm that affects a nonlinear mapping of input vectors into a higher dimensional feature space. It involves a dual formulation of governing equations and constraints.

SVM will construct a hyperplane in a high-dimensional features space as the decision surface between positive and negative patterns. SVM formulation approximates SRM (Structural Risk Minimization) principle by maximizing the margin of separation. Basic SVM is linear but it can also be used for non-linear data by using kernel function to first indirectly map non-linear data into linear feature space.

We have used RankSVM algorithm which is based on pairwise approach for extracting critical data patterns. [15], [20]

a) Margin Maximization

Let,

d_i^k represents i^{th} document retrieved from the document collection w.r.t query q_k.

$label_i^k$ and $label_j^k$ are the integer labels of documents d_i^k and d_j^k respectively w.r.t query q_k.

r be the number of features of the document used for learning a ranking function.

A feature vector of r features $\varphi(q_k,d_i^k) = (x_{i1}^k, x_{i2}^k,. . x_{is}^k, . .,x_{ir}^k)$ represented by x_i^k is created from each query-document pair (q_k,d_i^k) k= 1,2,. .,m and i=1,2,. .,n , where x_{is}^k is the value of the s^{th} feature of the i^{th} document w.r.t query q_k.

Then, mathematically ranking function R_f is represented by a weight vector *w* of size r that satisfies

$$\forall \{(d_i^k,d_j^k): label_i^k < label_j^k \in I\} \text{ w.r.t } q_k : R_f(d_i^k) > R_f(d_j^k) \Leftrightarrow w.\ \varphi(q_k, d_i^k) > w.\ \varphi(q_k, d_j^k)$$

SVM has to learn the values of the parameter *w* on a training sample. Here, learning is done pairwise i.e. it takes

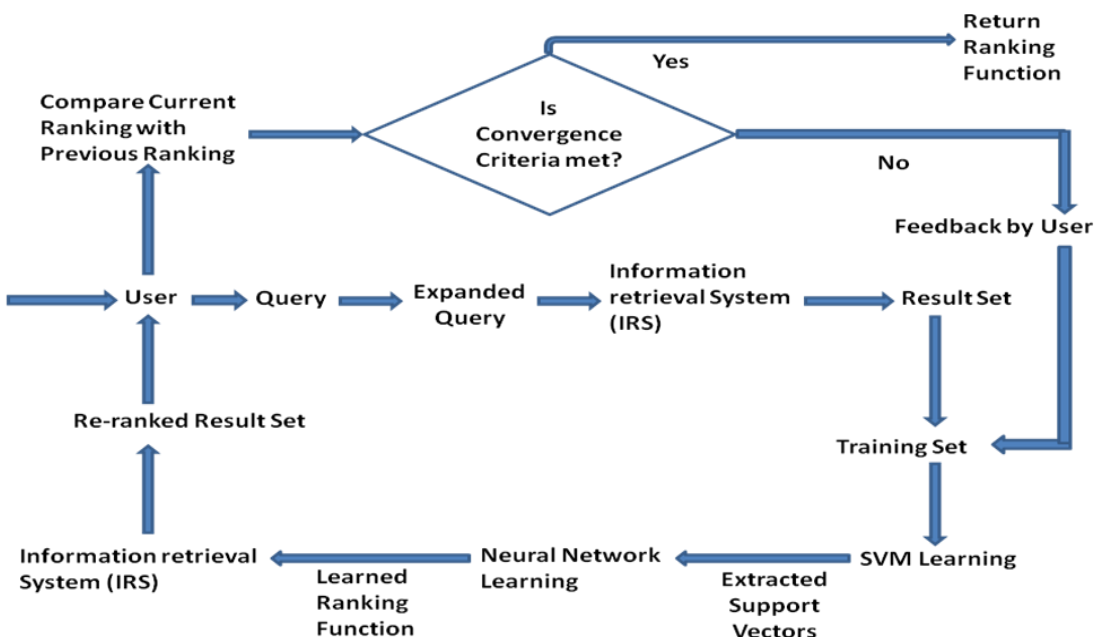


Figure 1. Architecture of Hybrid Learning System

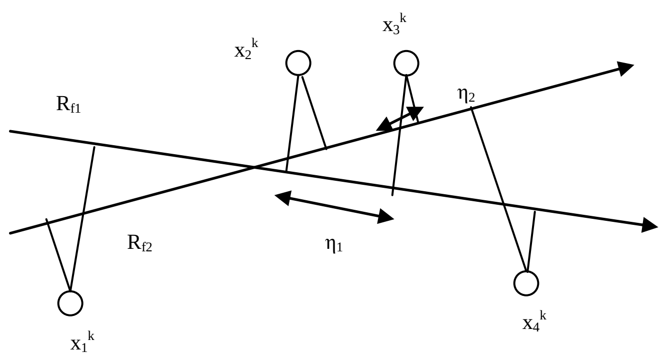


Figure 2. Ranking function as projecting data points

document pairs as input for training such that one document is more relevant than another. [7],[8],[22] A weight vector w is adjusted by a learning algorithm. The training set is denoted as triplet $T=\{(q_k,d_j^k), \varphi(q_k,d_j^k), label_i^k\}$ k=1,2,. . .,m and i=1,2,. .,n where, $label_i^k$ is the label of the i^{th} document w.r.t query k. $label_i^k < label_j^k$ means i^{th} document is more relevant than j^{th} document. Herein strict ordering is assumed.

The ranking model is a real valued function of features:

$R_f(q,d) = w.\varphi(q,d)$ where w denotes a weight vector

To rank all the documents retrieved w.r.t. query q_k, ranking model $R_f(q_k,d_j^k)$ gives score to each document d_j^k as their degree of relevance with respect to query q_k and sort the documents based on that score.

Let, two ranking vectors R_{f1} and R_{f2} and four feature vectors x_1^k ,x_2^k ,x_3^k and x_4^k of the four documents to be ranked. Ranking function can be viewed as the projecting data points of the four documents onto the separating hyperplane as shown in the figure 2. Let, η_1 and η_2 be the distance between the closest points on the hyperplane. From a geometric point of view, calculating the value of the parameters w means looking for a hyperplane called ranking vector that maximizes the distance between various data points of the documents to be ranked according to some criteria. The criterion is based on margin maximization i.e. to maximize the distance between closest points. The distance between these points is calculated as

$$\frac{w \| x_i^k - x_j^k \|}{\| w \|}$$

Given a training set of instance-label pairs, the SVM requires the solution of the following optimization problem:

$$\min_{w,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \sum_{j=1}^{n} \xi_{ij}$$

subject to : $\forall \{(x_i^k, x_j^k) : y_{ij}(w^T(x_i^k - x_j^k)) \geq 1 - \xi_{ij},$

$\forall (i, j) : \xi_{ij} \geq 0.$

where, i,j = 1,2,3 . . .n. and y_{ij} is an indicator variable to indicate order of i^{th} and j^{th} document and C = "capacity" is a tuning parameter for controlling the generalization ability of an SVM.

In the first part of the equation, it maximizes the margin or distance which is inversely proportional to $||w||$ by minimizing $||w||^2/2$ and second term is the sum of in-sample ranking errors ξ_{ij} times the parameter C. In most cases, $\xi_{ij}=0$, that means i^{th} and j^{th} documents are ranked correctly. Thus, SVM maximizes the margin width while minimizing errors. Thus, this problem is quadratic i.e. convex.

C weights in-sample errors and thus controls the generalization ability of an SVM. Higher is C, higher is the weight given to in-sample errors, and lower is the generalization of the learning model. By choosing a low C, the risk of overfitting an SVM on the training sample is reduced. So, the constant C is the soft margin parameter that controls the trade-off between the margin size and training error.

Using the method of Lagrange multipliers, we can obtain the dual formulation which is expressed in terms of variables α_i:

$$\text{maximize}_{\alpha} \sum_{ij} \alpha_{ij} - \frac{1}{2} \sum_{ij} \sum_{uv} \alpha_{ij}\, \alpha_{uv}\, y_{ij} y_{uv} ((x_i - x_j).(x_u - x_v))$$

subject to : C≥0 α_{ij} is a coefficient for a pairwise difference vectors $(x_i\text{-}x_j)$.[19]

We perform a non-linear mapping of the feature vector x onto a high-dimensional space that is hidden from the inputs or the outputs using Kernel function. As SVM solution depends only on the dot product of $(x_i\text{-}x_j)$ and $(x_u\text{-}x_v)$, operations in high dimensional space $\Phi(x)$ do not have to be performed explicitly if we find a function $K(x_i,x_j)$ such that $K(x_i,x_j)$ is equal to dot product of x_i and x_j. This function is called Kernel function.

Using Kernel function, the dual formulation can be expressed as:

$$\text{maximize}_{\alpha} \sum_{ij} \alpha_{ij} - \frac{1}{2} \sum_{ij} \sum_{uv} \alpha_{ij}\, \alpha_{uv}\, y_{ij} y_{uv} K(x_i - x_j, x_u - x_v)$$

subject to : $0 \leq \alpha \leq C$

where K(.) is a kernel function.

We have used the same optimization problem, except the dot product of $(x_i\text{-}x_j)$ and $(x_u\text{-}x_v)$ is replaced by the kernel $K(x_i\text{-}x_j, x_u\text{-}x_v)$. We have used linear Kernel function $K(x_i,x_j) = (x_i^T x_j)^2$ for mapping feature vector x onto a high-dimensional space.

The values of α's are obtained between 0 and C. Data points with non-zero α's are called support vectors which are critical for learning a ranking function. So, all those data point pairs with non-zero α's are extracted and then used to train NN.

b) Format of Training and Test file

The format of training file and test file is same. Details of all the documents are stored in a file in which each row represents one document in a LINE in the following format:

LINE —> L qid : QID F:FV F:FV . . . . F:FV # COMMENT

L —> <float>	// label
QID —> <integer>	// query identifier
F —> <integer>	// feature
FV—> <float>	// feature value
COMMENT —> <string>	// comment as line identifier

Each line contains the target label of the document, query identifier and each of the feature/value pairs are separated by a space character. Feature/value pairs must be ordered by increasing feature number. Features with value zero can be skipped. The target label defines the order of the examples for each query and is used to generate pairwise preference constraints. Two examples are considered for a pairwise preference constraint only if the value of "qid" is same. [14]

D. Learning a Ranking function Using NN

RankNet algorithm is used for learning a ranking function.[1] This algorithm is based on error back-propagation of the NN. Here also, the learning is done pairwise. All the relevant document w.r.t. a query in the training set are paired in such a way that odd numbered document is more relevant than the next even numbered document. E.g. Document positioned at row 1 is more relevant than document at row 2, and document at row 3 is more relevant than document at row 4 and so on. So, if we have 100 documents in the training set it will give 50 such pairs. Thus, in each pair first document is more relevant than second document.

Let the training set is denoted as triplet T= $\{(q_k,d_i^k), (\varphi(q_k,d_i^k),\varphi(q_k,d_j^k)), (label_i^k,label_j^k)\}$ k=1,2,. . .,m and i,j=1,2,. .,n where, $label_i^k$ is the label of the i^{th} document w.r.t query k. $label_i^k < label_j^k$ means i^{th} document is more relevant than j^{th} document. Herein strict ordering is assumed.

In this algorithm, learning consists of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, learning algorithm is given a set of pairs of documents (d_i^k,d_j^k), together with their labels indicating the order of the documents, and its effect propagates through the network layer by layer. A set of outputs is produced as the actual response of the network. During the backward pass, the synaptic weights are all adjusted in accordance with an error-correction rule to minimize the cost function. The cross entropy cost function is used as a loss function for training.[13]

The target probability P_{ij}^* of d_i being ranked higher than d_j is 1 for each document pair (d_i^k, d_j^k) in the training set as we have arranged all the document pairs accordingly. A logistic function used for mapping the output of the NN as a ranking function to a probability is as follows:

$$P_{ij} = \frac{e^{o_{ij}}}{1+e^{o_{ij}}}$$

where $o_{ij} = y_i - y_j$ where y_i is the output of the network for the i[th] document and $P_{ij} = P(d_i > d_j)$ i.e. the probability that the i[th] document is more relevant than j[th] document. The cross entropy is adapted as a loss function for training, which can be represented as:

$$CE_{ij} = CE(o_{ij}) = -P_{ij}^* \log P_{ij} - (1-P_{ij}^*) \log (1-P_{ij})$$
$$= -P_{ij}^* o_{ij} + \log(1+ e^{o_{ij}})$$

where CE_{ij} is the cross entropy loss of the pair (d_i^k, d_j^k), P_{ij}^* is the desired probability, and P_{ij} is the probability modeled by the NN.

Following steps are followed :

1) i[th] document and j[th] document pair along with their labels are presented to the input layer of the network. These inputs are propagated through the network until they reach the output units. This forward pass produces the predicted output in the form of the probability P_{ij}.
2) Because back propagation is a supervised learning algorithm, the desired output in the form of the desired probability P_{ij}^* is given as part of the training vector. The actual network output is then subtracted from the desired output and an error signal in the form of cross entropy loss is calculated.
3) This error signal is then the basis for the back propagation step, whereby the errors are passed back through the neural network by computing the contribution of each hidden processing unit and deriving the corresponding adjustment needed to produce the correct output. The connection weights are then adjusted and the neural network has just "learned" from an experience.

After the network has learned all the new documents can easily be scored by presenting them to the input layer and ranked them according to that score.

E. Features used for Learning

We have used the following features for learning a ranking function:

Term Frequency(TF): $\sum_{q_i \in Q\cap D} \log(c(q_i, D))$ where $c(q_i,D)$ is the frequency of the i[th] term of the query in document D

Inverse Document Frequency(IDF):
$\sum_{q_i \in Q\cap D} \log(idf(q_i))$ where $idf(q_i)$ is the inverse document frequency of the i[th] term of the query

Normalized Cumulative Frequency(NCF):
$\sum_{q_i \in Q\cap D} (\frac{log|C|}{c(q_i,C)})$
Normalized Term Frequency weighted by IDF(NTFI):
$\sum_{i=1}^{n} \log(1 + \frac{c(q_i,D)}{|D|} \times idf(q_i))$

F. Efficacy measure: Kendall's tau τ rank correlation coefficient

Let R^* be the optimal ranking of the documents in which all the documents are ranked according to user's preference. A new generated ranking function R_f is typically evaluated by how closely its ordering approximates optimal ordering of R^*. [11]

The ranking order of a set of training instances is optimized according to Kendall's Tau, τ :

$$\tau(R^*, R_f) = \frac{P-Q}{P+Q} = 1 - \frac{2Q}{\binom{n}{2}}$$

where,

P is the number of concordant pairs (two documents are ordered correctly)

Q is the number of discordant pairs (two documents are ordered incorrectly)

On a finite domain D of n documents, the sum of P (concordant pairs) and Q (discordant pairs) is $\binom{n}{2}$

G. Semantic searching using WordNet

The query keyword used for retrieval of documents is the most significant but not always sufficient. We have used synonyms of the keyword also for searching the corpus. We have created a WordNet containing synonyms of various words and thus stored semantic relations between various words. We have used those semantic relations for expanding the given query and to improve the retrieval effectiveness. Red-Black tree data structure is used to store various words along with their synonyms as shown in figure 3. Each node of the tree contains word along with its various synonyms. Each query term is expanded by the various synonyms before searching in the corpus.[17]

For example, If the query contains the term "estimate" then after expansion it will be replaced by terms "estimate, idea, approximate, appraisal".

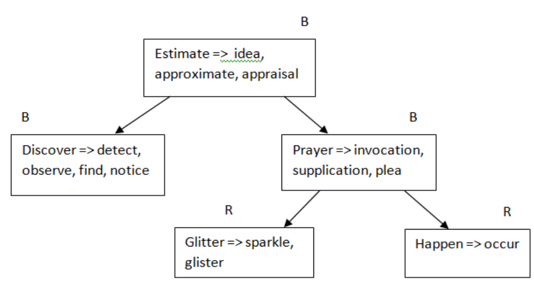


Figure 3. Sample Red-Black tree of Synonyms

IV. EVALUATION MEASURES

A. Recall

It is the measure of the ability of a system to present all relevant items.[3]

$$R = \frac{\text{Number of relevant documents retrieved}}{\text{Number of relevant documents in collection}}$$

B. Precision

It is the measure of the ability of the system to present only relevant items.

$$P = \frac{\text{Number of relevant documents retrieved}}{\text{Total number of documents retrieved}}$$

To evaluate ranked lists, precision can be plotted against recall after each retrieved document. To facilitate computing average performance over set of queries – each with different number of relevant documents – precision values for individual query are interpolated to a set of standard recall levels (0 to 1 in increments of .2). The standard rule to interpolate precision at standard recall level i is to use the maximum precision obtained for the query for any actual recall level greater than or equal to i.

Mathematically, Interpolated precision $P_{interpolated}$ at certain standard recall level i is defined as the highest precision found for any recall level $i' \geq i$:

$$P_{interpolated(i)} = \max P(i') \quad i' \geq i$$

C. Average Interpolated Precision (AIP)

It is the average of the interpolated precision at each standard recall point value for all queries together.

$$AIP_i(Q) = \frac{1}{|Q|}\sum_{j=1}^{|Q|} P_{interpolated(i)}(j) \quad i= 0.0, 0.1, 0.2, \ldots, 1.0$$

where Q is the set of queries $P_{interpolated(i)}$ is the interpolated precision at i^{th} recall value level,

$AIP_i(Q)$ is the average of the interpolated precision at i^{th} recall level for all the queries in set Q.

D. F-Measure (F)

It is harmonic mean of precision and recall. It is a single measure that trades precision versus recall.

$$F = \frac{2*Precision*Recall}{(Precision + Recall)}$$

E. Mean Average Interpolated Precision (MAIP)

It is the mean of all the average interpolated precisions calculated at all standard recall points.

$$MAIP = \frac{\sum AIP_i(Q)}{No\_recall\_points} \qquad i = 0.0, 0.1, 0.2, \ldots, 1.0$$

where $AIP_i(Q)$ is the average of the interpolated precision at i^{th} recall level for all the queries in set Q

No_recall_points are the number of standard recall points used.

V. EXPERIMENTAL EVALUATION

We have implemented Information Retrieval System in Java. For experiments, we have created English-Hindi test collection of about 50 documents extracted from gyannidhi corpus from CDAC. Details of English-Hindi parallel corpus are mentioned in table 3. We have also generated training set triplet of about 20 queries that are stored along with their relevant documents for training. Sample queries for both Hindi and English are shown in table 4. All the data are encoded in Unicode text. F-measure and AIP are used for evaluation.

TABLE III. PARALLEL CORPUS DETAIL

	English	Hindi
No. of Documents	50	50
No. of index terms	7896	12687
No. of queries	20	20
Average No. of terms/doc	152	236

TABLE IV. SAMPLE QUERIES

qid	English	Hindi
1	estimate the approximate age	आयु का अनुमान
2	pleasant sound	मधुर ध्वनि
3	hazardous journeys	जोखिम-भरी यात्राएं

VI. TABLE V. AIP VALUES FOR RANKNET AND OUR HYBRID LEARNER

Recall	AIP	
	RankNet	Hybrid Learner
0.2	0.3	0.4
0.4	0.45	0.54
0.56	0.52	0.65
0.7	0.63	0.7
0.83	0.74	0.8

TABLE VI. F-MEASURE VALUES FOR RANKNET AND HYBRID LEARNER

Recall	F-Measure	
	RankNet	Hybrid Learner
0.2	0.36	0.37
0.4	0.49	0.54
0.6	0.61	0.65
0.8	0.77	0.79

VII. RESULTS AND DISCUSSION

We now describe our experimental analysis of our learning system. We evaluated ranking functions for 20 queries by calculating AIP and F-Measure values on different recall values ranging from 0 to 1. Comparison of these values for RankNet and our hybrid is learner shown in table 4 and table 5 respectively. Their comparison is also shown graphically in figure 4 and figure 5 respectively. MAIP of retrieved documents using RankNet and our hybrid learner is 0.528 and 0.618 respectively. There is significant improvement in F-measure and AIP values using proposed hybrid learner.

VIII. CONCLUSION

Learning to Rank is applied to automatically learn a ranking function from the training data. In this paper, a new hybrid learner is introduced based on NN and SVM that gives better performance than learning using NN alone. SVM is used to extract support vectors from the training data which are critical for learning a ranking function. These support vectors contribute better in learning a ranking function. We have used semantic searching to improve search accuracy by using synonyms of the query term to retrieve more relevant results. For experiments, we have constructed English-Hindi IR data collection from Gyannidhi parallel corpus. This system works language independently. F-measure and AIP are used for evaluation. F-measure and AIP values of the retrieved documents are improved using the proposed hybrid learner.

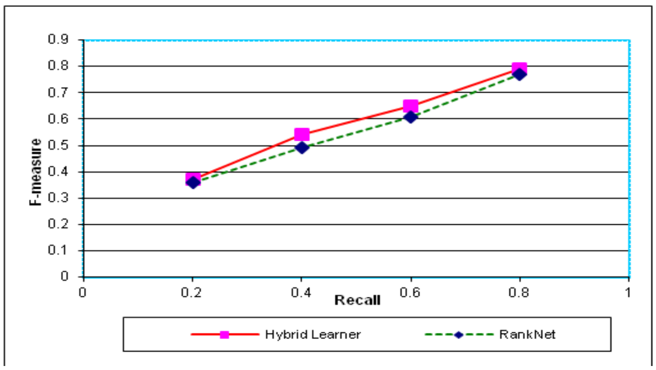


Figure 4. F-Measure using RankNet and Hybrid Learner

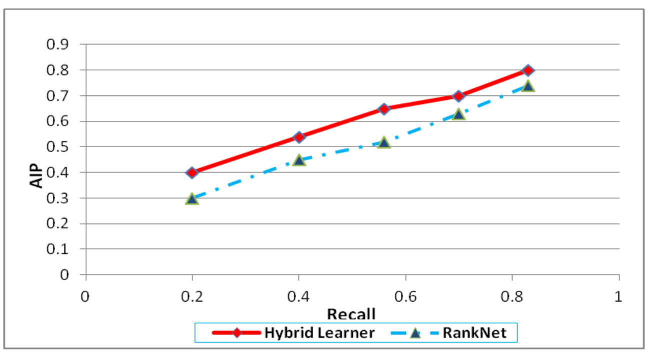


Figure 5. AIP using RankNet and Hybrid Learner

REFERENCES

[1] C. J. C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender. Learning to rank using gradient descent. In Proceedings of the International Conference on Machine learning, pages 89–96, 2005.

[2] Caruana, R., Baluja, S., & Mitchell, T. (1996). Using the future to "sort out" the present: Rankprop and multitask learning for medical risk evaluation. NIPS 8 (pp. 959-965).

[3] Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze. Introduction to Information Retrieval, Cambridge University Press,2008.

[4] Dekel, O., Manning, C., & Singer, Y. (2004). Log-linear models for label-ranking. NIPS 16.

[5] Fen Xia, Tie-Yan Liu, Jue Wang, Wensheng Zhang, Hang Li: Listwise approach to learning to rank: theory and algorithm. ICML 2008: 1192-1199.

[6] Harrington EF. Online ranking/collaborative filtering using the perceptron algorithm. In: Proceedings of the twentieth international conference on machine learning (ICML-2003), Washington, DC.

[7] Herbrich, R., Graepel, T., & Obermayer, K. Large margin rank boundaries for ordinal regression. Advances in Large Margin Classifiers, MIT Press, pages 115-132, 2000.

[8] Joachims, T.: Optimizing search engines using clickthrough data. In: Proc. ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (SIGKDD '02), pages 133–142, 2002.

[9] K.Crammer and Y. Singer. Pranking with ranking. In Advances in Neural Information Processing Systems (NIPS), 2001.

[10] M.-F. Tsai, T.-Y. Liu, T. Qin, H.-H. Chen, and W.-Y. Ma. Frank: a ranking method with fidelity loss. In SIGIR '07: Proceedings of the30th annual international ACM SIGIR conference on Research and development in information retrieval, pages 383–390, 2007.

[11] Mirella Lapata.Automatic Evaluation of Information Ordering: Kendall's Tau. In Association for Computational Linguistics.Volume 32,Number 4.2006.

[12] O. Chapelle, Y. Chang, and T.-Y. Liu. The Yahoo! Learning to Rank Challenge, 2010.

[13] Simon Haykin. Neural Networks and Learning machines, PHI Learning Private Limited, 2010.

[14] SVM-Light, http://svmlight.joachims.org/

[15] T. Joachims. Training linear svms in linear time. In KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, 2006.

[16] T. Qin, T.-Y. Liu, M. feng Tsai, X.-D. Zhang, and H. Li. Learning to search web pages with query-level loss functions. Technical Report, 2006.

[17] Thomas H. Cormen, Introduction to Algorithms, MIT Press. 2001.

[18] Tie-Yan Liu. Learning to Rank for Information retrieval, Springer.2008

[19] V. Kecman, Support Vector machines : Theory and Applications, Springer.2005

[20] Y. Cao, J. Xu, T.-Y. Liu, H. Li, Y. Huang, and H.-W. Hon. Adapting ranking SVM to document retrieval. In SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, 2006.

[21] Yoav Freund , Raj Iyer , Robert E. Schapire , Yoram Singer, An efficient boosting algorithm for combining preferences, The Journal of Machine Learning Research, 4, p.933-969, 12/1/2003

[22] Z. Cao, T. Qin, T.-Y. Liu, M-F. Tsai, H. Li.Learning to Rank: from pairwise approach to listwise approach.In Proceedings of 24th Annual International Conference on Machine Learning,pages:129- 136, 2007.

AUTHORS PROFILE

Pooja Arora was born in 1978 in Delhi, India. She took her Bachelor of Computer Science from Delhi University in 1998. Then she did her MCA from Banasthali Vidyapith Rajasthan, India. She has 10 years' experience in teaching MCA students in AKG engineering college, Uttar Pradesh India. Currently she is also pursuing Ph.D. under the guidance of Prof. Om Vikas and her area of research is information retrieval.

Prof. Om Vikas is former director IIITM, Gwalior India. He is also senior member IEEE, India.

Evaluation of Perception and Performance in ICT Related Courses

Annie O. EGWALI
Department of Computer Science,
University of Benin, Benin City.
Nigeria.

Efosa C. IGODAN
Department of Computer Science,
University of Benin, Benin City.
Nigeria.

***Abstract*— Some teaching methods adopted for disseminating Information Communication Technology Related Courses (ICTRC) in institutions of learning have been observed to be inadequate in bringing about the right perception and performance in students. In order to qualitatively establish the efficiency of tutoring ICTRC, this study investigates the effect of ICT resources on students' perception and performance in ICTRC. Two hundred and forty-eight (248) students in the Department of Computer Science/ICT from two universities: a Federal and a Private University from the South-South region of Nigeria were used. A pre-test, post-test control group and quasi experimental design were utilized. Findings revealed that teaching ICTRC without using ICT resources are not effective at empowering the students with the right perception and the students are not enabled to perform to the best of their ability. Some recommendations are proffered encouraging ICT aided teaching strategies in Nigeria.**

***Keywords*- Learning, ICT, Education, Perception, Performance, Examination.**

I. Introduction

According to [1], Information Communication Technology (ICT) is the acquisition, dissemination, processing and storage of numerical, vocal, textual and pictorial information by a microelectronics-based combination of computing and telecommunications. ICT plays a dominant role in our present environment enabling humans to understand the increasingly technological changing society. ICT provides learners with understanding, skills and scientific knowledge needed for scientific research, fostering technological and economic growth in the society, where they live thus improving the standards of living [9]. ICT enables learners to acquire problem-solving and decision-making skills that provides ways of thinking and inquiry which helps in radical changes.

With the rapid usage of ICT resources (i.e. internet), ICT based teaching-learning applications are considered an effective alternative to traditional teaching methods in that it affords students unlimited opportunities to demonstrate mastery of contents taught [6]. However within institutions of learning, new challenges have been generated, one of which is the teaching and learning of ICT Related Courses (ICTRC) like Computer Science, Software Engineering, Computer Education, Computer Engineering Etc. Lately, in some institutions, there has been a decline in academic achievement scores of students taking ICTRC and low performance in ICTRC at both the qualifying examination (SSCE) and placement examination like University Matriculation Examination - UME [2]. Some researchers confirm that there is also low enrolment for ICTRC, for students shun ICTRC as a discipline when given an option of courses to study in institutions.

Often times the teacher is blamed for the students' lack of interest in a course and poor performance. This is not farfetched because the teaching approach that a teacher adopts is a factor that affects students' perceptive power and performance in a course of study; however [8] and [3] postulated availability of teaching resources, capital resources, students attitude towards the teacher's teaching approach, the subject-matter and the teaching method as critical factors to the successful teaching and learning of ICTRC. Four factors critical to the effective learning of ICTRC in the classrooms and laboratories were identified by [4]: respondents' attitude, software selection, a proper utilization direction, and the capabilities of ICT professional educators.

There have been diverse research studies on students' effective learning of ICTRC. Reglin [11] studied the effects of teaching ICTRC on a sample of 53 minority prospective teachers, but focused on the issue of cooperativeness. A study conducted by [10] on a sample of 36 teachers centering on the use of ICT resources for ICT education based on knowledge acquisition. Another study by [7] focused on the educational use of ICT resources and students' performance while a research study by [5] compared the effectiveness of teaching agricultural science using ICT based learning against the regular teaching approach.

In this study, we posit that an ignored domain is the effect of gender as it affects single sex schools and coeducational schools. In the educational system, gender is taken into cognizance as it influences the curriculum, instructional materials, choice of career and general behavior of students and teachers alike. The relationship between gender and ICTRC is a vital area of research due to the fact that there are conflicting nature of results from researches that focus on gender and ICTRC.

The present study is interested in gender influence on the level of students' perception of teachers' effectiveness at teaching ICTRC and performance in ICTRC using ICT resources. Results from this study enable recommendations that promote ICT efficiency among students and lecturers in Nigeria and initiate some degree of resolution on the conflicting nature of learning ICTRC.

II. Statement Of The Problem

In ascertaining the level of students' performance in ICTRC, some ignored domains are the effect of gender as it affects single sex schools and coeducational schools and students' perception of teachers' effectiveness. Thus presently, there is a dearth of information relating to the level of students' perception of teachers effectiveness in teaching and students performance in ICTRC using ICT resources within the classroom and laboratories gender composition. It becomes crucial to provide information relating to the effects of classroom gender composition, which enables the evaluation of the consequences of disparities of the sex ratios and the perception of students as regards lecturers' effectiveness in teaching ICTRC via ICT resources. It is strongly believed that this aforementioned will provide the optimum level of performance of ICTRC within the student population.

III. Objectives Of The Study

The objectives of this study are:

- to compare the performance of male and female students taught ICTRC using ICT resources or the conventional lecture method.
- to investigate the effect of gender composition in students' performance in ICTRC when being taught with ICT resources or the conventional lecture method.
- to investigate students' perception of ICTRC taught by means of ICT resources as

IV. Hypothesis

The following null hypotheses were tested at 0.05 levels of significance:

- **Ho_1**: There is no significant difference in the performance of students taught ICTRC by means of ICT resources or the conventional lecture method.
- **Ho_2**: There is no significant difference in the effect of gender composition in students' performance in ICTRC taught by means of ICT resources or the conventional lecture method.
- **Ho_3**: There is no significant difference in students' perception of ICTRC taught by means of ICT resources or the conventional lecture method.

V. Research Methodology

The study design employs a pre-test, post-test control group and quasi experimental design to determine the significant effect of ICT resources on students' performance in ICTRC. The following three variables were utilized in the study.

(1) Independent Variable:
   a) Teaching via ICT resources
   b) Conventional lecture method

(2) Dependent Variables:
   a) Performance in ICTRC Test
   b) Perception of ICTRC taught via ICT resources.

(3) Moderator Variables:
   a) Gender (Male and Female)

B. The Population And Sample

The population of this study consisted of 248 randomly selected students offering ICTRC from university institutions in the South-South region of Nigeria. The stratified random sampling was used to select one Federal University (Uni-1) and one Private University (Uni-2) and to ensure evenness in institutions. Out of two institutions, some students were assigned as the experimental and control groups respectively.

C. Instrumentation and Validation

The following three instruments were designed and used for the study:

1) Instructional Plan for Teaching using Conventional Lecture Method (IPTCLM)

2) Instructional Plan for teaching using ICT Resources (IPITF).

3) Students' ICTRC Perception and Performance Test (SPPT).

a) ***Instructional Plan for teaching using Conventional Lecture Method (IPTCLM)****: To ensure consistency and uniformity, this teaching plan which involves teaching in the conventional lecture method, was used for all students in the control groups. It included the course objectives, instructional materials and teaching procedures of computer science concepts in the classroom. This plan was made available to teachers prior to the training sessions.*

b) ***Instructional Plan for teaching using ICT Resources (IPITF):*** *To ensure consistency and uniformity, this teaching plan includes subject area, objectives and the expected procedure to be followed by the teachers in the teaching of Computer Science concepts. This plan was made available to teachers prior to the training sessions and was used for all students in the experimental groups.*

Students ICTRC Perception and Performance Test (SPPT): This instrument is divided into two sections designed by the researchers in the areas of Computer Science. The first section of the test was designed to cover the areas of perception which includes the areas of knowledge and comprehension. The section has 50 multiple choice objective items with options A to E designed to test the students' perception of computer science based on the methods (via ICT resources or lecture method) employed in inculcating the subject area to them. The second section consists of 50 questions to test the students' performance in computer science. The maximum obtainable score was 100.

The face-validity and content-validity of the instruments was verified by three experts in the subject area. The instrument had a high Cronbach's alpha values, both for items within each subsection (range: 0.863 to 0.870) and for the instruments as a whole (0.951).

D. Data Collection Methods

For two weeks, the researchers familiarize themselves with the students and the educational strategies to employ for the control and experimental groups using the IPTCLM and IPITF instruments. In the next two weeks, the pre-test was conducted and the instruments SPPT was administered to all the participating students to assess their initial performance level before the treatment effect would takes place. For four more weeks, the IPITF was administered to the experimental group who were taught using ICT resources. PowerPoint presentations and Video-clips were used to present animations techniques, animation concepts and constructive procedures. The video-clips include real life pictures of computer animations. These animations help to emphasize the critical points of computer science concepts that are employed in real settings that allow meaningful learning. Concurrently, the IPTCLM was used to teach the same concepts as that of the experimental group, only that in this case, the students were subjected to the conventional lecture method of teaching. For both groups, the researchers carried out the same activities. In the last week of the study, at two different instances, (one instance for an institution), the post-test was administered on all the participating students. The students were examined and their scripts were marked and results compiled.

E. Method Of Data Analysis

The post-test data on students' performance in ICTRC were analyzed using means and standard deviation. A five point Likert scale was used to measure all the statements on students perception of ICTRC taught by means of ICT resources or the conventional lecture method (Strongly Disagree (SRD), Disagree (D), Somewhat Disagree (SD), Neither Agree or Disagree (NAD), Somewhat Agree (SWA), Agree (A) and Strongly Agree (SA)).

VI. Results And Discussions

TABLE 1: POST-TEST PERFORMANCE MEAN SCORE AND STANDARD DEVIATION OF SPPT

Variable	N		N	Mean	Std. Dev.
Experimental Group	Uni-1	Male	63	18.11	2.443
		Female	41	18.01	2.542
		Total	104	18.06	2.887
	Uni-2	Male	08	17.41	1.662
		Female	05	17.33	2.221
		Total	13	17.37	2.237
Control Group	Uni-1	Male	73	17.40	2.114
		Female	45	16.10	2.213
		Total	118	16.75	2.225
	Uni-2	Male	07	16.71	1.996
		Female	06	16.09	2.117
		Total	13	16.40	2.113
Gender	Uni-1	Male	136	18.22	2.167
		Female	86	18.13	2.213
		Total	222	18.18	2.201
	Uni-2	Male	15	17.46	1.887
		Female	11	17.33	1.954
		Total	16	17.40	1.913

The post-test mean score and standard deviation is shown in Table 1, which shows that Uni-1 students performed better than Uni-2 students. Generally, the experimental group who were being taught ICTRC with ICT resources performed better than students in the control group who were being taught using the conventional lecture method, therefore Ho_1 is rejected. It is also evident that the male students performed better than the female students in each of the groups, thus Ho_2 is rejected.

TABLE 2: EXPRESSIONS TO ESTABLISH STUDENTS' PERCEPTION OF ICTRC VIA ICT RESOURCES / LECTURE METHOD.

Questions	SRD	D	SD	NAD	SWA	A	SA
It did not take me long to answer the entire questions.	4	2	-	-	1	8	16
The manner of teaching was helpful	3	1	-	-	1	7	19
The allotted time for the test was too long	3	1	1	2	3	5	16
It was easy to remember the lesson taught	4	-	-	1	2	6	18
I think that the lesson was very well taught	-	3	1	-	6	1	20
My knowledge on the subject matter got better with practice.	3	4	-	-	-	2	22
I understood the entire lesson taught	15	8	1	1	1	-	5
It was easy to complete the entire posttest exercise.	4	4	-	-	1	5	17
Total	36	23	3	4	15	34	133

The students were asked the above eight questions to ascertain their perception of ICTRC taught by means of ICT resources or the conventional lecture method. The students' assessment is represented in Table 2. There was a significant difference on the eight questions as can be ascertained from the Chi-square analysis (Table 3).

TABLE 3: CHI-SQUARE ANALYSIS

X	fo	fe	(fo – fe)	$(fo – fe)^2$	$x^2 = \frac{(fo - fe)^2}{fe}$
Strongly Disagree	36	35.43	0.57	0.32	0.009
Disagree	23	35.43	-12.43	154.50	4.35
Somewhat Disagree	3	35.43	-32.43	1051.70	29.68
Neither Agree or Disagree	4	35.43	-31.43	987.84	27.88
Somewhat Agree	15	35.43	-20.43	417.38	11.78
Agree	34	35.43	-1.43	2.04	0.06
Strongly Agree	133	35.43	97.57	9519.90	268.70
Total	**248**	**248**			**342.46**

X^2 value = 342.46; Degree of freedom = 7 – 1 = 6; X^2 at .99 at 4 *d.f* = 16.812
X^2 at .95 at 4 *d.f* = 12.592; 342.46 > 16.812 > 12.592

The Chi square analysis (Table 3) showed that the differences in perception are not due to chance. We can thus conclude that student' perception and understanding of ICTRC got better with practice, for all 20 students who acknowledged that the lesson was very well taught were those taught using ICT resources. It is also this class of students that accepted the fact that the subject matter got better with practice and that the manner of teaching was helpful.

VII. CONCLUSION

Based on the findings in this study, it is evident that ICT-based resources and applications should drive the teaching-learning process of all ICTRC. Presently most educators adopt a theory-first approach and sometimes suspend teaching the application skills. For effective learning and performance, theory and applications should be interwoven and integrated. Or, at least, applications should be considered first, and then theory, to ensure that theory is related to real-world concepts and to enable students form conceptual relationships between theory and applications and create life-long learning experiences. Creating ICT-based case studies and interdisciplinary scenarios in ICTRC would also enable the students to solely define relevant concepts, identify theories and tools needed to solve conceptual problems, process data, report results, and providing all relevant documentation. To accomplish all the aforementioned and more, Departments within Faculties of ICTRC must work as a team. The team should discover, discuss and implement the interweaving of course materials.

VIII. RECOMMENDATIONS

The following recommendations were made based on the finding of the study:

- ICT teachers should use ICT resources in teaching ICTRC.
- ICT, IT and Science Faculties could collaborate to develop and teach courses for IT and its associated ICT.
- Seminars/worships should be organized for ICTRC teachers in school on the use of ICT resources.
- The government should set up laboratories and establish adequate ICT resources in all institutions of learning and make it possible for all disciplines to make use of it.

REFERENCES

[1] C. Adelman, "A Parallel Post-secondary Universe: The Certification System in Information Technology". (2000). Available at: http://www2.ed.gov/pubs/ParallelUniverse/universe.pdf

[2] D. Aduda, "Meeting gives Tips on How to Improve Science Subjects. *The Daily Nation, Nairobi,* Kenya": Nation Media group Ltd. September 2003, p.15.

[3] A. O. Egwali and V. V. N. Akwukwuma "Gender Inequality Issues in Educational Institutions: Goals for Millennium Development". *Proceedings of the 2009 Third World Organization for Women in Science African Regional Conference, TWOWS*' 09. November 16 - 20. (2009). Pg. 158 – 166. Abuja. Nigeria.

[4] D. J. Kadijevick, "Four critical issues of applying educational technology standards to professional development of ICT teachers". *Proceedings of the 2nd International Conference on the Teaching of ICT at the Undergraduate Level*, University of Crete. (2002).

[5] J. K. Kibett, & N. J. Kathuri, "Effects of Projected- Based Learning on Student Performance in Secondary School Agriculture". *Zimbabwe Journal of Educational Education*, (2005). 37(1), 18-24.

[6] T. L. Larkin, "Learning style in the classroom: a research guided approach". *Paper presented at the Annual Conference of International Conference on Engineering and Computer Education*, March 16 – 19, 2003, Sao Paulo, Brazil.

[7] E. Leuven, M. Indahl, H. Oosterbeek, and D. Webbink, "The Effect of Extra Funding for Disadvantaged Pupils on Achievement". *IZA Discussion Paper*. No. 1122. Bonn: Institute for the Study of Labor. (2004).

[8] H. R. Mills, "*Teaching and Training":* A handbook for instructors (3rd ed). London: Macmillan Publishers. (1991).

[9] O. Minishi, E. Muni, O. Okumu P. Mutai, G. Mwangasha, H. Omolo & F. Munyeke, "*Secondary Physics Form One" 3rd ed*. Kenya Literature Bureau. Nairobi. (2004).

[10] J., Moor, & R. Zazkis, "Learning ICT in a virtual classroom: Reflection on experiment". *Journal of Computers in ICT and Science Teaching,* (2000). *19 (2),* 89-113.

[11] G. L. Reglin, "The effects of individualized and cooperative computer assisted instruction on ICT achievement and ICT anxiety for prospective teachers". *Journal of Research on Computing in Education,* (1990). *22 (2),* 404-413.

AUTHORS PROFILE

Dr. (Mrs) Egwali Annie Oghenerukе, is presently a Senior Lecturer in the Department of Computer Science, Faculty of Physical Sciences, University of Benin. Nigeria. She obtained her B.Sc. degree in Computer Science in 1995 from the University of Benin, M.Sc. degree in Computer Science in 2001 from the University of Benin and PhD in degree in Computer Science in 2010 from the University of Benin. She lectured in the College of Education, Ekiadolor Benin for eight years during she engaged in various administrative and charitable functions. Her area of interests includes Information Technology, Software Engineering, Gender studies, E-commerce, Electronic Marketing and Software Security. To date, she has supervised several diploma, undergraduate and postgraduate students. She is a member of International Network for Women Engineers and Scientists (INWES), Nigerian Computer Society (NCS), Nigerian Association of Educationists for National Development (NAEND, National Association for the Advancement of Knowledge (NAFAK) and Third World Organizations of Women Scientists (TWOWS).

Igodan Charles Efosa, is currently a lecturer in the Department of Computer Science, University of Benin, Edo State, Nigeria. He obtained his B.Sc. degree in Computer Science in 2004 from University of Benin (UNIBEN) and M.Tech degree in Computer Science in 2012 from Federal University of Technology, Akure, Nigeria. He has worked as system engineer/consultant for 4years with NetCom Tech. Consortium and programme analyst in the University of Benin Teaching Hospital (UBTH) for 4years during which he has been teaching in various institutions. His interests are in Data mining, knowledge warehousing, expert system, information systems and databases.

Effects of Pronunciation Practice System Based on Personalized CG Animations of Mouth Movement Model

Kohei Arai [1]
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Mariko Oda [2]
Faculty of Engineering,
Kurume Institute of Technology University,
Kurume City, Japan

***Abstract*— Pronunciation practice system based on personalized Computer Graphics: CG animation of mouth movement model is proposed. The system enables a learner to practice pronunciation by looking at personalized CG animations of mouth movement model , and allows him/her to compare them with his/her own mouth movements. In order to evaluate the effectiveness of the system by using personalized CG animation of mouth movement model, Japanese vowel and consonant sounds were read by 8 infants before and after practicing with the proposed system, and their pronunciations were examined. Remarkable improvement on their pronunciations is confirmed through a comparison to their pronunciation without the proposed system based on identification test by subjective basis.**

Keywords- Pronunciation practice; mouth movement model; CG animation.

I. Introduction

There are many pronunciation practice systems [1] which allow monitoring voice waveform and frequency components as well as ideal mouth, tongue, and lip shapes simultaneously. Such those systems also allow evaluations of pronunciation quality through identification of voice sound. We developed pronunciation practice system (it is called "Lip Reading AI") for deaf children in particular [1]. The proposed system allows users to look at their mouth movement and also to compare their movement to a good example of mouth and lip moving picture. Thus users' pronunciation is improved through adjustment between users' mouth movement and a good example of movement derived from mouth movement model.

Essentially, pronunciation practice requires appropriate timing for controlling mouth, tongue, and lip shapes. Therefore, it would better to show moving pictures of mouth, tongue, and lip shapes for improvement of pronunciations [2]. Although it is not easy to show tongue movement because tongue is occluded by mouth, mouth moving picture is still useful for improvement of pronunciations. McGurk noticed that voice can be seen [3]. Some of lipreading methods and systems are proposed [4]-[6].

One of the key issues for improvement of efficiency of the pronunciation practice is personalization. Through experiments for the proposed "Lip Reading AI" with a number of examiners, it is found that pronunciation difficulties are different by examiner. Therefore, efficient practice needs a personalization. The proposed pronunciation practice system in the paper utilizes not only mouth movement of moving pictures but also personalization of moving picture by user.

The following section describes the proposed system followed by some experimental results with 8 examiners. Then effectiveness of the proposed system is discussed followed by conclusion.

II. Proposed Pronunciation Practice System

A. Lip Reading "AI"

Previously proposed Lip Reading "AI" allows comparison between learner's mouth movement and reference movement with moving picture in real time basis. An example of display image is shown in Fig.1.

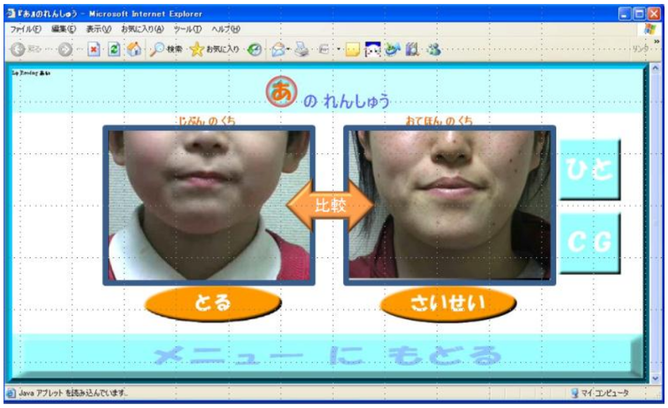


Figure 1 Example of display image of the previously proposed Lip Reading "AI".

B. CG Animation of Reference Moving Picture

In the system, real mouth images are used as reference movement of moving picture. Not only real mouth moving pictures, but also CG animation of mouth images can be used

---

1 http://www.advanced-media.co.jp/products/amivoicecallpronunciation.html
http://www.prontest.co.jp/soft/
http://www.english-net.co.jp/~pros/1/ppower/progfeat.htm
http://shop.alc.co.jp/course/hc/
http://www.smocca.co.jp/SMOCCA/English/HatsuonRyoku/index.html
http://sgpro.jp/demo/

as shown in Fig.2. Using CG animation of mouth moving picture, much ideal reference could be generated. "Maya"[2] of CG animation software (Fig.3) is used to create reference mouth moving picture. Then it can be personalized. Namely, resemble CG animation to the user in concern can be created as shown in Fig.4.

Figure 2 CG animation based reference moving picture for monitoring mouth movements.

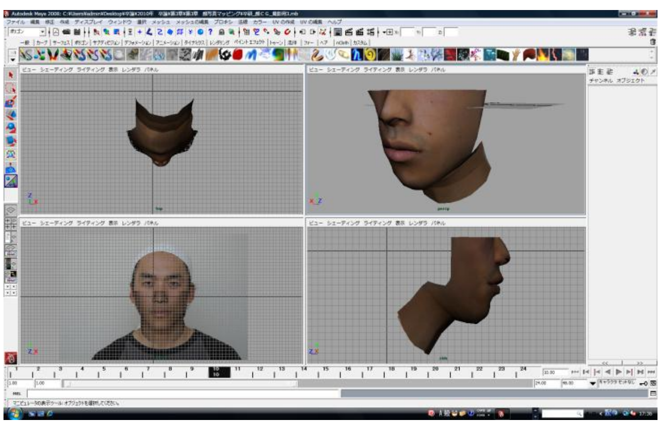


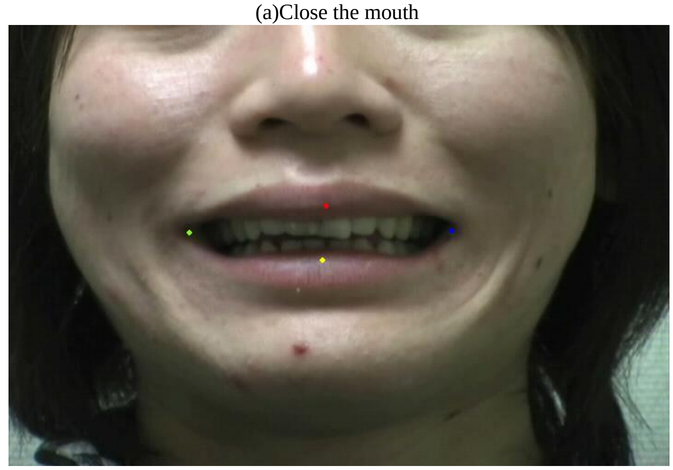
Figure 3 Example of Maya of CG animation software generated reference mouth and lip moving picture

In order to create resemble mouth moving picture to the user in concern, correct mouth movements are extracted from moving picture by using Dipp-MotionPro2D[3]. 12 of lecturers' mouth moving pictures are acquired with video camera and then are analyzed. Every lecturer pronounced "a", "i", "u", "e", and "o". Four feature points, two ends of mouth and middle centers of top and bottom lips are detected from the moving pictures. Four feature points when lecture close and open the mouth are shown in Fig.5 (a) and (b), respectively.

An example of motion analysis of four feature points when the lecture pronounces "a" is shown in Fig.6.

Figure 4 Resemble mouth and lip moving picture to the user in concern could be created with the CG animation software derived reference moving picture.

(a)Close the mouth



(b)Open the mouth

Figure 5 Four feature points when lecture close and open the mouth.

In the figure, red, yellow, light green, and blue lines show the top lip, the bottom lip, right end of mouth and left end of mouth, respectively. When the lecture pronounces "a", the bottom lip moves to downward direction remarkably while other three feature points (top lip and two ends of mouth) do not move so much. On the other hand, when the lecture pronounces "u", two ends of mouth moves so much in comparison to the other two (top and bottom lips) as shown in Fig.7.

[2] http://ja.wikipedia.org/wiki/Maya

[3] http://secure.shanon.co.jp/ipjbio2008/exhidir/BIO/ja/company/exhibitor_1032.html

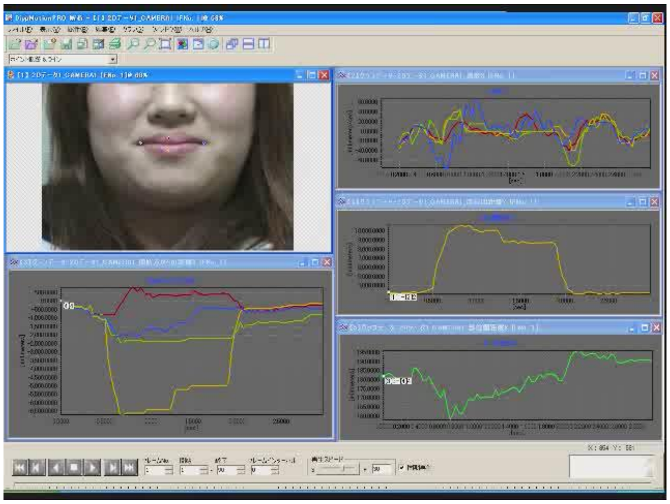


Figure 6 Example of motion analysis of four feature points when the lecture pronounces "a"

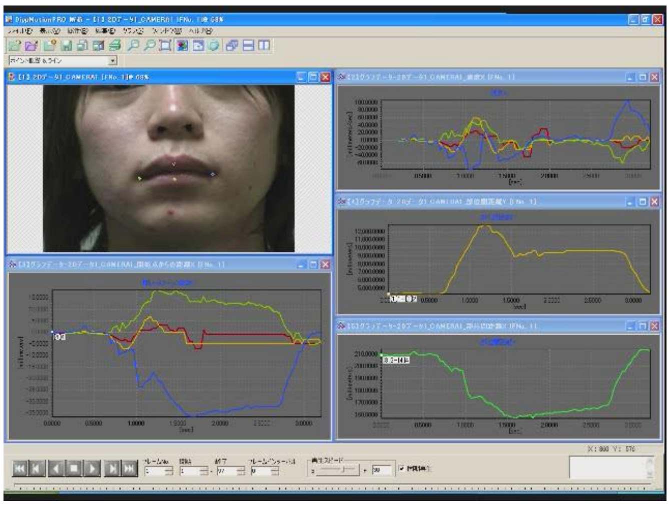


Figure 7 Example of motion analysis of four feature points when the lecture pronounces "u"

C. CG Animation Method

3D model with wireframe has to be created first as shown in Fig.8. Then rendering is followed by. Once basic 3D shape model is created, and then the target 3D models are created as shown in Fig.9.

After that, intermediate pictures are created with basic and the target 3D models as key frame method as shown in Fig.10.

During the process for creation of intermediate frames, chin movement has to be realistic. In this process, jaw bone is assumed as shown in Fig.11. Fig.11 (a) and (b) shows model derived lectures mouth and lip images when lecture closes and opens the mouth, respectively. On the other hands, Fig.11 (c) and (d) shows the assumed jaw bones when lecture close and open the mouth, respectively. The mouth can be opened when the jaw bone angle is increased.

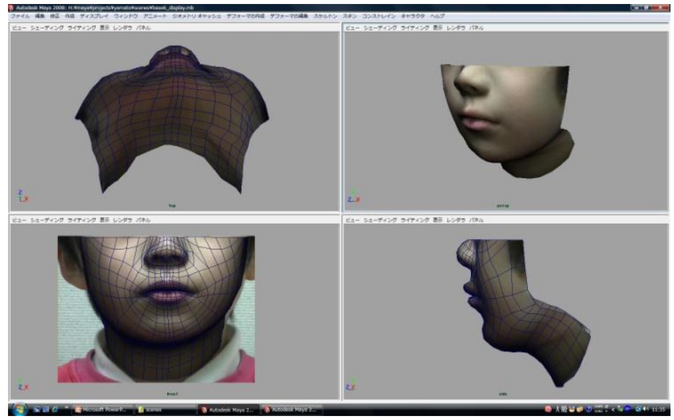


Figure 8 3D model with wireframe of mouth and lip

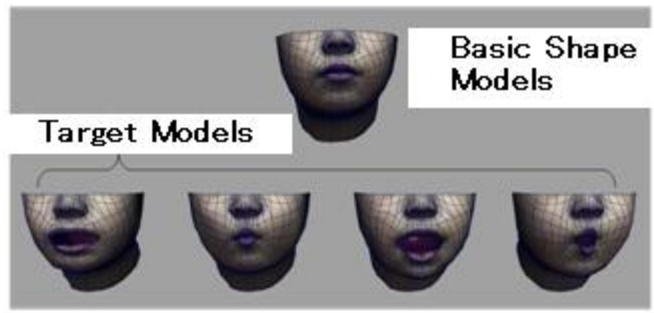


Figure 9 Target models of target 3D models are created from the basic 3D shape model.

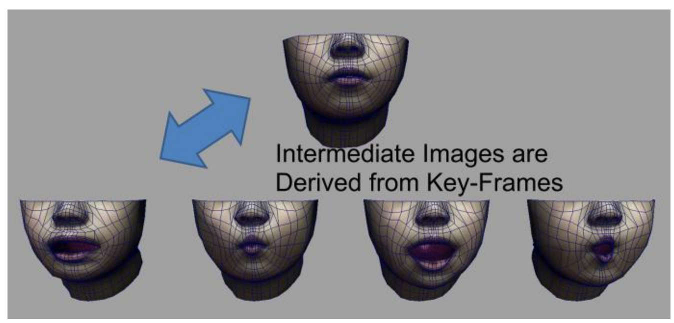
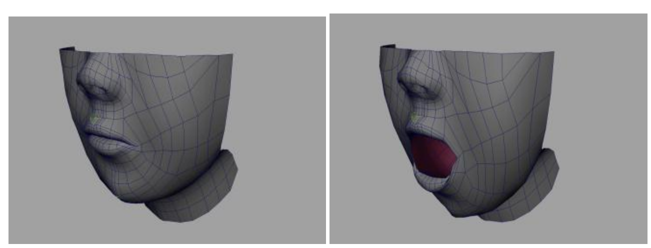
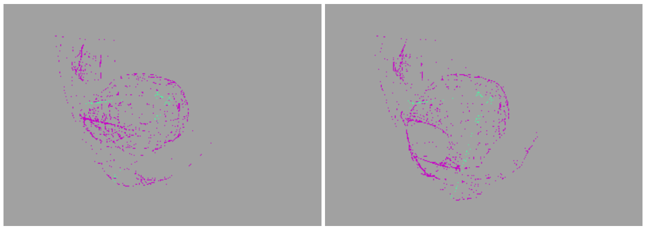


Figure 10 Intermediate pictures which are created with basic and the target 3D models as key frame method

Thus CG animated moving picture of mouth movements can be created as shown in Fig.4. All the feature points and picture for rendering are derived from the learners and lecturers so that CG animated moving picture is resemble to them.



(a)Close the mouth (b) Open the mouth



(c) Close the mouth (d) Open the mouth
Figure 11 Model derived lectures mouth and lip images together with assumed jaw bone when lecture closes and opens the mouth.

III. EXPERIMENTS

A. *Experiment Procedure*

8 kindergarten boys and girls (L1 to L8) whose age ranges from five to six are participated to the experiment. Pronunciation practice is mainly focused on improvement of pronunciation of vowels and consonants, /s/,/m/,/w/. Before and after pronunciation practice with the proposed system, 12 examiners identify their pronunciations with their voice only and with their moving picture only as well as with their voice and moving picture as shown in Fig.12. Thus identification ratio is evaluated together with mouth and lip shapes difference between before and after pronunciation practices. Pronunciation practice for vowels and consonants are conducted. Those are E1: vowels, E2: /a/-/sa/-/ma/-/ta/-/wa/, E3: /i/-/mi/, E4: /u/-/mu/, E5: /e/-/se/-/me/, E6: /o/-/so/-/mo/, respectively. All these pronunciations are Japanese.

Figure 12 Three types of pronunciation evaluations

B. *Experiment Results*

Table 1 shows experimental results of identification ratio for before and after the pronunciation practice.

TABLE 1 IDENTIFICATION RATIO FOR BEFORE AND AFTER THE PRONUNCIATION PRACTICES

	Moving Picture Only		Voice Only		Moving Picture and Voice	
Exercise	Before	After	Before	After	Before	After
E1	81%	87%	90%	94%	95%	98%
E2	70%	80%	87%	94%	92%	98%
E3	69%	73%	95%	95%	98%	99%
E4	70%	79%	93%	93%	98%	98%
E5	71%	78%	94%	99%	90%	99%
E6	64%	76%	92%	94%	91%	95%
Average	71%	80%	92%	95%	94%	98%

It is noticed that identification ratio for after pronunciation practice is improved by 3-9% from before pronunciation practice for all three cases of evaluation methods. Evaluation results with voice only show 3% improvement. This implies that their pronunciation is certainly improved.

Much specifically, pronunciation of /sa/ for L3 learner before pronunciation practice is 100% perfect while L8 learner has difficulty on pronunciation of /sa/ before pronunciation practice as shown in Table 2.

TABLE 2 IDENTIFICATION RATIO BEFORE PRONUNCIATION PRACTICE.

Learner	a	sa	ta	ma	wa	Unclear
L1	0.0%	87.5%	12.5%	0.0%	0.0%	0.0%
L2	12.5%	87.5%	0.0%	0.0%	0.0%	0.0%
L3	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
L4	0.0%	87.5%	12.5%	0.0%	0.0%	0.0%
L5	25.0%	62.5%	12.5%	0.0%	0.0%	0.0%
L6	0.0%	87.5%	0.0%	0.0%	0.0%	12.5%
L7	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
L8	12.5%	25.0%	37.5%	0.0%	0.0%	12.5%

In particular, pronunciation of /sa/ for L8 learner is used to confuse with /a/ (12.5%), /ta/ (37.5%), and unclear (12.5%) before pronunciation practice as shown in Table 3 (a). This situation is remarkably improved as shown in Table 3 (b). Identification ratio of pronunciation of /sa/ for L8 learner is changed from 25% to 100% perfect after the pronunciation practice.

TABLE 3 IDENTIFICATION RATIOS FOR BEFORE AND AFTER PRONUNCIATION PRACTICE OF PRONUNCIATION OF /SA/ FOR L8 LEARNER

(a) Before pronunciation practice

Sound (Sa)	a	sa	ta	ma	wa	Unclear
L8	12.5%	25.0%	37.5%	0.0%	0.0%	12.5%

(b) After pronunciation practice

Sound (Sa)	a	sa	ta	ma	wa	Unclear
L8	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%

Fig.13 shows one shot frame image of moving picture for mouth and lip when L8 learner pronounces /sa/ before and after the pronunciation practice. Although he could not open his mouth when he pronounces /sa/ before the pronunciation practice, he made almost perfect mouth shape for /sa/ pronunciation after the practice.

Another example for identification ratios for before and after pronunciation practice for L6 learner is shown in Table 4. L6 learner has difficulty on pronunciation of /a/ due to his mouth shape. Although identification ratio of /a/ is 100% before the practice when it is evaluated with voice only, it is 62.5% before the practice when it is evaluated with both voice and moving picture.

Fig.13 Moving picture for mouth and lip of L8 learner before and after the pronunciation practice of /sa/.

This implies that his mouth shape is resembled to that of /sa/, /wa/, and unclear even though his voice sound can be heard as /a/. It is improved to 100% perfect after the practice. Therefore, it may say that the practice is effective to improve not only voice but also mouth and lip shapes.

TABLE 4 IDENTIFICATION RATIOS FOR BEFORE AND AFTER PRONUNCIATION PRACTICE OF PRONUNCIATION /A/ FOR L6 LEARNER

(a) Voice only before practice

Sound (a)	a	sa	ta	ma	wa	Unclear
L6	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%

(b) Voice and moving picture before practice

Sound (a)	a	sa	ta	ma	wa	Unclear
L6	62.5%	12.5%	0.0%	0.0%	12.5%	12.5%

(c) Voice and moving picture after practice

Sound (a)	a	sa	ta	ma	wa	Unclear
L6	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%

His mouth and lip shapes before and after the pronunciation practice of /a/ are shown in Fig.14.

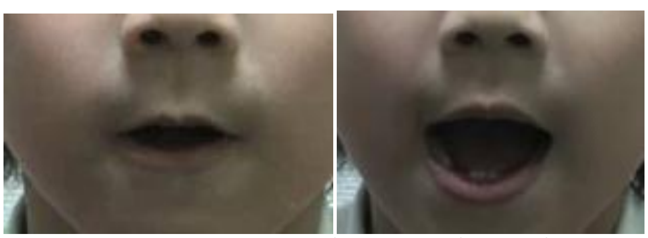


Figure 14 Mouth and lip shapes of L6 learner before and after the pronunciation practice of /a/

Another example is shown in Table 5. L3 learner has difficulty on pronunciation of /e/ due to his mouth and lip shapes. Although identification ratio of /e/ is 75% before the practice when it is evaluated with voice only, it is 62.5% before the practice when it is evaluated with both voice and moving picture. This implies that his mouth and lip shapes are resembled to that of /i/. Although it is improved to 87.5% after the practice when it is evaluated with voice only, it is improved to 100% perfect after the practice when it is evaluated with both of voice and moving picture. Therefore, it may say that the practice is effective to improve not only voice but also mouth and lip shapes.

His mouth and lip shapes before and after the pronunciation practice of /e/ are shown in Fig.15.

TABLE 5 IDENTIFICATION RATIOS FOR BEFORE AND AFTER PRONUNCIATION PRACTICE OF PRONUNCIATION /A/ FOR L6 LEARNER

(a) Voice only before practice

Sound (e)	a	i	u	e	O	Unclear
L3	0.0%	25.0%	0.0%	75.0%	0.0%	0.0%

(b) Voice and moving picture before practice

Sound (e)	a	i	u	e	o	Unclear
L3	0.0%	37.5%	0.0%	62.5%	0.0%	0.0%

(c) Voice only after practice

Sound (e)	a	i	u	e	o	Unclear
L3	0.0%	12.5%	0.0%	87.5%	0.0%	0.0%

(d) Voice and moving picture after practice

Sound (e)	a	i	u	e	o	Unclear
L3	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%

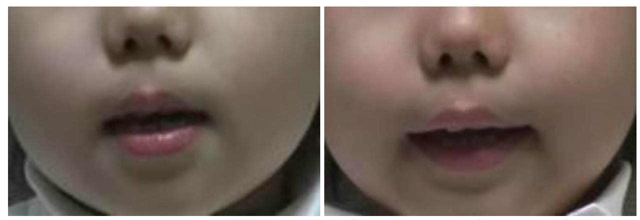


Figure 15 Mouth and lip shapes of L3 learner before and after the pronunciation practice of /e/

IV. CONCLUSION

Pronunciation practice system by means of mouth movement model based personalized CG animation is proposed. The system allows users to look at both users' mouth movement and model based CG animation of moving picture. Therefore, users practice pronunciation by looking at both moving pictures effectively. 8 infants examined pronunciation practices of vowels and consonants, in particular for /s/,/m/,/w/ by using the proposed system. Remarkable improvement (3-9%) on their pronunciations is confirmed.

ACKNOWLEDGMENT

Authors thank to Mr. Isao Narita and Mr. Hiroshi Kono of Faculty of Engineering, Kurume Institute of Technology for their great effort to experiments. Also authors thank to Professor Dr. Seio Oda of Fukuoka Institute of Technology, Junior College for his great supports and comments and suggestions to our research works.

REFERENCES

[1] Mariko Oda, Shun Ichinose, Seio Oda, Development of a Pronunciation Practice CAI System Based on Lip Reading Techniques for Deaf Children, Technical Report of the Institute of Electronics, Information, and Communication Engineers of Japan, vol.107, No.179 WIT2007-25(2007).

[2] Mariko Oda, Seio, Oda, and Kohei Arai, Effectiveness of an English /l/ -/r/ Pronunciation Practice CAI System Based on Lip Reading Techniques, Journal of Japan Society of Educational Technology, 26(2), pp.65－75(2002)

[3] McGurk H, Hearing lips and seeing Voices, Nature, 264, 746-748, 1976.
[4] Silsbee P., Computer lipreading for improved accuracy in automatic speech recognition, IEEE Trans. Speech & Audio Processing 4(5), 337-350, 1996
[5] Reiseberg D., Easy to hear, but hard to understand: A lipreading advantage with intact auditory stimuli, Hearing by Eye, 1987.
[6] Sumby W., Visual contributions to speech intelligibility in noise, J. Acoust. Soc. Amer. 26(2), 212-215, 1954

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science, and Technology of the University of Tokyo from 1974 to 1978 also was with National Space Development Agency of Japan (current JAXA) from 1979 to 1990.

During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post-Doctoral Fellow of National Science and Engineering Research Council of Canada. He was appointed professor at Department of Information Science, Saga University in 1990. He was appointed councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was also appointed councilor of Saga University from 2002 and 2003 followed by an executive councilor of the Remote Sensing Society of Japan for 2003 to 2005. He is an adjunct professor of University of Arizona, USA since 1998. He also was appointed vice chairman of the Commission “A” of ICSU/COSPAR in 2008. He wrote 30 books and published 332 journal papers.

Mariko Oda, She received BS degree in 1992 and MS degree in 1994 from Saga University. She was subsequently appointed assistant professor at Department of Information and Network Technology, Kurume Institute of Technology.

An Application of Intuitionistic Fuzzy in Routing Networks

Ashit Kumar Dutta
Dept. of Computer Science
Alquwaya college of Science, Shaqra University
Kingdom of Saudi Arabia

Abdul Rahaman Wahab Sait
Dept. of Computer Science
Alquwaya Community College, Shaqra University
Kingdom of Saudi Arabia

Abstract— **Routing is an important functional aspect of networks to transport packets from source to destination. A router sets up optimized paths among the different nodes in the network. In this paper the authors proposed a new type of routing algorithm which includes exchange of routing information for small amount of time and then halts for few hours but this information is used by each router to make its own routing decisions based on intuitionistic fuzzy logic during these hours.**

Keywords- component; Intuitionistic fuzzy set; fuzzy set; intuitionistic fuzzy routing; fuzzy routing; GV; IFV; MEO; IMEV;, IFLED.

I. Introduction

The growing usage of computer networks is requiring improvements in network technologies and management techniques so users will receive high quality service. As more individuals transmit data through a computer network, the quality of service received by the users begins to degrade. A more effective method for routing data through a computer network can assist with the new problems being encountered with today's growing networks.

Broadband ISDN (B-ISDN) was envisioned as a provider of higher bit rates to the user than N-ISDN. One of the key design objectives of B-ISDN is that the provision of a wide range of services to a broadband variety of users utilizing a limited set of connection types and multi-purpose user-network interfaces. The two prominent enabling technologies for the deployment of B-ISDN are fiber optics and ATM network architecture. ATM refers to switching and multiplexing techniques. As ATM has to support wide range services whose requirements vary over a wide range, the transport of cells must be at high speed. This calls for minimizing the processing time at the intermediate devices like router and the efficient methods for traffic management.

Increasing traffic loads will naturally lead to network delays, which will lead to other problems as well. These network delays can easily cause dropped sessions or lost data, not to mention dissatisfied users. It is impossible to stem this increasing traffic load. However, optimal routing of messages within a network can mitigate some of the difficulties of heavy traffic. Therefore, a more efficient method of routing needs to be developed to combat network delays.

Routing is an important functional aspect of networks to transport packets in general (or cells in ATM networks) from source to destination. A router sets up optimized paths among the different nodes in the network. An optimized path is that one which gives low mean packet delay and high network throughput. Many routing algorithms exist in the literature. All these can be broadly classified into static and dynamic algorithms. Dynamic routing algorithms make decision regarding the optimized paths independently of other routers based on the information exchanged among the adjacent routers. This exchange of routing information is carried out periodically increasing the traffic on the network.

In [5], we proposed a new type of routing algorithm which includes exchange of routing information for small amount of time and then halts for few hours but this information is used by each router to make its own routing decisions based on intuitionistic fuzzy logic during these hours. The new routing may be called "Intuitionistic Fuzzy Routing". The "fuzzy routing" is a special case of intuitionistic fuzzy routing as usual.

Dynamic routing algorithms exchange routing information periodically among the adjacent routers. This is called periodic updates. This period typically ranges from few tens of milliseconds to 1 or 2 minutes. If the updates are too frequent, congestion may occur. On the other hand, if updates are too infrequent, routing may not be efficient. Hence these dynamic algorithms add extra traffic due to the exchange of routing information among the routers to the network. Traffic due to ever increasing demand of new services is also growing. This is making the traffic (due to routing and user information) management issue a complex one and hence it is becoming a major field of research in present days networks like ATM. Hence the authors are making an attempt to reduce the traffic due to routers by embedding intuitionistic fuzzy intelligence into existing adaptive routers. These intuitionistic fuzzy routers are expected to increase the speed of routing as compared to conventional adaptive routers as frequent exchange of routing information is not required.

II. Existing Routing and Existing Fuzzy Routing

In today's Internet world, information, split into small blocks called packets or cells, is moved across some kind of networks and terminates at distance point. In this process, a data packet passes through a route or path identified by routers. Hence routing is an important issue to communicate among the users on different networks. The selection of optimized path between sender and receiver is a major design area of network layer of ISO's OSI reference model [1]. Many algorithms for routing are available in the literatures which fall in one of two major groups: non-adaptive and adaptive. Non-adaptive (also called static) algorithms do not base their routing decisions on measurements or estimates of the current traffic and topology.

Adaptive (also called dynamic) algorithms, in contrast, change their routing decisions to reflect changes in the topology and traffic. These adaptive algorithms decide the routing path based on the information they get from other routers. The various adaptive algorithms, available in the literature, differ in the way they get information (locally, from adjacent routers or from all routers), when they change the routes (e.g. every t msec, when the load changes, or when the topology changes), and what parameter is used for optimization (e.g. distance, number of hops, or estimated transit time, reliability, bandwidth, load etc). These adaptive routing algorithms are much applicable in the present scenario of changing networks both in size and service requirements.

Dynamic routing algorithms exchange routing information periodically among the adjacent routers. This is called periodic updates. This period typically ranges from few tens of milliseconds to 1 or 2 minutes. If the updates are too frequent, congestion may occur.

On the other hand, if updates are too infrequent, routing may not be efficient. Hence these dynamic algorithms add extra traffic due to the exchange of routing information among the routers to the network. Traffic due to ever increasing demand of new services is also growing. This is making the traffic (due to routing and user information) management issue a complex one and hence it is becoming a major field of research in present days networks like ATM.

Hence the authors are making an attempt to reduce the traffic due to routers by embedding intuitionistic fuzzy intelligence into existing adaptive routers. These intuitionistic fuzzy routers are expected to increase the speed of routing as compared to conventional adaptive routers as frequent exchange of routing information is not required.

Distance Vector Routing algorithm, originally proposed by Bellman[2], Ford and Folkerson[4] , is considered as conventional dynamic routing algorithm as it is being employed widely in today's networks like Routing Information Protocol (RIP) for IP, Cisco's Internet gateway Routing Protocol (IGRP), AppleTalk's Routing Table Maintenance Protocol (RTMP) etc.

In [7], the authors proposed a rule based fuzzy logic to find out the optimum path considering the crisp values of hop count metric. Further, testing and simulation study of the approach is carried out with different traffic loads. It is concluded that the approach decreases the processing overhead and provides a fair distribution of network traffic as compared to other traditional routing techniques.

In [6], intelligent fuzzy approach is proposed for routing the tagged cells. Simulation results showed an improvement in network utilization. The fuzzy routing reported by the authors [8,9] in queuing systems and computer network routing are different from the fuzzy routing proposed by us in this work.

In Distance Vector Routing, each router maintains a routing table containing one entry for each destination in the network. This entry tells the preferred outgoing line to use for that destination. The router knows the "distance" (number of hops, queue length or delay) to each of its neighbors. For example, consider that delay is used as a metric and assume that the router knows the delay to each of its neighbors. Once every t msec, each router sends to each neighbor a list of its estimated delays to each destination. It also receives a similar list from each neighbor.

Based on this information, a router can find out which estimate seems the best and updates its routing table. This routing table will be used by the router to route the packets for next T msec (T >> t), after which routing information will be exchanged again and this procedure is repeated. Thus for every t msec, routing information will be exchanged among the adjacent routers which leads to increased traffic on the network. But the advantage of the algorithm is that it updates routing information dynamically for every fixed time interval. In this paper, we propose a method to reduce the traffic due to exchange of routing information in Distance Vector Routing algorithm retaining its advantage mentioned above.

Our method is mainly based on fuzzy mathematics to deal with the uncertainties in traffic handling. We introduce some new terminologies in section-3 and then in section-4 we explain our method of fuzzy routing. The work reported here covers the theoretical aspects of fuzzy routing whereas implementation and simulations will be our future research plan.

III. Intuitionistic Most Expected Vector (IMEV)

In this section, we present some preliminaries on Intuitionistic Most Expected Vector. This notion of IMEV is used by us in [5] on intuitionistic fuzzy routing technique.

A. *Generated Vector (GV)*

Consider k number of n-dimensional vectors V1, V2, ... , Vk

$$V_i = \begin{bmatrix} v_{i1} \\ v_{i2} \\ . \\ . \\ v_{in} \end{bmatrix}, i = 1,2,3,\dots,k. \quad ,$$

where

For each j (j = 1, 2, . ., n), we will do here n number of extrapolations by using Newton's Backward Interpolation formula, for the n tables given by

1	2	3	4	. . . .	k
v_{1j}	v_{2j}	v_{3j}	v_{4j}	. . . .	v_{kj}

to calculate vk+1, j , j = 1, 2, 3, ……., n.

Thus a new vector is generated, which is

$$\overline{v_{k+1}} = \begin{bmatrix} \overline{v_{k+1,1}} \\ \overline{v_{k+1,2}} \\ --- \\ --- \\ --- \\ \overline{v_{k+1,n}} \end{bmatrix}$$

This vector we may call as a Generated Vector (GV) by V1, V2, …., Vk.

B. Intuitionistic Fuzzy Vector (IFV)

By a Intuitionistic Fuzzy Vector (IFV) $\tilde{V}_{k+1}$, we mean

$$\tilde{V}_{k+1} = \begin{bmatrix} \tilde{v}_{k+1,1} \\ \tilde{v}_{k+1,2} \\ --- \\ --- \\ --- \\ \tilde{v}_{k+1,n} \end{bmatrix}$$

where each of $\tilde{v}_{k+1,1}$, $\tilde{v}_{k+1,2}$, ... , and $\tilde{v}_{k+1,n}$ is a fuzzy number corresponding to the precise numbers $\overline{v_{k+1,1}}$, $\overline{v_{k+1,2}}$, . . . , and $\overline{v_{k+1,n}}$ respectively. We call it an IFV generated by the vectors V1, V2, ..., Vk and denote it by the notation.

IFV(V1, V2, ...,Vk) = V_{k+1}.

C. Most Expected Object (MEO)

Suppose that A is an ifs of a set X with membership function μA and non-membership function υA. By the term 'Most Expected Object (MEO)', we mean that element of X for which the membership value is maximum.

We denote it by MEO(A).

Thus, MEO(A) = xq ,

where xq ∈ X and

θ(xq) = max { θ(xi) : xi ε X },

where

$$\begin{cases} \theta(x_i) = \dfrac{\mu_A(x_i)}{\upsilon_A(x_i)}, & if \quad \upsilon_A(x_i) \geq \pi_A(x_i) \\ \quad\quad\;\; = \dfrac{\mu_A(x_i)}{\pi_A(x_i)}, & otherwise \end{cases}$$

D. Intuitionistic Most Expected Vector (IMEV)

Suppose that IFV (V1, V2, ..., Vk) = $\tilde{V}_{k+1}$, where

$$V_{k+1} = \begin{bmatrix} \tilde{v}_{k+1,1} \\ \tilde{v}_{k+1,2} \\ --- \\ --- \\ --- \\ --- \\ \tilde{v}_{k+1,n} \end{bmatrix}$$

The Most Expected Vector (IMEV) of $\tilde{V}_{k+1}$ is the vector

$$V_{k+1} = \begin{bmatrix} MEO(\tilde{v}_{k+1,1}) \\ MEO(\tilde{v}_{k+1,2}) \\ ------ \\ ------ \\ ------ \\ MEO(\tilde{v}_{k+1,n}) \end{bmatrix} = \begin{bmatrix} v_{k+1,1} \\ v_{k+1,1} \\ --- \\ --- \\ --- \\ v_{k+1,n} \end{bmatrix}, \quad \text{(say)}$$

which we denote by IMEV (V1, V2, ..., Vk).

IV. Intuitionistic Fuzzy Routing [5]

In [5], we presented our proposal for intuitionistic fuzzy routing. This is an application of Atanassov's[2,3] theory of IFS in routing. Suppose that the size of the network is r (the number of routers or nodes). We assume that all the routers are active and the metric is measured by delay.

Consider two choice parameters n and T (to be precisely understood later on).

Method:

Once every t msec, each router sends to each of its neighbor a "list of estimated delay" (LED) to reach the every router of the network. On sending n number of such lists, it halts for T hours.

After T hours, the router again sends a fresh set of LED and so on. Clearly each LED is a r-dimensional vector. Our work in this paper presents a method for generating a new type of LED which we call here Intuitionistic Fuzzy LED (IFLED).

For the sake of presentation of our method, let us consider the subnet as shown in Fig.1

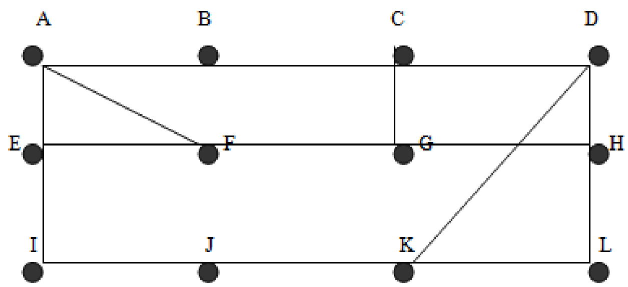


Figure 1: A Subnet

With no loss of generality consider any router, say F. It has three neighbors A, E and G. Therefore it can receive LED from these three neighbors only. Suppose that at some instant τ, and then at the regular instants (τ + t), (τ + 2t), ..., (τ + (n-1)t), the router F receives the following sets of three LEDs from A, E, and G respectively.

LED information received by the router F

At the instant	LED received from A	LED received from E	LED received from G
τ	V_0^A	V_0^E	V_0^G
τ+ t	V_1^A	V_1^E	V_1^G
τ+ 2t	V_2^A	V_2^E	V_2^G
…	…	…	…
τ+ (n-1)t	V_{n-1}^A	V_{n-1}^E	V_{n-1}^G

Each entry vector in the above table is r-dimensional (in this example r = 12).

$$Forexample(hypothetic\,al)V_3^C = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_r \end{bmatrix},$$

then it means that the optimal delay from C to the router R1 is x1, the optimal delay from C to the router R2 is x2, etc.

Now let us find the following three vectors:

(1) $IMEV\left(V_0^A, V_1^A, \ldots\ldots, V_{n-1}^A\right) = V_n^A$,(say)

(2) $IMEV\left(V_0^E, V_1^E, \ldots\ldots, V_{n-1}^E\right) = V_n^E$, (say)

(3) $IMEV\left(V_0^G, V_1^G, \ldots\ldots, V_{n-1}^G\right) = V_n^G$, (say).

The collection of these three vectors is called Intuitionistic Fuzzy LED (IFLED). During the computation of these three vectors, the router F will be functioning according to the previous IFLED. The new IFLED now will remain valid to the router F for T hours next, until the next IFLED is computed. With the help of this IFLED, routing will be done by F. Once an IFLED is computed, the previous IFLED gets replaced.

V. Conclusion

The work done in [5] is an application of Atanassov's Intuitionistic fuzzy set theory [2, 3] in network. We have proposed a new method of routing called intuitionistic fuzzy routing. We have considered Distance Vector Routing with intuitionistic fuzzy tools.

Consider the traffic situation on the roads of a city. During a time interval of T hours (say 30 minutes), the mean traffic on a road could be assumed to be almost equal. During the next 30 minutes, again it is almost equal with another mean value of traffic. This is the philosophy we have used in [5]. The intuitionistic fuzziness in the value of almost equal mean of traffic is to be dealt with proper membership function and non-membership function.

The method proposed is both dynamic as well as static. The method does not need continuous resource management traffic, but it works on updated information (updated every T hours). Thus a gain in the reduction of resource management traffic has been made possible. It is claimed that the method will improve the overall performance of the network.

The disadvantage of the method is that in case a catastrophic traffic of packets arises in any path, the method may not yield better performance.

The work for the construction of membership and non-membership functions for generating IFLED in intuitionistic fuzzy routing needs a rigorous study and survey, which is under our next research plan.

If fortunately, all the indeterministic functions are null functions (i.e. if there is no hesitation), intuitionistic fuzziness reduces to fuzziness and then the intuitionistic fuzzy routing will be reduced to 'fuzzy routing' as a special case[5]. The fuzzy routing reported by the authors [8,9,10] in queuing systems and computer network routing are different from the fuzzy routing proposed by us.

References

[1] A. S. Tanenbaum., Computer Networks, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[2] Atanassov, K., Intuitionistic Fuzzy Sets, Fuzzy Sets and Systems. 20(1986) 87-96.

[3] Atanassov, K., Intuitionistic Fuzzy Sets : Theory and Applications, Physica-Verlag, Heidelberg (1999).

[4] Bellman,R.E., Dynamic Programming, Princeton, NJ; Princeton University Press, 1957

[5] D. Driankov, H. Hellendoorn and M. Reinfrank, An Introduction to Fuzzy Control. Springer-Verlag, Berlin, New York, 1993.

[6] Ford,L.R. and Fulkerson,D.R., Flows in Networks, Princeton,NJ; Princeton University Press, 1962.

[7] Ka-Wing Wong, "Fuzzy routing control of service request messages in an individual computing environment", Proc. of the 1995 ACMSymposium on Applied Computing, Nashville (USA), 1995, pp 548-551.

[8] Runtong Zhang and Yannis A. Phillis, "Fuzzy Routing of Queueing Systems with Heterogeneous Servers", Proc. of IEEE Intl. Conference on Robotics and Automation Albuquerque, New Mexico, April 1997, pp. 2340-2345.

[9] Runtong Zhang and Yannis A. Phillis, "Fuzzy Service Rate Control of Queueing Systems", Proc. 1996 IEEE Intl. Conf. On Systems, Man and Cybernetics on Information, Intelligence and Systems (IEEE/SMC'96), October 14-17, 1996, Beijing, China.

[10] Sridhar Pithani and Adarshpal S. Sethi, "A fuzzy Set Delay Representation for Computer Network Routing Algorithms" Proc. Second Intl. Symposium on Uncertainty Modeling and Analysis, College Park, MD (April 1993), pp. 286-293.

[11] Zadeh, L. A., Fuzzy Sets, Information and Control. 8(1965) 338-353.

[12] Gasim Alandjani and Eric. E.Johnson,"Fuzzy routing in Ad hoc networks", IEEE 2003.

[13] Barolli L, Koyama A, Yamada T, and Yokoyama S, "An Intelligent fuzzy routing scheme for improving ATM Network Performace using Violation tagging function", IEEE Proceedings on Database and expert system application, Iss: 2000, pp. 5 -9.

[14] Dr.Shuchitha Upadhayay and Mini Sharma, " Reinforcement of a new fuzzy mixed metric approach through fuzzy routing Algorithms", IJCNS, Vol.8,No.2, Feb 2008, pp 271 – 276.

An SMS-SQL based On-board system to manage and query a database

Ahmed Sbaa
National School of Electricity and Mechanics, ENSEM
University Hassan II – Casablanca, Morocco

Rachid El Bejjet
National School of Electricity and Mechanics, ENSEM
University Hassan II- Casablanca, Morocco

Hicham Medromi
National School of Electricity and Mechanics
University Hassan II- Casablanca, Morocco

***Abstract*— Technological advances of recent years have facilitated the use of embedded systems. They are part of our everyday life. Thanks to them, electronic devices are increasingly present in our lives in many forms: Mobile phones, music players and managers have become the essential of modern life. Access to information anywhere at any time is increasingly a daily challenge of embedded system technology. Following an innovative idea, this paper describes an embedded system that can query any database through SMS commands to extend the consultation of data to mobile networks early generations. Based on a UNIX embedded system, the result of this work can serve as a standard consultation of databases through SQL-SMS Gateway which converts an SMS command in an SQL query. This system will open the database to the consultation via mobile without having to expose them to risks of online publication. While in the first part of this article we will discuss the state of the art of multi-agent systems and input systems onboard, the second part presents the architecture of our target system. In the third part we describe in detail the realized prototype. This article ends with a conclusion and an outlook.**

Keywords- SMS; SQL; UNIX; LINUX; Embedded System; Perl; GSM.

I. INTRODUCTION

Most of applications management, research and data consultation are using databases. However, it is not always easy to undertake a remote research because of the difficulty of generalizing access to all networks, by opening Internet risk which often requires huge amount of security investment or sometimes by the complexity of the publishing solutions implementation. Whether you're not connected to internet or do not have security platforms for publishing your databases, we propose through this paper an innovative solution based on a Unix embedded system providing a searchable database only via SMS.

This system opens the way for the standardization of a new exchange protocol which we call SMS-SQL that will find any database through an SMS command. In the first part of this article we will discuss the state of the art about embedded systems and multi agent systems used for our system modeling. In the second part we will present the architecture of the targeted system. The third part is devoted to the description of prototype as part of our research. This article ends with a conclusion and perspectives.

II. STATE OF THE ART

A. Multi-Agent System and Types of Agents

A multi-agent system (MAS) is a system composed of multiple interacting intelligent agents within an environment. Multi-agent systems can be used to solve problems that are difficult or impossible for an individual agent or a monolithic system to solve. Intelligence may include some methodic, functional, procedural or algorithmic search, find and processing approach.

Multi-agent systems consist of agents and their environment. Typically multi-agent systems research refers to software agents. However, the agents in a multi-agent system could equally well be robots, humans or human teams. A multi-agent system may contain combined human-agent teams.

Agents can be divided into different types:

Very simple like: passive agents or agent without goals (like obstacle, apple or key in any simple simulation); Active agents with simple goals (like birds in flocking, or wolf-sheep in prey-predator model) Or very complex agents (like cognitive agent, which has a lot of complex calculations).

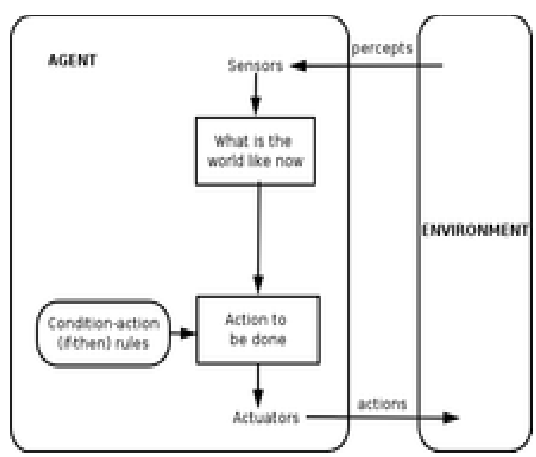


Figure 1. General reflex agent

B. *Characteristics of Agents*

The agents in a multi-agent system have several important characteristics:

Autonomy: the agents are at least partially autonomous

Local views: no agent has a full global view of the system, or the system is too complex for an agent to make practical use of such knowledge

Decentralization: there is no designated controlling agent (or the system is effectively reduced to a monolithic system).

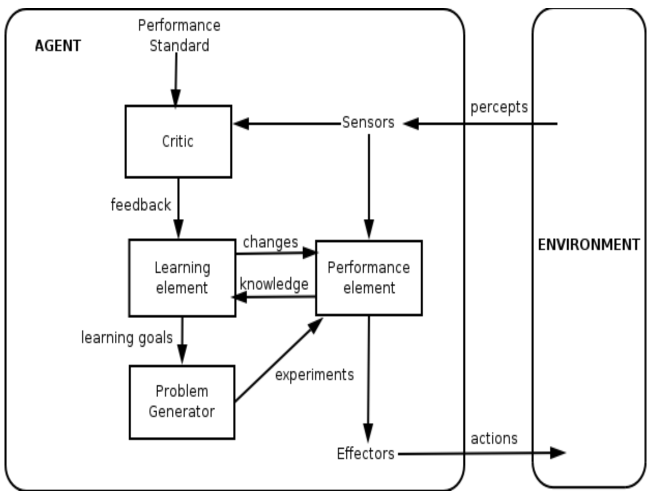


Figure 2. Learning Agent

C. *Embedded systems*

Embedded systems are devices or software and hardware components are intimately linked. The interest to embedded systems is increasingly evolving through the routine means of communication (Smartphone, switchboards etc...) and the ongoing need for facilitation of life through technology (MP3 players, storage data). In industry, embedded systems are often an indispensable choice for reasons of criticality of the features offered and the security risks (ABS brakes, alarm system and detection, aviation, etc ...).

For specific jobs, embedded systems bring several advantages to traditional systems by contribution based on normal computers. The main constraints to meet embedded systems are:

- The stability system: an embedded system is often dedicated to a specific operation, malfunctions must be mastered.
- Mastery of the security, integrity and access: with a minimum of features and services enabled, an embedded system is designed to be safer.
- Cost of production: a computer system can be produced through industrial processes that greatly limit the cost of production.
- Low power consumption: Unlike a conventional computer, an embedded system has the minimum resources tailored. Energy consumption is adjusted and optimized.
- Reactivated: the response time required in an embedded system often require real-time systems.
- Autonomy: embedded systems must operate without human intervention to perform automatic spots.

The operating system component is an essential building block in the design of any embedded system. Thanks to their major assets in terms of reliability, security, stability and effective resource management, it is not surprising to find unix and linux systems in most embedded systems. The Linux and Unix systems support multiple hardware platforms: ARM, X86, MIPS, and POWER PC. Hence the birth of several distributions for embedded systems. As examples, we find: Lineo Embedix, MontaVista, uCLinux and LEM.

III. DESCRIPTION OF THE TARGET SYSTEM

Our target system is designed to address the need to query a database via SMS queries. With a standard GSM, a user can formulate a specific syntax as a query and send SMS to a Gateway (SMS-Sql). The latter charge is to extract the query and translate it into SQL query which is sent to a database server. The answer to the user query will also be sent back by SMS.

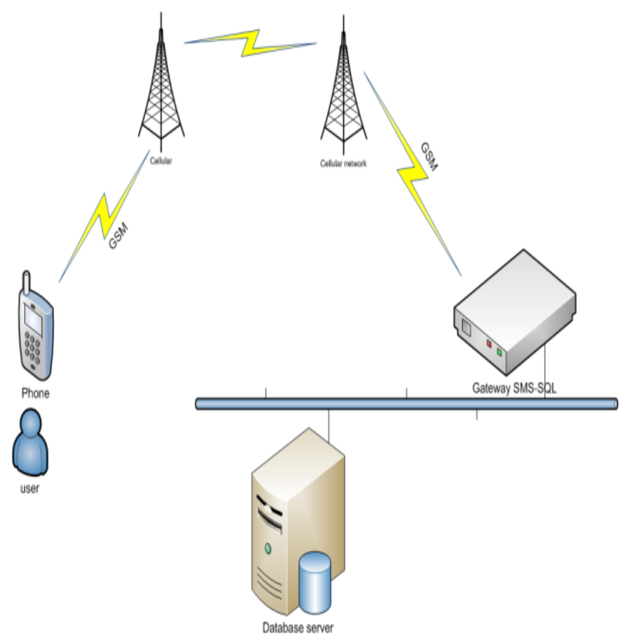


Figure 3. Operating diagram

The target system has the following features:

- Conversion of a text message to a SQL query
- Supports almost any database through the use of the correct driver
- Quick Setup "plug and play" on the network
- User authentication and numbers that sending SMS messages to maintain data security.

IV. SYSTEM MODELING AND ARCHITECTURE

A. *Multi-agent systems based modeling*

To better understand the details of how the target system, we modeled the interactions between different components using multi-agent systems:

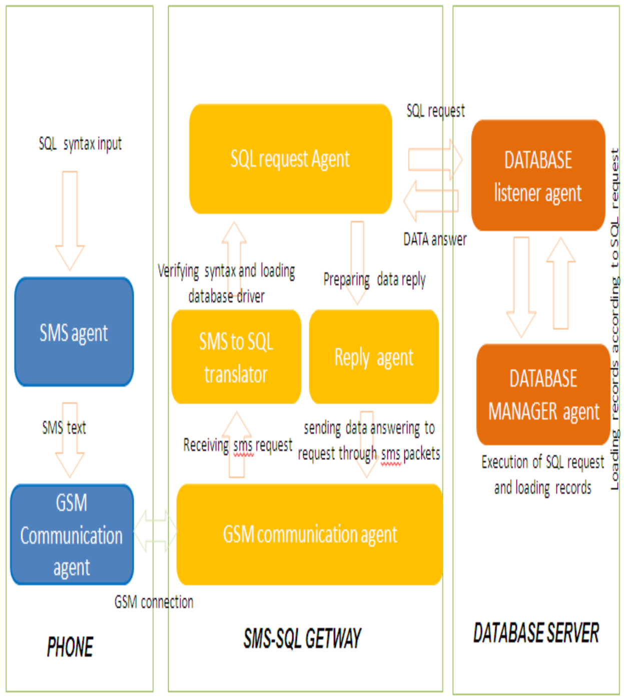


Figure 4. Diagram of System modeling

B. Target System Architecture:

The SQL-SMS Gateway is an embedded system where the OS is based on a UNIX kernel with device drivers and LAN modems. In the upper layers there are libraries for access to the GSM device for sending and receiving SMS queries. There are also drivers for querying various databases to consult.

This modular architecture provides better stability of the embedded system enjoying the benefits of the UNIX kernel and cascade functioning to isolate each layer separately to ensure safe operation of the system. In case of problems on a layer, it can be reloaded or reset without impacting the layers above or below.

The UNIX kernel is the most important component of the system. Its purpose is to control the material in a logical manner and provides services to low-level users to high level. It basically manages the control of network devices and memory to avoid system crashes and maximize resources. To avoid unexpected malfunctions of the system, applications do not access directly to GSM pilots. To do so, they must tease out specific libraries (libraries GSM). This will save us breakdowns due to unexpected or random workings of the system.

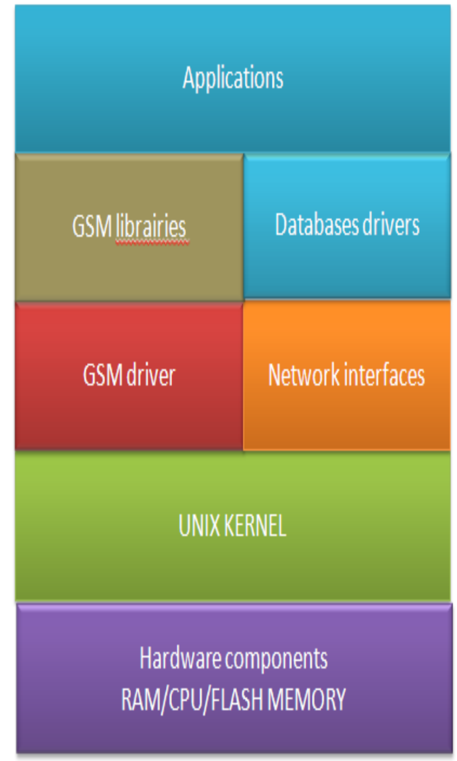


Figure 5. Diagram of System components

C. Scheme of procedures systems execution

The execution of tasks systems is depicted by the figure 6. Through a trigger, the system periodically consults the GSM interface to retrieve messages that will be inserted into a file called "receiving queue". SMS queries are translated into a SQL Query and sent to the server database. Answers to communicate to senders are stored in a queue called "sending queue."

This mode of operation can offer three major advantages to embedded system:

- Optimization of energy: indeed, if we maintain an ongoing connection to the GSM modem, you may lose energy in the initial connection to the cellular network.
- The use of queues helps maintain data integrity and retake the treatment in case of system malfunction.
- It is possible to implement prioritization in servicing the request based on the running time of each request at the level of criticality of the application.

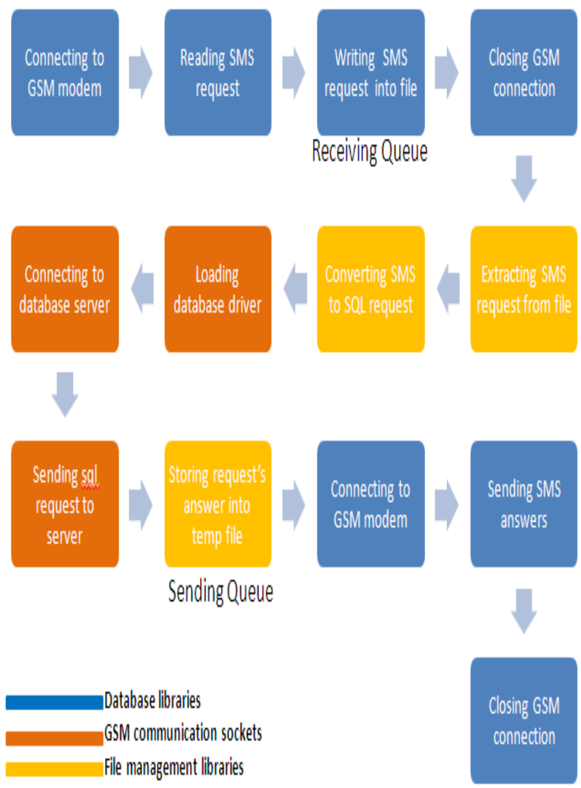


Figure 6. Diagram of Procedures executions

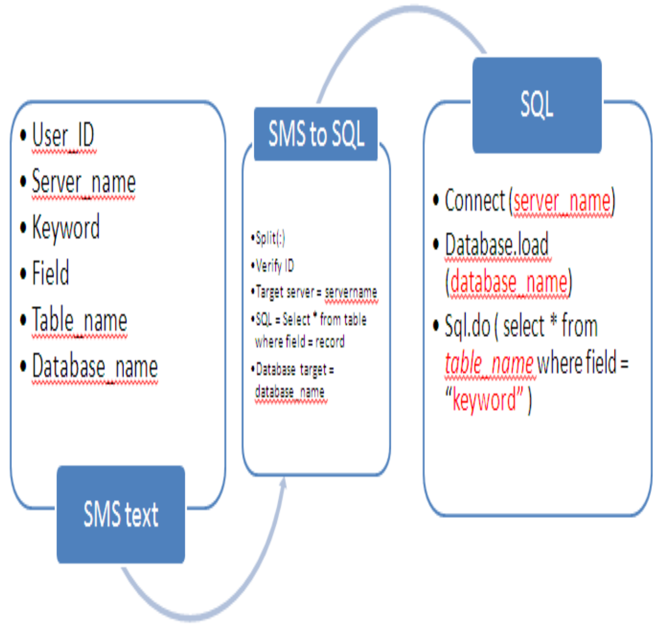


Figure 7. Diagram of the SMS-SQL Translator

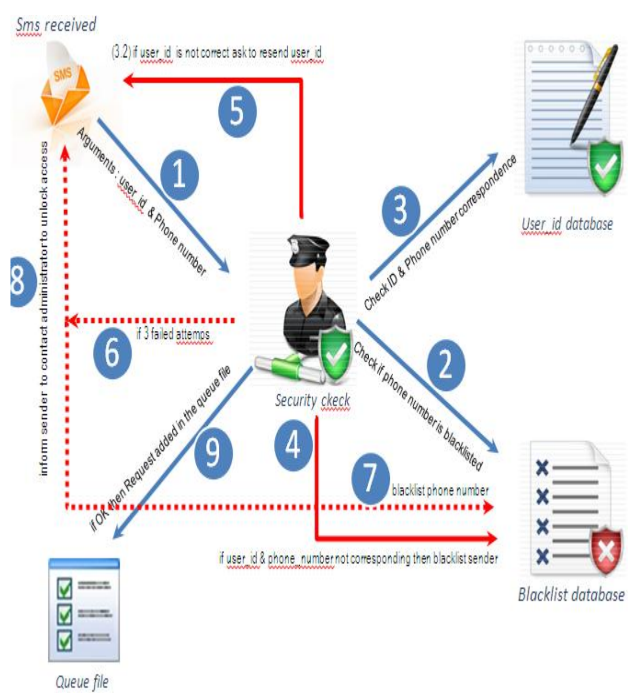


Figure 8. Diagram of security mechanism

D. SMS-SQL Translation

The user sends SMS containing arguments to be used for authentication, identification of the target server and the formalization of the SQL query. A SMS request has the following syntax:

USER_ID: Server_name: keyword: Field: Table_name: Database_name

The field extraction is expressed by the separator ':'.

This syntax requires that the ID is rather short and that keywords do not contain special characters.

E. Security measures

An embedded system is intended to perform secure spots. It is therefore necessary to implement safety rules to avoid major risks. Indeed, such a system is exposed to the following threats:

- Using a false identity to reveal confidential information (ID-spoofing)
- Attempts to access by brute force

The following diagram depicts the security mechanisms implemented to mitigate these threats.

V. DESCRIPTION OF THE PROTOTYPE

To implement the underlined concept, we chose several technological options for the prototype.

A. *Operating system²*

As noted in our target system architecture, the different layers of the system must be autonomous. This is a fundamental criterion for the choice of operating system. Our benchmark of Unix core revealed the stable and robust feature of the system. Indeed, NETBSD fits the following advantages over other OS:

- Portability to most hardware architectures

- Accuracy and quality codes

- Policy rise version very strict and demanding in terms of security and stability

- The kernel functions are highly optimized and therefore allow for better utilization and energy consumption.

GSM Modem:

We used a USB external modem GSM. In an industrial system, the majority of devices are typically integrated into the same card.

B. *Case of shipment*

We opted for a box Soekris NET 5501 which features the majority of interfaces and slots necessary. Its main features are:

- 433 to 500 Mhz AMD Geode LX single chip processor with CS5536 companion chip
- 512 Mbyte DDR-SDRAM soldered on board
- 4 Mbit BIOS/BOOT Flash
- Compact FLASH Type I/II socket
- USB 2.0 interface, one internal, one external port
- 10/100 Mbit Auto MDIX Ethernet ports, RJ-45, protected to 700W/40A Surge
- Power using external power supply is 6-25V DC, max 20 Watt, protected with TVS

Industrial embedded system industry can be significantly optimized until NANO size. Its size will rely primarily on the amount of requests to be processed and the nature of databases to query.

C. *Development libraries:*

To benefit from greater flexibility on the system we chose to use PERL to Unix native langage. This choice also offers several advantages:

- Rapid development of various systems common tasks
- Drivers database available for any DBMS (MYSQL, SQL SERVER ORACLE, etc. ...).

- GSM librairies book flexible and efficient.
- Better debugging system. In diagnosing system problems, it is critical to know at what level is the error. This will develop a circumvention procedure to make the system more reliable.

Figure 9. Soekris 5501 box

The following example shows a procedure that bypasses the failure of a GSM modem interface through the use of an emergency GSM interface.

```
#!/usr/bin/perl -w

use strict;
use Device::Gsm;

# GSM interface user for connecion with
Operator 1.
my $myPORT1='/dev/ttyHS1';

# GSM interface user for connecion with
Operator 2.
my $myPORT2='/dev/ttyHS1';

my $gsm = new Device::Gsm( port =>
$myPORT1);

if( $gsm->connect() ) {
# print "connected !\n";
} else {
print "interface  $myPORT is down \n";

my $gsm = new Device::Gsm( port =>
$myPORT2);

if( $gsm->connect() ) {
# print "connected !\n";
  } else {

print "interface  $myPORT is down \n";
exit 1;
}
```

In practice we found several deadlock situations for which we have developed workarounds:

- No GSM signal
- SIM card lock

- Saturation of the GSM modem
- Format sms modified by the operator
- Limiting the size of the SMS to 160 characters Etc ....

D. Database:

Needing a server database query we opted for a MySQL database hosted on a server UBUNTU. The following diagram summarizes all the components of the prototype:

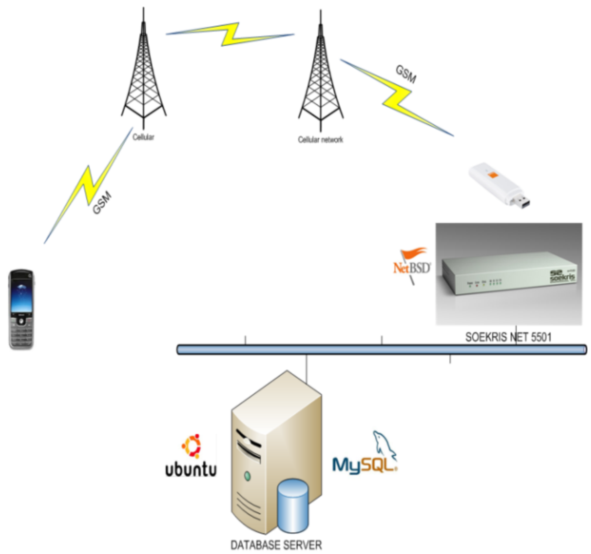


Figure 10. Diagram of the realized protocol

VI. CONCLUSION AND FURTHER WORK

Despite the advent of new communication protocols 3G, SMS remains a reliable protocol and accessible to a wider population. That's where the interest to extend its use to other areas. In this paper we designed an innovative embedded system for translating an SMS query to SQL one. Thanks to the prototype, our idea seems to offer more opportunities in the future. Indeed, this concept opens the way towards a standardization of a new protocol that allows making available a database research via SMS. The proposed embedded system can be appropriately used in several areas where industrial databases are often hindered by the complexity of the architecture.

REFERENCES

[1] M. Al Rababaa, « Hash Algorithms for Security on GSM System", IRECOS Journal, November 2009, Vol. 4. n.6, pp. 698-7039.

[2] A. Sbaa, R. El Bejjet, H. Medromi, "Design and implementation of an embedded system for administering a UNIX system through SMS commands (SMS-SHELL)", IRECOS Journal, July 2011, Vol.x, n.x.

[3] Alessandro Rubini et Jonathan Corbet, Linux Device Drivers; 3ème edition ; O'Reilly 2005.

[4] Karim Yaghmour, Building Embedded Linux Systems; O'Reilly 2003.

[5] Craigh Hollabaugh, Embedded Linux; Sams 2002

[6] "Smartphone definition from PC Magazine Encyclopedia". PC Magazine. http://www.pcmag.com/encyclopedia_term/0,2542,t=Smmrtphone&i=51537,00.asp . Retrieved 13 May 2010.

[7] Drogoul A., De la simulation multi-agent à la résolution collective de problèmes, Thèse de doctorat, Université Paris 6, 1993.

[8] Guessoum Z., Un environnement opérationnel de conception et de réalisation de systèmes multi-agents, Thèse de doctorat, Université Paris 6, mai 1996.

[9] Carabelea C., Boissier O., Florea A., «Autonomie dans les systèmes multiagents », JFSMA '03, 2003.

[10] J. Ferber: Les systèmes multi-agents, vers une intelligence collective, Paris, InterEditions, 1995.

[11] R. El Bejjet, H. Medromi, « A Generic Platform for a Multi-Agent Systems Simulation », IRECOS Journal, September 2010, Vol. 5. n. 5, pp. 505-509.

[12] Dave Pfaltzgraff, The Use of Linux in an Embedded System, December 1999,*Linux Journal*. Putting together a Card Access System.http://www.linuxjournal.com/lj-issues/issue68/3555.html

[13] Sam Williams, Open Season: Embedded Linux Comes to the Fore, August 17, 2000, *Upside Today*. The San Jose Linux World Conference and Expo activity is impressive. http://www.upside.com/Open_Season/399c565b0.html

[14] Stan Runyon, Linux Catches Eyes in Test World, January 4 2000, *EE Times*. National Instruments implement modular automated equipment with Linux. http://www.eet.com/story/OEG20000104S0027

[15] Doc Searls, The Next Bang: The Explosive Combination of Embedded Linux, XML and Instant Messaging", , September 2000, *Linux Journal*, http://www.linuxjournal.com/lj-issues/issue77/4195.html

[16] Embedded Systems Programming Magazine : http://www.embedded.com/mag.htm

[17] D. Kalinsky, R. Kalinsky ; « Introduction to I2C », Embedded.com. 2001. http://embedded.com/story/OEG20010718S0073

[18] D. Kalinsky, R. Kalinsky ; « Introduction to Serial Peripheral Interface », Embedded.com. 2002.

http://embedded.com/showArticle.jhtml?articleID=9900483

[19] M. Khemakhem, A. Belghith, « Agent Based Architecture for Parallel and Distributed Complex Information Processing », January 2007, Vol. 2. n. 1, pp. 38 – 44

AUTHORS PROFILE

[1]**A. Sbaa** is an engineer in computer science since 2002, from EMI an engineering school in Rabat, Morocco. In 2009 he got a Certificate of Information System Auditor (CISA). He joined the system architecture team of the ENSEM, Casablanca, Morocco.

As a Ph.D. Student since 2009, his current and previous research interest is about intelligent systems architecture based on Multi Agents Systems especially for mobile communication.

[2]**Rachid El Bejjet** is an engineer in computer science since 1995, from ENIM an engineering school in Rabat, Morocco. In 2009 he got his Master Business of Administration (MBA) in ENPC, Paris, France. In 2009 he joined the system architecture team of the ENSEM, Casablanca, Morocco. His actual main research interest is about platforms of Multi agents Systems' simulation and implementation.

[3]**Hicham Medromi** received the PhD in engineering science from the Sophia Antipolis University in 1996, Nice, France. He is a professor and responsible of the system architecture team of the ENSEM Hassan II University, Casablanca, Morocco. His actual main research interest concern Control Architecture of Mobile Systems Based on Multi Agents Systems. Since 2010 he is the Director of the ENSEM, Hassan II University, Casablanca, Morocco.