IJACSA

WHERE WISDOM SHARES

International Journal of Advanced Computer Science and Applications

Volume 2 Issue 4

April 2011

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

www.ijacsa.thesai.org

IJACSA Editorial

From the Desk of Managing Editor…

It is a pleasure to present our readers with the April 2011 Issue of International Journal of Advanced Computer Science and Applications (IJACSA).

With monthly feature peer-reviewed articles and technical contributions, the Journal's content is dynamic, innovative, thought-provoking and directly beneficial to the readers in their work.

The number of submissions have increased dramatically over the last issues. Our ability to accommodate this growth is due in large part to the terrific work of our Editorial Board.

Some of the papers have an introductory character, some of them access highly desired extensions for a particular method, and some of them even introduce completely new approaches to computer science research in a very efficient manner. This diversity was strongly desired and should contribute to evoke a picture of this field at large. As a consequence only 29% of the received articles have been finally accepted for publication.

With respect to all the contributions, we are happy to have assembled researchers whose names are linked to the particular manuscript they are discussing. Therefore, this issue may not just be used by the reader to get an introduction to the methods but also to the people behind that have been pivotal in the promotion of the respective research.

By having in mind such future issues, we hope to establish a regular outlet for contributions and new findings in the field of Computer science and applications. Therefore, IJACSA in general, could serve as a reliable resource for everybody loosely or tightly attached to this field of science.

And if only a single young researcher is inspired by this issue to contribute in the future to solve some of the problems sketched here or contribute to exiting methodologies and research work, the effort of all the contributors will be rewarded. In that sense we would like to thank all the authors and reviewers that contributed to this issue for their efforts and their collaboration in this project.

We hope to continue exploring the always diverse and often astonishing fields in Advanced Computer Science and Applications.

Thank You for Sharing Wisdom!

Managing Editor
IJACSA
Volume 2 Issue 4, April 2011
editorijacsa@thesai.org
ISSN 2156-5570 (Online)
ISSN 2158-107X (Print)

Editorial Board

IJACSA Reviewer Board

CONTENTS

Anthropomorphic User Interface Feedback in a Sewing Context and Affordances

Dr Pietro Murano
Computing, Science and Engineering
University of Salford
Salford, Gt Manchester, UK, M5 4WT

Tanvi Sethi
Computing, Science and Engineering
University of Salford
Salford, Gt Manchester, UK, M5 4WT

***Abstract*— The aim of the authors' research is to gain better insights into the effectiveness and user satisfaction of anthropomorphism at the user interface. Therefore, this paper presents a between users experiment and the results in the context of anthropomorphism at the user interface and the giving of instruction for learning sewing stitches. Two experimental conditions were used, where the information for learning sewing stitches was the same. However the manner of presentation was varied. Therefore one condition was anthropomorphic and the other was non-anthropomorphic. Also the work is closely linked with Hartson's theory of affordances applied to user interfaces. The results suggest that facilitation of the affordances in an anthropomorphic user interface lead to statistically significant results in terms of effectiveness and user satisfaction in the sewing context. Further some violation of the affordances leads to an interface being less usable in terms of effectiveness and user satisfaction.**

Keywords- anthropomorphism; affordances; user interface evaluation.

I. INTRODUCTION

The research area of the main author of this paper has been aiming to discover if anthropomorphism at the user interface improves interaction in terms of effectiveness and user satisfaction. While several other researchers have also been investigating similar issues for a number of years, overall there is still a lack of clear evidence to answer the basic question of whether anthropomorphism is effective and preferred by users. In previous studies by the main author (e.g. [11, 12]), it has been suggested that a lack of facilitation of the affordances at the user interface could be a primary reason for an interface not being effective and that the presence of anthropomorphism is more secondary in nature compared to the affordances being appropriately facilitated. These arguments were based on the observations of previous experiments on various prototypes and the data collected.

While it is acknowledged that context etc can play a role, the authors would argue that if the matter was so 'simply explained', the results of researchers across the world would be more closely aligned. However, this is not the case, the results overall tend to not follow a clear pattern. In some cases anthropomorphism appears to be better and in some cases the opposite seems true.

Anthropomorphism at the user interface typically involves some part of the user interface, taking on some human quality [5]. Some examples include a synthetic character acting as an assistant or a video clip of a human [2].

Therefore, this paper presents the next important stage of the authors' work. This stage has involved developing one anthropomorphic user interface that facilitates the affordances and a contrasting (but identical in content) non-anthropomorphic user interface that deliberately violates in a subtle manner the affordances (a future stage aims to try the opposite, i.e. a non-anthropomorphic interface that facilitates the affordances and an anthropomorphic interface that violates the affordances). This differs significantly from the main author's previous work, because the previous software prototypes developed and tested were not specifically designed to facilitate/violate the affordances. However their aim was to evaluate anthropomorphic feedback. Furthermore this research is novel in that as far as we know, no one has linked the affordances with anthropomorphic feedback and more practical hands-on activities such as sewing.

The next section will present some key related literature. Following the brief literature review, the details of the experiment will be presented along with the main statistically significant results. Then the paper will conclude with a discussion in terms of the results and the affordances.

II. KEY LITERATURE

One of the purposes of this section is to inform the reader about some of the current work carried out in connection with anthropomorphic feedback and to show that there is sometimes disagreement in the results of whether such feedback is more effective and preferred by users.

David, Lu, Kline and Cai [4] report the details of a study involving three experimental conditions in the context of a quiz about ancient history. The authors were looking into different anthropomorphic cues. These were the gender of a character and attitude and user perceptions about the character connected to quiz success. The overall results of their experiment indicated that anthropomorphic cues led to users believing the character to be less friendly, intelligent and fair. This result was linked with the male character and not with the female character. This suggests that anthropomorphic feedback (at least in some forms) may not be the best approach.

Also, Prendinger, Ma and Ishizuka [19] investigated the use of eye tracking for data collection. Prendinger et al specifically tested an animated character with gestures and voice, voice only and text. The context for this work involved showing users around an apartment via a computer monitor. Their main findings were that the character condition seemed to be better for directing 'attentional focus' to various objects on the screen. However the voice only condition fostered more attention on the part of the users towards 'reference objects' on the screen. They also observed that the text only condition induced participants to look at the text more than the character, in terms of gaze points. Finally, subjective aspects were inconclusive. Despite this study having some experimental flaws, such as having very small sample sizes, it does indicate that using an anthropomorphic entity is not necessarily better than other modes.

Furthermore, Qiu and Benbasat [20] investigated anthropomorphic agents in the context of an agent recommending products to users. They wanted to see the effects on users' 'social relationship' with an e-commerce system. In their study they observed that the anthropomorphic agent had a positive influence on users in terms of 'social presence', trust, enjoyment and future use of the 'system'. Therefore this study suggests that the anthropomorphic agent fostered positive interactions and thoughts from the part of a user.

However in an earlier study by McBreen, Anderson and Jack [15] the use of different embodied agents in different retail type domains was investigated. The retail domains they tested were cinemas, travel and banking. The agents they tested were female and male in each of the domains and formal/informal dress in each of the domains. Having conducted an experiment with this scenario, the basic results they obtained were that although participants rated the systems positively overall, no significant effects were found in participants' ratings for some applications (it must be noted however that exactly what they were rating was not clear from the paper). There were also no significant effects found for questions relating to how participants perceived the 'services' they used (e.g. likelihood of using the service in the future, convenience and ease of use etc.). Participants were also given the opportunity to select their preferred domain. This gave significant results showing the cinema domain to be the preferred one. The reason participants mainly chose the cinema domain was because of the issues of trust and errors. Overall, the cinema domain was seen as less of a problem should 'something' negative happen, e.g. it is better and more acceptable to miss a movie due to some error than say a flight to some destination. There were no significant effects for the gender and the type of agent for any of the domains tested with respect to the participants being asked if they approved of the voices used. Participants were also asked if they found the voices irritating. This gave a significant result showing that the female formal voice was preferred over the male formal voice. Also significance was found for the naturalness of the voices issue, where the female voice was considered to be more natural compared with the male voice. No differences were found for issues of agent friendliness, competence and domineering attitudes. However, there was a significant effect showing that participants had a preference for the formally dressed agent in the bank domain. The last significant effect was that participants expressed the opinion that the agents in the bank domain should be formal and that the agents in the cinema domain should be informal.

The study by McBreen et al [15] is not so clear cut as the study by Qiu and Benbasat [20]. McBreen et al found issues of lack of trust and no significance in terms of reusing the system at a future time. Therefore there are some contrasts in results with these two studies in the retail/e-commerce type context.

As stated at the outset, the main author of this work has also been working in this area for a number of years (e.g. [9, 10, 11, 12, 13, 14] and sometimes the results are not consistent with each other.

In a previous study conducted by the main author of this paper [10] in the context of PC building instructions, an anthropomorphic character condition was tested against a non-anthropomorphic text condition. For this experiment the main results for effectiveness (based on errors) were inconclusive. However the results for subjective satisfaction tended towards a preference for the anthropomorphic interface.

In contrast, another study by Murano [9] in the specific context of English pronunciation, showed with significant results that using an anthropomorphic feedback was more effective, where the feedback helped more with the users' self-correction process in pronunciation. However in terms of user satisfaction, the results were inconclusive. The experiment had Italian participants with imperfect English pronunciation. Tasks involving pronunciation exercises were used where either an anthropomorphic (video of a human) or non-anthropomorphic (guiding text and diagram) feedback was used to assist in the correction process.

This brief review of some of the work already carried out by others and the main author of this paper, indicates that the findings do not always agree with each other across the whole range of work and is therefore worthy of further study. The last part of this literature review will discuss the concepts of affordances so as to inform the reader about the concepts and more clearly see how they were used in the experiment detailed in subsequent sections of this paper.

The concept of affordances was initially devised by Gibson [7]. Gibson was the first researcher to systematically study and propose physical affordances. As the affordances in relation to a computer user interface are different to the affordances discussed by Gibson, a detailed consideration of Gibson's theory is beyond the scope of this paper.

However affordances have been reinterpreted for application to user interfaces. Norman [17, 18] and Hartson [8] are the main sources of the reinterpretations, with more lightweight contributions from Gaver [6] and McGrenere and Ho [16] where they started to apply affordances to computer systems and to decompose affordances into different components. For brevity this paper will only briefly review Hartson's contribution as in the authors'opinion Hartson's contribution is the most substantial.

Hartson [8] identifies cognitive, physical, functional and sensory affordances. His rationale is that when doing some

computer related task, the users are using cognitive, physical and sensory actions. Cognitive affordances involve 'a design feature that helps, supports, facilitates, or enables thinking and/or knowing about something' [8]. One example of this aspect concerns giving feedback to a user that is clear and precise. If one labels a button, the label should convey to the user what will happen if the button is clicked. Physical affordances are 'a design feature that helps, aids, supports, facilitates, or enables physically doing something' [8]. According to Hartson a button that can be clicked by a user is a physical object acted on by a human and its size should be large enough to elicit easy clicking. This would therefore be a physical affordance characteristic. Functional affordances concern having some purpose in relation to a physical affordance. One example is that clicking on a button should have some purpose with a goal in mind. The converse is that indiscriminately clicking somewhere on the screen is not purposeful and has no goal in mind. This idea is also mentioned in McGrenere and Ho [16]. Lastly, sensory affordances concern 'a design feature that helps, aids, supports, facilitates or enables the user in sensing (e.g. seeing, feeling, hearing) something' [8]. Sensory affordances are linked to the earlier cognitive and physical affordances as they complement one another. This means that the users need to be able to 'sense' the cognitive and physical affordances so that these affordances can help the user.

The prototype developed to deliver instruction on two sewing stitches and to also take into account the issues of facilitating and violating the affordances, was in terms of the affordances as interpreted by Hartson.

III. The Sewing Experiment

A. The Two Modes of Feedback and the Affordances

The anthropomorphic condition consisted of a video of a human describing how to perform each relevant stitch. Whilst the verbal description was taking place, the presenter also performed the actual stitch with a needle and fabric. This facilitated the affordances as described above, because the cognitive affordances were facilitated in that they supported the attempt of learning how to perform a stitch. Seeing the flow of a stitch occurring in a video aided this affordance. In this type of interface physical and functional affordances were less relevant. However the sensory affordances were also facilitated because the video amply helped the sense of seeing and hearing the instructions at the same time.

The non-anthropomorphic condition used a series of static diagrams showing the various stages of the stitches. Next to each diagram there was also some explanatory text for the user. The text was the same as the verbal description given in the anthropomorphic condition. This condition subtly violated the affordances because the series of diagrams and text did not facilitate as well the user in learning or knowing about the stitches due to the static and staged nature of displaying the steps to sew a stitch. This in turn would not have facilitated the sensory affordances because some aspects of 'flow' in performing a stitch would have been lost in this condition and therefore did not aid well the human ability of seeing or observation.

B. Hypotheses

The aim of the experiment was to check if anthropomorphic interface feedback would be more efficient and result in better user attitude, than non anthropomorphic interface feedback in the context of learning sewing. Also the anthropomorphic feedback was developed to facilitate the affordances, while the non-anthropomorphic feedback was developed to subtly violate some of the affordances. Testing the following hypotheses was a part of achieving the above aim.

Positive Hypothesis 1a: Participants will perform better in the easier sewing tasks after being instructed by the anthropomorphic feedback.

Null Hypothesis 1b: There will be no difference in performance for the easier sewing tasks regardless of feedback mode (anthropomorphic/non-anthropomorphic).

Positive Hypothesis 2a: Participants will perform better in the more difficult sewing tasks after being instructed by the anthropomorphic feedback.

Null Hypothesis 2b: There will be no difference in performance for the more difficult sewing tasks regardless of feedback mode (anthropomorphic/non-anthropomorphic).

Positive Hypothesis 3a: Participants will feel more positive while performing the easier sewing tasks after viewing the anthropomorphic form of instructions.

Null Hypothesis 3b: For the easier sewing tasks there will be no difference in positive perceptions on the participants' part.

Positive Hypothesis 4a: Participants will feel more positive while performing the more difficult sewing tasks after viewing the anthropomorphic form of instructions.

Null Hypothesis 4b: For the more difficult sewing tasks there will be no difference in positive perceptions on the participants' part.

Positive Hypothesis 5a: Participants will have a more positive attitude overall, using the anthropomorphic application.

Null Hypothesis 5b: There will be no difference in participants' overall positive perceptions.

C. Participants

There were 40 participants recruited. Participants were initially approached by the experimenter and were requested to participate in a software evaluation based on sewing. They were asked only once and not pressured to take part if they did not initially agree. This was because if they did not wish to take part in the experiment, pressuring them to do so could have negatively affected the results as the participants may have been unmotivated. Once initial agreement was granted by the participants, they were asked to complete a pre-experiment questionnaire which elicited various aspects of their past experiences.

For this experiment it was deemed to be necessary to have novices in terms of sewing skills as the prototypes developed were designed with novices in mind. Therefore the recruitment aimed to find participants with no hand or machine sewing

experience and no engagement in having extensively observed others sewing in the recent past. However participants with button sewing experience were allowed as it was considered that this skill would be quite common and difficult to completely filter out and that sewing a button is quite different to sewing a particular series of stitches in a particular 'pattern' on a piece of fabric. However, in order to be sure that button sewing experience did not affect the results, the data collected during the experiment was scrutinised in terms of the averages for performance variables. The results indicated that button sewing experience and lack of such experience did not greatly affect the averages (data not included for brevity).

It was a requirement to have participants that were fluent in English. Recruitment from the university population facilitated this aspect as native English speakers would be at the appropriate level and any non-native English speakers from overseas would have a minimum English level requirement for being granted a place to study at the university.

Computer experience was also a factor that was controlled. It was deemed to be important to have participants with a comfort level in using computers, applications and basic software installation. Also all participants had a minimum of between 1 and 3 years of experience and used a computer at least 2-4 times per week.

Since the tasks involved participants actually trying to sew some stitches with a needle and fabric, it was necessary to have participants with normal or corrected eyesight, no motor control impairments and no weakness in hands. Participants with such impairments could have had difficulty in using a sewing needle.

Furthermore, the participants chosen for this experiment were right handed. Left handed participants were excluded, as pilot tests were conducted with two left handed participants to see if the prototype instructions on sewing were suitable for left handed participants as well. The participants had trouble performing the tasks. One of them mentioned that any instruction which includes hand movements in a certain direction was hard to follow for left handed people. Lastly, the age of participants ranged from 20-35 years and each participant was randomly allocated to one of the two experimental conditions (anthropomorphic or non anthropomorphic).

D. Experiment Design

The participants were recruited randomly and were allotted one of the two applications, anthropomorphic or non-anthropomorphic. The experiment was a between users design. This design was chosen so as to avoid learning effects with the relevant sewing stitches to be used.

1) Anthropomorphic and Non-Anthropomorphic Applications: We developed two applications, one with anthropomorphic feedback and the other with non-anthropomorphic feedback. Both applications are Windows forms applications. Both began with the instructions on how to use the application and a briefing about the experiment itself. They then contained two sewing tasks each. Each Task is a tutorial with instructions on a type of sewing stitch followed by a test to perform the sewing stitch. Both the applications had the same two stitches instructed. The medium of instruction in the anthropomorphic feedback type application was a video of a person demonstrating the stitch. On the other hand the non-anthropomorphic application had the same instructions in text format with a graphical illustration. The instructions given in both applications for each of the stitches was exactly the same, even though the medium was a video or text. This was done to ensure that the level and accuracy of information in both types of application was exactly the same, so as not to create a bias for any one type of the application.

2) Pilot Testing: Before the prototypes and procedure were used in the actual experiment, some pilot testing was undertaken. Regarding the stitches used, a sewing professional was consulted regarding the suitability of the stitches and the ensuing discussion led to the use of back and chain stitch. These were chosen so that the prototype concepts could be demonstrated with contrasting stitches (i.e. one easier – back stitch, and the other more difficult – chain stitch). The instructions were taken from various websites e.g. [1, 3, 21] etc.

Furthermore, testing of the instructional content was carried out with a couple of users and these led to some improved phrasing in the instructions and diagrams initially used, for the non-anthropomorphic condition. The videos used in the anthropomorphic condition were also revised to be clearer, by altering the shooting angle. A change in the shooting angle also resulted in a better contrast with the fabric, needle and thread being used.

Once these changes had been carried out, the two conditions and experiment procedure were once again tested by two participants. The results then gave confidence that the actual experiment was ready to be carried out with real participants.

E. Variables

The independent variables were (1) the types of feedback presenting the information (anthropomorphic and non-anthropomorphic) and (2) Type of Task (performing back stitch and chain stitch), where the values obtained from the performance data were included in the analyses (i.e. not the tasks themselves).

The dependent variables were the participants' performance in carrying out the tasks and their subjective opinions.

The dependent measures were that the performance was measured by timing how long it took to complete 5 stitches, the number of correct stitches made, the number of partially correct stitches made, the number of incorrect stitches made and the number of revisits to the instructional material presented. Each stitch (back and chain stitch) was identified as being composed of several sub-stages. Therefore if a participant only achieved one or more sub-stages, but not all required sub-stages to complete a stitch, this would be categorised as a partially correct stitch. However all sub-stages completed correctly would be categorised as a correct stitch. An incorrect stitch would be one that had no correct sub-stages completed.

The subjective opinions were measured by means of a post-experiment questionnaire. The post-experiment questionnaire was divided into three main sections. The first section contained general questions regarding the appearance and organisation of the information of the prototypes, including elicitation of the participants' feelings during the interaction. The second section contained questions relating to the tasks. These questions concentrated on the participants' feelings during the stitching tasks and on the actual information relating to the instructional material. The third section was open ended in terms of participants expressing opinions for improvement for presenting such information. Furthermore, observable attitudes noted on the observation protocol used during the experiment, were also used as a measure.

F. Materials and Apparatus

The materials used were:

- A laptop on which the experiment took place. The screen resolution was 1200 by 800 pixels with the highest (32 bit) colour quality. The laptop speaker volume was kept to 50% and the volume of media player control in the anthropomorphic application (windows forms application) was kept on its default value.
- Threaded needle.
- Plain cloth mounted on a frame.
- Pre loaded applications (anthropomorphic and non-anthropomorphic) on the laptop.
- Observation protocol, pre-experiment and post-experiment questionnaires.
- A meeting room on the university campus was used for conducting the experiment.

G. Procedure and Tasks

All participants did the experiment individually in the presence of the experimenter alone. Each participant took around 25 minutes to perform the complete experiment.

Upon arrival to the venue, each participant was allotted the same desk in the room, a comfortable chair and the same laptop was given each time. These details would ensure the participants did the experiment with maximum and equal concentration and that no bias was introduced by treating participants differently. Just after being greeted in the room by the experimenter, as a method for briefing the participants on the experiment, they were asked to read a small note on a sheet of paper just as they entered. The briefing note contained some simple details about the study, its purpose and some procedural aspects the participant would need to expect. It also stated that the participants were not personally being evaluated, but that the evaluation was of the software. The experimenter did not give any instruction out verbally. This again tried to ensure the same treatment for all participants.

Before beginning the experiment the participants were given a pre-experiment questionnaire to ensure participants had the required profile (see participants section above). The pre-experiment questionnaire elicited basic demographic information, sewing experience skills, knowledge of using computers, English proficiency and possible impairments causing potential bias, e.g. hand weakness etc.

Then the participants were allotted one of the two applications randomly. Regardless of experimental condition, participants were treated the same, and asked to proceed in the same manner. There was no difference in the two experimental conditions except in the anthropomorphic feedback condition, the application had videos for the two tutorials and in the non-anthropomorphic application the tutorials consisted of text along with graphical illustrations. Therefore the instructions were exactly the same.

The participants allocated to the anthropomorphic condition also carried out a small sound check to test the volume settings. A sample piece of sound from one of the videos was played for the sound check which simply said ‘Good work you’ve just learned back stitch’. If required they could adjust the laptop volume to their needs if it wasn’t audible enough for them. As there was no sound in the non-anthropomorphic condition, this sound check step was not required.

The next stage was to commence the experiment and run the software. Firstly the software gave the participants an introduction about how to use the application. Then some information about the experiment was presented. It consisted of explaining about the two tasks they would do.

After the introduction was completed the participant clicked 'next' to begin task 1 and reach the tutorial page. Here they were asked to begin viewing or reading the tutorial on Back stitch and told that they could go through the tutorial as many times as they liked provided they told the experimenter. They were informed it would help the experiment if they gave the true figure and it was nothing to do with their performance. The alternative of showing text for an average reading time to ensure participants read the text only once was considered. However that could have led to an anxious feeling in participants and thus possibly creating bias. This was therefore not implemented. Further, the bottom of the window prompted the participants to verbally describe their feelings to the experimenter after going through the tutorial, of which the experimenter made a note. After going through the tutorial participants could go to the next window. Here they read instructions for the test. In the instructions participants were asked to try and repeat 5 continuous stitches given in tutorial 1. The participant could choose to go back and view the tutorial again on the condition that they would have to restart the test and do all 5 stitches each time they went back to view the tutorial. Before they could start the participants were given a plain piece of cloth and a threaded needle. The experimenter recorded the time to perform the test and a count was kept of the total number of times the tutorial was viewed. Also, the experimenter made a note of participants’ attitude when they were carrying out the task. After finishing task 1 participants pressed 'next' to reach the ‘Break’ window, where they were asked to have a short two minute break.

Task 2 began after the participants completed task 1. Task 2 was designed in exactly the same way as task 1. The task proceeded in the same way with the same rules as described above. Everything was the same except the stitch participants

were shown and asked to reproduce was the more difficult Chain stitch.

During each task, participants were discreetly observed and a specially designed observation protocol was used to aid this process. The protocol was used to record the time taken for each task, the number of correct, incorrect and partially correct stitches achieved and outwardly manifested participant attitudes. The protocol was particularly useful concerning the recording of stitch accuracy. This is because it is easier to see an incorrect stitch as it is being formed, rather than at the end of the process.

After finishing both tasks the participants were asked to complete a paper based post-experiment questionnaire (see Variables section for a brief description of the content of the questionnaire) using a 5 point Likert type scale. In all cases a 5 score was the most positive score that could be allocated.

H. Results

The data were analysed using a multifactorial analysis of variance (MANOVA) and when significance was found, the particular issues were then subjected to post-hoc testing using in all cases either t-tests or Tukey HSD tests. This was for confirmation of significance. For brevity the post-hoc test results are not presented here. Also for brevity, only the summary data concerning significant results are presented here. The distribution summary data is presented in Appendix I below.

However, the following tables in this section present the significant results for the MANOVA analysis. While all aspects in each table have their purpose and importance, the reader's main attention is drawn to the F Ratio in bold font of each table, where the discussion will tend to concentrate on these figures. Furthermore, the abbreviation DF (Degrees of Freedom) is used in each MANOVA table. Then the Experiment Discussion section below will discuss the results in light of the issues being investigated.

For the variables task 1 - number of correct stitches, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The anthropomorphic condition produced significantly more correct stitches than the non-anthropomorphic group (see table I).

TABLE I. MANOVA - TASK 1 - NO. OF CORRECT STITCHES, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	129.32008	25.8640	**12.6248**
Error	34	69.65492	2.0487	Prob > F
C. Total	39	198.97500		<.0001

For the variables Task 1 – number of incorrect stitches, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The anthropomorphic condition produced significantly fewer incorrect stitches than the non-anthropomorphic group (see table II).

TABLE II. MANOVA - TASK 1 – NO. OF INCORRECT STITCHES, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	104.48059	20.8961	**6.9726**
Error	34	101.89441	2.9969	Prob > F
C. Total	39	206.37500		0.0001

For the variables Time taken for task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The anthropomorphic condition completed the task significantly faster than the non-anthropomorphic condition (see table III).

TABLE III. MANOVA – TIME TAKEN FOR TASK 2, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	16.540933	3.30819	**4.3322**
Error	34	25.963264	0.76363	Prob > F
C. Total	39	42.504197		0.0037

For the variables number of correct stitches produced for task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The anthropomorphic condition significantly produced more correct stitches for the more difficult task, than the non-anthropomorphic condition (see table IV).

TABLE IV. MANOVA – TASK 2 NO. CORRECT STITCHES, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	131.43425	26.2868	**10.8907**
Error	34	82.06575	2.4137	Prob > F
C. Total	39	213.50000		<.0001

For the variables number of incorrect stitches produced for task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The anthropomorphic condition made significantly fewer incorrect stitches for the more difficult task, than the non-anthropomorphic condition (see table V).

TABLE V. TASK 2 NO. INCORRECT STITCHES, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	47.31334	9.46267	**3.6318**
Error	34	88.58666	2.60549	Prob > F
C. Total	39	135.90000		0.0097

For the variables number of tutorial visits for task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The anthropomorphic condition visited/viewed the tutorial significantly fewer times, than the non-anthropomorphic condition (see table VI).

TABLE VI. MANOVA – TASK 2- NO. TUTORIAL VISITS, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	15.223326	3.04467	**3.8805**
Error	34	26.676674	0.78461	Prob > F
C. Total	39	41.900000		0.0069

The above summary tables, concern the performance data that was recorded as part of the experiment. Now follows the summary data concerning the subjective opinions of the participants. As stated above, these opinions were elicited by means of a post-experiment questionnaire. In all cases these responses were on a five point Likert type scale where totally positive scores were ranked as 5 and totally negative scores were ranked as 1.

For the variables of having a feeling of clarity whilst going through the tutorial for task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition rated the tutorial with a significantly higher positive score for feelings of clarity in relation to the tutorial and the second task, than the non-anthropomorphic condition (see table VII).

TABLE VII. MANOVA – TASK 2 CLARITY OF FEELING WHILST GOING THROUGH TUTORIAL, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	28.471534	5.69431	**5.9337**
Error	34	32.628466	0.95966	Prob > F
C. Total	39	61.100000		0.0005

For the variables of having a feeling of satisfaction whilst going through the tutorial for task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition rated the tutorial with a significantly higher positive score for feelings of satisfaction in relation to the tutorial and the second task, than the non-anthropomorphic condition, i.e. the participants in the anthropomorphic condition felt significantly more satisfied (see table VIII).

TABLE VIII. MANOVA – TASK 2 – FEELINGS OF SATISFACTION WHILST GOING THROUGH TUTORIAL, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	24.728413	4.94568	**4.6391**
Error	34	36.246587	1.06608	Prob > F
C. Total	39	60.975000		0.0025

For the variables of feeling stimulated whilst going through the tutorial for task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition rated their feelings of being stimulated with a significantly higher positive score in relation to the tutorial and the second task, than the non-anthropomorphic condition i.e. the participants in the anthropomorphic condition felt significantly more stimulated (see table IX).

TABLE IX. MANOVA – TASK 2 – STIMULATING FEELING WHILST GOING THROUGH TUTORIAL, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	17.183223	3.43664	**5.1436**
Error	34	22.716777	0.66814	Prob > F
C. Total	39	39.900000		0.0013

For the variables of having a feeling of satisfaction after viewing the tutorial but before doing task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition rated their feelings of satisfaction with a significantly higher positive score in relation to the tutorial and the second task, than the non-anthropomorphic condition, i.e. the participants in the anthropomorphic condition felt significantly more satisfied after viewing the tutorial and just prior to actually doing the task (see table X).

TABLE X. MANOVA – TASK 2 – FEELING OF SATISFACTION AFTER VIEWING TUTORIAL BUT BEFORE DOING TASK, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	19.954554	3.99091	**4.0121**
Error	34	33.820446	0.99472	Prob > F
C. Total	39	53.775000		0.0057

For the variables of feeling stimulated after viewing the tutorial but before doing task 2, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition rated their feelings of being stimulated with a significantly higher positive score in relation to the tutorial and the second task, than the non-anthropomorphic condition i.e. the participants in the anthropomorphic condition felt significantly more stimulated after viewing the tutorial and just prior to actually doing the task (see table XI).

TABLE XI. MANOVA – TASK 2 – STIMULATING FEELING AFTER VIEWING TUTORIAL BUT BEFORE DOING TASK, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	24.067950	4.81359	**4.9546**
Error	34	33.032050	0.97153	Prob > F
C. Total	39	57.100000		0.0016

For the variables of, a feeling of clarity whilst carrying out the task, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition rated their feelings of clarity whilst carrying out task 2 with a significantly higher positive score, than the non-anthropomorphic condition i.e. the participants in the anthropomorphic condition felt significantly more feelings of clarity whilst doing the task (see table XII).

TABLE XII. MANOVA – TASK 2 – FEELING OF CLARITY DURING TASK, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	22.084087	4.41682	**3.7249**
Error	34	40.315913	1.18576	Prob > F
C. Total	39	62.400000		0.0085

For the variables concerning a feeling of satisfaction whilst carrying out the task, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The

participants in the anthropomorphic condition rated their feelings of satisfaction whilst carrying out task 2 with a significantly higher positive score, than the non-anthropomorphic condition i.e. the participants in the anthropomorphic condition felt significantly more satisfied whilst doing the task (see table XIII).

TABLE XIII. MANOVA – TASK 2 – SATISFYING FEELING DURING TASK, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	24.871758	4.97435	**5.2683**
Error	34	32.103242	0.94421	Prob > F
C. Total	39	56.975000		0.0011

For the variables concerning a feeling of satisfaction for the overall learning experience, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition rated their feelings of satisfaction for the overall learning experience with a significantly higher positive score, than the non-anthropomorphic condition (see table XIV).

TABLE XIV. MANOVA – TASK 2 – SATISFYING LEARNING EXPERIENCE, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	17.433564	3.48671	**3.5213**
Error	34	33.666436	0.99019	Prob > F
C. Total	39	51.100000		0.0113

For the variables concerning the perceived ability to remember the stitch learned, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition had a significantly higher positive score, than the non-anthropomorphic condition, i.e. participants in the anthropomorphic condition felt they would be able to remember the stitch they had learned (see table XV).

TABLE XV. MANOVA – TASK 2 – ABILITY TO RETAIN STITCH, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	23.035186	4.60704	**3.5812**
Error	34	43.739814	1.28647	Prob > F
C. Total	39	66.775000		0.0104

For the variables concerning the overall experience of carrying out the two sewing tasks, experimental condition, age range and gender, there is a significant ($p < 0.05$) difference. The participants in the anthropomorphic condition perceived the tasks overall to be significantly easier, than the non-anthropomorphic condition (see table XVI).

TABLE XVI. MANOVA – OVERALL EXPERIENCE OF DOING SEWING TASKS, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	14.827762	2.96555	**3.3445**
Error	34	30.147238	0.88668	Prob > F
C. Total	39	44.975000		0.0146

For the variables concerning the quantity of on-screen information, experimental condition, age range and gender, there is a significant ($p < 0.05$) difference. The participants in the anthropomorphic condition perceived the quantity of on-screen information to be significantly better, than the non-anthropomorphic condition (see table XVII).

TABLE XVII. MANOVA – ENOUGH ON-SCREEN INFORMATION, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	10.582722	2.11654	**2.4800**
Error	34	29.017278	0.85345	Prob > F
C. Total	39	39.600000		0.0510

For the variables concerning ease of understanding of the overall task instructions, experimental condition, age range and gender, there is a significant ($p < 0.01$) difference. The participants in the anthropomorphic condition perceived the overall task instructions to be significantly easier to understand, than the non-anthropomorphic condition (see table XVIII).

TABLE XVIII. MANOVA – EASY TO UNDERSTAND TASK INSTRUCTIONS, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	10.432167	2.08643	**5.4809**
Error	34	12.942833	0.38067	Prob > F
C. Total	39	23.375000		0.0008

For the variables concerning ease of understanding of the instructions for task 1, experimental condition, age range and gender, there is a significant ($p < 0.05$) difference. The participants in the anthropomorphic condition perceived the task 1 instructions (tutorial content) to be significantly easier to understand, than the non-anthropomorphic condition (see table XIX).

TABLE XIX. TUTORIAL 1 - EASILY UNDERSTOOD INSTRUCTIONS, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	7.533532	1.50671	**2.8083**
Error	34	18.241468	0.53651	Prob > F
C. Total	39	25.775000		0.0315

For the variables concerning ease of understanding of the instructions for task 2, experimental condition, age range and gender, there is a significant ($p < 0.05$) difference. The participants in the anthropomorphic condition perceived task 2 instructions (tutorial content) to be significantly easier to understand, than the non-anthropomorphic condition (see table XX).

TABLE XX. MANOVA – TUTORIAL 2 – EASILY UNDERSTOOD INSTRUCTIONS, EXPERIMENTAL CONDITION, AGE RANGE & GENDER

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	17.564110	3.51282	**3.3610**
Error	34	35.535890	1.04517	Prob > F
C. Total	39	53.100000		0.0142

I. Experiment Discussion

The results show a clear pattern suggesting that the anthropomorphic user interface was more effective and preferred by users. As expected the first task did not show many significant differences, because this involved a simpler stitch. However, crucially, despite the first task being simpler, the errors were significantly more in the non-anthropomorphic condition. The majority of significant results pertain to the second task which involved a more complicated stitch. Therefore the manner of presenting the instructional material had a more serious effect on the success of users and their perceptions about the interface.

The results give some confidence in accepting the stated positive hypotheses, which are reproduced below for convenience and are discussed in relation to the results presented in the previous section. Also the discussion will include the details of the affordances:

Positive Hypothesis 1a: Participants will perform better in the easier sewing tasks after being instructed by the anthropomorphic feedback. This clearly was the case as can be seen in tables 21 and 22, which showed with significance that the anthropomorphic condition incurred more correct stitches and fewer incorrect stitches. However the times taken for the first task were not significantly different.

Positive Hypothesis 2a: Participants will perform better in the more difficult sewing tasks after being instructed by the anthropomorphic feedback. This was shown to be true as can be seen in tables 23 to 26. These show that the anthropomorphic condition was significantly faster at completing the task, incurred more correct stitches, fewer incorrect stitches and fewer overall tutorial visits (which implies a more immediate/confident understanding of the stitch).

The data suggest that the facilitated affordances in the anthropomorphic condition did indeed have an effect. The cognitive affordances were facilitated by means of having a video of a human demonstrating the stitches. This better supported participants as they could see a stitch being performed in full flow, which is rather different to having to interpret a series of diagrams with accompanying text (as was the case in the non-anthropomorphic condition). Also the sensory affordances were better facilitated in the anthropomorphic condition. This is because the video of a human helped users to 'see' in full flow how a stitch should be made. This aspect of 'seeing' is very much a component of sensory affordances as discussed by Hartson [8]. While the non-anthropomorphic condition also had material for a user to 'see', it was deficient in terms of 'seeing' the flow of how to perform a stitch and therefore how to more easily go from one sub-stage to the next sub-stage so as to complete a full stitch and then a subsequent series of stitches.

Positive Hypothesis 3a: Participants will feel more positive while performing the easier sewing tasks after viewing the anthropomorphic form of instructions. This is less clear that the other hypotheses discussed in this section. However the authors cautiously accept this positive hypothesis because as can be seen in table 39, the participants in the anthropomorphic condition perceived the instructions for task 1 (i.e. the tutorial) to be significantly more easily understood, which shows a slightly more positive feeling. However not included in this paper, for brevity, the means of other non-significant results are in line with the anthropomorphic feedback being more effective and eliciting more positive user perceptions. Some examples include the number of tutorial visits being fewer for the anthropomorphic feedback and all the factors eliciting user perceptions had higher positive means for the anthropomorphic feedback (i.e. the non-anthropomorphic feedback was perceived more negatively).

Positive Hypothesis 4a: Participants will feel more positive while performing the more difficult sewing tasks after viewing the anthropomorphic form of instructions. This was also shown to be true as can be seen in tables 32 to 34. These suggest that participants in the anthropomorphic condition had more positive feelings of clarity and satisfaction. Also table 40 suggests that the instructions were significantly more easily understood under the anthropomorphic condition.

Positive Hypothesis 5a: Participants will have a more positive attitude overall, using the anthropomorphic application. The significant results also suggest this to be the case. Table 35 shows that participants felt more able to remember the stitch they had learned under the anthropomorphic condition. Also tables 36 to 38 show that positive perceptions were incurred for more general aspects, e.g. the experience, quantity of on-screen information and the task instructions being understandable. Furthermore, tables 27 to 31 suggest that participants in the anthropomorphic condition maintained a more positive attitude whilst going through the various stages of the experiment, i.e. from viewing the tutorial right through to carrying out and completing the task.

As can be clearly seen, the results for positive attitudes are strongly in favour of the anthropomorphic feedback. This is also linked with the cognitive and sensory affordances being better facilitated as discussed above. This is because users will tend to have a 'feel' (i.e. they may not consciously know why or how they 'feel' about something such as an interface) for a user interface that they perceive to have helped them well (or not). This is suggested by the results for the non-anthropomorphic condition, where it was consistently rated more negatively than the anthropomorphic feedback. Clearly the fact that overall users tended to make more errors, took longer for the tasks and viewed the tutorial more will have contributed to the more negative perceptions about the 'system'.

IV. Conclusions

The results agree with the expectation that the anthropomorphic feedback which facilitated more the affordances would have been more effective and preferred by users. The fact that the non-anthropomorphic feedback subtly violated the affordances resulted in more errors being committed and as one would expect, the users perceived that the feedback was not as good as it could be and therefore under several factors (see above) rated the non-anthropomorphic feedback significantly more negatively than the participants in the anthropomorphic condition.

When particularly the cognitive affordances and sensory affordances are not facilitated, the number of errors and the time taken to complete a task are significantly increased, particularly when the task has a degree of difficulty. The user perceptions are also more negative. They tend to feel less satisfied, less stimulated, less confident and understanding is perceived to be less. The results presented above indicate this and none of the results contradict this reasoning.

V. Further Work

As stated above, the anthropomorphic condition was deliberately developed to facilitate the affordances and the non-anthropomorphic condition deliberately and subtly violated the affordances and generally the expected results are borne out. The next stage in this work should continue to address the anthropomorphism issue and a future experiment will have an anthropomorphic condition that violates the affordances and a non-anthropomorphic condition that facilitates the affordances. This way forward should give further clear indicators that the affordances are possibly more crucial for usability than the actual anthropomorphic presence.

References

[1] Alternative Windows (2004) Backstitch, http://www.alternative-windows.com/stitches.htm, Accessed 2011.

[2] Bengtsson, B, Burgoon, J. K, Cederberg, C, Bonito, J, and Lundeberg, M. (1999) The Impact of Anthropomorphic Interfaces on Influence, Understanding and Credibility. Proceedings of the 32nd Hawaii International Conference on System Sciences, IEEE.

[3] CoatsCrafts (2011) How to Back Stitch, http://www.coatscrafts.co.uk, Accessed 2011.

[4] David, P., Lu, T., Kline, S. and Cai, L. (2007) Social Effects of an Anthropomorphic Help Agent: Humans Versus Computers, CyberPsychology and Behaviour, 10, 3. Mary Ann Liebert Inc.

[5] De Angeli, A, Johnson, G. I. and Coventry, L. (2001) The Unfriendly User: Exploring Social Reactions to Chatterbots, Proceedings of the International Conference on Affective Human Factors Design, Asean Academic Press.

[6] Gaver, W. W. (1991) Technology Affordances, Proceedings of the ACM, CHI 91, Human Factors in Computing Systems Conference, April 27 – May 2 1991, New Orleans, Louisiana, USA, p79-84.

[7] Gibson, J. J. (1979) The Ecological Approach to Visual Perception, Houghton Mifflin Co.

[8] Hartson, H. R. (2003) Cognitive, Physical, Sensory and Functional Affordances in Interaction Design, Behaviour and Information Technology, Sept-Oct 2003, 22 (5), p.315-338.

[9] Murano, P., (2002) Effectiveness of Mapping Human-Oriented Information to Feedback From a Software Interface, Proceedings of the 24th International Conference on Information Technology Interfaces, Cavtat, Croatia, 24-27 June 2002.

[10] Murano, P., Ede, C. & Holt, P. O. (2008) Effectiveness and Preferences of Anthropomorphic User Interface Feedback in a PC Building Context and Cognitive Load, 10th International Conference on Enterprise Information Systems, Barcelona, Spain, 12-16 June 2008. (c) - INSTICC.

[11] Murano, P. & Holt, P.O. (2010) Evaluation of an Anthropomorphic User Interface in a Telephone Bidding Context and Affordances, 12th International Conference on Enterprise Information Systems, Madeira, Portugal, 8-12 June. (c) - INSTICC.

[12] Murano, P., Malik, A. & Holt, P. O. (2009) Evaluation of Anthropomorphic User Interface Feedback in an Email Client Context and Affordances, 11th International Conference on Enterprise Information Systems, Milan, Italy, 6-10 May. (c) - INSTICC.

[13] Murano, P. (2005) Why Anthropomorphic User Interface Feedback Can be Effective and Preferred by Users, 7th International Conference on Enterprise Information Systems, Miami, USA, 25-28 May 2005. (c) - INSTICC

[14] Murano, P., Gee, A. & Holt, P. O. (2007) Anthropomorphic Vs Non-Anthropomorphic User Interface Feedback for Online Hotel Bookings, 9th International Conference on Enterprise Information Systems, Funchal, Madeira, Portugal, 12-16 June 2007. (c) - INSTICC.

[15] McBreen, H, Anderson, J. and Jack, M. (2000) Evaluating 3D Embodied Conversational Agents in Contrasting VRML Retail Applications. Proceedings 4th International Conference on Autonomous Agents, p. 39-45, ACM.

[16] McGrenere, J. and Ho, W. (1991) Affordances: Clarifying and Evolving a Concept, Proceedings of Graphics Interface, May 2000, Montreal, Canada.

[17] Norman, D. A. (1999) Affordance, Conventions, and Design, Interactions, May-June 1999, p.39-42.

[18] Norman, D. A. (2002) The Design of Everyday Things, Basic Books.

[19] Prendinger, H., Ma, C. and Ishizuka, M. (2007) Eye Movements as Indices for the Utility of Life Like Interface Agents: A Pilot Study, Interacting With Computers, 19, 281-292.

[20] Qiu, L. and Benbasat, I. (2009) Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems, Journal of Management Information Systems, 25(4), p 145-181, Sharpe Inc.

[21] Stitchclub (2011) How to Sew Chain Stitch, http://www.zincdesign.co.uk/stitchclub/Chain_stitch_final.pdf, Accessed 2011.

Authors Profile

Dr Pietro Murano is a Computer Scientist at the University of Salford, UK. Amongst other academic and professional qualifications he holds a PhD in Computer Science. His specific research areas are in Human Computer Interaction and Usability of software systems.

Tanvi Sethi obtained an MSc in Databases and Web Based Systems in 2008 and currently works in a professional capacity in the computing/software industry.

Appendix I - Distributions

TABLE XXI. Task 1 - No. of Correct Stitches

Anthropomorphic	
Mean	4.45
Std Dev	1.2763022
Std Err Mean	0.2853899
upper 95% Mean	5.0473278
lower 95% Mean	3.8526722
N	20
Non-Anthropomorphic	
Mean	1.1
Std Dev	1.7137217
Std Err Mean	0.3831998
upper 95% Mean	1.9020464
lower 95% Mean	0.2979536
N	20

TABLE XXII. Task 1 – No. of Incorrect Stitches

Anthropomorphic	
Mean	0.35
Std Dev	1.1367081
Std Err Mean	0.2541757
upper 95% Mean	0.8819958
lower 95% Mean	-0.181996
N	20
Non-Anthropomorphic	

Anthropomorphic	
Mean	3.4
Std Dev	2.1618705
Std Err Mean	0.4834089
upper 95% Mean	4.4117866
lower 95% Mean	2.3882134
N	20

TABLE XXIII. TIME TAKEN FOR TASK 2

Anthropomorphic	
Mean	1.6985
Std Dev	0.4186102
Std Err Mean	0.0936041
upper 95% Mean	1.8944156
lower 95% Mean	1.5025844
N	20
Non-Anthropomorphic	
Mean	2.887
Std Dev	1.1482119
Std Err Mean	0.256748
upper 95% Mean	3.4243797
lower 95% Mean	2.3496203
N	20

TABLE XXIV. TASK 2 – NO. OF CORRECT STITCHES

Anthropomorphic	
Mean	5
Std Dev	0
Std Err Mean	0
upper 95% Mean	5
lower 95% Mean	5
N	20
Non-Anthropomorphic	
Mean	1.5
Std Dev	2.1884866
Std Err Mean	0.4893605
upper 95% Mean	2.5242433
lower 95% Mean	0.4757567
N	20

TABLE XXV. TASK 2 – NO. OF INCORRECT STITCHES

Anthropomorphic	
Mean	0
Std Dev	0
Std Err Mean	0
upper 95% Mean	0
lower 95% Mean	0
N	20
Non-Anthropomorphic	
Mean	2.1
Std Dev	2.1980853
Std Err Mean	0.4915068
upper 95% Mean	3.1287356
lower 95% Mean	1.0712644
N	20

TABLE XXVI. TASK 2 – NO. OF TUTORIAL VISITS

Anthropomorphic	
Mean	1.45
Std Dev	0.6863327
Std Err Mean	0.1534687
upper 95% Mean	1.7712136
lower 95% Mean	1.1287864
N	20
Non-Anthropomorphic	
Mean	2.65
Std Dev	0.9880869

Anthropomorphic	
Std Err Mean	0.220943
upper 95% Mean	3.1124389
lower 95% Mean	2.1875611
N	20

TABLE XXVII. TASK 2 – CLARITY OF FEELING WHILST GOING THROUGH TUTORIAL

Anthropomorphic	
Mean	4.45
Std Dev	0.8255779
Std Err Mean	0.1846048
upper 95% Mean	4.8363824
lower 95% Mean	4.0636176
N	20
Non-Anthropomorphic	
Mean	2.85
Std Dev	1.0894228
Std Err Mean	0.2436024
upper 95% Mean	3.3598656
lower 95% Mean	2.3401344
N	20

TABLE XXVIII. TASK 2 – FEELINGS OF SATISFACTION WHILST GOING THROUGH TUTORIAL

Anthropomorphic	
Mean	4.4
Std Dev	0.680557
Std Err Mean	0.1521772
upper 95% Mean	4.7185105
lower 95% Mean	4.0814895
N	20
Non-Anthropomorphic	
Mean	3.15
Std Dev	1.3869694
Std Err Mean	0.3101358
upper 95% Mean	3.7991217
lower 95% Mean	2.5008783
N	20

TABLE XXIX. TASK 2 – STIMULATING FEELING WHILST GOING THROUGH TUTORIAL

Anthropomorphic	
Mean	4.5
Std Dev	0.606977
Std Err Mean	0.1357242
upper 95% Mean	4.784074
lower 95% Mean	4.215926
N	20
Non-Anthropomorphic	
Mean	3.4
Std Dev	1.0462967
Std Err Mean	0.2339591
upper 95% Mean	3.8896819
lower 95% Mean	2.9103181
N	20

TABLE XXX. TASK 2 – FEELING OF SATISFACTION AFTER VIEWING TUTORIAL BUT BEFORE DOING TASK

Anthropomorphic	
Mean	4.45
Std Dev	0.6863327
Std Err Mean	0.1534687
upper 95% Mean	4.7712136
lower 95% Mean	4.1287864
N	20

Anthropomorphic	
Non-Anthropomorphic	
Mean	3.2
Std Dev	1.2396944
Std Err Mean	0.2772041
upper 95% Mean	3.7801948
lower 95% Mean	2.6198052
N	20

TABLE XXXI. TASK 2 – STIMULATING FEELING AFTER VIEWING TUTORIAL BUT BEFORE DOING TASK

Anthropomorphic	
Mean	4.5
Std Dev	0.6882472
Std Err Mean	0.1538968
upper 95% Mean	4.8221096
lower 95% Mean	4.1778904
N	20
Non-Anthropomorphic	
Mean	3.2
Std Dev	1.2814466
Std Err Mean	0.2865402
upper 95% Mean	3.7997354
lower 95% Mean	2.6002646
N	20

TABLE XXXII. TASK 2 – FEELING OF CLARITY DURING TASK

Anthropomorphic	
Mean	4.4
Std Dev	0.88258
Std Err Mean	0.1973509
upper 95% Mean	4.8130601
lower 95% Mean	3.9869399
N	20
Non-Anthropomorphic	
Mean	3
Std Dev	1.213954
Std Err Mean	0.2714484
upper 95% Mean	3.5681479
lower 95% Mean	2.4318521
N	20

TABLE XXXIII. TASK 2 – SATISFYING FEELING DURING TASK

Anthropomorphic	
Mean	4.35
Std Dev	0.8127277
Std Err Mean	0.1817314
upper 95% Mean	4.7303683
lower 95% Mean	3.9696317
N	20
Non-Anthropomorphic	
Mean	3.2
Std Dev	1.2814466
Std Err Mean	0.2865402
upper 95% Mean	3.7997354
lower 95% Mean	2.6002646
N	20

TABLE XXXIV. TASK 2 – SATISFYING LEARNING EXPERIENCE

Anthropomorphic	
Mean	4.45
Std Dev	0.7591547
Std Err Mean	0.1697521
upper 95% Mean	4.8052953
lower 95% Mean	4.0947047
N	20

Anthropomorphic	
Non-Anthropomorphic	
Mean	3.25
Std Dev	1.1641577
Std Err Mean	0.2603136
upper 95% Mean	3.7948426
lower 95% Mean	2.7051574
N	20

TABLE XXXV. TASK 2 – ABILITY TO RETAIN STITCH

Anthropomorphic	
Mean	4.3
Std Dev	1.0809353
Std Err Mean	0.2417045
upper 95% Mean	4.8058933
lower 95% Mean	3.7941067
N	20
Non-Anthropomorphic	
Mean	3.05
Std Dev	1.234376
Std Err Mean	0.2760149
upper 95% Mean	3.6277058
lower 95% Mean	2.4722942
N	20

TABLE XXXVI. OVERALL EXPERIENCE OF DOING SEWING TASKS

Anthropomorphic	
Mean	4.55
Std Dev	0.6863327
Std Err Mean	0.1534687
upper 95% Mean	4.8712136
lower 95% Mean	4.2287864
N	20
Non-Anthropomorphic	
Mean	3.5
Std Dev	1.1470787
Std Err Mean	0.2564946
upper 95% Mean	4.0368493
lower 95% Mean	2.9631507
N	20

TABLE XXXVII. ENOUGH ON-SCREEN INFORMATION

Anthropomorphic	
Mean	4.55
Std Dev	0.7591547
Std Err Mean	0.1697521
upper 95% Mean	4.9052953
lower 95% Mean	4.1947047
N	20
Non-Anthropomorphic	
Mean	3.65
Std Dev	1.0399899
Std Err Mean	0.2325488
upper 95% Mean	4.1367302
lower 95% Mean	3.1632698
N	20

TABLE XXXVIII. EASY TO UNDERSTAND TASK INSTRUCTIONS

Anthropomorphic	
Mean	4.7
Std Dev	0.4701623
Std Err Mean	0.1051315
upper 95% Mean	4.9200428
lower 95% Mean	4.4799572
N	20
Non-Anthropomorphic	

Anthropomorphic	
Mean	4.05
Std Dev	0.8870412
Std Err Mean	0.1983484
upper 95% Mean	4.4651481
lower 95% Mean	3.6348519
N	20

TABLE XXXIX. TUTORIAL 1 – EASILY UNDERSTOOD INSTRUCTIONS

Anthropomorphic	
Mean	4.8
Std Dev	0.4103913
Std Err Mean	0.0917663
upper 95% Mean	4.9920691
lower 95% Mean	4.6079309
N	20
Non-Anthropomorphic	
Mean	4.05
Std Dev	0.9445132
Std Err Mean	0.2111996
Anthropomorphic	
upper 95% Mean	4.4920458
lower 95% Mean	3.6079542
N	20

TABLE XL. TUTORIAL 2 – EASILY UNDERSTOOD INSTRUCTIONS

Anthropomorphic	
Mean	4.45
Std Dev	0.8255779
Std Err Mean	0.1846048
upper 95% Mean	4.8363824
lower 95% Mean	4.0636176
N	20
Non-Anthropomorphic	
Mean	3.25
Std Dev	1.1641577
Std Err Mean	0.2603136
upper 95% Mean	3.7948426
lower 95% Mean	2.7051574
N	20

Digital Image Watermarking Technique Based on Different Attacks

Manjit Thapa
Department of Computer Science
Sri Sai College of Engg. & Tech,
acronyms acceptable
Badhani (Pathankot)

Dr. Sandeep Kumar Sood
Department of Computer Science
and Engineering
G.N.D.U.R.C,
Gurdaspur, India

A.P Meenakshi Sharma
Department of Computer Science
Sri Sai College of Engg. & Tech,
acronyms acceptable
Badhani (Pathankot)

***Abstract*— Digital watermarking is used to hide the information inside a signal, which cannot be easily extracted by the third party. Its widely used application is copyright protection of digital information. It is different from the encryption in the sense that it allows the user to access, view and interpret the signal but protect the ownership of the content. One of the current research areas is to protect digital watermark inside the information so that ownership of the information cannot be claimed by third party. With a lot of information available on various search engines, to protect the ownership of information is a crucial area of research. In recent year, several digital watermarking techniques are presented based on discrete cosine transform (DCT), discrete wavelets transform (DWT) and discrete fourier transforms (DFT). In this paper, we proposed an algorithm for digital image watermarking technique based on singular value decomposition; both of the L and U components are explored for watermarking algorithm. This technique refers to the watermark embedding procedure and watermark extracting procedure. Digital image watermarking techniques for copyright protection is robust. The experimental results prove that the quality of the watermarked image is good and that there is strong resistant against many attacks. The image watermarking techniques help to achieve artificial intelligence. Digital image watermarking is the most effective solution in this area and its use to protect the information is increasingly exponentially day by day.**

Keywords- Digital image watermarkin; copyright protection; Singular value decomposition; Watermark embedding procedure; Watermark extracting procedure.

I. INTRODUCTION

Digital watermarking is a technique that embeds data called watermark into a multimedia object such that watermark can be detected to make an assertion about the objects. It can be categorized as visible or invisible. Example of visible watermarking is the logo visible superimposed on the corner of television channel in a television picture. On the other hand, invisible watermark is hidden in the object, which can be detected by an authorized person. Such watermarks are used for suit the author authentication and detecting unauthorized copying. The novel technology of digital watermarking has been sponsored by many consultants as the best method for such multimedia copyright protection problem [1, 2]. O digital watermarking will have a variety of useful applications such as digital cameras, medical imaging, image databases, and video on demand systems, and many others. In recently years, many digital watermarking techniques have been proposed in the literature which is based on spatial domain technique and frequency domain technique. These techniques are used in watermark embedding algorithm and watermark extracting algorithm [3]. In 2005, Chen [4] proposed a singular value decomposition scheme that based on components of D and U. Without using DWT, DCT and DFT transforms. They showed that quality of watermarked image is good on their schemes. In 2007, Patra [5] introduced a novel digital watermarking method, which is based on single key image for extracting different watermarks. In this method, they used Arnold transform technique in watermark embedding and extraction, which is based on DWT and DCT algorithm. With the popularity of internet and availability of large storage devices, storing and transferring an image is simple and feasible. They showed that robustness of the algorithm against many signal processing operations. In 2010, Cox [6] suggested two blind, imperceptible and robust video watermarking algorithms that are based on singular value decomposition. Each algorithm integrates the watermark in the transform domain. They used the components of matrices such as U and V. Their schemes are shown to provide very good performance in watermarked video as compared to Chan [4].

Most of the domain transformation watermarking techniques works with DCT and DWT. However singular value decomposition (SVD) is one of the most powerful numeric analysis techniques and used in various requirements. These requirements can be organized and described as follows [7, 8].

In this paper, we will describe a digital image watermarking algorithm which is based on singular value decomposition technique. This paper is organized as follows. In Section 2, we introduce the SVD watermarking techniques briefly. In Section 3, we propose the embedding and extraction procedure. In Section 4, we evaluate the performance of watermark image. In section 5, we show the experimental results and Section 6 conclude the paper.

II. A REVIEW OF RELATED WORK

Singular value decomposition (SVD) is a mathematical technique based on linear algebra and used by factorization of a real matrix or complex matrix, with many useful applications in signal processing and statistics [9].

A. Singular Value Decomposition (SVD)

Singular value decomposition is one of a number of valuable numerical analysis tools which is used to analyze matrices. It can be appeared at from three jointly compatible points of view. On the other hand, we can see it as a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. At the same time, SVD is a method for identifying and ordering the dimensions along which data points demonstrate the most variation. This attach the third way of viewing singular value decomposition, which accepted the most variation, it's possible to find the best approximation of the original data points using less dimensions. Hence, SVD can be seen as a method for data reduction. In SVD transformation, a matrix can be decayed into three matrices that are having the same size as the original matrix. It is useful to establish a contrast with Gaussian elimination and its equation. Given A is a n × n square matrix, this matrix can be decomposed into three components, L, D and U, respectively such that

$$[L\ D\ U] = \text{SVD}(A),\ A^{'} = LDU^{T}$$

$$L^{-1} \text{ where } A = LDU.$$

$$= \begin{pmatrix} l_{1,1} & l_{1,2} & l_{1,n} \\ l_{2,1} & l_{2,2} & l_{2,n} \\ l_{3,1} & l_{3,2} & l_{3,n} \end{pmatrix} \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \sigma_{1,n} \\ \sigma_{21} & \sigma_{2,2} & \sigma_{2,n} \\ \sigma_{3,1} & \sigma_{3,2} & \sigma_{3,n} \end{pmatrix} \begin{pmatrix} u_{1,1} & u_{1,2} & u_{1,n} \\ u_{21} & u_{2,2} & u_{2,n} \\ u_{3,1} & u_{3,2} & u_{3,n} \end{pmatrix} \quad (1)$$

$$= \sum_{i=1}^{n} \sigma_i l_i u_i^{T}$$

Where the L and U components are real unitary matrices or complex matrices with small singular values, and the D component is an n × n diagonal matrix with larger singular value or eigen vector values entries which specify $\sigma_{1.1} >> \sigma_{2.2} \ldots\ldots\ldots\ldots \sigma_{k,k,\ k+1=} \ldots\ldots\ldots \sigma_{n.n} = 0$. $\sum$ are non zero matrix by diagonals of A. Reduced singular value decomposition is the mathematical technique underlying a type of document retrieval and word semblance method. These are also known as Latent Semantic Indexing or Latent Semantic Analysis. In this way, the three components of matrices L, D, and U specify $\mathbf{Au_i} = \sigma_i \mathbf{l_i}$ and $\mu_i^T A = \sigma_i u_i^T$.

Digital image watermarking techniques has some advantages that used singular value decomposition. Firstly, SVD transformation from the size of memory is not fixed and can be represented by a rectangle or a square. Secondly, SVD are increase in accuracy and decrease the memory requirement. Thirdly, digital images in singular values are less affected if general image watermark is executed. Fourth, singular value decomposition include by algebraic properties.

III. Proposed Watermarking Techniques

We proposed a singular value decomposition technique and quantization – based watermarking technique. The watermarking techniques can be represented by three components, L, D and U. It relies on row operations and column operations. Row operations involve pre-multiplying matrix and column operations involve post-multiplying matrix. The D component can be explored with a diagonal matrix. These techniques depend upon the watermark embedding procedure and watermark extraction procedure [12-13].

A. Watermark Embedding Procedure

The digital watermarking procedure can be followed by singular value decomposition techniques, which involve the characteristics of the D and U components. In the embedding procedure, the largest coefficients in D component were customized and used to embed a watermark. The adjustment was determined by the quantization method. We will start the procedure by applying SVD transformation on original image and to reconstruct the watermarked image. Because the largest coefficients in the D component can oppose with general image processing, the embedded watermark was not really affected. In this way, the quality of the watermarked image can be decomposed by quantization method. In our inspection, two important features of the D and U components are found. In the first feature, the number of non zero coefficients in the D component could be used to determine the complexity of a matrix. Commonly, the greater number of the non- zero coefficient can be specified by greater complexity. In the second feature, the relationship between the coefficients in the first column of the L component could be sealed, when usually image processing was presented as shown in figure 2. The watermarks embedding algorithm can be described as follows.

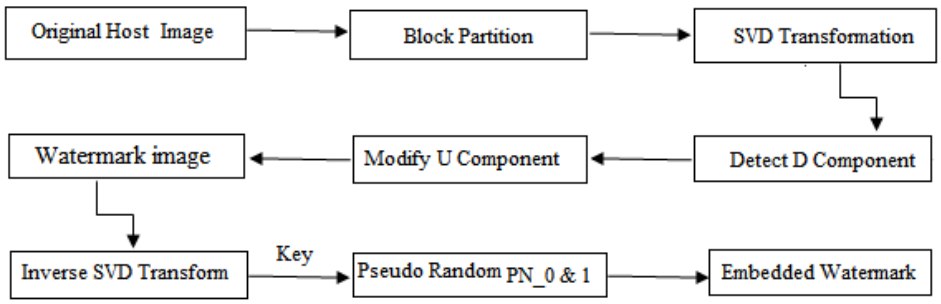


Figure 2. Watermark Embedding Algorithm

Step 1: Read the original image into blocks.

Step 2: Apply singular value decomposition (SVD) transformation.

Step 3: Extract the largest coefficient D (1, 1) from each D component and quantize by using a predefined quantization coefficients A. Suppose that S = D (1, 1) mod A.

Step 4: Perform embed the two pseudo-random sequences PN0, PN1, that is applied to the mid-band coefficients. If A is the matrix of the mid band coefficients of SVD transformed block, then embedding is done as follows:

If the watermark bit is 0 then, D (1, 1) = D (1, 1) + K/4 – A. so that [A< 3K/4] Otherwise, if the watermark bit is 1 then, D′ (1, 1) = D (1, 1) - k/4 +A so that [A<K/4]

Step 5: Apply the inverse of singular value decomposition transformation to reform the watermarked image.

B. Watermark Extracting Procedure

The watermark extracting procedure is similar to the watermark embedding procedure. Extraction procedure is the same as embedding one and pre-filtering is used before applying SVD transform to superior split watermark information from original image. The watermark extraction procedure is performed as described by the following steps. The first three steps of the watermark extracting procedure are same as the watermark embedding procedure except that the original image is replaced with the watermarked image [14]. Previously, an embedded block is detected according to the feature of the D component and PRNG, the relationship of the U component coefficients is observed. If a positive relationship is detected, the extracted watermark has assigned a bit value of 1. Otherwise, the extracted watermark has assigned a bit value of 0. These extracted bit values convert the original image SVD from the extracted watermark. The extracted watermark can be specified by original watermarked image and as shown in figure 3.

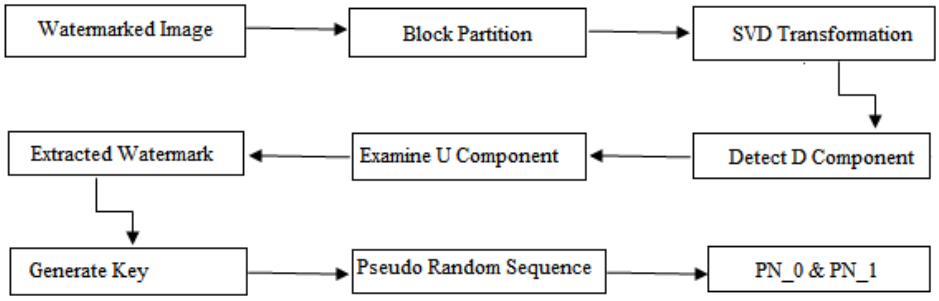


Figure 3. Watermark Extraction Algorithm

Step 1: Read the watermarked image into blocks.

Step 2: Apply the SVD transformation.

Step 3: Extract the largest coefficients D (1, 1) from each D component and quantize by using a predefined quantization coefficients A. Suppose that S = D (1, 1) mod A.

Step 4: Regenerate the two pseudo random sequences number using the same key, which is used in the watermark embedding procedure.

Step 5: For an extraction watermark bit valued of zero, if A<K/2. On the other hand, the extraction watermark bit value of one, if A>k/2.

Step 6: The watermark is restructured using the extracted watermark bits, and compute the similarity between the original watermark and extracted watermarks.

In this technique, the steady property of the largest D component coefficients resists the image processing was preserved. More ever, D component was a diagonal matrix in which only a small number of the coefficients could be used. In addition, the modification of the largest coefficients would cause a largest measure of image humiliation.

IV. Performance Evaluation

We evaluated the performance of the SVD image watermarking algorithms. The performance of the watermarking methods can be measured by imperceptibility and robust capabilities. Imperceptibility means that the superficial quality of the original image should not be distorted by the presence of watermark image. On the other hand, the robustness is a measure of the intentionally attacks and unintentionally attacks. It was found that the image quality measured by peak signal to noise ratio among the watermarked images was larger than 42 db [14-15]. This peak signal to noise ratio is defined as

$$PSNR_{db} = 10\log_{10}\left(\frac{MAX^2}{MSE}\right) \qquad (2)$$

$$= 20\log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right)$$

The PSNR is employed to evaluate the difference between an original image and watermarked image. For the robust capability, mean absolute error (MSE) measures the difference between an original watermark W and corresponding extracted watermark W^1 as shown by equation 3.

$$MSE(w, w^1) = \sum_{i=0}^{d} (\frac{w - w^1}{w}) \quad (3)$$

Generally, if PSNR value is larger than 40db the watermarked image is within acceptable degradation levels, i. e the watermarked is almost invisible to human visual system. A lower mean absolute error reveals that the extracted watermark W resembles the W^1 more closely. The strength of digital watermarking method is accessed from the watermarked image, which is further degraded by attacks and the digital watermarking performance of proposed method is compared with that of Chen [4]. If a method has a lower MSE (W, W^1), it is more robust.

V. EXPERIMENTAL RESULTS

The experimental results are simulated with the software MATLAB 7.10 version. We are using a 256×256 'Lena', 'facial', and 'Moon' as the gray scale original host image, and a 256×256 grey-scale image of the watermark image. The three images are shown in Fig. 4, 5 and 6 respectively. In the proposed method, we select the largest complexity of blocks; the original images can be separated into blocks of 4 × 4 pixels. Each block can be transformed into L, D, and U components by singular value decomposition. And then, a set of blocks with the same size as the watermark was selected, according to the feature of the D component. For an embedding watermark block, the relationship between the L component coefficients can be examined and the coefficients were modified, according to the watermark to be embedded. In our experiment, the original images and watermarked image quality shown by figure 4.

(a) (b) (c)

Figure 4. Three original images of 256 × 256 pixels (a) The original Lena Image (b) The original Facial Image (c) The original Moon Image.

(a) (b) (c)

Figure 5.Three watermarked images of 256×256 pixels (a) The watermarked Lena Image (b) The watermarked Facial Image (c) The watermarked Moon Image.

Table 1. The parameter values of attacked embedded watermarked image

Parameters	NO Attacks	Cropping Attacks	Pyramid Attacks	Rotation Attacks	Noise Attacks	Blurring Attacks	PSNR (DB)
E1	45.59	67.49	81.76	48.15	40.16	45.59	α = 0.3
E2	51.67	57.77	84.01	51.97	43.34	51.67	α = 0.3
E3	37.97	59.61	76.00	40.10	43.28	37.97	α = 0.3

Table2. The parameter values of attacked extraction watermarked image

Parameters	NO Attacks	Cropping Attacks	Pyramid Attacks	Rotation Attacks	Noise Attacks	Blurring Attacks	PSNR (DB)
E1	35.50	48.20	59.81	36.33	33.46	29.27	α = 0.4
E2	39.70	48.19	64.07	40.66	35.57	39.70	α = 0.4
E3	35.68	49.89	59.83	36.75	35.50	35.68	α = 0.4

The simulation results recommend that this algorithm can be robust against many different types of attacks such as no attacks, rotation attacks, noise attacks, and cropping attacks.

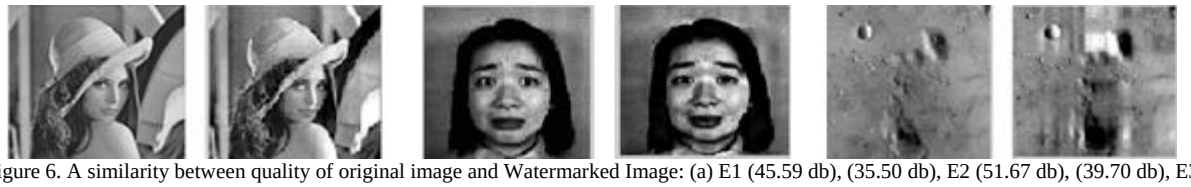


Figure 6. A similarity between quality of original image and Watermarked Image: (a) E1 (45.59 db), (35.50 db), E2 (51.67 db), (39.70 db), E3 (37.97 db), (35.68 db)

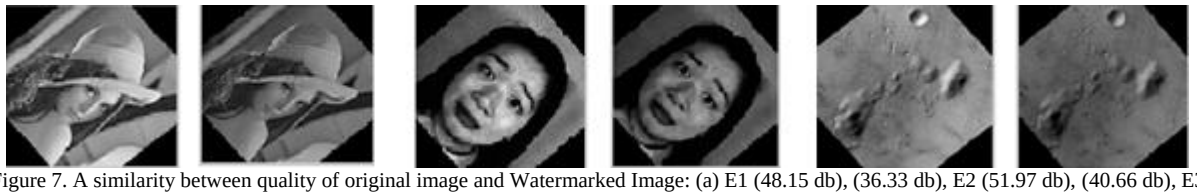


Figure 7. A similarity between quality of original image and Watermarked Image: (a) E1 (48.15 db), (36.33 db), E2 (51.97 db), (40.66 db), E3 (40.10 db), (36.75 db)

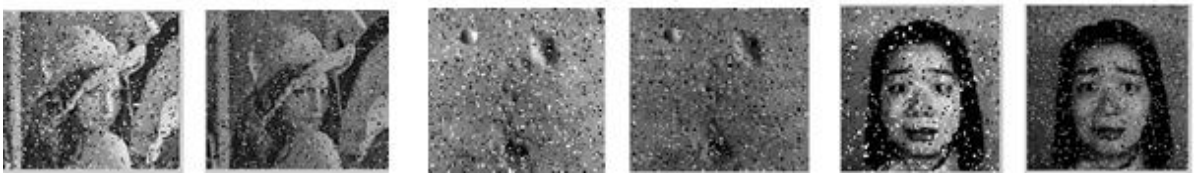


Figure 8. A similarity between quality of original image and Watermarked Image: (a) E1 (40.16 db), (33.46 db), E2 (51.67 db), (35.57 db), E3 (37.97 db), (35.50 db)

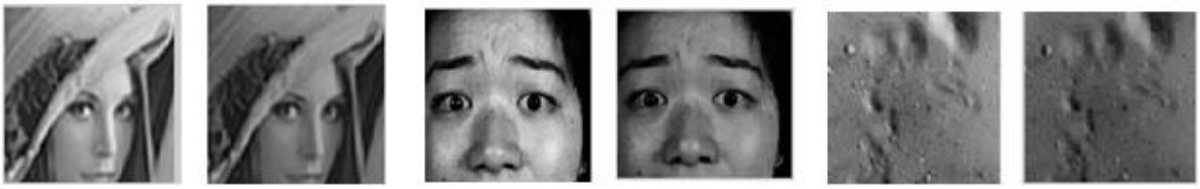


Figure 9. A similarity between quality of original image and Watermarked Image: (a) E1 (67.49 db), (48.20 db), E2 (57.77 db), (48.19 db), E3 (59.61 db), (49.89 db)

From the other data in table1 and table2, we can see the performance of our algorithm against the different geometrical attacks. Thus the proposed digital image watermarking algorithm can be used for protecting the copyrights of digital images.

VI. CONCLUSION

We introduced a digital image watermarking algorithm based on singular value decomposition. Digital image watermarking is one crucial area of research. Researchers have proposed various security techniques for to protect the ownership of digital information. It is used in security tools, security features and security parameter. We presented a technical discussion on digital watermarking techniques such as cropping attacks, rotation attacks, noise attacks and filter attacks. Digital watermarking can be utilized for aunthication of data, copyright protection and communication process. It provides a consistent performance on different original image and watermarked image in all the experiments.

The Experimental results prove that the quality of the watermarked image is better. Furthermore, the extracted watermark can be easily identified.

REFERENCES

[1] M. Barni and B. Bovid, "Digital Watermarking for Copyright Protection: A Communication Perspective", IEEE Communication Magazine, vol. 39, no. 8, pp. 90-91, 2001.

[2] A. Kumar and V. Santhi, "A Review on Geometric Invariant Digital Image Watermarking Techniques", International Journal of Computer Applications, vol. 12, no. 14, pp.31-36, 2010.

[3] F. Petitcolas, R. Anderson and M. Kuhn, "Attacks on Copyright Marking Systems in Information Hiding", LNCS, Berlin, vol. 1524, pp. 218-238, 1998.

[4] C. C. Chang and P. Tsai, "SVD-based Digital Image Watermarking Scheme", Pattern Recognition Letters, vol. 26, pp. 1577-1586, 2005.

[5] T. V. Nguyen and J. C. Patra, "A Simple ICA based Digital Image Watermarking Scheme", Digital Signal Processing, vol. 18, pp. 762-776, 2007.

[6] I. J. Cox and J. P. Linnartz, "Some General Methods for Tampering With Watermark", IEEE Journal on Selected Areas in Communications, vol. 16, pp. 587-593, 2010.

[7] Z. Bojkovic and D. Milovanovic, "Multimedia Contents Security :Watermarking Diversity and Secure Protocols", 6th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Service, TELSIKS, vol. 1, no. 3, pp. 377-383, 2003.

[8] S.J. Lee, S. H. Jung, "A Survey of Watermarking Techniques Applied to Multimedia", IEEE Transactions on Industrial Electronics, vol. 12 pp. 272-277, 2001.

[9] Y. Trank and W. Frank," Robust Image Watermarking in The Spatial Domain", Signal Processing, vol. 13, no 14, pp. 385-403, 1997.

[10] E. Koch and J. Zhao, "Robust Labels into Images for Copyright Protection", International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies, Vienna, pp. 1064-1087, 1985.

[11] P. K. Dhar and M.I. Khan, "A New DCT-based Watermarking Method for Copyright Protection of Digital Audio", International Journal of Computer Science & Information Technology (IJCSIT), vol. 2, no. 5, pp. 91- 97, 2010.

[12] M. Tsai and H. Hung, "DCT and DWT-based Image Watermarking Using Sub sampling," in Proc. of the IEEE Fourth Int. Conf. on Machine Learning and Cybernetics, China, pp: 5308-5313, 2005.

[13] S. C. Liu and S. D. Lin, "BCH code based robust audio watermarking in the cestrum domain," Journal of Information Science and Engineering, vol. 22, pp. 535–543, 2006.

[14] H. B. Kekre, D. Mishra, "Image retrieval using image hashing", Techno-Path: Journal of Science, Engineering & Technology Management, SVKM's NMIMS vol.2, n1,pp.230-240, 2010.

[15] A.H. Ali, M. Ahmad, Digital audio watermarking based on the discrete wavelets transform and singular value decomposition, Eur. J. Sci. Res. vol. 39 no. 1, pp. 6-21, 2010.

[16] W. Lu, H. Lu and F. L. Chung, "Feature Based Watermarking Using Watermark Template Match", Applied Mathematics and Computation, vol. 177, no. 1, pp. 886-893, 2011.

[17] Hamdy, S., El-messiry, H., Roushdy, M., & Kahlifa, E. (2010). Quantization Table Estimation in JPEG Images. International Journal of Advanced Computer Science and Applications - IJACSA, 1(6), 17-23.

[18] Babu, P. R. (2010). A Comprehensive Analysis of Spoofing. International Journal of Advanced Computer Science and Applications - IJACSA, 1(6), 157-162.

A Reliable Security Model Irrespective of Energy Constraints in Wireless Sensor Networks

D. Prasad
Department of Computer Engineering
Maharishi Markandeshwar University,
Mullana (Ambala), Haryana, India.
dprasadvns@gmail.com

Manik Gupta
Department of Computer Engineering
Maharishi Markandeshwar University,
Mullana (Ambala), Haryana, India.
manikjmu@gmail.com

R. B. Patel
Department of Computer Science,
D.C.R.U.S.T,
Murthal (Sonepat), Haryana, India.
patel_r_b@yahoo.com

***Abstract*- Wireless Sensor Networks (WSNs) are one of the most exciting and challenging research areas. It is an emerging technology that shows various applications both for public and military purpose. In order to operate these applications successfully, it is necessary to maintain privacy and secrecy of the transmitted data.**

In this paper, we have presented a Reliable Security Model (RSM) for WSNs. To incorporate the security, we are using four keys out of which two are static and remaining two are dynamic. One of the static key is obtained by composition of Q number of keys, and other is real-time MAC ID (RTMAC). Dynamic keys are computed on fly and keep on changing each time when the network is synchronized. In RSM, the synchronization time is less than the time required to compromise any node by an adversary, so that even if some nodes get compromised, the keying materials of the node have already been changed.

Keywords- Wireless sensor network (WSN); Sensor Node (SN); Base Station (BS); Static Keys; Dynamic Keys; Real-Time MAC ID (RTMAC).

I. INTRODUCTION

Wireless sensor networks (WSNs) are built up of sensor nodes (SNs), which consist of sensing, computing, communication, actuation and power components that cooperatively perform the task of collecting relevant data and monitor its surrounding for some change or event to occur [1]. Thus, two types of architectures were studied for WSNs. One for SNs itself and second for network architecture required for communication among the SNs. WSNs has its own features that not only differentiate it from other wireless networks but also craft the scope of wireless application to disaster relief, military surveillance, habitat monitoring, target tracking and in many civic, medical and security applications [2, 3, 20, 21]. Some features of WSNs that impose some limitation on WSNs and were kept in mind before developing this article are as following [11]:

A. Resource constraints

SNs have limited resources, including low computational capability, small memory, low wireless communication bandwidth, and a limited power battery.

B. Traffic characteristics

In WSNs, the primary traffic is in the upstream direction from the SNs to the sink node or BS, although the BS or sink nodes may occasionally generate certain downstream traffic for the purposes of query and control. In the upstream, this is a many-to-one type of communication. Depending on specific applications, the delivery of upstream traffic may be event-driven, continuous delivery, query-driven delivery, or hybrid delivery.

C. Small message size

Messages in sensor networks usually have a small size compared with the existing networks. As a result, there is usually no concept of segmentation in most applications in WSNs.

D. Sensor Location and Redundancy of Data

Position awareness of sensor network is important, since data collection is normally based on location. Also, there may be common phenomena to collect data, so there is a high probability that this data has some redundancy. There are three criteria that drive the common design issues for large-scale sensor networks; scalability (these networks might involve thousands of nodes), energy-efficiency (in particular, wireless communication can incur significantly higher energy cost than computation), and robustness (to environmental effects and node and link failures).

E. Network Lifetime

The time for which the network is operational or, put another way, the time during which it is able to fulfill its tasks (starting from a given amount of stored energy). It is not quite clear, however, when this time ends. Possible definitions are:

1) *Time to first node death:* When does the first node in the network run out of energy or fail and stop operating?
2) *Network half-life:* When have 50% of the nodes run out of energy and stopped operating?

F. Time to partition

When does the first partition of the network in two (or more) disconnected parts occur?

G. Addressing Schemes

Due to relatively large number of SNs, it is not possible to build global addressing schemes for the deployment of a large number of SNs as overhead of identity maintenance is high.

H. Density of nodes

In WSNs, the number of nodes per unit area i.e. the density of the network – can vary considerably. Different applications will have very different node densities. Even within a given application, density can vary over time and space because nodes fail or move; the density also does not have to be homogeneous in the entire network (because of imperfect deployment) and the network should adapt to such variations**.**

I. Maintainability

As both the environment of a WSNs and the WSNs itself change (depleted batteries, failing nodes, new tasks), the system has to adapt it by monitoring its own health and status to change operational parameters or to choose different trade-offs (e.g. to increase the interval of monitoring data and reduce quality when energy resource become scarce).

J. Node Deployment

Node deployment can be random, deterministic or self-organizing. For deterministic deployed networks the routes are pre-determined, however for random deployed networks and self-organizing networks route designation have been a challenging subject.

K. Energy consideration

Since the life-time of the WSNs depends on energy resources and their consumption by sensors, the energy consideration has a great influence on route design. The power consumed during transmission is the greatest portion of energy consumption of any node. Direct communication consumes more power than multi-hop communication; however the multi-hop communication introduces extra topology management and medium access control.

L. Miscellaneous applications

WSNs may be used in different environments supporting diverse applications, from habitat monitoring and target tracking to security surveillance and so on. These applications may be focused on different sensory data and therefore impose different requirements in terms of quality of service (QoS) and reliability. Thus sensor networks are application specific.

Now days, WSNs are not used only for security or social intention but also used for commercial purposes. Extensive research is going on in almost all fields of sensor network, including sensor design, communication protocol stack design, and operating system for sensors, security and management algorithm design. The design goals of WSNs are application specific, but share some common attributes like energy efficiency, scalability, robustness, network life time, fault tolerance, self-organization and data aggregation. Out of which, energy efficiency and security are more important. In recent years, the availability of cheap and tiny micro-sensors and low power wireless communication enabled the deployment of large quantity of wireless sensor nodes, which are scattered in the interested area.

In WSNs, routing, security and networks lifetime are seemed to be incompatible. But, in RSM, the balance of energy consumption among all the sensor nodes is maintained, in order to avoid the "hot spot" problem. Besides this, the availability of never lasting energy due to wireless power provides the security model more strength.

The main emphasis while designing and developing the protocol is uniform load distribution among SNs, in order to increase the network lifetime. Many protocols existing in the literature [22-25], minimize energy consumption on routing paths. Even though these many existing approaches increase energy efficiency, but technique such as; dynamic routing where, data is forwarded to nodes with the highest residual energy, may cause problem such as, unbounded delays. However, RSM is efficient enough to solve the security and energy issues, with dynamic routing and key management techniques that is not suffered from the traditional problems of unbounded delays and easy compromise of the nodes. Beside this, the network acts more efficiently in terms of energy with the help of wireless energy provided by the BS to the deployment area.

Rest of the paper is organized as follows. Section II summarizes the related works. In Section III, System model and protocol description is presented. Implementation of System Model is discussed in Section IV. In Section V, we analyze RSM in respect to energy, security and life of the network. System Model is compared with some existing techniques in Section VI followed by the results and discussions in Section VII. Finally, we conclude RSM and discuss the scope of future work in Section VIII.

II. RELATED WORKS

Security is a big issue, when WSNs are deployed in a hostile environment. Secret keys should be used to encrypt the exchanged data between communicating parties. In the Internet or traditional wireless networks, such as, cellular networks, most security protocols are based on asymmetric cryptography, such as; RSA or Elliptic Curve Cryptography (ECC) [6, 7] are not applicable, due to the high computational complexity, high-energy consumption and increased code storage requirements. Furthermore, due to unpredictable network topology and lack of infrastructure support, trusted-server based key distribution protocols are not suitable for WSNs either [5]. Research shows that key pre-distribution mechanism could be a practical method to solve the key distribution problem in WSNs. The basic idea of key pre-distribution scheme is preloading some secret keys into SNs, before they are deployed. After the deployment, the nodes discover shared keys for secure communications. It is divided into 3 phases; i.e. Key distribution, Shared key discovery and Path-key establishment. During these phases, secret keys are generated, placed in sensor nodes, and each sensor node searches the area in its communication range, to find another node to communicate. A secure link is established, when two nodes discover one or more common keys (this differs in each scheme), and communication is done on that link between those two nodes. For this purpose, various keying techniques are being used. Some of the common key management schemes are as follows [8]:

A. Single Network-wide Key Establishment

In this technique, a single key is preloaded into all the nodes of the network. After deployment, every node in the

network can use the same key for the message encryption and decryption. The main advantages of this technique are minimal storage requirements and avoidance of complex protocols. However, such keys provide enough space to adversaries for accessing the network. If any node is compromised in this scheme, then entire network is compromised.

B. Pair-wise Key Establishment

In pair-wise key establishment, every node in the network has to stores n-1 key pairs. This can be eradicated when we use a Trusted BS, also called centralized key distribution center (KDC), to send the session keys for the communication, between any two nodes. This scheme has small memory requirement and perfectly controlled node replication, it is resilient to node capture and possible to revoke key pairs. The drawbacks of this scheme are that it is not scalable and the base station becomes the target of attacks.

C. Dynamic Key Management

Dynamic key management system was proposed by Eltoweissy et al., also called exclusion-based system (EBS) [26]. Here, various nodes and methods are disclosed for automatically disseminating key node contact information in a network. Some of the advantages of using a dynamic key management scheme are: improved network survivability and better support for the network growth.

The issue in creating a dynamic key management system is being able to make it secure and efficient. The EBS assigns each node k keys from a key pool of size k+m. If node capture is detected, re-keying occurs throughout the network. A disadvantage to this EBS scheme is that if even a small number of nodes in the network are compromised, information about the entire network could be uncovered by an adversary.

D. Public Key Schemes

A public key scheme employs a pair of different but associated keys. One of these keys is released to the public while the other, the private key, is known only to its owner. WSNs have mostly been using symmetric key and other non-public-key encryption schemes [27]. A drawback to these schemes is that they are not much flexible, but they are computationally faster. With limited memory, computing and communication capacity, and power supply, sensor nodes cannot employ sophisticated cryptographic technologies, such as; typical public key cryptographs. The use of public key cryptography on WSNs has not been tested enough to rule it out completely.

E. Q-Composite Random Key Management

In this scheme, BS generates key sets from the key pool, for each node. SNs needs to have at least Q number of keys in common to establish a communication link, rather than only one key, which enhances the security level as compared to those schemes having single common key.

Besides the key management techniques, we have used RTMAC [28] for real time data streaming in WSNs. RTMAC is a collision free TDMA based MAC protocol which uses channel reutilization technique based on the network topology to reduce its latency. It maximizes spatial channel reuse by avoiding false blocking problem of RTS/CTS exchange based wireless MAC protocols. RT-MAC reduces contention duration for control packets to facilitate faster traveling of data packets; thus, it reduces end-to-end delay of data packet transmission and hence facilitates periodic delivery of data packets as well as fast reporting of an alarming event.

Each of the above WSN key management scheme consists of three main components [9]:

(1) key establishment (2) key refreshment (3) key revocation

Key establishment is about creating a session key between the parties that need to communicate securely with each other. Key refreshment prolongs the effective lifetime of a cryptographic key, whereas, key revocation ensures that an evicted node is no longer to able to decipher the sensitive messages that are transmitted in the network.

The reliable security model proposed in this article is scalable, secure and energy efficient. It enhances the security level of Q-composite keys scheme by changing the keying material every time, the network is synchronized. To increase the scalability and life of the network, RSM is introducing one more parameter (i.e. distance between SNs) in the Q-composite key scheme, for the establishment of communication link between SNs. Beside the energy efficient method, the model also introduces the concept of power beam [12] to supply power wirelessly to the BS as well as to the network throughout the life of the rechargeable batteries and the laser diode arrays and thus increasing the lifetime of the network almost to infinity.

Due to the recent efforts of MIT and Intel Co. [14, 15, 16], wireless electricity comes into revolutionary phase, which motivate us to think upon and work over the concept so as to make the WSN's life never lasting. Though this concept is motivating to work upon future applications, but still it is facing problem related to the transmission range, 3 to 8 meters. So, due to this problem as well as introduction of new hardware structure to the SNs, the current technology introduced by MIT and Intel Co. limits anyone to work upon WSN. But, if one thinks about the principle of power transmission in space to the space vehicles or space based solar power satellites (SPS), as mentioned in [17, 18] or to recharge batteries of satellites in geo-stationary orbits as in [19] made us to think beyond the concept of simple wireless electricity, as discussed in [14-16]. Though the concept of SPS is capable enough to provide unlimited power to the SNs in the deployment area as well as to the flying BS over it, but a large burden of extra cost is also involved in it as includes a large number of SNs and the BS to get charged through the satellites, which are governed by any third party and hence the point of security also arises. So, it would be more cost effective and more secured, if we are able to apply the same principle from the land itself rather than space, at the deployment area by the first party itself. This can be possible with Microwave power transmission (MPT) or Laser Power Beaming.

III. System Model & Protocol Description

We have divided this RSM into two phases i.e. pre-deployment phase and post-deployment phase.

A. Pre-Deployment Phase

In the pre-deployment phase, the network is deployed with an assumption that it is free from adversarial attacks during the setup phase. The pre-deployment phase deals with the following activities:

- Delivering wireless energy to the BS and SNs in the deployment area as in [12].
- Random deployment of SNs in the deployment area from the flying BS.
- Cluster formation among the sensor nodes in the deployment area.
- Generating a strong security model for the sensor nodes at their respective cluster ends.

Initially the power is supplied to the BS and SNs in the deployment area as in [12]. In RSM, the BS is considered to be any flying object like UAVs, which remains over the top of the deployment area. Though it may change its position, but it can compute the Localization ID (LID) of the sensor nodes in the deployment area with its own reference [4] to incorporate more security with respect to the GPS system.

After the BS receives energy to the threshold level, the static sensor nodes are randomly deployed from the BS in the deployment area. On the basis of the node Localization ID (LID), the BS divides the deployment area into various clusters. After the successful completion of the deployment phase of the SNs, the power supply phase in the deployment area begins after some time, since the SNs are initially completely charged before deployment.

Besides regular power supply to the deployment area, the BS also generates key sets from the key pool, for each node. By using the concepts of Q-composite keys, the nodes needs to have at least Q number of keys in common to establish a communication link rather than only one key, which enhances the security level of the network. But, the drawback with Q-composite is that if the Q number of keys is common between two nodes, which are far away from each other, then to establish communication link between such nodes is a bad idea, because to make communication between these nodes is much energy consuming. Keeping this in mind, the security model imposes a new constraint of distance on Q-composite concept. In RSM, communication link between such nodes will be established, only if the distance between such nodes is less than or equal to some threshold value D_0. The value of D_0 is guided by the deployment area and the density of nodes within that area.

ID_1	LID_{12}	$RTMAC_1$	$K_{11}, K_{12}, K_{13}, \ldots, K_{1K}$
ID_2	LID_{22}	$RTMAC_2$	$K_{21}, K_{22}, K_{23}, \ldots, K_{2K}$
ID_3	LID_{35}	$RTMAC_3$	$K_{31}, K_{32}, K_{33}, \ldots, K_{3K}$
:	:	:	
:	:	:	
:	:	:	
ID_N	LID_{N9}	$RTMAC_N$	$K_{N1}, K_{N2}, K_{N3}, \ldots, K_{NK}$
Sensor ID	Location ID	RTMAC	Sensor Key Sets

Figure 1. Data structure representing node information

In RSM, the BS generates N key sets of K keys in each, from the key pool, for the node to be deployed in the deployment area, and maintains a list containing node ID, node Localization ID (LID), Real-time MAC ID (RTMAC) and key sets assign to the node, as shown in Figure 1. RTMAC allows sensors to go to sleep when they are not communicating and hence it conserves energy. In RSM, nodes having Q keys in common, are known as logical neighbors, nodes having distance less than or equal to D_0 between them, are known as physical neighbors and nodes satisfying both criteria, are known as actual neighbors.

1) Finding Logical Neighbor (LN): The method of finding Logical neighbors is straightforward. To find the Logical neighbors of any node, BS compare all the keys in the key set, assign to the node under consideration, with all the keys in the key sets assign to other nodes one by one, and whenever BS find any node having Q keys in common with the node under consideration, it store the ID of that node in the 'LNbr' list. The BS finds all logical neighbors for each node and stores them in list 'LNbr', as shown in Figure 2(a). If any entry in the list 'LNbr' remains empty then BS assign new key set to the corresponding nodes and repeat the process.

2) Finding Physical Neighbor (PN): To find the physical neighbors of any node, BS compute the distance of the node under consideration with all its logical neighbors and store the ID of all those node, which distance is less than or equal to D_0 in 'PNbr' list as shown in Figure 2(b).

Finding Actual Neighbor (AN): With the help of these two lists, for each nodes BS finds all those nodes, which falls within distance D_0 and having at least Q keys in common and store them in 'Nbr' list, as shown in Figure 2(c). Any empty location in list 'Nbr' indicate that the corresponding node falls far away from the remaining node, which chance is very rare, because of dense deployment, and even if it happened, we can ignore it, because such nodes are very few in numbers.

In this way, all the exhaustive operations are managed by the BS itself rather than the SNs in the deployment area. Hence, we can achieve more security even in large sized network without any loss of energy as in the conventional scheme of the Q-Composite scheme.

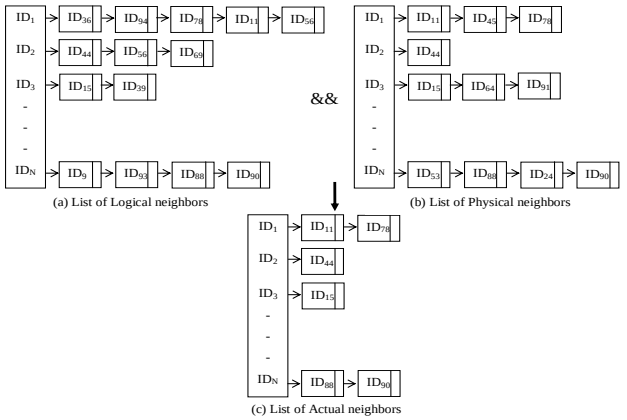


Figure 2. List of various neighbors maintained at BS.

B. *Post-Deployment Phase*

Once the BS obtains the actual neighbors for all nodes, it constructs two network graphs where, edges represents secure link and nodes represents sensor nodes in the network. The graphs are obtained separate for both odd as well as even RTMAC values, stored in the sensor nodes, which work on switching basis. The network graph which is active will sense the data but the inactive network will be in sleep mode to preserve energy and to provide illusion of dead network to the adversary if he was trying to capture a network, which suddenly enters into sleep mode by transferring the last sensed packet to the immediate physical neighbor, present in the alternate network. From the network graphs, the BS obtain two minimum spanning trees as shown in the Figure 3, and designate one of the node as cluster head in the active tree and set it to communicate with base station. This delegation must be on rotation basis, otherwise the energy of the node communicating continuously with the BS will be depleted soon and the whole network will be disconnected. To rotate the delegation, BS can choose any scheduling scheme; RSM introduced in this paper is using the scheme presented in GANM [10]. BS computes link keys between a node and all of its neighbors by applying some hash function, as shown in the algorithm. It also computes a timer value, as shown in the algorithm, to synchronize the network.

Here, both the trees so obtained are the representations within the same cluster units and the corresponding position nodes in odd and even networks are actually the physical neighbors (PNbr) of each other. Here, the dotted lines show the traversing of tree to rotate the delegation. It can be seen in figure 3 that when the delegation reached to the last node of the odd minimum spanning tree, it gets shifted to the even minimum spanning tree and vice-versa; thus giving illusion to any continuously observing adversary that the network becomes dead after sometime, which in fact is in sleeping mode for security and energy efficiency purpose.

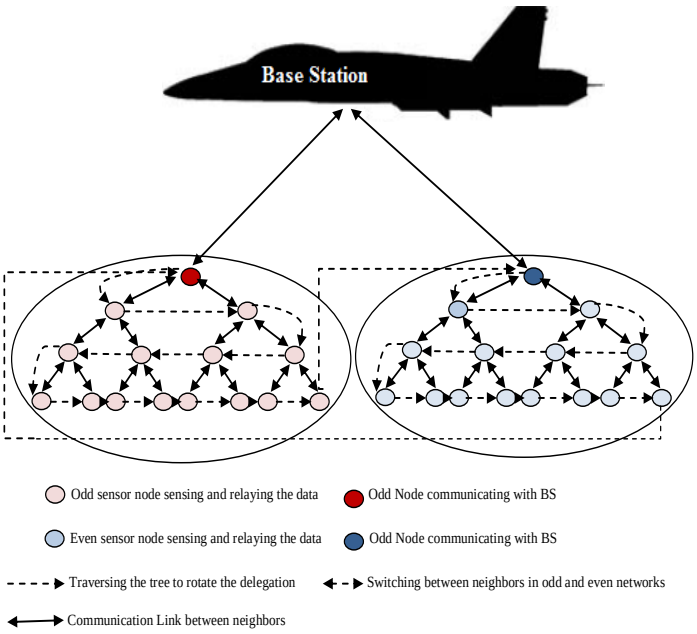


Figure 3. Communication network of our protocol

During this process of shifting the delegation from odd to even tree or vice-versa, each of the nodes in the tree which is going to be in sleep mode will send its last sensed data packet to its corresponding physical neighbor in other tree which will now act as active tree. Since, the physical neighbor in active tree will replace the corresponding node in the tree in sleep mode hence it is assumed that it will now sense the data in approximately the same sensing limits.

Once the link keys and timer value are computed, BS constructs N packets; one for each node, containing node ID, RTMAC, set of link keys for that node and timer value, as shown:

Cluster ID	Node ID	Neighbors Link Keys	Timer Value	RTMAC

Synchronization Packet

BS broadcast these packets in the network. Nodes in the area receive only the packet meant for them, store this information within its memory, and ignore other packets. On receiving the above packet, timer within the node is triggered and the whole network gets synchronized. Once the network is synchronized, nodes within the cluster start sensing the surrounding and send the sensed data, in the following format to its parents, where, these data are aggregated and this aggregated data is forwarded to their parent; this process is continued and finally aggregated data reaches to the node communicating with BS through which data is reached to the BS.

Source ID	Neighbors Link Keys	Timer Value	Data	RTMAC

Data Packet

This whole process of resynchronization is repeated after the regular interval of time in order to enhance the security level of each cluster within the network, by generating unpredictable key values, in the least possible interval. To

enhance the security level further, link keys and timer values may be encrypted before their transmission in the data packet.

IV. Implementation of System Model

In our network model, we have assumed that all sensors are distributed in an evenly randomized manner in a polygon region, and the network has the following properties:

1) There exists a unique BS, located at the top of the network with a maximum height of 1 mile due to the power constraint of power beam [13].
2) Each SN has a unique identity.
3) All sensors cannot move after being deployed.
4) Network is homogeneous i.e. all sensor nodes are equivalent, having the same energy, computing and communication capacity.
5) Location of nodes is obtained using virtual co-ordinate system as in [4].
6) The transmitter can adjust its amplifier power based on the transmission distance.
7) Each sensor node consists of photo-voltaic cells for charging the power both from the sun as well as from the beam director embedded at the BS.

A. Algorithm

List 'LNbr' is an array of pointer, in which locations are pointing to the link list of ID of logical neighbors of the node under consideration (i.e. nodes with Q keys in common, can be represented as $CKeys_{I,J}$ for nodes I and J), 'PNbr' is an array of pointer, in which locations are pointing to the link list of ID of those nodes whose physical distance is less than D_o, 'Nbr' is an array of pointer, in which locations are pointing to the nodes allowed to communicate with each other, Nbr[O] and Nbr[E] are the neighbor lists for odd and even MST respectively and 'Iso' is list of all those nodes, which are isolated from the network. Besides this, a list RTMAC[] is maintained to store the RTMAC numbers of the sensor nodes respectively.

The function SECURE_LINK(Iso[I]) is used to establish a secure link among the isolated nodes present in the list 'Iso[I]'. However, the function RESYNCHRONIZE( ) is used to resynchronize the entire network in the synchronization time, 'T_{sync}' so that, each delegate node, 'Dlgt' in the list communicates with the BS on round robin basis. The synchronization time, 'T_{sync}' must be less than the threshold time, 'T_0' (by some tolerance value ε) is the time taken by an adversary to capture any node in the network. Moreover, T_{BS} and timer[I] are the timers maintained at the Base Station and at each node in the network; acting as their respective dynamic keys.

For each cluster in the deployment area, the following algorithm will be executed concurrently, so as to find the dynamic cluster head that will act as delegate node to communicate with the BS.

1) While((Nbr[1])||(Nbr[2])||(Nbr[3])||...||(Nbr[E])||…||(Nbr[O]) = NULL) repeat steps 2 to 5
2) Initialize C :=1.
3) For I:=1 to N // Here, N=O+E.
   If (Nbr[I] = NULL) add its ID to Iso[C]; C:=C+1.
4) For I:=1 to C
   Generate new key set and replace the key set in KSets[I] corresponding to node Iso[I] by new set.
5) Call SECURE_LINK (Iso[I]).
6) Establish two way communication links by the link key as: K:= HASH $\{k_1||k_2||\ldots\ldots||k_Q\}$.
7) Call MINIMUM_SPANNING_TREE for the graph obtained in step 6.
8) Traverse the Tree constructed in step 7 and store the nodes in Dlgt[I].
9) Initialize E[I]:= E[1].
10) while ((E[I] <= N/2) || (O[I] <= N/2)) repeat step 11 to 15
11) Temp:= T_{sync}:= T_0 - ε
12) while(Temp > 0)
   a) Delegate node Dlgt[I] to communicate with BS.
   *b)*Temp: = Temp -1.
13) Call RESYNCHRONIZE ( ).
14) I:=I+1; Temp:= T_{sync}
15) If (E[I] == E[N/2]) then
   Set O[I]:= O[1] and transfer latest sensed data packet to corresponding PNbr in Odd network tree .
   Else If (O[I]== O[N/2]) then
   Set E[I]:= E[1] and transfer latest sensed data packet to corresponding PNbr in Even network tree .

RESYNCHRONIZE ( )

1) Temp:= T_{sync}
2) while (Temp >= 0) repeat steps 3 to 6
3) If (Temp == T_{sync})
   a) Generate two way communication links between each pair of nodes by the link key as:
   K:= ((HASH $\{k_1||k_2||\ldots\ldots||k_k\}$+
   LID[Dlgt[I]])*T_{BS})
   b) x:= (int | (T_0-HASH{n||LID[Dlgt[I]]})/2|).
   c) Set T_{BS}:= x.
   d) for I:=1 to N
      Set timer[I]:= x.
4) T_{BS}:= T_{BS} + 1.
5) timer[I]:= timer[I] + 1.
6) Temp:= Temp-1.

SECURE_LINK (Iso[I])

1) Initialize E = 0 and O = 0.
2) for I:=1 to N repeat steps 3 to 5.
3) for J:=1 to N repeat step 4 to 5.
4) If ((I!=J) && (($CKeys_{I,J}$>=Q) && (|LID_I-LID_J|<=D_0) && (RMAC[I]%2==0)))
   a) Add the IDs of the nodes in the neighbor list (Nbr[E]).
   b) E++
5) If ((I!=J) && (($CKeys_{I,J}$>=Q) && (|LID_I-LID_J|<=D_0) && (RMAC[I]%2!=0)))
   a) Add the IDs of the nodes in the neighbor list (Nbr[O]).
   b) O++

V. Analysis of the System Model

In WSN routing, energy and security are the three primary factors that should be kept in mind, before designing any protocol. It is a general myth that efficient routing, security and networks lifetime are seemed to be incompatible, but RSM trying to balance all these parameters. All these aspects are considered in development of RSM.

We are using four keys for communication; out of which two are static (i.e. ID of node and RTMAC) and remaining two are dynamic, which are computed by applying hash function; as given in the algorithm. These two dynamic keys are changed every time, when the network gets resynchronized. So, in RSM, if some node gets compromised, it will be identified in the next synchronization. RSM resynchronize entire network in the time less than T_0, where, T_0 is the time required to compromise any node by an adversary. Shorter is the value of T_0, higher is the security level of the network.

In some protocols, highest residual energy nodes are identified within the network and all data to the BS are routed through that node, which may causes problem, such as, unbounded delays. However, rather than checking nodes with highest residual energy, RSM delegate a node to communicate with BS on rotation basis, which is selected based on GANM [10], and we kept this rotation time less than T_0, so that, even if somehow an adversary is able to capture it, its effect could be minimized.

Energy is considered to be most important factor to enhance the life of the network. In RSM, wireless energy is provided to both the BS as well as to the SNs in the deployment area; moreover, communication link between two nodes is established only if the distance between these two nodes is less than D_0 and they satisfied the key criteria of Q composite keys. In RSM, all the exhaustive operations to set up the network are running at the BS, which saves energy of SNs a lot. In Q-composite scheme, there is no restriction of distance between two nodes and if communication link is established between two nodes, which are far away with each other, then much more energy is required to communicate with each other, as compared to RSM. Also, by increasing the value of Q in Q-composite scheme, more energy gets dissipated to match more number of keys, but RSM make it possible to enhance security with larger values of Q and match the keys at the BS itself rather than at the deployment area, as in regular fashion.

VI. Comparison with Existing Techniques

Table I shows the comparison of RSM with Q-Composite and Peer-to-Peer key magangement schemes which are considered as strong techniques for security existing till now.

TABLE I. Comparisons of RSM, Q-Composite and Peer-to-Peer

Attributes	*RSM*	*Q- Composite*	*Peer-to-Peer*
Feasible Key Set Size that can be Achieved	High	Small	Single Key is involved
Feasible Value of Q	Very High	High	Not Applicable
Security Level	Very High	High	High
Number of Keys Required to Establish Communication Among Nodes	3 Keys(1 static key in composition of Q keys and 2 dynamic keys)	Only 1 Key (composition of Q keys)	1Key (static)
Suitable Network Size	Very Large	Small	Large
Energy Consumption at SNs	Low	High	Low
Dynamic Keys	Yes	No	No
Network Scalability	Yes	No	No
Key Refreshment	Yes	No	No
Hot Sink Problem	Hot sink problem can never exist as the data aggregating node/sink node gets changed dynamically after short spans.	Hot sink problem may exist here, as the upstream communication path remains constant.	Hot sink problem may exist here, as in peer to peer communication also the upstream path remains constant.
Network Setup Time	Very High	High	Less

VII. Results and Discussion

Simulation is done using Matlab as the plotting software, as well as the calculation engine to plot results for the energy consumed by SNs with the increasing threshold distance value required to identify the physical neighbors, the effect of key set size on the time taken to establish a secure link, comparison between RSM and Q-Composite random key pre distribution and finally we plot the change in total flux intensity with respect to time, considering the total flux, considering both the flux due to solar energy and power beam.

The following are the simulation parameters considered for the implementation of the developed scheme:

- Deployment area is 100m X 100 m.
- The distance between the BS and the network is taken as 125m.
- Size of message is 80 bytes.
- Free space attenuation coefficient (E_{fs}) is 10 pJ/bit/m^2.
- Multipath attenuation coefficient (E_{mp}) is 0.0013 pJ/bit/m^4.
- Electronic power (E_{elec}) is 50 nJ/bit.
- Size of node ID 4 bytes.
- Size of MAC 8 bytes.

For realistic, our simulation uses the first order radio model as the communication model. Equation (1) and (2) represent

the energy dissipation, when a SN sends or receives an l -bit message.

$$E_{recieve} = l \times E_{elec} \quad (1)$$

$$E_{trans} = \begin{cases} l \times (E_{elec} + E_{fs} \times d^2), if d \leq \sqrt{\frac{E_{fa}}{E_{mp}}} \\ l \times (E_{elec} + E_{mp} \times d^4), if d > \sqrt{\frac{E_{fs}}{E_{mp}}} \end{cases} \quad (2)$$

Figure 4 shows the establishment of secure link among the randomly deployed static sensor nodes within the deployment area.

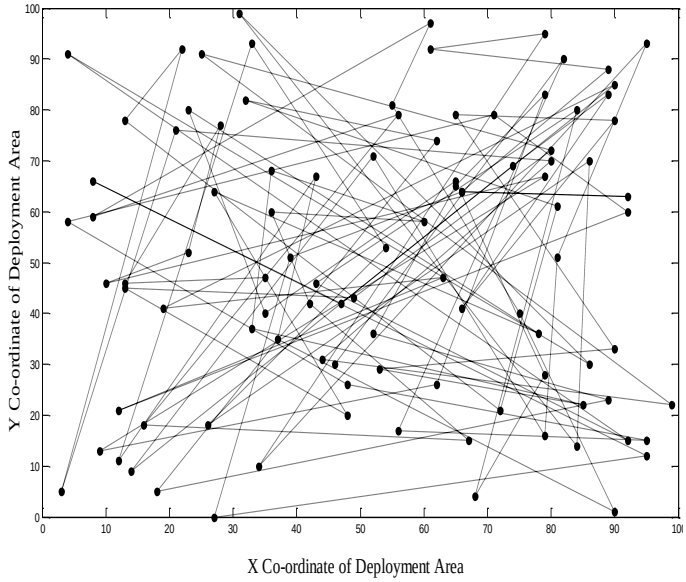


Figure 4. Deployment of SNs in the deployment area of 100m X 100m

Figure 5 shows that the times taken for the establishment of secure link decreases with the increasing key set size assigned to the sensor nodes. It is observed that the lifetime of a SNs decreases as the distance between two nodes increases as shown in Figure 6, and finally, Figure 7 shows the comparison between the traditional Q- composite random key pre-distribution technique and RSM. It can be observed that the energy consumption for the secure link establishment in Q-composite random key pre-distribution scheme gets increased with the increasing size of Q. However, in RSM, the energy consumption remains constant, since all the exhaustive tasks are managed by the BS rather than SNs itself.

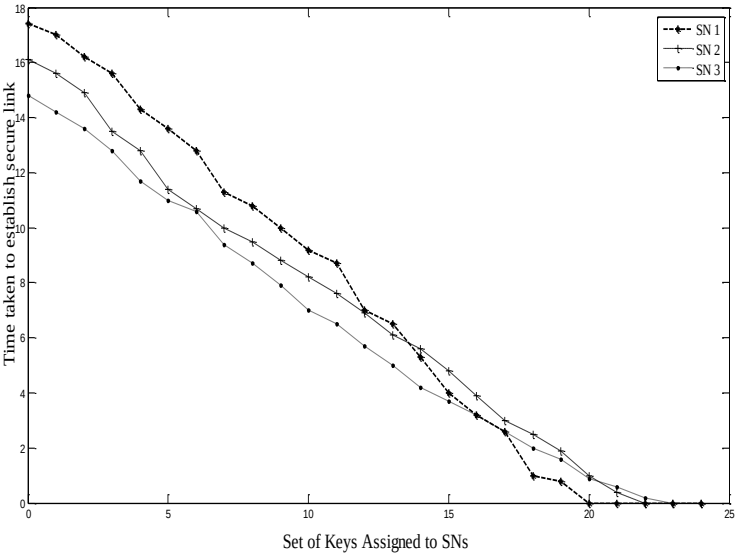


Figure 5. Effect of Key Set Size on Secure Link Establishment

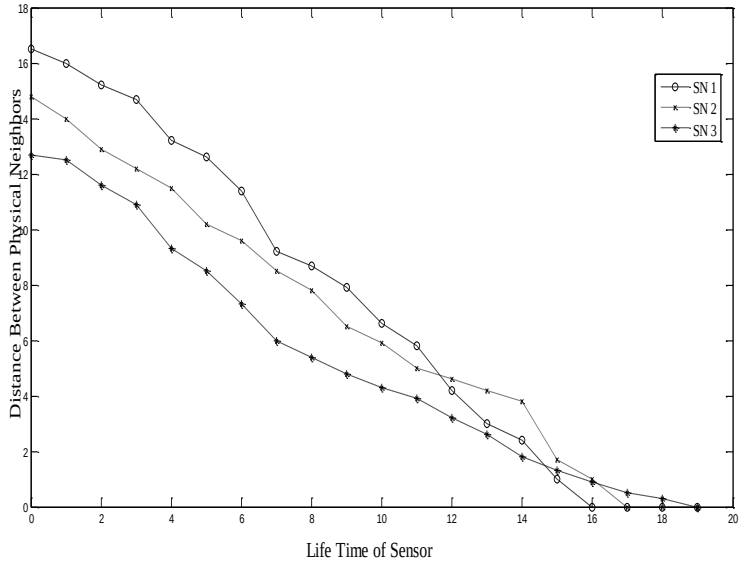


Figure 6. Effect of Distance between the Neighbors on the Lifetime of SNs

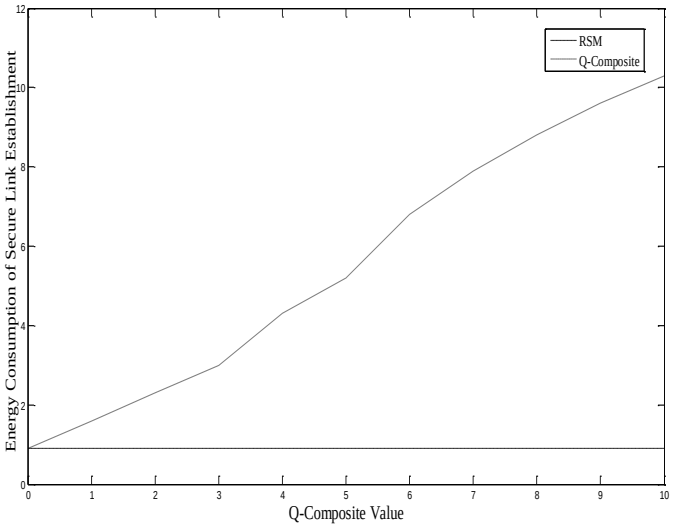


Figure 7. Comparison of Q-Composite Scheme with RSM

VIII. Conclusion and Scope of Future Work

In this article, we have presented a system model (RSM) for WSNs. The design of RSM is motivated by the observation of Q- composite scheme. To enhance the security, RSM keeps on changing keying materials every time network gets resynchronized.

Some of the advantages of RSM are as follows:

1) RSM ensures high energy at the BS end, rather than any assumption in fictions.

2) RSM provides ample amount of energy to the BS as well as to the SNs with a very low investment at the corporate end.

3) High energy by RSM may also ensure continuous sensing of data rather than periodic sensing as in general techniques of WSN.

As a future work, one can work upon some good robotics technology to introduce Mobile Sensor Nodes by taking advantage of high energy and homogeneous distribution of SNs even though they have been spread randomly from the BS as well as fault revoking can be done easily at the deployment end.

References

[1] Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: A survey”, Computer Networks Journal, Elsevier Science, Vol. 38(4):393-422, No. 4 pp 393–422, March 2002.

[2] Alan Mainwaring, Joseph Polastre, Robert Szewczyk, and David Culler. “Wireless sensor networks for habitat monitoring”, In First ACM International Workshop on Wireless Sensor Networks and Applications, 2002.

[3] Robert Szewczyk, Joseph Polastre, Alan Mainwaring, and David Culler. “Lessons from a sensor network expedition”, In First European Workshop on Wireless Sensor Networks (EWSN'04), January 2004.

[4] Ajay Kr. Gautam (Member IEEE), and Amit Kr. Gautam, “Accurate Localization Technique using Virtual Coordinate System in Wireless Sensor Networks”, International Journal of Recent Trends in Engineering, Vol 2, No. 5, November 2009

[5] S. A. Camtepe and B. Yener, “Key distribution mechanisms for wireless sensor networks: a survey,” Department of Computer Science, Rensselaer Polytechnic Institute, Tech. Rep. TR-05-07, March 23 2005.

[6] J.-P.Kaps, “Cryptography for ultra-low power devices”, Ph. D. thesis, at Worcester Polytechnic Institute, 2006.

[7] Heo, J., Hong, “Efficient and authenticated key agreement mechanism in low-rate WPAN environment”, International Symposium on wireless pervasive computing, pp. 1-5, Phuket, Thailand 16 – 18 January 2006, IEEE 2006.

[8] A Survey of Key Management Schemes in Wireless Sensor Networks.Yang Xiao, Venkata Krishna Rayi, Bo Sun, Xiaojiang Du, Fei Hu, and Michael Galloway; Computer Communications, Special Issue On Security On Wireless Ad Hoc and Sensor Networks.

[9] Key Management Building Blocks for Wireless Sensor Networks; Yee Wei Law, Jeroen Doumen and Marimuthu Palaniswami: The University of Melbourne, Australia, University of Twente, The Netherlands.

[10] Devendra Prasad, R. B. Patel, Ajay Kr. Gautam “A Reconfigurable Group Aware Network Management Protocol for Wireless Sensor Networks”, in proceeding of the IEEE International Conference on Advance Computing(IACC), Patiala, India, 6-7 March 2009.

[11] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: A survey”, Computer Networks Journal, Elsevier Science, Vol. 38, No. 4 ,pp 393–422, March 2002.

[12] Manik Gupta, D. Prasad, R.B. Patel, “FREEDOM: Fault Revoking and Energy Efficient Protocol for the Deployment of Mobile Sensor Nodes in Wireless Sensor Networks”, International Journal of Computer Science Engineering and Applied Research, Vol. 1, No. 1, pp 1-9, November 2010.

[13] T.J. Nugent and J.T. Kare “Laser Power for UAVs”, Laser Motive White Paper- Power Beaming for UAVs, NWEN, March 2010.

[14] http://technologyreview.com/energy/18836/page1/

[15] http://www.mit.edu/~soljacic/

[16] http://www.gizmag.com/intel-researchers-working-to-commercialise-wireless-power-sources/9858/picture/50110/

[17] W. Neil Johnson, Keith Akins, James Armstrong, Kwok Cheung, Glen Henshaw , Steven Huynh, Paul Jaffe,Matthew Long, Michael Mook, Michael Osborn, Robert Skalitzky, and Frederick Tasker, Jill Dahlburg, Michael N. Lovelette, David Huber, Mark Dorsey, Donald Gubser, Philip Jenkins, Scott Messenger, John Pasour, Robert Walters, Nathan Smith, Wayne Boncyk, Michael Brown, Robert Bartolo and Keith Williams “Space-based Solar Power: Possible Defense Applications and Opportunities for NRL Contributions” October 23, 2009.

[18] Glaser, P.E., "The Future of Power From the Sun," Intersociety Energy Conversion Engineering Conference (IECEC), IEEE publication 68C-21- Energy, 1968, pages 98-103.

[19] HerbertW.Friedman, “Near-Term Feasibility Demonstration of Laser Power Beaming” SPIE's International Symposium on Optoelectronic and Microwave Engineering Los Angeles, California January 25-27, 1994.

[20] Matt Welsh, Dan Myung, Mark Gaynor, and Steve Moulton. “Resuscitation monitoring with a wireless sensor network”. In Supplement to Circulation: Journal of the American Heart Association, October 2003.

[21] G.L. Duckworth, D.C. Gilbert, and J.E. Barger. “Acoustic counter-sniper system”, In SPIE International Symposium on Enabling Technologies for Law Enforcement and Security, 1996.

[22] S. Bandyopadhyay, E. Coyle, "An energy-efficient hierarchical clustering algorithm for wireless sensor networks", in Proceedings of the IEEE INFOCOM 2003, San Francisco, IEEE Computer Press, July 2003, pp. 1713 -1723.

[23] H. O. Tan, "Power efficient data gathering and aggregation in wireless sensor networks", SIGMOD Record, 2003, 32(4): 66 -71.

[24] Y. Tang, M. Zhou, X. Zhang, "Overview of Routing Protocols in Wireless Sensor Networks", Journal of Software, 2006, 17(3):410-421

[25] O. Younis, S. Fahmy, "Distributed clustering in Ad hoc sensor networks: A hybrid, energy-efficient approach", in Proceedings of the IEEE INFOCOM 2004. Hong Kong: IEEE Computer Press, 2004, pp. 630-640.

[26] M. Eltoweissy, M. Moharrum, R. Mukkamala, "Dynamic key management in sensor networks," IEEE Communications Magazine, Vol. 44, No. 4, April 2006, pp. 122- 130.

[27] A. S. Wander, N. Gura, H. Eberle et al., "Energy Analysis of Public-Key Cryptography for Wireless Sensor Networks," in Proceeding of the Third IEEE International Conference on Pervasive Computing and Communications (PERCOM), 2005.

[28] Brajendra Kumar Singh, Kemal Ertugrul Tepe "Feedback based real-time MAC (RT-MAC) protocol for wireless sensor networks", in Proceedings of the 28th IEEE conference on Global Telecommunications (GLOBECOM'09), 2009.

AUTHORS PROFILE

Devendra Prasad is an Associate Professor in the Department of Computer Science and Engineering, M.M. University, Haryana, India. Devendra Prasad is in teaching and Research & Development since 1996. He has supervised several M. Tech, and M. Phil Thesis. Devendra Prasad received his B.E. degree from Kumaon University, Nainital, India in 1995, M.Tech. Degrees from Kurukshetra University, Kurukshetra, India in 2007. He is currently enrolled as a PhD student in the Department of Computer Science and Engineering at M. M. University, Haryana, India. His research area is security, Fault tolerant and data dissemination, in wireless sensor networks.

Manik Gupta is a member of IAENG-International Association for Engineers and currently enrolled as a student of Masters of Technology in Computer Science in the Department of Computer Science, Maharishi Markandeshwar University, Mullana, Ambala, India. He had been awarded for Best Research Paper, in December, 2010. He also worked as Software Developer for one year after completing his Bachelors Engineering in Computer Science from University of Jammu. His research area is Security, Energy Efficiency, Fault Tolerance, Fault Revoking and Mobility in Wireless Sensor Networks.

Dr. R. B. Patel received PhD from IIT Roorkee in Computer Science & Engineering, PDF from Highest Institute of Education, Science & Technology (HIEST), Athens, Greece, MS (Software Systems) from BITS Pilani and B. E. in Computer Engineering from M. M. M. Engineering College, Gorakhpur, UP. Dr. Patel is in teaching and Research & Development since 1991. He has supervised 30 M. Tech, 7 M. Phil and 1 PhD Thesis. He is currently supervising 3 M. Tech, and 8 PhD students. He has published more than 100 research papers in International/National Journals and Refereed International Conferences. He had been awarded for Best Research paper many times in India and abroad. He has written numbers books for engineering courses (These are "Fundamentals of Computing and Programming in C", "Theory of Automata and Formal Languages", "Expert Data Structures with C," "Expert Data Structures with C++," "Art and Craft of C" and "Go Through C". His research interests are in Mobile & Distributed Computing, Mobile Agent Security and Fault Tolerance, development infrastructure for mobile & Peer-To-Peer computing, Device and Computation Management, Cluster Computing, Sensor Networks, etc.

Annotations, Collaborative Tagging, and Searching Mathematics in E-Learning

Iyad Abu Doush
Department of Computer Sciences
Yarmouk University
Irbid, Jordan

Faisal Alkhateeb
Department of Computer Sciences
Yarmouk University
Irbid, Jordan

Eslam Al Maghayreh
Department of Computer Sciences
Yarmouk University
Irbid, Jordan

Izzat Alsmadi
Department of Computer Information Systems
Yarmouk University
Irbid, Jordan

Samer Samarah
Department of Computer Information Systems
Yarmouk University
Irbid, Jordan

***Abstract*—This paper presents a new framework for adding semantics into e-learning system. The proposed approach relies on two principles. The first principle is the automatic addition of semantic information when creating the mathematical contents. The second principle is the collaborative tagging and annotation of the e-learning contents and the use of an ontology to categorize the e-learning contents. The proposed system encodes the mathematical contents using presentation MathML with RDFa annotations. The system allows students to highlight and annotate specific parts of the e-learning contents. The objective is to add meaning into the e-learning contents, to add relationships between contents, and to create a framework to facilitate searching the contents. This semantic information can be used to answer semantic queries (e.g., SPARQL) to retrieve information request of a user. This work is implemented as an embedded code into Moodle e-learning system.**

Keywords- Semantic Web; MathML; Adaptive e-learning; Folksonomies; Collaborative tagging.

I. INTRODUCTION

In recent years, we have observed the fast evolution of the Internet. The advent of the Internet has significantly enhanced availability of technical content, by making millions of documents available on-line. The introduction of sophisticated Learning Management Systems (LMSs), such as Blackboard and Moodle, has widely increased the opportunities for a geographically diverse population of students to access educational programs.

Recent statistics (Whitehead, 2009) indicate that more than 3.2 million people are involved in some form of online education. Modern LMSs are designed to provide sophisticated organization of the various components of a course.

The literature provides several proposals aimed at adding semantic meaning/markups to different parts of the online course (e.g., (Bateman, 2007; Bateman, Brooks, Mccalla, & Brusilovsky, 2007; Stojanovic, Staab, & Studer, 2001)). We strongly believe that this can have a profound impact in facilitating the use of LMSs.

Partitioning the course page into logically related components will allow the student to select specific parts of the learning contents (Enagandula, Juthani, Ramakrishnan, Rawal, & Vidyasagar, 2005). This would minimize the learning time and direct the user to focus on the parts that s/he is interested in.

Following a research trend that has developed in the last years, we explore the introduction of technologies drawn from the field of the semantic web to aid in the process of semantic searching of LMSs. We rely on three principles: ontologies, RDFa annotations, and folksonomies.

Ontology is one of the main elements of the semantic web. An ontology is used to formalize the concepts in a domain and to describe their mutual relationships (Kim, Kim, & Park, 2005). Ontologies can be used to build adaptable e-learning systems: using ontologies, human or software agents can automatically understand the meaning of the e-learning content and process it appropriately (Kim et al., 2005). For example, ontologies enable semantic-based search of e-learning content (Chang, Ham, Moon, Choi, & Cha, 2007).

The process of collaborative tagging produces a set of tags that can be used to describe a resource (Al-Khalifa & Davis, 2006). This tagging process, made popular by sites like del.icio.us, can be employed to generate grassroot ontologies, commonly called folksonomies. Folksonomies add semantics to resources through a social markup process, enabling the use of tag clouds to describe entities and determine relevancy. The use of folksonomies allows the addition of semantics without the aid of individual manual indexers or automated keyword generators (Al-Khalifa & Davis, 2006).

The tagging process does not create a strict taxonomy for objects, but it enables the user to employ his own keywords to categorize objects. Gruber (Gruber, 2005) mentioned that, in

collaborative tagging, there is no explicit links between entities, and no standard form is used for presenting the data. This limitation requires introducing a formal ontology along with folksonomies to formalize tags.

Resource Description Framework (RDF) data can be embedded inside XHTML as RDFa ("RDFa Primer: Bridging the Human and Data Webs, http://www.w3.org/TR/xhtml-rdfa-primer/." 2010). The RDFa annotations are used for making parts of the web page foldable into a more detailed information (i.e., according to the vocabulary and the relations of the used RDF). Standard extractors for RDFa can be used to retrieve the annotations in the web page (e.g., ("RDFa Distiller and Parser, http://www.w3.org/2007/08/pyRdfa/." 2010; "rdfquery, RDF processing in your browser, http://code.google.com/p/rdfquery/." 2010)).

Mathematical contents represent a particular challenge for searching the contents in e-learning - e.g., formulas, mathematical symbols, and abbreviated function names. The W3C recommendation for encoding mathematics on the web is called MathML("Mathematical Markup Language, http://www.w3.org/Math/." 2010). There are two types of MathML: presentation MathML and contents MathML. Presentation MathML describes the visual appearance of the mathematical expression by using 2-dimensional layout and formatting of the mathematical expression. On the other hand, content MathML encodes the meaning or the mathematical semantic of the expression.

The encoding of mathematical expressions using presentation MathML can help in capturing the conceptual structure of the mathematical expression. This can help in having a common notation which can be used for the underlying search of the mathematical contents.

Adding semantics to the e-learning contents on the web can provide several benefits to the users:

1. Provide a more accessible contents for the blind and visually impaired individuals, as the contents can be read by screen readers.
2. Easier searching for technical and educational materials.
3. The contents can be explained (e.g., using other students annotations or from relations between other contents).
4. Help people with learning disabilities in navigating the e-learning contents (e.g., providing information about a concept in the navigated e-learning as a tooltip).

In this paper we propose a novel framework for adding semantic information to the e-learning contents. The e-learning contents can be highlighted, and annotations and tags can be added by the student. A defined ontology is used to categorize user annotations and tags. Another part of the framework is the automatic insertion of semantic information to the mathematical equation. The mathematical contents in e-learning are annotated by semantic information using RDFa.

II. BACKGROUND

A. Searching Mathematical Contents

A review of related work shows that developing searchable mathematical expressions imposes a number of requirements. An extensive literature exists, dealing with various aspects of this problem. Some of these relevant studies are discussed next.

Munavalli and Miner (Munavalli & Miner, 2006) introduce a math aware search engine which search mathematical contents. The system analyzes MathML mathematical expressions into text math fragments. In this system the user enters the math query using graphical equation editor.

According to Youssef (Youssef, 2006) the mathematical search purpose is: 1) allowing the user to perform fine grained search on mathematical data 2) Allow users to enter the math query naturally and easily using the symbols and notations applied by mathematicians and scientists.

In another work Guidi and Schena (Guidi & Schena, 2003) introduce a math query language for RDF metadata repository called MathQL. Asperti et al. (Asperti, Padovani, Coen, & Schena, 2001) presented HELM, a framework that uses XML technology for building structures contents in a logical manner. The purpose is to use the system as a library for indexing and retrieving mathematical documents.

Altamimi and Youssef (Altamimi & Youssef, 2008) presented a math query language that enable users to express their information needs intuitively yet precisely. The new math query language offers an alternative way to describe mathematical expressions that are more consistent and less ambiguous than conventional mathematical notations. In addition, the language goes beyond the Boolean and proximity query syntax found in standard text search systems. It defines a powerful set of wildcards that are deemed important for math search. These wildcards provide precise structural search and multi-levels of abstractions.

Hijikata et al. (Hijikata, Hashimoto, & Nishida, 2009) presented a search engine for MathML objects using the structure of mathematical formulas. The system makes the inverted indices by using the Document Object Model (DOM) structure of the MathML object. It also proposes three types of indexes:

One type is constructed from some paths of the DOM structure and expressed in XPath.

The second type is constructed by encoding the nodes in the same level in DOM structure.

The third type is a hybrid method from the other two types.

B. Semantic Web and E-learning

An approach based on a student model and an ontology in order to personalize an eLearning system is proposed in (Gomes, Antunes, Rodrigues, Santos, & Barbeira, 2006; Pah, Stoica, Cacovean, & Popa, 2008). The ontology is used to map the student knowledge to course concepts allowing a better access to her/his progress and to adapt contents and navigation structure to a particular student.

Soylu et al . (Soylu, Kuru, Wild, & Mdritscher, 2008) proposed a framework for harvesting learning objects from web-based content. The framework is based on a lightweight application profile and a microformat for learning objects using well-known learning object metadata standards in order to address mainly interoperability and reusability of learning content. They also describe a web service to extract learning objects from different web pages, and provide an SQI target as a retrieval facility using SPARQL and XSL transformation.

Henze et al. (Henze, Dolog, & Nejdl, 2004) proposed an approach for personalization of e-Learning systems by using semantic web and shows how the semantic web resource description formats can be utilized for automatic generation of hypertext structures from distributed RDF annotations. Several ontologies have been utilized corresponding to the components of an adaptive hypermedia system: a domain ontology (describing the document space, the relations of documents, and concepts covered in the domain of this document space), a user ontology (describing learner characteristics), and an observation ontology (modeling different possible interactions of a user with the hypertext).

In another work Stojanovic et al. (Stojanovic et al., 2001) presented an approach for implementing the eLearning scenario using Semantic Web technologies. The goal of the proposed framework is to provide a flexible and personalized access to these learning Materials, the proposed eLearning scenario exploit ontologies in three ways: for describing the semantics (content) of the learning materials (this is the domain dependent ontology), for defining the learning context of the learning material and for structuring the learning materials in the learning courses.

C. Collaborative Tagging and E-learning

In another work Stojanovic et al. (Stojanovic et al., 2001) presented an approach for implementing the eLearning scenario using Semantic Web technologies. The goal of the proposed framework is to provide a flexible and personalized access to these learning Materials, the proposed eLearning scenario exploit ontologies in three ways: for describing the semantics (content) of the learning materials (this is the domain dependent ontology), for defining the learning context of the learning material and for structuring the learning materials in the learning courses.

Collaborative tagging (also known as folksonomy , social classification, social indexing, and social tagging) is a system for collaboratively creating and managing tags to annotate and categorize contents. The term folksonomy was first used by Thomas Vander Wal in a discussion on an information architecture mailing list. It is a combination of "folk" and "taxonomy". (Folks => done by people, Taxonomy => classification of items into groups) (Peters & Becker, 2009).

One of the main objectives of collaborative tagging is to make a collection of information increasingly easy to search, discover, and navigate over time. Thus, collaborative tagging has a main role in developing semantic web and information retrieval systems.

Collaborative tagging became popular on the Web around 2004 as part of social software applications such as social bookmarking and photograph annotation. Tagging, which is one of the defining characteristics of Web 2.0 services, allows users to collectively classify and find information. Some websites include tag clouds as a way to visualize tags in a folksonomy (Mathes, 2004; Peters & Becker, 2009).

Bateman et el. (Bateman, 2007) outline their experience with applying collaborative tagging in e-learning systems to supplement more traditional metadata gathering approaches. They state that metadata is best created if it focuses on a particular goal, is contextualized to a particular user, and is created in an ambient manner by observing the actions and interactions of students in learning environments. Consequently, collaborative tagging seems to be a leading method by which we can collect learner-centric metadata. Using tags enables useful resource organization and browsing techniques, and the viewing of tags used on a webpage can give a learner some idea of its importance and its content, it may help a learner in finding the exact point of interest within the page.

According to (Bateman, 2007) collaborative tagging systems have potential to be a good fit with e-learning systems, because of the following:

1. Learning managements systems currently lack sufficient support for self organization of learning content.
2. Collaborative tagging has potential to further enrich peer interactions and peer awareness centered around learning content.
3. Tagging, by its nature is a reflective practice, which can give students an opportunity to summarize new ideas, while receiving peer support (through viewing other learners' tags; tag suggestions).
4. The information provided by tags provides insight on learner's comprehension and activity, which is useful for both educators and administrators.

Bateman et el. (Bateman, Brooks, & Mccalla, 2006) have suggested merging the use of ontologies with collaborative tagging through a new approach and they have called it CommonFolks. This approach should help to reduce the effort required in the production of useful metadata, while maintaining the expressiveness inherent in lightweight ontologies, thus opening the door to a better quality of metadata and authoring by those not traditionally involved in metadata creation.

In (Bateman, 2007), Bateman states that the interest in investigating the use of tagging in e-learning revolves around several points of interest. Some of these points can be summarized as follows:

- Organization and Annotation: learning management systems (LMSs) do not properly support self organization and annotation of learning content. However, students usually use a number of organization and annotation techniques. These include writing notes, creating marginalia in

books, highlighting text, and bookmarking pages. How can we make these traditional capabilities available to students using digital learning materials? Are there methodologies more suitable for such tasks in an online environment?

- In this regard, tagging and note-taking can be linked together, since tags represent an aspect to be used in the tagger's recall process. Tagging provides a straightforward method for self-organization and most tagging interfaces provide fields for longer sentence based notes to be taken (Bateman, 2007).
- Metadata Collection: according to the learning object paradigm, online educational content can be collected, aggregated, and packaged for delivery to learners. However, the major difficulty in accomplishing this vision is the lack of meaningful metadata describing learning objects. Collaborative tags can be considered as a form of metadata, which could supplement the needs for detailed learning object descriptions (Bateman, 2007).
- Knowledge Gain: tagging gives students the chance to summarize new ideas while receiving peer support. Tagging can be considered as an action of reflection, where the tagger can summarize a series of thoughts into one or more tags, each of which stands on its own to describe some aspect of the resources based on the tagger's experiences and beliefs (Bateman, 2007).
- Pedagogical Reflection: in e-learning there is a lack of direct interaction with learners that inform instructors about the understanding of new concepts by them. Collaborative tags, created by learners can help instructors in predicting their students' progress e.g. tags that are out of context could represent a misconception (Bateman, 2007).

III. METHODOLOGY

We have developed an ontology and infrastructure for the annotation of the course content in Moodle. The proposed ontology consists of activities and entities involved in the e-learning process (e.g., lesson, quiz, and student). The ontology defines the relationships between different classes in the e-learning ontology (e.g., an assignment is solved by a student). The ontology can be collaboratively extended by the users of the system by introducing new classes that are associated with other pre-defined classes in the ontology.

The ontology is structured in two parts: an upper ontology and a bottom ontology. Intuitively, the upper ontology is used to describe the overall structure of a generic course in Moodle. The upper ontology describes the components of a course that are common across all courses; as such, this is a static ontology. The bottom ontology describes the components that are course-specific and domain-specific. It is an inherently dynamic ontology.

The fundamental idea is to get the help from students in the class to introduce a taxonomy for a specific online course. Students participating in the same class are encouraged to annotate the course contents and share such annotations with other students. The benefit from integrating the two approaches (i.e., bottom-up and top-down) is to create a comprehensive semantic structure of a course, which can be used by students to understand the content of the course and facilitate the semantic searching of the contents.

For example, a student who reads the online course material can add annotations to explain specific parts of the lecture notes (e.g., describe a figure or a mathematical formula) in the process of studying the material. These annotations can be used by other students; the tagging of a figure can, for example, provide a textual documentation to look for the figure in the online course.

We use the current web page in Moodle (e.g. quiz, lesson...) to select the ontology class for the highlight. We will suggest the category (class) for the user tags using synonyms and other user's tags. WordNet synonyms are used to infer concepts related to the students' tags. Synonyms are used to group tags as instances of the same ontology concept.

An equation editor is used to create mathematical equations encoded in presentation MathML. The mathematical contents are then annotated using RDFa with other set of semantic information (i.e., link, value, description, and category) which will be explained next in the paper. The main goal is to provide the user with different search schemes (i.e., keywords, equation structure, equation category, and equation uses).

A. Annotation System Use Scenario

To understand how the system operates, let us consider a student, who wants to search an online course in Moodle to review lecture notes for a math course. (e.g., a mathematical formula is annotated with a link that explains it, or a paragraph in lecture notes is tagged as a potential quiz question).

The user can use the category of the tag to find its relation with other course materials (e.g., a specific part of the lecture notes is tagged as an explanation for a mathematical formula in the notes the student is browsing). When the student encounters graphical content in the page he could look for the annotations associated with it to get explanation about its contents and how it relates to the content of the lecture notes (e.g., a graph for mathematical equation). The use of the taxonomy provided by the ontology can be seen also when the student wants to access specific parts of the page (e.g., geometry → area of a triangle).

B. The Top Ontology

We built different classes in the top ontology (Figure1) on the basis of various components of Moodle (e.g., quiz, homework). The Activity class is the super class of all the activities (Lesson, Glossary, Assignment, Wiki, Forum, and Resource) the student can perform in the e-learning system. Smaller components within the e-learning activity are defined (e.g., Post in a forum, Question in a quiz) to give a category for smaller components of the annotated information.

The interactions between the different components in the LMS are presented as relationships between different concepts

(e.g., Quiz ConsistOf Questions, Assignment isSolvedBy Student). The ontology concepts and the relationships between them are used to describe the structure of a course and as main concepts to be refined by the students with their bottom-ontology.

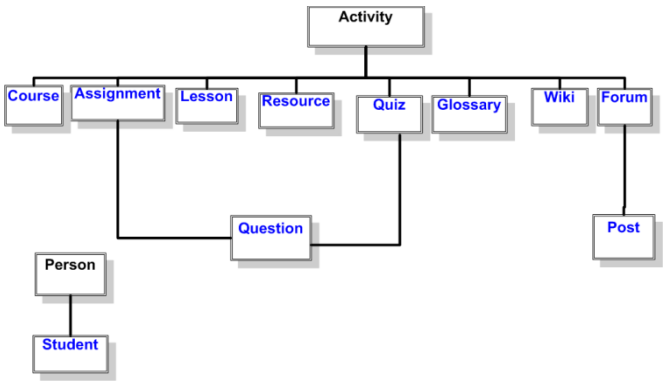


Figure1- Graph of a Fragment of the Upper Ontology

C. The Bottom Ontology

The user can tag a course component (e.g., a text fragment, an image, a formula) in the online course. The predefined set of upper ontology concepts can be used as the categories for the current annotated resource, or the user can define new concepts that are used to refine and expand the upper-ontology. Each tag introduced by the student is presented as an instance of the concept selected or defined by the student.

For example, a student can define a new concept (e.g., trigonometry formula) to annotate an equation in a math course lesson. The annotated equation will become an instance of this new concept. The user can add a tag or a set of tags (e.g., math, formula) to the annotated equations which are linked to the selected class via the has-tag relationship. As another example, the student can annotate a post in a forum as an explanation for a specific lesson; in this case, the annotated post will be treated as an instance of the Post class, and the connection between the Post and Lesson classes will be identified using the relationship defined between the two classes.

D. Generating the Bottom Ontology

Collaborative tagging among different students of an online course can help in extending the upper-ontology and create new connections between different learning contents. Users are allowed to generate tag clouds, by freely tagging the content encountered in Moodle.

The tag cloud for a component of a Moodle class will provide domain specific classification of the elements of the course as well as the level of ``agreement'' across students concerning the description of the course entities. The tags frequency can be used as an agreement measure (based on the algorithm in (Heymann & Garcia-Molina, 2006)). The most agreed upon tags will flow into the bottom ontology, as new concepts. Using the proposed system (Figure2) the user can highlight a resource in the online course (e.g., text fragment) and then the tool interface can be used to select then category (concept) of the annotated resource or to create a new concept.

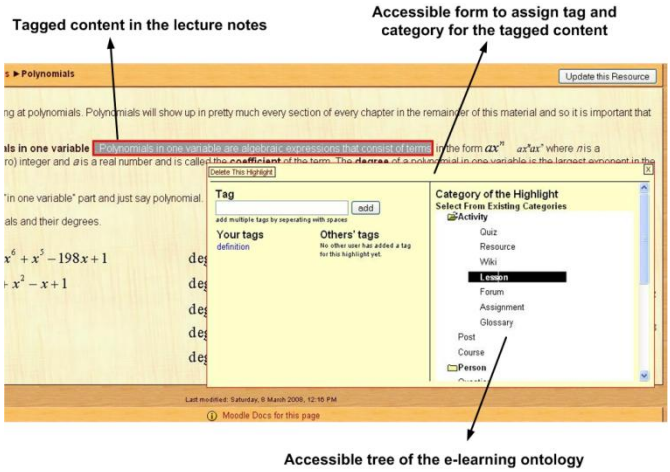


Figure2- A Screenshot of the Proposed System

The construction of the tags and the expansion of the ontology is user-driven. The user can select the concept for the highlighted content and assign a tag for the content. The user can navigate the tags (i.e., perform keyword search) in order to retrieve the set of items associated with that tag.

E. A Mathematical Search Use Scenario

The e-learning content creator uses an equation editor inside the learning management system to generate mathematical contents. The content creator adds other information (i.e., description) to the mathematical contents. Other information is also added automatically to the mathematical contents (i.e., category, value, source, and link). The created contents are then annotated with this information using RDFa.

A student who is interested in searching for integral examples in lessons only can then search the proposed system by using the keyword "integral" and tick the checkbox lessons. The results can be obtained using extracted RDFa annotations. The student can also specify a specific type of integrals to search for by using the equation editor to enter a specific integral and search for it in specific place in the e-learning contents (e.g., previous quizzes). The system returns a URL that points the user exactly where the searched mathematical formula is located in the e-learning content.

F. Mathematical Contents with Semantic Information

The equation editor can be used by the mathematical contents generator in e-learning to create a mathematical equations encoded in presentation MathML. The presentation MathML can be visually rendered by the machines but they are not understandable by them (Asperti et al., 2001). Presentation MathML does not provide adequate semantic information (Kohlhase & Sucan, 2006).

The encoding of mathematical expressions using common RDFa annotations can help in capturing the conceptual structure and remove any ambiguity and inconsistency related to the use of presentation MathML in encoding the mathematical expression. This can help in having a common notation which can be used for the underlying search of the mathematical contents.

The goal of this part of the system is to provide the student with a mathematical content with semantic encoding. This can lead to a more accurate search of the mathematical contents in e-learning. Using this scheme the system will search for the agreed vocabulary used in RDFa.

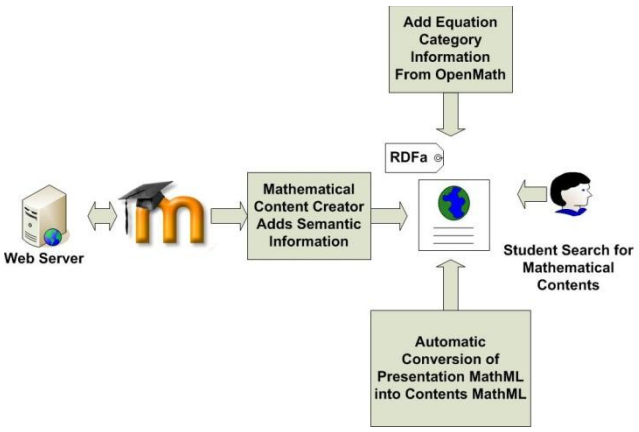


Figure3- The System for Adding Semantic Information to the Mathematical Contents.

- The proposed methodology consists of the following major phases (see Figure3):
- Mathematical content creator in e-learning uses an equation editor to build a mathematical equation. The equation is encoded using presentation MathML.
- The contents creator adds other parts (e.g., description) to the content as RDFa annotations.
- In order to allow the user to do a fine grain search on the page level the presentation MathML is annotated using RDFa.
- The RDFa extractors will be used next to extract the information about the mathematical expressions along with the URI of the mathematical expression.
- The extracted RDFa annotations can be compared then with the user query to retrieve mathematical contents with closer similarity.

We will have a place where the mathematical vocabulary (i.e., classes and properties) located, and use this vocabulary to annotate the content MathML. The content MathML from the previous example can be annotated as follows:

```
<div xmlns:m="http://example.com/math/vocab#">
  <div about="/math101/lesson1/powerEquation">
    <math>
      <apply >
        <power property="m:binary-operator"/>
          <apply>
            <plus  property="m:unary-operator"/>
                <ci property="m:identifier">x</ci>
                    <cn
                    property="m:number">3</cn>
            </apply>
          <cn property="m:number">2</cn>
        </apply>
      </math>
    </div>
</div>
```

Using this annotation scheme the user can find the exact position of the needed mathematical expression in the web page.

TABLE1- The Added RDFa Annotations to the Presentation MathML.

RDFa Field	Meaning
Anchor link	A link to the equation in the web page (e.g., www.example.com/lesson/lesson3.htm#equation2)
Value	The equation encoding in Content MathML
Source	The activity in e-learning in which the equation is found (e.g., Lesson)
Category	The mathematical equation category according to OpenMath (e.g., cosine is a trigonometric function)
Description	The uses of the mathematical equation (e.g., distance between 2 points)

The following is a description of different fields in RDFa annotation (see Table1). The anchor link points where is the mathematical equation located in the web page. This link is composed of the URL address of the e-learning web page along with anchor ID for the mathematical equation in the web page. The anchor ID is a serial number automatically generated for each annotated mathematical equation in the current web page. The link can be used for a direct access to the location of the mathematical equation in the web page.

The source of the mathematical equation is extracted from the URL address of the current component where the mathematical equation was added (e.g., lesson, quiz, wiki). This information can be used later on to perform a more directed search by searching only in specific sections of the e-learning contents (e.g., search for quadratic equation in lessons only).

The value represents the content MathML encoding of the mathematical equation. Kohlhase and Sucan (Kohlhase & Sucan, 2006) mentioned that usually math web search uses content MathML as its basis. This encoding is obtained using a converter which converts from presentation MathML into contents MathML (for details see (Doush, Alkhateeb, & Maghayreh, 2010)). The use of content MathML can help in

achieving a more structured unambiguous searching of the mathematical contents.

The category field represents the equation classification according to the OpenMath content dictionary ontology. The OpenMath ontology can be used to infer the mathematical equation category according to the classes and relations used in the Content Dictionary ontology. The reasoning can be performed according to the structure of the content MathML of the mathematical equation.

The use of mathematical equation is saved in the description field. Why scientists and mathematicians use the mathematical equation is entered by the mathematical contents creator according to the common uses of the equation in the field of study. The description can be helpful in cases where the user don't know the structure of the mathematical equation but s/he knows what the equation can be used for.

G. Indexing of Mathematical Equations

Indexing is the process of collecting, analyzing, and storing data to facilitate fast and accurate information retrieval (finding relevant documents for a search query). Without an index, the search engine has to scan all of the documents in the corpus, which would require a considerable amount of time (Baeza-yates & Ribeiro-Neto, 1999).

The indexing of the mathematical equations (see Figure3) needs to allow the user to perform different types of searching:

- When the user is searching for a mathematical expression s/he usually are looking for equations with similar structure and do not care about the variables used in the mathematical equation. This means that the retrieval system needs to return the mathematical equation in the search results even if it does not contain any of the variable names used in the search. For example, assume that the user query was "X = Y + Z - W" then the search should be performed on the equation "term = term + term - term".
- In some cases the user needs to retrieve the mathematical equation with exact term or variable names. In this case we have to give the user the option to search for the mathematical equation with exact variable names.
- Sometimes the user does not know the structure of the mathematical equation, but s/he knows what is the common name of the equation. Searching for a common name of the mathematical equation (e.g., Newton equation) needs to retrieve the mathematical equations with a structure similar to a defined equation with that common name.
- In some cases the user is interested in searching for equations under a large category (e.g., polynomial equations). The classification of equations into these categories relies on a mathematical ontology called OpenMath Content Dictionaries.

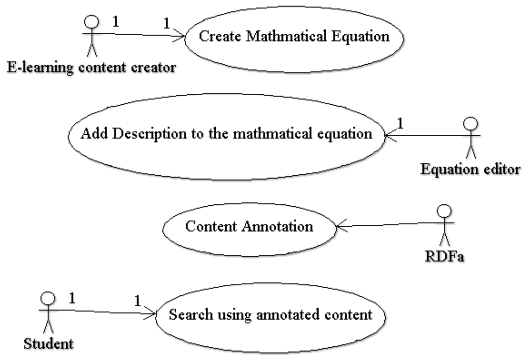


Figure4- A Use Case Diagram for the Indexing of Mathematical Equation.

The indexing of mathematical equations need to be performed by giving weights to the different operations applied to the equation terms (e.g., +, -). Order of evaluation is used to evaluate the weighting of terms in the mathematical equation.

The order of the operation in the equation affects the weight value used for the indexing as the order of the operations is an important criterion used when searching for the mathematical equations. The weight represents the value to be searched for when we are looking for a mathematical equation with similar structure and order of operations.

Another weight for the same equation can be calculated by including the variable names in the weight calculations. This weight value can be used for exact equation matching between the query and the indexed equations.

The user should have the ability to search for mathematical equations with the exact terms. The search needs to identify similarity distance between the searched equation and the indexed equations. The number of matched operations with correct order can be used as a weight to perform a partial matching for the mathematical operation.

In our context the documents represent a collection of mathematical expressions, and the query represents the user need to retrieve one or more expressions with certain characteristics of interest.

In textual documents, words and phrases can be used as index terms, but for mathematical expressions, what will be the index terms? From the representation of mathematical expressions (described in earlier sections) we can see that the index terms can involve the operators in the expressions, variable names, and other words used to describe the meaning or the uses of the mathematical expression.

The Vector Space Model (VSM) of information retrieval (Baeza-yates & Ribeiro-Neto, 1999) can be used in our context. In this case, both of the mathematical expression and the query are represented as vectors where each vector consists of a set of index terms' weights. The weight of an index term in the vector

space model depends on its frequency in the collection of the documents (for more details see (Baeza-yates & Ribeiro-Neto, 1999).

To decide whether a given mathematical expression is relevant to a given query or not we have to measure the similarity between the query and the expression. In the VSM the similarity is represented by the cosine of the angle between the two vectors representing the expression and the query. The system must return as a response to a query a list of expressions ordered according to their similarity with the query, with the highest comes first. This is similar to what is done in traditional information retrieval systems.

The question to be addressed is how the user will enter his query? A simple solution is to use a simple textbox and allow the user to use a flag followed by a set of index terms that can be used in expressions.

For example the user can enter the following query:

Math: x y z + *

This query indicates that the user is looking for an expression that contains some or all of the above terms.

The only thing we need here during indexing is a procedure that scans the representation of the expression and tokenizes it where each token can be a variable name or an operator. These tokens along with any other words on the description of the expression represent the set of the index terms. The weight of these terms can be evaluated using the same method used in the VSM.

The user can enter the query directly as described above or we can allow him to use an equation editor to enter the query and then a procedure is used to tokenize it as it was indicated earlier. Then the weight of the terms in the query can be evaluated using the same method used in the VSM. Once we have the two vectors it is easy to calculate the similarity to determine whether the given expression is relevant to the query or not.

In this method the following query is identical to the one shown in the above example

Math: x + y * z

This is because the method does not consider the order of operator evaluation in the expression. If we want to consider this case then there are two possible solutions. The first one will require some changes in the way of evaluating the weight for the index terms. This is due to the fact that the current version of the VSM considers the frequency of the index terms in any document not its order. Using this solution the index terms with the same order of the query will have larger similarity than other index terms.

The second solution is simpler and described next. The index terms can be either variable names or pairs where each pair consists of an operator and a number representing its order of evaluation.

For example, given the following expression

x + y * z

The index terms will be x, y, z, (+,2), (*,1)

Now how the user will enter the query? There are two choices similar to the ones highlighted above. The first choice is to use a textbox and ask the user to enter the operators in his query as pairs.

For example if the user enters the following query

Math: (+,1) (*,2)

Then we understand that he is looking for an expression where the first operator evaluated is + and/or the second operator is *. In this case the similarity between this query and the above expression is zero.

The above explained method may not be easy to use and it may lead to many mistakes, as it is hard for some users to write their queries in the paired-format according to the order of the evaluation. Another solution for entering the query is to allow the user to write the query in it is normal format using a text or equation editor and have the parser generate the paired-format.

H. Searching for Mathematical Equation

Using the proposed searching scheme the users can apply more semantic search. For example the user can search for algebraic equations in quizzes only. Also, by applying the semantic queries the system can answer specific user's queries, for example the user can search for the lessons with Newton's equations.

Searching in the system can be performed in two modes:

- Searching using the search box by turning on the math search for the equations (e.g., Math: polynomial equations, and then the user specifies using a check box that the search is on the lessons).
- Searching by using the mathematical equation editor to enter the mathematical expression in the query. This query is converted into weighted query which is then searched for in the index. The user can specify if s/he is interested in partial matching of the query, and also if s/he is interested in exact matching with the terms and variable names of the equation.

SPARQL (Prud'hommeaux & Seaborne, 2008) is a W3C recommendation language developed in order to query RDF data. A simple SPARQL query is expressed using a form resembling the SQL SELECT query:

SELECT B FROM UWHERE P

Where U is the URL of an RDF graph G to be queried, P is a SPARQL graph pattern (i.e., a pattern constructed over RDF graphs with variables) and B is a tuple of variables appearing in P. Intuitively, an answer to a SPARQL query is an instantiation of the variables of B by the terms of the RDF graph G such that the substitution of the values to the variables of P yields to a subset of the graph G.

The following SPARQL query modeling this information:

SELECT *

```
FROM  <RDF data>
WHERE  {
        ?resource hasLink ?Link .
        ?resource hasValue ?Value .
        ?resource hasSource ?Source .
        ?resource hasCategory ?Category .
        ?resource hasDescription NewtonEquation .
        }
```

Could be used to search for annotations of a resource containing Newton's equation. The link (anchor link) of the resource will be returned to allow the user directly to access the page that contains it.

It should be noticed that the resource description and the value are determined from the user query. More precisely, the searching process is achieved in two phases. In the first phase, the user enters keywords to be searched. These keywords are used to search the description and the value of the resource (e.g., equation) to be retrieved. While in the second phase, the value and the description of the resource are used in a SPARQL query to retrieve the annotations of the resource.

The information obtained in the second phase is extracted from RDF document that matches the query. The RDF document represents the fields (as shown in Table1) retrieved using RDFa extractors.

IV. System Design

The first prototype of the system has been realized by including our ontology and annotation components (accessible OATS) into Moodle. In order to make the annotation process possible, our tool is included inside Moodle in a place visible to all web pages of the course. An accessible modified version of the Open Annotation and Tagging System (OATS) (Bateman, Farzan, Brusilovsky, & McCalla, 2006) has been developed.

The new OATS has been made accessible, by replacing all user interactions with keyboard controls and aural messages to perform the highlighting and tagging. This has been applied in order to make the system accessible for people with disabilities. The e-learning upper-ontology is opened by accessible OATS and is available to the user for the annotation process.

The e-learning ontology provides the concepts taxonomy needed to classify the information (e.g., math → algebra → addition quiz) in the e-learning system. The relations defined in the ontology can be used to infer new information about the learning materials. For example, the has_answer relation between a quiz and a lesson can be used to imply a link between the instances of these classes.

The system structure is depicted in Figure5. The implementation relies on several software components:

- Connector between Apache and Tomcat: to integrate Apache and Tomcat mod-proxy-ajp, a connector to connect Tomcat and Apache Httpd, is used.
- REST web services (Servlets and AJAX): a REST architecture is used to update the tagging and annotation system database.
- JENA API: this semantic web API is used to maintain the e-learning ontology.
- WAI ARIA: the different classes for Moodle ontology are displayed to the students using an accessible tree structure built using WAI ARIA.

We propose also adding an open source equation editor (e.g., MathCast ("MathCast, an open source equation editor. http://mathcast.sourceforge.net/home.html," 2010)) to Moodle. The equation editor is used to add mathematical equations to the e-learning contents with presentation MathML encoding.

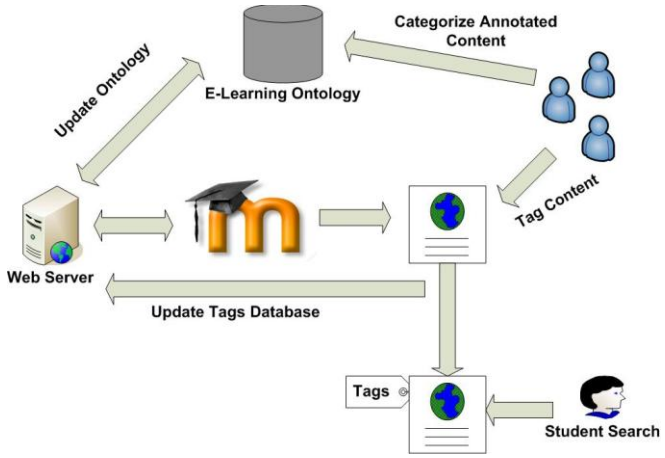


Figure5- The System Architecture.

V. Conclusion and Future Work

The proposed system uses collaborative tagging performed by the students in an online course to help other students. It relates different course components to each other, clarifying course materials by offering additional explanations, and providing connections between different course components. The context of the highlighted text can be used to select the ontology class for classifying the highlighted contents.

We presented investigation aimed at improving the process of searching for mathematical contents within e-learning systems. The goal of the proposed solution is to identify exactly where the mathematical expression is located in the web page by using RDFa annotations.

Using the proposed system the presentation MathML in the e-learning contents are automatically annotated with RDFa annotations using a pre-defined vocabulary to embed the semantics of the mathematical expression. The user query is then matched with the extracted RDFa annotations and the user is pointed to the list of URLs that have the mathematical formula.

The future work will include a comparison between the search using our proposed mathematical encoding and the regular text search. A user study for system evaluation is currently prepared to test the search enhancement when using the semantic search in e-learning. Another future direction is testing the system usability.

VI. REFERENCES

[1] Al-Khalifa, H. S., & Davis, H. C. (2006). Measuring the Semantic Value of Folksonomies. Innovations in Information Technology, 1-5.

[2] Altamimi, M. E., & Youssef, A. (2008). A Math Query Language with an Expanded Set of Wildcards. Mathematics in Computer Science, 2(2), 305-331.

[3] Asperti, A., Padovani, L., Coen, C. S., & Schena, I. (2001). HELM and the Semantic Math-Web. TPHOLs '01: Proceedings of the 14th International Conference on Theorem Proving in Higher Order Logics, 59-74.

[4] Baeza-yates, R., & Ribeiro-Neto, B. (1999). Modern Information Retrieval.

[5] Bateman, S. (2007). Collaborative tagging : folksonomy, metadata, visualization, e-learning. University of Saskatchewan, Canada.

[6] Bateman, S., Brooks, C., & Mccalla, G. (2006). Collaborative Tagging Approaches for Ontological Metadata. in Proc. Workshop on Applications of Semantic Web Technologies for e-Learning, at the 4th International Conference on Adaptive Hypermedia and Adaptive Web-based Systems, 3-12.

[7] Bateman, S., Brooks, C., Mccalla, G., & Brusilovsky, P. (2007). Applying Collaborative Tagging to E-Learning. In Proceedings of the 16th International World Wide Web Conference (WWW2007).

[8] Bateman, S., Farzan, R., Brusilovsky, P., & McCalla, G. (2006). OATS: The Open Annotation and Tagging System Proceedings of I2LOR 2006.

[9] Chang, B., Ham, D.-h., Moon, D.-s., Choi, Y. S., & Cha, J. (2007). Educational information search service using ontology. Seventh IEEE International Conference on Advanced Learning Technologies (ICALT). , 414 - 415.

[10] Doush, I. A., Alkhateeb, F., & Maghayreh, E. A. (2010). Towards Meaningful Mathematical Expressions in E-Learning. The International Conference on Intelligent Semantic Web Services and Applications.

[11] Enagandula, V., Juthani, N., Ramakrishnan, I. V., Rawal, D., & Vidyasagar, R. (2005). BlackBoardNV: a system for enabling non-visual access to the blackboard course management system. Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility, 220 - 221.

[12] Gomes, P., Antunes, B., Rodrigues, L., Santos, A., & Barbeira, J. (2006). Using Ontologies for eLearning Personalization. Proceedings of 3rd Learning Conference, Portugal

[13] Gruber, T. (2005). Ontology of Folksonomy: A Mash-up of Apples and Oranges. http://tomgruber.org/writing/mtsr05-ontology-of-folksonomy.htm.

[14] Guidi, F., & Schena, I. (2003). A Query Language for a Metadata Framework about Mathematical Resources. The 2nd International Conf. Mathematical K nowledge Management, 105-118.

[15] Henze, N., Dolog, P., & Nejdl, W. (2004). Reasoning and ontologies for personalized e-learning in the semantic web. Educational Technology and Society, 7, 82-97.

[16] Heymann, P., & Garcia-Molina, H. (2006). Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems. http://heymann.stanford.edu/taghierarchy.html: Stanford University.

[17] Hijikata, Y., Hashimoto, H., & Nishida, S. (2009). Search Mathematical Formulas by Mathematical Formulas. Human Interface and the Management of Information. Designing Information Environments, 404-411.

[18] Kim, T.-J., Kim, M.-C., & Park, G.-L. (2005). On Employing Ontology to e-Learning. Proceedings of the Fourth Annual ACIS International Conference on Computer and Information Science, 334-339.

[19] Kohlhase, M., & Sucan, I. A. (2006). A Search Engine for Mathematical Formulae. Proceedings of Artificial Intelligence and Symbolic Computation, AISC'2006, 241-253.

[20] MathCast, an open source equation editor. http://mathcast.sourceforge.net/home.html. (2010).

[21] Mathematical Markup Language, http://www.w3.org/Math/. (2010).

[22] Mathes, A. (2004). Folksonomies - cooperative classification and communication through shared metadata.

[23] Munavalli, R., & Miner, R. (2006). MathFind: a math-aware search engine. SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, 735-735.

[24] Pah, I., Stoica, F., Cacovean, L. F., & Popa, E. M. (2008). Using ontology in electronic evaluation for personalization of e-Learning systems. Proceedings of the 8th conference on Applied informatics and communications (AIC'08), 332-337.

[25] Peters, I., & Becker, P. (2009). Folksonomies : indexing and retrieval in Web 2.0. Berlin: De Gruyter/Saur.

[26] Prud'hommeaux, E., & Seaborne, A. (2008). SPARQL Query Language for RDF, http://www.w3.org/TR/rdf-sparql-query/. (Recommendation): W3C.

[27] RDFa Distiller and Parser, http://www.w3.org/2007/08/pyRdfa/. (2010).

[28] RDFa Primer: Bridging the Human and Data Webs, http://www.w3.org/TR/xhtml-rdfa-primer/. (2010).

[29] rdfquery, RDF processing in your browser, http://code.google.com/p/rdfquery/. (2010).

[30] Soylu, A., Kuru, S., Wild, F., & Mdritscher, F. (2008). e-Learning and Microformats: a Learning Object Harvesting Model and a Sample Application Proceedings of the of Mupple'08 Workshop 57-65.

[31] Stojanovic, L., Staab, S., & Studer, R. (2001). eLearning based on the Semantic Web. In Proceedings of the World Conference on the WWW and the Internet (WebNet 2001).

[32] Whitehead, J. (2009). Challenges Of Online Education. http://www.articlesnatch.com/Article/Challenges-Of-Online-Education/115265.

[33] Youssef, A. (2006). Roles of Math Search in Mathematics. Mathematical Knowledge Management, 2-16.

Dimensionality Reduction technique using Neural Networks – A Survey

Prof. Mrs. Shamla Mantri
Dept. of Computer Engg
MIT College of Engineering,
Pune University, India

Nikhil S. Tarale
Student, Dept. of Computer Engg
MIT College of Engineering,
Pune University, India

Sudip C. Mahajan
Student, Dept. of Computer Engg
MIT College of Engineering,
Pune University, India

***Abstract*— A self-organizing map (SOM) is a classical neural network method for dimensionality reduction. It comes under the unsupervised class. SOM is a neural network that is trained using unsupervised learning to produce a low-dimensional, discretized representation of the input space of the training samples, called a map. SOM uses a neighborhood function to preserve the topological properties of the input space. SOM operates in two modes: training and mapping. Using the input examples, training builds the map. It is also called as vector quantization.**

In this paper, we first survey related dimension reduction methods and then examine their capabilities for face recognition. In this work, different dimensionality reduction techniques such as Principal component analysis [PCA], independent component analysis [ICA] and self-organizing map [SOM] are selected and applied in order to reduce the loss of classification performance due to changes in facial expression. The experiments were conducted on ORL face database and the results show that SOM is a better technique.

Keywords- Principal component analysis [PCA]; Independent component analysis [ICA]; self-organizing map [SOM]; Face recognition.

I. INTRODUCTION

Biometrics refers to the study of methods for uniquely recognizing human based upon one or more intrinsic or behavioral characteristics. Biometrics is used to identify the input sample when compared to a template used in cases to identify specific people by certain characteristics.

Face recognition is an important part of today's emerging biometrics and video surveillance markets. Face recognition can benefit areas of airport security, access control, driver's license, passports; homeland defense, customs and immigration etc. face recognition has been a research area for almost 30 years with significant increased research activity since 1990.

This has resulted in successful algorithms and the introduction of commercial products. The benefit of using a computer system for face recognition would be its capacity to handle large amount of data and the ability to do a job in a predefined repeated manner.

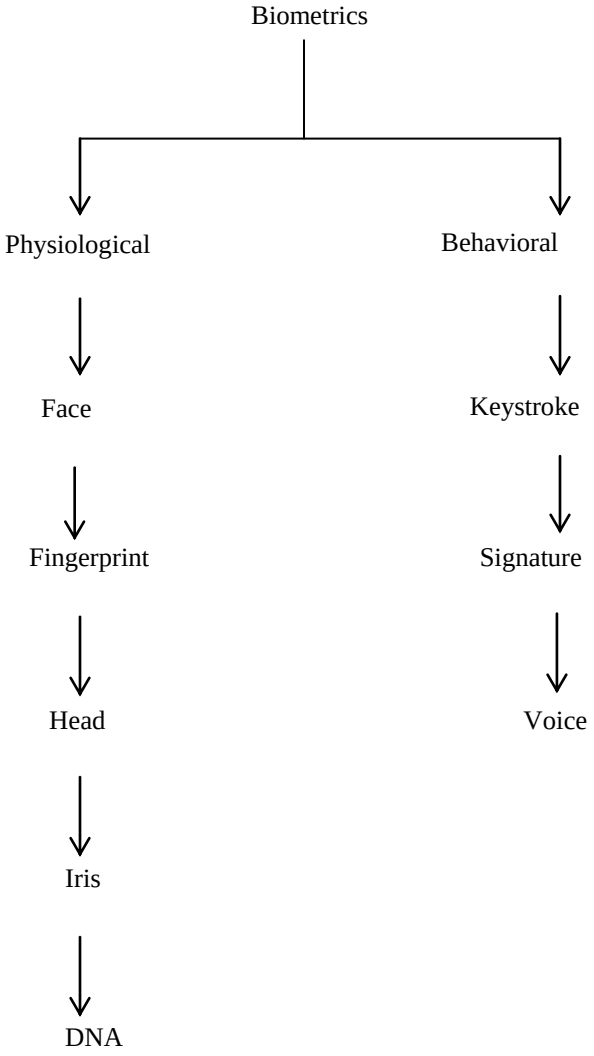


Fig 1. Biometrics

Suppose we have a collection of points of n-dimensional real vectors drawn from an unknown probability distribution but the situation in most of the cases is where dimensions are very large.

This leads one to the methods of dimensionality reduction that allows one to represent data in lower dimension space. The steps for face recognition in [3] are as follow:

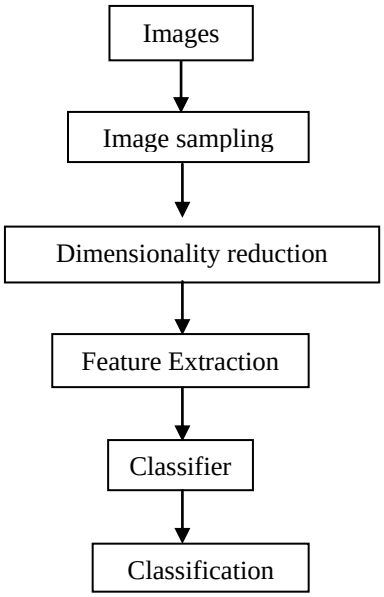


Figure 2. Face recognition

1. Selection and sampling of the image from the database;
2. Dimensionality reduction;
3. Feature extraction;
4. Classifier

A. *Selection and Sampling:*

Sampling is selection of those points which are required to represent the given image. It is mapping of the image from a continuum of points in space to a discrete set.

B. *Dimensionality and reduction*

In this phase, the dataset i.e. the images are reduced to minimum size by sampling. Thus, a reduced data set is obtained.

C. *Feature Extraction*

This process deals with extracting patterns from the data by using techniques such as classification, regression, segmentation or Deviation detection.

D. *Classifier*

Classification involves mapping data into one of several predefined or newly discovered classes.

In practical situation one is often forced to use linear or even sub-linear techniques. Principal component analysis [PCA], Independent component analysis [ICA] and Self-organizing mapping are the popular form of linear techniques. Using the SOM as a feature extraction method in face recognition applications is a promising approach, because the learning is unsupervised, no pre-classified image data are needed at all. When high compressed representations of face images or their parts are formed by SOM, the final classification produced is fairly simple, needing only a moderate number of labeled training samples.

In this paper, we have introduced face recognition algorithms based on this consideration.

Technically, a principal component can be defined as a linear combination of optimally-weighted observed variables.

II. Principal Component Analysis

Principal Component Analysis is a variable reduction procedure.

It is useful for removing redundancy (i.e. some variables are correlated to each other) among data and to reduce the observed variables into smaller no of principal components that will account for most of the variance in the observed variables.

Principal Component Analysis provides mapping of n-dimensional data space into m-dimensional data space, a set of n sample images {I1, I2...IN} taking values of n-dimensional is considered and there mapping from original n-dimensional image spaced to m-dimensional feature space where m<n. the new feature vector YK is defined by the following linear transformation $Y_K = \varphi^T I_k$

where φ is a matrix with orthogonal columns $[e_1,e_2,..,e_m]$, the Eigen vectors covariance matrix

$$R = \sum_{k=1}^{N} \left(\mathbf{I_k} - \mathbf{I}_{mean}\right)\left(\mathbf{I_k} - \mathbf{I}_{mean}\right)^{\mathbf{T}}$$

where I_{mean} is the mean image of the samples. Only m no of n-dimensional eigenvectors of R corresponding to m largest eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are chosen.

III. Independent Component Analysis

ICA is a statistical and computational technique for revealing hidden factors that underlie sets or random variables, measurements, or signals. ICA is superficially related to principal component analysis and factors analysis.

The ICA algorithm aims at finding S_i component as independent as possible so that the set of observed signals can be expressed as a linear combination of statistically independent components.

We have a random vector

$X = (x_1, x_2, x_3,\ldots, x_m)^T$, and an m-dimensional vector for observation and original source vector with n independent components.

$$S = (s_1, s_2,\ldots,s_n)^T$$

The linear transformation can be modeled as

$$X = As$$

Here, A is an m*n non-singular matrix.

The original source vector S recovered by following transformation:

$$Y = W x ;$$

where $Y = (y_1, y_2, \ldots, y_n)^T$ is an n dimensional output vector, and W is an n*m weight matrix. In the information theoretic technique like ICA, various objective functions based on information theoretic concepts such as negentropy, minimization of mutual information, maximum entropy, maximum likelihood have been used for source separation problem. This paper follows maximum entropy based ICA method for face recognition [5], the weight update rule for which is [6].

$$\Delta W = \eta (I + (1 - 2z)y^T) W$$

Where **z** is the output of nonlinearity (logistic function) used. ICA has been performed on both the Architectures (I & II) as proposed in [5].

IV. Self-Organizing Map

T. Kohonen introduced self-organizing map[1]. It is unsupervised learning process, which learns the distribution of a set of patterns without any class information. It has the property of topological preservation.

SOMs have also been successfully used in dimensionality reduction and feature selection for face space representations.

Algorithm

PCA is applied as it generates a set of orthogonal axes of projections known as principal components or eigen vectors. PCA is applied to weight matrix generated by mapping the image onto lower dimensional space using SOM. Only the Eigen vectors for large values are considered and those for smaller values are ignored. The steps are as follows:-

Step 1. A face image of size $m \times m$ was divided into sub-blocks of size $b \times b$ resulting in total of p = (m*m) / (b*b) blocks each of which gives q =b*b number of elements, concatenation of which gives a vector to represent one block resulting in a matrix $X = [X_1, X_2, \ldots, X_p]$ of size $q \times p$. This gives a stream of training vectors $\{X_i\}^p_{i=1}$

Step 2. Consider 2-dimensional ($s \times s$) map of neurons each of which is identified as index $jk, j,k = 1,2,\ldots,s$.

The *j k th* neuron has an incoming weight

$W_{jk} = (w_{1,jk}, \ldots, w_{q,jk})$ at instant *i*. The value of neighborhood function around the winning neuron as h_{jk} at instant *i*. Initialize weight W_{jk}, neighborhood h_{jk} and the learning rate η_o.

Step 3. Pick a sample vector X_i at random and present it to a two dimensional $(s \times s)$ map of neurons with a total of $z = s * s$ neurons.

Step 4. Find out best matching (wining neuron) using following distance criterion

$$\| X_i - W_{JK(i)} \| = \min_{jk} \{ \| X_i - W_{jk(i)} \| \}$$

where W_{jk} is the best matching weight vector.

Step 5. Update the synaptic weight vectors of only the winning cluster

$$W_{jk(i+1)} = W_{jk(i)} + \eta_i (X_{(i)} - W_{jk(i)}) \qquad jk \; \varepsilon \; h_{JK(i)}$$

Step 6. Update learning η_i and the neighborhood $h_{jk(i)}$

Step 7. Continue with step 3 until no noticeable changes in the feature map are observed. Finally a matrix *M* of size $z \times q$ is obtained.

Step 8. Compute the Eigen vectors and Eigen values of the covariance matrix M^TM, sort the Eigen vectors and retain the Eigen vectors corresponding to highest values.

Step 9. Calculate the KL coefficients (M^T * Eigenvectors) and retain them.

Step 10. Repeat above steps for all training images.

Step 11. Reconstruct the images at the time of recognition match with the test image using nearest neighbor classifier.

V. Training and Test Data

For the face recognition experiment, we partitioned the ORL database into a training set and a testing set. The partition is done as follow:

First, k images are selected for test; the remaining images i.e. (10 - k) are used for training set and for computing the projected matrix. All the ten images in the training and test sets were projected to a dimension reduced space.

VI. Experimentation

Here the eigenvectors of the weight matrix were found. PCA was then applied to the transpose of the weight matrix and the Eigen vectors corresponding to the largest eigenvalues were retained for reconstruction of the image. The table shows the results for PCA, SOM and PCA+SOM

TABLE I. Recognition Rate of the Face Recognition System For Varying Number of Classes

Recognition Rate (%)			
Method	***Number of Classes***		
	10	20	40
SOM(5×5)	94.06	90.72	89.92
PCA	93.39	90.25	89.51
SOM+PCA	77.75	72.08	62.64

The above table and the figure given below shows that SOM performs better than PCA and PCA + SOM.

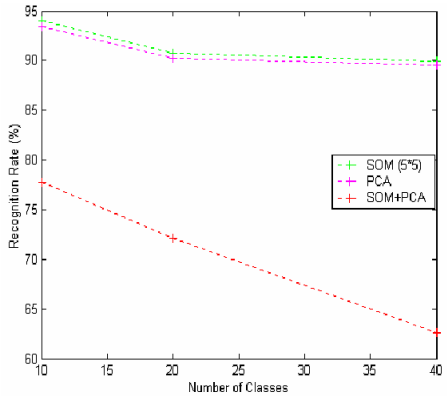


Figure 3. Recognition Rate as a function of number of classes

We used the AT&T database [10] for the face recognition experiments. Here experiments were done with nearly 400 images with variation of 40 persons. A preview image of the Database of Faces is as shown in Figure 4.

Figure 4. Example of faces in AT&T database.

The experimental observations of the experiments performed on datasets are shown in Table

TABLE II. EXPERIMENTAL OBSERVATIONS

Criteria	Values
Training images	30
Testing images	370
Learning coefficient	0.01
Iterations	50
Recognition Rate	85.5%

VII. CONCLUSION

The SOM algorithm is a typical dimensionality reduction technique which has good properties to preserve topological relationships even in lower dimensional space. The algorithm is very suitable for using K nearest neighbor classifier. The experimental results show that our proposed algorithm performs better and faster in real data set.

An efficient system for face recognition using SOM has been proposed. Firstly, this system provides a general integration of multiple feature-sets using multiple self-organizing maps. Secondly, with the help of compressed feature vector, SOM is trained to organize all face images in database.

The highest average recognition rate of 85.5% is obtained for 40 persons' 400 images of AT&T database, where the training is done on 30 images only and tested on remaining images. Thus, the SOM method is an efficient face recognition process.

REFERENCES

[1] T. Kohonen, "The Self -Organizing Map", IEEE, VOL. 78, NO. 9, SEPTEMBER 1990.

[2] Santaji Ghorpade, Jayshree Ghorpade, Shamla Mantri, "Patter Recognition using Neural Networks", International Journal of Computer Science and Technology (IJCST), Vol 2, No 6, December 2010

[3] Santaji Ghorpade, Jayshree Ghorpade, Shamla Mantri,"Neural Networks for face recognition Using SOM", IJCST Vol1, issue 2, December 2010

[4] Jayshree Ghorpade, Siddhant Agarwal, "SOM and PCA Approach for Face Recognition – A Survey", International Journal of Computer Trends and Technology – March to April Issue – 2011

[5] M. S. Barlett, "Face image analysis by unsupervised learning and redundancy reduction," *Ph.D. Dissertation*, University of California, San Diego, 1998.

[6] A. J. Bell and T. J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution", *Neural Computation*, Vol. 7, No. 6, pp. 1129-1159,1995.

[7] Dinesh Kumar, C. S. Rai and Shakti Kumar, "Dimensionality Reduction using SOM based Technique for Face Recognition", Journal of Multimedia, VOL. 3, NO. 1, MAY 2008.

[8] A. S. Raja and V. JosephRaj, "A New Multimodal Technique Without Subject's Cooperation Using Neural Network Based Organizing Maps", International Conference on Electronics and Information Engineering (ICEIE 2010).

[9] Qiu Chen, Koji Kotani, Feifei Lee and Tadahiro Ohmi, "Face Recognition Using Self – Organizing Maps"

[10] AT&T Laboratories Cambridge, The database of faces at http://www.cl.cam.ac.uk/research/dtg/attractive/facesataglance.html.

[11] Li chaoyang, lin fang, xie yinxiang,"Face recognition using Self-organizing features map and support vector machines",

[12] M.A.kerin,T.J stomham."face recognition using digital neural network with self-organising capabilities"

[13] Firegoire Lefebvre and Christophe Garcia," A probabililistic self-organizing map for Face Recognition".

[14] Dey, A. K., & Saha, S. (2010). A Method of Genetic Algorithm (GA) for FIR Filter Construction : Design and Development with Newer Approaches in Neural Network Platform. International Journal of Advanced Computer Science and Applications - IJACSA, 1(6).

[15] Trivedi, J. A. (2011). Framework for Automatic Development of Type 2 Fuzzy , Neuro and Neuro-Fuzzy Systems. International Journal of Advanced Computer Science and Applications - IJACSA, 2(1), 131-137.

A Novel approach for Implementing Security over Vehicular Ad hoc network using Signcryption through Network Grid

Vijayan R
Network and Information Security Division
School of Information Technology and Engineering
VIT University
Vellore, India

Sumitkumar Singh
Master of Technology
School of Information Technology and Engineering
VIT University
Vellore, India

***Abstract*— Security over Vehicular ad hoc network and identifying accurate vehicle location has always been a major challenge over VANET. Even though GPS system can be used to identify the location of the vehicle they too suffer from major drawbacks. A novel approach has been suggested by the author wherein the VANET is made more secured by using Signcryption technique and at the same time unique approach of using Network Grid to flawlessly identify the location of the vehicle has been proposed.**

Keywords- Network Grid; Computation Server; Vehicular Node; Public/Private keys.

I. INTRODUCTION

The invention of Vehicular ad hoc network has eased the burden of communication over Vehicle to Vehicle communication and Vehicle to Interface communication. But VANET as compared to Mobile ad hoc network or MANET is highly dynamic and unsecured. Providing security and at the same time preventing the current transmission to attain loss due to frequent path breakage over VANET has always been a major challenge. Moreover, the computation cost required over VANET should be less as compared to MANET. So, even if a security protocol is to be implemented over VANET care must be taken so that the computation cost doesn't increase. A unique public key cryptography technique Signcryption as proposed by Zengh [1] has been suggested in this paper. Signcryption is a cryptography technique which combines the two step of Digital signature and Encryption in one step and hence reduces the computation time up to great extent as compared to Signature-then-Encryption.

The Digital Signature for the automobiles has been proposed in [2]. The major drawback of this technique involves higher computation cost as more number of machine cycles is required for the computation as compared to Signcryption process. The Signcryption provides an entire feature to enhance the security measures like Confidentiality and Integrity of the message. Different types of attacks over VANET have be stated in [4] [5]. The detailed process of using Signcryption will be explained in the proceeding sections. As now we have dealt with the security part we have proposed a technique which can be used to prevent the data transmission loss due to frequent path breakage that often appears over VANET. Many papers have suggested the use of GPS [6] system over VANET but GPS system too faces some serious drawbacks which can be dangerous in some situations. [3] have suggested the advantages of using DSRC therefore DSRC shall be used in this model. Some of the drawbacks of GPS System are the Cost which has to be minimized when being used over VANET, Inaccuracy since not all the GPS devices are updated hence the system cannot state the exact updated road conditions, the network coverage can be weak in challenging locations like between obstacles like tall buildings or sparse coverage areas and many more. In order to minimize these drawbacks we have proposed a unique technique of dividing the entire geographical locations in to network grid and assigning every intersection a unique network ID. An Infrastructure based model has been suggested over this network wherein the vehicle to vehicle communication will be established through the Roadside units (RSU) which are in turn connected to the Computation server (CS). The Details of this process will be explained in the proceeding sections.

II. NETWORK MODEL ENTITIES

The proposed network model as in Fig: 1. consists of Computation Server (CS), Roadside Unit (RSU) and nodes model.

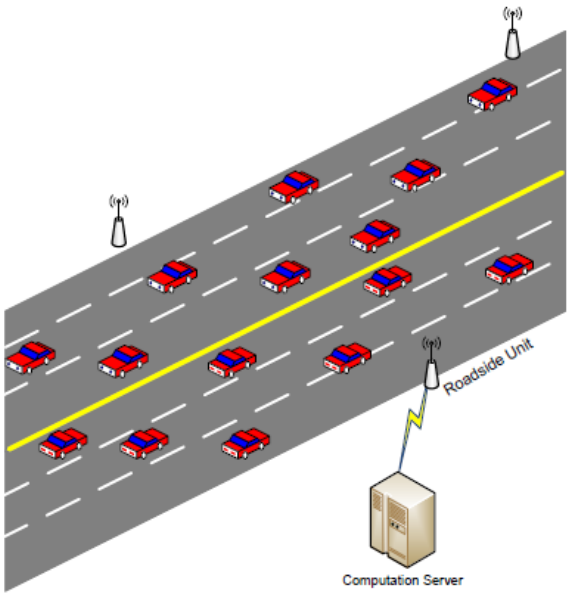


Fig:1 Represents Network model Comprising of Computation Server, Vehicles as nodes and Roadside Units as Access Point

The following are the entities that participate in this network.

A. *Computation Server*

Computation server is responsible for generating unique public and private key as requested by the Source node. The request from the source node is forwarded to the Computational Server through the RSU. The CS is also responsible for selecting the suitable RSU for sending data to the vehicles over the network based on the computation of the current velocity of the vehicle which is received in the request packet from the vehicle.

B. *Roadside Unit*

The RSU acts as an access point between CS and the vehicle. The RSU is responsible for transmitting the requested packet from the vehicle to the CS where the further computation is done.

C. *Vehicular Node*

The Vehicular nodes are the entities which will exchange messages through the CS. Every node possesses a unique Vehicular ID (VID) through which it is identified. In case of source node, the source node will be aware of the destination nodes unique id (DID). Based on the DID the source request the CS for the transmission. The detailed process of transmission is explained in proceeding section.

III. Proposed Network Model

The proposed network model uses Signcryption as the security model whereas Network grid to identify the location of the vehicles. In our model every vehicle is supposed to register with the service provider before entering into the network. Once registered, every vehicle will be assigned a unique Vehicle Identity (VID). This ID can be the vehicles chassis number or the vehicle number. After the registration every vehicle will be identified by its VID. During the process of registration, the vehicle will also be given the IDs of the other vehicles allowed to communicate within the network. For a source vehicle the destination vehicle will be identified by its destination ID (DID) which itself will be the VID of the destination. After getting assigned by Unique identification number the vehicle is now allowed to transmit within the network. Our model follows proactive routing technique wherein every node is required to transmit beacon at regular interval to identify its location to other nodes within the network. These beacons are also received by the CS which can be used to transmit the packet to the node which is not within the transmission range of the Source node. The initial transmission requires the Source vehicle to send REQUEST packet to CS which will contain the Current Velocity (CV), Vehicle ID (VID), Destination ID (DID) and Time (T) at which the message was created. Once the Request packet is received by CS the initial computation is done over the REQUEST message. First, the VID and DID is verified and if found to be true the further computation is done. Public and private keys are now generated and are transmitted to the Source and destination nodes. The transmission is done by evaluating the CV and forwarding the packet to appropriate RSU. Our model follows a unique network grid model which allows the CS to accurately identify the location of the node and forward it accordingly.

The detailed step of Network Grid will be explained in proceeding section. Once the keys are received by the source node the process of Signcryption [1] is carried out. The detailed step of Signcryption is explained in proceeding section. After the Signcryption, the processed packet is transmitted to the CS by encrypting it with the Shared Key between the Source node and the CS. The processed packet also contains the current velocity of the source node. After receiving the encrypted packet from Source node the CS decrypts the packet and sends the packet to destination node depending upon its location over Network Grid. The destination node now receives the packet and Unsigncrypt the message to obtain the original message.

IV. The Network Grid

A unique network grid is followed in this paper. The entire geographical region is divided into grid as in Fig 2. In highways or in the urban areas the roads are normally divided into Lanes.

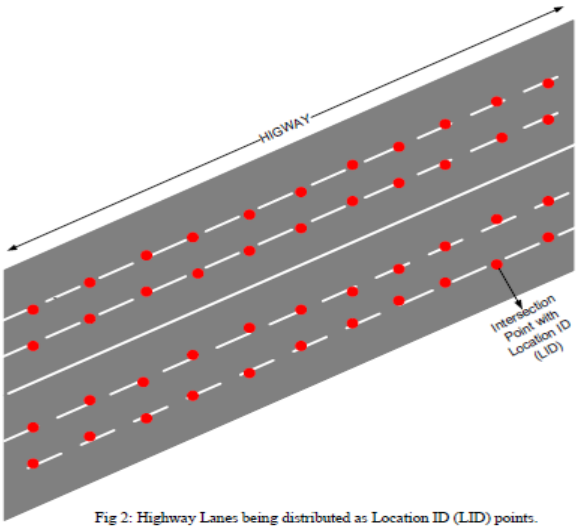


Fig 2: Highway Lanes being distributed as Location ID (LID) points.

Therefore we can make use of such geographical features and divide the Lanes into Network Grids as in Fig.2

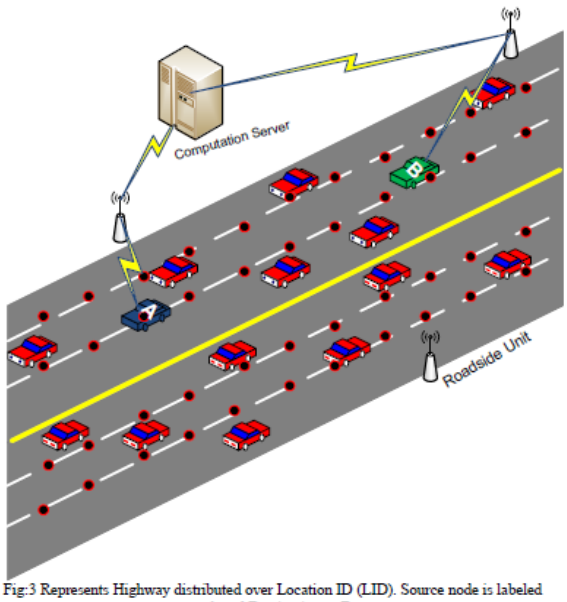


Fig:3 Represents Highway distributed over Location ID (LID). Source node is labeled as A and Destination as B.

As in fig 3, the CS receives the packet from a particular intersection. Every RSU are kept at a required distance from the intersections based on its transmission range. Consider a scenario as in Fig 3, where the source node needs to send packet to destination node. The source node first sends the packet by encrypting the packet with unique shared key between Source and CS. The REQUEST packet will contain CV, VID, DID and T. After verifying the IDs the CS will now calculate the CV of the vehicle. This is required since we are working over vehicular network and the vehicle may cross the transmission range of the current RSU after the computation at the CS. During the process of calculation of the CV of the vehicle the CS also evaluates the Time T at which the message was generated. As in Fig 4: which shows a typical propagation delay of the packet which is transmitted from the vehicle at a specific position to the RSU. Based on this delay of message transmission and Current velocity of the message the location of the Vehicle can be identified. As mentioned earlier on that our model follows proactive routing hence every node knows about the current location of the neighboring nodes and also the CS is aware about the nodes as the packets are also received by the RSU. Once the LID of destination node, as requested by the source, is identified the keys are transmitted to the destination node and source node. Once the Keys are received the process of Signcryption is carried out as explained in proceeding section.

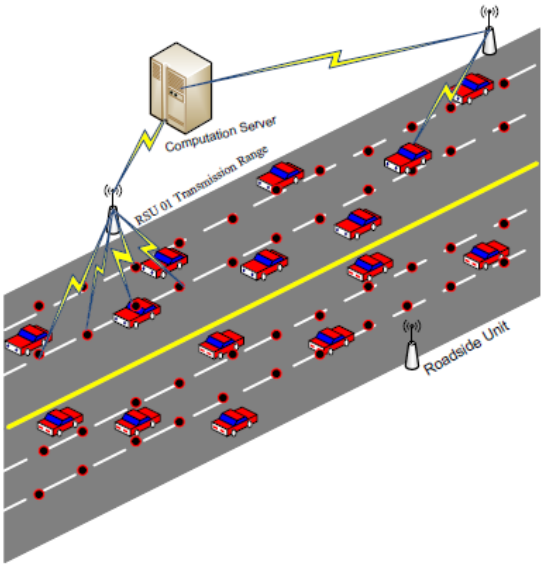


Fig 4: RSU 01 having transmission range over Locations. The message received from last lane node is assumed to have a propagation delay of 30ms whereas message received from middle Lane is assumed to have propagation delay of 60ms.

V. Security Features

The Source node in our model first checks for the location of destination and creates a REQUEST packet containing CV, VID and DID. This whole packet is encrypted by the shared key SKC used between Source node and CS. On receiving the request from the source node for the transmission to the desired destination the CS first checks the VID of the source node. If the ID is forged the node is immediately removed from the network. Thus the source node is termed as malicious or attacker node and is informed to all the nodes within the network about this node. Every node on receiving this message will update their database and remove the node from their destination list. After the verification, CS generates Public key and private key pairs for Source node and destination node and Shared key SKN which will be used by source to encrypt and transmit Cipher text ‘c’, value ‘r’ and ’s’ and by destination to decrypt to obtain the message.. The type of encryption used will depend upon the level of security required by the network. All the keys and values that are generated by the CS are sent to source and destination by encrypting it with their SKC. After receiving the Keys the Signcryption process is carried out at the Source node. The Algorithm 1 represents the detailed process carried out during the transmission. The Notations used during a message transmission are shown in TABLE – I

A. Signcryption at Source

Once the sender has received the PU_a and PR_a it can now perform Signcryption [2] over the message that is to be sent to the destination. It is understood that destination has received the PU_b and PR_b and is ready to receive the message from source node.

Algorithm used for evaluating the message at VANET server from Source Node for Single Mode of Transmission

Algorithm 1:

1. E(SKC[REQUEST(CV,VID,DID,T)]) from source node 'a'
2. Search(VID==VID)
3. **IF** found (VID==VID) **Then {**
4. Generate PU_a,PR_a, PU_b, PR_b
5. **E**(SKC[RPLY(ID_{new},PU_a,PR_a,SKN)]) to Source node
6. **E**(SKC[SEND(PU_b,PR_b,SKN)]) to Destination node
7. **} Else** (Remove Source node from the network)

Following steps for Signcryption are carried out as described in [1] by the source node. The Signcryption process is then followed by the process of Unsigncryption as described in [1].

- Source node selects random value 'x' where x is in the range of (1,…,q-1). This chosen random value 'x' will be used in further Hash function.
- The source now selects PU_b and random value x to compute Hash function out of it. This creates a 128bit string. K = H (PU_b mod p) where 'p' is a large prime number.
- The 128bit key obtained is divided into two halves K1 and K2.
- Source now uses AES encryption technique and encrypts the message using Key K1 to produce Cipher C=E(K1[m])
- It is now followed by one-way keyed Hash function over message 'm' with Key K2 to produce 'r' where r = KH(m).
- Now the sum of PR_a and 'r' is calculated and a modulo is performed over the sum with value 'q' where 'q' is the prime factor of (p-1) to produce 'result' which is then divided by the random value 'x' which produces a value 's'

Now, three different values have been produced that are c, r and s. The source node can now encrypt these three values using Advanced Encryption Standard using Sk_{ab} and transmit them to the destination node.

B. Unsyncrption at destination

After the signcryption at the source is completed the destination node now possess c, r and s. using these values the destination now decrypts the message.

- After receiving the values c, r and s the destination node now decrypts the message to obtain the original message.
- The destination receives three values that are c, r and s. The destination now uses r, s, PUa, PRb, p and g to compute a hash to produce 128bit result where 'g' is an integer with the order q modulo p chosen randomly from (1,….p-1).
- The Hash function then produces Key K = H ((PUa * gr)s X PRb mod p). This Hash function now produces a key of 128bits. This 128 bit key is now divided into two halves to produce two 64 bit key and these are identical to the keys that are generated during Signcryption process by source node.
- Destination node now uses Key K1 to decrypt Cipher 'c' to get the original message m = D(K1[c]).

TABLE I. NOTATIONS USED DURING THE SIGNCRYPTION AND UNSIGNCRYPTION PROCESS

Symbol	*Process*
REQUEST	Request from source node
RPLY	Reply from CS to source node
SEND	Send key from CS to
DSPLY	Display message
E (…)	Encryption of Message
D (…)	Decryption of Message
PU_a	Public key for source node 'a'
PR_a	Private Key for source node 'a'
PU_b	Public Key for destination node 'b'
PR_b	Private Key for destination node 'b'
SKN	Shared Key between Source node 'a' and Destination node 'b'.
SKC	Shared key between source node 'a' and CS
LID	Location ID
VID	Vehicle Identity
T	Time

ACKNOWLEDGMENT

I owe a great many thanks to a great many people who helped and supported me during the writing of this paper. My deepest thanks to Lecturer, **Vijayan R** the Guide and author of this paper for guiding and correcting various documents of mine with attention and care. He has taken pain to go through the paper and make necessary correction as and when needed.

REFERENCES

[1] Yuliang Zheng, "Digital Signcryption or How to Achieve Cost(Signature & Encryption)<<Cost(Signature) + Cost(Encryption)" 1997 in CRYPTO '97 Proceedings of the 17th Annual International Cryptology Conference on Advances in Cryptology

[2] Dr. iur.Lutz Gollan, Prof. Dr. sc. Christoph Meinel "Digital Signature in Automobiles" 2002 in Systemics, Cybernetics and Informatics (SCI)

[3] Arijit Khan, Shatrugna Sadhu, and Muralikrishna Yeleswarapu, "A comparative analysis of DSRC and 802.11 over Vehicular Ad hoc Networks" http://www.cs.ucsb.edu/~arijitkhan/cs276.pdf

[4] J.T. Isaac1 S. Zeadally J.S. Ca´mara, "Security attacks and solutions for vehicular ad hoc networks" 2010 in IEEE-Communications IET ,Volume 4 issue 7, 1751-8628

[5] Maxim Raya and Jean-Pierre Hubaux,"Securing vehicular ad hoc networks" 2005 3rd ACM workshop on Security of ad hoc and sensor networks (SASN)

[6] Jason Chao, Yong-qi Chen, Wu Chen, Xiaoli Ding, **Zhilin Li**, Nganying Wong and Meng Yu**, 2001.** An Experimental Investigation into the Performance of GPS-based Vehicle Positioning in Very Dense Urban Areas, Journal of Geospatial Engineering, 3(1): 59._-66.

[7] Journal, I., Science, A. C., & Hod, M. (2011). A Survey on Attacks and Defense Metrics of Routing Mechanism in Mobile Ad hoc Networks. International Journal of Advanced Computer Science and Applications - IJACSA, 2(3), 7-12.

[8] Suri, P. K. (2011). A Novel Approach to Implement Fixed to Mobile Convergence in Mobile Adhoc Networks. International Journal of Advanced Computer Science and Applications - IJACSA, 2(1).

[9] Suri, P. K. (2011). Simulation of Packet Telephony in Mobile Adhoc Networks Using Network Simulator. International Journal of Advanced Computer Science and Applications - IJACSA, 2(1), 87-92.

[10] Indukuri, R. K. R. (2011). Dominating Sets and Spanning Tree based Clustering Algorithms for Mobile Ad hoc Networks. International Journal of Advanced Computer Science and Applications - IJACSA, 2(2).

AUTHORS PROFILE

Prof. Vijayan R: An Assistant Professor (senior) at SITE, VIT University, Vellore, India is a Research scholar and currently pursuing his research work on network and Information security systems. He has published number of papers in various different journals and presented number of papers in International conference.

Mr.Sumitkumar Singh is currently pursuing his Master of Technology in Information Technology Specializing in Networking from SITE, VIT University, Vellore, India. His current research work involves security over Vehicular Ad Hoc Network. Sumitkumar has published number of papers and presented papers in different conferences.

Performance Analysis of MIMO-OFDM System Using Singular Value Decomposition and Water Filling Algorithm

Md. Noor-A-Rahim[1], Md. Saiful Islam[2], Md. Nashid Anjum[3], Md. Kamal Hosain[4], and Abbas Z. Kouzani[5]
[1,3]Electronics & Communications Engineering, Khulna University of Engineering and Technology
Khulna, Bangladesh.
[4]Electronics and Telecommunications Engineering, Rajshahi University of Engineering and Technology
Rajshahi, Bangladesh-6204.
[2,5]School of Engineering, Deakin University, Victoria 3217, Australia

***Abstract*—In this paper, MIMO is paired up with OFDM to improve the performance of wireless transmission systems. Multiple antennas are employed both at the transmitting as well as receiving ends. The performance of an OFDM system is measured, considering multipath delay spread, channel noise, Rayleigh fading channel and distortion. In this paper, bits are generated and then mapped with modulation schemes such as QPSK, 8PSK, and QAM. Then, the mapped data is divided into blocks of 120 modulated data where a training sequence of the data is inserted both at the beginning and ending parts of the block. The equalization is used to determine the variation to the rest of data. The singular value decomposition (SVD) and water filling algorithm have been employed to measure the performance of the MIMO-OFDM integrated systems. Therefore, the capacity is increased by transmitting different streams of data through different antennas at a same carrier frequency. Any intersymbol interference (ISI) produced after the transmission is recovered by using spatial sampling integrated with the signal processing algorithm. Furthermore, the performance remains the same with different combinations of transmitting and receiving antennas.**

Keywords-MIMO; OFDM; ISI; SVD; Water filling algorithm

I. INTRODUCTION

Wireless communication systems face high level of ISI which originates from multipath propagation and inherent delay spread. A multicarrier based technique such as orthogonal frequency division multiplexing (OFDM) can be used for extenuating ISI to improve capacity and spectral efficiency (bps/Hz) in a wireless system [2]. In addition, MIMO systems are promising techniques to increase performance with acceptable bit error rate (BER) by using a number of antennas [3]. A MIMO-OFDM system transmits OFDM modulated data from multiple antennas at the transmitter. Data transmitted with subcarriers at different antennas are mutually orthogonal. The receiver extracts different data stream from different subcarriers after OFDM demodulation and MIMO decoding. Flat fading MIMO algorithms reduce computational requirement and make MIMO-OFDM attractive for mobile applications [1].

Although MIMO was first introduced in 1990s, it is still an active research area. Adaptive technology has been used in MlMO systems to get high spectral efficiency. For a time-varying channel, we need a simple algorithm which adaptively adjusts transmit parameters for high capacity [4]. A novel power allocation and adaptive modulation algorithm for water filling can be created by combining the adaptive technology with cell planning, which can reduce interfere to the adjacent cell. Adaptive SVD algorithm uses a two-step recursive method, is an important technique to exploit the full capability of MIMO-OFDM systems. Adaptive power control based on water filling, and rate adaptation maximizes the spectral efficiency of MIMO-OFDM in discrete and upper bounded alphabets [5]. Whereas, Y. Jiang, et al. have investigated that two dimensional water filling power allocation algorithm with the SVD for MIMO-OFDM system provide higher capacity than that of its uniform power allocation counterpart with one dimensional water filling power allocation scheme [6].

Co-Channel Interference (CCI) between transmitted and received substreams of single-frame TDMA data can be minimized by OFDM-MIMO systems with V-BLAST signal processing at the receiver [7]. A custom hardware for computing SVD can reduce its computing time [8]. Generalized SVD based on beamforming can be used in MIMO broadcasting, and in MIMO relaying [9]. W. Liejun has proposed improved water filling algorithm for determining the optimal transmit powers and larger throughputs in MIMO-OFDM system [10].

This paper focuses on high data rate wireless communication due to such benefits as inter-symbol interference (ISI) free transmission, high spectral efficiency, and reduced equalization complexity. The paper shows that OFDM-MIMO system with SVD and water filling algorithm is robust against the signal-to-noise variations, and its performance is superior to a single antenna system.

II. BASIC MIMO-OFDM SYSTEM

Orthogonal frequency division multiplexing (OFDM) is a popular wireless multicarrier transmission technique. It is a promising candidate for next-generation wired and mobile wireless systems. The basic principle of OFDM is to split a high rate data stream into a number of lower rate streams so that the lower data rate can be transmitted simultaneously over

a number of subcarriers. In OFDM, the amount of dispersion in time, caused by multipath delay spread, is decreased due to the increased symbol duration for lower rate parallel subcarriers. The spectrum of OFDM is more efficient because of the use of closer channels space. Interferences are prevented by making all the carriers orthogonal to one another. MIMO systems utilize space-multiplex by using antenna arrays to enhance the efficiency in the used bandwidth. These systems are defined by spatial diversity and spatial multiplexing. Spatial diversity is known as Rx- and Tx-diversity. Signal copies are transferred from another antenna, or received at more than one antenna. With spatial multiplexing, the system carries more than one spatial data stream over one frequency, simultaneously.

In an N subcarriers MIMO-OFDM system, the individual data streams are first passed through an OFDM modulator. Then, the resulting OFDM symbols are launched simultaneously through the transmit antennas. In the receiver side, the individual received signals are passed through OFDM demodulator. The outputs of the OFDM demodulator are decoded and rearranged to get the desired output. Fig. 1 shows the schematic diagram of a basic MIMO-OFDM system.

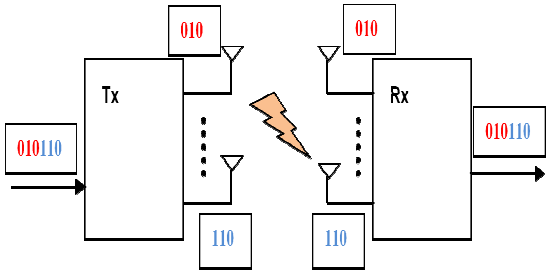


Figure 1. A basic MIMO-OFDM system [1].

III. THEORY

A. Spatial Multiplexing

The transmission of multiple data streams over more than one antenna is called spatial multiplexing [11]. The advantage of spatial multiplexing is linear capacity gains in relation to the number of transmit antennas.

B. Spatial Diversity

Spatial diversity improves the signal quality and achieves a higher signal to noise ratio at the receiver side. The principle of diversity relies on the transmission of structured redundancy. This redundancy can be transmitted at any time, from any antenna, over any frequency or at any polarization. Two kinds of spatial diversity need to be considered: (i)Tx-Diversity, where a signal copy is transmitted through more than one antenna, (ii) Rx-Diversity, where the received signal is multiple evaluated [12].

C. MIMO Channel Matrix

The MIMO system has multiple links and operates on the same frequency whereas the non-MIMO system is linked over multiple channels by several frequencies. The challenge of this technology is the separation and the equalization of the signal in all paths. The channel model includes the channel matrix H with the direct and the indirect channel components. The direct components represent the channel flatness whereas the indirect components stand for the channel isolation. Consider an OFDM symbol of N sub-symbols and cyclic prefix, P, the length of which is less than the last significant tap delay. The sent signal is 'x' and the received signal is 'y'. A time-invariant and narrowband channel is assumed. The output y can be expressed in a matrix format as follows [13]:

$$y = H\,x\,\eta \quad (1)$$

where y is the received vector, x is the transmitted vector and η is complex AWGN vector.

D. Capacity

Claude Elwood Shannon [14] developed the following equation for the theoretical channel capacity:

$$c_{SISO} = f_g \log_2(1 + S/N) \quad (2)$$

It includes the transmission bandwidth fg and the signal-to-noise ratio. Most channel capacity improvements are based on bandwidth extensions or modulations. The Shannon capacity of MIMO Systems additionally depends on the number of antennas. For MIMO the capacity is given by the following equation:

$$c_{MIMO} = Mf_g \log_2(1 + S/N) \quad (3)$$

where M is the minimum of MT (number of transmitting antennas) or MR (number of receiving antennas) and represents the number of spatial streams. For example, a 2 x 3 system can only support two spatial streams, which is also true for a 2 x 4 system. Asymmetrical antennas constellations are referred to as receive or transmit diversity. In that case, the capacity CTx/Rx grows logarithmically with the number of antennas [15, 16].

$$c_{Tx/Rx} = f_g \log_2\left(1 + M\left(\frac{S}{N}\right)\right) \quad (4)$$

E. SNR Threshold

In the proposed transmission control scheme, the threshold value identifies whether multiple transmit antennas should be used or not. Therefore, it is obvious that selecting the right threshold value has significant effect on the performance of the MIMO system. A small threshold increases the complexity without achieving high multiplexing gain [17]. On the other hand, a large threshold causes loosing multiplexing gain due to not using the multiple antennas.

F. Water Filling

Water filling refers to a technique whereby the power for the spatial channels are adjusted based on the channels gain. The channel with high gain and signal to noise ratio (SNR) is given more power. More power maximizes the sum of data rates in all subchannels. The data rate in each subchannel is related to the power allocation by Shannon's Gaussian capacity formula C = B log(1 + SNR). However, because of the capacity is a logarithmic function of power, the data rate is usually insensitive to the exact power allocation. This motivates the search for simpler power allocation schemes that can perform close to the optimal. The water filling algorithm is based on an iterative procedure, as described below [18].

Filling Algorithm: VEC is the absolute or relative level of noise in LINEAR units at different frequencies, space or

whatever bins. PCON is a total power constrain given in the same units as the VEC. TOL is an acceptable tolerance in the units of VEC. WLINE indicates the WATERLINE level in units of VEC so that:

abs(PCON – SUM (MAX (WLINE – VEC,0))) <= TOL.

The algorithm is built in such a way that PCON >= SUM (MAX (WLINE – VEC, 0)) and never PCON < SUM (MAX (WLINE – VEC, 0)). VEC must be a row vector representing a noise level. PCON and TOL must be scalars in the same units as VEC.

(i) Find the length (N) of VEC.
N = length (vec)
(ii) Determine the initial waterline level, WLINE.
wline = min(vec) + pcon/N
(iii) Measure the total power of current waterline.
ptot = sum(max(wline - vec,0))
(iv) Repeat step (a) & (b) until
abs (pcon - ptot) > tol
a) wline = wline + (pcon - ptot)/N
b) ptot = sum(max(wline - vec,0))
The water filling algorithm is shown in fig. 2.

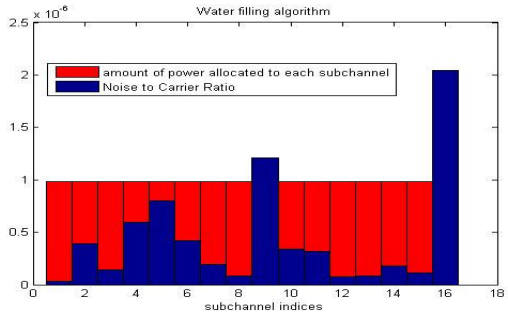


Figure 2. Water filling algorithm.

where U and V are unitary matrices and V^h is the hermitian of V. U has dimension of R x R and V has dimension of T x T. $\sum$ is a T x R matrix. If T = R, then $\sum$ become a diagonal matrix. If T > R, it is made of R x R diagonal matrix followed by T – R zero columns. If T < R, it is made of T x T diagonal matrix followed by R – T zero rows. This operation is called the singular value decomposition of H [9, 19].

In case, where T ≠ R, the number of spatial channels become restricted to the minimum of T and R. If the number of transmit antennas is greater than the receive antennas (T > R), U will be an R x R matrix, V will be a T x T matrix and $\sum$ will be made of a square matrix of order R followed by T-R zero columns [9, 19].

IV. MODELLING AND SIMULATION

A. OFDM Simulation Model

The OFDM system is modeled using MATLAB and is shown in Fig. 3[20]. The data to be transmitted on each carrier is differentially encoded with previous symbols, then mapped into a Phase Shift Keying (PSK) format. The data on each symbol is then mapped to a phase angle based on the modulation method such as QPSK. The serial data stream is formatted into the word size required for transmission, e.g. 2 bits/word for QPSK, and shifted into a parallel format. Zero padding has used in our system to increase sampling rates for better resolution of signals. After the required spectrum is worked out, an inverse Fourier transform is performed to find the corresponding time waveform. The guard period is then added to the start of each symbol. After the guard has been added, the symbols are then converted back to a serial time waveform. This is then the base band signal for the OFDM transmission.

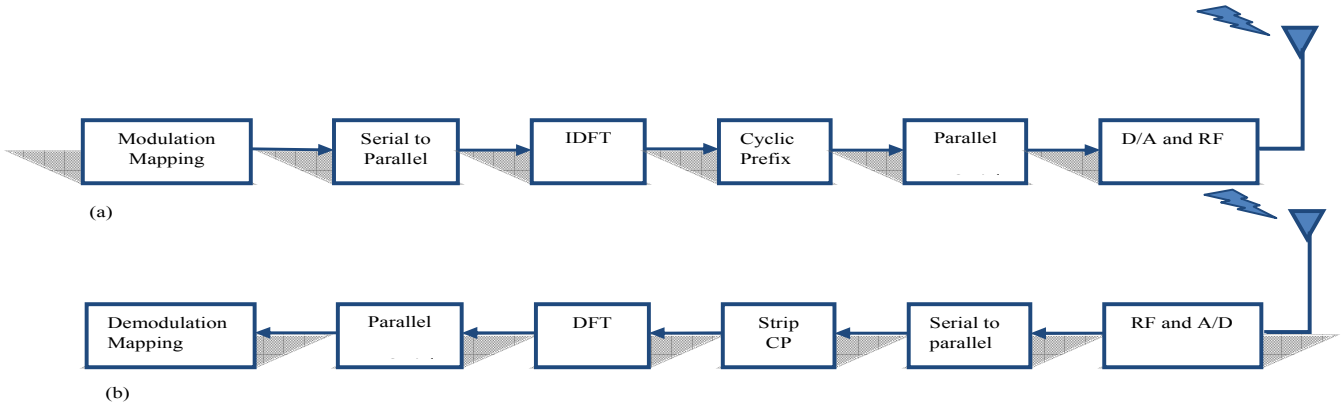


Figure 3. OFDM model used for simulations: (a) OFDM transmitter and (b) OFDM receiver

G. Singular Value Decomposition

The SVD technique decouples the channel matrix in spatial domain in a way similar to the DFT decoupling the channel in the frequency domain. The channel matrix H is the T x R channel matrix. If H has independent rows and columns, SVD yields:

$$H = U \sum V^h \quad (5)$$

A channel model is then applied to the transmitted signal. The model allows for the signal to noise ratio (SNR), multipath, and peak power clipping to be controlled. The SNR is set by adding a known amount of white noise to the transmitted signal. Multipath delay spread is then added by simulating the delay spread using an FIR filter. The length of the FIR filter represents the maximum delay spread, while the coefficient amplitude represents the reflected signal magnitude.

The receiver basically does the reverse operation to the transmitter. In the receiving side, the model recovers the input data, and performs an analysis to determine the transmission error rate. Table I represents the OFDM system parameters used for the simulation.

TABLE I. OFDM SYSTEM PARAMETERS USED FOR SIMULATION

Parameter	Value
Carrier Modulation	QPSK
FFT Size	128
Number of Carriers	120
Guard Time	32 samples (25%)
Guard Period Type	Full cyclic extension of the symbol

B. OFDM Simulation Result

The simulation model accepts inputs as text or audio files, binary, sinusoidal, or random data. The channel simulation allows examination of common wireless multipath channel characteristics such Rayleigh fading channel with various Doppler Effect. Fig. 4 represents data streams after modulation mapping (at transmission end). Fig. 5 shows amplitude and phase spectrum of transmitted OFDM signal. Fig. 6 and Fig. 7 depict amplitude and phase spectrum of received OFDM signal with Dophlar spread 15 and 90 respectively. And, Fig. 8 illustrates the received signal spectrum after FFT.

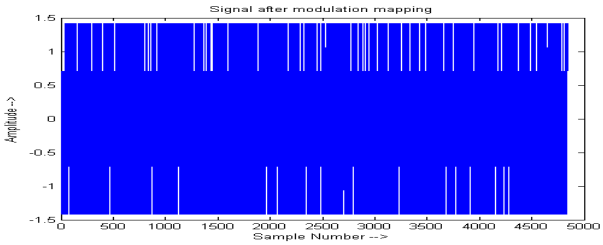


Figure 4. Data streams after modulation mapping (At transmission end).

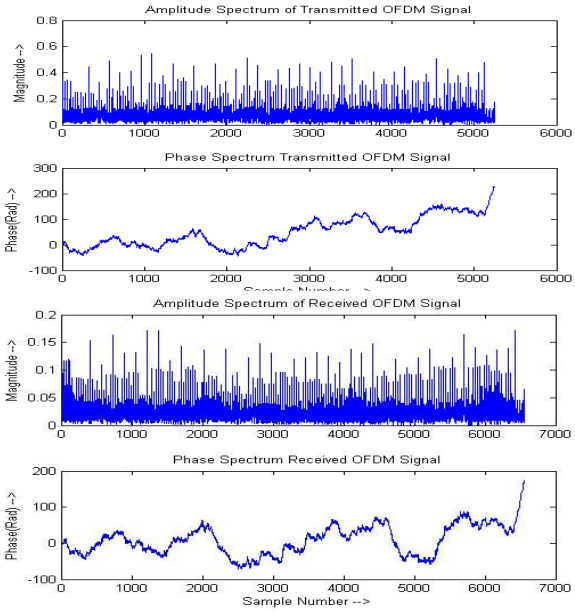


Figure 6. Amplitude and phase spectrum received OFDM signal (Dophlar spead=15).

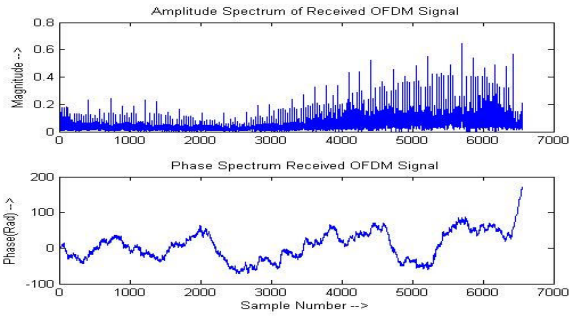


Figure 7. Amplitude and phase spectrum received OFDM signal (Dophlar

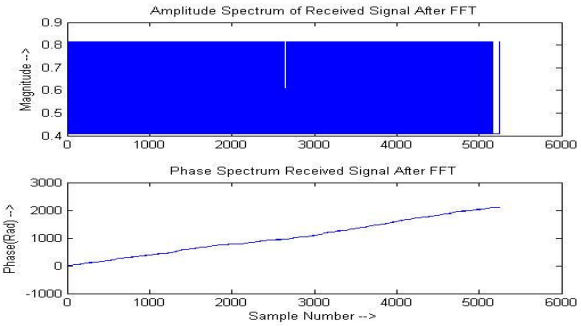


Figure 8. Received signal spectrum after FFT.

C. MIMO- OFDM Simulation Modelling

In this simulation, a highly scattered environment is considered. The capacity of a MIMO channel is analyzed with the antenna configuration as shown in Table II. Each channel is considered as a parallel flat fading channel. The power in a parallel channel (after decomposition) is distributed as water filling algorithm. Channel matrix H is measured using Rayleigh distribution function. This simulation computes channel capacity and PDF (probability density function) of elements in SVD of matrix H, by varying the SNR from -10 dB to 20 dB, where 104 iterations are performed.

TABLE II. TRANSMITTING AND RECEIVING ANTENNA COMBINATION

Combination	No. of tx antenna	No. of rx antenna
1	1	1
2	2	2
3	2	3
4	3	2
5	4	4

D. MIMO-OFDM Simulation Results

The channel capacity vs SNR curve of different MIMO systems is shown in Fig. 9. Fig. 10 represents the power spectral density (PDF) vs SNR. These graphs depict that the 4 x 4 MIMO systems provides better channel capacity and PDF than other combinations. This indicates that a higher order MIMO system increases the system performance. It is interesting to note that the system performance remains almost the same when the number of transmitter and receiver antennas is altered (2x3 MIMO and 3x2 MIMO systems).

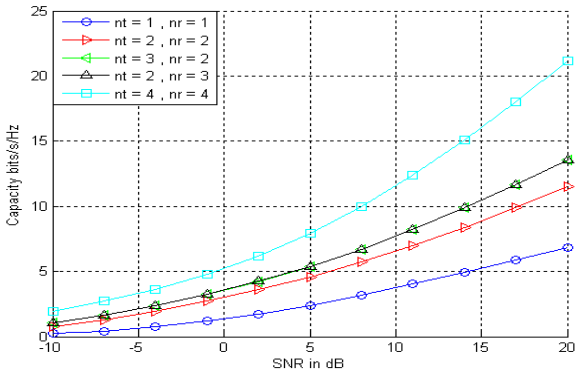


Figure 9. Comparison of channel capacity of different MIMO system.

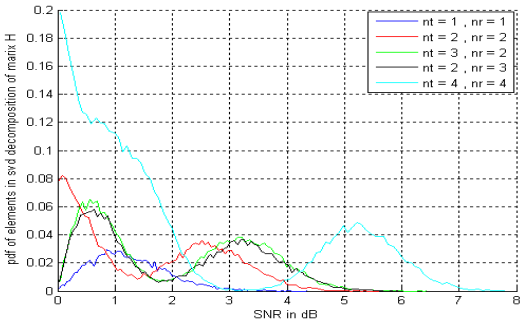


Figure 10. Comparison of PDF of different MIMO system.

V. CONCLUSION

An OFDM link was demonstrated through computer simulations and practical tests performed on a low bandwidth base-band signal. Four main performance criteria were tested, which include OFDM's tolerance to multipath delay spread, channel noise, peak power clipping and start time error. Several other important factors affecting the performance of OFDM have only been partly measured. In the past, there were a lot of problems with multiple wave propagation which led to creation of ISI. The MIMO has helped to reduce the ISI problem. With the implementation of MIMO-OFDM, the probability that the transmission arrives at the receiver with little or no error is greatly increased compared to the rest of the transmission techniques. In this system, the capacity is increased significantly by transmitting the different streams of data through different antennas at a same carrier frequency.

REFERENCES

[1] H. Bolcskei, "MIMO-OFDM wireless systems: basics, perspectives, and challenges," *Wireless Communications, IEEE,* vol. 13, pp. 31-37, 2006.
[2] R. Rao, "Impact of Phase Noise in MIMO-OFDM Systems," 2007.
[3] Y. Li and N. Sollenberger, "Adaptive antenna arrays for OFDM systems with cochannel interference," *Communications, IEEE Transactions on,* vol. 47, pp. 217-229, 2002.
[4] A. Goldsmith and S. Chua, "Variable-rate variable-power MQAM for fading channels," *Communications, IEEE Transactions on,* vol. 45, pp. 1218-1230, 2002.
[5] J. Huang and S. Signell, "The application of rate adaptation with finite alphabet in MIMO-OFDM," 2005, pp. 946-949.
[6] Y. Jiang, *et al.*, "Two-dimensional water-filling power allocation algorithm for MIMO-OFDM systems," *SCIENCE CHINA Information Sciences,* vol. 53, pp. 1242-1250, 2010.
[7] E. Whu, "MIMO-OFDM Systems for High Data Rate Wireless Networks," *Proj. Report for EE360 Advanced Wireless Networks: MIMO-OFDM Wireless Networks*.
[8] Y. Wang, *et al.*, "Singular Value Decomposition hardware for MIMO: State of the art and custom design," Cancun, 2010, pp. 400-405.
[9] D. Senaratne and C. Tellambura, "Generalized singular value decomposition for coordinated beamforming in MIMO systems," Miami, FL, 2010.
[10] W. Liejun, "An Improved Water-filling Power Allocation Method in MIMO OFDM Systems," *Information Technology Journal,* vol. 10, pp. 639-647, 2011.
[11] A. Gorokhov, *et al.*, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *Signal Processing, IEEE Transactions on,* vol. 51, pp. 2796-2807, 2005.
[12] V. R. Anreddy and M. A. Ingram, "Capacity of measured ricean and rayleigh indoor MIMO channels at 2.4 GHz with polarization and spatial diversity," in *Wireless Communications and Networking Conference, 2006. WCNC 2006. IEEE*, 2006, pp. 946-951.
[13] S. Venkatesan, *et al.*, "Capacity of a Gaussian MIMO channel with nonzero mean," 2004, pp. 1767-1771.
[14] T. Kan, *et al.*, "MIMO Channel Capacity of a Measured Radio Channel for Outdoor Macro Cellular Systems at 3GHz-Band," in *Vehicular Technology Conference Fall (VTC 2009-Fall), 2009 IEEE 70th*, 2009, pp. 1-5.
[15] S. Sumei, *et al.*, "Precoding for Asymmetric MIMO-OFDM Channels," in *Communications, 2006. ICC '06. IEEE International Conference on*, 2006, pp. 3117-3122.
[16] J. Wang, *et al.*, "Capacity of MIMO-OFDM system with multi-path fading channel," Beijing, 2009.
[17] A. Maaref and S. Aissa, "Combined adaptive modulation and truncated ARQ for packet data transmission in MIMO systems," 2005, pp. 3818-3822.
[18] G. Scutari, *et al.*, "The MIMO iterative waterfilling algorithm," *Signal Processing, IEEE Transactions on,* vol. 57, pp. 1917-1935, 2009.
[19] G. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *Numerische Mathematik,* vol. 14, pp. 403-420, 1970.
[20] I. Khan, *et al.*, "Capacity and performance analysis of space-time block coded MIMO-OFDM systems over Rician fading channel," *International Journal of Electrical and Computer Engineering,* 2009.

FPGA Based Cipher Design & Implementation of Recursive Oriented Block Arithmetic and Substitution Technique (ROBAST)

Rajdeep Chakraborty
Dept. of Computer Science & Engineering,
Netaji Subhash Engineering College, Garia, Kolkata-700152,
West Bengal, India.

JK Mandal, Professor,
Dept. of Computer Science & Engineering,
University of Kalyani, Nodia, West Bengal, India.

***Abstract* - Proposed FPGA based technique considers a message as a binary string on which ROBAST is applied. A block of n-bits is taken as an input stream, where n ranges from 8 to 256 – bit, then ROBAST is applied in each block to generate intermediate stream, any one intermediate stream is considered as a cipher text. The same operation is performed repeatedly on various block sizes. It is a kind of block cipher and symmetric in nature hence decoding is done in similar manner. This paper also presents an efficient hardware realization of the proposed technique using state-of-the-art Field Programmable Gate Array (FPGA). The technique is also coded in C programming language and Very High Speed Integrated Circuit Hardware Description Language (VHDL). Various results and comparisons have been performed against industrially accepted RSA and TDES. A satisfactory results and comparisons are found.**

Keywords – VHDL; FPGA; RTL; Block Cipher; Session key and Private Key; Cryptography; Symmetric/Private key cryptosystem.

I. INTRODUCTION

Transmission of sensitive electronic information [7] from and all around the globe has emphasizes the need of fast & secure network [2,3,5]. For achieving this secrecy, integrity and confidentiality, cryptographic techniques [1,2,3,14] are the tools. To achieve high performance it is highly recommended to implement the cryptographic techniques in hardware. A promising solution that combines high flexibility with the speed and physical security of Application Specific Integrated Circuits (ASIC) [7,8,9,10] is the implementation of cryptographic technique on state-of-the-art re-configurable devices such as Field Programmable Gate Array (FPGA) [7,8,9,10]. Sub-Section A, discussed the framework of the scheme along with private/symmetric key cryptosystem [4,7].

Private Key / Symmetric Cryptosystem

The aim is to develop an efficient crypto hardware. The figure 1 illustrates the conventional encryption model [2,7]. The main objective is to convert the intelligent plain text [2,3,4,5,7], X, to a non-sense cipher text [2,3,4,5,7], Y, using a single key [2,3,4,5,7], K. This process is the encoding/encryption [1,2,14], E, and the decoding/decryption [1,2,14], D, is performed similarly in case of symmetric or the opposite in other private key algorithms. The Session Key [1,2,14], K must be sent through a secured channel and cipher text, Y, may be sent through unsecured channel. Cryptanalyst [1,2,7,14] are the entity who attempts to discover plain text, X, and or key, K.

The Section II illustrates the principle of ROBAST, Section III gives the key generation process, result and simulations are given in Section IV, A brief analysis is given in Section V, Section VI draws the conclusion and finally the list of the references are given.

II. PRINCIPLE OF ROBAST

The message can be considered as blocks of bits with different block size [1,7,14] like 8, 16, 32, 64, 128 & 256 bits. The rules to be followed for generating a cycle are as follows:

1. Consider any source stream [1,2,7,14] containing finite number of bits (where $N=2^n$, n =3 to 8) and divide it into two equal parts.
2. Make the source stream into paired form so that a pair can be used for the operation.
3. Perform the modulo-4 addition [6,13] between the first and second pair, second and third pair, and so on of the source stream, to obtain the first intermediate block.
4. The same process is repeated recursively [7,11,12] between second and first, third and second, fourth and third and so on of the source stream, to generate the next intermediate block.

This process is repeated until the source stream is generated. After a finite number of iterations source stream is regenerated. So, decryption is basically the iteration of the same process. In this proposed technique the modulo addition with substitution and permutation is given but to enhance the security further other arithmetic operations has also been implemented in this technique. Sub-Section A illustrates the scheme numerically and that of Sub-Section B outlines the implementation issues.

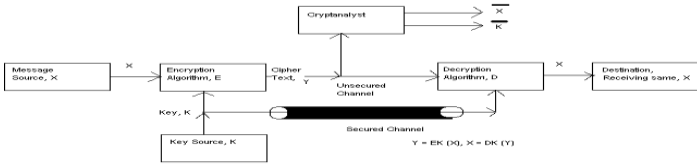


Figure 1: Model of Conventional Cryptosystem

A. Example

Consider the block S = 10010011 of size 8 bits. The Flow diagram to show how positions of the bits of S and the different intermediate blocks can be reoriented with the key values to complete the cycle is shown in figure 2. In this diagram, each arrow indicates positional orientation of a bit during iteration. Therefore the final cipher text is S'=00011001.

B. Implementation of ROBAST

The technique executes modulo addition between two blocks, the first iteration performs in forward basis and then backward operation is performed. Next, final permutation is done to get the final cipher text. This technique has been implemented in C [11,12,13] and then feasibility study has been performed. Finally, FPGA [8,9,10] based implementation has been done in VHDL [8,9,10]. In both implementation, the technique takes input from file as a source stream and encryption is performed. The cipher text generated is finally written in another file [7,10,11,12]. The data blocks (8, 16, 32, 64, 128 and 256-bits) from the input file have been stored in array. Then encryption is performed and also stored in array. The reading and writing of data from and in file is based on 8-bit ASCII codes [7,11,12]. XilinX [8,9,10] software has been used for writing codes in VHDL. The encryption/decryption entity input bit vector (16-bit), output bit vector (16-bit), key bit vector (8-bit) and EN_DN signal. If EN_DN = 1 then encryption is performed else decryption is performed. Figure 3 gives the main ROBAST entity coded in VHDL.

The above operations discussed are substitute technique [6,7,13] followed by permutation technique [6,7,13] has been performed by orientation of bits based on the session key. Therefore, these resultant blocks of stream can be considered as cipher text.

III. The Key Generation Process

The key generation process depends on block size, iteration of each block and final permutation performed. Thus, in the proposed scheme, eight rounds have been considered, each for 2, 4, 8, 16, 32, 64, 128, and 256-block size. As mentioned in Section II, each round is repeated for a finite number of times and the number of iterations will form a part of the encryption-key. Although the key may be formed in many ways, for the sake of brevity it is proposed to represent the number of iterations in each round by a 16-bit binary string. The binary strings are then concatenated to form a 128-bit key for a particular key. Example in Sub-Section A illustrates the key generation process. Sub-Section B describes the modulo addition, which is an important operation in the technique and should be taken into account while forming the session key.

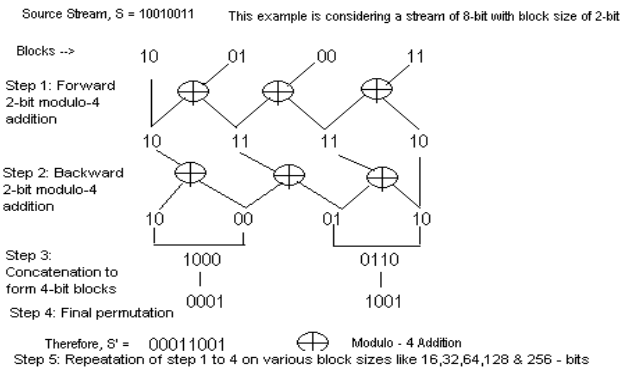


Figure 2: Flow Diagram of proposed technique, ROBAST

```
library std;
library ieee;
use ieee.std_logic_arith.all;
use work.pack.all;
use std.textio.all;
use ieee.std_logic_TEXTIO.all;
entity ROBAST_VHDL is

Port ( input_bits : in  BIT_VECTOR (16 downto 1);
output_bits : out  BIT_VECTOR (16 downto 1);
key_bits : in  BIT_VECTOR (8 downto 1);
EN_DN : in  BIT);
end ROBAST_VHDL;

architecture Behavioral of ROBAST_VHDL is

begin

process(EN_DN)

variable varin_bits,varout_bits: bit_vector(16 downto 1);

begin

if (EN_DN='1')then
varin_bits:=input_bits;
AA: ROBAST_Encryption(varin_bits,key_bits,varout_bits);
output_bits<=varout_bits;
else
BB: ROBAST_Decryption(varin_bits,key_bits,varout_bits);
output_bits<=varout_bits;
end if;

end process;

end Behavioral;
```

Figure 3: ROBAST Entity and its function

A. Example

Consider a particular session where the source file is encrypted using iterations for block sizes 2, 4, 8, 16, 32, 64, 128, and 256 bits, respectively. Table I shows the corresponding binary value [7,8,10,13] for the number of iterations in each round. The binary strings are concatenated together to form the 128-bit binary string:

1100001101100101110000101100111010111111001101101010110110011011101101001010101000010000111000010000001010110010000000000001001000.

This 128-bit binary string will be the encryption-key for this particular session. During decryption, the same key is taken to iterate each round of modulo-subtraction for the specified number of times and reverse permutation.

B. Modulo Addition

An alternative method for modulo addition is proposed here to make the calculations simple. The need for computation of decimal equivalents of the blocks is avoided here since it may generate large decimal integer values for large binary blocks. The method proposed here is just to discard the carry out of the MSB after the addition to get the result. For example, if we add 1101 and 1001 we get 10110. In terms of decimal values, 13+9=22. Since the modulus of addition is 16 (24) in this case, the result of addition should be 6 (22-16=6). Discarding the carry from 10110 is equivalent to subtracting 10000 (i.e. 16 in decimal). So the result will be 0110, which is equivalent to 6 in decimal. The same is applicable to any block size.

IV. Results and Simulation

Any cryptographic technique is to be accepted, a satisfactory results are very much required. Proposed technique has been tested for feasibility both in terms of algorithmic parameters and cryptographic parameters. Sub-Section A gives the time complexity results [7,11,12,13], Sub-Section B tests for non-homogeneity using Chi-Square values [1,6,14,15] and degree of freedom [1,6,14,15], Sub-Section C illustrates the frequency distribution [1,6,14,15] of ASCII characters [7,11,12], Sub-Section D test for cryptanalysis using avalanche ratio [2,3,4,5] and finally Sub-Section E gives the FPGA-based simulation results [8,9,10]. All these results are against well known and industrially accepted RSA and TDES [1,2,3,4,13,14]. For the shake of brevity 20 (twenty) sample files of different types has been taken for these results. The Section V briefly analyses all these results.

A. Time Complexity

Time complexity is based on encryption time and decryption time [1,2,14]. Encryption time is the time required to encrypt a source file and decryption time is the time to decrypt the cipher text file to get the original file. Table II gives the time complexities and Figure 4 illustrates the same. This test is in terms of efficient algorithmic parameter.

B. Tests for Non-Homogeneity

Test for non-homogeneity has also been done using Chi-Square value and degree of freedom; this is one of the important cryptographic parameters. Chi square value is the statistical value between source file and encrypted files, which gives the difference. Degree of freedom in the character distribution of the above said files. Table III gives the Chi-Square value and Figure 5 illustrates the same.

C. Frequency Distribution

The frequency distribution is the distribution of the all-256 ASCII characters in the respective files. This is also a cryptographic parameter, which measures the degree of cryptanalysis. Figure 6 illustrates the various frequency distribution results found after implementation of respective algorithms/techniques.

D. Avalanche Ratio

The avalanche ratio is the ratio between the modified results to the original result. The avalanche ratio is obtained by modifying 2-3 bits/bytes in the encryption key as well as in source files. It's a strong cryptographic parameter and this may be conceptualize with the avalanche occurs in hill area. Table IV gives the avalanche ratio values of ROBAST.

TABLE I. Representation of Number of Iterations in Each Round by Bits

Round	Block Size	Number of Iterations	
		Decimal	Binary
8.	256	50021	1100001101100101
7.	128	49870	1100010011001110
6.	64	48950	1011111100110110
5.	32	44443	1010110110011011
4.	16	46250	1011010010101010
3.	8	4321	0001000011100001
2.	4	690	0000001010110010
1.	2	72	0000000001001000

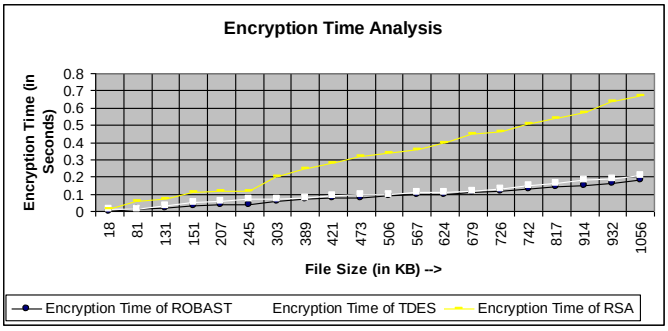


Figure 4 (A): Encryption Time

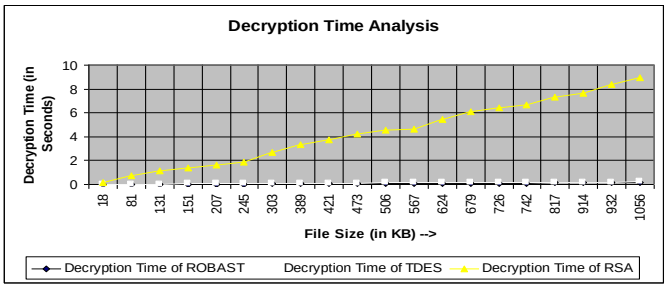


Figure 4 (B): Decryption Time

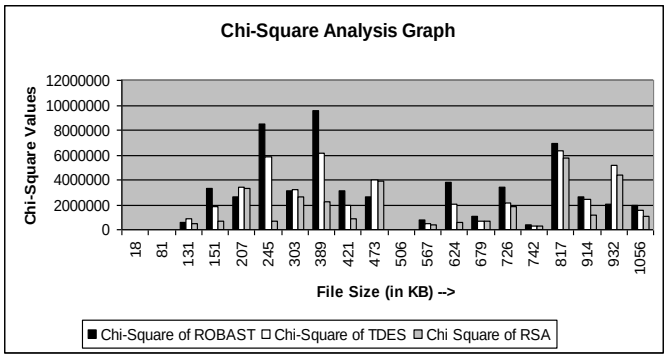


Figure 5: Chi-Square Analysis Graph

TABLE II. TIME COMPLEXITY ANALYSIS

Serial no.	File Name	File Size (in Kilo Bytes)	Encrypted time (in second)			Decryption time (in second)		
			ROBAST	TDES	RSA	ROBAST	TDES	RSA
01	Poppy.jpg	18	0.00	0.01	0.01	0.01	0.02	0.15
02	07.jpg	81	0.01	0.01	0.06	0.02	0.03	0.71
03	Sqmapi.dll	131	0.02	0.03	0.07	0.03	0.03	1.15
04	Jview.exe	151	0.03	0.05	0.11	0.03	0.06	1.36
05	Gender.txt	207	0.04	0.06	0.12	0.04	0.07	1.61
06	Pod.exe	245	0.04	0.07	0.12	0.04	0.08	1.86
07	Devices.txt	303	0.06	0.07	0.20	0.05	0.09	2.71
08	Dtliteui.dll	389	0.07	0.08	0.25	0.06	0.10	3.34
09	Vssapi.dll	421	0.08	0.09	0.28	0.07	0.11	3.73
10	Names.txt	473	0.08	0.10	0.32	0.08	0.11	4.25
11	Photo000.jpg	506	0.09	0.10	0.34	0.08	0.13	4.54
12	Uninst.exe	567	0.10	0.11	0.36	0.10	0.14	4.67
13	Iexplore.exe	624	0.10	0.11	0.40	0.11	0.14	5.43
14	Cordic.pdf	679	0.11	0.12	0.45	0.11	0.16	6.10
15	Iedvtool.dll	726	0.12	0.13	0.46	0.12	0.17	6.40
16	Guide.pdf	742	0.13	0.15	0.51	0.12	0.18	6.67
17	Setuplog.txt	817	0.14	0.16	0.54	0.14	0.18	7.34
18	Adobearm.exe	914	0.15	0.18	0.57	0.15	0.19	7.68
19	De10.txt	932	0.16	0.19	0.64	0.16	0.20	8.39
20	Omat.doc	1056	0.18	0.21	0.67	0.19	0.23	8.92

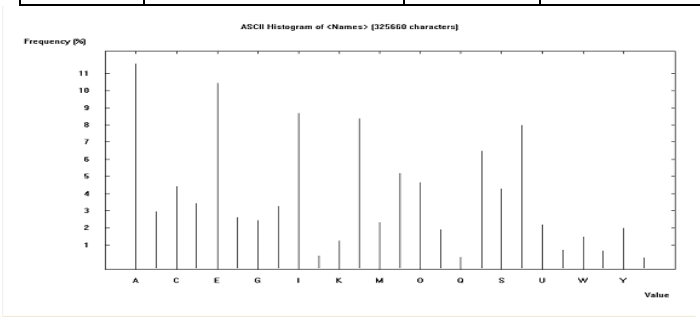


Figure 6 (A): Frequency distribution of Source File

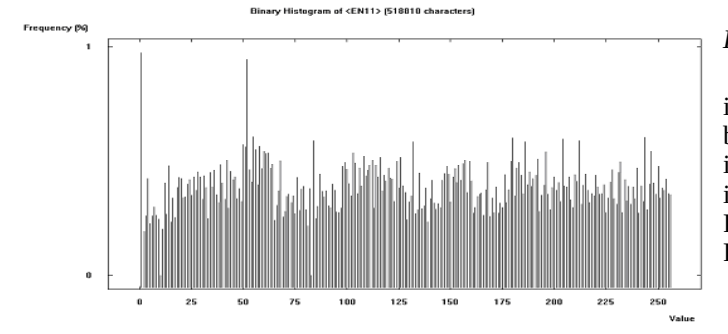


Figure 6 (B): Frequency distribution of ROBAST encrypted file

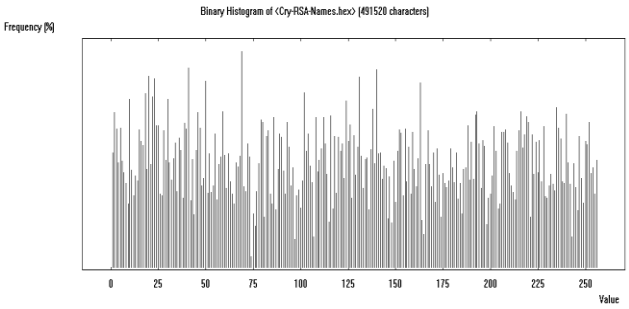


Figure 6 (C): Frequency Distribution of RSA encrypted file

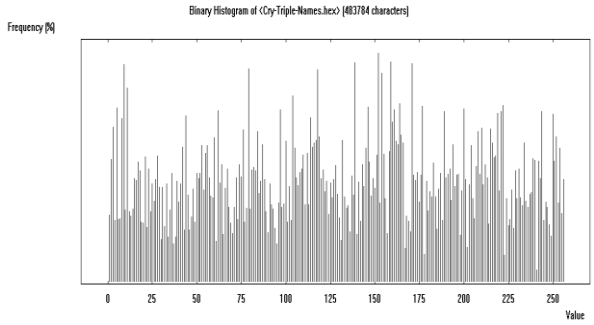


Figure 6 (D): Frequency distribution of Triple-DES encrypted file

E. FPGA-Based Simulation Result

This Section gives some of the results found after implementing the proposed technique in VHDL. This code has been simulated and synthesized in XilinX. The main objective is to find an efficient FPGA-based cryptographic technique for implementation in embedded systems. The Figure 7 gives the RTL schematic [8,9,10] of the proposed technique and the Figure 8 gives the chip diagram for Spartan 3E [8,9,10].

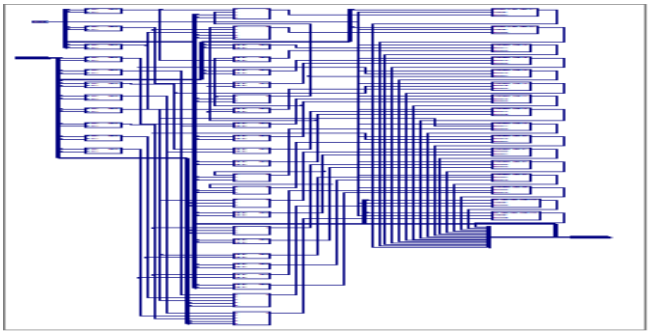


Figure 7: RTL Schematic

TABLE III. CHI-SQUARE AND DEGREE OF FREEDOM VALUES

Serial no.	File Name	File Size (in Kilo Bytes)	Chi-Square Value			Degree of Freedom		
			ROBAST	TDES	RSA	ROBAST	TDES	RSA
01	Poppy.jpg	18	6472	4572	5668	253	255	255
02	07.jpg	81	4407	2943	2654	253	255	255
03	Sqmapi.dll	131	560357	890752	447984	253	255	255
04	Jview.exe	151	3307374	1847893	685963	253	255	255
05	Gender.txt	207	2679799	3426290	3318506	253	93	84
06	Pod.exe	245	8495675	5810912	694410	253	255	255
07	Devices.txt	303	3131296	3220112	2667664	254	66	88
08	Dtliteui.dll	389	9559993	6190253	2216429	253	255	255
09	Vssapi.dll	421	3102369	1980059	906300	253	255	255
10	Names.txt	473	2590855	4044603	3896171	253	88	90
11	Photo000.jpg	506	38465	31719	30353	253	255	255
12	Uninst.exe	567	776122	512668	342450	253	255	255
13	Iexplore.exe	624	3799155	2043250	588049	253	255	255
14	Cordic.pdf	679	1065255	684198	686392	253	255	255
15	Iedvtool.dll	726	3422000	2192824	1845040	253	255	255
16	Guide.pdf	742	420469	320825	311524	253	255	255
17	Setuplog.txt	817	6904009	6340148	5737525	253	255	255
18	Adobearm.exe	914	2625926	2458497	1196585	253	255	255
19	De10.txt	932	2043522	5194261	4407281	251	86	11
20	Omat.doc	1056	1968558	1516848	1082800	253	255	255

TABLE IV. AVALANCHE RATIO OF ROBAST ENCRYPTED FILES

Serial no.	File Name	File Size (in Kilo Bytes)	Avalanche Ratio of ROBAST encrypted files (in %)
01	Poppy.jpg	18	96.25
02	07.jpg	81	99.69
03	Sqmapi.dll	131	99.93
04	Jview.exe	151	99.96
05	Gender.txt	207	97.72
06	Pod.exe	245	99.84
07	Devices.txt	303	98.22
08	Dtliteui.dll	389	99.97
09	Vssapi.dll	421	99.98
10	Names.txt	473	99.95
11	Photo000.jpg	506	99.91
12	Uninst.exe	567	99.98
13	Iexplore.exe	624	99.99
14	Cordic.pdf	679	99.70
15	Iedvtool.dll	726	99.97
16	Guide.pdf	742	99.56
17	Setuplog.txt	817	99.56
18	Adobearm.exe	914	99.98
19	De10.txt	932	06.83
20	Omat.doc	1056	99.73
Average Avalanche Ratio			94.84
The avalanche ratio is obtained by modifying 2-3 bits/bytes in the encryption key as well as source files			

V. ANALYSIS OF THE RESULTS

Analyzing all the results presented in the result Section(s), following are the points obtained on the proposed technique:

1. The encryption time and decryption time varies linearly with the file sizes. Also the time complexity of ROBAST is quite less than RSA, but it's slight less than TDES.
2. Considering the Chi-Square values, the proposed technique, ROBAST, is most non-homogeneous than that of RSA and TDES. But, there is no substantial result found in terms of degree of freedom because all three (ROBAST, RSA, TDES) have almost same value.
3. Result for the frequency distribution illustrates the ASCII characters are well distributed in ROBAST. The well distribution was also found for RSA and TDES. So, the frequency distribution result is at par with that found in Chi-Square and degree of freedom values.
4. A very good result has been obtained in the avalanche ratio of the proposed technique. The average avalanche ratio is 94.84. So, cryptanalysis is quite difficult.
5. The RTL diagram signifies that the proposed technique has been successfully implemented in VHDL and the same is illustrated for Spartan 3E FPGA. If we closely look at, there are 29 Look-Up-Tables (LUT s) [8,9] used for this technique.

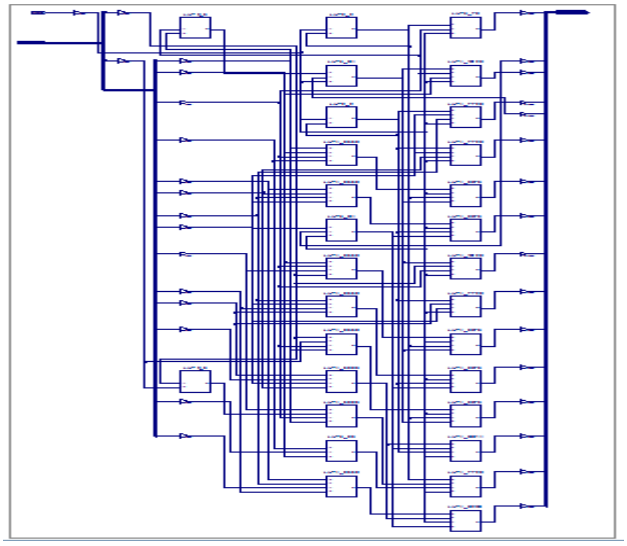


Figure 8: Spartan 3E Schematic

Sub-Section A. gives the application of this proposed technique, ROBAST and along with future scope of work.

A. Application of ROBAST

The following are some of the avenues of application for the proposed technique, ROBAST: -

1. Since the proposed system is simple, fast and low power consumption based crypto solution, it can be used in various embedded systems.
2. This technique may be used to develop electronic codebooks.
3. It's can also be used to develop a private network and master-key-based applications.
4. This proposed FPGA-based system may be used in hardware applications such as switch, gateways and routers.

The FPGA implementation of Vertex series with increased block length and also with low computational complexity is the future scope of the work.

VI. Conclusion

The proposed technique given here is easily implemented in high-level language and in VHDL. This technique is very easy and it's implemented in FPGA-based systems, the goal of fast execution and strong cryptanalysis requirements are also obtained here. This technique can be fabricated in chip to be used in embedded systems. The main goal of the author(s) is to develop an efficient FPGA-based crypto hardware and this paper is the first step towards this.

Acknowledgment

The authors express their deep sense of gratitude to The Department of Computer Science and Engineering, Netaji Subhash Engineering College, Garia, Kolkata, West Bengal, India and The Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nodia, West Bengal, India.

References

[1] Rajdeep Chakraborty, Dr. J.K.Mandal, "A Microprocessor-based Block Cipher through Rotational Addition Technique (RAT)", ICIT – 2006 18-21 December, 2006, Bhubaneswar, INDIA.

[2] W. Stallings, Cryptography and Network Security: Principles and Practices, Prentice Hall, Upper Saddle River, New Jersey, USA, Third Edition, 2003.

[3] B. Schneier. Applied Cryptography. John Wiley & Sons Inc., New York, New York, USA, 2nd edition,1996.

[4] U.S. Department of Commerce/National Institute of Standard and Technology. FIPS PUB 197, Specification for the Advanced Encryption Standard (AES), November 2001. Available at http://csrc.nist.gov/encryption/aes.

[5] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. Handbook of Applied Cryptography. CRC Press, Boca Raton, Florida, USA, 1997.

[6] A.M. Goon, M.K. Gupta, B. Dasgupta, Fundamentals of Statistics, Vol. 1, The World press Ltd.

[7] Dictionary of Computers and Information Technology Terms, 1st edition, low point, Kolkata, India.

[8] FPGA- Based System Design by W. Wolf, Pearson Education.

[9] Embedded Core Design with FPGA's by Z. Navavi, TMGH.

[10] AVHDL: Premier by J. Bhasker, Pearson Education

[11] Programming in C by Balaguruswamy, India.

[12] Pointers in C by Y Kanitkar, India

[13] The software cryptographic tools for educational purpose available at http://www.cryptool.com/

[14] S. Mal, J.K. Mandal, S. Dutta, "A Microprocessor Based Encoder for Secured Transmission", Proceedings of the National Conference on Intelligent Computing on VLSI, Kalyani Govt. Engg. College, 16-17 February, 2001, pp 164-169.

[15] Number theory home page for secured key generations http://www.numbertheory.org/ntw/web.html

[16] Pasha, A., & Gafoor, A. (2011). Transparent Data Encryption- Solution for Security of Database Contents. *International Journal of Advanced Computer Science and Applications - IJACSA, 2*(3), 25-28.

[17] Meshram, C. (2010). Modified ID-Based Public key Cryptosystem using Double Discrete Logarithm Problem. *International Journal of Advanced Computer Science and Applications - IJACSA, 1*(6).

[18] Nath, J. (2011). Advanced Steganography Algorithm using Encrypted secret message. *International Journal of Advanced Computer Science and Applications - IJACSA, 2*(3).

AUTHORS PROFILE

RAJDEEP CHAKRABORTY, born on 23rd August ' 1978, did his Bachelor of Engineering (B.E.) in Computer Science and Engineering (CSE) from Utkal University, Bhubaneswar, India at 2002, then he did his Master of Technology (M. Tech) in Information Technology (IT) from Sikkim Manipal University of Health Medical and Technological Sciences (SMUHMTS), Gangtok, India at 2004, he is presently persuing PhD in Computer Science and Engineeing (CSE) from University of Kalyani, Kalyani, India, in the field of Cryptography. Presently he is assistant professor in the department of Computer Science and Engineering (CSE), Netaji Subhash Engineering College, Kolkata, India. He has almost 6 years of teaching and research expirence and the total number of his publication is eight, all in internatinal conferences and journals. The average makrs throughout his carrer is 71.0%.

JK MANDAL did his M. Tech in Computer Science & Engineering at University of Calcutta, then did his PhD in CSE, at Jadavpur University in the field of Data Compression and Error Correction techniques. Presently he is professor in Computer Science and Engineering department, University of Kalyani, India. He is also a life member of Computer Society of India (CSI) since 1992 and life member of Cryptology Research Society India. Moreover, he is also serving as a dean, faculty of Engineering, Technology and Managemat, at University of Kalyani. The fields of his work are Network Security, Steganography, Remote Sensing, GIS application and Image Processing. He has 23 years of Teaching and Research Experience and seven scholars are awarded PhD and now eight scholars are pursuing PhD under his guidance. The total number of his publication is 140.

Create a Virtual Mannequin Through the 2-D Image-based Anthropometric Measurement and Radius Distance Free Form Deformation

Sheng-Fuu Lin
Institute of Electrical Control Engineering
National Chiao Tung University
Hsinchu, Taiwan(ROC)

Shih-Che Chien
Institute of Electrical Control Engineering
National Chiao Tung University
Hsinchu, Taiwan(ROC)

***Abstract*—3-D human body models are used in a wild spectrum of applications, such as film and entertainment industry, that require images of human replicas, but the computer generated models of human body generally do not adequately model the complex human morphology. These models do not reflect the realistic anthropometric data and are not specific enough for commercial use. This paper presents an approach to adjust virtual mannequins through the use of anthropometric measurement data, which are obtained from 2-D image-base measurement. In this approach, a novel method which used the Chinese medicine acupuncture theory for fast position locating and human body slice model to approach circumferences is proposed to 2-D image-based anthropometric measurement. The measurement data are used in grouping 3-D scanned body objects into clusters. The virtual mannequins are then adjusted by using the measurement data of the standard model that belong to its cluster. In this way, the realistic accurate virtual mannequins are created.**

Keywords-human body; anthropometric measurement; free form deformation deformation

I. INTRODUCTION

3-D human body models are used in a wild spectrum of applications, such as film and entertainment industry, that require images of human replicas. In film and entertainment industry, the 3-D human body models, virtual mannequin, are employed in visualization and animation environment. Although there are many techniques in 3-D virtual environment, depending on the functionalities of each element, efficient shading, skinning, and motion algorithms, to simulate the physical properties, most of these techniques are focus on human body models. Because of it can be used to create the personalized virtual mannequins and imitate the movement of human. Internet applications such as 3-D games and advertisement videos contain personalized human body model. In particular, computer generated models of human body usually do not adequately model the complex human morphology and stand for a lot of applications.

The method used for creating or adjusting virtual human body model has been evolved. Two major steps, anthropometric measurement and adjustable virtual mannequin, are always utilized for creating the personalized 3-D human body model. The anthropometric measurements are provided the personalized body measurement data, such as height, chest circumference, waist circumference…etc, for mannequin adjustment. The adjustable virtual mannequin model provides a platform for adjusting 3-D human body model to approach the realistic body.

Anthropometry is a since of measurements which is used in order to establish the physical geometry, mass property, and strength capabilities of the human body [1]. In traditional anthropometric measurement methods, the measurement can be done by using a simple instrument such as tape and without complex measurement pre-setting [2]. Although the traditional anthropometry measurement is easy and convenient to use, the traditional measurement relies on manual operations that are inefficient and prone to errors. 2D image-based anthropometric measurement methods adopt two or more photographic, and through image process and geometric transformation, the anthropometric measurement data can be obtained from captured images. Meunier and Yin [3] proposed an anthropometric measurement system that can generate body measurements from two-dimension images. Hung et al. [4] used geometric shapes (ellipse and rectangular) to approximate the shape of critical part circumference. The ellipse shape was used to approximate neck, wrist, and palm circumferences, and the major and minor axes length were obtained from the front and side views.

Nowadays, the 3-D laser scanning technique makes it possible to digitalize the complete surface of human body and provides much richer information about the body shape than the traditional anthropometric measurements [5]. CAESAR, which stands for Civilian American and European Surface Anthropometry Resource, is the first large scale 3-D anthropometry survey project [6]. This approach provides a standard model of digitalized human body shape and opens up opportunities to extract new measurements for quantifying the body shape. The first attempt in processing 3-D anthropometric data for analyzing the body shape is to extract traditional anthropometric measurements from the scanned data [7]. Although working with the 3-D surface data has the advantage of being able to perform repeated measurements without the subject being present, the 3-D scanning equipment is not available in anytime and anywhere and the collected 3-D scanning data still needs complicated analysis to calculate human body anthropometric data.

The creating methodologies of 3-D human body can be divided into three parts, 3-D scanning, reconstruction, and example based [8][9]. The 3-D scanning method is described in past paragraph, which used precise optoelectronic measurement device for body shape scanning. In unknown human body model creating, the example based method is appropriated. In this paper, we concern about how to create virtual mannequin that more approaching realistic and saving more processing time.

The most of example based methods are adopted the space deformation methods. The space deformation method is the process of mapping some of vertices from one space to the other. Because of space mapping methodology, the wide range deformation of 3-D object can be developed. Barr's [10] presented a work on parametric shape and Scheeper's [11] is also used the parametric shape for modeling human body muscle. Sederberg and Parry proposed the Free Form Deformation (FFD) method that embedded the 3-D object into a cubic box, which had several control points to influence vertices of object, deforming the box will also deform the underlying object. Borrel [12] proposed a variant FFD for different type of object parts.

The adjustment of virtual human body model to fit real measurement data provides an important challenge to the research community. This paper presents an approach that adopts the benefit and avoids the disadvantage of 3-D scanned human model. The 2-D image-based anthropologic measurement [13], that using Chinese medicine acupuncture theory for position locating and human body slice model for circumference approaching, is applied for body measurement. In this approach, we propose a method Radius Distance Free Form Deformation (RD-FFD) for geometric deformation. The Radius Distance system is a coordinate in polar coordinate system (r,θ). Basically the radius distance is the scalar Euclidean distance between a point and origin of the system of coordinates. In RD-FFD system, the model divided into slices and a radius of each slice is scaled to achieve the deformation effect to instead of using polygonal methods. Since the deformation scales are along with the radius direction in radius distance system, the detail of human body deformation could be exhibited and help to save more computing time.

This paper is organized as follows. Section 2 provides a 2D image-based anthropometric measurement that used two side images, front and side, which captured by camera, and the human body slice model is supplied to approximate the critical circumference shape. In section 3, the method Radius Distance Free Form Deformation (RD-FFD) for geometric deformation is proposed. The experimental results and discussion are showed in section 4. Section 5 concludes this paper. The flowchart of this system illustrated in Fig. 1.

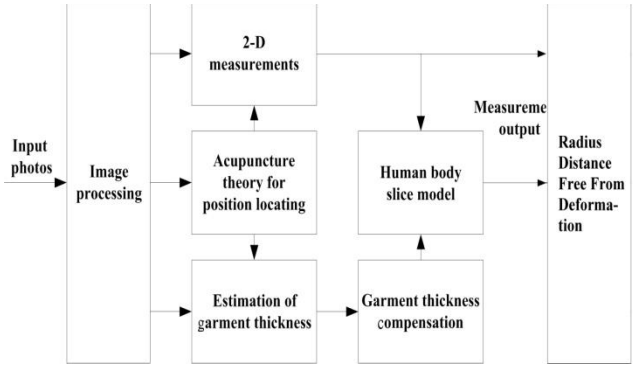


Figure 1. The flowchart of this approach

II. 2D IMAGE-BASED ANTHROPOMETRIC MEASUREMENT SYSTEM

The method of 2-D image based anthropometric measurement includes 2-D image preprocessing, Chinese medicine acupuncture theory for position locating, and human body slice model. At the first, the image preprocessing is used to process the images captured from camera. Secondly, the Chinese medicine acupuncture theory is adopted to locate the critical position because of its distinctive acupuncture point locating method of human body. After locating the critical position, the 2-D anthropometric data can be obtained by direct measurement. In order to acquire the circumference, the human body slice model is used to approximate the circumference shape. At the last, an efficient compensation system that aims to reduce the influence of subject wearing clothes for measurement is presented.

A. 2-D image preprocessing

The camera is set on tripod with about one meter off the ground, and dual-axis bubble levels on tripod help to keep the camera level. The system captures front and side images of objects standing with some required postures, as show in Fig. 2(a). The extraction of human body shape is performed by skin color extraction and automatic image segmentation [14][15], which uses color edge extraction and seeded region growing (SRG) to do image segmentation. Then, the full body silhouette shape can be obtained, as show in Fig. 2(b).

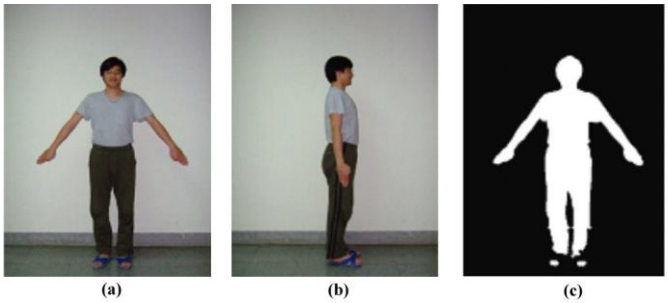


Figure 2. (a)(b) The postures of front and side view. (c) The human body segmentation of front view image.

B. Chinese medicine acupuncture theory for position locating

Acupuncture is the procedure of inserting filiform needles into various acupuncture points on the body to relieve pain or for therapeutic purposes. In treatment procedural, the therapist should find out the correct acupuncture point position corresponding to pain before inserting needle [16]. Hence, the acupuncture point location method is developed to locate the acupuncture points on human body fast and accurately. Five body measurement data (shoulder length, chest circumference, waist circumference, hip circumference, and leg length) are selected for discussion because there are the most commonly used in manufactory and general application, such as clothing size. For locating the corresponding position, six acupuncture points (Lian-Quan, Tian-Tu, Chien-Yu, Shan-Zhong, Shen-Que, and Qu-Gu) are taken to locate the critical positions and show as Fig. 3.

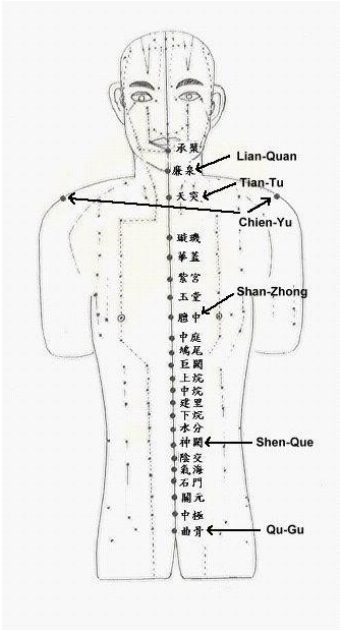


Fig. 3 Illustrate the position of acupuncture points on human body.

The measurement data can be separated into two groups, direct measurement (linear) and indirect measurement (circumference). In linear part, the measurement is to calculate the distance between two acupuncture points: shoulder length is the distance between two Chien-Yu acupuncture points or to calculate the distance between the acupuncture point and extremity: leg length is the distance between the Qu-Gu acupuncture point and sole of foot. Circumference cannot be measured directly using front and side view images only and must therefore be calculated using some form of mathematical model. In this paper, the human body slice model is adopted to approximate the circumference shape, and the circumference could be obtained through the calculation with two parameters, the corresponding width of front and side view image.

C. Human body slice model

Circumference cannot be measured directly using only 2D measurement, and therefore the mathematical model is applied to approximate the corresponding circumference. In this paper, the human body slice model which using Bezier curve [17] is used to approximate the circumference. In this paper, the cubic Bezier curve is adopted for basic curve to patch the piecewise curve. The cubic Bezier curve content with four control points: P0, P1, P2, and P3, which existing on the same plane of curve. The cubic Bezier curve B(t) starts from P0 to P1 and arrive at P3, and the function can be denoted as follow:

$$\begin{aligned} B(t) &= (1-t)^3 P_0 + 3(1-t)^2 t P_1 + 3(1-t)t^2 P_2 + t^3 P_3 \ , \ t \in [0,1]. \\ &= [t^3 \ t^2 \ t \ 1] \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix} \\ &= TB_c P \end{aligned} \quad (1)$$

The 3-D human body scanning data is the body digitalized data, and the slice data can be extracted with the plane which is concerned. The Taiwan human body bank (TAIBBK) [18][19], which stands for Industrial Technology Research Institute, Tshing Hua University, and Chang Gung University, is a large scale 3-D anthropometric measurement project. About 1100 civilians, between the age of 19 and 65 in Taiwan, were scanned. In generally, the shape of human body is virtually bilateral reflection symmetry with respect to the center mirror plane. Furthermore, the shapes of chest, waist, and hip are also symmetrically in slice human model. It also means that the circumference can be obtained, if the curve length of corresponding half-shape is known. Fig. 4(a) shows the whole and half chest slice shape of standard male build type, and confirms the symmetry of slice shape model. the piecewise polynomial curve based on cubic Bezier curve is used to approximate the curve of half-shape, and then the curve length can be calculated by using definite integral. Fig. 4(b) shows the curve approximation with two cubic Bezier curves, and two sets of control points are marked with blue circle and star. The red curve is actual curve of chest slice shape and the blue dotted curve is the approximation curve, which is similar to manual measurement.

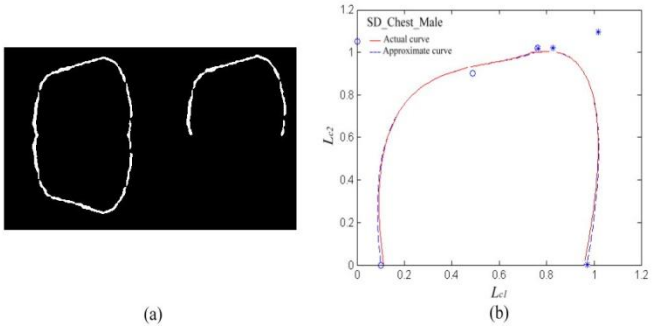


Figure 4. (a) The whole and half chest shape of standard male build type. (b) The approximation curve and actual curve of chest

Then, the expression of piecewise polynomial curve can be denoted as follow:

$$C(t) = C_{bz1} + C_{bz2} = TB_c(P_1 + P_2) \quad (2)$$

where P1 and P2 are the control point sets of anterior curve (Cbz1) and posterior curve (Cbz2) respectively. Now, we take

the chest circumference measurement for example. The expression of approximation curve, showed in Fig. 3(b) and followed from Eq.(4) is showed as:

$$C_{chest}(t) = TB_c(P_{c1} + P_{c2}) \tag{3}$$

$$p_{c1} = \begin{bmatrix} 0.1 & 0 \\ 0 & 1.05 \\ 0.49 & 0.88 \\ 0.75 & 1 \end{bmatrix} \begin{bmatrix} L_{c1} & 0 \\ 0 & L_{c2} \end{bmatrix} \tag{4}$$

$$P_{c2} = \begin{bmatrix} 0.75 & 1 \\ 0.82 & 1 \\ 1.14 & 1.11 \\ 0.97 & 0 \end{bmatrix} \begin{bmatrix} L_{c1} & 0 \\ 0 & L_{c2} \end{bmatrix} \tag{5}$$

where L_{c1} and L_{c2} are the width of chest in side image and half width of chest in front image respectively.

D. The compensation system for measurement

The compensation system provides an efficient method to avoid the error calculation which including garment thickness, therefore the subject has been required for no-wearing clothes or only wearing light underwear in measurement. Two process steps, garment thickness estimation and measurement compensation, are adopted for garment thickness compensation. Our approach is based on the method Lin [13] proposed efficient garment thickness compensation strategy. The fuzzy inference system is used to infer the thickness of garment and the back-propagation neural network (BPNN) is used to structure compensation system.

III. VIRTUAL MANNEQUIN DEFORMATION

The Radius Distance system is a polar coordinate system. Basically the radius distance is the scalar Euclidean distance between a point and origin of the system of coordinates. Therefore, the complexity of human body object surface construction and manipulation can often be transferred to sweep-based surface model. The sweep-based surface has many merits in shape deformation and animation of arbitrary 3D objects. Human body object is especially adequate in sweep-based surface, because the human torso and limbs exhibit a cylindrical topology and the underlying skeleton provides natural axes.

For more precision deformation, the human body object will be separated into several parts, left upper arm, left lower arm, right upper arm, right lower arm, chest part, waist part, hip part, right thigh, right shank, left thigh, and left shank, belong to the major joints. Through sweep-based surface and segmentation, the each contour is attached to the proximal joint and defined in the local coordinate system. For connecting each part, we construct wire frame of each part to connect to neighbor parts, and then we can join them together after deformations of each part. It is clear to observe that the wire in wire frame is the skeleton of human body. Fig. 5 (a) shows the each segmented part of human body and Fig. 5(b) shows the weep-based surface with the wire frame and the major joints are showed as solid blue circles.

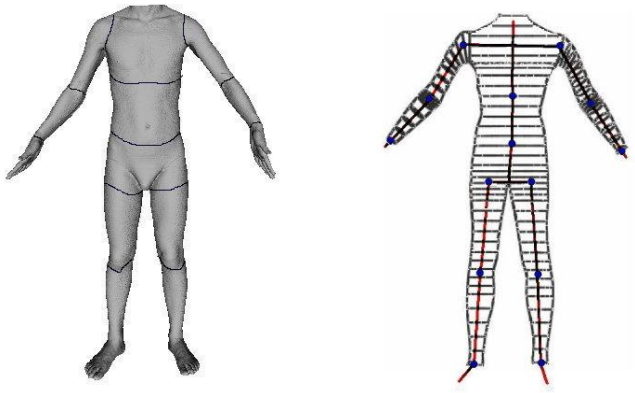


Figure 5. (a) shows the body segmented of human parts and (b) shows the weep-based surface with the wire frame

In this approach, the measurements Body Mass Index (BMI) is used in grouping 3-D body scans, which are obtain using the precise optoelectronic laser scanning measurement device from TAIBBK that described in past paragraph. The virtual mannequins are adjusted by using the measurement data of standard mannequin model. For rough separated of different type of builds, there are 120 human body objects (50 males and 50 females) were selected in TAIBBK database, and the males and females were divided into six clusters on the basis of BMI (BMI<18.5, 18.5 ≤ BMI< 22.5, 22.5 ≤ BMI< 25, 25 ≤ BMI<30, 30 ≤ BMI<35, and BMI ≥ 35) respectively. Each cluster chooses one standard model for represented and deformation to create new mannequin at same cluster.

A. Radius Distance Free Form Deformation

In this approach, we extend the deformation method, Cylinder Free Form Deformation (CFFD), by using a constrain region instead of a constraint control point and used a novel deformation function to simplify the position computing. The radius distance coordinate contained with the scalar distance between a point and origin (or center) and the azimuth angle between the reference direction and the line from the origin is applied, and it is suitable for Cylinder Free Form Deformation. Fig. 6 is illustrated the cylinder coordinate and Radius Distance Free Form Deformation, and shows the *s* layers of deformation control (blue circles).

In each layer, the control point are unified distribute with the same angle between the neighbors on the circle contour. The control point P_{ij} is defined as the *i*-th layer and at the *j*-th control point. In our approach, the number of control point on each layer is fixed for 8 points and the cross section point of contour and y axis is assigned for P_{i0}. The radius distance of each control layer is defined as R and the distance between each layer is fixed as h, so the height H of local deformation part is $(s\text{-}1)h$. The slices of object are show as red circles and the vertexes are denoted as solid red circle. The radius distance r_{pq} of vertex v_{pq} is defined as the distance between the point c_p, which is the cross section point of axis and slice, and vertex v_{pq}.

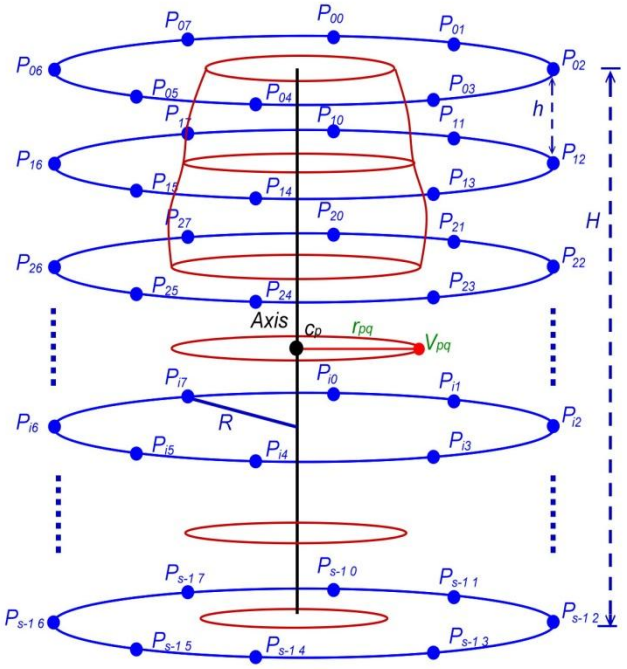


Figure 6. Illustrate the cylinder coordinate and Radius Distance Free Form Deformation.

In sweep-based surface, the object was cut into *n* slices and the vertex v_{oq} belong to these slice is denoted as the *p*-th layer and *q*-th vertex. The point v_{oq} in cylinder coordinate can be written as:

$$\begin{aligned} v_{pq} &= (r_{pq}, \theta_q, h_p) \\ &= (r_{pq}, \frac{2q}{m}\pi, p*h) \quad p \in [0, n-1] \ \ \& \ q \in [0, m-1] \end{aligned} \tag{6}$$

where *m* is the number of chosen vertexes in *p*-th layer.

For simplified calculation of parametric set of deformation, we choose the vertex locate on contour and have the same azimuth angle with control point. So the number of chosen vertex in each sweep layer is also 8 in training model. Fig. 7 shows the *p*-th slice contour and *i*-th control plane and designated as red and blue circles respectively. The inner red circle is the contour of standard model and outer circle is the contour of training data, and then the movement $\delta_{v_{pq}}$ of vertex v_{pq} can be calculated as:

$$\vec{\delta}_{v_{pq}} = v'_{pq} - v_{pq} = (r'_{pq} - r_{pq}) \overrightarrow{u_q} / |\overrightarrow{u_q}| \tag{7}$$

where $\overrightarrow{u_q}$ is the unit vector of *q*-th direction.

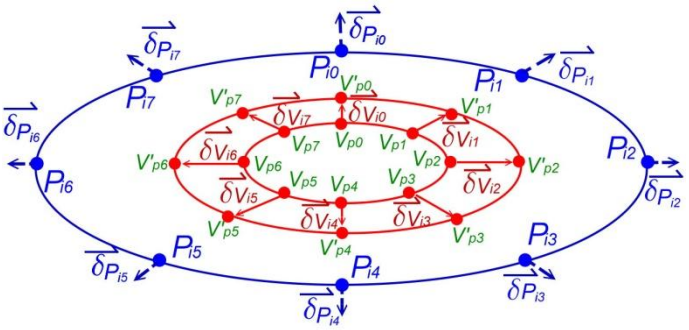


Figure 7. Illustrate the Cylinder Free Form Deformation.

In this paper, we adopt an example-based deformation method that adjusted by using the measurement data of standard mannequin model. So the parametric sets of free form deformation are obtained by training of each cluster that will be done at the beginning. In training model, the function of control point P_{ij} movement is defined as:

$$\begin{aligned} \overrightarrow{\delta_{P_{ij}}} &= \sum_{p=0}^{n-1}\sum_{q=0}^{m-1} f(v_{pq}) = \sum_{p=0}^{n-1}\sum_{q=0}^{m-1} D_{ij}(p,q) E_{ij}(p,q) * \overrightarrow{u_j} \\ &= (C_r - 1) a_{ij} e^{|1-C_r|} * \overrightarrow{u_j} \end{aligned} \tag{8}$$

where $D_{ij}(p,q)$ is a volume of the distance ratio of vertex v_{pq} to control point P_{ij} and can be written as :

$$D_{ij}(p,q) = 1 - \frac{d(v_{pq}, P_{ij})}{D_{\max}} \tag{9}$$

where $d(v_{pq}, P_{ij})$ is the distance between v_{pq} and P_{ij} and can be denoted as follow

$$\begin{aligned} d(v_{pq}, p_{ij}) = [&(R\sin\theta_j - r_{pq}\sin\theta_q)^2 + (R\cos\theta_j - r_{pq}\cos\theta_q)^2 \\ &+ ((i-1)h - (p-1/n-1)H)^2]^{1/2}. \end{aligned} \tag{10}$$

The definition of $D_{\max}$ is the longest distance in this deformation part and can be assigned for the distance between two control points which belong to top and down plane and has the difference azimuth angle π .

$$D_{\max} = [(2R)^2 + H^2]^{1/2} = [(2R)^2 + ((s-1)h)^2]^{1/2}. \tag{11}$$

The formula $E_{ij}(p,q)$ is the effective component of vertex v_{pq} to control point P_{ij}, hence the $E_{ij}(p,q)$ can be explained for the inner product of the vector of vertex v_{pq} movement to the unit vector of control point P_{ij}. The formula $E_{ij}(p,q)$ is showed as:

$$E_{ij}(p,q) = \overrightarrow{\delta_{v_{pq}}} \bullet \overrightarrow{u_{ij}} = (v'_{pq} - v_{pq}) \bullet \overrightarrow{u_j} \tag{12}$$

where $\overrightarrow{\delta_{v_{pq}}}$ is delta vector of vertex v_{oq} from standard model

contour to trainning model contour. $\overrightarrow{u_j}$ is the unit vector of *j*-th control point. In the equation (8), *f* is deformation function where $f : R^3 \to R^3$ and a_{ij} is the free form deformation parameter set of control point P_{ij}. C_r is the circumference ratio of standard model and training model, and defined as:

$$C_r = L \Big/ L_{std} \tag{13}$$

where L is the circumference of training, and L_{std} is the circumference of standard model.

The parametric set A of control point can be acquired from (8) and be denoted as:

$$P = L(C_r) * A * U = (C_r - 1)e^{|1-C_r|} * A * U \tag{14}$$

where U is the unit vector set of each orientation.

After each cluster trained, the parametric sets of each segmented parts are attained. Furthermore, the complete parametric sets of each cluster are obtained through training each cluster. Therefore, the virtual mannequins can be created just only by anthropometric measurement data and Radius Distance Free Form Deformation. The deformation (creation) function of RD-FFD is similar as function (8) and be defined as:

$$\begin{aligned} \overrightarrow{\delta_{v_{pq}}} &= \sum_{i=0}^{s-1}\sum_{j=0}^{m-1} G(p_{ij}) = \sum_{i=0}^{s-1}\sum_{j=0}^{m-1} D_{pq}(i,j)E_{pq}(i,j) * \overrightarrow{u_p} \\ &= v_{pq\_std} - v_{pq\_def} \end{aligned} \tag{15}$$

where v_{pq_std} is the position of vertex v_{pq} on standard model and v_{pq_def} is the position of vertex v_{pq} after deformation.

Deforming a segment without any filtering stage will result in discontinuous passed at the boundaries of the regions. The cosine-tape window filters are performed on the boundary vectors of each segment. Then, the overlap and discontinuous vector will be bounded in a sensible value to smooth the conjunction surface.

IV. Experimental Results and Discussion

This section describes the experiment results including the accuracy of 2-D image-based anthropometric measurement and demonstrates of RD-FFD by anthropometric measurement. The accuracy part adopts statistic methods to calculate the correlations between proposed system and traditional measurement system. The precision analysis is designed to repeated measurements by proposed and manual method, and aimed to show the repeatability of system.

A. Accuracy

The accuracy of the 2D image-based measurement system was calculated by comparing 2D image-based measurement with the manual measurement, or called traditional method, taken by anthropometric [20]. The test sample composed of 155 subjects (87 males and 68 females) that have worn short sleeve T-shirt and been measured both with traditional method and with 2D image-based method. Each subject was measured once by each method. The Pearson correlation coefficients between manual and 2D image-based measurements also list in TABLE I. The error rate (E) is the absolute value of difference of measurement and actual value (manual measurement) against the actual value. It is computed using the follow equation

$$E = \frac{\left| x_{measurement} - x_{actual} \right|}{x_{actual}} \times 100\% \,. \tag{16}$$

TABLE I. Accuracy results of simplex garment (short sleeve T-shirt)

Measurement (mm)	Males	(*n*=87)	Females	(*n*=68)
	Corr.	E_{avg}	Corr.	E_{avg}
Shoulder length	0.92	2.13%	0.97	1.98%
Chest circumference	0.94	1.79%	0.96	1.83%
Waist circumference	0.96	1.75%	0.97	1.72%
Hip circumference	0.96	1.85%	0.95	1.86%
Leg length	0.94	1.86%	0.94	1.89%

The test sample composes with 17 males and 11 females and be required to wear 4 kinds of garment, short T-shirt, long T-shirt, thin jacket, and thick jacket, for measurement. In order to assess the accuracies of 2D image-based measurement for each type of garment, the subjects were measured once with each type of garment for 2D system measurement and once with short sleeve T-shirt for manual measurement. The results of average error rate and Pearson correlation coefficients are listed in TABLE II.

TABLE II. Accuracy results of manifold garment

Measurement(mm) (*n*=46)	Corr.	E_{avg}	Corr.	E_{avg}
	Shoulder length		**Hip circumference**	
2-D system I	0.91	1.92%	0.96	1.88%
2-D system II	0.92	1.93%	0.95	1.93%
2-D system III	0.91	2.11%	0.92	1.77%
2-D system IV	0.90	2.29%	0.93	2.09%
	Chest circumference		**Leg length**	
2-D system I	0.96	1.77%	0.96	1.92%
2-D system II	0.95	1.81%	0.95	1.97%
2-D system III	0.95	2.37%	0.97	2.89%
2-D system IV	0.90	4.43%	0.96	2.13%
	Waist circumference			
2-D system I	0.98	1.66%		
2-D system II	0.96	1.78%		
2-D system III	0.97	3.29%		
2-D system IV	0.88	4.58%		

In first experiment results, the Pearson correlation coefficient show that the closely perfect positive correlations were calculated and exhibited these two measurement data have similar. The results of average error rate were less than 2.5 %, and showed the perfect accuracy of 2D image-based measurement.

The second experiment performed to compare the 2D image-based measurement for each type of garment. Although the challenge of four type of garment was added in this

experiment, the results of Pearson correlation coefficient exhibit the highly positive correlation of 2D image-based measurement and manual measurement. No matter what type of garment was worn, the average error rates of each dimension are less than 5 %. It can be proved that the compensation system is feasible to reduce the influence of garment thickness on measurement.

B. The demonstrate of Virtual mannequins deformation

In human body segmentation, the body is separated into eleven segments and each segment has it individual deformation that belongs to the ratio Cr. But in our 2-D image anthropometric measurement the chest circumference, waist circumference, and hip circumference can be attained, but the significant value right and left upper and lower arm, thigh, and shank circumferences are not available. In order to create a complete virtual mannequin, these anthropometric measurement data are acquired from the manual direct measurement. Fig. 8(b) is a created virtual mannequin with the anthropometric measurement data in TABLE III, and Fig. 8(b) is the standard model at the cluster of meal at BMI range ($25 \leq$ BMI<30).

TABLE III. THE ANTHROPOMETRIC MEASUREMENT DATA OF STANDARD MODEL AND CREATED MODEL.

(cm)	**Height**	**Weight**	**BMI**	**Right upper arm**	**Right lower arm**	**Left upper arm**	**Left lower arm**
Std.	172	74.8	25.28	31.5	29	31.4	29.4
Created	174	78.4	25.89	32	29.5	31.8	29.9
	Chest cir.	**Waist cir.**	**Hip cir.**	**Right thigh**	**Left thigh**	**Right shank**	**Left shank**
Std.	94.3	91.2	99.2	52.6	52.4	39.5	39
Created	97.2	92.3	99.7	53.2	53	41	41.5

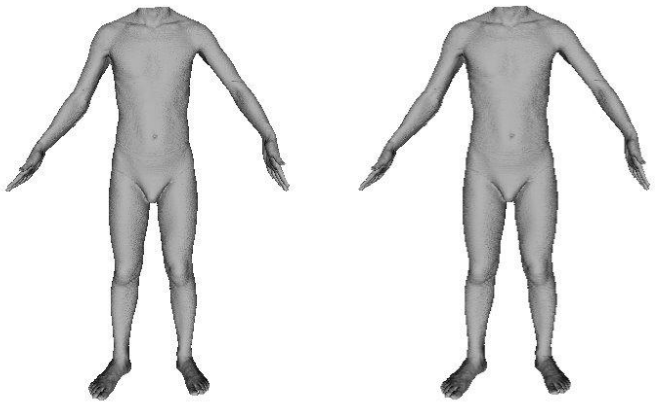


Fig. 8 demonstrates (a) the standard model and (b) The created virtual mannequin.

V. CONCLUSION

In this paper, we introduce a 2-D image-based anthropometric measurement system to cut out the human body silhouette and calculate the direct distances and circumferences. This method provides a more friendly measurement environment, because the conditions of subject for measurement are less than other attempts. The RD-FFD provides a fast and efficient method to create a virtual mannequin, and the example-based deformation reduces the incidence of the distortion. The cylinder coordinate also provide a good geometric space for RD-FFD. The goal of this thesis is to propose a system that have an easy and precise anthropometric measurement, and create a virtual mannequin by using the anthropometric measurement data. In the future work, the other circumference measurement will be done and the automatic segmentation and conjunction are the new challenge to take.

REFERENCES

[1] J. A. Roebuch, Jr., Anthropometric Methods: Designing to Fit the Human Body. Santa Monica, CA: Hum. Factors and Ergonom. Soc., 1995.

[2] C. C. Gordon, T. Churchill, C. E. Clauser, B. Bradtmiller, J. T. McConville, I. Tebbetts, and R. A. Walker, 1988 anthropometric survey of US army personnel: methods and summary statistic. NATICK/TR-89/044. US Army Natick Research, Development, and Engineering Center, Natick, MA, 1989.

[3] P. Meunier, and S. Yin, "Performance of a 2D image-based anthropometric measurement and clothing sizing system," Applied Ergonomics, vol. 31, pp. 445-451, October 2000.

[4] C. Y. Hung, P. Witana, and S. Goonetilleke, "Anthropometric measurements from photographic images," 7th Int. Proc. on Work with computing system, pp. 104-109, June 2004.

[5] E. Paquet and H. L. Viktor, "Adjustment of virtual mannequins through anthropometric measurements, cluster analysis, and content-based retrieval of 3-D body scans," IEEE Trans. on Instrumentation and Measurement, vol. 56, pp.1924 - 1929, October 2007.

[6] K. Robinette, H. Daanen, and E. Paquet, "The CAESAR project: A 3-D surface anthropometric survey," 2nd Int. Conf. on 3-D digital imaging and modeling, Ottawa, Ont. , Canada, pp. 380-386, October 1999.

[7] D. Burnsides, M. Boehmerk, and K. Robinette, "3-D landmark detection and identification in the CAESAR project," 3rd Int. Conf. on 3-D Digital imaging and modeling, Quebec City, Que. , Canada, pp. 393-398, May 2001.

[8] H. Seo and N. Magnenat-Thalman. An example-based approach to human body manipulation. Graph. Models, vol. 66, no. 1, pp. 1-23, 2004.

[9] J. Li and Y. wang. Automatically construct skeletons and parametric structures for polygonal human bodies. Computer Graphics International, 2007.

[10] A. Barr., Superquadrics and angle-preserving transformations. IEEE Computer Graphics and Applications, vol. 1, no. 1, pp. 11-23, 1981.

[11] F. Scheepers, R. E. Parent, W. E. Carlson, and S. F. May. Anatomy-based modeling of the human musculature. In SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, pp. 163–172, New York, NY, USA, 1997.

[12] P. Borrel and A. Rappoport. Simple constrained deformations for geometric modeling and interactive design. ACM Trans. Graph., vol. 13, no. 2, pp. 137–155, 1994.

[13] S. F. Lin, S. C. Chien, and K. Y. Chiu, "The 2D Image-Based Anthropologic Measurement by Using Chinese Medical Acupuncture and human body slice model," International Journal of Computer Science and Information Security," vol. 8, no. 8, pp. 20-29, Nov. 2010.

[14] J. F. David and K. Y. Yan, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," IEEE Trans. on image processing, vol. 10, pp. 1454-1466, October 2001.

[15] A. Albiol, L. Torres, and E. J. Delp, "Optimum color spaces for skin detection," IEEE Int. Conf. Image processing, Thessaloniki , Greece, vol. 1, pp. 122-124, October 2001.

[16] P. J. Shen and K. W. Wu, Massage for pain relief: A step-by-step guide, morning star, Taipei, 2003.

[17] A. Watt and H. Watt, Advanced animation and rendering techniques: theory and practice, Addison Wesley, NY, 1992.

[18] Taiwan Human Body Bank (TAIBBK), http://3d.cgu.edu.tw/DesktopDefault.asp, accessed September. 2009.

[19] C. Y. Yu, Y. H. Lo and W. K. Chiou, "The 3D scanner for measuring body surface area: a simplified calculation in the Chinese adult," Applied Ergonomics, Vol.34, pp.273-278, 2003.

[20] A. Chamberland, R. Carrier, F. Forest, and G. Hachez, Defence and civil institute of environmental medicine, Toronto, Ontario, 1997.

AUTHORS PROFILE

Sheng-Fuu Lin was born in Tainan, R.O.C., in 1954. He received the B.S. and M.S. degrees in mathematics from National Taiwan Normal University in 1976 and 1979, respectively, the M.S. degree in computer science from the University of Maryland, College Park, in 1985, and the Ph.D. degree in electrical engineering from the University of Illinois, Champaign, in 1988. Since 1988, he has been on the faculty of the Department of Electrical and Control Engineering at National Chiao Tung University, Hsinchu, Taiwan, where he is currently a Professor. His research interests include image processing, image recognition, fuzzy theory, automatic target recognition, and scheduling.

Shih-Che Chien was born in Chiayi, R.O.C., in 1978. He received the B.E. degree in electronic engineering from the Nation Chung Cheng University, in 2002. He is currently pursuing the M.E. and Ph.D. degree in the Department of Electrical and Control Engineering, the National Chiao Tung University, Hsinchu, Taiwan. His current research interests include image processing, image recognition, fuzzy theory, 3D image processing, intelligent transportation system, and animation.

Coordinate Rotation Digital Computer Algorithm: Design and Architectures

Naveen Kumar
Electronics & Communication Engineering
University College of Engineering
Punjabi University, Patiala
Punjab, India

Amandeep Singh Sappal
Electronics & Communication Engineering
University College of Engineering
Punjabi University, Patiala
Punjab, India

***Abstract*— COordinate Rotation DIgital Computer (CORDIC) algorithm has potential for efficient and low-cost implementation of a large class of applications which include the generation of trigonometric, logarithmic and transcendental elementary functions, complex number multiplication, matrix inversion, solution of linear systems and general scientific computation. This paper presents a brief overview of the developments in the CORDIC algorithm and its architectures.**

Keywords- CORDIC Algorithms; CORDIC Architectures; FPGA.

I. INTRODUCTION

FIRST described in 1959 [1], CORDIC algorithm is an iterative algorithm, which can be used for the computation of trigonometric functions, multiplication and division. Last half century has witnessed a lot of progress in design and development of architectures of the algorithm for high-performance and low-cost hardware solutions. CORDIC algorithm got its popularity, when [2] showed that, by varying a few simple parameters, it could be used as a single algorithm for unified implementation of a wide range of elementary transcendental functions involving logarithms, exponentials, and square. During the same time, [3] showed that CORDIC technique is a better choice for scientific calculator applications.

The popularity of CORDIC was very much enhanced thereafter primarily due to its potential for efficient and low-cost implementation. With the advent of low cost, low power FPGAs, this algorithm has shown its potential for efficient and low-cost implementation. CORDIC algorithm can be widely used in as wireless communications, Software Defined Radio and medical imaging applications, which are heavily dependent on signal processing. Some other upcoming applications are:

- Direct frequency synthesis, digital modulation and coding for speech/music synthesis and communication;
- Direct and inverse kinematics computation for robot manipulation;
- Planar and three-dimensional vector rotation for graphics and animation.

Although CORDIC may not be the fastest technique to perform these operations, yet it is attractive due to the simplicity and efficient hardware implementation.

The development of CORDIC algorithm and architecture has taken place for achieving high throughput rate and reduction of hardware-complexity as well as the latency of implementation. Latency of implementation is an inherent drawback of the conventional CORDIC algorithm. Angle recoding schemes and higher radix CORDIC have been developed for reduced latency realization. Parallel and pipelined CORDIC have been suggested for high-throughput computation.

This paper presents an overview of the development of CORDIC algorithm. The paper is organized as follows: Section II discusses the basics of CORDIC algorithm, different CORDIC architectures are discussed in Section III. The conclusion along with future research directions are discussed in Section IV.

II. DEFINITION OF CORDIC

The CORDIC is very simple and iterative convergence algorithm that reduces complex multiplication, greatly simplifying overall hardware complexity. This serves as an attractive option to system designers as they continue to face the challenges of balancing aggressive cost and power targets with the increased performance required in next generation signal processing solutions. The basic principle underlying the CORDIC-based computation, and present its iterative algorithm for different operating modes and planar coordinate system.

A. Overview of CORDIC Algorithm

CORDIC algorithm has two types of computing modes Vector rotation and vector translation. The CORDIC algorithm was initially designed to perform a vector rotation, where the vector V with components (X,Y) is rotated through the angle θ yielding a new vector V' with component (X',Y') shown in Fig. 1.

$$V' = [R][V] \quad (1)$$

where **R** is the rotation matrix:

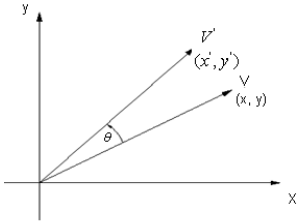


Figure 1: Vector Rotation

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (2)$$

$$R = \begin{bmatrix} \frac{1}{\sqrt{1+\tan^2\theta}} & -\frac{\tan\theta}{\sqrt{1+\tan^2\theta}} \\ \frac{\tan\theta}{\sqrt{1+\tan^2\theta}} & \frac{1}{\sqrt{1+\tan^2\theta}} \end{bmatrix} \quad (3)$$

By factoring out the cosine term in (3), the rotation matrix **R** can be rewritten as

$$R = \left[\left(1+\tan^2\theta\right)^{-1/2}\right]\begin{bmatrix} 1 & -\tan\theta \\ \tan\theta & 1 \end{bmatrix} \quad (4)$$

and can be interpreted as a product of a scale-factor $K = \left[\left(1+\tan^2\theta\right)^{-1/2}\right]$ with a pseudo rotation matrix R_c, given by

$$R_c = \begin{bmatrix} 1 & -\tan\theta \\ \tan\theta & 1 \end{bmatrix} \quad (5)$$

In vector translation, rotates the vector V with component (X, Y) around the circle until the Y component equals zero as illustrated in Fig. 2. The outputs from vector translation are the magnitude X' and phase θ', of the input vector V.

After vector translation, output equations are:

$$X' = K_i\sqrt{\left(X^2+Y^2\right)} \quad (6)$$

$$Y' = 0 \quad (7)$$

$$\theta' = a\tan\left(\frac{Y}{X}\right) \quad (8)$$

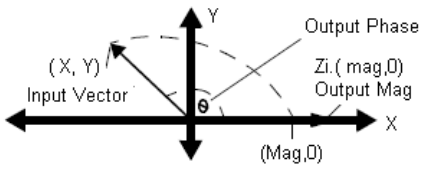


Figure 2: Vector Translation

To achieve simplicity of hardware realization of the rotation, the key ideas used in CORDIC arithmetic are to decompose the rotations into a sequence of elementary rotations through predefined angles that could be implemented with minimum hardware cost and to avoid scaling, that might involve arithmetic operation, such as square-root and division. The second idea is based on the fact the scale-factor contains only the magnitude information but no information about the angle of rotation.

B. Generalized CORDIC Algorithm

After few years, Walther found how CORDIC iterations could be modified to compute hyperbolic functions [2] and reformulated the CORDIC algorithm in to a generalized and unified form which is suitable to perform rotations in circular, hyperbolic and linear coordinate systems. The unified formulation includes a new variable m, which is assigned different values for different coordinate systems. The generalized CORDIC is formulated as follows:

$$\begin{aligned} x_{i+1} &= x_i - m\sigma_i.2^{-i}.y_i \\ y_{i+1} &= y_i + \sigma_i.2^{-i}.x_i \\ w_{i+1} &= w_i - \sigma_i.\alpha_i \end{aligned} \quad (9)$$

Here $\sigma_i = \begin{cases} sign(w_i) & \text{for rotation mode} \\ -sign(w_i) & \text{for vectoring mode} \end{cases}$

III. CORDIC ARCHITECTURES

CORDIC computation is inherently sequential due to two main bottlenecks firstly the micro-rotation for any iteration is performed on the intermediate vector computed by the previous iteration and secondly the (i+1)th iteration could be started only after the completion of the ith iteration, since the value of α_{i+1} which is required to start the (i+1)th iteration could be known only after the completion of the ith iteration. To alleviate the second bottleneck some attempts have been made for evaluation of σ_i values corresponding to small micro-rotation angles [4]. However, the CORDIC iterations could not still be performed in parallel due to the first bottleneck. A partial parallelization has been realized in [4] by combining a pair of conventional CORDIC iterations into a single merged iteration which provides better area-delay efficiency. But the accuracy is slightly affected by such merging and cannot be extended to a higher number of conventional CORDIC iterations since the induced error becomes unacceptable [5]. Parallel realization of CORDIC iterations to handle the first bottleneck by direct unfolding of micro-rotation is possible, but that would result in increase in computational complexity and the advantage of simplicity of CORDIC algorithm gets degraded [6]. Although no popular architectures are known to us for fully parallel implementation of CORDIC, different forms of pipelined implementation of CORDIC have however been proposed for improving the computational throughput [7].To handle latency bottlenecks, various architectures have been developed and reported in this review. Most of the well-known architectures could be grouped under bit parallel iterative CORDIC, bit parallel unrolled CORDIC , bit serial iterative CORDIC and

pipelined CORDIC architecture which we discuss briefly in the following subsections.

A. Bit Parallel Iterative CORDIC Architecture

The vector Rotation CORDIC structure is represented by the schematics in Fig. 3. Each branch consists of an adder-subtractor combination, a shift unit and a register for buffering the output. At the beginning of a calculation initial values are fed into the register by the multiplexer where the MSB of the stored value in the z-branch determines the operation mode for the adder-subtractor. Signals in the x and y branch pass the shift units and are then added to or subtracted from the unshifted signal in the opposite path. The z branch arithmetically combines the registers values with the values taken from a lookup table (LUT) whose address is changed accordingly to the number of iteration. For n iterations the output is mapped back to the registers before initial values are fed in again and the final sine value can be accessed at the output. A simple finite-state machine is needed to control the multiplexers, the shift distance and the addressing of the constant values.

When implemented in an FPGA the initial values for the vector coordinates as well as the constant values in the LUT can be hardwired in a word wide manner. The adder and the subtractor component are carried out separately and a multiplexer controlled by the sign of the angle accumulator distinguishes between addition and subtraction by routing the signals as required. The shift operations as implemented change the shift distance with the number of iterations but those require a high fan in and reduce the maximum speed for the application. In addition the output rate is also limited by the fact that operations are performed iteratively and therefore the maximum output rate equals 1/n times the clock rate.

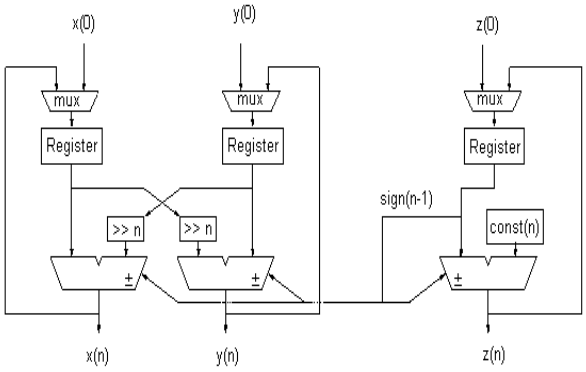


Figure 3: Iterative CORDIC

B. Bit Parallel Unrolled CORDIC Architecture

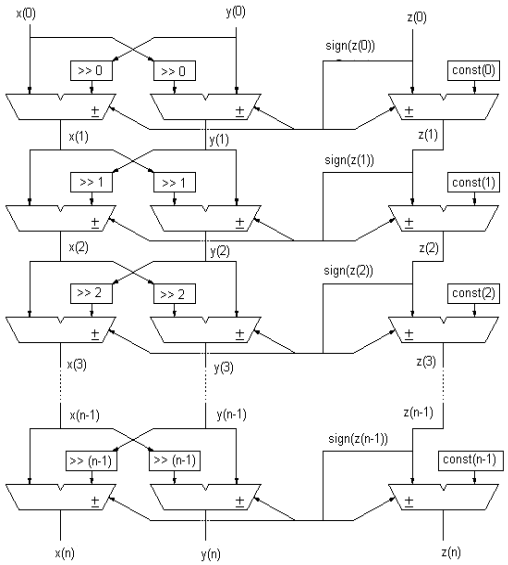


Figure 4: Unrolled CORDIC

Instead of buffering the output of one iteration and using the same resources again, one could simply cascade the iterative CORDIC, which means rebuilding the basic CORDIC structure for each iteration. Consequently, the output of one stage is the input of the next one, as shown in Fig. 4, and in the face of separate stages two simplifications become possible. First, the shift operations for each step can be performed by wiring the connections between stages appropriately. Second, there is no need for changing constant values and those can therefore be hardwired as well. The purely unrolled design only consists of combinatorial components and computes one sine value per clock cycle. Input values find their path through the architecture on their own and do not need to be controlled. As we know, the area in FPGAs can be measured in CLBs, each of which consist of two lookup tables as well as storage cells with additional control components. For the purely combinatorial design the CLB's function generators perform the add and shift operations and no storage cells are used. This means registers could be inserted easily without significantly increasing the area. Pipelining ads some latency, of course, but the application needs to output values at 48 kHz and the latency for 14 iterations equals 312.5 μs which are known to be imperceptible. However, inserting registers between stages would also reduce the maximum path delays and correspondingly a higher maximum speed can be achieved.

C. Bit Serial Iterative CORDIC Architecture

Both, the unrolled and the iterative bit-parallel designs, show disadvantages in terms of complexity and path delays

going along with the large number of cross connections between single stages. To reduce this complexity one could change the design into a completely bit-serial iterative architecture. Bit-serial means only one bit is processed at a time and hence the cross connections become one bit-wide data paths. Clearly, the throughput becomes a function of In spite of this the output rate can be almost as high as achieved with the unrolled design. The reason is the structural simplicity of a bit-serial design and the correspondingly high clock rate achievable. Fig. 5 shows the basic architecture of the bit serial CORDIC processor.

$$\frac{\textit{clock}\ \text{rate}}{\textit{number}\ \text{of iterations} \times \text{word length}}$$

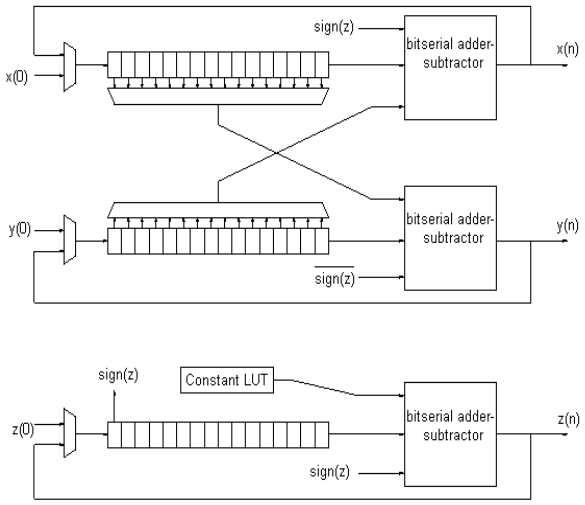


Figure 5: Bit-serial CORDIC

D. D. Pipelined CORDIC Architecture

Since the CORDIC iterations are identical, it is very much convenient to map them into pipelined architectures. The main emphasis in efficient pipelined implementation lies with the minimization of the critical path. The earliest pipelined architecture that we find was suggested in 1984. Pipelined CORDIC circuits have been used thereafter for high-throughput implementation of sinusoidal wave generation, fixed and adaptive filters, discrete orthogonal transforms and other signal processing applications [8].

IV. CONCLUSION

CORDIC algorithm can be implemented by using simple hardware through repeated shift-add operations. This feature makes it attractive for a wide variety of applications. Moreover, its applications in several diverse areas including signal processing, image processing, communication, robotics and graphics apart from general scientific and technical computations have been explored. In the last half century, several algorithms and architectures have been developed to speed up the CORDIC algorithm by reducing its iteration counts and through its pipelined implementation.

ACKNOWLEDGMENT

The authors would thanks the reviewers for their help in improving the document.

REFERENCES

[1] J. E. Volder, “The CORDIC trigonometric computing technique,” IRE Transactions on Electronic Computers, vol. EC- 8, pp. 330–334, Sept. 1959.

[2] J. S. Walther, “A unified algorithm for elementary functions,” in Proceedings of the 38th Spring Joint Computer Conference, Atlantic City, NJ, 1971, pp.379–385.

[3] D. S. Cochran, “Algorithms and accuracy in the HP-35,” Hewlett-Packard Journal, pp. 1–11, June 1972.

[4] S. Wang, V. Piuri, and J. E. E. Swartzlander, “Hybrid CORDIC algorithms,”IEEE Transactions on Computers, volume 46, no. 11, pp. 1202–1207, November1997.

[5]S. Wang and E. E. Swartzlander, “Merged CORDIC algorithm,” in IEEE International Symposium on Circuits Systems (ISCAS’95),1995, volume 3, pp.1988–1991.

[6] B. Gisuthan and T. Srikanthan, “Pipelining flat CORDIC based trigonometric function generators,” Microelectronics Journal, volume 33, Pp.77–89, 2002.

[7] E. Deprettere, P. Dewilde, and R. Udo, “Pipelined CORDIC architectures for fast VLSI filtering and array processing,” in IEEE International Conference on Acoustic, Speech, Signal Processing, ICASSP’84, March 1984, volume 9, pp.250–253.

[8] D. E. Metafas and C. E. Goutis, “A floating point pipeline CORDIC processor with extended operation set,” in IEEE International Symposium on Circuits and Systems, ISCAS’91, June 1991, volume 5, pp. 3066–3069.

AUTHORS PROFILE

Naveen Kumar received the Bachelor of Technology (B.TECH) degree in 2009. Currently he is pursuing Master of Technology (M.Tech) in Electronics & Communication from Punjabi University Patiala, India.

Amandeep Singh Sappal has submitted his Ph.D. in Electronics & Communication at Punjabi University Patiala and presently he is working as an Assistant Professor in Punjabi University Patiala, India. He has published more than 25 papers in reputed journals and conferences. He is reviewer of prestigious journals like Elsevier and Springer etc. Presently he is guiding 5 M.tech students.

Managing Knowledge in Development of Agile Software

Mohammed Abdul Bari
Department of Computer Science, College of Science & Arts
University of Al-Kharj
Wadi Al-Dawasir-11991, Kingdom of Saudi Arabia

Dr. Shahanawaj Ahamad
Department of Computer Science, College of Science & Arts
University of Al-Kharj
Wadi Al-Dawasir-11991, Kingdom of Saudi Arabia

***Abstract*— Software development is a knowledge-intensive work and the main attention is how to manage it. The systematic reviews of empirical studies presents, how knowledge management is used in software engineering and development work. This paper presents how knowledge is used in agile software development and how knowledge is transferred to agile software using agile manifesto. It then argues for the need to scale agile development strategies in knowledge management to address the full delivery. The paper explores the eight agile software scaling factors with knowledge management and their implication for successfully scaling of agile software delivery to meet the real world needs of software development organization.**

Keywords- Knowledge management; Agile software; Scaling factor; Agility; Knowledge capturing

I. INTRODUCTION

Knowledge management is "A method that simplifies the process of sharing, distributing, creating, capturing and understanding the company knowledge [1]. Argyris [2] define "Knowledge is a fluid mix of framed experience, values, contextual information and expert insight that provide a frame work for evaluation and incorporating new experience and new information. According to Nonoka and Takeuchie [3] Knowledge passes through different modes of conversion , which makes the knowledge more refined and spreads it across different layers in an organization.

II. KNOWLEDGE MANAGEMENT IN SOFTWARE DEVELOPMENT

Software development is a knowledge intensive activity. The main assets of software development are not manufacturing plants, building and machines but the knowledge held by the employees and development culture of organization. Software development has long recognized the need for managing knowledge so that the community could learn from the knowledge management. As the field of software engineering matures, there is an increase demand for empirically validated results and not just the testing of technology [4].

Companies developing information system have fail to learn effective means for problem solving to an extent that they have learned to fail [5]. The main differences between methods are they are plan based or traditional, which rely primary on managing explicit knowledge or agile method [6]. There has been much discussion in software development, how to manage knowledge reusing life cycle experience which is gain by processing and producing software development projects which is often referred as experience factory [7] which is stored in experience base, by storing generalizing, tailoring and formalizing experience so that it is easy to reuse. In May 2002 issue of IEEE software [8] was devoted to knowledge management in software engineering, giving several example of knowledge management, applications in software companies. In 2003, the book "Managing Software Engineering Knowledge "[9] was published touching various range of topics from identifying why knowledge management is important in software engineering and development [10].

III. KNOWLEDGE MANAGEMENT IN AGILE SOFTWARE

A. Agile Software development

It consists of set of practice for software development, which has been created by experience practitioners [11]. In Williams and Cockburn [12] stated that agile development is "about feedback and change ". Agile software development techniques have taken the industry by storm, nearly 76% of software organization reported in 2009 that they had one or more agile software underway [13]. According to Agile manifesto 2001[14], it underlines 12 basic principles which are given below:

1. The organization highest priority is to satisfy the customer by continuous delivery of software.
2. Welcoming the changes in requirement even at the later part of development.
3. Delivering the software frequently.
4. Business people and developers must work together throughout the project.
5. Build project around individuals, give them environment and support they needed.
6. Face to face conversation in team with developers.
7. Working software is only means to progress.
8. Agile process prototype sustainable developed.
9. Continuous attention result is excellent product.
10. The act of maximizing the amount of work done.
11. Self-organized the team with requirement, architectures and others.
12. Give regular interval to the project team.

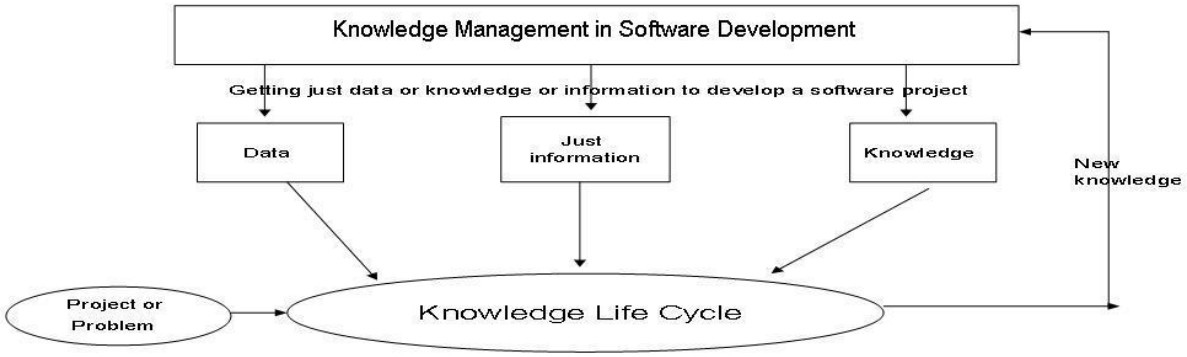


Figure 1: Knowledge Life Cycle

B. Agile Methods

- Agile Modeling (AM): it is a practice based methodology for modeling and documentation of software based system. It is planned to be a collection of values, principal and practice for modeling software that can be applied on software development project in flexible manner. [15]
- Agile Unified Process (AUP): it is simplified version of RUP (Rational Unified Process). It describe a simple, easy to understand approach to developed business application software using agile technology and concept [16]
- Dynamic System Development Method (DSDM): It is based upon Rapid application development methodology [17]. In 2004 DSDM become generic approach to project management and solution delivery. It emphasized continuous user /customer involvement [18].
- Essential Unified Process (EssUP): It was invented by Ivar Jacobson [19] which is an improvement of Rational Unified Process. It uses use case, iterative development, architecture driven development, team practice and process practice which is borrowed from RUP (Rational

  Unified Process) [19]. The main idea here is that you can pick those practices that are applicable to your situation and combine them in to yours own process.
- Extreme Programming (XP): It is a software development methodology which is planned to improve software quality by changing customer requirement. It frequently releases a short development cycle which is intended to improve productivity and introduce check point where the new customer requirement can be adopted.[20]
- Open Unified Process (OpenUP): It is an open source process developed within conceals foundation. It preserves the essential characteristics of RUP/unified process [21] which include incremental development, use case and scenarios deriving development. [22].
- Scrum: It is an iterative increment methodology for project management often seen in agile software development. Although it is mainly used for management of software development. It can also be used to run software maintenance.[23]
- Velocity: It measure the productivity in agile software development .Velocity tracking is an act of measuring said velocity. The velocity is calculated by counting the number of unit of work completed in certain interval, determined at the start of the project. [24]
- Feature Driven Development (FDD): It is an iterative and incremental software development process. FDD blends a number of industry recognized best practices like domain object modeling, developing by feature, individual class, feature teams, inspections, configuration management, regular builds and visibility of progress and result in to cohesive whole. These practices are all driven from client-valued feature perspective. Its main purpose is to deliver tangible, working software repeatedly in timely manner. [25]

IV. KNOWLEDGE MANAGEMENT IN AGILE SOFTWARE

The knowledge is capture from agile software, it is also capture from market research, surrounding area and scientific method, kept in knowledge management box (which consist of people, rules, method (old and new) files etc.). Whenever an organization gets a project the developers will search their knowledge box, they learn from it before starting of project or during the making and also after making the project. Sometime an innovative method is developed for a particular project , once the project is developed the method is send to knowledge management box so that it can used at later part for different project .

V. AGILE SOFTWARE IN SCALE

In early days agile software development techniques were small and relatively straight forward. Today the picture has changed and organization want to apply agility software techniques to a broader set of project. They are dealing with problems which requires large teams, distribute work force many more. They are eight scaling factor that define agility software.

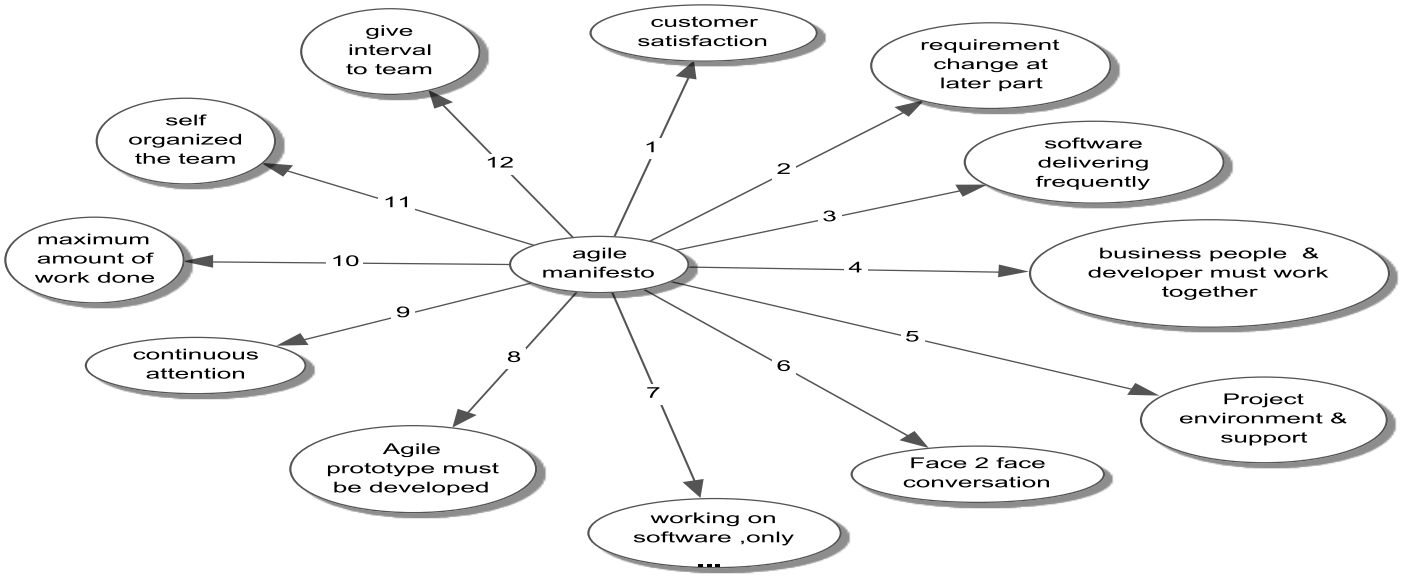


Figure 2: Agile Manifesto

- Team Size: Agility process work very good when the team size is small, as the team size increases communication risks increases and co-ordination become more difficult. In order to learn something from the past, they have to go through the knowledge management capture process which is created by previous projects.
- Geographical Distribution: when the team size is distributed in different countries, effective collaboration become more difficult, more challenging and more error likely to occur and it became more difficult to capture the knowledge from knowledge management capture process.
- Regulatory Compliance: Issues such as ISO 9000[26], these mandates bring requirement of their own, this means, the formality of the work has to increase.
- Domain Complexity: some project team find themselves addressing a straight forward problem, more complex domain require greater emphasis on exploring and experimenting [27].
- Organization Distribution: many project teams includes members from different division, different partner companies or from some external service firms. The more organizationally distributed teams, the more the relationship will be contractual.
- Technical Complexity: Some applications are more complex than others. It is easy to achieve high level quality if you're building a new system from scratch but it is not easy to develop a new application with the existing agility software.
- Organization complexity: your existing organization structure and culture may reflect waterfall [28] values, which increases the complexity of adopting and scaling agile strategies within your organization.
- Enterprise Discipline: Many organizations want to have common infrastructure platform to lower the cost, reduce time, and improve consistency, that is very difficult if project team focus only on their immediate needs.

VI. Future Scope

The scope of this study suggested that agile software development is effective and suitable for many situation and environment. However, at present only few empirically validated studies can be found to support the claims. More ever, the frequent releases of new agile software development methods also bring confusions rather than clarity. The urgent need now (more than new model) is to adopt a few particular methods which can be used by software professional, projects and organizations to choose a particular method to produce a right product at a right time.

VII. Conclusion

Knowledge management may provide important contribution in developing software. Today, there is no doubt that organizations have exploited their potential to create knowledge by focusing on their developing software members by not only focusing on externally developed information, knowledge or data. Studies have shown that traditional plan-driven software development methodologies are not used in practice. Many organizations have been successful at adopting agile software development approaches. Agile scaling model provides a road map for complexities which occur when adopting and tailoring agile method. The knowledge management used in agile software development method provide a novel way of approaching software development problem's , while also maintaining that the method are by no means capable of solving all problems.

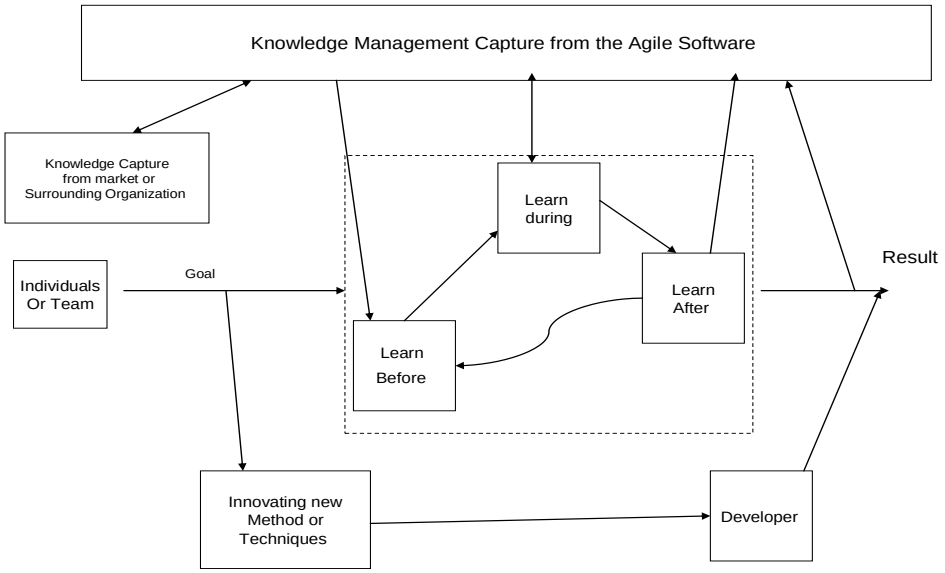


Figure 3: Knowledge Management in Agile Software

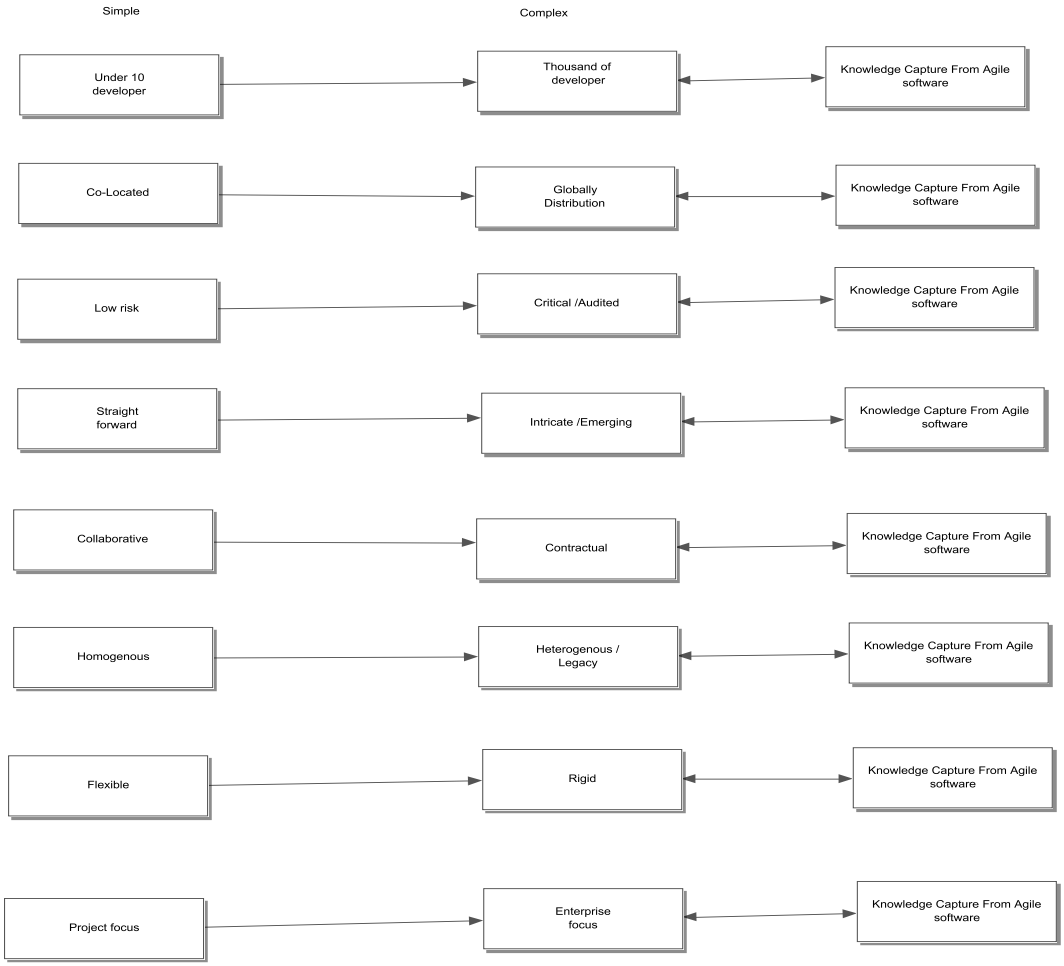


Figure 4: Scaling Factor in Agility

REFERENCES

[1] T.H.Davenprot, L.prusak," *Working Knowledge: how organizations Manage what they Know*", Harvard Business School Press, Boston, USA, 1998.

[2] Argyris C.,"*Knowledge for Action* ". (1993), San Francisco, CA: Jossey-Bass.

[3] I.Nonaka, H.Takeuchi.," *The Knowledge-Creating Company*" , Oxford University Press 1995.

[4] Finn Olav Bjornson , Torgeir Dingsoyr , " *Knowledge management in software engineering : A systematic review of studied concepts , finding and search methods used* ", (2008), International Journal of Information and Software Technology .

[5] K.Lyytinen, D.Robey, "*Learning failure in information systems development* ", (1999), Information system journal.

[6] S.Nerur, V. Baligepally," *Theoretical reflections on agile development methodologies* ", (2007), Communications of the ACM 50 79-83.

[7] V.R.Basili, G.Caldiera, H.D. Rombach," *The experience factory*: ,"(1994) in JJ Marciniak (Ed),Encyclopedia of Sofware Engineering ,J.John WILEY , New York.

[8] M. Lindvall , I.Rus , " *Knowledge management in software engineering* ", (2002), IEEE software .

[9] H.D Doran," *Agile knowledge management in practice"* ,(2004) in :Proceeding of the Sixth International Workshop on Learning Software Organization , Springer Verlag, Banff, Canada.

[10] M. Lindvall, I.Rus," *Knowledge Management for Software Organization*", (2003) in A. Aybuke et al.(Eds.), Managing Software Engineering Knowledge , Springer Verlag, Berlin..

[11] D.Kolb ," *Experiential Learning: Experience as the Source of Learning and Development* ",(1984), Prentice Hall, Englewood Cliffs,USA.

[12]S. Koening, "*Integrated process and knowledge management for product definition , development and delivery* ,(2003) in : Proceeding of the IEEE International Conference on Software-Science , Technology & Engineering .

[13] Dobb's Journal's July 2009 State of the IT Union Survey - www.ambysoft.com/surveys/state-OfITUnion200907.html

[14] Principles Behind the Agile Manifesto - www.agilemanifesto.org/principles.html

[15] Scott w. Ambler, "*Effective practice for modeling and documentation* "(.2007)
http://www.agilemodeling.com/.

[16] Scott w.Ambler." *The Agile Unified Process* ", (2009) http://www.ambysoft.com/unifiedprocess/agileUP.html

[17] Casemaker Totem,"*what is Rapid Application Development*?",(2000)
http://www.casemaker.com/download/products/totem/rad_wp.pdf

[18] Benjamin J.J.Voigt, Dr.M.Glinz,"*Dynamic System Development Method*", (2004), Department of Information Technology, University of Zurich, Retrieve on.

[19] Ivar Jacobson ," *Essential Unified Process* " (2010), http://en.wikipedia.org/wiki/Essential_Unified_Process

[20] Kent Beck ,Cynthia Andres ," *Extreme Programming Explained*",(2004), Addison –Wesley Professional.

[21] Wikipedia ,"*IBM Rational Unified Process* ",(2003). http://en.wikipedia.org/wiki/RUP.

[22] Wikipedia ," OpenUP ", (2009) http://en.wikipedia.org/wiki/Open_Unified_Process

[23] Mike Cohn," *Succeeding with Agile: Software Development Using Scrum*" ,(2009),The Addison –Wesley Series.

[24] Jeremy Weiskotten," *Velocity : Measuring and Planning an Agile Project* ,(2009,
http://agilesoftwaredevelopment.com/blog/jeremy/velocity-measuring and-planning-agil

[25] Wikipedia,"*Feature Driven Development* ", (2009), http://en.wikipedia.org/wiki/Feature_Driven_Development

[26] Hongyi Sun, "*Total quality management, ISO 9000 certification and performance improvement*",(2000) International Journal of Quality & Reliability Management, Vol. 17 Iss: 2, pp.168 – 179.

[27] Kruchten, P. (2009). "*The Context of Software Development* "- http://pkruchten.wordpress.com/2009/07/22/the-context-of-software-development/

[28] Ambler, S.W.," *Agile Modeling: Effective Practices for Extreme Programming and the Unified Process*", (2002) New York: Wiley Press.

[29] Dobb's Journal's July 2009 State of the IT Union Survey - www.ambysoft.com/surveys/state-OfITUnion200907.html

[30] Dobb's Journal's 2008,"*Project Success Survey*" - www.ambysoft.com/surveys/success2008.html

[31] Pekka Abrahamsson ,Outi Salo , Jussi Ronkainen & Juhani Warsta ," *Agile software development method*", (2002), VVT Publication

AUTHORS PROFILE

Mr. Mohammed Abdul Bari is an Information System Architect and expert in handling software process improvement. His research area includes Business Process Reengineering, Process Modeling, Information System Redesign and Reengineering. He did B.E. in Computer Science & Engineering from Bangalore University, INDIA and M.S. in Information Systems from London South Bank University, United Kingdom, currently pursuing Ph.D. in Computer Science from University of Newcastle, District Columbia, U.S.A.

Dr. Shahanawaj Ahamad is an active academician and researcher in the field of Software Reverse Engineering with experience of ten years, working with Al-Kharj University's College of Science & Arts in Wadi Al-Dawasir, K.S.A. He is the member of various national and international academic and research groups, member of journal editorial board and reviewer. He is currently working on Legacy Systems Migration, Evolution and Reverse Engineering, published more than twenty papers in his credit in national and international journals and conference proceedings. He holds M. Tech. followed by Ph.D. in Computer Science major Software Engineering, supervised many bachelor projects and master thesis, currently supervisor of Ph.D. theses.

Efficient Retrieval of Text for Biomedical Domain using Data Mining Algorithm

Sumit Vashishta
Computer Science department
Samrat Ashok Technological Institute
Vidisha, M.P. INDIA

Dr. Yogendra Kumar Jain
Computer Science department
Samrat Ashok Technological Institute
Vidisha, M.P. INDIA

***Abstract*—Data mining, a branch of computer science [1], is the process of extracting patterns from large data sets by combining methods from statistics and artificial intelligence with database management. Data mining is seen as an increasingly important tool by modern business to transform data into business intelligence giving an informational advantage. Biomedical text retrieval refers to text retrieval techniques applied to biomedical resources and literature available of the biomedical and molecular biology domain. The volume of published biomedical research, and therefore the underlying biomedical knowledge base, is expanding at an increasing rate. Biomedical text retrieval is a way to aid researchers in coping with information overload. By discovering predictive relationships between different pieces of extracted data, data-mining algorithms can be used to improve the accuracy of information extraction. However, textual variation due to typos, abbreviations, and other sources can prevent the productive discovery and utilization of hard-matching rules. Recent methods of soft clustering can exploit predictive relationships in textual data. This paper presents a technique for using soft clustering data mining algorithm to increase the accuracy of biomedical text extraction. Experimental results demonstrate that this approach improves text extraction more effectively that hard keyword matching rules.**

Keywords-Data mining; Biomedical text extraction; Biomedical text mining.

I. INTRODUCTION

This paper aims to use data mining techniques to extract text from biomedical literature with reasonably high recall and precision. In recent years, along with development of bioinformatics and information technology, biomedical technology grows rapidly. With the growth of the biomedical technology, enormous biomedical databases are produced. It creates a need and challenge for data mining. Data mining is a process of the knowledge discovery in databases and the goal is to find out the hidden and interesting information [3]. The technology includes association rules, classification, clustering, and evolution analysis etc. Clustering algorithms are used as the essential tools to group analogous patterns and separate outliers according to its principles that elements in the same cluster are more homogenous while elements in the different ones are more dissimilar [2]. Furthermore, data mining algorithms do not need to rely on the pre-defined classes and the training examples while classifying the classes and can produce the good quality of clustering, so they fit to extract the biomedical text better. A major challenge for information retrieval in the life science domain is coping with its complex and inconsistent terminology. In this paper we try to devise an algorithm which makes word-based retrieval more robust. We will investigate how data mining algorithms based on keywords affects retrieval effectiveness in the biomedical domain. We will try to answer the following research question in this paper "How can the effectiveness of word-based biomedical information retrieval be improved using data mining algorithm?"

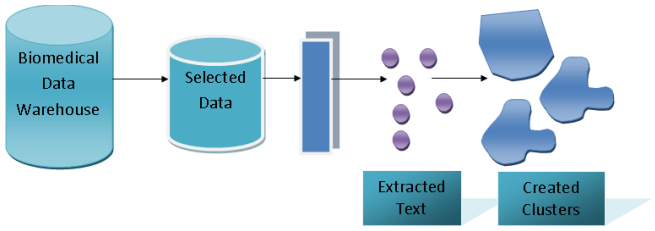


Figure 1: Text extraction from Biomedical literature base

II. BACKGROUND

Biomedical text extraction refers to text mining applied to texts and literature of the biomedical and molecular biology domain. It is a rather recent research field on the edge of natural language processing, bioinformatics, medical informatics and computational linguistics.

There is an increasing interest in text mining and information extraction strategies applied to the biomedical and molecular biology literature due to the increasing number of electronically available publications stored in databases.

The main developments in this area have been related to the identification of biological entities (named entity recognition), such as protein and gene names in free text, the association of gene clusters obtained by microarray experiments with the biological context provided by the corresponding literature, automatic extraction of protein interactions and associations of proteins to functional concepts (e.g. gene ontology terms). Even the extraction of kinetic parameters from text or the subcellular locations of proteins have been addressed by information extraction and text mining technology.

The optimal retrieval of a literature search in biomedicine depends on the appropriate use of Medical Subject Headings,

descriptors and keywords among authors and indexers. We hypothesized that authors, investigators and indexers in four biomedical databases are not consistent in their use of terminology in Complementary and Alternative Medicine.

The increasing research in Complementary and Alternative Medicine and the importance placed on practicing evidence-based medicine require ready access to the biomedical scientific literature. The optimal retrieval of a literature search in biomedicine depends on the appropriate use of Medical Subject Headings, descriptors and keywords among authors, indexers, and investigators [4]. It has been recognized that available online databases for biomedical domain differed in their thesaurus construction and indexing procedures, making effective and efficient searching difficult [5].

In this paper we try to employ an algorithm that extracts the biomedical texts fro the biomedical database based on the some data mining algorithm. Our approach first identifies the keywords contained in the biomedical database and then clustering these keywords to group all the text that fall into the category of the given keyword i.e. if that keyword is being used for searching the returned cluster for that particular keyword will contain all the text corresponding to that keyword.

III. METHOD

Text mining is defined as the automatic discovery of new, previously unknown, information from unstructured textual data. This process is done in three steps: information retrieval, information extraction and data mining. A primary reason for using data mining for biomedical text is to assist in the analysis of collections of the available biomedical text. Biomedical data is vulnerable to co linearity because of unknown interrelations. The analysis in this paper will be augmented by using experiment-based approach.

Before data mining algorithms can be used, a target data set will be assembled. As data mining can only uncover patterns already present in the data, the target dataset must be large enough to contain these patterns. Pre-process is essential to analyze the multivariate datasets before clustering or data mining. The target set is then cleaned. Cleaning removes the observations with noise and missing data.

The biomedical data available with us is first put into a data warehouse. Before putting the data in the data warehouse the keyword extraction algorithm is used to find out the keywords from the full text. This keyword extraction uses partial parser to extract entity names (gene, protein names etc). This parser uses linguistic rules and statistical disambiguity to achieve greater precision.

The data is then organized into clusters. Clustering is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data. The clusters will be created based on the keywords extracted from our biomedical text. These clusters will be created using fuzzy C mean algorithm. The fuzzy c-means algorithm is one of the most widely used soft clustering algorithms. It is a variant of standard k-means algorithm that uses a soft membership function. Fuzzy C-Means (FCM) clustering algorithm is one of the most popular fuzzy clustering algorithms. FCM is based on minimization of the objective function Fm(u, c):

$$F_m(u,c) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m d^2(x_k, c_i)$$

FCM computes the membership *uij* and the cluster centers *cj* by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m x_i}{\sum_{i=1}^{N} u_{ij}^m}$$

where *m*, *the fuzzification factor* which is a weighting exponent on each fuzzy membership, is any real number greater than 1, *uij* is the degree of membership of *xi* in the cluster *j*, *xi* is the *i*th of d-dimensional measured data, *cj* is the dimension center of the cluster, *d2(xk,ci)*is a distance measure between object xk and cluster center ci, and ||*|| is any norm expressing the similarity between any measured data and the center.

The FCM algorithm involves the following steps:

1. Set values for c and *m*
2. Initial membership matrix U= [*uij*], which is U(0) (|i| = number of members, |j| = number of clusters)
3. At *k-step:* calculate the centroids for each cluster through equation (2) if k ≠ 0. (If k=0, initial centroids location by random)
4. For each member, calculate membership degree by equation (1) and store the information in U(k)
5. If the difference between U(k) and U(k+1) less than a certain threshold, then STOP; otherwise, return to step 3.

IV. PROPOSED MODEL

Clustering is the process of organizing objects into groups whose members are similar in some way. It can be considered the most important unsupervised learning problem which deals with finding a structure in a collection of unlabeled data. A cluster is therefore a collection of objects which are “similar” between them and are “dissimilar” to the objects belonging to other clusters. Hard clustering is the techniques in which any pattern can be in only one cluster at any time. Soft clustering is the technique which permits patterns to be in more than one cluster at any time. There are various clustering approaches that can be applied to cluster the biomedical keywords extracted from full text articles, some of them are k-means, k-median, Hierarchical Clustering Algorithm, Nearest Neighbor Algorithm etc. Here we are using modified fuzzy C mean clustering algorithm.

Here the proposed algorithm is responsible for extracting keywords present in the full text biomedical article store these keywords in a relation. Then the actual work of algorithm begins, it starts clustering of keywords. The algorithm initially picks some keywords that are extracted. It groups the full text articles based on these keywords. It means each cluster contains only those articles which contain that keyword as their part. Then it starts using fuzzy C mean clustering to combine the clusters together on some similarity measure. Here we combine two clusters if their similarity measure is greater than or equal to a specified threshold value. The proposed Algorithm repeats this process until no more changes are made to the clusters. Finally the proposed algorithm stores all the clusters in an xml file. Here our motive to extract all the full text articles which may be relevant for the user providing the search string, for this out of all clusters the cluster with largest number of articles is our target.

V. Proposed Algorithm

The proposed algorithm will take a complete list of all the biomedical articles and the output will be the XML files containing the clusters created using fuzzy c mean algorithm on keywords.

Input: List of full text biomedical articles.

Output: XML files containing the created clusters.

Algorithm

1. Read the next article in the list of biomedical text
2. Read the full text article
3. Extract the keywords from the article using KEA algorithm
4. Refer to the biomedical lexicon and discard the irrelevant keywords
5. Put the data in following relation so that the full text can be retrieved later using keywords only

Article UID	Article Name	Keywords	Full text	Source

6. Go to step 1 and repeat till all the articles in the list of biomedical articles are processed.
7. Use the fuzzy c-means algorithm to create clusters on keywords.
8. Save the article clusters in form of an XML file(containing articles IDs).

Note: The relation created step 6 will be used at the time of retrieval. Whenever the biomedical database is searched for any word the cluster containing the matching keywords is returned. The respective full text and other details corresponding to the returned cluster can be retrieved using this relation.

VI. Result

The experiments were performed on the test application developed in .Net 2.0. The database contains all the article entries populated manually from the web resources like "http://www.medilexicon.com" and few more, starting with letter 'A'.

The search was performed using the traditional keyword based search algorithm and compared with the proposed algorithm. The snapshot for asset of search results is shown in Figure 2.

Given the same data for text extraction, the proposed algorithm seems to be retrieving approximately 69% more relevant search results than the keyword based searching. Figure 3 illustrates the improvement achieved using the proposed algorithm.

	List of matching articles found	
Search Keyword	Keyword based search	Proposed algorithm
abarognosis	42	71
abasia	23	39
abasia-astasia	34	57
abasic	32	54
abatement	42	71
abatic	5	8
abaxial	53	90
Abbé	43	73
Abbé condenser	44	74

Figure 2: Comparison of results using traditional and proposed algorithm

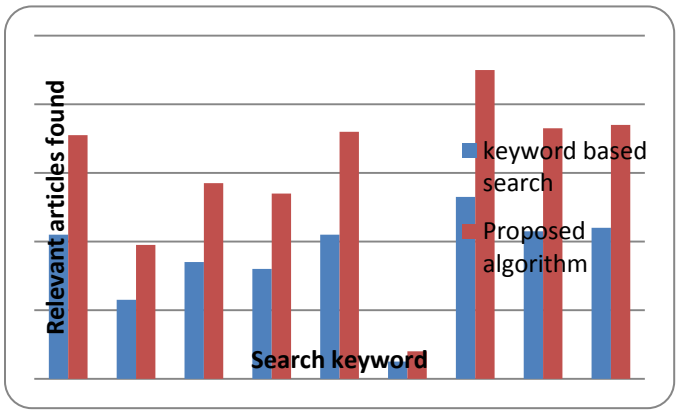


Figure 3: Improved text extraction using proposed algorithm

VII. Conclusion

Extraction of text from biomedical literature is an essential operation. Given that there have been many text extraction methods developed; this paper presents a novel technique that employs keyword based article clustering to further enhance the text extraction process. The development of the proposed algorithm is of practical significance; however it is challenging to design a unified approach of text extraction that retrieves the relevant text articles more efficiently. The proposed algorithm, using data mining algorithm, seems to extract the text with contextual completeness in overall, individual and collective forms, making it able to significantly enhance the text extraction process from biomedical literature.

ACKNOWLEDGMENT

This research is supported by the Computer Science and Engineering department, SATI, Vidhisha.

REFERENCES

[1] Clifton, Christopher (2010). "Encyclopedia Britannica: Definition of Data Mining". Retrieved 2010-12-09.

[2] Han, J., & Kamber, M., Data Mining Concepts and Techniques. CA : Morgan Kaufmann, 2001.

[3] Badgett RG: How to search for and evaluate medical evidence. Seminars in Medical Practice 1999, 2:8-14, 28.

[4] Richardson J: Building CAM databases: the challenges ahead. J Altern Complement Med 2002, 8:7-8.

[5] Kantardzic, Mehmed (2003). Data Mining: Concepts, Models, Methods, and Algorithms. John Wiley & Sons. ISBN 0471228524. OCLC 50055336

[6] Miller, H. and Han, J., (eds.), 2001, Geographic Data Mining and Knowledge Discovery, (London: Taylor & Francis).

[7] Manu Aery, Naveen Ramamurthy, and Y. Alp Aslandogan. Topic identification of textual data. Technical report, The University of Texas at Arlington, 2003.

[8] Pavel Berkhin. Survey of clustering data mining techniques. Technical report, Accrue Software, San Jose, CA, 2002.

[9] Cecil Chua, Roger H.L. Chiang, and Ee-Peng Lim. An integrated data mining system to automate discovery of measures of association. In Proceedings of the 33rd Hawaii International Conference on System Sciences, 2000.

[10] George Forman. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res., 3:1289-1305, 2003.

[11] Rayid Ghani. Combining labeled and unlabeled data for text classification with a large number of categories. In IEEE Conference on Data Mining, 2001.

[12] George Karypis and Eui-Hong Han. Concept indexing: A fast dimensionality reduction algorithm with applications to document retrieval and categorization. Technical report TR-00-0016, University of Minnesota, 2000.

[13] Jerome Moore, Eui-Hong Han, Daniel Boley, Maria Gini, Robert Gross, Kyle Hastings, George Karypis, Vipin Kumar, and Bamshad Mobasher. Web page categorization and feature selection using association rule and principal component clustering. In 7th Workshop on Information Technologies and Systems, 1997.

[14] Sam Scott and Sam Matwin. Text classification using wordnet hypernyms. In Proceedings of the COLING/ACL Workshop on Usage of WordNet in Natural Language Processing Systems, Montreal, 1998.

[15] Michael Steinbach, George Karypis, and Vipin Kumar. A comparison of document clustering techniques. In KDD Workshop on Text Mining, 2000.

[16] Andreas Weingessel, Martin Natter, and Kurt Hornik. Using independent component analysis for feature extraction and multivariate data projection, 1998.

[17] Robert Nisbet (2006) Data Mining Tools: Which One is Best for CRM? Part 1, Information Management Special Reports, January 2006.

[18] Dominique Haughton, Joel Deichmann, Abdolreza Eshghi, Selin Sayek, Nicholas Teebagy, & Heikki Topi (2003) A Review of Software Packages for Data Mining, The American Statistician, Vol. 57, No. 4, pp. 290–309.

[19] R. Agrawal et al., Fast discovery of association rules, in Advances in knowledge discovery and data mining pp. 307–328, MIT Press, 1996.

[20] Kumar, V. (2011). An Empirical Study of the Applications of Data Mining Techniques in Higher Education. *International Journal of Advanced Computer Science and Applications - IJACSA, 2*(3), 80-84.

[21] Jadhav, R. J. (2011). Churn Prediction in Telecommunication Using Data Mining Technology. *International Journal of Advanced Computer Science and Applications - IJACSA, 2*(2), 17-19.

[22] Devi, S. N. (2011). A study on Feature Selection Techniques in Bio-Informatics. *International Journal of Advanced Computer Science and Applications - IJACSA, 2*(1), 138-144.

Authors Profile

Sumit Vashishta is a research scholar pursuing M.Tech in Computer Science & Engineering from Samrat Ashok Technological Institute Vidisha M.P India. He secured degree of B.E. in IT from Rajiv Gandhi Technical University, Bhopal (M.P.) India in 2006.
E-mail-sumitvbpl@gmail.com

Dr. Yogendra Kumar Jain presently working as head of the department, Computer Science & Engineering at Samrat Ashok Technological Institute Vidisha M.P India. The degree of B.E. (Hons) secured in E&I from SATI Vidisha in 1991, M.E. (Hons) in Digital Tech. & Instrumentation from SGSITS, DAVV Indore(M.P), India in 1999. The Ph. D. degree has been awarded from Rajiv Gandhi Technical University, Bhopal (M.P.) India in 2010. Research Interest includes Image Processing, Image compression, Network Security, Watermarking, Data Mining. Published more than 40 Research papers in various Journals/Conferences, which include 15 research papers in International Journals.
Tel:+91-7592-250408,
E-mail: ykjain_p@yahoo.co.in

Design and Performance Analysis of Microstrip Array Antennas with Optimum Parameters for X-band Applications

Md. Tanvir Ishtaique-ul Huque[1], Md. Kamal Hosain[2], Md. Shihabul Islam[3], and Md. Al-Amin Chowdhury[4]
Dept. of Electronics and Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh.

***Abstract*—This paper demonstrates simple, low cost and high gain microstrip array antennas with suitable feeding techniques and dielectric substrate for applications in GHz frequency range. The optimum design parameters of the antenna are selected to achieve the compact dimensions as well as the best possible characteristics such as high radiation efficiency, high gain, etc. In this paper different microstrip array antennas such as series feed, corporate feed and corporate- series feed are designed, simulated, analyzed and compared regarding to the antenna performances. The designed antennas are 4x1, 4x1, and 4x2 arrays. The optimum feeding system is decided based on the various antenna parameters that are simulated. The simulation has been performed by using SONNET version V12.56 simulator which is a commercially available antenna simulator. The designed antennas provide return losses in the range of -4.21dB to -25.456dB at frequencies around 10GHz by using Taconic TLY-5 dielectric substrate with permittivity, εr= 2.2 and height, h =1.588 mm. The gain of these simulated antennas is found about 15dB and side lobe label is maintained lower than main lobe. Since, the resonance frequency of these antennas is around 10GHz, these antennas are suitable for X-band applications such as satellite communication, radar, medical applications, and other wireless systems.**

Keywords-microstrip antenna; array antenna; corporate-series feed array; corporate feed array; series feed array

I. INTRODUCTION

Modern wireless communication system requires low profile, light weight, high gain, and simple structure antennas to assure reliability, mobility, and high efficiency characteristics. Microstrip antenna satisfies such requirements. The key features of a microstrip antenna are relative ease of construction, light weight, low cost and either conformability to the mounting surface or, an extremely thin protrusion from the surface. This antenna provides all of the advantages of printed circuit technology. These advantages of microstrip antennas make them popular in many wireless communication applications such as satellite communication, radar, medical applications, etc[1]. The limitations of microstrip antennas are narrow frequency band and disability to operate at high power levels of waveguide, coaxial line or even stripline. Therefore, the challenge in microstrip antenna design is to increase the bandwidth and gain[2].

Different array configurations of microstrip antenna can give high gain, wide bandwidth and improved efficiency. The distribution of voltages among the elements of an array depends on feeding network. Suitable feeding network accumulates all of the induced voltages to feed into one point [3]. The proper impedance matching throughout the corporate and series feeding array configurations provides high efficiency microstrip antenna[4]. Power distribution among antenna elements can be modified by corporate feed network. The corporate feed network can steer beam by introducing phase change[5].

The choosing of design parameters (dielectric material, height and frequency, etc) is important because antenna performance depends on these parameters. Radiation performance can be improved by using proper design structures [6]. The use of high permittivity substrates can miniaturize microstrip antenna size[7]. Thick substrates with lower range of dielectric offer better efficiency, and wide bandwidth but it requires larger element size[8]. Microstrip antenna with superconducting patch on uniaxial substrate gives high radiation efficiency and gain in millimeter wave lengths [9]. The width discontinuities in a microstrip patch reduces the length of resonating microstrip antenna and radiation efficiency as well [10].

Different radar systems such as synthetic aperture radar (SAR), shuttle imaging radar, remote sensing radars, and other wireless communication systems operate in L, C and X bands. Microstrip antenna is the first option for this high frequency band such as X-band due to its low cost, light weight, and robustness [11]. This article provides a way to choose the design parameters of antennas to achieve the desired dimensions as well as the characteristics for the effective radiation efficiency. This paper also compare the characteristics of series feed, corporate feed and corporate-series feed microstrip array antennas to get optimum feeding system. These designed antennas are potential candidate for the X-band wireless applications due to the simplicity in structure, ease of fabrication and high gain and high efficiency.

II. MICROSTRIP ANTENNA DESIGN

Microstrip patch antennas consist of very thin metallic strip (patch) placed on ground plane where the thickness of the metallic strip is restricted by $t << \lambda 0$ and the height is restricted

by $0.0003\lambda_0 \leq h \leq .05\lambda_0$ [12-14]. The microstrip patch is designed so that its radiation pattern maximum is normal to the patch. For a rectangular patch, the length L of the element is usually $\lambda_0/3 < L < \lambda_0/2$. There are numerous substrates that can be used for the design of microstrip antennas and their dielectric constants are usually in the range of $2.2 \leq \varepsilon_r \leq 12$. To implement the microstrip antennas usually Fr-4 ($\varepsilon_r = 4.9$), Rogers TMM 4($\varepsilon_r = 4.5$), Taconic TLY-5 ($\varepsilon_r = 2.2$), Alumina (96%) ($\varepsilon_r = 9.4$), Teflon(PTFE) ($\varepsilon_r = 2.08$), Arlon AD 5 ($\varepsilon_r = 5.1$) dielectric materials are used as the substrate[1, 12, 13].

The Performance of the microstrip antenna depends on its dimension. Depending on the dimension the operating frequency, radiation efficiency, directivity, return loss and other related parameters are also influenced. For an efficient radiation, the practical width of the patch can be written as [12, 13, 15]

$$w = \frac{1}{2 f_r \sqrt{\mu_0 \varepsilon_0}} \times \sqrt{\frac{2}{\varepsilon_r + 1}} \quad (1)$$

and the length of the antenna becomes

$$L = \frac{1}{2 f_r \sqrt{\varepsilon_{eff}} \sqrt{\varepsilon_0 \mu_0}} - 2\Delta L \quad (2)$$

where

$$\Delta L = 0.41 h \frac{\varepsilon_{eff} + 0.3}{\varepsilon_{eff} - 0.258} * \frac{\left(\frac{w}{h} + 0.264\right)}{\left(\frac{w}{h} + 0.8\right)} \quad (3)$$

and

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2\sqrt{1 + 12\frac{h}{w}}} \quad (4)$$

where λ is the wave length, fr is the resonant frequency, L and W are the length and width of the patch element respectively and ε_r is the dielectric constant. In the following Fig. 1 shows an antenna that has been designed to cover operating frequency of 10 GHz and the quarter wavelength transformer method is used to match the impedance of the patch element with the transmission line [12, 13].

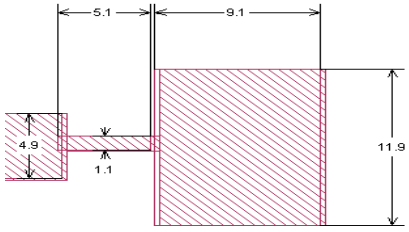


Figure 1. Single element microstrip patch antenna.

III. Parameters Analysis of Microstrip Antenna

Performance of the microstrip antenna depends on the used dielectric patch material, operating frequency and height of the substrate. As the antenna dimension is bounded by all of these parameters, hence the radiation efficiency as well as the directivity is also influenced. Thus in order to get better performance of a microstrip antenna we need to maintain the value of all of these parameters within a desired threshold level.

Fig. 2 illustrates that the radiation efficiency is near about to independent to relative dielectric constant for 500MHz operating frequency. Radiation efficiency declines rapidly with dielectric constant for 10GHz frequency. In summary, with the increasing value of relative dielectric constant, the radiation efficiency of the microstrip antenna is decreased. Thus it needs to choose the dielectric material having a lower dielectric constant near about the air ($\varepsilon_r = 1$) to get higher efficiency.

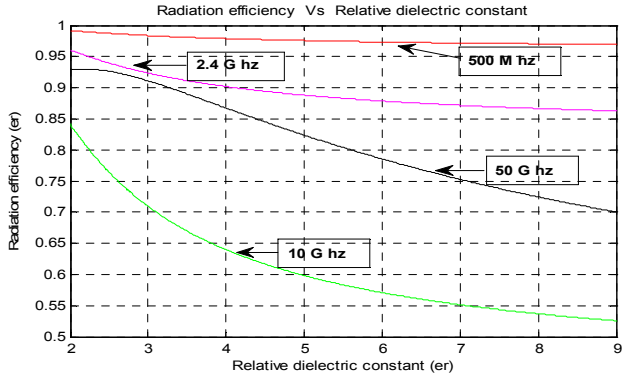


Figure 2. Effect of the dielectric material on the radiation efficiency at different operating frequencies.

Fig. 3 shows that radiation efficiency is the lowest at around 30GHz operating frequencies. When the operating frequency is below 10GHz it provides higher radiation efficiency and more higher radiation efficiency can also be achieved by increasing the operating frequency above 50GHz. Teflon (PTFE) is the material which has more efficiency than others.

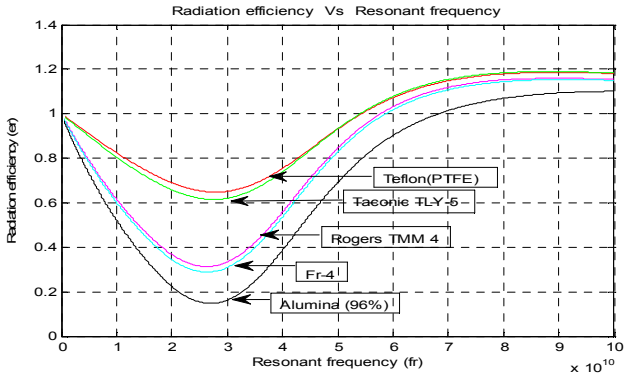


Figure 3. Effect of the operating frequency on the radiation efficiency for different dielectric materials.

Fig. 4 describes that radiation efficiency goes down with the increase of the height of substrate. Hence, the lower the value of the height of the substrate as well as its relative dielectric constant, the higher radiation efficiency can be achieved. Fig. 5 demonstrates that at the frequency less than 10GHz, with the growing of the height of the substrate the radiation efficiency is gradually reduces but at the frequency

greater than 30GHz it shows the reverse trend where with the increasing of the height of the substrate the radiation efficiency is also increased progressively. For the 500MHz operating frequency, radiation efficiency is almost independent on the height of the dielectric material.

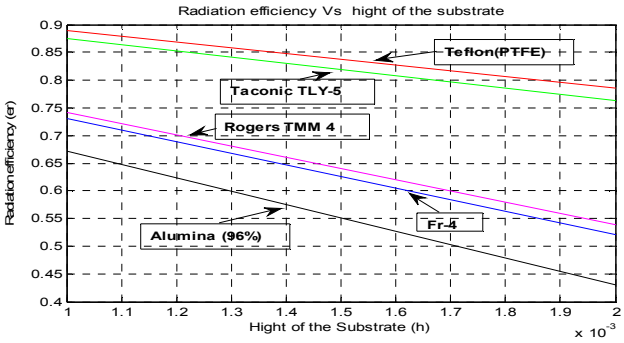


Figure 4. Effect of the radiation efficiency on the height of the dielectric material considering the relative dielectric constant.

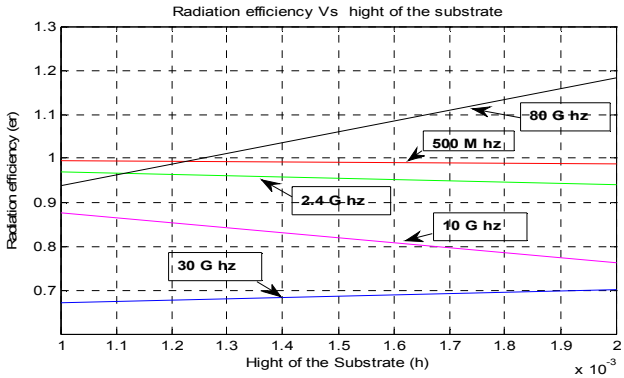


Figure 5. Effect of the radiation efficiency on the height of the dielectric material at different operating frequencies.

Thus after observing the above figures it can be concluded that to get a microstrip antenna having higher radiation efficiency following considerations are required

- The relative dielectric constant of the dielectric material should be less than 3 ($\varepsilon_r \leq 3$) in order to get higher radiation efficiency and directivity.
- The operating (resonant) frequency of the microstrip antenna should be less than 10GHz or higher than 50GHz (10GHz $\geq f_r \geq$ 50GHz) in order to get improved radiation efficiency. If less than 10GHz operating frequency range is used, the size of the patch will be increased proportionally and that opposes the fundamental compactness criteria of the microstrip antenna. Therefore it needs to use the operating frequency in the range above 50GHz to get enhanced efficiency.
- The height (h) of the substrate should be near to 1mm to get the higher radiation efficiency. Again, at the higher operating frequency, for the lower height of the substrate, surface wave increases that in turn increases the losses and reduces the radiation efficiency. Thus when the operating frequency is greater than 30GHz then the radiation efficiency of the microstrip antenna increases with the increasing height of the substrate.

Therefore, for the operating frequency less than 10GHz, the height should be h ≤ 1.5 mm. For the operating frequency greater than 30GHz, the height should be h ≥ 1.5 mm.

IV. Microstrip Array Antennas and Feed Networks

Microstrip antennas are used not only as single element but also very popular in arrays. Main limitation of microstrip is that it radiate efficiently only over a narrow band of frequencies and they can't operate at the high power levels of waveguide, coaxial line, or even stripline [2]. This can be minimized with the help of various array configurations, feeding methods, dielectric materials and ground planes. Antenna arrays are used to scan the beam of an antenna system, to increase the directivity, gain and enhance various other functions which would be difficult with single element antenna. In the microstrip array, elements can be fed by a single line or multiple lines in a feed network arrangement[12, 13]. Based on their feeding method the arrays are classified as

- Series feed network
- Corporate feed network
- Corporate-series feed network toolbar.

A. Microstrip Series Feed Network

A series feed microstrip array, as shown in Fig. 6, is formed by interconnecting all the elements with high impedance transmission line and feeding the power at the first element. Here two successive patch elements are matched by using quarter wavelength transformer method. Since, the feed arrangement is compact, the line losses associated with this type of array are lower than those of the corporate feed type [13].

The main beam direction and the scan sensitivity can be calculated from the following equations [13, 16]

$$d \sin\theta + \sqrt{\varepsilon} l = \lambda = \frac{c}{f} \quad (5)$$

$$\frac{\partial \theta}{\partial \lambda} = \frac{c}{\partial f^2 \cos\theta} \quad (6)$$

where d is the element spacing, l is the length of transmission line joining the successive elements, c is the velocity of light, f is the operating frequency, and θ is the beam-pointing angle measured from the broadside direction. For series feed linear array, we consider all excitation amplitude are same. The E-plane radiated fields for a single element patch can be expressed as [12, 17]

$$E = j \frac{k_0 W V_0 e^{-jk_0 r}}{r\pi} \left\{ \frac{\sin(\frac{k_0 h}{2}\cos\varphi)}{\frac{k_0 h}{2}\cos\varphi} \right\} \cos\left(\frac{k_0 L_e}{2}\sin\varphi\right) \quad (7)$$

where W is the width of the patch antenna, L_e is the extended length, $V_0 = hE_0$ is the voltage across radiating slot of the patch h

is the substrate height and r is the far field distance from the antenna.

The array factor can be written as

$$FA = \frac{\sin(N\pi d_x(u-u_0))}{\sin(\pi d_x(u-u_0))} \quad (8)$$

where, u=sinθ, u0 =sinθ0, dx is the element spacing and N is the number of elements. Combining array factor and element voltage radiation pattern we get the total element normalized power [18] radiation pattern that is

$$20\log(|E/FA|) \quad (9)$$

The main limitation of the series feed arrays is the large variation of the impedance and beam-pointing direction over a band of frequencies [13].

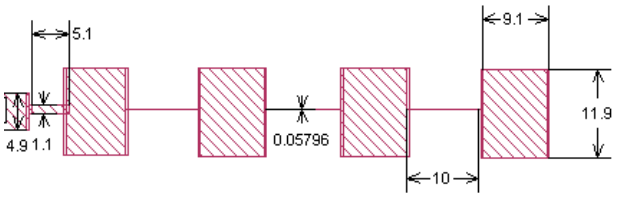


Figure 6. 4-elements series feed microstrip array antenna.

B. Microstrip Corporate Feed Network

Another popular microstrip antenna feeding system is the corporate feeding. Corporate feed arrays are general and versatile. This method has more control of the feed of each element and is ideal for scanning phased arrays, multi beam arrays. The phase of each element can be controlled using phase shifters while amplitude can be adjusted using either amplifiers or attenuators [12, 18]. The corporate feed network is used to provide power splits of 2n (i.e. n = 2; 4; 8; 16; etc.). This is accomplished by using either tapered lines or using quarter wavelength impedance transformers [13, 17]. Here, in the Fig. 7, the patch elements are connected by using the quarter wavelength impedance transformer method.

The radiated field equation of it is similar to that of the series feed array and the array factor as given in [18, 19] as

$$FA = \frac{\sin^2(N\pi(d_x/\lambda)\sin\theta)}{N^2\sin^2(\pi(d_x/\lambda)\sin\theta)} \quad (10)$$

We can get the normalized power radiation pattern by combining the element radiation pattern and array factor [20].

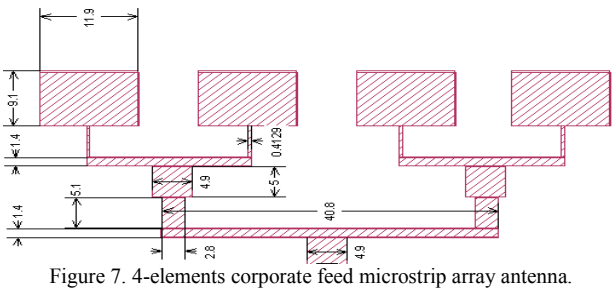


Figure 7. 4-elements corporate feed microstrip array antenna.

C. Microstrip Corporate-Series Feed Network

The combination of series feed and corporate feed are frequently used for array antennas [16, 19] to get benefits of both feeding networks. An 8-elemants array antenna can be constructed by using this method as shown in Fig. 8. It is a two dimensional rectangular planar array whose aperture illumination can be separated into two orthogonal planes such as the horizontal and vertical planes and the radiation pattern may then be written as the product the radiation patterns in these two planes. The array factor of this antenna with element spacing in the x and y direction of dx and dy respectively as given in [18] is

$$FA = \frac{\sin^2(N\pi(d_x/\lambda)\sin\theta_a)\times\sin^2(M\pi(d_y/\lambda)\sin\theta_a)}{(N^2\sin^2(\pi(d_x/\lambda)\sin\theta_a))\times(M^2\sin^2(\pi(d_y/\lambda)\sin\theta_a))} \quad (11)$$

where N = number of vertical elements of array that gives rise to the azimuth angle, θa and M = number of horizontal elements of array that gives rise to the elevation angle, θe. Multiplying the above equation with element radiated field gives normalized power radiation pattern.

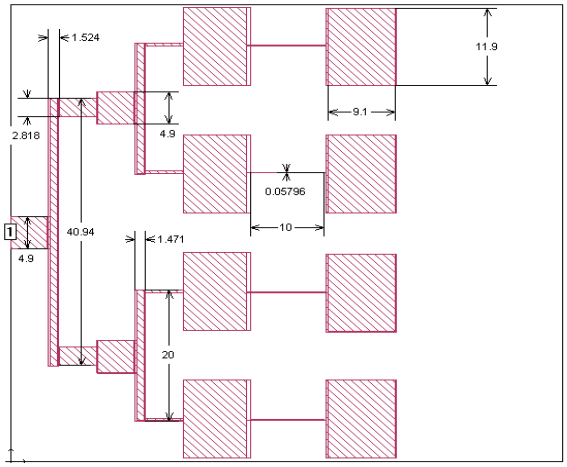


Figure 8. 8-elements corporate-series feed microstrip array antenna

V. Simulation Result & Discussion

A. Series Feed Array

In this design, it is considered that the substrate permittivity of the antenna is 2.2 (Taconic TLY-5), height is 1.588 mm ,and resonance frequency of the antenna is 10 GHz. Fig. 9 illustrates the current distribution of 4-elements series feed microstrip array antenna. It is apparent that current distribution is near about the same in each element. Fig. 10 shows that the return loss is -4.21dB at 10GHz and it is maximum of -6.26dB at 10.9GHz. Since the return loss is higher in lower frequency band, therefore antenna efficiency is lower at these frequencies. The maximum antenna efficiency can be obtained at 10.9GHz frequency. The simulated gain and directive gain of the antenna, according to fig. 11, are 11.97dB and 27.21dB respectively at $\theta = -25^0, \varphi = 0^0$ for the operating frequency 10 GHz.

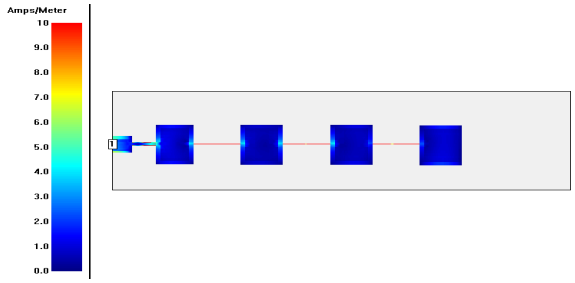


Figure 9. Current distribution of the 4-elements series feed microstrip array antenna.

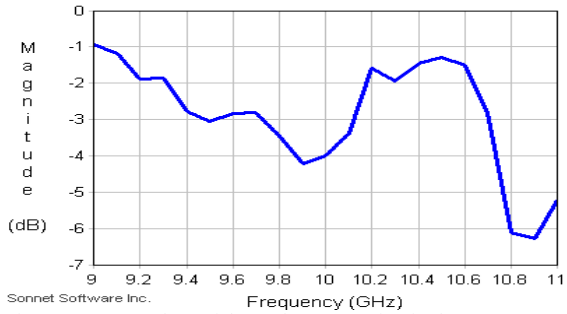


Figure 10. Return loss of the 4-elements series feed array antenna.

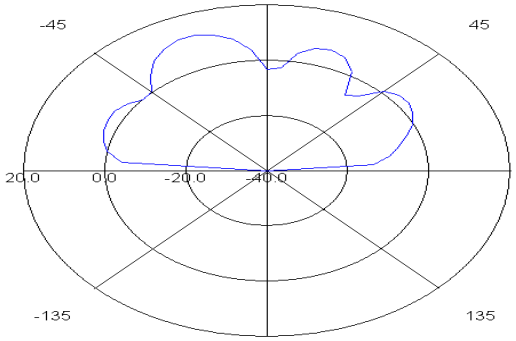


Figure 11. Radiation (polar plot) pattern of the 4-elements series feed array antenna.

B. *Corporate Feed Array*

Fig. 12 describes the current distribution of 4-elements corporate feed microstrip array antenna. In this array network, two successive patch elements as well as their corresponding transmission lines are matched by using quarter wavelength transformer method. Here, the substrate permittivity of the antenna is 2.2 (Taconic TLY-5), height is 1.588 mm and resonance frequency of the antenna is 10GHz. Fig. 13 presents that the maximum return loss is -25.456dB at 10GHz. Return losses increases for both lower and higher frequencies. Fig. 14 shows that the simulated gain and directive gain of the antenna are 14.14dB and 19.245dB respectively at θ=00, φ=00 for the operating frequency 10GHz.

C. *Corporate-Series Feed Array*

Fig. 15 illustrates the current distribution of 8-elements corporate-series feed microstrip array antenna. It is clear from the figure that the amount of current declines at the edge of elements. The same substrate as in series feed antenna of Taconic TLY-5 is used. The resonance frequency of the designed antenna is 10GHz. Fig. 16 demonstrates that return loss is -7.55dB at 10GHz and the peak value of return loss is -17.96dB at 9.6GHz. The simulated gain and directive gain of the antenna, as shown in fig. 17 are 17.48dB and 20.001dB respectively at θ=00, φ=00 for the operating frequency of 10GHz.

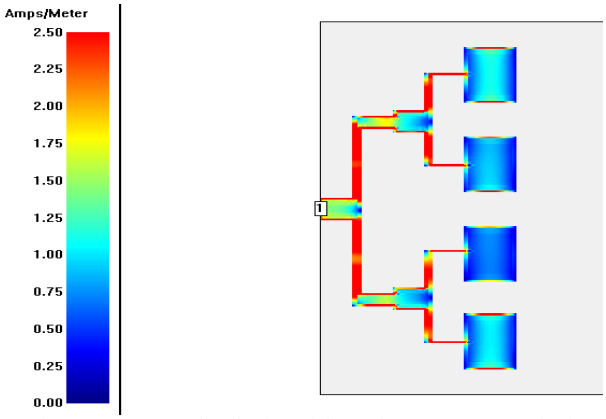


Figure 12. Current distribution of the 4-elements corporate feed microstrip array antenna.

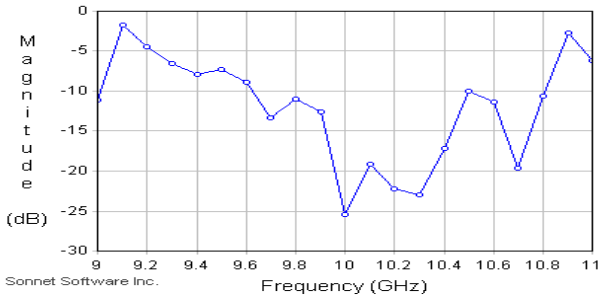


Figure 13. Return loss of the 4-elements corporate feed array antenna.

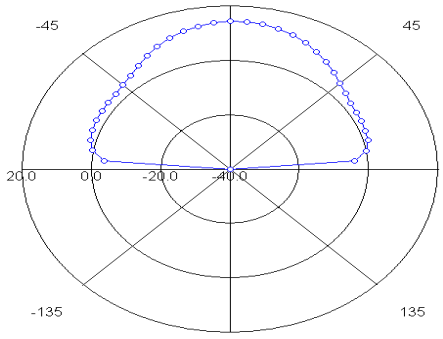


Figure 14. Radiation (polar plot) pattern of the 4-elements corporate feed array antenna.

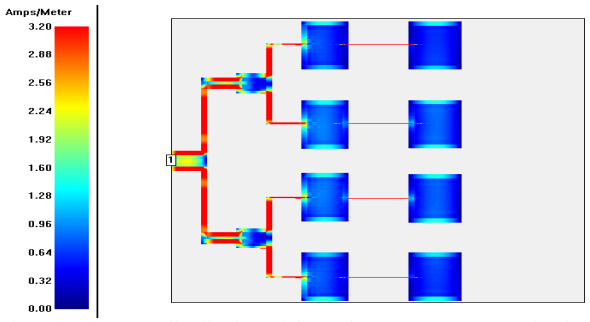


Figure 15. Current distribution of the 8-elements corporate-series feed microstrip array antenna.

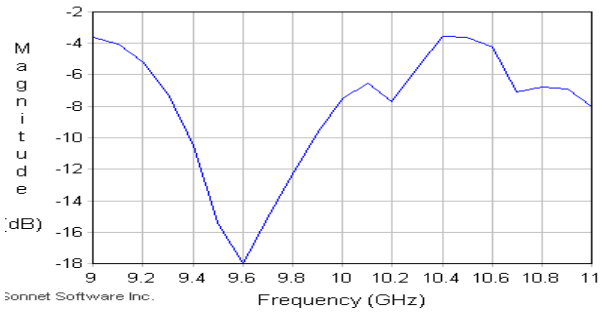


Figure 16. Return loss of for the 8-elements corporate-series feed array.

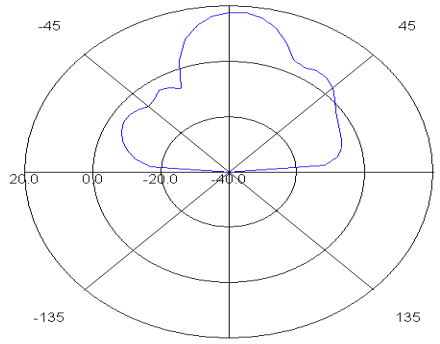


Figure 17. Radiation (polar plot) pattern of the 8-elements corporate-series feed array antenna

D. Comparison Among Three Differently Feed antennas

Table 1 compares three types of feed antennas. Series feed array is a simple and compact feeding method. As the feed arrangement is compact, the line losses associated with this type of array are lower than those of the corporate feed type. But this feeding gives a poor directive gain (HPBW = 140.30) with a large variation of beam-pointing directions. The corporate feed array has more control of the feed of each element as shown in fig. 12 whereas the current density of all elements are same near about 1 amp/meter. The peak value of return loss for corporate feeding is -25.456dB that is very much lower than that of series feed array of -4.21dB. As a result corporate feeding antennas have higher radiation efficiency than series feed antennas. In addition, corporate feeding provides better directivity and reduces the beam fluctuations over a band of frequencies.

The corporate-series feed array is a combination of these two feeding methods. This combined feeding offers HPBW of 400-600 and simulated gain of 17.48dB that are highest among three feeding methods. Therefore, it presents not only better directivity performance as the corporate feed array but also achieve a reduction of transmission line associated loss as the series feed array which can be easily realized by observing and comparing the current distribution of the fig. 9, fig. 12 and fig. 15 for series feed, corporate feed and corporate-series feed array antennas respectively.

TABLE I. COMPARISON AMONG THREE DIFFERENTLY FEED ANTENNAS

Performance Parameter	Feeding Types		
	Series feed (4x1)array	Corporate feed (4x1)array	Corporate-series (4x2)feed array
Physical Dimension	One	One	Two
HPBW	About 140.3^0	About 95.6^0	40^0-60^0
Simulated gain(dB)	11.97	14.14	17.48
Transmission line loss	Higher	Lower	Moderate
Return loss(dB)	-4.21	-25.456	-7.55

VI. CONCLUSION

These designed antennas are very simple, cost effective and high efficiency for the applications in GHz frequency ranges. The optimum design parameters (i.e. dielectric material, height of the substrate, operating frequency) are used to achieve the compact dimensions and high radiation efficiency. On the other hand, corporate-series feed antenna merges the advantages of both series and corporate feed antennas. The combined feeding antenna is planar array, therefore, this antenna can control the beam shape in both planes and provides more directivity and radiation efficiency. The operating frequency of all our designed antennas is about 10GHz which is suitable for X-band applications. It would also be possible to design an antenna operating in any other frequency bands by changing the design parameters. In future, we will investigate the spiral arrays with different feeding techniques which seem to be having more improved performances for both the series feed and corporate feed networks. At the same time, we will merge more than two patch elements operating at different frequencies by using quarter wavelength transformer method within an array network configuration to get multiband support.

REFERENCES

[1] M. T. I. Huque, et al., "Design and Simulation of a Low-cost and High Gain Microstrip Patch Antenna Arrays for the X-band Applications," in International Conference on Network Communication and Computer – ICNCC 2011, New Delhi, India., March 21-23, 2011.

[2] R. Mailloux, et al., "Microstrip array technology," Antennas and Propagation, IEEE Transactions on, vol. 29, pp. 25-37, 1981.

[3] H. Cheng-Chi, et al., "An aperture-coupled linear microstrip leaky-wave antenna array with two-dimensional dual-beam scanning capability," Antennas and Propagation, IEEE Transactions on, vol. 48, pp. 909-913, 2000.

[4] K. Gi-Cho, et al., "Ku-band high efficiency antenna with corporate-series-fed microstrip array," in Antennas and Propagation Society International Symposium, 2003. IEEE, 2003, pp. 690-693 vol.4.

[5] A. Abbaspour-Tamijani and K. Sarabandi, "An affordable millimeter-wave beam-steerable antenna using interleaved planar subarrays," Antennas and Propagation, IEEE Transactions on, vol. 51, pp. 2193-2202, 2003.

[6] A. Boufrioua and A. Benghalia, "Effects of the resistive patch and the uniaxial anisotropic substrate on the resonant frequency and the scattering radar cross section of a rectangular microstrip antenna," Aerospace science and technology, vol. 10, pp. 217-221, 2006.

[7] K. C. Lo and Y. Hwang, "Microstrip antennas of very high permittivity for personal communications," 1997, pp. 253-256 vol.

[8] J. M. Rathod, "Design Development of Antenna for TV Transmission for Connecting Outdoor Broadcasts Van to the Studio for Rural Areas," International Journal of Computer and Electrical Engineering, vol. 2, pp. 251-256, 2010.

[9] O. Barkat and A. Benghalia, "Radiation and resonant frequency of superconducting annular ring microstrip antenna on uniaxial anisotropic media," Journal of Infrared, Millimeter and Terahertz Waves, vol. 30, pp. 1053-1066, 2009.

[10] L. Choon Sae and T. Kuo-Hua, "Radiation efficiency of electrically small microstrip antennas with width discontinuities," Antennas and Propagation, IEEE Transactions on, vol. 53, pp. 871-873, 2005.

[11] M. F. Islam, et al., "Dual band microstrip patch antenna for SAR applications," Australian Journal of Basic and Applied Sciences, vol. 4, pp. 4585-4591, 2010.

[12] C. A. Balanis. (2005). Antenna theory : analysis and design (3rd ed.).

[13] R. Garg, Microstrip antenna design handbook. Boston, Mass. [u.a.]: Artech House, 2001.

[14] M. T. I. u. Huque, et al., "Design and performance analysis of the rectangular spiral microstrip antenna and its array configuration," in Antennas Propagation and EM Theory (ISAPE), 2010 9th International Symposium on, 2010, pp. 219-221.

[15] T. A. Milligan, et al. (2005). Modern antenna design. Available: http://dx.doi.org/10.1002/0471720615

[16] M. M. Alam, et al., "Design and performance analysis of microstrip array antenna," Session 5AP, p. 846.

[17] R. J. Mailloux, "Electronically Scanned Arrays," Synthesis Lectures on Antennas, vol. 2, pp. 1-82, 2007.

[18] H. J. Visser, "Array and Phased Array Antenna Basics," ed: John Wiley & Sons.

[19] M. I. Skolnik, Introduction to radar systems, 3rd ed. New York,: McGraw-Hill, 2000.

[20] W. L. Stutzman, "Estimating directivity and gain of antennas," Antennas and Propagation Magazine, IEEE, vol. 40, pp. 7-11, 1998.

AUTHORS PROFILE

Md. Tanvir Ishtaique-ul Huque was born in 1988 in Bangladesh. He received his B.Sc. Engineering degree from the Rajshahi University of Engineering & Technology (RUET) in 2010. Now he is working as a part time teacher in the Dept. of Electronics and Telecommunication Engineering of RUET. His research interests include the antenna application of the wireless body area network(WBAN) and next generation wireless communication system.

Md. Kamal Hosain was born in 1984 in Bangladesh. He received his B.Sc. Engineering degree from the Khulna University of Engineering & Technology (KUET), Bangladesh in 2006. Now, he is working as a Lecturer in the Dept. of Electronics and Telecommunication Engineering (ETE) of RUET. His research interests include the antennas and its medical applications.

Md. Shihabul Islam was born in 1987 in Bangladesh. He received his B.Sc. Engineering degree from the Rajshahi University of Engineering & Technology(RUET) in 2010 and now he is working as a system engineer of Technology Division in Grameen Phone Ltd. Which is a part of Telenor Group. His research interests include the antenna application and wireless sensor network.

Md. Al-Amin Chowdhury was born in 1988 in Bangladesh. he has completed his B.Sc. in Electronic and Telecommunication Engineering from Rajshahi University of Engineering & Technology(RUET) in 2010. He has keen interest to research on the optical fiber, different types of antennas. He wants to do his further study in USA on the communication field. World is becoming closer and closer due to the remarkable achievements in the communication field. He wants to receive the sound and proper knowledge in communication field so that he can contribute to the next generation demands in the communication sectors.

Backpropagation with Vector Chaotic Learning Rate

A.M. Numan-Al-Mobin, Mobarakol Islam,
Md. Rihab Rana, Md. Masud Rana
Dept. of Electronics & Communication Engineering,
Khulna University of Engineering & Technology,
Khulna – 9203, Bangladesh

Kaustubh Dhar,Tajul Islam,Md. Rezwan,M. Hossain
Dept. of Electronics & Communication Engineering,
Khulna University of Engineering & Technology,
Khulna – 9203, Bangladesh

***Abstract*—In Neural Network (NN) training, local minimum is an integrated problem. In this paper, a modification of standard backpropagation (BP) algorithm, called backpropagation with vector chaotic learning rate (BPVL) is proposed to improve the performance of NNs. BPVL method generates a chaotic time series as Vector form of Mackey Glass and logistic map. A rescaled version of these series is used as learning rate (LR). In BP training the weights of NN become inactive, after arrival of local minima in the training session. Using integrated chaotic learning rate, the weight update accelerated in the local minimum region. BPVL is tested on six real world benchmark classification problems such as breast cancer, diabetes, heart disease, australian credit card, horse and glass. The proposed BPVL outperforms the existing BP and BPCL in terms of generalization ability and also convergence rate.**

Keywords—Neural network; backpropagation; BPCL; BPVL chaos; generalization ability; convergence rate.

I. INTRODUCTION

Gradient based methods are one of the most widely used error minimization methods used to train back propagation networks. The BP training algorithm is a supervised learning method for multi-layered feed forward neural networks [1, 20]. BP have been applied to a wide variety of problems, including pattern recognition, signal processing, image compression, speech recognition etc due to its most appealing features and adaptive nature [2].It is essentially a gradient descent local optimization technique which involves backward error correction of the network weights. Despite the general success of BP in learning the neural networks, several major deficiencies are still needed to be solved [3, 4]. First, the BP algorithm will get trapped in local minima especially for non-linearly separable problems [4]. Having trapped into local minima, BP may lead to failure in finding a global optimal solution [6] second; the convergence rate of BP is still too slow even if learning can be achieved [5]. Furthermore, the convergence behavior of the BP algorithm depends very much on the choices of initial values of connection weights and the parameters in the algorithm such as the learning rate and the momentum. Improving the training efficiency of neural network based algorithm is an active area of research and numerous papers have been proposed in the literature. Early days BP algorithm saw further improvements.

There are many improvement and variations of BP with goals of enlarged speed of convergence, avoidance of local minima and improvement in the network's ability to generalize. BP trains the NNs with constant values of learning rate (LR) and momentum factor. When the LR and momentum factor are made adaptive with the training, the performance of BP is increased [8, 9]. The speed of convergence is accelerated for BP algorithm by using adaptive accuracy of weights, instead of using fixed weight accuracy [10]. Dynamic LR of BP algorithm is optimized by using an efficient method of deriving the first and second derivatives of the objective function with respect to the LR [11]. Another approach that does not require the second order derivative has been proposed to optimize LR. In this case, a set of recursive formula is formed which accelerate the convergence of BP with remarkable savings in running time [12]. Reducing the number of patterns in the active training set effectively increases training efficiency, and accordingly, permits training with a larger pattern set [13]. Many modifications that have been proposed to improve the performance of BP have focused on solving "flat spot" [7] problem to increase the generalization ability. However, their performance is limited due to the error overshooting problem. A novel approach called 2P-MGFPROP has been introduced to overcome the error overshooting problem and hence it speeds up the convergence rate. C. C. Cheung enhanced this approach by dividing the learning process into multiple phases, and different fast learning algorithms are assigned in different phases to improve the convergence rate [15]. All these methods require much computational effort and these do not guarantee good generalization ability for all the cases.

BPCL (Backpropagation with Chaotic Learning Rate) escapes the NN training from premature saturation with chaotic LR. In BPCL to generate chaos Logistic map is use, which is very fast. For this reason we search another learning criteria which is perform much better than BPCL. In this paper, a modified BP algorithm, called 'Backpropagation with Vector Chaotic Learning Rate' (BPVL) is proposed which is a vector form of Mackey glass and Logistic map chaotic time series . BPVL is applied on several benchmark problems including breast cancer, diabetes, heart disease, Australian credit card & horse. For all the problems, BPVL outperforms BP & BPCL in terms of generalization ability and convergence rate.

The rest of the paper is organized as follows. The proposed BPVL is described in section II. Section III includes the experimental studies. The discussion on BPVL is presented in section IV. Section V. concludes the paper.

II. BPVL

BPVL follows the standard BP as well as BPCL. In BPCL the learning rate is chaotic but in BPVL the learning rate is a vector. The Vector is formed by two chaotic time series, one is logistic map [18, 19] and another is Mackey Glass[22,23].

The rescaled LR (RLR) is produced by the following eq^n :

$$V(t)=[iC(t) + jM(t)]$$

$$RLR = LR * abs(V(t)) - \beta$$

β are user specified parameters. β is always kept less than LR close to zero.

C (t) is logistic map time series & M (t) is Mackey Glass time series. However, proposed BPVL is explained as follows.

A feed-forward neural network with one input layer, one hidden layer, and one output layer is shown in Fig. 1. Let, the numbers of input, hidden and output units are *I, J* and *K* respectively. w_{ij} is the network weight that connects input unit *i* and hidden unit *j*, and w_{jk} is the network weight that connects hidden unit *j* and output unit *k*. The number of training examples is *N* and any arbitrary *n*-th training example is $\{x_{n1}, x_{n2}, \ldots\ldots, x_{nI}; y_{n1}, y_{n2}, \ldots\ldots, y_{nK}\}$, where x_n is the input vector and y_n is the target vector. h_{nj} and o_{nk} are the outputs of hidden unit *j* and output unit *k* for the *n*-th training example. Δ represents the difference between the current and new value of the network weights. The consecutive steps of BPVL are given below.

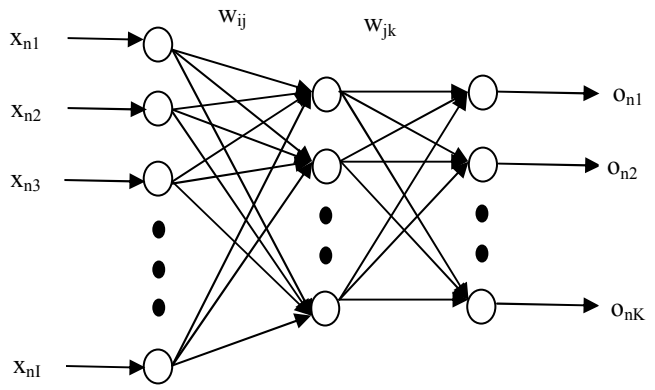


Figure 1. A feed-forward neural network with single hidden layer.

Step 1) Initialize the network weights in an arbitrary interval (-w, +w). Set the iteration number, ITE = 0.

Step 2) Compute the Rescaled LR by using following equation:

$$V(ITE)=abs[iC(ITE) + jM(ITE)]$$

$$RLR\ (ITE) = LR * V(ITE) - \beta$$

Step 3) Compute h_{nj} and o_{nk} for the n-th training example by using the following:

$$h_{nj} = f\left(\sum_{i=1}^{I} w_{ij}\, x_{ni}\right)$$

$$o_{nk} = f\left(\sum_{j=1}^{J} w_{jk}\, h_{nj}\right)$$

Where, the sigmoid function $f(x) = 1/(1 + e^{-x})$ is used as the activation function. Compute the changes of the weights, Δw_{ij} and Δw_{jk} for the n-th example using the following:

$$\Delta w_{jk} = -RLR\ (ITE)\frac{\partial E_n}{\partial w_{jk}}$$

$$= -RLR\ (ITE)\ \delta_{nk} h_{nj}$$

$$\Delta w_{ij} = -RLR\ (ITE)\delta_{nj} x_{ni}$$

Where,

$$E_n = \frac{1}{2}\sum_{k=1}^{K}[o_{nk} - y_{nk}]^2$$

$$\delta_{nk} = (o_{nk} - y_{nk}) o_{nk}(1 - o_{nk})$$

$$\delta_{nj} = h_{nj}\left(1 - h_{nj}\right)\sum_{k=1}^{K}\delta_{nk} w_{jk}$$

Update w_{ij} and w_{jk} by adding Δw_{ij} and Δw_{jk} with them correspondingly. When this process is repeated for all the training examples, iteration is completed. ITE is increased by one.

Step 4) Check the termination condition. If the termination condition is fulfilled, the training is stopped and the trained NN is tested; otherwise go to **step 2**.

III. EXPERIMENTAL STUDIES

A. Characteristics of Benchmark datasets

BPVL is applied on six benchmark classification problems – breast cancer, diabetes, heart disease, Australian credit card, horse, and glass identification. The datasets are collected from the University of California at Irvine (UCI) Repository of the machine learning database [16] and PROBEN1 [17]. Some characteristics of the datasets are listed in TABLE I. For example, cancer is a 2 class problem having total examples of 699 with 9 attributes. Other problems are arranged in a similar fashion. The total examples of a dataset are divided into – training examples, validation examples, and testing examples. The training examples are used to train the NN, validation examples are used to terminate the training process and the trained NN is tested with the testing examples.

TABLE I. CHARACTERISTIC OF BENCHMARK DATASETS.

Datasets	Number of		
	Total Examples	*Input Attributes*	*Output Classes*
Breast Cancer	699	9	2
Diabetes	768	8	2
Heart Disease	920	35	2
Credit Card	690	51	2
Horse	364	58	3
Glass	214	9	6

B. Experimental Process and Comparison

A feed-forward NN with single hidden layer is taken for each problem. The numbers of input and output units of the NN are equal to the number of attributes and number of classes of the dataset respectively. The number of hidden units is taken arbitrarily for several datasets. The weights of the NN are initially randomized in the interval (-1, +1). The training is stopped when the mean square error on the validation set increases in consecutive five iterations. In order to check the generalization performance of trained NN, the 'testing error rate' (TER) is computed. TER is the ration of the number of

classified testing examples to the number of testing examples. The experimental results are reported in terms of TER and number of required iterations. Mean and SD indicate the average and standard deviation values of 20 independent trials. BPVL is compared with standard BP & BPCL. To make a fair comparison, the LR of BP is selected in such a way that, this fixed value of learning rate is always an intermediate point in the range [min (RLR) & max (RLR)]

1) Breast Cancer: This problem has 699 examples. The first 350 examples are used as training, 175 examples as validation and last 174 examples for testing. A NN of 9-4-2 (nine input units, four hidden units, and two output units) is taken. The experimental results are listed in TABLE II. When RLR is -0.03 to 0.15, BPVL requires 75.40 iterations to obtain TER of 0.7115(dB), BPCL requires 87.85 iterations to obtain TER of 1.0037(dB) In this range, the LR of BP is considered at 0.06. The TER with BP is 1.9033(dB) and BP requires 152.65 iterations. Here TER is taken in dB 100 times of orginal error Hence BPVL obtains good generalization ability than that of BPCL as well as BP and the number of required iteration with BP is about two times of BPVL and Less than BPCL. Fig. 2 shows the training error of BPVL,BPCL and BP with respect to iteration..

TABLE II. TERS AND NUMBER OF REQUIRED ITERATIONS WITH BP,BPCL AND BPVL FOR CANCER PROBLEM OVER 20 INDEPENDENT RUNS.

Exp. No.	Algorithm	RLR/LR	TER		Iteration	
			Mean(dB)	*SD*	*Mean*	*SD*
01.	BP	.06	1.9033	0.0026	152.65	40.0003
	BPCL	-0.03 to .15	1.0037	0.0029	87.85	23.0636
	BPVL		0.7115	0.0012	75.40	9.6409
02.	BP	0.11	1.2057	0.0026	99.60	18.8981
	BPCL	-0.03 to 0.25	0.7555	0.0017	34.50	5.5812
	BPVL		0.2938	0.0012	32.20	6.4156

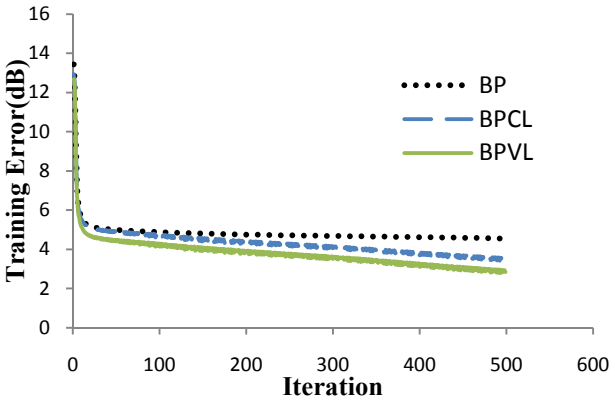


Figure 2. Training error vs. iteration of BP,BPCL and BPVL for cancer problem.

2) Diabetes: This problem has 768 examples .The first 384 examples are used for training, 192 examples for validation and the last 192 examples are used for testing. Here the size of NN is 8-4-2. The results for different LRs are reported in TABLE III. It is shown that BPVL has fast convergence rate than BPCL & BP for all the cases.

TABLE III. TERS AND NUMBER OF REQUIRED ITERATIONS WITH BP, BPCL AND BPVL FOR DIABETES PROBLEM OVER 20 INDEPENDENT RUNS.

Exp. No.	Algorithm	RLR/LR	TER		Iteration	
			Mean(dB)	*SD*	*Mean*	*SD*
01.	BP	.06	13.9058	0.0276	255.95	81.0108
	BPCL	-0.03 to 0.15	13.8650	0.0104	109.90	61.9725
	BPVL		13.8003	0.0066	91.40	15.0176
02.	BP	0.11	13.8970	0.0087	163.65	48.6141
	BPCL	-0.03 to 0.25	13.8686	0.0097	81.70	18.0710
	BPVL		13.7401	0.0070	63.50	6.5459

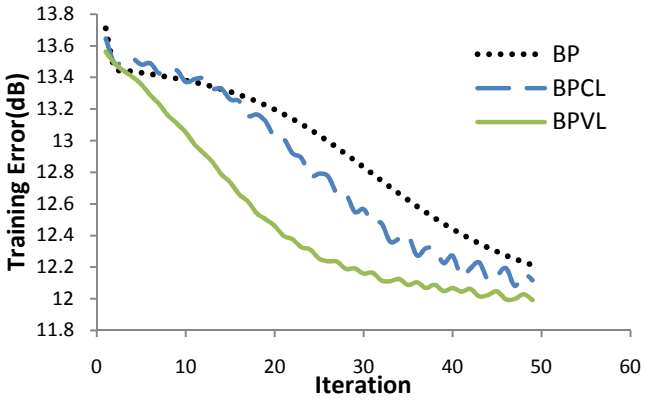


Figure 3. Training error vs. iteration of BP,BPCL and BPVL for diabetes problem.

3) Heart Disease: The size of NN is considered as 35-4-2. This problem has 920 examples. The first 460 examples are used to train the network, 230 examples for validation and the trained network is tested with last 230 examples. The obtained average results are reported in TABLE IV. BPVL outperforms BP & BPCL in all the cases. For example, when RLR is -0.03 to 0.15, TER and iteration of BPVL are 13.0125(dB) and 21.50, while these are 13.1952(dB) and 57.75 for standard BP & 13.0535(dB) and 31.50 for BPCL. The convergence curve of BPVL is also better as shown in Fig. 4.

TABLE IV. TERS AND NUMBER OF REQUIRED ITERATIONS WITH BP, BPCL AND BPVL FOR HEART PROBLEM OVER 20 INDEPENDENT RUNS.

Exp. No.	Algorithm	RLR/LR	TER		Iteration	
			Mean(dB)	*SD*	*Mean*	*SD*
01.	BP	.06	13.1952	0.0085	57.75	26.5191
	BPCL	-0.03 to 0.15	13.0535	0.0092	31.50	17.3882
	BPVL		13.0125	0.0057	21.50	7.8835
02.	BP	0.11	13.1218	0.0096	32.30	16.7186
	BPCL	-0.03 to 0.25	13.0557	0.0086	16.40	7.4659
	BPVL		13.0146	0.0084	13.95	4.6419

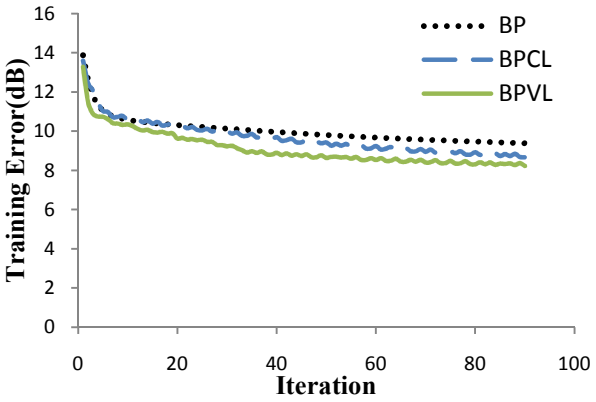


Figure 4. Training error vs. iteration of BP,BPCL and BPVL for heart problem.

4) Australian Credit Card: A 51-4-2 NN is trained with first 345 examples. The training is stopped with 173 validation examples and the NN is tested with last 172 testing examples. TABLE V shows the comparative results between BPCL and BP. When RLR is -0.03 to 0.15, TER and iteration of BPVL are 11.4489(dB) and 22.40, while these are 11.8011(dB) and 41.90 for standard BP 11.5776(dB) & 31.80 for BPCL respectively. These results ensure that the generalization ability of BPVL is better than that of BP & BPCL for credit card problem. Fig. 5 shows the competitive convergence curves of BPVL,BPCL and BP.

TABLE V. TERS AND NUMBER OF REQUIRED ITERATIONS WITH BP, BPCL AND BPVL FOR CARD PROBLEM OVER 20 INDEPENDENT RUNS.

Exp. No.	Algorithm	RLR/LR	TER		Iteration	
			Mean(dB)	*SD*	*Mean*	*SD*
01.	BP	.06	11.8011	0.0112	41.90	8.1234
	BPCL	-0.03 to 0.15	11.5776	0.0089	31.80	4.8846
	BPVL		11.4489	0.0107	22.40	2.9899
02.	BP	0.11	11.8013	0.0096	23.95	6.9700
	BPCL	-0.03 to 0.25	11.4737	0.0072	19.10	4.1901
	BPVL		11.1461	0.0810	14.65	3.7319

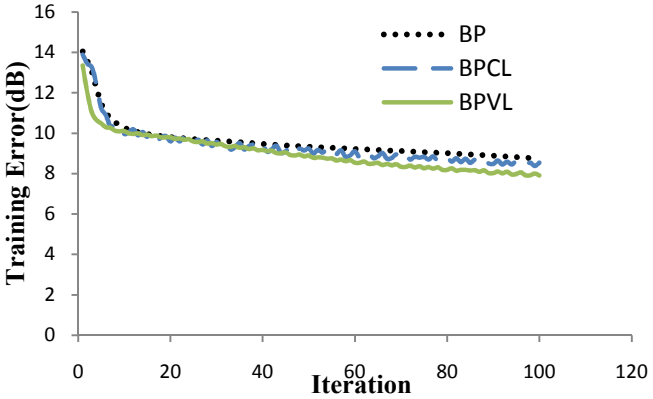


Figure 5. Training error vs. iteration of BP, BPCL and BPVL for card problem.

5) Horse: The horse dataset has total 364 examples. The first 182 examples are used for training, 91 examples for validation and last 91 examples for testing. A 58-7-3 NN is trained. The numerical results are reported in TABLE VI and the error curves are shown in Fig. 6. When RLR is -0.02 to 0.15, TER and the number of required iterations of BPVL are 11.5776(dB) and 30.40 while these are 14.6419(dB) and 47.05 for standard BP,. 14.5163(dB) and 38.5 for BPCL respectively.

TABLE VI. TERS AND NUMBER OF REQUIRED ITERATIONS WITH BP, BPCL AND BPVL FOR HORSE PROBLEM OVER 20 INDEPENDENT RUNS.

Exp. No.	Algorithm	RLR/LR	TER		Iteration	
			Mean(dB)	*SD*	*Mean*	*SD*
01.	BP	.06	14.6419	0.0185	47.05	13.3958
	BPCL	-0.02 to 0.15	14.5163	0.0271	38.55	13.9373
	BPVL		14.4576	0.0135	30.40	7.0092
02.	BP	0.11	14.6090	0.0301	36.20	12.3150
	BPCL	-0.02 to 0.25	14.3965	0.0201	24.90	6.7742
	BPVL		14.2911	0.0233	23.55	12.214

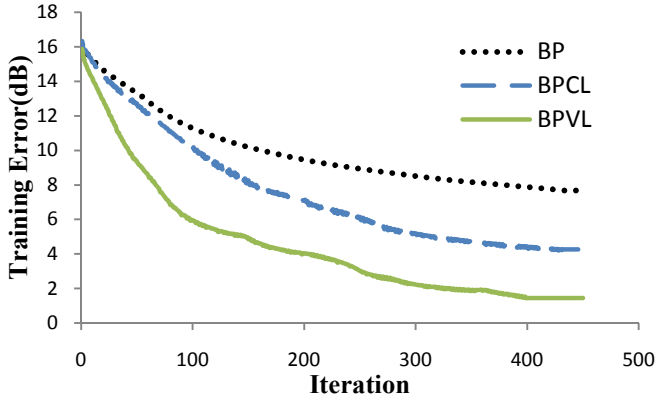


Figure 6. Training error vs. iteration of BP, BPCL and BPVL for horse problem.

6) Glass: The glass dataset has total 214 examples. The first 107 examples are used for training, 54 examples for validation and last 53 examples for testing. A 9-6-6 NN is trained. The numerical results are reported in TABLE VII and the error curves are shown in Fig. 7. When RLR is -0.02 to 0.15, TER and iteration of BPVL are 15.9295(dB) and 131.60, while these are 16.0239(dB) and 186.55 and 16.2407(dB) and 410.50 for BPCL and standard BP respectively. Here the numbers of required iterations for all examples with BP are about four times than that of with BPVL. Fig. 7 shows that BPCL has better convergence rate than that of BP.

TABLE VII. TERS AND NUMBER OF REQUIRED ITERATIONS WITH BP,BPCL AND BPVL FOR GLASS PROBLEM OVER 20 INDEPENDENT RUNS.

Exp. No.	Algorithm	RLR /LR	TER		Iteration	
			Mean(dB)	SD	Mean	SD
01.	BP	0.06	16.2407	0.1307	410.50	171.98
	BPCL	-0.02 to 0.15	16.0239	0.1367	186.55	79.09
	BPVL		15.9295	0.0680	131.60	31.20
02.	BP	0.11	16.1109	0.0719	360.00	175.12
	BPCL	-0.02 to 0.25	15.4986	0.0670	91.12	21.60
	BPVL		15.4370	0.0256	80.51	15.78

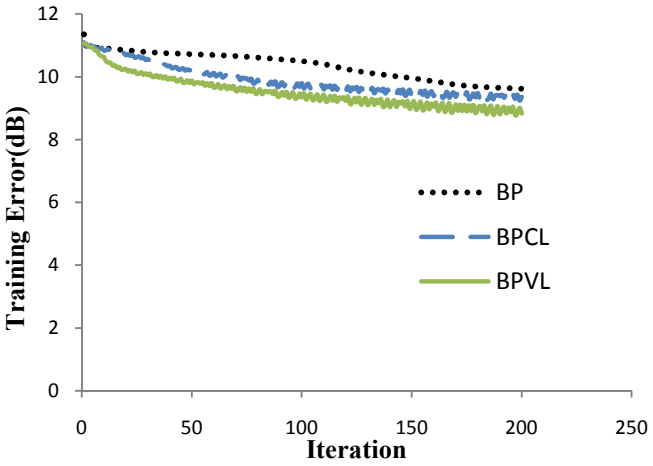


Figure 7. Training error vs. iteration of BP, BPCL and BPVL for glass problem.

IV. DISCUSSION

The difference between BPCL and BPVL is that BPCL makes the LR chaotic during the training by logistic map while BPVL trains the network by a Vector learning rate, which is a vector form of two chaotic time series, Mackey glass and Logistic map. Logistic map is the fastest chaotic time series; that is why consecutive learning rate difference is high in BPCL. To optimize this problem vector learning rate is used in BPVL. The training error curve of BPVL is improved and sometimes it is marginally improved than that of BPCL & BP. However, BPCL is proposed to achieve a good generalization ability, not to improve the convergence, although BPVL shows fastest convergence too.

The NN training is terminated based on validation examples. When the mean square error on the validation set starts to increase, BPVL stops the training. There is a possibility to fall the validation error into local minima and so the mean square error is checked in five consecutive iterations.

V. CONCLUSION

In this paper, a supervised training algorithm, called BPVL is proposed. BPVL works on the learning parameter of training. Standard BP trains NNs with a constant value of LR. On the other hand, biological systems such as human brain involve chaos. To fill up this gap, BPVL trains NNs with LR which is a vector chaotic time series. The chaotic series is generated by using the complex vector form of the Mackey Glass and logistic map and a rescaled version of this is intentionally incorporated with the training. Due to the nonlinear error surface of real world problem, BP training often falls into premature saturation that leads the training to a non updating zone. Since most of the time, the LR is positive and sometimes it is negative, BPVL resolves this problem.

BPVL is applied on six benchmark classification problems to observe the training performance. BPVL is capable to train NNs with better generalization ability and also faster convergence rate than that of standard BP and BPCL.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable suggestion.

REFERENCES

[1] M. K. S. Alsmadi, K. B. Omar, and S. A. Noah, "Back Propagation Algorithm: The Best Algorithm Among theMulti-layer Perceptron Algorithm", International Journal of Computer Science and Network Security, vol. 9, no. 4, April 2009.

[2] X. Zhou, J. Zheng, and K. Mao, "Application of Modified Backpropagation Algorithm to the Prediction of the Chloride Ion Concentration in Cracked Concrete", In the Proc. of 3rd International Conference on Natural Computation, pp. 257-261, 2007.

[3] S. Haykin, "Neural Networks, A Comprehensive Foundation", New York 10022: IEEE Society Press, Macmillan College Publishing, *1994.*

[4] S. E. Fahlman, "An empirical study of learning speed in back propagation networks," tech. rep., CMU-CS-88-162, Carnegie Mellon University, Pittsburgh, PA., 1988.

[5] M. A. Otair and W. A. Salameh, "Speeding Up Back-Propagation Neural Networks", Proceedings of the Informing Science and IT Education Joint Conference, 2005.

[6] M. Gori and A. Tesi, "On the problem of local minima in backpropagation," IEEE Trans. Pattern Anal. Machine Intell., vol. 14, pp. 76–86, 1992.

[7] J. E. Vitela and J. Reifman, "Premature saturation in backpropagation networks: Mechanism and necessary conditions," Neural Netw., vol. 10, no. 4, pp. 721–735, 1997.

[8] C. C. Yu, and B. D. Liu, "A Backpropagation Algorithm With Adaptive LR And Momentum Coefficient", In the Proc. of International joint Conference on neural network, pp. 1218-1223, 2002.

[9] S. Ng, C. Cheung, and S. Leung, "Magnified Gradient Function With Deterministic Weight Modification in Adaptive Learning", IEEE trans. neural netw, vol. 15, no. 6, 2004.

[10] C. Gonzalo, Q. Tian, Y. Fainman, and S. H. Lee, "Fast Convergence of the Backpropagation Learning Algorithm by Using Adaptive Accuracy of Weights", Proceedings of the 35th Midwest Symposium on Circuits and systems, vol. 2, pp. 1221-1224, 1992.

[11] X. H. Yu, "Dynamic LR Optimization of the Backpropagation Algorithm", IEEE Transactions on Neural Networks, vol. 6, no. 3, May 1995.

[12] X. H. Yu and G. A. Chen, "Efficient estimation of dynamically optimal LR and momentum for backpropagation learning", In the Proc. of IEEE International Conference on Neural network, vol. 1, pp. 385-388, 1995.

[13] E. M. Strand and W. T. Jones, "An Active Pattern Set Strategy for Enhancing Generalization While Improving Backpropagation Training Efficiency", In Proc. of International Joint Conference on neural network, vol. 1, pp. 830-834, 1992.

[14] C. C. Cheung and S. C. Ng, "Backpropagation with Two-Phase Magnified Gradient Function", Proceedings of IEEE World Congress on Computational Intelligence (WCCI), Hong Kong, 2008.

[15] Chi-Chung Cheung, "The Multi-Phase Method in Fast Learning Algorithm", Proceeding of international join conference on neural networks, Atlanta Georgia, USA, June 14-19, 2009.

[16] A. Asuncion and D. Newman, ''UCI Machine Learning Repository", Schl. Inf. Comput. Sci., Univ. California, Irvine, CA, 2007.

[17] L. L. Prechelt, "PROBEN1-A set of neural network benchmark problems and benchmarking rules", Technical Report 21/94, Faculty of Informatics, University of Karlsruhe, 1994.

[18] Junshan Gao, Baiyu Sun and Jiaxiang Yang, "The Construction of the Cubic Logistic Chaotic Function Family", In Proc. of Sixth World Congress on Intelligent Control and Automation, 2006.

[19] Olivares, E. I. Vazquez-Medina, R. Cruz-Irisson, and M. Del-Rio-Correa, "Numerical calculation of the lyapunov exponent for the logistic MAP", 12th International Conference on Mathematical Methods in Electromagnetic Theory, 2008.

[20] Rumelhart, D.E., G.E. Hinton and R.J. Williams, "Learning Internal Representations by Error Propagation", Parallel Distributed Processing,MIT ,pp: 318-362, 1986.

[21] Mobarakol Islam, Md. Rihab Rana, S.U. Ahmed, Md Shahjahan, "Training Neural Network with Chaotic Learning Rate", In the proc. of International Conference on Emerging Trends in Electrical and Computer Technology, March 23-24, 2011.

[22] M. Farzad, H. Tahersima and H. Khaloozadeh, "Predicting the Mackey Glass Chaotic Time Series Using Genetic Algorithm", Proceeding of International Joint Conference, SICE-ICASE, pp. 5460 – 5463, 2006.

[23] W.C Mead, R.D Jones, Y.C. Lee, C.W. Barnes, G.W. Flake, L.A. Lee and M.K. O'Rourke, "Using CNLS-net to predict the Mackey-Glass chaotic time series" In the proc. of International Joint Conference on Neural Networks(IJCNN) , vol.2, pp. 485 – 490, 1991.

[24] Gupta, S., Doshi, V., Jain, A., & Iyer, S. (2010). Iris Recognition System using Biometric Template Matching Technology. *International Journal of Advanced Computer Science and Applications - IJACSA*, *1*(2), 24-28. doi: 10.5120/61-161.

A. M. Numan-Al-Mobin received the B.Sc. degree in Electronics and Communication Engineering from Khulna University of Engineering & Technology, Bangladesh, 2008. He was a Rollout Junior Engineer, July 2008 to October 2008, in Orascom Telecom Bangladesh Limited (Banglalink). He is currently working as a lecturer in the same department of the university. His current research interests include electromagnetic, antennas and propagation, wireless communication and signal processing.

Mobarakol Islam received the B.Sc. degree in Electronics & Communication Engineering from Khulna University of Engineering & Technology,Bangladesh, 2011. He is presently working in the Samsung Bangladesh Research & Development Centre as Software Engineer-1. His research interest include machine learning, pattern recognition, neural networks, image processing, telecommunication and feature selection.

Md. Rihab Rana received B.Sc. degree in Electronics & Communication Engineering from Khulna University of Engineering & Technology,Bangladesh, 2011. His research interest include machine learning,neural networks, feature selection and telecommunication.

Md. Masud Rana is currently the Assistant Professor at Department of Electronics & Communication Engineering, Khunla University of Engineering & Technology, Bangladesh. He received the B.Sc in Electronics & Communication Engineering and M.Eng. in Electronics & Radio Engineering degrees from Khunla University of Engineering & Technology and Kyung Hee University, Korea, respectively. His current research interests include broadband digital communications systems, channel estimation and tracking, interference prediction and cancellation, equalization, digital signal processing in communications, OFDM, cooperative communications as well as single and multicarrier transmission. He is a member of the Institute of Electronics Engineers of Korea (IEEK), Institute of Engineers Bangladesh (IEB), live member of Bangladesh Electronics Society (BES), and student member of IEEE.

Kaustubh Ratan Dhar received B.Sc. degree in Electronics & Communication Engineering from Khulna University of Engineering & Technology,Bangladesh in 2011. His reaearch field interests include Neural Networks, Telecommunication,Numerical computation of electromagnetic fields, integrated optical devices and waveguide materials

Md. Tajul Islam received B. Sc. degree in Electronics & Communication Engineering from Khulna University of Engineering & Technology,Bangladesh, 2011. His reaearch field is Neural Networks,Telecommunication and VLSI chip design.

Md. Rezwanul Haque received B. Sc. degree in Electronics & Communication Engineering from Khulna University of Engineering & Technology,Bangladesh, 2011. His reaearch field is Neural Networks,Telecommunication and VLSI chip design.

Md. Moswer Hossain received B. Sc. degree in Electronics & Communication Engineering from Khulna University of Engineering & Technology,Bangladesh, 2011. His reaearch field is Neural Networks, feature selection and Telecommunication .

A study on classification of EEG Data using the Filters

V. Baby Deepa
Department of Software Engineering,
M. Kumarasamy College of Engineernig, Karur,
Tamil Nadu, 639 113

Dr. P. Thangaraj
Department of Computer Science & Engineering,
Bannari Amman Institute of Technology,
Sathyamangalam, Erode, 638 401

***Abstract*— In the field of data mining, classification of data is being a difficult task for further analysis. Classifying the EEG data would require more efficient algorithms. In this paper the classification filters such as Fast Hartley Transform (FHT) and Chebyshev filters are used to classify the EEG data signals. In a bulk data set of EEG signals, the signals are classified into many channels. Though various filters are available for classification, FHT with Chebyshev and FT tree only are taken to know the efficiency in classifying the EEG data signals. When these filters are applied to the data instances the percentage of correctly classified instances is high. Based on the experimental result it is suggested that these filters could be used for the enhancement of classification of EEG data.**

Keywords- EEG (Electro-encephalogram); BCI (Brain Computer Interface; FHT (Fast Hartley Transform); Chebyshev filters; FT tree.

I. INTRODUCTION

The EEG data set obtained from BCI is used for classification. In the EEG signals there would be a cluster of features. It is vital to extract good features from that cluster. Classification of EEG dataset involves much careful effort. Identifying and extracting good features from the signals is a crucial step in the design of BCI.

The features extracted from EEG are not relevant and do not describe well the neurophysiologic signals employed, the classification algorithm which will use such features will have trouble in identifying the class of these features i.e., mental state of the user. Consequently the correct recognition rates of mental states will be very low, which will make use of the interface not convenient or even impossible for the user. It is sometimes possible to use raw signals as the input of the classification algorithm, it is recommended to select and extract good features in order to maximize the performances of the system by making easier the task of subsequent classification algorithm.

According to researchers, it seems that the choice of a good preprocessing and feature extraction method have more impact on the final performances than the choice of a good classification algorithm. In the section II the classification filters such as Fast Hartley Transform, Chebyshev filters and FT are described. This section gives a clear picture of how the classification filters function. The filters are applied to the data during the experimentation and the results are given in the section III. Finally it has been concluded that a combination of FHT, Chebyshev and FT tree have the potential to enhance the classification of EEG data.

II. CLASSIFICATION FILTERS

A. Discrete Hartley transform (DHT)

A Discrete Hartley transform (DHT) [1] is a Fourier-related transform of discrete, periodic data similar to the discrete Fourier transform (DFT), with analogous applications in signal processing and related fields. Its main distinction from the DFT is that it transforms real inputs to real outputs, with no intrinsic involvement of complex numbers. Because there are fast algorithms for the DHT analogous to the fast Fourier transform (FFT), the DHT was originally proposed by R. N. Bracewell in 1983 as a more efficient computational tool in the common case where the data are purely real. It was subsequently argued, however, that specialized FFT algorithms for real inputs or outputs can ordinarily be found with slightly fewer operations than any corresponding algorithm for the DHT.

B. The Fast Hartley Transform (FHT)

EEG data are inherently real valued, yet most general Fourier transform algorithms [8] accept complex valued input and return complex valued output. The generality of these algorithms is also their weakness, for in the process of transforming real data they perform twice as many operations (arithmetic, address and transfer) as is necessary. Since the Fourier transform is commonly used in the analysis of real signals, special versions of almost every transform algorithm have been developed to deal more efficiently with real data. Unfortunately, when it comes to inverse transformation, *another* special version of the algorithm is required to efficiently transform the complex output back into the real sequence.

The Hartley transform distinguishes itself from its close cousin, the Fourier transform [5], by being real valued; it produces real output from real input. Even so, it provides the same phase and amplitude information about the data as the Fourier transform. The Hartley transform may also be computed using a ‘fast’ algorithm which requires O(Nlog2N) operations. Finally, the fast Hartley transform (FHT) [2, 3] is twice as fast as a complex valued FFT, requiring virtually the same number of operations as the real valued FFT algorithms.

Formally, the discrete Hartley transform [4, 6, 7] is a linear, invertible function $H : \mathbf{R}^n \rightarrow \mathbf{R}^n$ (where **R** denotes the set of real numbers). The N real numbers $x_0, ..., x_{N-1}$ are transformed into the N real numbers $H_0, ..., H_{N-1}$ according to the formula

$$H_k = \sum_{n=0}^{N-1} x_n \left[\cos\left(\frac{2\pi}{N} nk\right) + \sin\left(\frac{2\pi}{N} nk\right)\right] \qquad k = 0, \ldots, N-1$$

where π is Pi. The combination

$$\cos(z) + \sin(z) = \sqrt{2}\cos(z - \pi/4)$$

is sometimes denoted cos(z), and should be contrasted with the $e^{-iz} = \cos(z) - i\sin(z)$ that appears in the DFT definition (where i is the imaginary unit).

C. Chebyshev filters

These are analog or digital filters [9] having a steeper roll-off and more pass band ripple (type I) or stop band ripple (type II) than Butterworth filters. Chebyshev filters have the property that they minimize the error between the idealized and the actual filter characteristic over the range of the filter, but with ripples in the pass band. This type of filter is named in honor of Pafnuty Chebyshev because their mathematical characteristics are derived from Chebyshev polynomials.

Because of the passband ripple inherent in Chebyshev filters, filters which have a smoother response in the passband but a more irregular response in the stopband are preferred for some applications.

1) Type I Chebyshev filter:

These are the most common Chebyshev filters. The gain (or amplitude) response as a function of angular frequency ω of the nth order low pass filter is

$$G_n(\omega) = |H_n(j\omega)| = \frac{1}{\sqrt{1 + \varepsilon^2 \, T_n^2(\omega/\omega_0)}}$$

where ε is the ripple factor, ω_0 is the cutoff frequency and $T_n()$ is a Chebyshev polynomial of the nth order.

The passband exhibits equiripple behavior, with the ripple determined by the ripple factor ε. In the passband, the Chebyshev polynomial alternates between 0 and 1 so the filter gain will alternate between maxima at $G = 1$ and minima at $G = 1/\sqrt{1+\varepsilon^2}$. At the cutoff frequency ω_0 the gain again has the value $1/\sqrt{1+\varepsilon^2}$ but continues to drop into the stop band as the frequency increases. This behavior is shown in the diagram. The common definition of the cutoff frequency to −3 dB does *not* hold for Chebyshev filters.

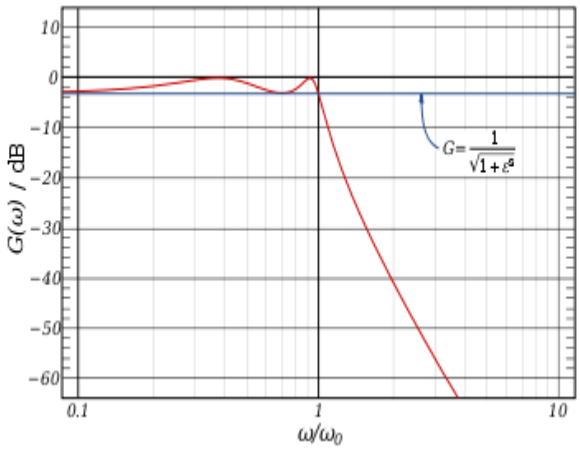
The order of a Chebyshev filter [10] is equal to the number of reactive components (for example, inductors) needed to realize the filter using analog electronics.

The ripple is often given in dB:

$$\text{Ripple in dB} = 20\log_{10}\frac{1}{\sqrt{1+\varepsilon^2}}$$

so that a ripple amplitude of 3 dB results from ε =1

An even steeper roll-off can be obtained if we allow for ripple in the stop band, by allowing zeroes on the $j\omega$-axis in the complex plane. This will however result in less suppression in the stop band. The result is called an elliptic filter, also known as Cauer filters.



The frequency response of a fourth-order type I Chebyshev low-pass filter with ε = 1.

2) Type II Chebyshev Filter:

This is also known as inverse Chebyshev, this type is less common because it does not roll off as fast as type I, and requires more components. It has no ripple in the passband, but does have equiripple in the stopband. The gain is:

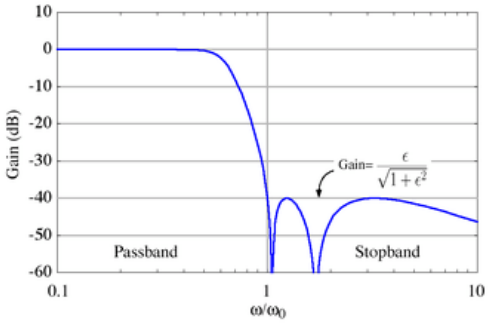
$$G_n(\omega, \omega_0) = \frac{1}{\sqrt{1 + \frac{1}{\varepsilon^2 T_n^2(\omega_0/\omega)}}}.$$

In the stop band, the Chebyshev polynomial will oscillate between 0 and 1 so that the gain will oscillate between zero and 1

$$\frac{}{\sqrt{1+1/\varepsilon^2}}$$

and the smallest frequency at which this maximum is attained will be the cutoff frequency ω_0. The parameter ε is thus related to the stopband attenuation γ in decibels by:

$$\varepsilon = \frac{1}{\sqrt{10^{0.1\,|\gamma|} - 1}}$$



The frequency response of a fifth-order type II Chebyshev low-pass filter with ε = 0.01

III. EXPERIMENTAL RESULTS

FHT with Chebyshev filter and FT tree classification

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	90	53.5 %
Incorrectly Classified Instances	78	46.4 %
Kappa statistic		0.0672
Mean absolute error		0.4713
Root mean squared error		0.6592
Relative absolute error		94.458 %
Root relative squared error		131.9763 %
Total Number of Instances	168	

=== Confusion Matrix ===

a b <-- classified as

39 41 | a = hand

37 51 | b = foot

IV. CONCLUSION

Based on the experimental results it is concluded that while classification is done on the EEG dataset, it is suggested that the classification filters such as Fast Hartley Transform (FHT) and Chebyshev filters may be used for the better classification which may make researchers do further analysis with utmost confidence.

REFERENCES

[1] R. N. Bracewell, "Discrete Hartley transform," *J. Opt. Soc. Am.* **73** (12), 1832–1835 (1983).

[2] R. N. Bracewell, "The fast Hartley transform," *Proc. IEEE* **72** (8), 1010–1018 (1984).

[3] R. N. Bracewell, *The Hartley Transform* (Oxford Univ. Press, New York, 1986).

[4] R. N. Bracewell, "Computing with the Hartley Transform," *Computers in Physics* **9** (4), 373–379 (1995).

[5] R. V. L. Hartley, "A more symmetrical Fourier analysis applied to transmission problems," *Proc. IRE* **30**, 144–150 (1942).

[6] H. V. Sorensen, D. L. Jones, C. S. Burrus, and M. T. Heideman, "On computing the discrete Hartley transform," *IEEE Trans. Acoust. Speech Sig. Processing* ASSP-33 (4), 1231–1238 (1985).

[7] H. V. Sorensen, D. L. Jones, M. T. Heideman, and C. S. Burrus, "Real-valued fast Fourier transform algorithms," *IEEE Trans. Acoust. Speech Sig. Processing* ASSP-35 (6), 849–863 (1987).

[8] Miodrag Popović and Dragutin Šević, "A new look at the comparison of the fast Hartley and Fourier transforms," *IEEE Trans. Signal Processing* 42 (8), 2178-2182 (1994).

[9] Pierre Duhamel and Martin Vetterli, "Improved Fourier and Hartley transform algorithms: application to cyclic convolution Of real data," *IEEE Trans. Acoust. Speech Sig. Processing* ASSP-35, 818–824 (1987)

[10] Daniels, Richard W. (1974). *Approximation Methods for Electronic Filter Design*. New York: McGraw-Hill. ISBN 0-07-015308-6.

[11] Williams, Arthur B.; Taylors, Fred J. (1988). *Electronic Filter Design Handbook*. New York: McGraw-Hill. ISBN 0-07-070434-1

[12] Wavelet Time-frequency Analysis of Electro-encephalogram (EEG) Processing. (2010). International Journal of Advanced Computer Science and Applications - IJACSA, 1(5), 1-5.

[13] Karyati, C. M., & Muslim, A. (2011). Effect of Thrombi on Blood Flow Velocity in Small Abdominal Aortic Aneurysms from MRI Examination. International Journal of Advanced Computer Science and Applications - IJACSA, 2(3), 13-18.

AUTHORS PROFILE

V.Baby Deepa, received her Bachelor's and Master's in Computer Science from Barathidasan University, Trichy and did her M.Phil. as well in the same university. She has 12 years of teaching experience and is Assistant professor in the faculty of Software Engineering, she is serving as the head for the same faculty in M.Kumarasamy College of Engineering, Karur. She has presented more than 15 papers on various topics including national, international conference and journals. She is a research scholar of Anna University Chennai and her research area is Fuzzy and Data Mining.

Dr.P.Thangaraj**,** did his graduation and post graduation in Mathematics at Madras University.He completed his M.Phil degree in the year 1993 from Bharathiar University. He completed his research work on Fuzzy Metric Spaces and awarded Ph.D degree by Bharathiar University.He completed the post graduation in Computer Applications at ICNOU in 2005. He completed his Master of Engineering degree in Computer Science in the year 2007 at Vinayaka Missions University. Currently he is a professor and Head of Computer Science and Engineering in Bannariamman College of Engineering and Technology. His current area of research interest is in Fuzzy Metric Spaces and Data Mining.

An Advanced Technology Selection Model using Neuro Fuzzy Algorithm for Electronic Toll Collection System

D.R.Kalbande
Asst. Professor, Computer Engineering
Research Scholar, University of Mumbai
k_dhananjay@yahoo.com

Priyank Singhal
Student, Computer Engineering
University of Mumbai, India
singhal.priyank@gmail.com

Nilesh Deotale
Lokmanya Tilak College of Engineering
University of Mumbai, India
mailtonilesh@indiatimes.com

Sumiran Shah
Student, Computer Engineering
University of Mumbai, India
sumiranshah@gmail.com

G.T.Thampi
Principal, TSEC
University of Mumbai
gtthampi@yahoo.com

***Abstract*— Selecting an optimum advanced technology system for an organization is one of the most crucial issues in any industry. Any technology system which makes business process more efficient and business management more simplified is one of the important Information System (IS) to the organization. The comprehensive framework is a three-phase approach which introduces two main ideas, one is the adopting of the McCall software quality model which is extracted from technology management essentials, and use the factors of McCall software quality model to be some of the technology selection criteria. Another major point is implementing and proposing a model based on this research using Neuro-Fuzzy algorithm to evaluate advance technology selection. This paper includes the concept of a new multi attribute selection process which combines both the Fuzzy logic (linguistic) and Neural network (integral valuation) methodology to evaluate or estimate a technology for Electronic Toll Collection System. Managers will be able to use this model for selecting a new technology in to their organization.**

Keywords- Neural Network; Fuzzy Logic; Technology Integration; Technology Evaluation.

I. INTRODUCTION

A. Electronic toll collection:

Electronic toll collection (ETC), an adaptation of military "identification friend or foe" technology, aims to eliminate the delay on toll roads by collecting tolls electronically. It is thus a technological implementation of a road pricing concept. It determines whether the cars passing are enrolled in the program, alerts enforcers for those that are not, and electronically debits the accounts of registered car owners without requiring them to stop [16 [17].

ETC is generally broken up into three pieces; automatic vehicle classification (AVC), automatic vehicle identification (AVI), and violation enforcement (VE).

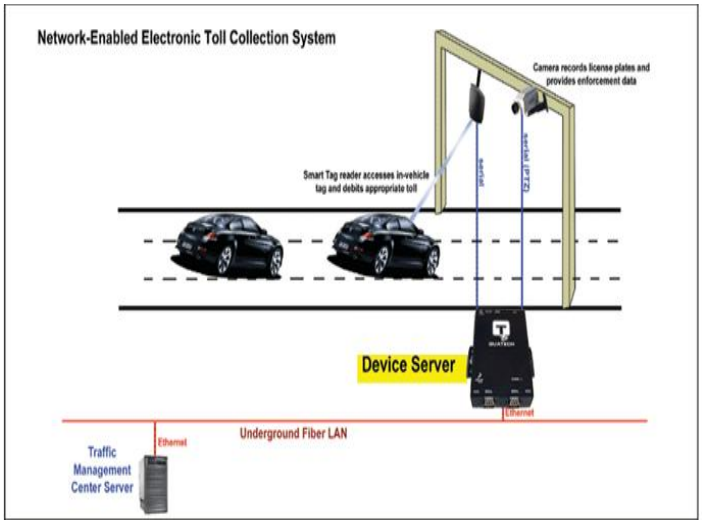


Fig1: Electronic toll collection

B. Automatic Vehicle Classification (AVC):

Classification of the vehicle is extremely important for those plazas where the fare is dependent on the number of axles on the vehicle, a very common case throughout the E-Z Pass ETC. Sensors, called treadles, are embedded in the roadway to count the axles and determine the tire width. Additional sensors similar to the motion sensors found on automated doors detect the presence of the vehicle and help distinguish and individualize the vehicles. Such sensors use the latest

technologies magnetic induction loops, treadles, and laser imaging [12].

C. Automatic Vehicle Identification (AVI):

The AVI component of the system consists of the RFID transponder such is located in the automobile and the equipment to communicate with the transponder located at the toll plaza and the License Plate Recognition (LPR) subsystem, a good primer of which can be found at License Plate Recognition - A Tutorial. While the toll plaza RFID transponder equipment is generally called a reader, in most ETC systems it can also write information to the vehicle transponder such as the time, date, location and vehicle class of the transaction [20].

D. Violation Enforcement (VE):

Violation enforcement consists of using the identification elements gathered from the AVC and AVI components along with additional information such as license plate and vehicle images to allow authorities to collect from and/or prosecute those who violate the electronic toll plaza. Typical ETC violations are [22]:

- Use of electronic toll collection lanes without a vehicle transponder,
- Insufficient funds in the associated account for identified transponders,
- Use of a transponder from a low-toll vehicle such as a car with two axles in a high-toll vehicle such as a tractor trailer.

List of Electronic Toll Collection Systems in India:

Name of Roadway	**Type of Roadway**	**Owned by**	**Operated by**	**Location**
NH-6 toll road	Highway	NHAI	TollTrax Toll Collection System	Kharagpur, India
Delhi Gurgaon Expressway	Highway	NHAI	Metro Electronic Toll Collection Systems	Delhi, India

Advantages of ETC [21]:

- Increases patron convenience and safety with nonstop payment
- Improves traffic flow
- Reduces patron commute time
- Reduces traffic congestion
- Lowers patron fuel use
- Reduces emissions which are a major cause of pollution
- Reduces need for new roads
- Reduces operating costs for toll authorities
- Provides proven reliability and unparalleled accuracy

II. NEURO-FUZZY MODEL

A. Fuzzy logic

A fuzzy logic model with its fundamental input-output relationship consists of four components namely; the fuzzifier, the inference engine, the defuzzifier, and a fuzzy rule base.

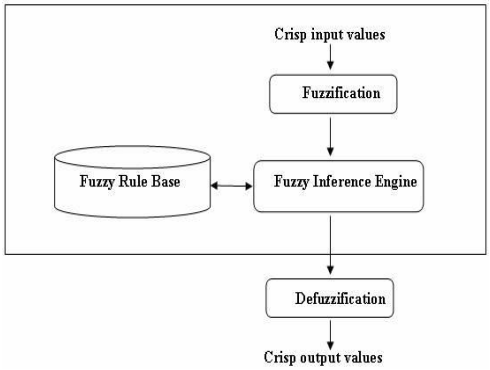


Fig 2: Structure of a Fuzzy Logic Model

In the fuzzifier, crisp inputs are fuzzified into linguistic values to be associated to the input linguistic variables. After fuzzification, the inference engine refers to the fuzzy rule base containing fuzzy IF-THEN rules to derive the linguistic values for the intermediate and output linguistic variables. Once the output linguistic values are available, the Defuzzifier produces the final crisp values from the output linguistic values [9] [10].

A basic FLC system is shown in above figure, which comprises four principal components:

- A fuzzy interface (FI), which is somewhat like an A/D converter in digital control.
- A decision-making logic (DML), which is like a digital controller.
- A defuzzification interface (DFI), which functions like digital theorems.
- A knowledge base (KB), which comprises knowledge of application domain and control goals to be met.

B. Neural Networks:

An artificial neural network (ANN), usually called "neural network" (NN), is a mathematical model or computational model that tries to simulate the structure and/or functional aspects of biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data [1].

To capture the essence of biological neural systems, an artificial neuron is defined as follows:

- It receives a number of inputs (either from original data, or from the output of other neurons in the neural network). Each input comes via a connection that has a strength (or *weight*); these weights correspond to synaptic efficacy in a biological neuron. Each neuron also has a single threshold value. The weighted sum of the inputs is formed, and the threshold subtracted, to compose the *activation* of the neuron (also known as the post-synaptic potential, or PSP, of the neuron).
- The activation signal is passed through an activation function (also known as a transfer function) to produce the output of the neuron.

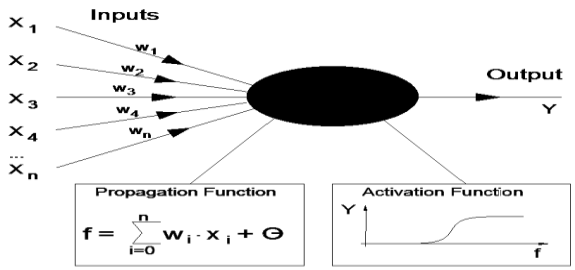


Fig 3: Structure of Neural model

If the step activation_function is used (i.e., the neuron's output is 0 if the input is less than zero, and 1 if the input is greater than or equal to 0) then the neuron acts just like the biological neuron described earlier (subtracting the threshold from the weighted sum and comparing with zero is equivalent to comparing the weighted sum to the threshold). Actually, the step function is rarely used in artificial neural networks, as will be discussed. Note also that weights can be negative, which implies that the synapse has an inhibitory rather than excitatory effect on the neuron: inhibitory neurons are found in the brain. The input, hidden and output neurons need to be connected together [2][3].

C. NEURO-FUZZY Model:

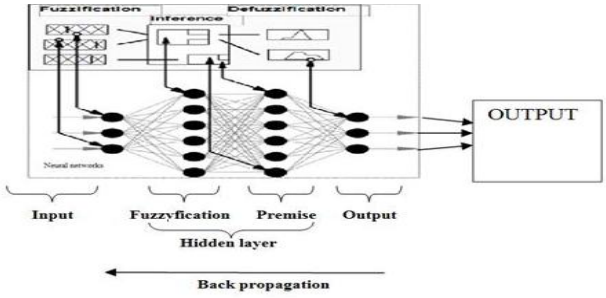


Fig 4: Structure of Neuro-Fuzzy model

The required data is collected through intensive survey in the destined sector. Interview sessions are conducted and depending on the characteristics of particular technology its attributes are identified [4].

We store the data collected in a database and as per the industry requirement [5]. The Fuzzy logic is then applied to the output of activation function which then generates linguistic weights using the predefined fuzzy sets.

It consists of 3 layers:

- Fuzzification
- Fuzzy rules (Inference engine)
- Defuzzification

In a fuzzification layer each neuron represents an input membership function of the antecedent of a fuzzy rule In a fuzzy inference layer fuzzy rules are fired and the value at the end of each rule represents the initial weight of the rule, and will be adjusted to its appropriate level at the end of training.

In the defuzzification layer each neuron represents a consequent proposition and its membership function can be implemented by combining one or two sigmoid functions and linear functions. The weight of each output link here represents the centre of gravity of each output membership function of the consequent.

The links between the premises and consequences of fuzzy rules are stored in these weights. The most common approach is to use so-called Fuzzy Associative Memories (FAMs). A FAM is a fuzzy logic rule with an associated weight. A mathematical framework exists, that maps FAMs to neurons in a neural net. The synapses of the neuron then are modeled as weights. The strength of the connection between an input and a neuron is noted by the value of the weight [11].

Negative weight values reflect inhibitory connections, while positive values designate excitatory connections [Haykin]. The next two components model the actual activity within the neuron cell. We now apply the input function to the neuron. An adder sums up all the inputs modified by their respective weights. This activity is referred to as linear combination. Finally, an activation (transfer) function controls the amplitude of the output of the neuron. An acceptable range of output is usually between 0 and 1, or -1 and 1.

We can also train our neural network by applying learning rule to the neurons which can be derived from training set to produce optimal output. After getting the corresponding output the adjustment is made in the connection weights and the membership functions in order to compensate the error and produce a new control signal if the output generated satisfies the target then it is fed to the fuzzifier. Else we detect the error present in one of the layers by checking any anomalies present in the weight matrix prepared and using back propagation we correct the error by adjusting the weights accordingly.

The output can then be analyzed for integrating a particular technology. In this model new attributes can be added for a new technology and thus widens our scope for technology adoption [6].

III. ADVANTAGES & DISADVANTAGES OF ETC TECHNOLOGIES STUDIED

1. Radio Frequency Identification(RFID)
2. Microwave CEN 278
3. ISO CALM Active Infrared
4. Global Positioning Systems(GPS)

A. RFID [15][18]:

Advantages of RFID:

1. RFID technology is used in a variety of industries for inventorying, making for efficient, cost-effective
2. RFID tags are very simple to install.
3. The RFID tags can store data up to 2 KB whereas other systems (like bar code) have the ability to read just 10-12 digits.
4. It improves security of the vehicle. RFID tags are placed inside the car and an alarm is installed at the doors.
5. The RFID tags are made rugged and robust so that they can be used in any harsh environments and temperatures.
6. The RFID tags are unique. This makes the system highly reliable and error-free. The RFID system reduces the labor cost of the company by providing a good tracking system.

Disadvantages of RFID:

1. The main disadvantage of RFID systems is high cost. The RFID system is costly as compared to other automatic identification systems. The cost can increase further, if the RFID system is designed for a specific application.
2. It is difficult for an RFID reader to read the information in case of RFID tags installed in liquids and metal products. The problem is that the liquid and metal surfaces tend to reflect the radio waves, which makes the tags unreadable. The tags have to be placed in various alignments and angles for taking proper reading. This is a tedious task when the work involves big firms.
3. Interference has been observed if devices such as forklifts and walkie-talkies are in the vicinity of the distribution centres. The presence of mobile phone towers has been found to interfere with RFID radio waves.
4. The USA and Europe, for instance, have different range of frequencies that allow RFID tags to function. This makes it mandatory for international shipping companies and other organizations to be aware of the working pattern of other nations also, which can be very time-consuming.
5. RFID technology has been referred to as invasive technology. Consumers are apprehensive about their privacy. The customer can be tracked and his personal information can be collected by the RFID reader.
6. There is no way in which damaged tags can be tracked and replaced by tags that are intact.
7. Although the tags do not require line-of-sight communication, they can be read within a specified range only.

B. CEN 278 MICROWAVE:

Advantages of CEN 278:

1. Speed up toll collection.
2. Smoother traffic flow.
3. Increasing the capacity of motorways.
4. Increased road safety
5. Enable police co-operation in the transport sector
6. Increase supply-chain efficiency.
7. Reduce congestion, delays and accidents

Disadvantages of CEN 278:

1. Rather expensive vehicle units
2. If a mandatory device is not possible for vehicles, an additional manual registration system is necessary. In such a dual system, however, the costs for operation and checks increase considerably
3. No standards
4. Additional operating costs from GSM communication
5. Units can only be installed in trained workshops

C. ISO CALM Active Infrared:

Advantages of ISO CALM [13][14]**:**

1. Data rate of 2Mbit/s
2. Vehicle speeds up to 200km/h.
3. V2I communication of 50 meters
4. Interference free multilane free flow communication
5. Lane specific communication zones with sharp Boundaries
6. Few milliseconds latencies and communication setup delays
7. Distance measurement with meter accuracy
8. License free wireless spectrum

Disadvantages of ISO CALM:

1. No support for Interoperability
2. Does not have wide vendor support as compared to other technologies as it is patented.
3. More expensive
4. Difficult to implement in Indian conditions

D. GLOBAL POSITIONING SYSTEMS (GPS)

Advantages of GPS [19]:

1. Road infrastructure, which can be expensive and often infeasible due to space constraints, is no longer needed.
2. It comprises greater flexibility in defining or changing payment systems.
3. Toll areas can be changed easily by redefining "virtual" toll areas.
4. A GPS system streamlines supply chains and truck movements. The system can track goods at any point of time and accurately predict when goods will reach their destination.

5. GPS systems are used to detect structural problems in buildings and roads and to predict disasters like earthquakes and so on. The scientific applications of a GPS system are many.

Disadvantages of GPS:

1. It is expensive.
2. People pay more attention on tracking rather than road. Hence there may be more risk of an accident.
3. It requires more power and needs batteries (handheld ones).

It needs good care and handling. The maintenance parameter value is high.

IV. RESEARCH FINDING

A. Existing System

At present, there exists no system which provides any automation in neuro-fuzzy prediction field [7]. In India, the prediction techniques used are raw and not much research has been done in the deciding advanced technology selection attributes. The technology selection done through other techniques are not only tedious but also less efficient. The following problem areas remain uncovered in such an approach [8]:

1) Inefficient allocation

There arises a problem of deciding proper selection attributes for a particular technology and allocating importance to them (attributes) owing to the existence of ample parameters. Also, new technologies are regularly being made available in the market which complicates the decision-making process (New Parameter may get added).

2) No Database approach

There is no such database oriented approach present for technology selection and in the present scenario selection process is handled by top management level as per their own will with no any such survey or research being done for effective and selection of optimal technologies as per the organization's requirements.

3) Constraint problem

During selection of technologies manually, all the constraints are hard to be followed and maintained. Technology-selection itself being a constraint, the problem has to be dealt with efficiently and technically.

B. Salient Features with Constraints

Our research aims at overcoming the above limitations and including several new enhancements.

Some of the features incorporated in the research are,

1) For every selection criteria the following rules are applied:
   - The characteristics of technology to be evaluated must be identified.
   - Maximum number of attributes (12-15) to be decided and evaluation be done on that basis.
   - Surveying of the industries in particular sector.
   - The input (neuron-weights) and output (linguistic) format be decided and the design of the neuro-fuzzy model be made as per the available thesis.
   - Identifying fuzzy rules
   - Error detection and correction be done using Back-propagation algorithm.
   - Training the neural network by creating a training set.
2) The technology selected has following characteristics.

Technology-Related	User-Related	Vendor-Related
Flexibility	Customization	After Sales Support & Training
Implementability	Reporting & Analysis Features	Maintenance
System Requirements	Integration with Other Applications	Cost
Real-time Changes		Vendor Credentials
Back-up System		Financing Options

Fig 9: Technology Selection attributes

3) The user can select certain technology he wants to decide on from different options available in the database. The system will compare all the attributes and will provide him with the optimal solution in the user-friendly linguistic form.
4) The model being web-based, can also entertain queries being submitted. Thus many organizations all around the world can benefit from this.
5) Email will then be sent to the respected party after the query gets analysed.

C. GOALS OF THE DECISION MAKING MODEL

The neuro-fuzzy model under development is meant to satisfy following goals:

- Generate a semi-automatic model that can satisfy all the constraints we have specified.
- Flexibility in the decision-making process with quick and accurate prediction of feasible technology options available.

The model will be made available online (administrator will have right to access) to identify proper fuzzy rules and using neural networks to tune the membership functions.

V. PROPOSED METHODOLOGY

A. Mathematical model:

Mathematical description:

Fuzzy sets:

Step 1: Let us consider no. of ETC technologies T1, T2, T3, T4 etc

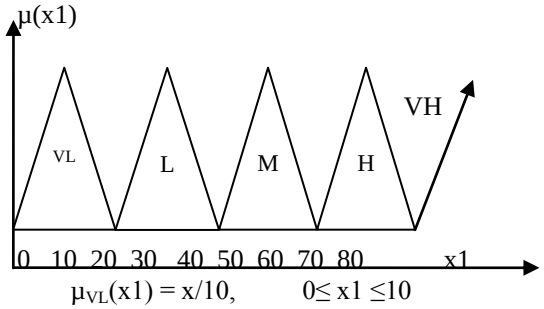
Let us consider criteria for the selection of technologies as given below

1. Co=cost
2. Vs = Vendor support
3. In= Interoperability
4. Pc=Power consumption
5. Rn=Range

Let us consider parameters for the every criterion are:

VL -very low, L - low
M – medium, H- high
VH – very high

STEP 2:
Let us consider cost (x1) = {VL, L, M, H, VH}



$\mu_{VL}(x1) = x/10, \quad 0 \le x1 \le 10$

$(20-x1)/10, \quad 10 \le x1 \le 20$

$\mu_L(x1) = (x1-20)/10, \quad 20 \le x1 \le 30$

$(40-x1)/10, \quad 30 \le x1 \le 40$

$\mu_M(x1) = (x1-40)/10, \quad 40 \le x1 \le 50$

$(60-x1)/10, \quad 50 \le x1 \le 60$

$\mu_H(x1) = (x1-60)/10, \quad 60 \le x1 \le 70$

$(80-x1)/10, \quad 70 \le x1 \le 80$

$\mu_{VH}(x1) = (x1-80)/20, \quad 80 \le x1$

Similar equations for x2, x3, x4, x5, x6 and x7.

TECHNOLOGY → / OBJECTIVES ↓	T1	T2	T3	T4
CO	L	M	H	M
In	M	VH	H	M
Vs	M	M	H	H
Pc	L	M	VH	VH
Rn	M	H	VH	H

STEP 4: DEFUZZIFICATION

Variables	Criteria	Value
X1	Co	50
X2	Vs	45
X3	In	90
X4	Pc	75
X5	Rn	70

$\mu_M(x1) = 1 \quad 40 \le x1 \le 50$

$1 \quad 50 \le x1 \le 60$

$\mu_M(x2) = ½ \quad 40 \le x2 \le 50$

$3/2 \quad 50 \le x2 \le 60$

$\mu_{VH}(x3) = ½ \quad 80 \le x3$

$\mu_H(x4) = 3/2 \quad 60 \le x4 \le 70$

$½ \quad 70 \le x4 \le 80$

$\mu_H(x5) = 1 \quad 60 \le x5 \le 70$

$1 \quad 70 \le x5 \le 80$

Now we will apply weights to these defuzzified values and send it as an input to neural network for calculating the RMSE Error using Back propagation Algorithm.

Step 4:

B. Back propagation Algorithm:

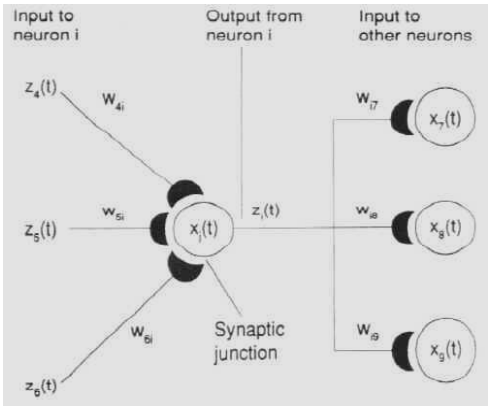


Fig 10: Back propagation Algorithm

A unit in the output layer determines its activity by following a two step procedure.

- First, it computes the total weighted input x_j, using the formula:

$$X_j = \sum_i y_i W_{ij}$$

where y_i is the activity level of the j^{th} unit in the previous layer and W_{ij} is the weight of the connection between the i^{th} and the j^{th} unit.

- Next, the unit calculates the activity yj using some function of the total weighted input. Typically we use the sigmoid function:

$$y_j = \frac{1}{1+e^{-x_j}}$$

- Once the activities of all output units have been determined, the network computes the error E, which is defined by the expression:

$$E = \frac{1}{2}\sum_i \left(y_i - d_i\right)^2$$

where y_j is the activity level of the j^{th} unit in the top layer and d_j is the desired output of the j^{th} unit.

The back-propagation algorithm consists of four steps:

1. Compute how fast the error changes as the activity of an output unit is changed. This error derivative (EA) is the difference between the actual and the desired activity.

$$EA_j = \frac{\partial E}{\partial y_j} = y_j - d_j$$

2. Compute how fast the error changes as the total input received by an output unit is changed. This quantity (EI) is the answer from step 1 multiplied by the rate at which the output of a unit changes as its total input is changed.

$$EI_j = \frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} \times \frac{dy_j}{dx_j} = EA_j y_j \left(1 - y_j\right)$$

3. Compute how fast the error changes as a weight on the connection into an output unit is changed. This quantity (EW) is the answer from step 2 multiplied by the activity level of the unit from which the connection emanates.

$$EW_{ij} = \frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial W_{ij}} = EI_j y_i$$

4. Compute how fast the error changes as the activity of a unit in the previous layer is changed. This crucial step allows back propagation to be applied to multilayer networks. When the activity of a unit in the previous layer changes, it affects the activities of all the output units to which it is connected. So to compute the overall effect on the error, we add together all these separate effects on output units. But each effect is simple to calculate. It is the answer in step 2 multiplied by the weight on the connection to that output unit.

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial y_i} = \sum_j EI_j W_{ij}$$

By using steps 2 and 4, we can convert the EAs of one layer of units into EAs for the previous layer. This procedure can be repeated to get the EAs for as many previous layers as desired. Once we know the EA of a unit, we can use steps 2 and 3 to compute the EWs on its incoming connections.

Step 5:
The technology comprising of minimum error as compared to the target value will be selected and will be chosen as the best optimum technology.

VI. CONCLUSION

The research demonstrates the use of dynamic neuro-fuzzy model which enhances the prediction capability of the model and hence gives accurate estimation for adoption and selection of new technology. We combine the strengths of both neural networks and fuzzy logic through our model. The process is relatively simple, supports creation of high level pedagogical strategies and can be easily adapted to individual technological preferences. Compared to neural networks, the neuro fuzzy methods provide models which can be interpreted by human beings. The model is in the form of familiar if-then rules implying easy selection with the operators' (expert) rules.

Technology selection is a complex process, and requires the understanding of the stages prior to (e.g. technology scanning) and stages after (e.g. technology implementation) technology selection decisions. It is a multi-dimensional process requiring the analysis of a wide range of internal and external factors. Most technology selection decisions in industry have been largely limited to the analysis of the financial or economic factors. The research extends this approach to cover a wide range of criteria that may affect technology selection decisions. It is expected that the ongoing case studies will result in a comprehensive list of key criteria that can be applied to any technology selection decision.

We observed that there was as error of 1% between the output produced by actual and the empirical model. Hence with the help of more exhaustive surveys and acquisition of more accurate values we can improve the optimality of the model even further.

REFERENCES

[1] Neural Networks and Fuzzy Systems http://fuzzy.cs.uni-magdeburg.de/papers.html

[2] SpringerLink - Neural Processing Letters, Volume 4, Number 2 http://www.springerlink.com/content/p2u0465437514453/

[3] Artificial Neural Networks http://www.learnartificialneuralnetworks.com/

[4] Abraham Ajith, 2001. Neuro Fuzzy Systems: State-of-the-art Modelling Techniques, School Of Computing & Information Technology Press, Australia, 2001

[5] Buckley J.J. and Hayashi Y., 1994. Fuzzy neural networks: A survey, Fuzzy Sets and Systems (1994) 1-13.

[6] Studies in Fuzziness and Soft Computing (Publisher: Springer Berlin / Heidelberg)

[7] Review of Business Research Papers Vol.2. No.4. December 2006, Pp. 39-50 eBusiness Process-Personalization using Neuro-Fuzzy Adaptive Control for Interactive Systems by Zunaira Munir, Nie Gui Hua, Adeel Talib and Mudassir Ilyas.

[8] The Use of Artificial Neural Networks for Technology Selection in the Presence of Both Continuous and Categorical Data Reza Farzipoor Saen Department of Industrial Management, Faculty of Management and Accounting, Islamic Azad University-Karaj Branch, P.O. Box: 31485-313, Karaj, Iran.

[9] Using Fuzzy Neural Networks and Analytic Hierarchy Process for Supplier Classification in e-Procurement by Arpan Kumar Kar XLRI School of Business & Human Resources, India.

[10] Report on Quantitative measurement of advanced manufacturing technology transfer from foreign-based companies to local companies, from the graduate school of computer and engineering management, Thailand.

[11] 2007 Report on scope and objectives of EFC Standards, from NNI. Netherlands

[12] Road Transport and Traffic Telemetic www.cen.eu/cen/Sectors/Sectors/.../Automobile/.../n1910PRkitforWG1.pdf

[13] Information about ISO CALM www.efkon.com/index.pl/iso-calm

[14] CALM IR Coopers; www.coopers-ip.eu/fileadmin/results/ITSWORLD01.pdf

[15] RFID Journal www.rfidjournal.com

[16] Institute of Transportation Studies www.calccit.org/.../Electronic_toll_collection/electron_toll_collection

[17] SANS Institute Reading Room
http://www.sans.org/reading_room/whitepapers/threats/electronic-toll-collection_1424

[18] RFID in Pervasive Computing
http://www.perada.eu/documents/articles-perspectives/an-introduction-to-rfid-technology.pdf

[19] 2007 Report on 'A methodology for selection and evaluation of advanced manufacturing technologies', ICIM Journal.
Information about GPS

[20] Improving Metropolitan Transportation Efficiency with FAST Miles http://www.calccit.org/itsdecision/serv_and_tech/Electronic_toll_collection/electron_toll_collection_report.html

[21] Ahmad, Y., & Husain, S. (2010). Applying Intuitionistic Fuzzy Approach to Reduce Search Domain in an Accidental Case. International Journal of Advanced Computer Science and Applications - IJACSA, 1(4).

[22] Trivedi, J. A. (2011). Framework for Automatic Development of Type 2 Fuzzy , Neuro and Neuro-Fuzzy Systems. International Journal of Advanced Computer Science and Applications - IJACSA, 2(1), 131-137.

[23] Rahman, F., & Kumar, C. (2010). Loss Reduction in Distribution System Using Fuzzy Techniques. International Journal of Advanced Computer Science and Applications - IJACSA, 1(3), 15-19.

[24] Electronic Toll Collection System for Heavy Goods
http://www.roadtraffic-technology.com/projects/lkw-maut/

[25] Electronic Toll Collection
http://www.thalesgroup.com/Pages/Solution.aspx?id=9364

Energy-Efficient, Noise-Tolerant CMOS Domino VLSI Circuits in VDSM Technology

Salendra.Govindarajulu[1], Dr.T.Jayachandra Prasad[2], C.Sreelakshmi[3], Chandrakala[4], U.Thirumalesh[5]
[1]Associate Professor, ECE, RGMCET, Nandyal, JNTU, A.P.State, India
[2]Principal, RGMCET, Nandyal, JNTU, A.P.State, India
[3,4,5]PG Student, ECE, RGMCET, Nandyal, JNTU, A.P.State, India

***Abstract*— Compared to static CMOS logic, dynamic logic offers good performance. Wide fan-in dynamic logic such as domino is often used in performance critical paths, to achieve high speeds where static CMOS fails to meet performance objectives. However, domino gates typically consume higher dynamic switching and leakage power and display weaker noise immunity as compared to static CMOS gates. Keeping in view of the above stated problems in previous existing designs, novel energy-efficient domino circuit techniques are proposed. The proposed circuit techniques reduced the dynamic switching power consumption; short-circuit current overhead, idle mode leakage power consumption and enhanced evaluation speed and noise immunity in domino logic circuits. Also regarding performance, these techniques minimize the power-delay product (PDP) as compared to the standard full-swing circuits in deep sub micron CMOS technology.**

Also the noise immunity of the CMOS Domino circuits with various techniques and keepers are analyzed. Various noise sources are considered and noise immune domino logic is proposed.

Keywords- Dynamic; Domino; Noise Margin; Very Deep submicron technology; High speed; Power consumption; Power delay product (PDP).

I. INTRODUCTION

Dynamic domino logic circuits are widely used in modern digital VLSI circuits. These dynamic circuits are often favoured in high performance designs because of the speed advantage offered over static CMOS logic circuits. The main drawbacks of dynamic logic are a lack of design automation, a decreased tolerance to noise and increased power dissipation. However, domino gates typically consume higher dynamic switching and leakage power and display weaker noise immunity as compared to static CMOS logic circuits. In this paper novel energy-efficient domino circuit techniques are proposed.

This paper is organized as follows. In section II, Dual-rail domino circuit with self-timed precharge scheme is proposed.

The pseudo-footless dynamic circuit technique is presented in section III. Section IV describes performance evaluation results of energy-efficient dual-Vt domino logic. Section V describes the Noise immune domino logic. Then conclusions are presented in section VI.

II. DUAL-RAIL DOMINO FOOTLESS CIRCUIT WITH SELF-TIMED PRECHARGE SCHEME (DRDFSTP):

Conventional domino circuits:

In this section, several conventional domino circuits with their own clocking schemes are briefly reviewed.

A. Dynamic DCVSL Footed Circuit (DDCVSLF):

Fig.1 shows AND/NAND dynamic DCVSL Footed circuit. One of the disadvantages of this kind of domino circuit is that the existence foot transistor slows the gates somewhat, as it presents an extra series resistance. Moreover, simultaneous precharge may cause an unacceptable IR-drop noise.

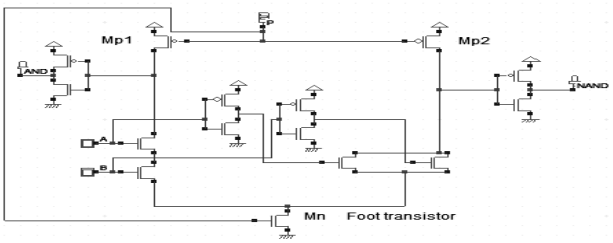


Fig.1. Dynamic DCVSL AND/NAND Footed gate

B. Dynamic DCVSL Footless Circuit (DDCVSLFL):

Fig.2 shows AND/NAND dynamic DCVSL Footless circuit. Two benefits come from the usage of footless domino gates: improved pull-down speed and reduced precharge signal load. Main disadvantage is simultaneous precharge will cause short-circuit current.

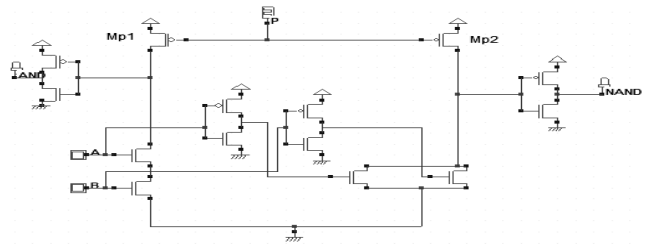


Fig. 2. Dynamic DCVSL AND/NAND Footless gate

C. Delayed-Reset Domino Circuit (DRDC):

Fig.3 illustrates the delayed-reset domino AND/NAND circuit [3]. However, the use of delay elements, together with the need of both footed and footless cell libraries tends to increase design complexity.

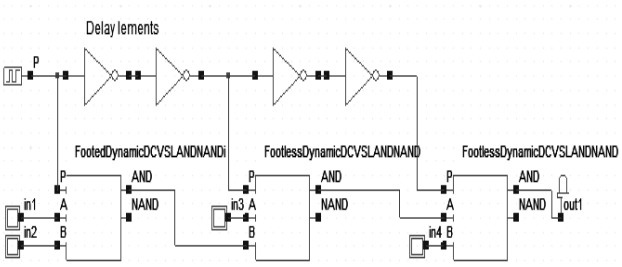


Fig:3. The delayed-reset domino AND/NAND circuit

D. Dual-Rail Data-Driven Dynamic Logic (D^4L):

D^4L circuit uses input signals instead of precharge signal for correct precharge and evaluation sequencing [5]. Correspondingly, clock-buffering and clock-distribution problems can be eliminated. Furthermore, the foot transistor can be eliminated without causing a short-circuit problem. A D^4L two-input AND/NAND gate is shown in Fig.4.

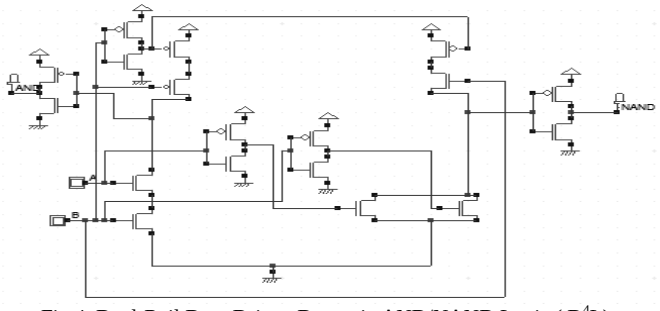


Fig.4. Dual-Rail Data-Driven Dynamic AND/NAND Logic (D^4L)

Dual-Rail Domino Footless Circuit with Self- Timed Precharge Scheme (DRDFSTP):

The presence of the foot transistor in the conventional dynamic DCVSL circuit shows the gate somewhat, as it presents an extra series resistance. To safely remove the transistor, two constraints must be met: (1) gate changes to evaluation phase before valid input come; (2) gate changes to precharge phase only after inputs change to zero. We propose a footless duail-rail domino circuit with self-timed precharge scheme to realize a high performance footless domino circuit while meeting the constraints mentioned above. It is expected that the peak of precharge current could be reduced due to the self-timed precharge scheme. Fig. 9 shows the AND/NAND gate of the proposed footless dual-rail domino circuit with self-timed precharge scheme. The self-timed precharge control logic consists of static CMOS inverter whose source of NMOS transistors are tied to input signals, which generate sub-precharge signals (PC1-PC4) from precharge signal P in cases of the corresponding input signals are zero. The PMOS precharge tree above the pull down network (PDN) is used for precharging the corresponding gate.

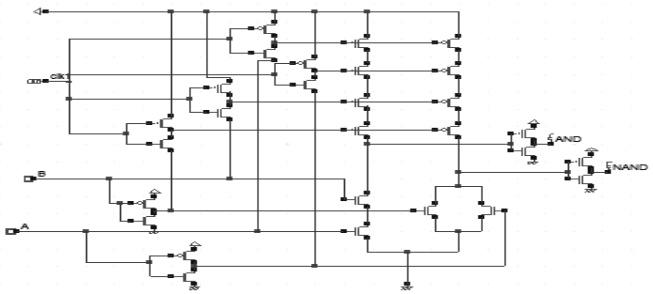


Fig:5. Dual-rail footless domino AND/NAND gate with self-timed precharge scheme.

Simulation results:

In this work, we have implemented a Dynamic DCVSL circuit, Dual-Rail Data-Driven Dynamic Logic and a proposed circuit Dual-Rail Domino Footless Circuit with Self-Timed Precharge Scheme. The results of simulation are shown in the below TABLES1-3.

Table1. AND/NAND GATE

Technique	Power (µw)	CriticalDelay (ns)	PDP (10^{-15} w-s)	Area (µ.sqm)
DDCVSLF	7.6	0.088	0.6688	69.62
DDCVSLFL	152	0.025	3.8	65.41
DRDC	205	0.137	28.085	252.9
D^4L	72.555	0.111	8.053606	93.3
DRDFSTP	7.676	0.042	0.322392	177.6

Table2. OR/NOR GATE

Technique	Power (µw)	CriticalDelay (ns)	PDP (10^{-15} w-s)	Area (µ.sqm)
DDCVSLF	7.58	0.087	0.65946	74.82
DDCVSLFL	145	0.090	13.05	66.59
DRDC	220	0.403	88.66	290
D^4L	10.163	0.112	1.138256	78.48
DRDFSTP	7.583	0.042	0.318486	30.18

Table3. XOR/XNOR GATE

Technique	Power (µw)	Critical Delay (ns)	PDP (10^{-15} w-s)	Area (µ.sqm)
DDCVSLF	11.7	0.032	0.3744	99.2
DDCVSLFL	99.023	0.032	3.1687	92.17
DRDC	231	0.091	21.021	391.9
D^4L	16.802	0.029	0.487258	100.5
DRDFSTP	11.642	0.04	0.46568	200.13

III. PSEUDO FOOTLESS DOMINO CIRCUIT (PF-DOMINO)

Footed domino circuit with a global clock:(FD)

Fig. 6 shows the most conventional domino circuit, which comprises of footed domino gates driven by a common clock

buffer. One of the disadvantages of this kind of domino circuit is that it should be constructed with only true-logic gates. Moreover, simultaneous precharge may cause an unacceptable IR-drop noise.

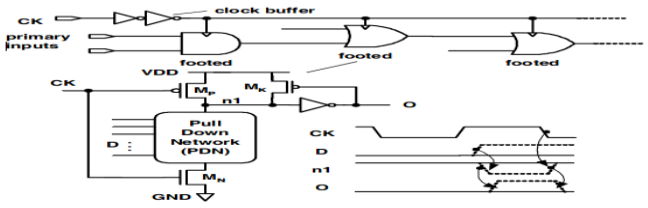


Fig. 1 The footed domino gate

Fig.6. The footed domino gate

Footless domino circuit with delayed clocks:(DR-domino)

Fig.7 illustrates the delayed-reset domino circuit (DR-domino). The DR-domino circuit does not improve the logic construction flexibility because it still accepts true logic gates only.

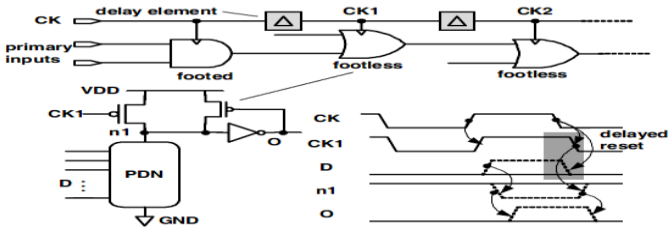


Fig. 2 The DR-domino circuit

Fig.7. The DR-domino circuit

Footed domino circuit with delayed clocks:

In order to improve the logic construction flexibility, the Clock-Delayed domino (CD-domino) circuit, shown in Fig. 8, is proposed to allow the usage of both positive and negative logic gates within a block. To achieve this flexibility, the clock rising edge of a gate should be delayed until all the incoming data settle. However, the delayed evaluation and the footed gates degrade the performance of the whole circuit seriously.

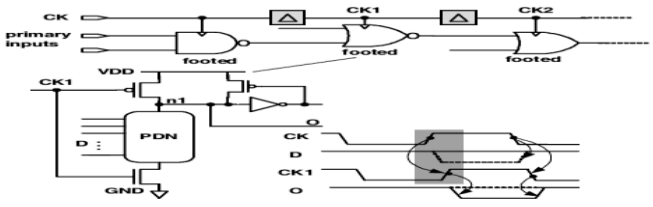


Fig. 3 The CD-domino circuit

Fig: 8. The CD-domino circuit

In this work, we start from adopting an improved delayed-evaluation clocking style to preserve the logic construction flexibility, but add new circuit techniques to remove the other origin of speed limitation, i.e. the usage of footed gates.

Pseudo footless domino circuit :(PF-domino):

The pseudo footless domino circuit (PF-domino) is shown in Fig. 9. Basically, the circuit structure of the PF-domino is exactly the same with that of the CD-domino circuit. The differences lie in two aspects. First, all the logic gates used are pseudo-footless (PF) dynamic gates (as the inserted gate shows), rather than footed gates. Second, an enhanced self-timed delayed-evaluation clocking scheme is used to replace the simple clock-delayed scheme used in the CD-domino circuit. These two techniques are introduced in the following step by step.

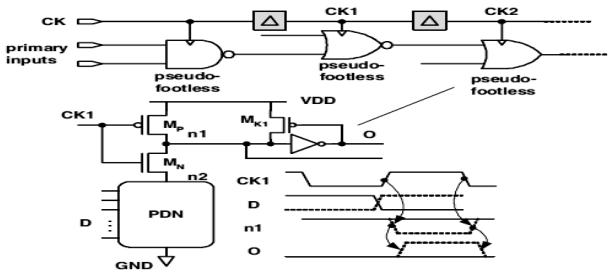


Fig. 4 The PF-domino circuit with the primitive PF gate

Fig.9. The PF- domino circuit with primitive PF gate

The pseudo-footless dynamic gates:

The pseudo-footless dynamic circuit technique was first proposed. The PF gate inserted in Fig. 9 is the primitive version used , which is quite similar to a typical footed domino gate except that M_N is pulled up beneath M_P. The preferred PDN function is NOR. Such an arrangement is beneficial for both speed and power. First, for the dynamic part, only a small output node is precharged, and then the discharged charge, if necessary, is much smaller than that of a conventional footed gate. Second, we require that all the data inputs be ready before the clock rises up. Then, before the evaluation phase, most charges in the PDN have been discharged, which results in a very high-speed discharge in the evaluation phase. This mechanism is also the name "pseudo-footless" comes from.

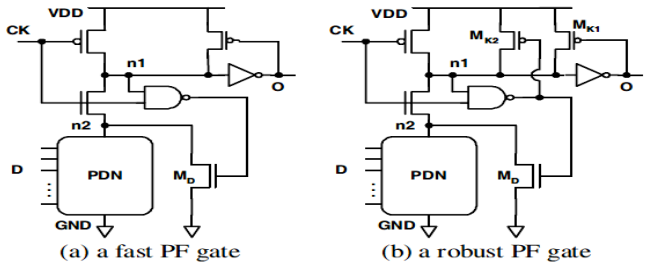


Fig. 5 Derivatives of the primitive PF gate

Fig.10. Derivatives of the primitive PF gate

When used in a general domino environment, the PDN may realize a complicated large-fan-in function. The increased capacitance at node *n2* will slow down the discharge. The circuit shown in Fig. 10(a) is proposed for speeding in such a condition. The transistor M_D is added in parallel with the PDN and is activated in the *precharge phase* to deplete the charge at *n2* in advance. During evaluation, M_D is initially disabled because *n1* is high. If *n1* is being pulled down, M_D will be turned on to help discharge. This gate is called a fast PF gate. When the capacitance of *n2* is much larger than that of *n1*, we need to consider the problem of charge sharing. In this case, we can use the gate shown in Fig. 10(b), a robust PF gate, where a second keeper M_{K2} is added to replenish the charge to *n1* when it is subject to a voltage fluctuation due to a charge sharing condition. The output loading and the fin-in number are the dominant factors that determine the performance of PF gates.

Hence, we need to find out which type of the PF gate is the best choice for each loading and fan-in combination. First, different PF gates with different fan-in numbers are designed and characterized for various loading conditions. And second, the fastest circuit without the charge sharing effect is considered to be the best choice.

The enhanced self-timed delayed-evaluation:

The delay element is the key component for the speed, as explained in the following. If a gate receives all non-inverted inputs, the arrival time of the clock rising edge will not cause malfunction. In this case, the clock signal is usually designed to arrive ahead of the data inputs so that a higher speed can be obtained. For a gate with at least one pull-down path controlled by inverted inputs, the clock signal should be delayed until all the data inputs settle to avoid an unrecoverable error. An enough margin of this delay must be kept to face the PVT variations. In the CD-domino circuit, a simple buffer-type delay element is mentioned, which asks for a quite large margin of the delay and causes remarkable performance degradation. We propose to use a more robust self-tracking.

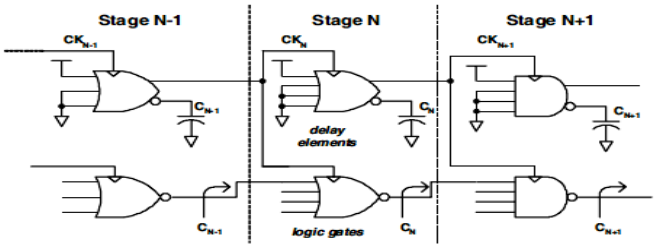


Fig. 6 The proposed robust self-tracking scheme

Fig. 11. The proposed robust self-tracking scheme

Simulation results:

Using the above techniques OR2 gate, AND2 gate, XOR2 gate are implemented. These design styles are compared by performing detailed transistor-level simulations on benchmark circuits using DSCH3 and Microwind3 CAD tool for 65 nm technology.

Table4.AND2 Gate

Technique	Power(µw)	Delay(ns)	PDP(*10-15)	Area(sq.µm)	Noise Immunity(mv)
FD	10.045	0.045	0.452025	29.64	10
DLRFLD	201.00	0.045	9.045	94.82	40
DLRFD	10.047	0.045	0.452115	28.54	30
PSFLD	10.006	0.050	0.5003	26.31	30
Fast PSFLD	160.00	0.064	10.24	58.08	40
Robust PSFLD	159.00	0.100	15.9	57.51	60

Table5.OR2 Gate

Technique	Power(µw)	Delay(ns)	PDP(*10-15)	Area(sq.µm)	Noise Immunity(mv)
FD	5.7760	0.044	0.254144	31.92	50
DLRFLD	94.589	0.044	4.161916	21.43	70
DLRFD	5.7660	0.044	0.253704	28.54	100
PSFLD	5.5600	0.047	0.261320	25.94	100
Fast PSFLD	110.00	0.066	7.260000	47.35	120
Robust PSFLD	111.00	0.102	11.32200	53.51	120

Table6. XOR2 Gate

Technique	Power(µw)	Delay(ns)	PDP(*10-15)	Area(sq.µm)	Noise Immunity(mv)
FD	33.709	0.082	2.764138	83.26	10
DLRFLD	63.697	0.049	3.121153	62.97	10
DLRFD	18.078	0.049	0.885822	56.17	10
PSFLD	1.4290	0.052	74.308	53.84	10
Fast PSFLD	166.00	0.066	10.956	83.42	10
Robust PSFLD	167.00	0.073	12.191	89.43	90

IV. ENERGY-EFFICIENT DUAL-V_T DOMINO LOGIC

A. *Standard single threshold (low-V_t) voltage*

In this, all standard low-threshold voltage transistors (V_t = 0.4 volts) are used in implementing the bench mark circuits and are simulated using DSCH and Microwind 3.1.

B. *Standard single threshold (high-V_t) voltage*

In this, all standard high-threshold voltage transistors (V_t = 0.7 volts) are used in implementing the bench mark circuits .

C. *Standard dual threshold voltage*

This Dual Threshold CMOS (DTCMOS) design technique uses fast low threshold voltage (LTV) and slow high threshold voltage (HTV) devices. Thus, the aim of DTCMOS is to maximize the gain in leakage at the HTV devices without worsening the performance of the circuit. In this, the PMOS and NMOS transistors in the output inverter are used with high V_t and remaining are used with low V_t devices.

D. *Modified dual-V_t technology*

This technology is the proposed technology, which is a modification of standard dual-threshold technology. In standard dual-V_t technology, the transistors of the output inverter circuit in CMOS domino logic are introduced with high-V_t transistors. In this modified dual-V_t technology, only the pull-down transistor is introduced with the standard high-V_t transistor and

the pull-up transistor is introduced with standard low-V_t transistor.

Simulation results:

In this work, we implemented benchmark circuits using the above four technologies. The figure of merit used to compare these technologies is Power-Delay Product (PDP). The benchmark circuits implemented in this work are and2, or2, or8, or16, xor2, 16-bit adder, 16-bit comparator, D-Latch, 4-bit LFSR which are given below from Table1-9. The OR2 gate is illustrated for the proposed technologies which are given below in Figures 12,13, 14, 15 .

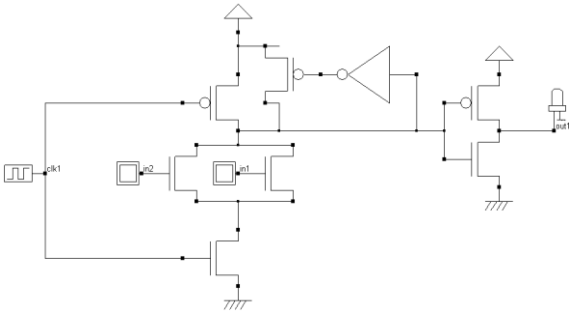


Fig.12. OR2 Standard Low-V_t

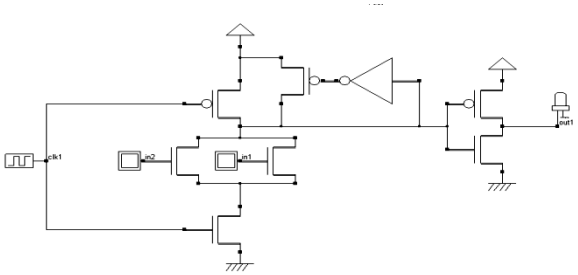


Fig.13. OR2 Standard High-V_t

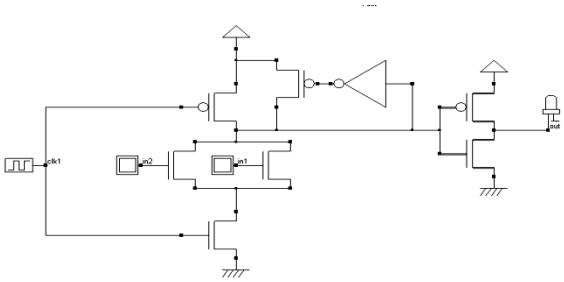


Fig.14. OR2 Standard Dual-Vt

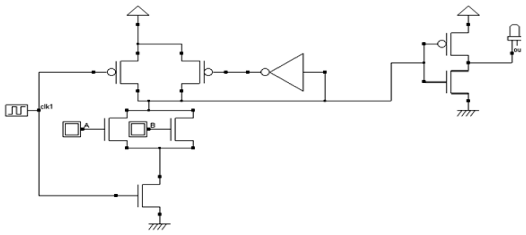


Fig.15. OR2 Modified Dual-Vt

Table7.16 Bit Adder

Technique	Power(mw)	Delay(ns)	PDP (10^-12 w-s)	Area (µ.sqm)
Standard low V_t	6.359	55.416	352.39	10841.32
Standard high V_t	6.270	54.943	344.49	12518.60
Standard Dual V_t	6.262	56.219	352.04	11294.89
Modified Dual V_t	6.270	60.967	382.26	12950.92

Table8.16 Bit Comparator

Technique	Power(mw)	Delay(ns)	PDP (10^-12 w-s)	Area (µ.sqm)
Standard low V_t	6.648	74.557	495.654	18294.44
Standard high V_t	6.619	71.155	470.974	19835.87
Standard Dual V_t	6.637	67.751	449.663	18505.20
Modified Dual V_t	6.634	80.915	536.790	19306.26

Table9.D Latch

Technique	Power(mw)	Delay(ns)	PDP (10^-12 w-s)	Area (µ.sqm)
Standard low V_t	0.189	0.307	0.058	259.86
Standard high V_t	0.221	0.389	0.085	273.16
Standard Dual V_t	0.223	0.429	0.095	291.04
Modified Dual V_t	0.221	0.352	0.077	259.86

Table10.4 Bit LFSR

Technique	Power(mw)	Delay(ns)	PDP (10^-12 w-s)	Area (µ.sqm)
Standard low V_t	3.748	3.192	11.963	2682.39
Standard high V_t	4.037	3.583	14.464	2861.61
Standard Dual V_t	4.033	3.735	15.063	2733.37
Modified Dual V_t	4.008	3.532	14.156	2795.82

Table11.OR8 gate

Technique	Power(µw)	Delay(ns)	PDP (10^-15 w-s)	Area (µ.sqm)
Standard low V_t	0.892	0.088	0.078	80.25
Standard high V_t	0.990	0.143	0.141	82.16
Standard Dual V_t	0.789	0.130	0.102	81.06
Modified Dual V_t	0.845	0.095	0.080	81.46

Table12.OR2 gate

Technique	Power(µw)	Delay(ns)	PDP (10^-15 w-s)	Area (µ.sqm)
Standard low V_t	1.443	0.064	0.092	32.16
Standard high V_t	1.726	0.119	0.205	32.41
Standard Dual V_t	1.212	0.106	0.128	32.96
Modified Dual V_t	1.355	0.071	0.096	32.23

V. NOISE IMMUNE DOMINO LOGIC CIRCUITS

In DOMINO gates, noise immunity is sacrificed for high performance. The DC noise margin of DOMINO gates is equal to the threshold voltage of pull-down transistors. Unlike static CMOS gates, the charge lost from dynamic node due to noise cannot be restored in DOMINO gates. This makes DOMINO gates more vulnerable to noise than static CMOS gates. A keeper is used to restore any loss of charge from the dynamic node. An analytical noise model for DOMINO gates where the effect of keeper is taken into account is considered.

Noise Margin:

The maximum voltage amplitude of extraneous signal that can be algebraically added to the noise-free worst-case input level without causing the output voltage to deviate from the allowable logic voltage level.

A typical n-type domino CMOS logic gate as shown in Fig. 9, consists of clock controlled transistor M1 and M2, a pull-down n-type transistor network, and an output driver. The operation of a domino CMOS logic gate can be divided into two phases. In the pre charge phase when the clock CLK is low, the dynamic node is charged to logic high through

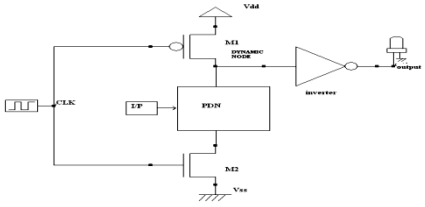


Fig. 16 domino logic

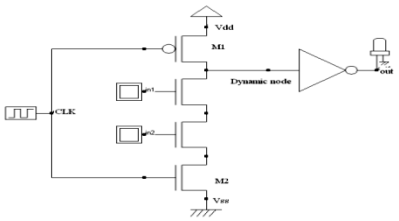


Fig.17. two input and gate

M1 and the output of the not gate is low. The evaluation phase starts when the clock goes high. In this phase, M1 is OFF and M2 is ON. The dynamic node discharges or retains its charge depending on the inputs to the pull-down network. A two input AND gate is illustrated in Fig.17.

Noise sources in dynamic logic circuits can be broadly classified into two basic types:

1) Gate internal noises, including charge sharing noise, leakage noise etc.,
2) External noises, including input noise, power and ground noise, and substrate noise.

Domino Noise Model:

Fig.18 describes the noise model for DOMINO gates.

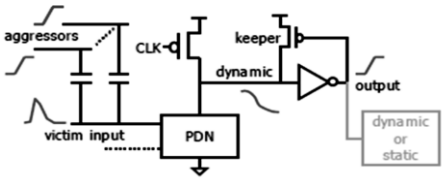


Fig.18. Crosstalk noise model for domino gates

Domino Noise Margin:

In order to obtain an analytical solution for noise margin for DOMINO gates, consider the current model for the PDN NMOS transistor.

We define the DOMINO noise margin as

$$DNM_{DOMINO} = \frac{NM_{inv}.C_d + \left|\frac{1}{2}.T.I_{k\_Max}\right|}{g_m}$$

Note that the keeper effect does not contribute to any extra computational cost since T is obtained from the already available input noise pulse and $I_{k\text{-}max}$ can also be pre-characterized.

Circuit Techniques for Noise immune Domino Logic:
Internal nodes precharging: (PCIN)

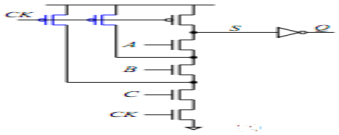


Fig.19. internal nodes Precharchig

A simple effective way to prevent the charge sharing problem is to precharge the internal nodes in the pull-down

network along with precharging the dynamic node. An example of dynamic 3-input AND gate using this technique is illustrated in Fig.19. Finally, it is noted that techniques based on precharging internal nodes alone are not very effective against external noises.

Pull-up Technique with PMOS: (PPTQ)

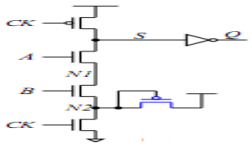


Fig.20. pull-up technique with PMOS

The pull-up technique, shown in Fig. 20, employs a PMOS transistor at node N2 forming a resistive voltage divider with the bottom clock controlled transistor. One major drawback of this technique is the DC power consumption in the resistive voltage divider. Furthermore, since the voltage level at the dynamic node S can never get lower than the voltage at node N2, the voltage swing at node S is not rail-to-rail. When the size of the PMOS pull-up transistor is large in an effort to aggressively raise gate noise immunity, the gate output may also not have a rail-to-rail swing.

NMOS Pull-up Technique: (NPTQ)

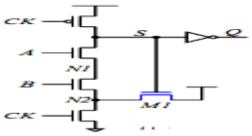


Fig.21. NMOS pull-up technique

An improved method, shown in Fig. 21, employs a pull-up transistor with feedback control. Here an NMOS transistor M1 is used to pull up the voltage of an internal node. This design allows the pull-up transistor to be shut off when the voltage of the dynamic node goes low, therefore, the dynamic node S undergoes rail-to-rail voltage swing. Also, the DC power consumption problem is partially solved.

Feedback NMOS Mirror Technique: (MRTQ)

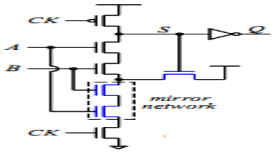


Fig.22. Feedback NMOS Mirror technique

The mirror technique employs a feedback controlled NMOS transistor similar to the NMOS pull-up technique. In addition, it duplicates the pull-down network in an effort to further reduce DC power consumption and to further improve gate noise tolerance. A 2-input dynamic AND gate designed using the mirror technique is shown in Fig. 22. However, this technique significantly lengthens the discharge path in the pull-down network, which potentially leads to slower circuit or considerably increased circuit active area when the transistors are aggressively sized.

NMOS Two Transitor Technique: (TTTQ)

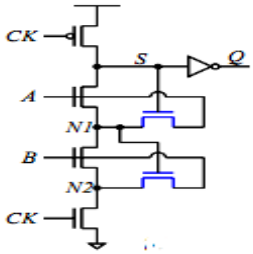


Fig.23. NMOS Two transistor technique

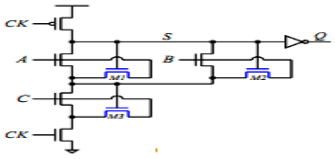


Fig.24. A 3-input OR-AND gate

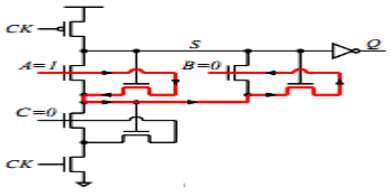


Fig.25. Direct conducting path.

The NMOS two transistor technique adopts NMOS pull-up transistors at all internal nodes to further improve dynamic gate noise immunity. In addition, the drain nodes of the pull-up NMOS transistors are connected to the inputs instead of to the power-supply network, as illustrated in Fig.23. As an example, in Fig.24, we show a 3-input OR-AND gate implementing the logic function of (A+ B).C. Assume input A is high while inputs B and C are low. The dynamic node S stays high because C is low and there is no discharging path to the ground. Under such scenario, there is a DC conducting path between the two inputs A and B, as illustrated in Fig.25.

Complementary weak P-Network Technique: (CPNTQ)

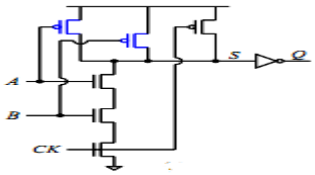


Fig.26.Complementary weak p-network technique

The basic principle of this class of techniques is to construct a weak complementary p-network to prevent the dynamic node from being floating in the evaluation phase. One such technique is illustrated in Fig. 26. In additional to the silicon area overhead associated with the pull-up network, a major drawback of this technique in practice is its ineffectiveness in dealing with very wide logic gates, for example, wide OR gates, where dynamic logic styles really outshine static CMOS logic gate in performance.

Inverter Technique: (CMITQ)

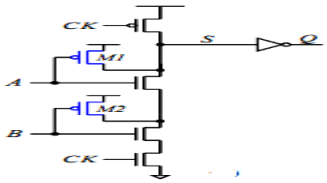


Fig.27(a). inverter circuit

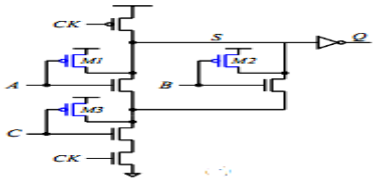


Fig.27(b) A 3- input OR-AND gate.

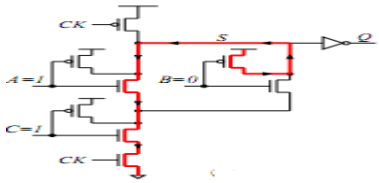


Fig.27(c) Direct conducting path

PMOS transistors can also be employed at a per transistor level, as shown in Fig. 27. This technique is known as inverter technique.

Inverter Gated Technique: (GCMITQ)

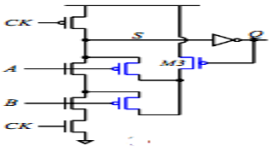


Fig: 28. inverter gated technique

In Fig: 27(a), for example, if input A stays high and input B falls from high to low during the evaluation phase, the dynamic node may be reset to high by the pull-up PMOS transistor M2. With a view to solve this false reset problem, an additional transistor M3 is used shown in Fig. 28, it is called inverter gated technique.

Three Transistor Technique: (TTRTQ)

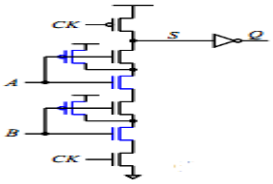


Fig.29. Three transistor technique

Figure 29 illustrates a noise-tolerant 2-input AND gate using a triple transistor technique, where each NMOS transistor in the pull-down network of a simple dynamic logic gate is replaced by three transistors.

Noise immune logic using different keepers:

Domino Always on Keeper (DAOK):

Always On Keeper uses 'weak'-PMOS device between the output node and V_{DD} as shown in Figure 30. As the gate is connected to GND, this PMOS device will always be turned ON. So, even in the evaluation phase, the output node will be connected in some capacity to V_{DD}. The PMOS 'keeper,' has the effect of maintaining the output node charge even at slower clock speeds.

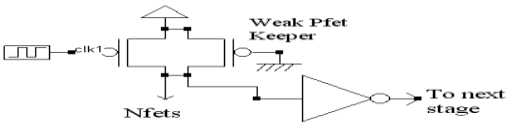


Fig.30. Domino always on keeper

Although this configuration has advantages, it does introduce another PMOS device into each stage and also causes excess power dissipation due to possibility of the connection from V_{DD} to GND through the NMOS devices and the PMOS keeper.

Domino Feedback Keeper (DFBK):

The use of a keeper PMOS in dynamic logic could be further improved by connecting the gate of the keeper not to GND, but to the output node of the inverter stage as shown in Figure 31.

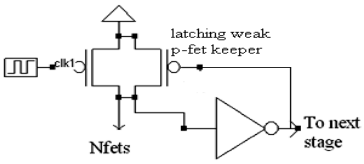


Fig: 31.Domino Feedback keeper

The keeper would now function as a latch cutting off whenever the output of the inverter is high. In this way, power dissipation is significantly reduced whenever a pull-down path to GND has been formed in the NMOS logic block since this would make the input to the inverter low and thus the output of the inverter high. When the output of the inverter is low however, as would be the case if no pull-down path to ground was formed in the NMOS logic block, the keeper PMOS would turn on and maintain the output high charge on the precharge node even at reduced clock speeds or an idle.

Domino Standard Keeper (DSTDK):

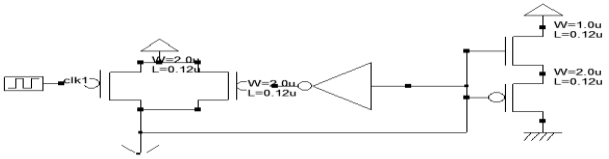


Fig.32. Standard keeper

Domino Modified Feedback Keeper (DMDFBK):

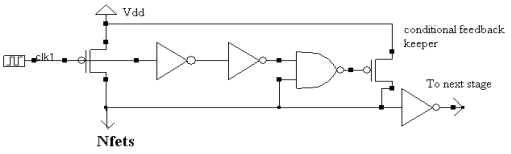


Fig.33.Domino modified feedback keeper

The Conditional Feedback Keeper is the keeper consists of two not gates and a NAND gate and a PMOS transistor. The conditional feedback keeper provides two delays by using two not gates in order to retain the voltage at the dynamic node when the pull down network is off during the evaluation phase.

Domino Modified Feedback High Performance Keeper (DMDFBKHP):

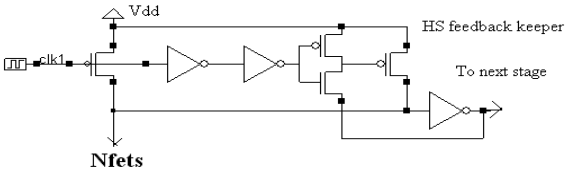


Fig.34 Modified Feedback Keeper High Performance

The Modified feedback keeper high performance is termed as high speed feedback keeper, the keeper consists of two not gates and CMOS inverter and a PMOS transistor. The Modified feedback keeper high performance provides two delays by using two not gates in order to retain the voltage at the dynamic node when the pull down network is off during the evaluation phase.

Simulation and Implementation Results:

The simulation results are given in below Tables13-21.

OR8 (65nm Technology):

Table13. OR8 gate

S. No	Technique	Power Dissipation(µw)	Propagation Delay (ns)	PDP (10^-18 w-s)	Noise Margin(mv) [power(µw)]	Area (µ. sqm)	No of Symbols
1	PCIN	1.813	0.040	72.52	60[2.52]	86.07	23
2	PPTQ	1.943	0.044	85.49	70[3.46]	75.70	25
3	NPTQ	126	0.044	5544	120[148]	79.21	25
4	MRTQ	3.308	0.077	254.71	160[18.92]	133.74	33
5	TTTQ	0.033	1.163	38.37	230[0.243]	111.72	31
6	CPNTQ	54.941	0.110	6043	300[93.80]	121.74	31
7	CMITQ	68.611	0.256	17564	250[93.19]	123.49	39
8	GCMITQ	14.939	0.063	941.15	350[65.10]	113.65	32
9	TTRTQ	4.071	0.047	191.33	230[8.76]	117.92	47

Table14. OR8 gate

S.NO	CIRCUIT TECH	PROPAGATION DELAY(nsec)	POWER DISSIPATION (µW)	PDP ($\times 10^{-18}$)	NOISE MARGIN(mV) [POWER (µw)]	AREA (µ-mts^2)	NO.OF SYMBOLS
1	DAOK	0.049	89.033	4362	200[89.339]	63.03	25
2	DFBK	0.086	7.129	613	230[42.685]	63.24	24
3	DSTDK	0.088	7.762	683	230[43.743]	74.14	28
4	DMDFBK	0.063	0.988	62	1600[218]	74.60	27
5	DMDFBKHP	0.060	9.473	568	1700[220]	80.63	28

AND2 (65nm Technology):

Table15. AND2 gate

S.NO	CIRCUIT TECH	PROPAGATION DELAY(nsec)	POWER DISSIPATION (µW)	PDP ($\times 10^{-18}$)	NOISE MARGIN(mV) [POWER (µw)]	AREA (µ-mts^2)	NO.OF SYMBOLS
1	DAOK	0.049	89.033	4362	200[89.339]	63.03	25
2	DFBK	0.086	7.129	613	230[42.685]	63.24	24
3	DSTDK	0.088	7.762	683	230[43.743]	74.14	28
4	DMDFBK	0.063	0.988	62	1600[218]	74.60	27
5	DMDFBKHP	0.060	9.473	568	1700[220]	80.63	28

OR2 (65nm Technology):

Table16. OR2 gate

S.NO	CIRCUIT TECH	PROPAGATION DELAY(nsec)	POWER DISSIPATION (µW)	PDP ($\times 10^{-18}$)	NOISE MARGIN(mV) [POWER (µw)]	AREA (µ-mts^2)	NO.OF SYMBOLS
1	DAOK	0.025	0.172	4.3	250[67.497]	21.17	13
2	DFBK	0.050	0.173	8.65	220[54.589]	21.36	12
3	DSTDK	0.064	7.264	464	220[67.237]	30.17	16
4	DMDFBK	0.034	0.551	18	740[167]	31.89	15
5	DMDFBKHP	0.036	0.539	19	950[182]	35.19	16

OR4 (65nm Technology):

Table17. OR4 gate

S.NO	CIRCUIT TECH	PROPAGATION DELAY(nsec)	POWER DISSIPATION (µW)	PDP ($\times 10^{-18}$)	NOISE MARGIN(mV) [POWER (µw)]	AREA (µ-mts^2)	NO.OF SYMBOLS
1	DAOK	0.033	76.042	2500	100[74.599]	34.27	17
2	DFBK	0.062	0.733	45	220[27.158]	34.27	16
3	DSTDK	0.072	1.019	73	330[50.026]	45.50	20
4	DMDFBK	0.047	0.742	34	1500[210]	47.29	19
5	DMDFBKHP	0.044	0.743	32	1700[220]	80.63	28

XOR2 (65nm Technology):

Table18. XOR2 gate

S.NO	CIRCUIT TECH	PROPAGATION DELAY(nsec)	POWER DISSIPATION (µW)	PDP ($\times 10^{-18}$)	NOISE MARGIN(mV) [POWER (µw)]	AREA (µ-mts^2)	NO.OF SYMBOLS
1	DAOK	0.034	21.788	740	120[6.323]	42.08	17
2	DFBK	0.048	21.791	1045	140[17.187]	42.08	16
3	DSTDK	0.069	22.206	1532	140[18.81]	50.71	20
4	DMDFBK	0.051	1.755	89	380[6.360]	53.07	19
5	DMDFBKHP	0.048	19.507	936	400[17.402]	56.59	20

8-Bit MUX(65nm Technology):

Table19. 8-Bit MUX

S.NO	CIRCUIT TECH	PROPAGATION DELAY(nsec)	POWER DISSIPATION (µW)	PDP ($\times 10^{-18}$)	NOISE MARGIN(mV) [POWER (µw)]	AREA (µ-mts^2)	NO.OF SYMBOLS
1	DAOK	0.058	53.854	3123	230[138]	135.58	40
2	DFBK	0.063	53.92	3396	230[138]	135.55	39
3	DSTDK	0.073	13.183	962	100[18.117]	151.03	43
4	DMDFBK	0.039	13.251	516	100[19.730]	149.92	42
5	DMDFBKHP	0.031	13.075	405	100[19.217]	160.82	43

16-Bit MUX(65nm Technology):

Table20. 16-Bit MUX

S.NO	CIRCUIT TECH	PROPAGATION DELAY(nsec)	POWER DISSIPATION (μW)	PDP ($\times 10^{-18}$)	NOISE MARGIN(mV) [POWER (μw)]	AREA (μ-mts^2)	NO.OF SYMBOLS
1	DAOK	0.090	25.463	2291	180[107]	392.3	72
2	DFBK	0.104	25.536	2655	120[38.303]	395.2	71
3	DSTDK	0.105	13.510	1418	70[18.424]	419.2	75
4	DMDFBK	0.095	13.440	1276	70[16.836]	419.8	74
5	DMDFBKHP	0.097	13.462	1305	70[17.039]	428.7	75

Table21.4 Input OR gate

S.NO	Technique	Power Dissipation(μw)	Propagation Delay(ns)	PDP (10^-18 w-s)	Noise Margin(mv) [power(μw)]	Area (μ.sqm)	No of Symbols
1	PCIN	0.810	0.024	19.44	60[1.83]	32.19	15
2	PPTQ	204	0.028	5712	270[260]	38.67	17
3	NPTQ	53.74	0.028	1504	200[60.71]	40.29	17
4	MRTQ	2.129	0.045	958.05	250[37.55]	63.14	21
5	TTTQ	39.52	0.097	3833	50[40.08]	55.35	19
6	CPNTQ	34.47	0.086	2964	220[71.50]	98.43	19
7	CMITQ	34.57	0.192	6638	210[67.97]	93.03	23
8	GCMITQ	2.546	0.039	99.29	430[73.24]	59.41	20
9	TTRTQ	1.922	0.031	59.58	280[75.08]	81.37	27

Table22..2 Input AND gate

S.NO	Technique	Power Dissipation(μw)	Propagation Delay(ns)	PDP (10^-18 w-s)	Noise Margin(mv) [power(μw)]	Area (μ.sqm)	No of Symbols
1	PC	0.431	0.01	81.8	350[	26.	12
	IN		9	9	45.55]	20	
2	PPTQ	64.236	0.025	1605	300[80.81]	25.16	13
3	NPTQ	43.006	0.037	1591	470[71.49]	24.66	13
4	MRTQ	0.558	0.032	17.85	360[28.30]	34.81	15
5	TTTQ	0.391	0.034	13.29	60[4.48]	34.98	13
6	CPNTQ	0.397	0.045	17.86	260[25.81]	29.62	13
7	CMITQ	13.764	0.043	591.85	160[22.19]	30.64	15
8	GCMITQ	0.485	0.030	14.55	340[38.29]	35.23	14
9	TTRTQ	0.780	0.031	24.18	150[21.67]	23.34	17

VI. CONCLUSIONS

This work consists of four parts. In section II the circuits Dynamic DCVSL footed circuit, Dynamic DCVSL footless circuit; Dual-Rail Data-Driven Dynamic Logic and Dual-rail Footless domino gate with self-timed precharge scheme are successfully implemented using CMOS domino logic. The proposed circuits have offered an improved performance in power dissipation, speed and noise tolerance when compared with standard domino circuit. In section III, Pseudo footless domino circuit is proposed. The proposed circuit offers better performance. In section IV, energy-efficient domino logic is presented. Among the four techniques, the standard dual V_t and modified dual V_t offer better performance. In section V, an attempt has been made to simulate the noise immunity of the benchmark domino circuits with different techniques and keeper transistors which are the basic building blocks for high performance. The proposed circuits have offered an improved performance in power dissipation and noise tolerance when compared with standard domino circuit. As it is observed from the results, the DMDFBK and DMDFBKHP have lower PDP, high noise immunity. Hence, it is concluded that the proposed designs will provide a platform for designing high performance and low power digital circuits and high noise immune digital circuits such as, processors and multipliers.

REFERENCES

[1] L. G. Heller, W. R. Griffin, J. W. Davis, and N. G. Thoma, “Cascode voltage switch logic: A differential CMOS logic family,” in Proc. IEEE Int. Solid-State Circuits Conf., pp. 16-17, 1984.

[2] P. Ng, P. T. Balsara, and D. Steiss, “Performance of CMOS Differential Circuits,” IEEE J. of Solid-State Circuits, vol. 31, no. 6, pp. 841-846, June 1996.

[3] P. Hofstee, et al., “A 1 GHz Single-Issue 64b PowerPC Processor,” in Proc. IEEE Int. Solid-State Circuits Conf., pp. 92-93, 2000.

[4] J. Wang, S. Shieh, C. Yeh, and Y. Yeh, “Pseudo-Footless CMOS Domino Logic Circuits for High-Performance VLSI Designs,” in Proc. Int. Symp. on Circuits and Systems, vol. 2, pp. 401-404, 2004.

[5] R. Rafati, A. Z. Charaki, G. R. Chaji, S. M. Fakhraie, and K. C. Smith, “Comparison of a 17b Multiplier in Dual-Rail Domino and in Dual-Rail

D^3L (D^4L) Logic Styles," in Proc. Int. Symp. on Circuits and Systems, vol. 3, pp. 257-260, 2002.

[6] S. Mutoh et al., "1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS," IEEE J. Solid-State Circuits, vol.30, pp. 847–854, Aug. 1995.

[7] V. Kursun and E. G. Friedman, "Domino logic with dynamic body biasedkeeper," in Proc. Eur. Solid-State Circuits Conf., Sept. 2002, pp.675–678.

[8] "Variable threshold voltage keeper for contention reduction in dynamic circuits," in Proc. IEEE Int. ASIC/SOC Conf., Sept. 2002, pp.314–318.

[9] S. Borkar, .Low Power Design Challenges for the Decade,. Proceedings of the IEEE/ACM Design Automation Conference, pp. 293-296, June 2001.

[10] P. Srivastava, A. Pua, and L. Welch, .Issues in the Design of Domino Logic Circuits, Proceedings of the IEEE Great Lakes Symposium on VLSI, pp. 108-112, February 1998.

[11] G. Balamurugan and N. R. Shanbhag, .Energy-efficient Dynamic Circuit Design in the Presence of Crosstalk Noise,. Proceedings of the IEEE International Symposium on Low Power Electronics and Design, pp. 24-29, August 1999.

[12] S.Govindarajulu, Dr.T.Jayachandra Prasad "Design of High Performance Dynamic CMOS Circuits in Deep submicron Technology" International Journal of Engineering Science and Technology, Vol.2 (7), 2010, pp.2903-2917, ISSN:0975-5462

[13] S.Govindarajulu, Dr.T.Jayachandra Prasad et.al. "Low Power, Reduced Dynamic Voltage Swing Domino Logic Circuits" Indian Journal of Computer Science and Engineering, 2010 pp.74-81, ISSN:0976-5166.

[14] S.Govindarajulu, Dr.T.Jayachandra Prasad "Energy efficient Reduced Swing Domino Logic Circuits in 65 nm Technology" International Journal of Engineering Science and Technology, Vol.2 (6), 2010, pp.2248-2257, ISSN:0975-5462.

[15] S.Govindarajulu, Dr.T.Jayachandra Prasad et.al. "Design of High Performance Arithmetic and Logic Circuits in DSM Technology" International Journal of Engineering and Technology, Vol.2 (4), 2010, pp.285-291, ISSN:0975-4024.

[16] S.Govindarajulu, Dr.T.Jayachandra Prasad et.al. "High Performance VLSI Design Using Body Biasing in Domino Logic Circuits" International Journal of Computer Science and Engineering, Vol.2, No.5, 2010 pp.1741-1745, ISSN:0975-3397.

[17] S.Govindarajulu, Dr.T.Jayachandra Prasad et.al. "Design of Low Power, High Speed, Dual Threshold Voltage CMOS Domino Logic Circuits with PVT Variations" International Journal of Electronic and Engineering Research, Vol.2, No.5, 2010 pp.619- 629, ISSN:0975-6450.

AUTHORS PROFILE

[1]Salendra.Govindarajulu:- He is working as an Associate Professor in the Dept. of Electronics & Communication Engg. at RGMCET, Nandyal, Andhra Pradesh, India. He presented more than 25 International/National Technical Papers. He is a Life Member of ISTE, New Delhi. His interest includes Low Power VLSI CMOS design.

[2]Dr.T.Jayachandra Prasad:- He is working as a Principal and Professor in the Dept. of Electronics & Communication Engg. at RGMCET, Nandyal Andhra Pradesh, India. He presented more than 48 International/National Technical Papers. He is Life Member in IE (I), CALCUTTA, Life Member in ISTE, NEW DELHI, Life Member in NAFEN, NEW DELHI, and IEEE Member. His interest includes Digital Signal Processing.

E-Shape Micro strip Patch Antenna on Different Thickness for pervasive Wireless Communication

Neenansha Jain
Department of ECE
NIST, RGTU
Bhopal, India

Anubhuti Khare
Department of ECE
UIT, RGPV
Bhopal, India

Rajesh Nema
Department of ECE
NIST, RGTU
Bhopal, India

***Abstract*—In this Paper Presents the result for different standard thickness values, and the result is performed by thickness of 31 mile, Ku- band frequency 12GHz are gives the best result. The antenna has become a necessity for many applications in recent wireless communications, such as Radar, Microwave and space communication. The proposed antenna design on different thickness and analyzed result of all thickness between 1GHz to 15GHz frequency, When the proposed antenna design on a 31 mil RT DUROID 5880 substrate from Rogers-Corp with dielectric constant of 2.2 and loss tangent of 0.0004. At 12GHz the verify and tested result on IE3D SIMULATOR are Return loss = -23.08dB, VSWR = 1.151, Directivity = 11dBi, Gain = 4dBi, 3 dB beam width = 35.5575 degrees, Mismatch loss= -0.0289842dB is very low, Efficiency= 65.3547%, All results shown in Simulation results. The Return losses and VSWR results shown in Table 1, Table2 respectively.**

Keywords- Micro strip antenna; IE3D SIMULATOR; Dielectric; Patch width; Patch Length; Losses; strip width; strip length.

I. INTRODUCTION

Ahmed H. Reja [1] proposed Study of Micro Strip Feed Line Patch Antenna experimentally increase the Return Loss -33.6dB at 2.5GHz frequency and VSWR is 1.5 by using CAD (Microwave office 2000 version 3.22) for RT DUROID 5880. Santanu Kumar Behera and Y. Choukiker [2] proposed Design and Optimization of Dual Band Micro Strip Antenna using Practical Swarm Optimization maximize the return loss for dual band Frequency at 2.4GHz is -43.95dB and at 3.08GHz is -27.4dB. A A Deshmukh and G Kumar [3] proposed compact L Shape patch broadband Microstrip antenna experimentally increase bandwidth up to 13.7%. Z M Chen [4] further increase bandwidth of this antenna up to 23.7% - 24.43%. K F Lee [5] proposed U Shape slot shorting post small size Microstrip Antenna and increase bandwidth up to 42%. S C Gao [6] used uniplanar photonic band gap structure for enhancing band width and gain. M Khodier[7] New wideband stacked microstrip antennas for enhancing band width. Major issue for micro strip antenna is narrow Bandwidth.

II. MATHMATICAL ANALYSIS

Theoretical analysis and calculations from of all dimensions will be obtained;

The width of the patch element (W) is given by.

$$W = \frac{c}{2 f_o \sqrt{\frac{(\varepsilon_r + 1)}{2}}}$$

Substituting c = 3x10^8 m/s, ε r = 2.2, and f o = 5 GHz, then

W =2.3717cm or 933.74 mile.

The effective of the dielectric constant (εreff) depending on the same geometry (W, h) but is surrounded by a homogeneous dielectric of effective permittivity εreff, whose value is determined by evaluating the capacitance of the fringing field.

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-\frac{1}{2}}$$

Substituting ε r = 2.2, W =2.3717cm, and h =0.1575cm, then

ε reff = 2.1074cm or 829.69mile.

The effective length (L eff) is given

$$L_{eff} = \frac{c}{2 f_o \sqrt{\varepsilon_{reff}}}$$

Substituting c = 3x10^8 m/s, εreff = 2.0475cm, and f o = 5 GHz, then L_{eff} = 2.0665cm or 813.6 mile.

The length extension (ΔL) is given by:

$$\Delta L = 0.412h\frac{(\varepsilon_{reff} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\varepsilon_{reff} - 0.258)\left(\frac{W}{h} + 0.8\right)}$$

Substituting E_{reff} = 2.1074cm, W= 2.3717cm, and h =0.0787cm, then ΔL= 0.041469cm or 16.3266mile.

The actual length (L) of patch is obtained by:

$$L = L_{eff} - 2\Delta L$$

Substituting ΔL= 0.041469cm, and L_{eff} = 2.0665cm, then L=1.9835cm or 780.92mile.

III. ANTENNA DESCRIPTION

The results of proposed E-Shaped Multiband micro strip patch antenna verified in IE3D Simulator with optimization.

A. *Proposed Antenna on 31mil RT DUROID 5880 substrate:*

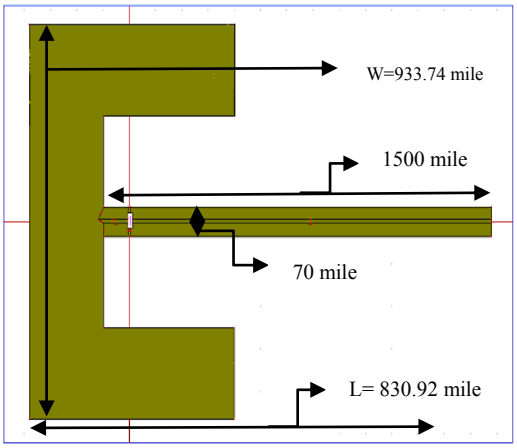


Fig. 1Block Diagram of Proposed Antenna

The Proposed antenna has:-
Proposed Patch length = 780.92+50miles
Proposed Patch Width = 933.74miles
Strip Path Length= 1500miles
Strip Path Width= 70miles
Cut width = 300miles
Cut depth = 300 miles

IV. RESULT AND DISSCUSSIONS

A. *Comparison of Different Micro strip Patch Antenna in Different thickness by using optimization in IE3D Simulator for RT DUROID 5880 Substrate.*

1) *Thickness when h=15mile*

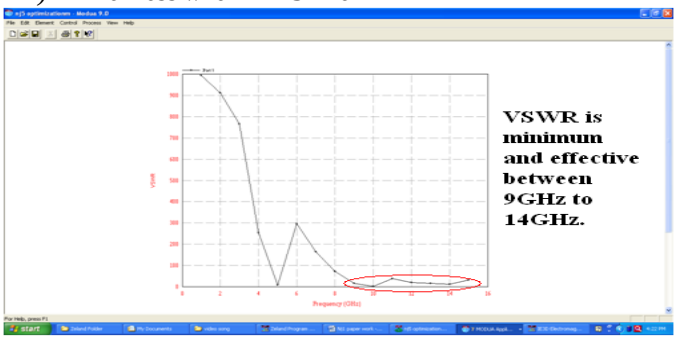


Fig. 2 VSWR Vs Frequency (in GHz)

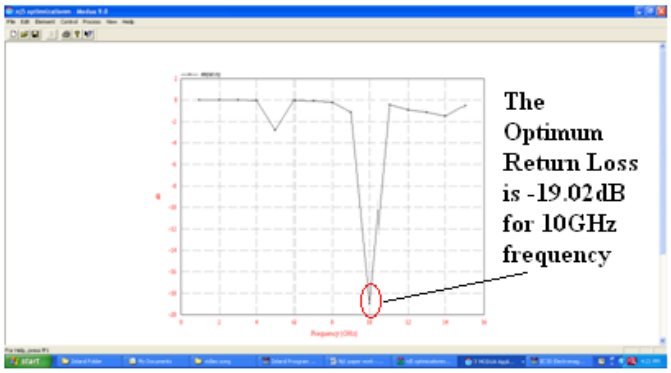


Fig. 3 Return Loss Vs Frequency (in GHz)

2) *Thickness when h=20mile.*

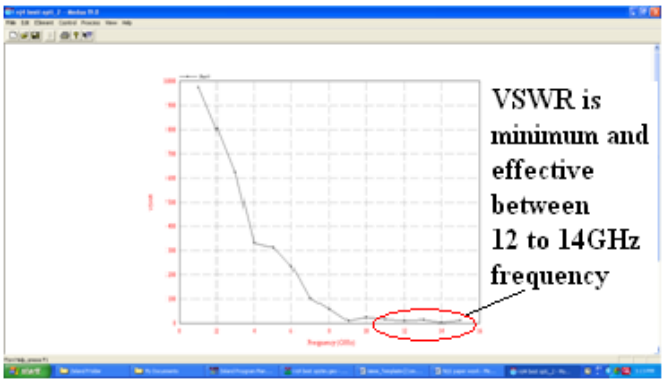


Fig. 4 VSWR Vs Frequency (in GHz)

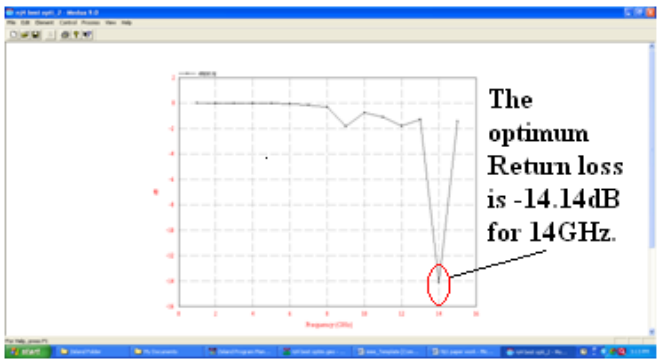


Fig. 5 Return Loss Vs Frequency (in GHz)

3) *Thickness when h=31mile*

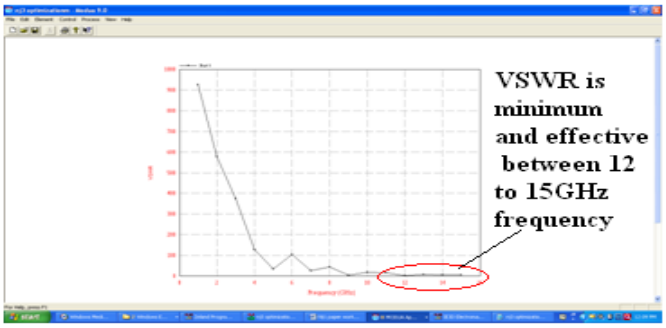


Fig. 6 VSWR Vs Frequency (in GHz)

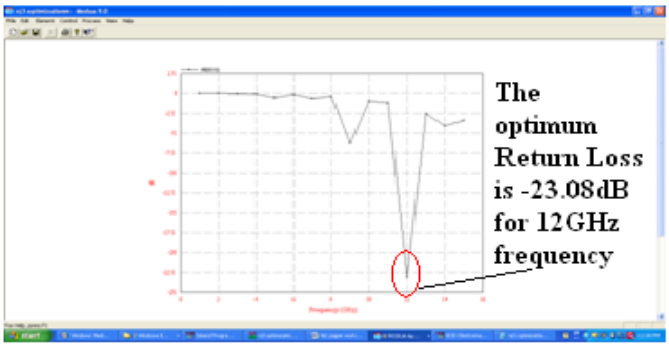


Fig. 7 Return Loss Vs Frequency (in GHz)

4) Thickness when h=62mile

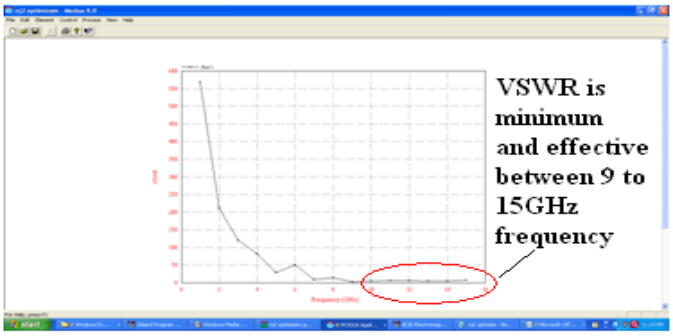


Fig. 8 VSWR Vs Frequency (in GHz)

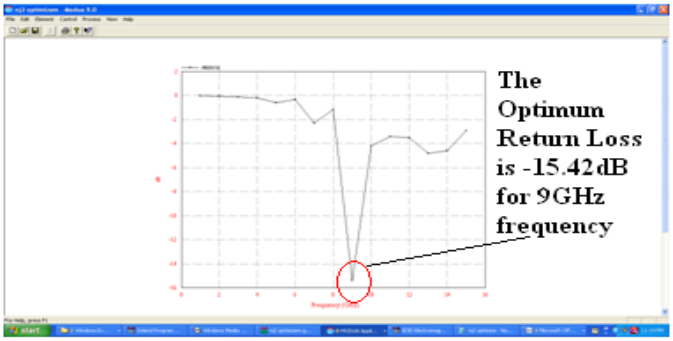


Fig. 9 Return Loss Vs Frequency (in GHz)

5) Thickness when h=125mile

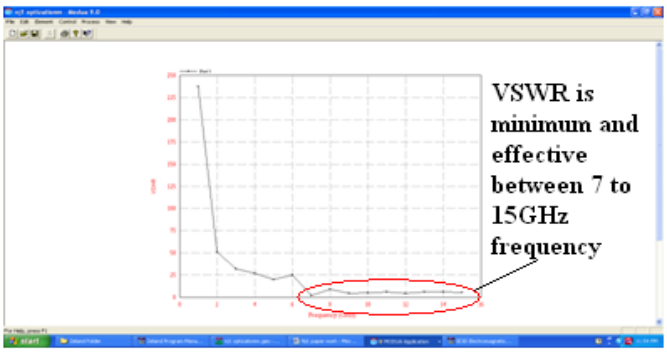


Fig. 10 VSWR Vs Frequency (in GHz)

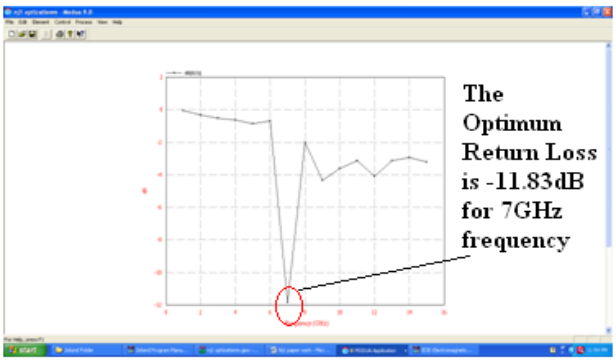


Fig. 11 Return Loss Vs Frequency (in GHz)

B. Best Result Simulated Micro strip Patch Antenna in IE3D Simulator for 31mil RT DUROID 5880 Substrate

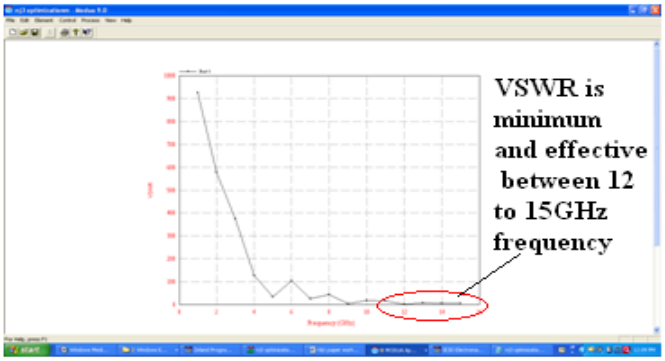


Fig. 12 VSWR Vs Frequency (in GHz)

For proposed design the value of VSWR is effective between 12GHz to 15GHz, for this value return loss is minimum. At 12GHz return loss is -23.08dB and VSWR is 1.151, At 9GHz VSWR is 2.909, 13GHz VSWR is 6.687, At 14GHz VSWR is 4.311, at 15GHz VSWR is 5.145.

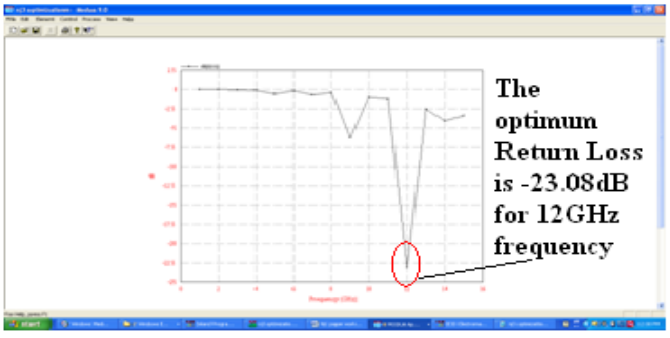


Fig. 13 Return Loss Vs Frequency (in GHz)

The frequency at 10GHz return losses is -17.71, at 11GHz return losses is -1.222dB, and at 13GHz return losses reduce very significantly -23.08dB.

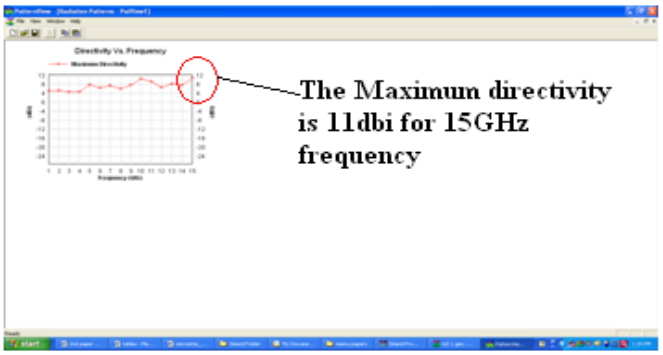


Fig. 14 Directivity Vs Frequency (in GHz)

At 10GHz frequency Directivity is 11dBi, at 12GHz Directivity is 7dBi, at 13GHz Directivity is 8dBi, and at 15 GHz Directivity is 11dBi.

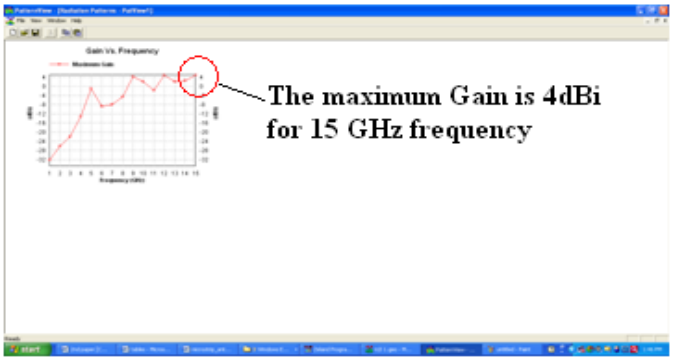


Fig. 15 Gain Vs Frequency (in GHz)

At 10GHz Frequency Gain is 3dBi, at 12GHz Gain is 5dBi, at 13GHz Gain is 2dBi, and at 15GHz Gain is 4dBi.

C. Radiation Pattern for 13GHz Frequency:

Study of different Azimuth pattern and Elevation pattern in IE3D. Analyzed radiation characteristic of antenna at 13 GHz shown in figure.

1) 2D Polar Radiation pattern

a) Elevation Pattern

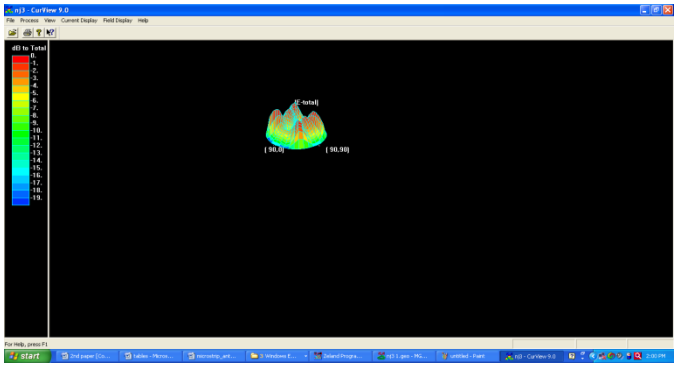


Fig. 16 Elevation Pattern of E Maximum

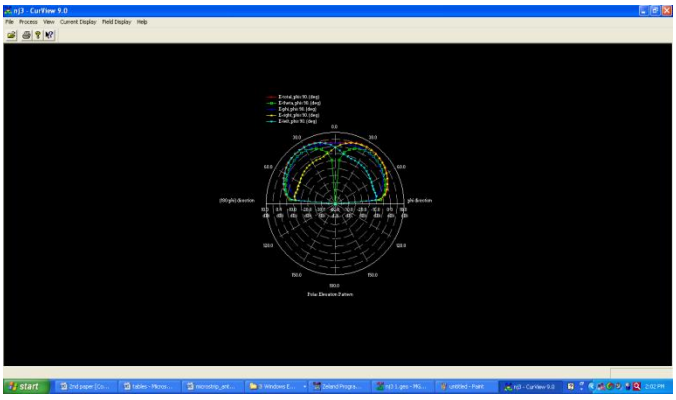


Fig. 17 Elevation Pattern of E Total, E Right, E Left, E Theta, E Phi at Phi=90 (deg)

b) Azimuth Pattern

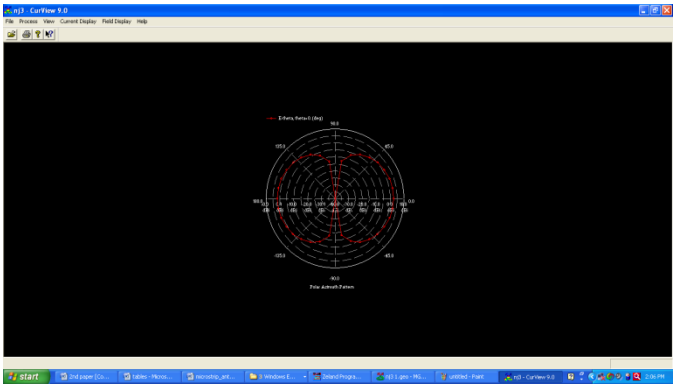


Fig. 18 Azimuth Pattern of E Theta=0(deg)

2) Axial Ratio Pattern

a) Elevation Pattern

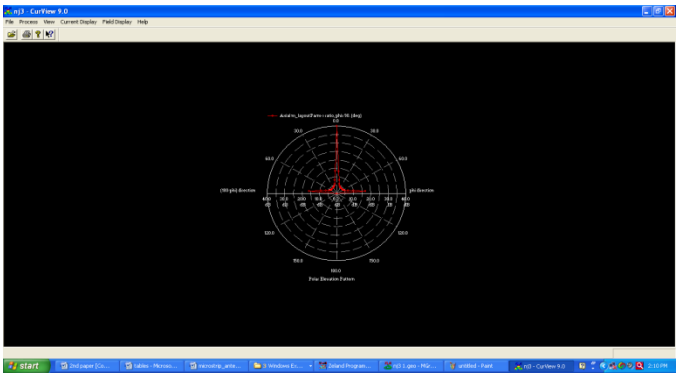


Fig. 19 Axial Pattern of Phi= 90(deg)

b) Azimuth Pattern

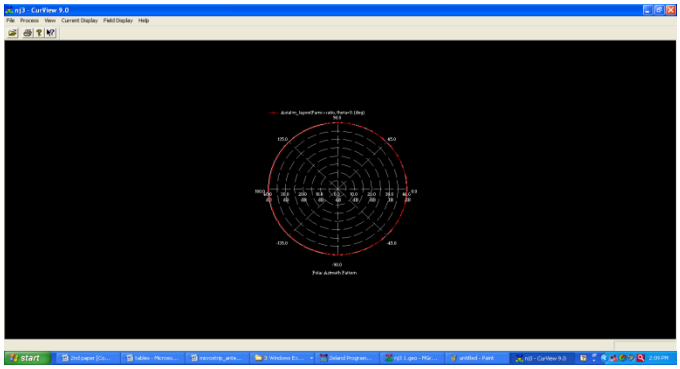


Fig. 20 Axial Pattern of theta = 0(deg)

3) 3D Pattern Display

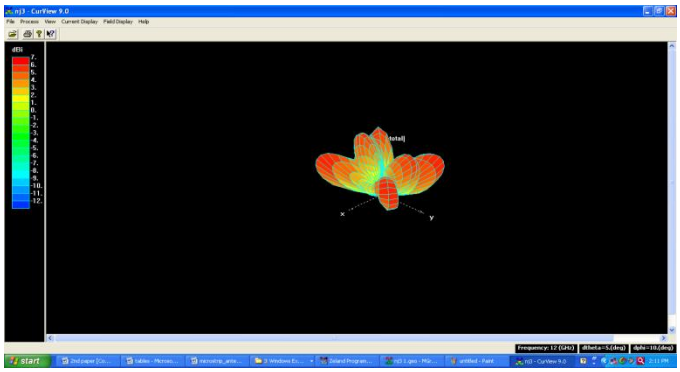


Fig. 21 Elevation Pattern at E-total

4) 2D Radiation Pattern

a) Elevation Pattern

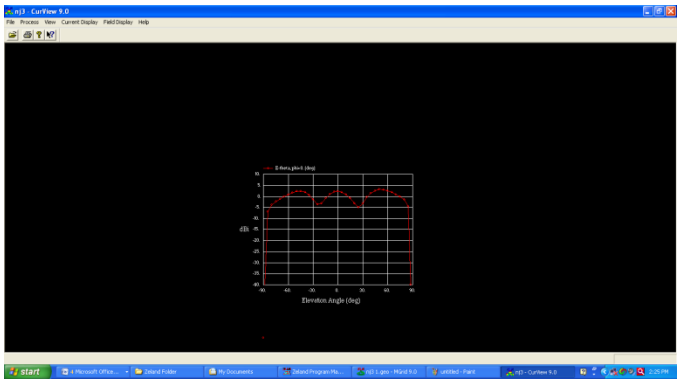


Fig. 22 Elevation Pattern at E-theta at phi=0(deg)

b) Azimuth Pattern

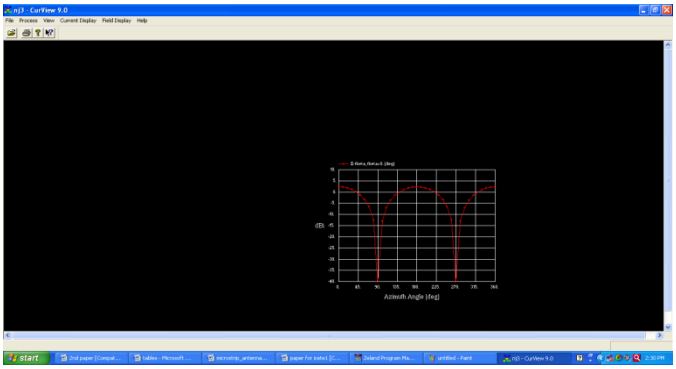


Fig. 23 Azimuth Pattern at E-total at theta=0(deg)

V. CONCLUSION

Micro strip antennas have become a rapidly growing area of research. Their potential applications are limitless, because of their light weight, compact size, and ease of manufacturing. One limitation is their inherently narrow bandwidth. However, recent studies and experiments have found ways of overcoming this obstacle. A variety of approaches have been taken, including modification of the patch shape, experimentation with substrate parameters, Most notably mobile communication systems where many frequency ranges could be accommodated by a single antenna. We here design simple and low costlier patch antenna for pervasive wireless communication by using different patch length. The transmission line model seems to be the most instructive in demonstrating the bandwidth effects of the changing the various parameters. When the proposed antenna design on a 31mil RT DUROID 5880 substrate from Rogers-Corp with dielectric constant of 2.2 and loss tangent of 0.0004. The proposed antenna give best result when antenna has standard thickness is 31 mile and after optimization addition of extra length is 50mile patch length and some little changes of patch width and more feed line length. The proposed frequency range 12GHz (Ku Band) and Analysis Radiation Characteristics of micro strip Antenna by IE3D Simulator. The results of proposed designing are effective between 1GHz-15GHz. proposed antenna simulated in IE3D Simulator. The optimum results of proposed antenna verify and tested in IE3D SIMULATOR. The simulated results of IE3D at 13GHz is Return loss = -23.08dB, VSWR = 1.151, Directivity = 11 dBi, Gain = 4dBi, 3 dB beam width = 35.5575degrees, Mismatch loss= -0.0289842dB is very low, Efficiency= 65.3547%, Total Radiated Power= 0.00649199W, Average Radiated Power= 0.000516616W/s and Input Radiated Power at ports= 0.009973348.The proposed 31mil RT DUROID 5880 substrate E-Shaped multiband micro strip antenna effective work on 12GHz(Ku Band) the proposed antenna work very effectively for pervasive wireless communication.

TABLE1: Frequency (in GHz) Vs Return Losses in dB[s(i j)] for Different Thickness

Freq GHz	dB[S(1,1)] for h=15mile	dB[S(1,1)] for h=20mile	dB[S(1,1)] for h=31mile	dB[S(1,1)] for h=62mile	dB[S(1,1)] for h=125mile
1	-1.748e-002	-1.784e-002	-1.88e-002	-3.056e-002	-7.317e-002
2	-1.909e-002	-2.163e-002	-3.011e-002	-8.247e-002	-0.3412
3	-2.273e-002	-2.793e-002	-4.655e-002	-0.1457	-0.5498
4	-6.832e-002	-5.278e-002	-0.1371	-0.2139	-0.6493
5	-2.835	-5.589e-002	-0.535	-0.6108	-0.8795
6	-5.886e-002	-7.422e-002	-0.1688	-0.3482	-0.6976
7	-0.1057	-0.1715	-0.6865	-2.329	-11.83
8	-0.2456	-0.309	-0.3989	-1.206	-2.023
9	-1.135	-1.828	-6.225	-15.42	-4.345
10	-19.02	-0.7607	-1.024	-4.22	-3.638
11	-0.4599	-1.109	-1.222	-3.419	-3.147
12	-0.9343	-1.808	-23.08	-3.531	-4.105
13	-1.149	-1.305	-2.617	-4.851	-3.142
14	-1.502	-14.14	-4.104	-4.615	-2.962
15	-0.5566	-1.461	-3.42	-2.908	-3.22

TABLE2: Frequency (in GHz) Vs. VSWR for Different Thickness

Freq GHz	VSWR for h=15mile	VSWR for h=20mile	VSWR for h=31mile	VSWR for h=62mile	VSWR for h=125mile
1	993.9	973.8	924.1	568.5	237.4
2	910.1	803.1	576.9	210.6	50.92
3	764.1	621.9	373.2	119.2	31.61
4	254.3	329.1	126.7	81.22	26.77
5	6.183	310.8	32.48	28.45	19.77
6	295.2	234.1	102.9	49.9	24.92
7	164.3	101.3	25.32	7.505	1.688
8	70.73	56.22	43.56	14.42	8.627
9	15.33	9.536	2.909	1.408	4.081
10	1.252	22.85	16.99	4.198	4.844
11	37.78	15.69	14.24	5.146	5.581
12	18.61	9.646	1.151	4.988	4.311
13	15.15	13.33	6.687	3.674	5.588
14	11.6	1.489	4.311	3.852	5.922
15	31.22	11.91	5.145	6.029	5.456

ACKNOWLEDGMENT

The Authors would like to thanks Principal & H.O.D, Electronics Department of NRI.Engg.College, Patel Nagar for their support and Encouragements, and Electronics Department of NRI Engg.College, Patel Nagar for given testing and development facility for this work.

REFERENCES

[1] Ahmed H. Reja "Study of Micro Strip Feed Line Patch Antenna", Antennas and Propagation International Symposium, vol. 27, pp. 340-342 December 2008.

[2] Sahntanu Kumar Behera and Y. Choukiker, "Design and Optimization of Dual Band Micro Strip Antenna Using Practicle Swarm Optimization Technique," in Springer Science+Business Media, LLC, pp. 1346-1354, 2010.

[3] A. A. Deshmukh and G. Kumar, "Compact broadband gap-coupled shorted L-shaped microstrip antennas," in IEEE Antennas and Propagation International Symposium, vol 1, (Baltimore, Maryland), pp. 106–109, IEEE, July 2001.

[4] Z. M.Chen and Y.W.M. Chial, "Broadband probe-fed L-shaped plate antenna," Microwave and Optical Technology Letters, vol. 26, pp. 204–206, 1985.

[5] K. F. Lee, K. M. Luk, K. F. Tong, Y. L. Yung, and T. Huynh, "Experimental study of the rectangular patch with a U-shaped slot," in IEEE Antennas and Propagation International Symposium, vol. 1, (Baltimore, Maryland), pp. 10–13, IEEE, July 1996.

[6] S. C. Gao, L. W. Li, M. S. Leong, and T. S. Yeo, "Design and analysis of a novel wideband microstrip antenna," in IEEE Antennas and Propagation International Symposium,vol. 1, (Boston, Massachusetts), pp. 90–93, IEEE, July 2001.

[7] M. Khodier and C. Christodoulou, "A technique to further increase the bandwidth Of stacked microstrip antennas," in IEEE Antennas and Propagation International Symposium, vol. 3, (Salt Lake City, Utah), pp. 1394–1397, IEEE, July 2000.

[8] A.Shackelford, K. F. Lee, D. Chatterjee, Y. X. Guo, K. M. Luk, and R. Chair, "Smallsize wide bandwidth microstrip patch antennas," in IEEE Antennas and Propagation International Symposium,vol. 1, (Boston, Massachusetts), pp. 86–89, IEEE, July 2001

[9] F. Yang, X. -X. Zhang, X. Ye, and Y. Rahmat-Samii, "Wide-Band E Shaped Patch Antennas for Wireless Communications," in IEEE Trans. Antennas Propagation, vol. 49, no. 7, pp. 1094-1100, July. 2001.

[10] K. -L. Wong and W. -H. Hsu, "A Broad-Band Rectangular PatchAntenna with a Pair of Wide Slits," IEEE Trans. Antennas Propagation, vol. 49, no. 9, pp. 1345-1347, Sept. 2001.

[11] Tong K.F., Wong T.P.: "Circularly polarized U-slot antenna", IEEE Trans. Antennas Propagation, 2007, 55, (8), pp. 2382–2385 .

[12] Salonen P, "Dual-band E-shaped patch wearable textile antenna,"Antennas and Propagation Society International Symposium, 2005 IEEE. 2005; 1A:466,469 Vol. 1A.

[13] Murad NA, "Microstrip U-shaped dual-band antenna. Applied Electromagnetics,", 2005 APACE 2005 Asia-Pacific Conference on. 2005:4 pp..

[14] Tong K.F., Wong T.P.: "Circularly polarized U-slot antenna", IEEE Trans. Antennas Propagation, 2007, 55, (8), pp. 2382–2385.

[15] Tanaka T., Houzen T., Takahashi ITO K, "Circularly polarized Printed antenna combining slots and patch", IEICE Trans. Communication., 2007, E90-B, (3), pp. 621–628.

[16] M. T. Islam, M. N. Shakib and Norbahiah Misran,"High gain microstrip patch antenna", European Journal of Scientific Research,vol.32 No.2, pp.187-193,2009.

[17] C.A. Balanis, "Antenna theory", John Wiley, 1982, pp 727-734.

[18] Nasimuddin Z.N. Chen "Aperture coupled asymmetrical c-shaped slot microstrip antenna for circular polarization", IET Microwave Antennas Propag. , Vol. 3, Iss. 3, pp. 372–378, 2009.

[19] Shivnarayan & Babu R Vishvakarma "Analysis of notch-loaded patch for dual-band operation", Indian Journal of Radio & Space Physics.Vol.35, pp.435-442.

[20] Mohammad A. A. Subhi H. Ahmad A. K. and Juma S. M. "Cavity model analysis of rectangular micro strip antenna operating in TM03 mode", IEEE proc. pp. 0-2218-2223, 2006.

[21] S. K Satpathy, Vijay Srinivasan, K P Ray and G Kumar, "Compact microstrip antennas for personal mobile communication", IEEE proc.pp. 245-248, 1998.

[22] Kuo, J.S. and K.L., Wong, 2001. "A compact microstrip antenna with meandering slots in the ground plane", Microwave and Opical Technology Letters 29(2), pp. 95-97.

[23] Sze, J.Y. and K.L., Wong, 2000. "Slotted rectangular microstrip antenna for bandwidth enhancement", IEEE Transactions on Antennas and Propagation 48, pp. 1149-1152.

[24] Targonski, S.D., R.B., Waterhouse, and D.M., Pozar, 1998. "Design of wide-band aperture stacked patch microstrip antennas", IEEE Transactions on Antennas and Propagation 46(9), pp. 1245-1251.

AUTHORS PROFILE

Dr. Anubhuti khare (BE, MTECH, PHD) working as a Professor in Electronics and communication department UIT RGPV, Bhopal (M.P.).
Email:anubhutikhare@gmail.com
Phone no: 09425606502,

Rajesh Nema (BE, MTECH, PHD Pursuing) Working as a Assistant Professor in Electronics and communication department NIIST ENGG College Bhopal. The degree of B.E. secured in Electronics and Communication engineering. He secured M.Tech in Electronics and Communication engineering MANIT University. He is currently pursuing PHD in Electronics and Communication engineering
Author Address: Rajesh Nema E7/128 Ashoka society arera colony Bhopal M.P India Pin code 462016
Email: rajeshnema2010@rediffmail.com
Phone no: 09893216819

Miss Neenansha Jain (BE, MTECH (P)) MTECH student in Electronics and communication department NIIST ENGG College Bhopal(M.P.). She is secured of B.E. in ECE from Shravanabelagola Karnataka, VTU University India in 2007.
Author Address: Neenansha Jain A-157 Indrapuri, J.K.Road, Bhopal M.P India Pin code 462021
Email : neenanshajain2011@gmail.com
Phoneno: 09406582186

ICT for Education

A Platform for Modernization of computer science teaching methods in secondary schools in Cameroon

Marcellin Nkenlifack1, Raoul Nangue2, Bethin Demsong3, Victor Kuate Fotso3
[1]LAIA - IUTFV - University of Dschang, LIMMS - National Polytechnic, Cameroon
[2]LAIA - IUTFV of Bandjoun - University of Dschang
[3]IUTFV of Bandjoun - University of Dschang
Email: marcellin.nkenlifack@gmail.com

***Abstract*—This paper presents the modeling, design and implementation of a learning platform in Cameroon. This platform contains structured knowledge acquisition modules as well as teaching, learning and assessment modules to promote a constructive learning. The objective of this project is to show how ICT can be used to improve teaching and learning with modern digital tools. This will result into enabling teaching of the official computer curriculum in secondary schools at the national level. A second objective of the project is to put in place a set of ICT based administrative and managerial tools that will guide the day-to-day activities of a secondary school. This will aid in generating informations accessible from all levels of the educational system and all its related departments and partners. This project will aim at promoting ICT accessibility in the educational system while contributing to reducing the digital divide.**

Keywords-Elearning, ICT; Management Tools; Secondary school; Platform

I. INTRODUCTION

The new planetary economy is imposing a new type of education. Thus, the installation of platforms for the teaching of computer science in schools is not the matter of mode but it presents a quite range of advantages. This issue in tackled worldwide and researches are been carried on [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]; but the real question that raises is "which platform should be used for which pedagogy and by which pedagogues" [3] ? From studies and experiences [1], [14], [15], [16], it is clear that the interests of learners and parents is visible when teaching techniques and strategies impart not only knowledge but greatly increases chances of success in official exams. Henceforth ICT for education becomes a vital teaching and learning platform worth in secondary and high schools in Cameroon. This new method will help cultivating group work to the learners' right from childhood. It presents numerous advantages such as harmonization of official curricula nationwide, availability of statistics and decision making tools inter alia.

This project has been the focus of all education strategies in our country, be it, ministerial, national or regional. It will contribute to the consolidation of education as the target mission of the state as well as it will improve on the quality and outcomes of our education policy. The target population for this method is made up of students, teachers, teaching staffs and decision makers.

This project goes beyond professionalization of teaching – learning strategies, standardization and harmonization of teaching methods as well as its evaluation to tackle areas such as:

- Providing operational pedagogic teaching tools like teacher's guides textbook, exercises for self evaluation, examination type exercises

- Providing e-learning tutorial learning or evaluation

- Providing a teaching and learning content enriching interface

- Providing at the level of the platform managerial tools for staffs.

Technical solutions are put in place in pilot schools gradually as they occur.

The issue tackled in the present article will have applications and a real impact on the development of applied research. The work is organized in six steps: We are respectively going to discuss on the context in section 2, the methodology in section 3, the expected results in section 4 and the perspectives in section 5 before concluding in section 6.

II. CONTEXT

The non existence of a legal provision is generally the first difficulty encountered in development technology in Africa that is a lack of adequate rules and regulations. Researches have proven that there is a favorable national and international context. This can be seen through a range of legal texts meant to implement and reinforce the teaching of ICT in training schools. The focus on the institutional framework and operational input of the project discussed on the following line will enlighten us on this context.

A. Institutional Framework

- The 18th January 1996 constitution of the republic of Cameroon states that every citizen has the right to education.

- Relying on the 1995 Estates General on Education, The law N°98/004 of 14th April 1998 states the directives of

education in Cameroon and gives priority to basic education and new methods of teaching.

- The 27 January 2007 agreement signed between Cameroon government and an international structure (Cisco Networking Academy) emphasizes on training specialists in ICT. It states that secondary schools should provide online notes to their students; that a high frequency backbone platform on IT should be developed; that a taskforce of Network Academies should be realized. The members of this taskforce should be the representatives of the Ministries in charge of Education, those of the Ministry in charge Telecommunication, those of the Ministry in charge Research as well as international organizations such as PNUD ("programme des Nations unies pour le Développement"), UIT ("Union Internationale des Telecommunications"), FDNUF ("Fond de développement des Nations Unies pour les Femmes") and USAID.

- The decree N° 2002/004 of 04th January 2002 creating in the Ministry of Secondary education (MINESEC) the general Inspectorate of Pedagogy in charge of teaching computer science in secondary schools.

- Ministerial order N°3745/P/63/MINEDUC/CAB of 16th June 2003 making the teaching of computer science compulsory in secondary education.

- The directives of the Minister of Higher Education (MINESUP) assigning Universities and Internet Academies the mission of patroning secondary and high schools in mastering ICT, precisely in terms of teachers' training in computer science, defining the training curricula conform to the norm or technical assistance in multimedia centers.

- Minister of Secondary Education texts and circulars related to rules and regulations in training in computer science in Secondary and High schools in Cameroon.

B. Operationnal Input of the Project

This project will facilitate the learning process and enhance creativity to the learners. It will also give equal chances to a maximum of learners to have access to knowledge through new methods of teaching based on ICT and facilitating ipso facto their professional insertion. It intends to provide lessons based both on the rigor of international curricula and official national program of teaching computer science, on standardization and harmonization of the methods of teaching this science, the use of new forms of teaching using an interactive or multimedia technique. This technique integrates operational pedagogic teaching tools like teacher's guides textbook, exercises for self evaluation, examination type exercises; an objective type of evaluation relying on a powerful evaluation management system, a progressive content frequently updated in order to take into consideration technological development and a wide opening in terms of bibliography. Thus it requires a specified methodology.

III. METHODOLOGY

A. Process

Sensitization is placed on a high table in our process of putting in place this policy. It is worth for the start of the training of the target population made up of pupils, students, teachers and leaders on the interest of ICT in acquiring or transmitting knowledge and decision making in their various schools.

We have at the core of our development a platform. The implementation of this platform follows a set of steps to be followed in relation to whether the school possesses computer laboratories or not. Many referential documents related to this issue have been elaborated. We can name a few:

* For schools already equipped:

- A technical forms indicating equipments available in the school to be filled by the school,

- The leveling at the required standard of leaders

- Elaboration of a guiding plan

- Signing of an agreement the engages the two parties

- Training of main trainers and school leaders or administrators

- Implementation of the platform in the school and training of local learners

- Extension to other training domains such as science, mathematics, languages, etc…

*For non equipped schools

- Putting in place of computer science laboratory with a network

- Connection of administrative and pedagogic leaders with the network

- The following up of steps described for equipped schools as mentioned above.

B. Components of the Project

Table 1 presents the different components of the project with expected results and realization indicators as well while table 2 indicates the physical display programming of works, equipments and services concerned.

TABLE I. COMPONENTS & COAST OF THE PROJECT

N°	Components	Realization	
		Results	Indicators
1	institutional and structural support, putting in place and rehabilitation of equipments, training of trainers and administrators recruited	- multimedia Centers rehabilitated or constructed - training of trainers and administrators	- Number of multimedia centers rehabilitated or constructed - Number of trainers - Number of platform administrators recruited and formed
2	Development of the Platform "ICT for Education" Development of the management software called "SIGES"	- functional Platform with pedagogical programs and evaluation modules - SIGES software realized and validated	- Functionalities and conviviality platform - Number of applicative strata offered - Training on elaborated Modules - SIGES software developed and delivered
3	Display, creation and permanent enrichment of contents	Basis of knowledge and Scenarii created	-number of schools -Quality and number of documents and media -Number of scenarii - Number of tests
4	Technical and scientific research projects on education applied to ICT	Research modules to facilitate teaching	-number of research projects positively competed and evaluated -availability of the cooperative and piloting platform
5	Management of projects, local personnel and evaluation mission Consultations	- diffusion of information and management of resources - missions maint for actors -control and evaluation Structures	- Diffusion of budgetary packages - assessment of activities - Missions defined and texts diffused - Number of technical controls made

TABLE II. PHYSICAL PROGRAMMING OF THE PROJECT DISPLAY

Phases & works to be carried on	2007-2008	2009	2010	2011-2012	2012-2013
works	Studies of the context and institutional Framework	Studies of technical and organizational solutions Installation of basic equipments	Display equipments and solutions Beginning of experimentation Training	Installation of equipments Experimentation Assessment and preparation of a plan of over generalization	Installation of equipments Progressive Generalization and follow up training
Equipments	RAS	Basic equipments for development Passive equipments of the pilot sites	Active equipments and internet connection of pilot sites	Passive equipments of other sites	Active equipments and internet connection of other sites
Consultants/studies Services	Studies of Context and institutional framework	Prior studies of technical and organizational solutions	Research and realization of scenarii and basic modules	Validation and feedback continuation of researches in other modules	Continuation of development of complementary modules

C. Actors of the System

The system presents five main groups or categories of actors. The table 3 that follows illustrates the role played by each group of actors.

- Teachers' role is to create evaluations and provide solutions on the platform.

- Students listen to the lectures on line and can discuss among themselves or with the lecturers.

- Leaders on their part make consultations on statistics or use the school management module, referred to by its French acronym SIGES ("Système Intégré de Gestion des Etablissements Secondaires/Supérieurs").

- The platform administrators take care of parameters and maintenance of the platform.

TABLE III. DIFFÉRENTS ACTORS OF THE SYSTEM (FROM INTRANET/INTERNET)

Categories of actors	Activities	Comments
Teachers as producers Teachers as tutors	-pedagogical follow up - Tutorial-evaluations on line - putting multimedia contents on line	
Researchers developers	- Creation/Optimization of Solutions - Conception of scenarii	Project development Laboratories
Platform administrators	Administration and Maintenance of the system	Multimedia rooms (high schools and colleges)
Learners on the scene	- follow up of lectures on line - Discussion with tutors and other learners	Multimedia rooms (high schools and colleges or at home
Distants learners Professionals from entreprises Populations and Families	- follow up of lectures on line - Discussion with tutors and other learners	Offices Residences Cybercafés etc.

D. Technologies used and global architecture

We relied on a set of techniques and tools offered and already demonstrated by some e-learning and web service

authors [1] [16], [6], [2], [8], [14], [7], [9]. We are putting into action a set of applications integrated in the platform called ICTE "ICT for Education". The one line teaching environment combines the two categories of tools generally found in ICTE i.e synchone and asynchone tools, notably with:

- Web access to information available such as lectures backups, Workshop sheets, Labs, numerical animations and videos

- Access to powerful tools integrated meant for collaboration, sharing and classical communication

- A shared working panel, announcement space, messaging, etc.

Our strategy is mainly based on implementation and display of applications using free softwares or open sources. This enables to put the project into practice and it guaranties:

- Secured Linux servers (level of access, FireWall, Backup)

- Development tools such as Apache-MySQL-PHP, LDAP, WebMail-IMAP-SMTP, etc…

- A files server (SMB) which secures the stockage of teachers and learners directories

- Resources are stocked in the platform based on the kernel claroline. Its code has deeply been modified in order to adapt it to our needs. These modifications include inter alia evaluation module, multiple connection on the platform, the chat, harmonization of its genuiness right away from LDAP directory whose role in to centralize in put resources, inter operability and portability cap city of the system. The platform of resources hosts various back ups on various formats.

The type of the architecture is client-server. It respects the "3-tiers" model inspired from [3], [4], [20]. The first stratum is the user interface while the second layer is the applicative stratum and the third is the data stockage layer.

The client enters in the server through the network (Intranet) using its navigator.

The lines that follow briefly present the sub-system that build up the ICT for Education system:

- The Distant Learning System : it offers a resources library to pupils and students i.e multimedia backups, exercises and didactic materials

- The Learning Management System" or SIGES: it helps learners to be registered on line, to consult their training program and their results on line, to submit and follow up the claims on their results on time, to generate automatically their documents such as school ID carte, their transcripts and school certificates, to consult any statistics on their school population and all types of infrastructures available as well as their functionalities. The SIGES system is not the concern of this article.

- The mail & SMS system manages communication data between various actors. Messages are automatically dispatched. The documents managed by the SIGES sub-system are also transmitted by e-mail. The SMS module generates alerts sent to tutors and learners automatically. (eg: submission of a document)

- The users of the global system can use the same account to login to every service offered by ICTE. This is made possible thanks to the LDAP (Lightweight Directory Access Protocol) universal directory.

It is evident that the access and exchanges between various modules and services make use of a set of footbridge that serve as « software bus ».

E. Modeling

Our system is modeled using the UML language [17], an universal language that enables to give a model, then clearly and readably brings out the different static and dynamic elements thanks to powerful objects oriented tools or models used. We are simply going to an extract of this modelisation; details are presented by other authors [18]. Figure 1 presents the use cases diagram of platform.

We observe that a student can participate to practical works (Labs) and evaluations (quiz and exams) or can consult the notes after login the system. Similarly, a teacher can lecture, can create and organize TP and quiz. The ICT pedagogic coordinator administers the system meanwhile the school principal can consult statistics and other elements as well.

Figure 2 presents on of the sequence diagrams describing an exchange of information between actors of the system.

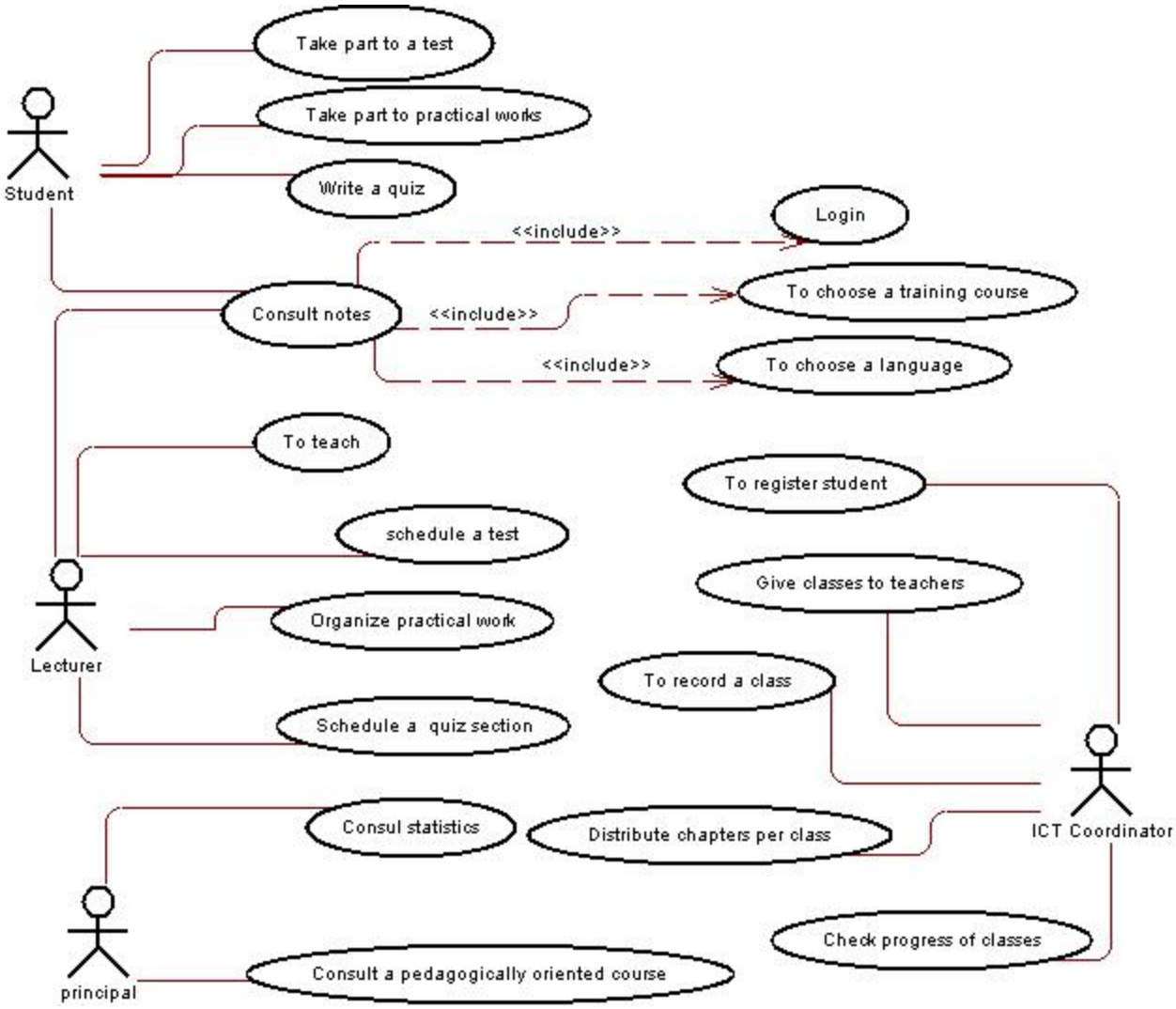


Figure 1. Use Cases Diagram of the platform

This diagram of figure 2 describes the distribution of activities between actors of the system. From the diagram, we can see that before any operation, the ICT coordinator much first of all parameter the system to open access to other actors to various the activities to be carried on. When this methodology is rigorously followed, the outcomes appear very clearly.

IV. RESULTS

A. *Presentation of results*

This project which is of national scale, offers a framework of exchange and online training to our students and teachers. It enables to extend training to a larger population. It can be a possible solution to the lack of teachers in our schools. Beyond that, trainers can henceforth play their role independently to their geographical position. The technology put in place facilitates the updating of contents with a larger opening to re usage. We are gradually going to proceed through sensitization of potential target afar population via media and foldouts or initiation and implementation of school and university days dedicated to ICT.

We can point out many advantages offered by the display of ICTE, namely:

- More famous to schools using the platform

- Professionalization of trainings offered and better professional insertion of graduates since emphasis are put to practical works

- Standardization and harmonization of computer science official programs in high schools and colleges in Cameroon.

- Increase of chances of success of learners at computer science official exams

- Availability of tools for a quality control for leaders at the level of the training platform

- Easy acquisition and maintenance of computer material through the variety of network partners

- Progressive evolvement towards the ratio of one computer for one student

- Generalization of the usage of ICT through management tools (registration, administration, library management, etc.)

- Training of 150 teachers from different schools in using ICT to teach computer science on line (1st phase of the project)

- Display of the training platform in 100 schools (1st phase)

- Production and distribution of 500 000 self-learning CDROMs and DVD (1st phase)

- Integration of a multilingual system in the training platform

- Development and possession of e-learning habits to learners

- Production of operational pedagogic tools such as teaching guides, Labs, individual rehearsal exercises, evaluations, etc.

- A guarantee of the progress of contents for, they are constantly updated.

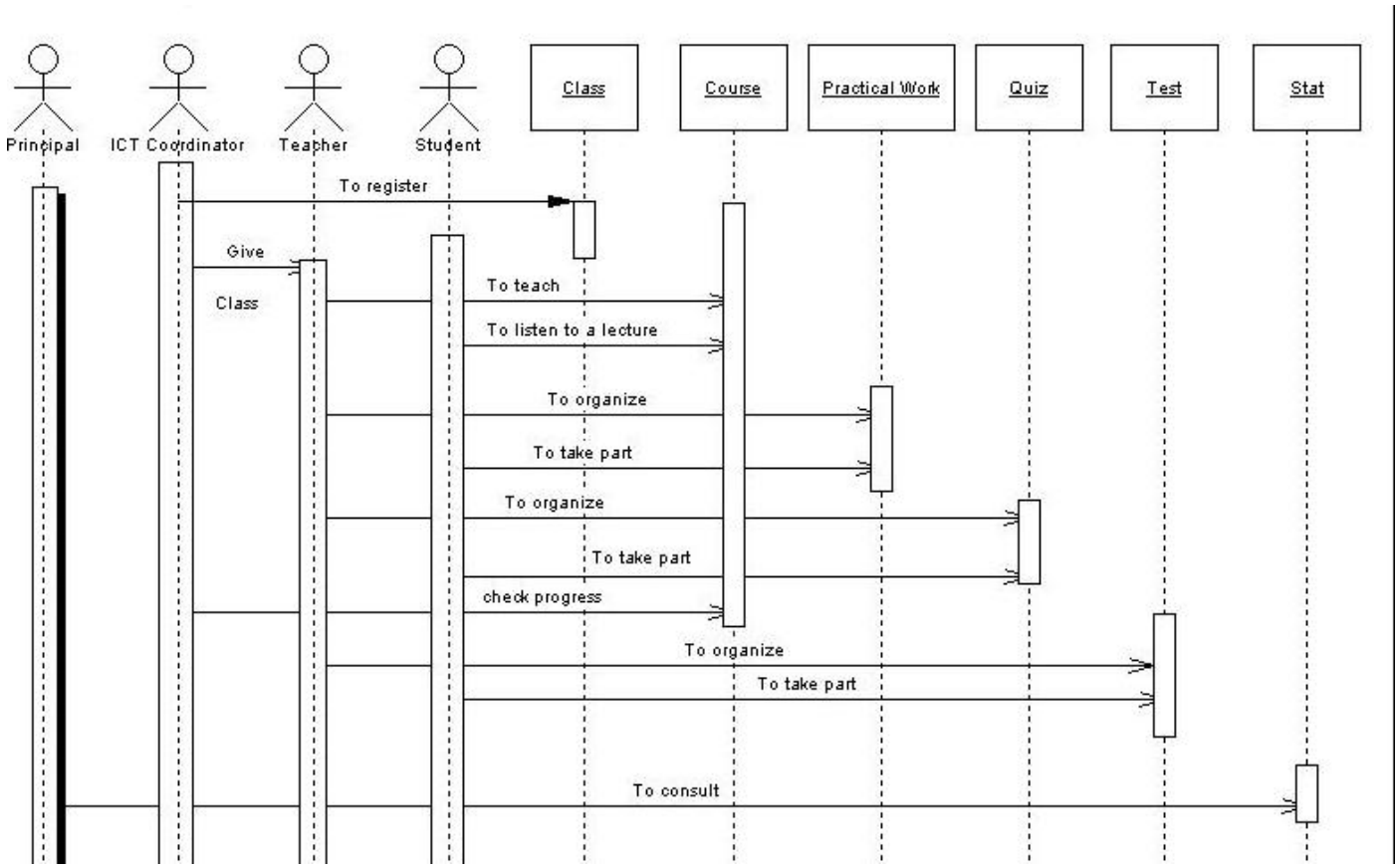


Figure 2. Sequence Diagram of activities on the platform

Figure 3 presents one of the pages of indexes integrated in a lecture to fluctuate evaluations.

Figure 3. Page example for loading Assesment

Figure 4 is a capture of an evaluation screen in the form of a quiz in the platform. However, it is worth noticing that evaluations are many and can bare various forms.

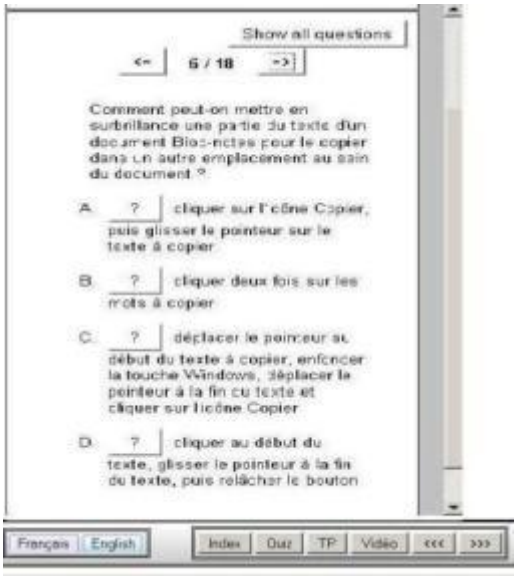


Figure 4. Example of evaluation in the form of a Quiz

The interface that presents this section of a lecture is described in figure 5. This interface shows a section of a lecture with an image associated. The image is indicated in the text by the key "1". In the right window, a click on a key (if they are many keys) enables to load the image required.

Figure 5. Example of interface for the presentation of a section of a lecture

As the management tool is concerned, we have developed the SIGES application which is a computer system based on an architecture distributed and adapted to the management of academics as well as other services via network (Intranet or Internet). As example, figure 6 shows the situation on payment of school fees for a section. This situation enables to have a global overall view of students who have paid or not at that section.

Figure 6. Example of the situation of payment of school fees after administrative registration in SIGES

The ICTE project has a number of components and should progressively be implemented all over the country in its subsequent phases.

Table 4 presents a quantity of expected outputs as well as expected results at the end of the first phase of the project.

TABLE IV. EXPECTED OUTPUT AND EXPECTED PRODUCTS

Main outputs (goods or services)	Quantity or Values	Target Beneficiary			
		Male	Female	youth	others
Training of trainers and leaders	900	300	300	300	
Training of the platform administrators	300	100	100	100	
Production and distribution of CDROMs and DVD	500 000	150 000	150 000	200 000	All the regions
Display of the ICTE platform	300 schools with (30 during the first experimental)	Total access within the school	Total access within the school	Total access within the school	And average of 30 per region (high schools and Colleges)
Realization of official piloting site and administration of the entire system	Server accessible via Internet or Intranets put in place (Delegations, Ministries)	unlimited for leaders	unlimited for leaders	unlimited for leaders	
Support of research through mobilization of many researchers	Many teams Planned at different levels of the project	10 researchers from different disciplines in ICT and Sciences of education	5 researchers from different disciplines in ICT and Sciences of education	30 developers And integrators	1 team for piloting and tens of corresponding consultants

As shown on table 4, the project will mobilize teams of researchers put into action to promote training of all social layers to ICT and will also contribute to the development of the country.

B. *Projections : Deployment in schools*

The envisaged projections of a progressive putting in place of the project in the various regions of Cameroon are presented in table 5.

One can remark a global spread in 6 years to cover a great part of the national territory during the two years of the first phase.

TABLE V. PROJECTIONS OF DEPLOYMENT OF THE SYSTEM IN THE COUNTRY

	Number of schools where the platform is deployed	
		Potentiels
October 2008	1	15
October 2009	2	15
October 2010	3	30
October 2011	10	45
October 2012	30	60
October 2013	90	150

V. PERSPECTIVES AND POSSIBLE EXTENSIONS

This project will strengthen or contribute to a better promotion of what is granted in the domain and locales predispositions towards emergency and development of technology.

Considering the techniques put in action or envisaged, the project opens new perspectives with chance of generalization of the usage of ICT through the development of a complementary management tools in other domains where the mastery of the usage of technologies displayed (health, agriculture, virtual libraries, administration, personnel and salaries, etc.).

From our experience, many perspectives are offered:

- We have started by working on physical identification of the students who can be connected to the platform and participate at evaluations through videoconference on the system. Due to the difficulties that poor countries meet because of the poor standard of bandwith. We can equally envisage a synchronous system using simplified videos. This system should make an intelligent choice of images to be posted taken them either from a data base with fixed predefined images or by making a selecting of a number of images that varies per second to make an on time light animation from a numerical camera. It is the problem of adaptation of the video flux to the bandwidth in a dynamic way. At this point, this article [10] could serve as a base.

- Considering the scarcity of teachers and the number of the students on roll that increases every day, we ought to integrate and possess the MLearning [10], [12] in order to bring closer all the actors of the system or between themselves with alerts indicating either training or tutorial ship and management.

VI. CONCLUSION

In this article, while presenting our ICT for Education platform which is progressively taking place in Cameroon, we have intended to confirm a said [19], [21] that online teaching is inevitably an interesting solution to the problem exhibited by various traditional methods of teaching. The said platform brings a better interaction and a great number of possibilities of follow up of students independently to the time and place where the actors are situated.

Finally, this project will ensure a systematic usage of ICT to master computer science and one line teaching in a greatest number of schools and training centers in Cameroon.

Actually, related researches are normally evolving and will be the focus of our next articles. All the TDR (“Termes de Référence”) and various studies considered necessary have been realized:

- Study of impact,

- Technical studies,

- Studies of risks,

- Studies on economic impacts,

- Studies and financial programming of works (this works will not be presented in this article).

The elaboration of the remaining offer file (DAO - "Dossier d'Appel d'Offre") is on.

Concerning the progress of realization:

- Every remaining task has been planned

- Teams of correspondents are in action

- Operational teams are also in action

- Teams of trained researchers increase for, it is a matter of an important project that requires a variety of competences pluridisciplinary human resources. The results obtained are gradually integrated in the platform.

This platform is actually been used in a number of schools in the West region of Cameroon, precisely in "Government High School Baham" (*http://www.cisco-ra-iutfv-online.net/LyBiBah*), "Government High School Bafang", "Government High School Bayangam" and "Collège Saint Thomas de Bafoussam". It actually enables students to:

- Personally learn their notes in computer science without needing a help of a teacher

- Self-evaluate their level of understanding by answering related questions provided by the system

- Share their knowledge with other through an integrated forum in the system

The project offers good opportunities and perspectives, namely:

- Promotion of social, human, economic and technological development

- Fight against poverty through progress using ICT and new methods of teaching

- Develop our chances of establishing our integration in the global world by reducing the gap between rich and poor countries

One of our trump cards that we play in this project is our open source softwares. They give us:

- A better Independence (technical, financial aspects, etc.)

- A great ease in displaying

- More chances in terms of security tools

- More facilities to update (GPL) the tools used

- More facilities of maintenance

With this we are convinced that this project will enable overcoming a great number of challenges in future. It will also insure a long lasting development with a mastery of ICT in the sense of the millennium development objectives.

ACKNOWLEDGMENT

This project has been conceived thanks to the expertise and support, be it direct or indirect of a number of partners and contribution of many structures already associated or to be associated to the project, notably :

- French Cooperation through the project COMETES ("Coordination et Modernisation des Etablissements Technologiques de l'Enseignement Supérieur"), since 2004

- AUF (Agence Universitaire de la Francophonie),

- PNUD (Programme des Nations Unies pour le Développement),

- MINESUP (Ministère de l'Enseignement Supérieur du Cameroun)

- MINESEC (Ministère des Enseignements Secondaires) à travers ses Délégations Régionales et Départementales

- MINRESI (Ministère de la Recherche Scientifique et Technique)

- MINPOSTEL (Ministère des Postes et Télécommunications)

- University of Dschang

- IUTFV Bandjoun (Internet Academy)

- Research Laboratories: LIMSS (National Polytechnic), LAIA (IUTFV).

REFERENCES

[1] Nkenlifack M., Nangue R., Noulamo T. and Kwonche A. (2009), "Les TICE au service de la Formation Ouverte à Distance à l'Université de Dschang : Implémentation de SIEL (Système Intégré d'Enseignement en Ligne basé sur Internet)", Journal Langue et Communication, N° 07 novembre 2009, Revue scientifique internationale de recherche multidisciplinaire, ISSN 1560-3407

[2] Talla N., Tonye E., Dipanda A. and Ewoussoua L. (2010), "A model of Distance Learning of Technologies for Developing countries: Case of the Master (M2) in Telecommunications at the National Advanced School of Engineering in Cameroon", 10th African Conference on Research in Computer Science and Applied Mathematics CARI'2010, Côte d'Ivoire, Yamoussoukro, October 18 – 21, 2010

[3] Boyom S., Essome S., Takoudjou A. and Kamdem D. (2005), "Campus virtuel: exploit technologique mais pour quelle pédagogie, par quel pedagogue ?", in Akono A., Tonye E., Dipanda A., Kokou Y. (ed.), Proc. Int. Conf. On Signal & Image Technology and Internet Based Systems, IEEE SITIS'2005, Yaounde, Cameroon, ISBN 2-9525435-0.

[4] Manlescu I., Brambilla M., Ceri S., Fraternali P. (2005), "Model Driven Design and Deployment of Service-Enabled Web Applications", ACM Transaction on Internet Technology, August 2005, Vol. 5, Number 3.

[5] Koum G., Yekel A., Tampolla S. and Sanbong T. (2005), "Vocal Interaction in Web User Interface involving natural Language Processing", in Akono A., Tonye E., Dipanda A., Kokou Y. (ed.), Proc. Int. Conf. On Signal & Image Technology and Internet Based Systems, IEEE SITIS'2005, Yaounde, Cameroon, ISBN 2-9525435-0.

[6] Drissi M. and Talbi M. (2009), "Dispositif de la formation à distance pour préparer les étudiants universitaires marocains à suivre des cours scientifiques en français – FOSEL (français sur objectifs spécifiques en ligne)", Revue africaine de didactique des sciences et des mathématiques, Numéro 4, 15 décembre 2009 : http://www.radisma.infodocument.php?id=687.

[7] Dahbi A., El-kamoun N. and Berraissoul A. (2009), "Conception d'un système hypermédia d'enseignement adaptatif centré sur les styles

d'apprentissage : modèle et expérience", International Journal of Technologies in Higher Education, vol 6 (1), 2009, PP 55-71.

[8] Cacheux C. (2009), "Analyse des usages des espaces numériques de travail dans l'enseignement secondaire usages prescrits : adhesion ou resistance des usagers ?", Journal ISDM n°37-2010 spécial Numérique (s) : Défis, Enjeux et Perspectives, Actes du VIe Colloque Jeunes Chercheurs Praxiling, Montpellier, les 25 et 26 juin 2009.

[9] Kandel R. (2010), "De la conception des supports informatiques a la conception des dispositifs didactiques", Journal ISDM n°37-2010 spécial Numérique (s) : Défis, Enjeux et Perspectives, Actes du VIe Colloque Jeunes Chercheurs Praxiling, Montpellier, les 25 et 26 juin 2009.

[10] Trifonova A. and Ronchetti M. (2004), "A General Architecture of M-Learning", Journal of Digital Contents, Vol. 2, Issue1, 2004, ISSN: 1696-313X

[11] Raynaud J., Martel C., Villiot-Leclercq E., Gerbe O., Jullien J. and Camarero R. (2009), "Pour un système intégré de gestion du processus d'éducation et de formation", Revue internationale des technologies en pédagogie universitaire, vol. 6 (2-3) 2009, www.ritpu.org.

[12] Wishart J., McFarlane A. and Ramsden A. (2005), "Using Personal Digital Assistants (PDAs) with Internet Access to Support Initial Teacher Training in the UK", mLearn 2005, 4th World conference on mLearning, Cape Town, South Africa, 25 –28 October 2005. Available: http://www.mlearn.org.za/CD/papers/Wishart.pdf

[13] Assude T., Bessieres D., Combrouze D. and Loisy C. (2010), Conditions des genèses d'usage des technologies numériques dans l'éducation , Revue STICEF, Volume 17, 2010, ISSN : 1764-7223, mis en ligne le 10/072010, http://sticef.org

[14] Nkenlifack M., Nangue R., Tchokomakoua M. (2009), "Projet TICLAC : TIC pour la Modernisation de l'Enseignement des Langues et Cultures nationales dans les établissements, Conférence internationale : ASAP 2009 sur la "Diversité culturelle et Internet multilingue en Afrique", 2 au 5 décembre 2009, Hôtel Hilton, Yaoundé-Cameroun

[15] Nkenlifack M., Noulamo T. and Nangue R. (2006), "Contribution des TIC au Développement de la Formation Ouverte à Distance à l'IUT Fotso Victor de l'UDS : Déploiement de SIEL (Système Intégré d'Enseignement en Ligne basé sur Internet)", Proc Int. Conf. on Sustainable Engineering Development In Africa - African Solutions for African Problems -, June 4-8, 06, Yaounde, Cameroon.

[16] Fogue M. and Nkenlifack M. (2006), "Formation Ouverte à Distance : Nouvelle façon d'apprendre et d'enseigner, ETUDE DE CAS SUR LA DIVERSIFICATION DE L'ENSEIGNEMENT SUPÉRIEUR ET L'ADAPTATION AU MARCHÉ", Conf. Int., Thème "L'enseignement Supérieur au coeur des Stratégies de Développement en Afrique Francophone. Mieux Comprendre les Clefs du Succès", 13-15 Juin 2006, Ouagadougou, Burkina Faso : http://siteresources.worldbank.org/EDUCATION/Resources/278200-1121703274255/1439264-1137083592502/Presentation_IUT_Dschang.ppt

[17] Web site of the OMG (Object Management Group) - manuel de référence UML 2.0, //www.omg.org, (consulté en 2008).

[18] Owono A., Mbaya A., Guepi L., Kegninkeu A. and Guegang P. (2008), "Projet ICT for Education : Mise en œuvre du Curriculum ITS pour les nouvelles formes d'enseignement de l'informatique", Mémoire de fin d'études de DUT Informatique de Gestion, IUT FV de l'Université de Dschang, Cameroun, 2008.

[19] Abid-Zarrouk B. (2010), "Une analyse de l'efficacité interne des modes d'enseignements par correspondance, en présentiel et en ligne dans le cadre de la préparation au DAEU", Revue STICEF, Volume 17, 2010, ISSN : 1764-7223, mis en ligne le 19/07/2010, http://sticef.org

[20] Nkenlifack M., Noulamo T. and Nangue R. (2006), "Système Intégré d'Enseignement en Ligne basé sur Internet au service de la Formation Ouverte à Distance à l'Univ. de Dschang", Actes Conf. Int. Euro Graduation @ccess, Yde-Cameroun, 23-24 mai 2006 : http://www.euro-graduation-access.org

[21] Mvoto C. (2010), "Appropriation des innovations dans les écoles normales supérieures : une étude des besoins, des avantages et contraintes de l'intégration des TIC", frantice.net, Numéro 1 - juillet 2010. Récupéré du site de la revue : http://www.frantice.netdocument.php?id=125. ISSN 2110-5324

AUTHORS PROFILE

Dr. Marcellin Julius NKENLIFACK is the Head of Computer Science Department, Institute of Technology, University of Dschang, and an Associate Professor. He received M.A. degrees in Computer Science, followed by Ph.D. in Computer Engineering and Control from National Polytechnic Institute, University of Yaounde I. He had been a visiting researcher at Institut Scientifique et Polytechnique Galilee, Universite de Paris 13 (2001) and SUPELEC - Rennes, France (2003). He has published more than 20 papers in reputed International journals and Conferences in the field of Software Engineering, Computer Applications in Industry and Engineering, Object oriented Modeling and Simulation, Meta-modeling, UML, Hybrid Control Systems, Distributed Control, Computer in Education, E-learning. He has been awarded a Lecturer's award for 2003-2007 period for exceptional contributions towards developing e-learning at University of Dschang. He is an Editorial Board Member / Associate Editor / Reviewer of many internationals journals and conferences.

Raoul Calvain Nangue is teaching and researcher at Laboratory LAIA of the IUTFV of the University of Dschang. It is also Doctorant in the "Department off Information and Communication Technology, School off ICT, Nelson Mandela Metropolitan University" in South Africa.

Bethin Demsong *is a Head of Service of the General Affairs, IUT Fotso Victor of the University of Dschang, Cameroun, studying with the cycle of Master in Science of the Language at the University of Dschang,Cameroun. He is a French and English Languages teacher.*

Victor Kuate Fotso is working in the Data-processing Cell with the IUT Fotso Victor of the University of Dschang, it is titular of Bachelor in Computer Science and Network, and postulates with an inscription in Master in Computer Science.

Application Of Extended Kalman Filter For A Free Falling Body Towards Earth

Leela Kumari. B
Department of Electronics and communication Engineering
G.V.P. College of Engineering,
Visakhapatnam, India
Email:leela8821@yahoo.com

Padma Raju. K
Department of Electronics and communication Engineering
J. N.T.U.KAKINADA
Kakinada, India
Email:padmaraju_K@yahoo.com

Chandan .V.Y.V
Department of Electronics and communication Engineering
G.V.P. College of Engineering
Visakhapatnam, India
Email:chandanvyv964@gmail.com

Sai Krishna. R
Department of Electronics and communication Engineering
G.V.P. College of Engineering
Visakhapatnam, India
Email: rskrishna89@gmail.com

V.M.J. Rao
Department of Electronics and communication Engineering
G.V.P. College of Engineering
Visakhapatnam, India
Email: vmjrao8217@gmail.com

***Abstract*— State estimation theory is one of the best mathematical approaches to analyze variants in the states of the system or process. The state of the system is defined by a set of variables that provide a complete representation of the internal condition at any given instant of time. Filtering of Random processes is referred to as Estimation, and is a well-defined statistical technique. There are two types of state estimation processes, Linear and Nonlinear. Linear estimation of a system can easily be analyzed by using Kalman Filter (KF) and is used to compute the target state parameters with a priori information under noisy environment. But the traditional KF is optimal only when the model is linear and its performance is well defined under the assumptions that the system model and noise statistics are well known. Most of the state estimation problems are nonlinear, thereby limiting the practical applications of the KF. The modified KF, aka EKF, Unscented Kalman filter and Particle filter are best known for nonlinear estimates. Extended Kalman filter (EKF) is the nonlinear version of the Kalman filter which linearizes about the current mean and covariance. The estimation can be linearised around the current estimate using the partial derivatives to compute estimates even in the face of nonlinear relationships.. The EKF has been considered the standard in the theory of nonlinear state estimation. This paper deals with how to estimate a nonlinear model with Extended Kalman filter (EKF). The approach in this paper is to analyze Extended Kalman filter where EKF provides better probability of state estimation for a free falling body towards earth.**

Keywords- Kalman filter; Extended Kalman filter; free fall body; apriori information.

I. INTRODUCTION

Filtering and estimation are two of the most important tools of engineering. Whenever the state of a system needs to be estimated from noisy sensor information, state estimator is employed to produce the best estimate of the true system state. When the system dynamics and observation models are linear, the minimum mean squared error (MMSE) estimate can be computed using the Kalman filter.

Control of any process modeling, obtained from a priori knowledge of certain observable parameters is standard practice for Engineers. For many of the applications simple models with linear approximations around a design point suffice the requirement. Since all the natural phenomena are non-linear, it is very important to study the nonlinear models and their control for the following reasons:

1) Some systems have a linear approximation that is non-controllable near interesting working points. Linearization is ineffective even locally for such cases.

2) Even if the linearized model is controllable one may wish to extend the operational domain beyond the validity domain into nonlinear region for better prediction.

3) Some control problems are external to the process and cannot be answered by a linearly approached model.

The success of the linear model in identification or in control has its cause in the good understanding of it. A better mastery of invariants of nonlinear models for some transformations is a prerequisite to a true theory of nonlinear identification and control. And all nonlinear systems are supposed to have a state space of finite dimension. State Estimation techniques are handled by filtering technique models for performance.

A common approach to overcome this problem is to linearize the system before using the Kalman filter, resulting in an extended Kalman filter. This linearization does however pose some problems, e.g. it can result in nonrealistic estimates [1, 2] over a period of time. The development of better estimator algorithms for nonlinear Systems has therefore attracted a great deal of interest in the scientific community,

because the improvements will undoubtedly have great impact in a wide range of engineering fields. The EKF has been considered the standard in the theory of nonlinear state estimation. This paper deals with how to estimate a nonlinear model with Extended Kalman filter (EKF). The approach in this paper is to analyze Extended Kalman filter where EKF provides better probability of state estimation for a free falling body towards earth.

II. Linear and Nonlinear Models

Kalman Filter (KF), Extended KF (EKF), Unscented KF (UKF) and Particle filter (PF) are models popularly used for state estimation.

The traditional Kalman Filter is optimal only when the model is linear. . The practical application of the KF is limited because most of the state estimation problems like tracking of the target are nonlinear. If the system is linear, the state estimation parameters like the mean and covariance can be exactly updated with the KF.

The EKF works on the principal that a linearized transformation is approximately equal to the true nonlinear transformation.

In this paper, EKF for State Estimation have been considered for their relative performance levels and to give an idea as to their applications with sample State Estimation case study.

A. State space models

A state space model is a mathematical model of a process, where *state* **x** of a process is represented by a numerical vector. State-space model actually consists of two sub models: the *process model*, which describes how the state propagates in time based on external influences, such as input and noise; and the *measurement model*, which describe how measurements **z** are taken from the process, typically simulating noisy and/or inaccurate measurements.

B. Linear State Space Model

A linear state-space model assumes the functions *F* and *H* are linear, in both state and input. The functions can then be expressed by using the matrices, *B* and *H*, reducing state propagation calculations to linear algebra. Overall this results in the following state-space model:

$$\mathbf{x}_k = F_k\mathbf{x}_{k-1} + B_k\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \quad (1)$$

$$\mathbf{z}_k = H_k\mathbf{x}_k + \mathbf{v}_k \quad (2)$$

Where

u is process input
w is state vector
v is measurement noise vector
k is the discrete time

The above expressions (1) and (2) govern state propagation and measurements respectively.

Linear model is easier both to calculate and analyze. This enables modelers to investigate properties such as controllability, observability and frequency response [11].

Linear state models are either based on inherently linear processes, or simply linearized versions of a nonlinear process by means of a first order Taylor approximation.

C. Nonlinear State Space Model

The most general form of state-space models is the Nonlinear model. This model does typically consist of two functions, *f* and *h*:

$$\mathbf{x}k = f(\mathbf{x}k\text{-}1, \mathbf{u}k\text{-}1, \mathbf{w}k\text{-}1) \quad (3)$$

$$\mathbf{z}k = h(\mathbf{x}k, \mathbf{v}k) \quad (4)$$

III. Extended Kalman Filter

A. Back ground- State estimation

State estimation concerns the problem of estimating the probability density function (pdf) for the state of a process which is not directly observable. This involves both predicting the next state (based on the current state) and applying corrections (based on measurement model).

Estimator**:** Estimator is a tool that predicts the future behavior of a model from the available information.

The Estimator uses knowledge about the evaluation of the variable, the probabilistic characterization of the various random factors and the prior information.

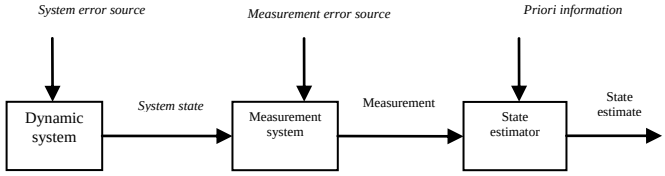


Figure 1: Mathematical view of state estimation

Different estimators:

- Recursive Bayesian Estimation
- Kalman Filter (KF)
- Extended KF (EKF)
- Unscented KF (UKF) and
- Particle filter (PF)

B. Recursive Bayesian Estimation(RBE) State Space

The most general form of state estimation is known as Recursive Bayesian Estimation [12]. This is the optimal way of predicting a state pdf for any process, given a system and a measurement model. RBE works by simulating the process, while at the same time adjusting it to account for new measurements z, taken from the real process. The calculations are performed recursively in a two-step procedure. First, the next state is predicted by extrapolating the current state onto next time step, using state propagation belief $p(x_k|x_{(k-1)})$ obtained from function f. Secondly, this prediction is corrected using measurement likelihood $p(z_k|x_k)$ obtained from function h, taking new measurements into account. Unfortunately, this method does not scale very well in practice, mainly due to the large state space for multidimensional state vectors. Calculating the prior probability of each point in this state space involves a

multidimensional integral, which quickly becomes intractable as the state space grows. Computers are also limited to calculation of the pdf in discrete point in state space, requiring a discretization of the state space. This technique is therefore mainly considered as a theoretic foundation for state estimation in general. Bayesian estimation by means of computers is only possible if either the state space can be discretized, or if certain limitations apply for the model.

C. Kalman filter

The problem of state estimation can be made tractable if we put certain constrains on the process model, by requiring both 'f' and 'h' to be linear functions, and the Gaussian and white noise terms 'w' and 'v' to be uncorrelated, with zero mean. Put in mathematical notation, we then have the following constraints (5) and (6).

As the model is linear and input is Gaussian, we know that the state and output will also be Gaussian [13]. The state and output pdf will therefore always be normally distributed, where mean and covariance are sufficient statistics. This implies that it is not necessary to calculate a full state pdf anymore, a mean vector ˆx and covariance matrix P for the state will suffice.

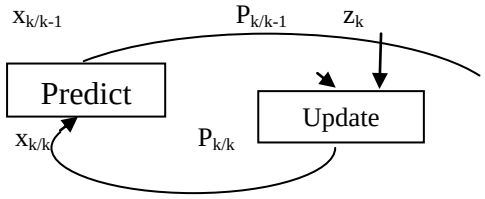


Figure 2: *Kalman filter loop*

The recursive Bayesian estimation technique is then reduced to the Kalman filter, where f and h is replaced by the matrices F, B and H. The Kalman filter is, just as the Bayesian estimator, decomposed into two steps: predict and update.

The Kalman filter is quite easy to calculate, due to the fact that it is mostly linear, except for a matrix inversion. It can also be proved that the Kalman filter is an optimal estimator of process state, given a quadratic error metric [14, 15].

Most processes in real life are not linear, and therefore need to be linearised before they can be estimated by means of a Kalman filter. So the practical applications of the KF are limited and so modified KF, aka EKF is generally used.

Different from KF, EKF deals with nonlinear process model and nonlinear observation model. In the extended Kalman filter, the state transition and observation models need not be linear functions of the state, but may be differentiable functions [4, 5, 6, 7, 8]. The nonlinear process model (from time *k* - 1 to time *k*) is described as

$$X_k = f(x_{k-1}, u_{k-1}) + w_{k-1} \quad (5)$$

$$Z_k = h(x_k) + v_k \quad (6)$$

where x_{k-1}, x_k are the system state (vector) at time k-1; k, f is the system transition function, u_k is the control, w_k is the zero-mean Gaussian process noise $w_k \sim N(0;Q)$, h is the observation function and v_{k+1} is the zero-mean Gaussian observation noise $v_{k+1} \sim N(0;R)$.

The function *f* can be used to compute the predicted state from the previous estimate and similarly the function *h* can be used to compute the predicted measurement from the predicted state. However, *f* and *h* cannot be applied to the covariance directly. Instead a matrix of partial derivatives (the Jacobian) is computed.

At each time step the Jacobian is evaluated with current predicted states. These matrices can be used in the Kalman filter equations. This process essentially linearizes the non-linear function around the current estimate.

D. Predict and update equations

Predicted state

$$\hat{x}^{-}_{k!k-1} = f(\hat{x}^{-}_{k-1!k-1}, u_{k-1}) \quad (7)$$

Predicted estimate covariance

$$P_{k!k-1} = F_{k-1} P_{k-1!k-1} F^{T}_{k-1} + Q_{k-1} \quad (8)$$

Updating state

Innovation(or residual) covariance

$$y^{\sim}_{k} = z_k - h(\hat{x}_{k!k-1}) \quad (9)$$

$$S_k = H_k P_{k!k-1} H^{T}_{k} + R_k \quad (10)$$

Optimal Kalman gain

$$K_k = P_{k!k-1} H^{T}_{k} S^{-1}_{k} \quad (11)$$

Updated state estimate

$$\hat{x}_{k!k} = \hat{x}_{k!k-1} + K_k y^{\sim}_{k}) \quad (12)$$

Updated estimate covariance

$$P_{k!k} = (1 - K_k H_k) P_{k!k-1} \quad (13)$$

Where the state transition and observation matrices are defined to be the following Jacobians

$$F_{k-1} = \frac{\partial f}{\partial x} | \hat{x}_{k-1|k-1}, u_{k-1} \quad (14)$$

$$H_k = \frac{\partial h}{\partial x} | \hat{x}_{k|k-1} \quad (15)$$

E. Continuous-time extended Kalman filter Model

$$x(t) = f(x(t),u(t)) + w(t),$$
$$w(t) \sim N(0,Q(t)) \quad (16)$$
$$z(t) = h(x(t)) + v(t),$$
$$v(t) \sim N(0,R(t)) \quad (17)$$

Initialize

$$\hat{x}(t_0) = E[x(t_0)], P(t_0) = Var[x(t_0)] \quad (18)$$

Predict-Update

$$\hat{x}(t_0) = f(\hat{x}(t_0), u(t) + K(t)(z(t) - h(\hat{x}(t_0))) \quad (19)$$

$$P(t) = F(t)P(t) + P(t)F(t)^T - K(t)H(t)P(t) + Q(t) \quad (20)$$

$$F(t) = \frac{\partial f}{\partial x} \mid x^{\wedge}(t), u(t) \quad (21)$$

$$K(t) = P(t)H(t)^T R(t)^{-1} \quad (22)$$

$$H(t) = \frac{\partial h}{\partial x} \mid x^{\wedge}(t) \quad (23)$$

Unlike discrete-time extended Kalman filter, the prediction and update steps are coupled in continuous-time extended Kalman filter [9, 10].

F. Continuous- discrete extended Kalman

Most physical systems are represented as continuous-time models while discrete-time measurements are frequently taken for state estimation via a digital processor. Therefore, the system model and measurement model are given by

$$x(t) = f(x(t),u(t)) + w(t),$$
$$w(t) \sim N(0,Q(t)) \quad (24)$$

$$z(t) = h(x_k) + v_k,$$
$$v_k \sim N(0,R_k) \quad (25)$$

where, $x_k = x(t_k)$

Initialize

$$\hat{x}_{0|0} = E[x(t_0)], P_{0|0} = Var[x(t_0)] \quad (26)$$

Predict

$$\begin{cases} \dot{\hat{\mathbf{x}}}(t) = f\left(\hat{\mathbf{x}}(t), \mathbf{u}(t)\right) \\ \dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}(t)^{\top} + \mathbf{Q}(t) \end{cases}, \text{with} \begin{cases} \hat{\mathbf{x}}(t_{k-1}) = \hat{\mathbf{x}}_{k-1|k-1} \\ \mathbf{P}(t_{k-1}) = \mathbf{P}_{k-1|k-1} \end{cases}$$
$$\Rightarrow \begin{cases} \hat{\mathbf{x}}_{k|k-1} = \hat{\mathbf{x}}(t_k) \\ \mathbf{P}_{k|k-1} = \mathbf{P}(t_k) \end{cases} \quad (27)$$

where,

$$F(t) = \frac{\partial f}{\partial x} \mid x^{\wedge}(t), u(t) \quad (28)$$

Update

$$K_k = P_{k|k-1}H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \quad (29)$$

$$\hat{x}_{k!k} = \hat{x}_{k!k-1} + K_k (z - h(\hat{x}_{k!k-1})) \quad (30)$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \quad (31)$$

where,

$$\boldsymbol{H}_k = \frac{\partial h}{\partial x} \mid \hat{\boldsymbol{x}}_{k|k-1}$$

The update equations are identical to those of discrete-time extended Kalman filter.

IV. Modelling Example for a Free Falling Body Towards Earth

Consider the problem of estimating various states of a free falling body towards earth. Range measurements are corrupted by additive Gaussian noise. The state estimation cannot be accurately explained by KF since nonlinearities are exhibited by forces that act on the body, and the measuring device is located at an altitude h and the horizontal range between the measuring device and the body is M.

The trajectory parameters are the altitude above the earth's surface (h), velocity (v) and ballistic coefficient (K).

A. Examples of objects not in free fall

- Flying in an aircraft: there is also an additional force of lift.
- Standing on the ground: the gravitational acceleration is counteracted by the normal force from the ground.
- Descending to the Earth using a parachute, which balances the force of gravity with an aerodynamic drag force (and with some parachutes, an additional lift force).

An initially-stationary object which is allowed to fall freely as shown in fig.3, under gravity drops a distance which is proportional to the square of the elapsed time. This image, spanning half a second, was captured with a stroboscopic flash at 20 flashes per second. During the first 1/20th of a second the ball drops one unit of distance (here, a unit is about 12 mm); by 2/20ths it has dropped at total of 4 units; by 3/20ths, 9 units and so on.

Under normal earth-bound conditions, when objects move owing to a constant gravitational force a set of dynamical equations describe the resultant trajectories. For example, Newton's law of universal gravitation simplifies to $F = mg$, where m is the mass of the body. This assumption is reasonable for objects falling towards earth over the relatively short vertical distances of our everyday experience, but is very much untrue over larger distances, such as spacecraft trajectories. Please note that in this article any resistance from air (drag) is neglected [3].

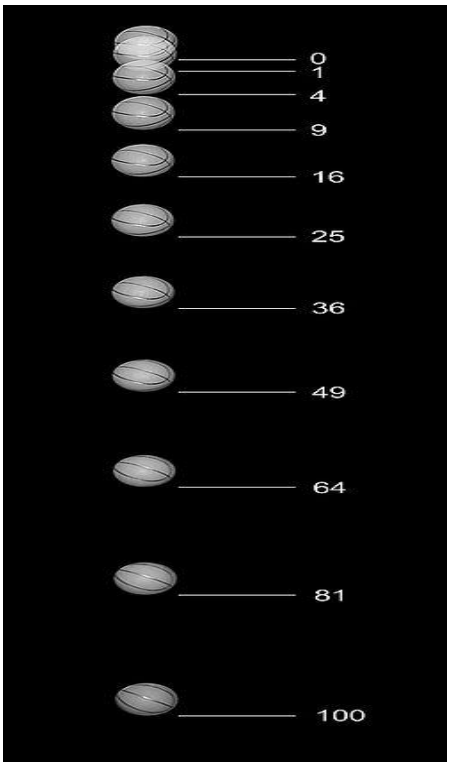


Fig3: Free fall body

B. Development of the filter model

The state space model of Particle filter is given by

$$X_{k+1} = f(x_k, u_k, w_k) \quad (32)$$

$$z_k = h(x_k, v_k) \quad (33)$$

$$f(x_k, u_k, w_k) = \begin{pmatrix} f_1(x_k, u_k, w_k) \\ f_2(x_k, u_k, w_k) \\ \cdots \\ \cdots \\ f_n(x_k, u_k, w_k) \end{pmatrix} \quad (34)$$

The falling body state vector be X_S (k) where

$$Xs(k) = [x1(k) \quad x2(k).....xn(k)]^T \quad (35)$$

$$Xs(k) = [h(k) \quad v(k) \quad k]^T \quad (36)$$

$$\dot{h}(k) = v + w1 \quad (37)$$

$$\dot{v}(k) = \rho_0 e^{(-h/k)} * (v^2/2k) - g + w2 \quad (38)$$

$$\dot{K} = W_3(k) \quad (39)$$

$$Z = h + V \quad (40)$$

ρ_0 –Air density at sea level
g –Acceleration due to gravity
W (k) –Process noise
V –Measurement noise

Let R (k) is the horizontal distance given by

$$R(k) = \sqrt{M^2 + h((t) - a)^2} + v_k \quad (41)$$

V. RESULTS

A. Results

Fig .4

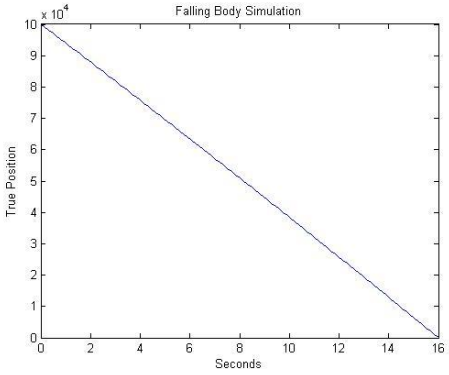


Fig .5

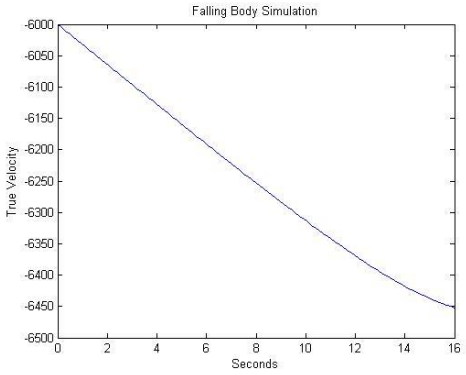


Fig .6

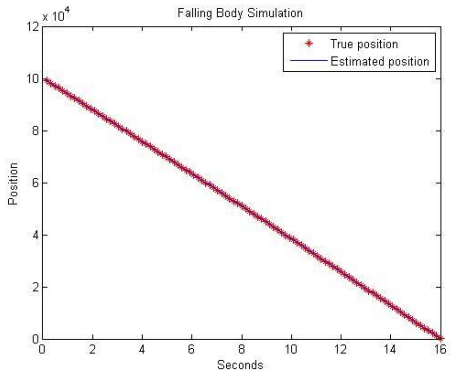


Fig .7

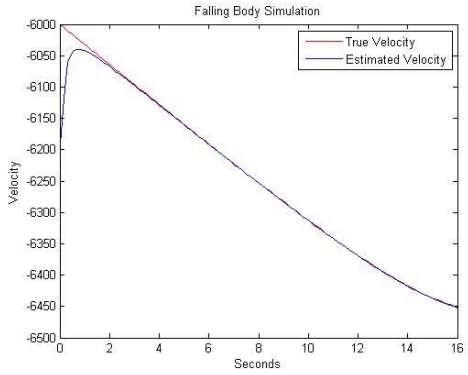
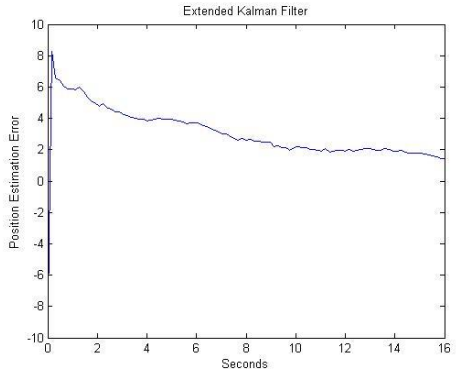
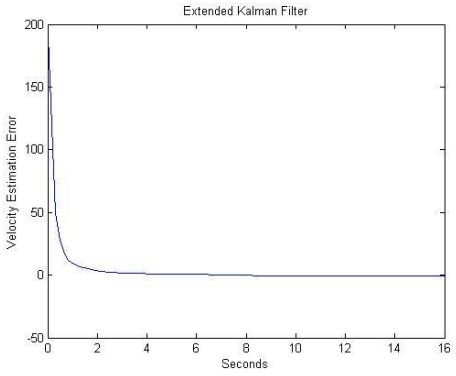


Fig .8



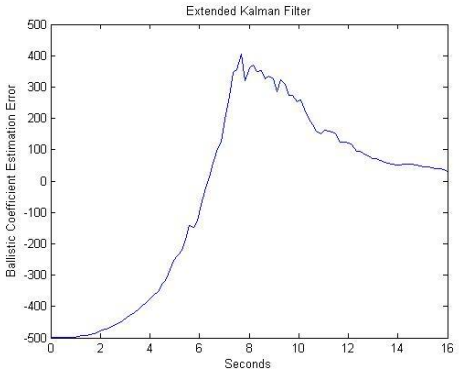
ERROR = 1.8 feet

Fig .9



ERROR = 0feet/sec

Fig .10



ERROR = 31

The figures 4,&5 plots the true values of the two parameters Altitude, and velocity respectively and the figures 6 & 7 plots the comparision between true and estimated parameters Altitude, and velocity respectively. the figures 8,9,& 10 show plots of estimation errors.

VI. CONCLUSIONS

This paper deals with the performance of Extended Kalman filter which is an extension to the basic Kalman filter. Even though the Kalman filter is simple to implement, it is not able to provide accurate results because it mainly deals with pure linear models. So in order to improve the performance, Extended Kalman filter is implemented for free fall body towards Earth. This model is applied to the non-linear processes. This model deals with linearization of the non-linear process, same as that of the linear approximation. The linearization is upto first order approximation. The linearization of non-linear model can be achieved using Jacobian matrix, which is related to the number of updations. As a result, there will be much more accurate results when compared to that of the Kalman filter. This is applied to a realistic example like tracking a freely falling body. EKF is implemented for a specific application, tracking a freely falling body through the atmosphere for 100 Monte-Carlo simulations and the result shows that EKF provides better probability of state estimation compared to Kalman filter.

VII. FUTURE SCOPE

The Extended Kalman filter linearizes the nonlinear model through a single point altogether. In addition, if the initial estimate of the state is wrong, or if the process is modeled incorrectly, the filter may quickly diverge, owing to its linearization. Another problem with the extended Kalman filter is that the estimated covariance matrix tends to underestimate the true covariance matrix and therefore risks becoming inconsistent in the statistical sense without the addition of "stabilizing noise". Linearization of nonlinear system with Unscented Kalman filters and particle filters may provide better probability of state estimation.

ACKNOWLEDGMENT

The authors thank Sri. G.T.Rao, Associate Professor, Department of E.C.E, G.V.P.College of Engineering, for his valuable suggestions in completing this paper. The authors also thank Dr.N.Balasubrhamanyam, Head of the Department

E.C.E, G.V.P.College of Engineering for the support given in presenting this paper.

REFERENCES

[1] Simon Julier and Jeffrey Uhlmann. A new extension of the kalman filter to nonlinear systems. *Int. Symp. Aerospace/Defense Sensing, Simul. And Controls, Orlando, FL*, 1997.

[2] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. A novel approach to nonlinear/non-Gaussian Bayesian state estimation. In IEE Proceedings on Radar and Signal Processing, volume 140, pages 107{113, 1993.

[3] Leela Kumari B,Padma raju.K,Application of particle filter for a free fall body towards earth,In IJAEA Proceedings ,Volume II,pages 195{199,2011}

[4] Fredrik Orderud Sem Sælands vei 7-9, NO-7491 Trondheim Comparison of Kalman Filter Estimation Approaches for State Space Models with Nonlinear Measurements.

[5] Simon Julier and Jeffrey Uhlmann. A new extension of the kalman filter to nonlinear systems. Int. Symp. Aerospace/Defense Sensing, Simul. And Controls, Orlando, FL, 1997

[6] Cox, H., "*On the Estimation of State Variables and Parameters for Noisy Dynamic Systems*," IEEE Transactions on Automatic Controls, Vol. AC-9, 1964, pp. 5-12.

[7] M. Athans, R. P. Wishner and A. Bertolini. Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements. IEEE Transactions on Automatic Control, TAC-13(6):504{518, October 1968.

[8] F. E. Daum. New Exact Nonlinear Filters. In J. C. Spall, editor, Bayesian Analysis of Time Series and Dynamic Models, chapter 8, pages 199{226. Marcel Drekker, Inc., 1988.

[9] Simon Julier and Jeffrey Uhlmann. A new extension of the kalman filter to nonlinear systems. *Int. Symp. Aerospace/Defense Sensing, Simul. And Controls, Orlando, FL*, 1997

[10] Extended Kalman Filter_Shoudong *ARC Centre of Excellence for Autonomous Systems (CAS) Faculty of Engineering and Information Technology,University of Technology Sydney April 23,2010.*

[11] Chi-Tsong Chen. *Linear System Theory and Design, third edition*. Oxford University Press, 1999.

[12] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, February 2002.

[13] N. Bergman, "Recursive Bayesian estimation: Navigation and tracking applications," Ph.D. dissertation, Linköping Univ., Linköping, Sweden,1999..

[14] A. Doucet, "On sequential Monte Carlo methods for Bayesian filtering," Dept. Eng., Univ. Cambridge, UK, Tech. Rep., 1998.

[15] Leela Kumari B,Padma raju.K, Analysis of state space estimation models for linear and nonlinear systems ,In National conference DEVICE-2011 Proceedings ,pages 80{85,2011}

AUTHORS PROFILE

B.Leela Kumari received the B.Tech degree in Electronics and Communication Engineering from J.N.T.University, and M.Tech degree in Radar and Microwave Engineering from Andhra University College of Engineering. She has 8 years of teaching experience and is Associate Professor of Electronics and Communication Engineering, G.V.P.College of Engineering, Visakhapatnam. She has published 7 technical papers in National/International Journals/Conference proceedings.Her research interests include Signalprocessing, State Estimation, tracking and particle filters. She registered for Ph.D. from J.N.T.U. Kakinada.

K.Padma Raju received B.Tech from Nagarjuna University, M. Tech from NIT Warangal, Ph. D from Andhra University, India and Post Doctoral Fellowship at Hoseo University, South Korea. He has worked as Digital Signal Processing Software Engineer in Signion Systems Pvt. Ltd., Hyderabad, India, before joining Jawaharlal Nehru Technological University Kakinada, India.

He has 17 years of teaching experience and is Professor of Electronics andCommunication Engineering, Jawaharlal Nehru Technological University Kakinada, India. Presently working as Director, Industry Institute Interaction, Placements & Training, Jawaharlal Nehru Technological University Kakinada, India.He worked as Research Professor at Hoseo University, South Korea during 2006-2007. He has published 30 technical papers in National/International Journals/Conference proceedings and guiding 06 research students in the area of Antennas, EMI/EMC and Signal Processing.His fields of interest are Signal Processing, Microwave and Radar Communications and EMI/EMC.

V.Y.V.Chandan received the B.Tech degree in Electronics and Communication Engineering from J.N.T.University. His research interests include Signal processing, State Estimation.

R.Sai Krishna received the B.Tech degree in Electronics and Communication Engineering from J.N.T.University, His research interests include Signal processing, State Estimation.

V.M.J.Rao received the B.Tech degree in Electronics and Communication Engineering from J.N.T.University, His research interests include Signalprocessing, State Estimation.