Key Brain Circuits

Contents

Contents

Contents

Contents

Contents

A Novel Scheme to Eliminate Common Mode Voltage in Multilevel Inverters

Nasim Rashidi-rad, Abdolreza Rahmati, Adib Abrishamifar

Abstract—Nowadays, multilevel voltage source inverters offer several advantages compared to their conventional two-level inverters. In these inverters, by synthesizing several levels of dc voltages, the staircase output waveform is produced. The structure of this waveform will have lower total harmonic distortion which leads to an approach to a desired sinusoidal waveform. Achieving higher output voltage and lower stress on power switches are other advantages of theses inverters. But in multilevel inverters the problem of common mode voltage which had been found in conventional two level inverters can still be considered as a major issue which leads to motor bearing failures. Therefore to eliminate these voltages proposing some methods seems to be necessary. This paper proposes a generalized method to generate pulse width modulation signals in multilevel inverters that have an odd number of levels. The main idea of this method to generate these signals for an n-level inverter is based on a freely selectable modulation method of an (N+1)/2 level imaginary inverter. This method which leads to eliminate common mode voltages of the n-level inverter can be extended to higher levels.

Index Terms— Common mode voltage, Phase voltage, Line voltage, N level inverter.

———————————— ◆ ————————————

1 INTRODUCTION

Multilevel inverters are receiving increased attention recently, especially for use in high power applications. This increased attention is probably due to the fact that the output waveforms are much improved over those of the two level inverters.

The staircase waveform of multilevel inverters which is composed of several levels of dc voltages, will lead to higher output voltage and lower stress on power switches. Furthermore by increasing the levels of output voltage, the waveform contains lower harmonic contents that will lead to reduce the requirements of output filter.

Therefore multilevel inverters have been selected as a preferred power inverter topology for high voltage and high power applications [1-4].

All multilevel PWM inverters such as conventional two level inverters generate common-mode voltage within the motor windings. This voltage may result in motor and drive application problems [5].

This paper presents a method in multilevel inverters, which is based on space vector diagram in an imaginary inverter. This scheme completely eliminates common-mode voltages and will be applicable for any general circuit configuration.

————————————————

- *N. Rashidi-rad, Iran University of Science and Technology, Iran, E-mail: Nasim83.iust.iust@gmail.com*
- *A. Rahmati, Iran University of Science and Technology, Iran, E-mail: rahmati@iust.ac.ir*
- *A. Abrishamifar, Iran University of Science and Technology, Iran, E-mail: Abrishamifar@iust.ac.ir*

The following sections of the paper present the algorithm of this method and simulation results in a five level inverter.

2 PRINCIPLES OF THE PROPOSED METHOD

In an N-level inverter, each phase voltage can produce an N-level staircase waveform. Therefore line voltages of this inverter which is the subtraction of two phase voltages will have 2N-1 levels.

$$V_{ab} = V_a - V_b$$
$$V_{bc} = V_b - V_c \qquad (1)$$
$$V_{ca} = V_c - V_a$$

The average of line voltages with 120 degree phase difference will be zero at each moment.

$$V_{ab} + V_{bc} + V_{ca} = (V_a - V_b) + (V_b - V_c) + (V_c - V_a) = 0 \quad (2)$$

Therefore if line voltages of one imaginary inverter will be used as phase voltages of another inverter, the momentary average of the phase voltage of the second inverter will be always zero. To achieve this, the following equations are necessary to be valid in the system:

$$if \ V_{Line \ (N_I \ Level)} = V_{Phase \ (N \ Level)} \ \Rightarrow$$

$$\begin{cases} 2N_I - 1 = N \\ \sqrt{3}(N_I - 1)V_{C(N_I \ Level)} = (N-1)V_{C(N \ Level)} \end{cases} \qquad (3)$$

$$\Rightarrow V_{C(N_I \ Level)} = \frac{(N-1)V_{C(N \ Level)}}{\sqrt{3}(N_I - 1)}$$

Therefore it can be concluded that if line voltages of an (N+1)/2 level inverter will be used as phase voltages of an N level inverter, then the common mode voltage of the N level inverter will become zero.

In other words, each voltage vector in the space vector diagram of the (N+1)/2 level inverter will be corresponded to the voltage vectors with zero common mode voltage in an N level inverter. Therefore the number of all voltage vectors in an (N+1)/2 level inverter is equal to the number of voltage vectors with zero common mode voltage in an N level inverter.

It should be noted that since (N+1)/2 is an positive integer number, N cannot become an even number; this means that for an even N, an N level inverter cannot have zero common mode voltage.

In the following paragraphs, the relations between number of vectors and switching states for an N and (N+1)/2 level inverters has been presented.

Based on table 1, the number of switching states in a three phase N level inverter will be N^3. Among these states, the number of voltage vectors will be calculated based on following relations.

$$N_{V(N=1)} = 1, \quad N_{V(N)} = N_{V(N-1)} + 6(N-1) \implies$$
$$N_{V(N=2)} = 7, N_{V(N=3)} = 19, ... \tag{4}$$

Therefore the number of voltage vectors with zero common mode voltage in an N level inverter can be calculated as:

$$N_{Vectors\ With\ Vcom=0(N)} = N_{V((N+1)/2)} \tag{5}$$

3 RESULTS

This scheme has been explained for a 5level modular inverter, which has been utilized to drive a 400 volt motor. Therefore the space vector diagram which has been used, belongs to the 3level inverter and $V_{C_{TOTAL}} = 600v$.

$$if \quad V_{Line(3level)} = V_{Phase(5level)} :$$

$$2V_{C(3level)} = \frac{4V_{C(5level)}}{\sqrt{3}} \implies V_{C(3level)} = \frac{2V_{C(5level)}}{\sqrt{3}} \tag{6}$$

$$V_{C_{TOTAL}} = (N-1)V_{C(N\ Level)}, \quad N = 5 \implies$$

$$V_{C(5level)} = 150v \implies V_{C(3level)} = 173.2v$$

TABLE I.   RELATIONS BETWEEN NUMBER OF VECTORS AND SWITCHING STATES

N	NUMBER OF SWITCHING STATES IN AN N LEVEL INVERTER	NUMBER OF VOLTAGE VECTORS IN AN N LEVEL INVERTER	NUMBER OF VOLTAGE VECTORS WITH VCOM=0 IN AN N LEVEL INVERTER	NUMBER OF VOLTAGE VECTORS IN AN (N+1)/2 LEVEL INVERTER	NUMBER OF SWITCHING STATES IN AN (N+1)/2 LEVEL INVERTER
1	1	1	1	1	1
3	27	19	7	7	8
5	125	61	19	19	27
7	343	91	37	37	64
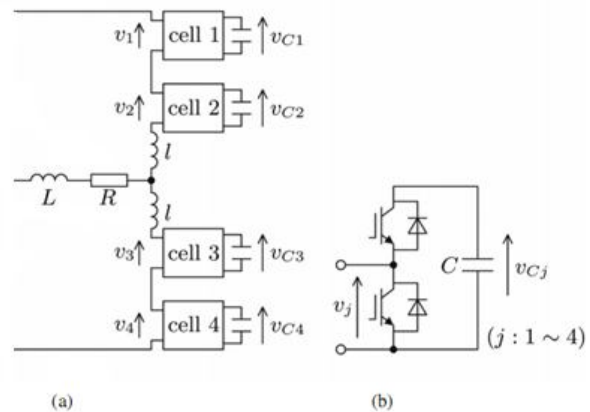
Figure 1.   Circuit configuration of (a) one phase leg of a five level modular inverter, (b) a half-bridge cell

According to the proposed scheme, phase voltages of the 5level inverter will be composed by line to line voltages of the 3level inverter. So in these equations V_a, V_b and V_c are phase voltages and:

$$V_{a(5level)} = V_{a(3level)} - V_{b(3level)}$$
$$V_{b(5level)} = V_{b(3level)} - V_{c(3level)} \tag{7}$$
$$V_{c(5level)} = V_{c(3level)} - V_{a(3level)}$$

The definition of the common mode voltage is:

$$V_{com(5level)} = \frac{V_{a(5level)} + V_{b(5level)} + V_{c(5level)}}{3} \implies \tag{8}$$

Therefore:

$$V_{com(5level)} = 0 \tag{9}$$

Considering the elimination of common mode voltage in the 5level inverter, it can be concluded that it won't be dependent to the modulation method which is utilizing in the imaginary 3level inverter.

Another conclusion is that redundant switching states as shown in figure 2, will have similar results. For instance if the selected modulation method in 3level inverter will be based on delivering nearest vector to the reference vector, election each of switching states of (2,1,1) or (1,0,0) will lead to zero common mode voltage and equal output voltages in the 5-level inverter. Furthermore among three switching states for zero vector, the vector of (1, 1, 1) has been selected.

In this stage, selecting a suitable modulation method in the imaginary 3level inverter, that has a rather low harmonic distortion and a good linearity relationship is highly important. But it should be noted that reducing or eliminating of common mode voltage in the selection of the modulation method won't be considered.

Therefore one of the major advantages of the proposed method is its independency in selection of modulation method of the (N+1)/2 level inverter.

Since elimination of common mode voltage in SPWM methods usually is more difficult than SVM methods, the proposed scheme can be utilized to eliminate the common mode voltage in SPWM methods. Figure 4 has shown phase and line and also common mode voltage in the 5level inverter utilizing an SPWM method.

It should be noted that the selective carrier based modulation is PD (Phase Disposition) method, which has the same phases of two adjacent carriers [6].

(a)

(b)
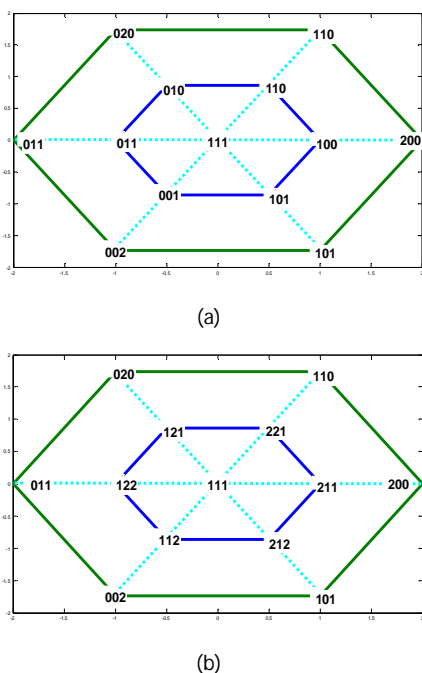
Figure 2. Considering of redundant switching states in space vector diagrams of the 3level inverter
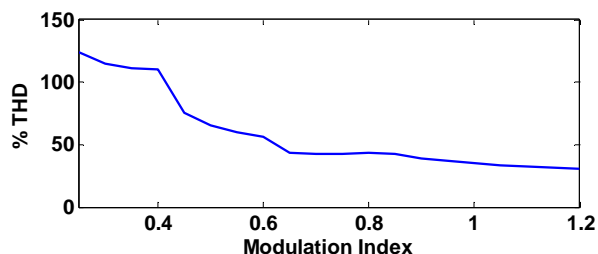
Figure 3. Total harmonic distortion in the 5level inverter, for both diagrams of figure 2

(a)

(b)

(C)
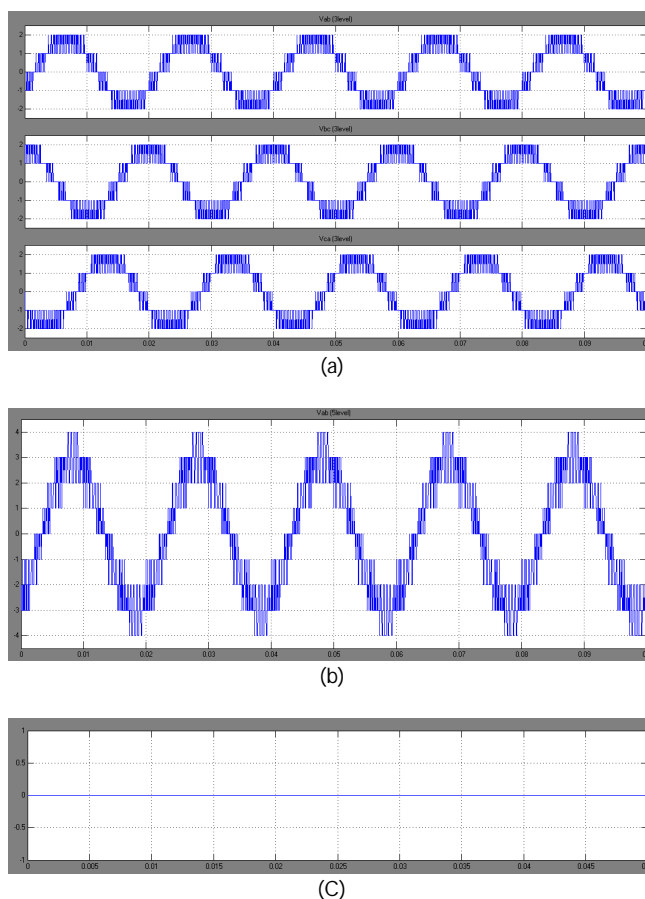
Figure 4. (a) phase voltages, (b) line voltage, and (c) common mode voltage in the 5level inverter by an SPWM method (PD)

4 CONCLUSION

This paper has proposed a novel scheme to generate pulse width modulation signals for multilevel inverters. This scheme which can be used in inverters with an odd number of levels will reduce the common mode voltage to zero.

The main idea of the scheme is based on zero momentary average of line voltages in each inverter. Therefore by

utilizing the line voltages of an imaginary inverter as phase voltages of the real inverter, the common mode voltage of the real inverter will be zero.

Furthermore it can be concluded that the number of all voltage vectors in the imaginary inverter will be equal to the number of voltage vectors with zero common mode voltage in the real inverter.

Independency of this scheme in selection of SPWM or SVM methods for the imaginary inverter is one of its major advantages.

The proposed method has been shown for an inverter with 5level voltage. It should be noted that this method can be easily extended to higher levels.

REFERENCES

[1] Nikola Celanovic, DushanBoroyevich, "A Fast Space-Vector Modulation Algorithm for Multilevel Three-Phase Converters", IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS, VOL.37, NO.2, MARCH/APRIL 2001

[2] Aziz, J.A, Salam, Z. "A PWM Strategy for the Modular Structured Multilevel Inverter Suitable for Digital Implementation", 2002 IEEE.

[3] Sanmin Wei and Bin Wu, Fahai Li and Congwei Liu "A General Space Vector PWM Control Algorithm for Multilevel Inverters", 2003 IEEE.

[4] "Investigation and Comparison of Multi-Level Converters for Medium Voltage Applications", Seyed Saeed Fazel, Berlin 2007

[5] Haoran Zhang, Annette von Jouanne, Shaoan Dai, Alan K.Wallace, Fei Wang, "Multilevel Inverter Modulation Schemes to Eliminate Common-Mode Voltages", IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS, VOL. 36, NO. 6, NOVEMBER / DECEMBER 2000.

[6] Zhou Jinghua, Li Zhengxi, "Research on Multi-carrier PWM Modulation Strategies of Three-level Inverter"

Physical Behavior of Eigenvalues and Singular Values in Matrix Decompositions

Azmol Huda, R. L. Aguiar

Abstract— An apposite as well as realistic treatment of eigenvalue and singular value problems are potentially of interest to a wide variety of people, including among others, design engineers, theoretical physicists, classical applied mathematics and numerical analysis who yearn to carry out research in the matrix field. In real world, it is extensively used but at sometimes scantily understood. This paper focuses on building a solid perception of eigenvalue as well as singular value with their substantial meanings. The main goals of this paper are to present an intuitive experience of both eigenvalue and singular value in matrix decompositions throughout a discussion by largely building on ideas from linear algebra and will be proficiently to gain a better perceptive of their physical meanings from graphical representation.

Index Terms—Eigenvalue, singular value, matrix decomposition, orthonormal basis, linear mapping.

———————————————— ◆ ————————————————

1 INTRODUCTION

EIGENVALUE as well as singular value has widespread application in diverse fields of empirical science from mathematics to neuroscience because they are a straightforward, non-parametric value that extracts pertinent information from a large matrix. With nominal effort they make available a roadmap to divulge the hidden, simplified structures in a large matrix that frequently lie beneath it. Eigenvalues take part in a significant role in situations where the matrix is a transformation form one vector space onto itself. The primary paradigms of it are systems of linear ordinary differential equations. The eigenvalues of a matrix be capable of keeping up a correspondence to frequencies of vibration, or critical values of stability factors or energy level of atoms. The most application of eigenvalue probably in the field of dimension reduction.

Singular values also play a vital role where the matrix is a transformation from one vector space to a different vector space, possibly with a dissimilar dimension. Systems of over or undetermined algebraic equations are the most important examples. The term "singular value" relates to the distance between a matrix and the set of singular matrices. From the seminal study of eigen and singular value, there has been a lot of research work related to this field [1], [2], [3], and incredibly common to all mathematicians or engineers although nearby there is very little works straightforwardly in attendance to the relation of eigenvalues and singular values simultaneously with their graphical properties, in an efficient and representative way. In paper [4], the author addresses a number of numerical issues of singular value arising in the study of models of linear systems with its applications. Numerical computation of the characteristic values of a real symmetric matrix and numerically stable, fairly fast technique for achieving the singular values are discussed in paper [5] and [6]. Authors in paper [7], describe the solution of large scale eigenvalue problems. In paper [8] and [9], the authors illustrate that the largest eigenvalue

of the sample covariance matrix have a tendency to a limit under certain conditions and the limit of the cumulative distribution function of the eigenvalues is likely to determine by using a technique of moments. In paper [10], C.B. Moler, presents a graphical depiction, but the singular value decomposition as well as the relation with the eigenvalues is not discussed in a broader consideration. In this paper, we put in plain words an instictive feel of the physical meaning of eigenvalues and singular values along with emphasizing more on their graphical consequence with numerical example in a meticulous way.

The rest of this paper is structured as follows. The notations used all the way throughout of this paper are given in section 2. We discuss the eigenvalue and singular value matrix decompositions in section 3. Relation between eigenvalue and singular value and their geometrical elucidation are given in section 4 and 5 respectively. Graphical representations are presented in section 6. Concluding remarks are presented in section 7.

2 BACKGROUND NOTATIONS

Symbol	Meaning
A	: square or rectangular matrix of order n over a field F ;
λ_i	: eigenvalues of matrix A ;
Λ	: n-by-n diagonal matrix with the λ_j on the diagonal ;
X	: denotes n-by-n diagonal matrix whose j-th column is x_i ;
U, V	: orthogonal or unitary matrices;
σ_i	: singular values of A;
Σ	: n-by-n diagonal matrix with the σ_r on the diagonal;
A^T	: transpose matrix of A;
A^H	: Hermitian matrix of A;
$\Re$	: set of real numbers;

3 EIGENVALUE AND SINGULAR VALUE DECOMPOSITIONS

3.1 Eigenvalue Decompositions

Let A be a square matrix of order n over a field F. A scalar $\lambda \in F$ is an eigenvalue of a matrix A if there exists a non-zero column vector x for which

$$Ax = \lambda x \qquad (1)$$

The eigenvalue-eigenvector equation for a square matrix can be written as $(A - \lambda I)x = 0, \ x \neq 0$ where I denotes the identity matrix of order n. This implies that $(A - \lambda I)$ is singular and hence that $\det(A - \lambda I) = 0$. This is christened the characteristic equation or characteristic polynomial of A. The degree of polynomial is the order of the matrix and an n-by-n matrix has n eigenvalues, counting repeated roots. Suppose $\lambda_1, \lambda_2, ..., \lambda_n$ be the eigenvalues of a matrix A and $x_1, x_2, ..., x_n$ be a set of corresponding eigenvectors. Let Λ denote the n-by-n diagonal matrix with the λ_j on the diagonal, and let X denotes the n-by-n matrix whose j-th column is x_j, then we can write

$$AX = X\Lambda \qquad (2)$$

From equation (2), it is indispensable to put Λ on the right side so that each column of X can be multiplied by its resultant eigenvalues. A noteworthy key conjecture is that, this statement is not factual for all matrices. At this juncture, we assume that the eigenvectors are linearly independent, afterward the inverse of matrix X exists and

$$A = X\Lambda X^{-1} \qquad (3)$$

with the non singular matrix X. This is well-known as the eigenvalue decomposition of matrix A. When it exists, it consents us to investigate the properties and characterization of A by exploring the diagonal elements of Λ. For illustration, repeated matrix powers can be put across in terms of powers of scalars i.e.

$$A^k = X\Lambda^K X^{-1} \qquad (4)$$

But, if the eigenvectors of A are not linearly independent, then such a diagonal does not exist and the powers of A give us an idea or an evidence of a more complex behavior. If T is any non singular matrix, then $A = TQT^{-1}$ is known as similarity transformation and A and Q are said to be similar. If $Ax = \lambda x$ and $x = Ty$ then $Qy = \lambda y$. So a similarity transformation preserves eigenvalues. Commonly, the eigenvalue decomposition is an attempt to find a similarity transformation to diagonal form.

3.2 Singular value Decompositions

Let A be a rectangular matrix of order m-by-n and the rank of AA^T is r. Therefore, AA^T is a square symmetric matrix of order m-by-m. Let us define some more quantities: let $V = \{v_1, v_2, ... v_r\}$ be the set of orthonormal $m \times 1$ eigenvectors with associated eigenvalues $\{\lambda_1, \lambda_2, ... \lambda_r\}$ for the symmetric matrix AA^T. Therefore, $(AA^T)v_i = \lambda_i v_i$. Also let, $U = \{u_1, u_2, ... u_r\}$ be the set of $n \times 1$ vectors defined by

$$u_i \equiv \frac{1}{\sigma_i} Av_i \ \text{where} \ u_i.u_j = \begin{cases} 1 & if \ i = j \\ 0 & otherwise \end{cases}$$

and σ_i is related with the eigenvalues by the relation $\sigma_i = \sqrt{\lambda_i}$. This makes an unexpected property, $\|Av_i\| = \sigma_i$. A scalar $\sigma \in F$ is called a singular value of a rectangular matrix A, with a pair of singular vectors U and V if it satisfies the relation

$$AV = \sigma U \qquad (5)$$

This can be written as in matrix form $AV = U\Sigma$ or $A^H U = V\Sigma^H$ where Σ stands for the n-by-n diagonal matrix with the σ_r on the diagonal. The superscript H stands for Hermitian transpose. The mathematical intuition behind the construction of the matrix is that we craving to express all n scalar equations in just one equation. The ultimate form of SVD is thick and it is uncomplicated to comprehend this process in graphically. The complete structures of singular value decomposition are described in Figure 1.

Figure 1.(a) is the basic form of singular value decomposition of a matrix, whereas, Figure 1.(b) and 1.(c) are the form of singular value decomposition of the matrix when $m < n$ and $m > n$ respectively. It yields that singular vectors can constantly be chosen to be perpendicular to each other, so the matrices U and V, whose columns are the normalized singular vectors, satisfy $U^H U = I$ and $V^H V = I$. In other words, U and V are orthogonal, if they are real and unitary, if they are complex. From equation (4), we can effortlessly derive

$$A = U\Sigma V^H \qquad (6)$$

with the diagonal matrix Σ and orthogonal or unitary matrices U and V. This is well-known as the singular value decomposition or SVD, of matrix A.
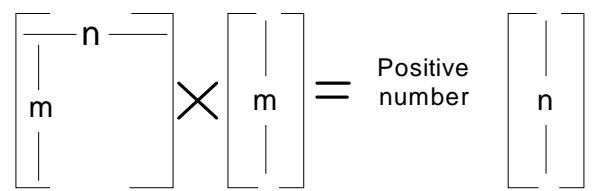
Figure. 1(a) General form of SVD
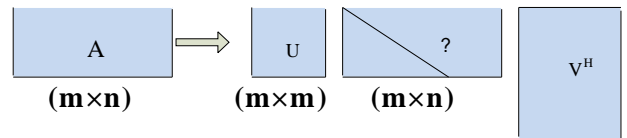
Figure 1.(b) when m<n
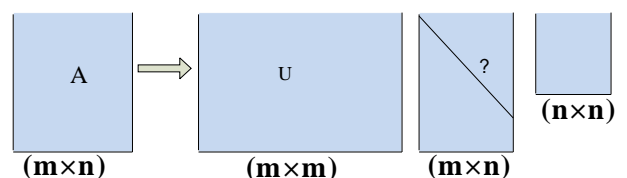
Figure 1.(c) when m>n

Fig.1. Structures of singular value decompositions.

4 RELATION BETWEEN EIGENVALUE AND SINGULAR VALUE DECOMPOSITIONS

In our earlier section, we have established that $A = U\Sigma V^H$. Now we can obtain the relation between eigenvalues and singular values by a trouble-free computation.

$$A^H A = \left(U\Sigma V^H\right)^H \left(U\Sigma V^H\right) = V\Sigma^H U^H U\Sigma V^H = V\left(\Sigma^H \Sigma\right)V^H$$

$$AA^H = \left(U\Sigma V^H\right)\left(U\Sigma V^H\right)^H = U\Sigma V^H V\Sigma^H U^H = U\left(\Sigma\Sigma^H\right)U^H \qquad (7)$$

The right hand side of these two equations expresses the eigenvalue decompositions of the left hand sides. It is evidently seen that, the squares of the non-zero singular values of A are equal to the non-zero eigenvalues of their $A^H A$ and AA^H. Furthermore, the columns of U are eigenvectors of AA^H and the columns of V are eigenvectors of $A^H A$.

5 GEOMETRICAL INTERPRETATION OF EIGENVALUE AND SINGULAR VALUE

Let us take a closer look carefully of equation (1), $Ax = \lambda x$ and let us now ask whether there are any vectors which are not changed in direction by the deformation? The answer may be found easily from the equation (1). We could make some quick concluding remarks of equation (1). The square matrix A can be thought of as a transformation matrix. The multiplication of the matrix A, on the left of a vector **x**, the answer is another vector that is transformed from its original position. The transformed vector **x** does not change its direction, only changes its magnitude.

Moreover, it is effortlessly seen that, even if we scale the vector **x**, by a few amount before we multiply it, we at a standstill get the same multiple of it as a result. This is for the reason that, if we scale by some amount, all we are doing is, building it longer, not change its direction. Perhaps the preeminent way to think of an eigenvector **x** of a matrix A is that it represents a direction which remains invariant under multiplication by A. The corresponding eigenvalues of A are then the representation of A in the subspace spanned by the eigenvector **x**. The benefit of having this representation is that multiplication by A is reduced to a scalar operation along the eigenvector. For instance, from the equation (4), $A^k = X\Lambda^K X^{-1}$ we can state that the effect of a power of A along **x** can be determined by taking the subsequent power of A. From the point of linear algebra, eigenvalues are pertinent of a square n-by-n matrix A which is reflection of as a mapping of n dimesional space onto itself. Here, we attempt to find a basis for the space so that the matrix is converted into diagonal. Even if A is real, this basis might be complex. In fact, if the eigenvectors are not linearly independent, such a basis does not subsist.

On the contrary, SVD is relevant to a possibly rectangular m-by-n matrix A which is thought of as mapping n-space onto m-space. Eigenvalue decomposition is applicable only for square matrix; in contrast, the singular value decomposition is possible for any rectangular matrix. In SVD, we strive to find one change of basis in the domain and a typically different change of basis in the range with the intention that the matrix becomes diagonal. Such bases always exist and real if A is real. Usually, the transforming matrices are orthogonal or unitary so as they preserve lengths and angles and do not magnify errors. The geometrical outlook of the SVD decompositions can be summarized as follows: for every linear map $\Omega : F^n \to F^m$ of the field F one can find orthonormal bases of F^n and F^m such that Ω maps the i-th basis vector of F^n to a non-negative multiple of the i-th basis vector of F^m. With respect to these bases, the map Ω is therefore represented by a diagonal matrix with non-negative real diagonal entries. Finally, let we look a more visual flavor of singular values and singular value decompositions, at least when it works on a real vector space. If S is a sphere of radius one in $\Re^n$, after that the linear map $\Omega : F^n \to F^m$ maps this sphere onto an ellipsoid in $\Re^m$ and usually, nonzero singular values are simply the lengths of the semi-axes of this ellipsoid.

Figure 2.(a) Figure 2.(b)

Figure 2.(c) Figure 2.(d)

Figure 2.(e) Figure 2.(f)
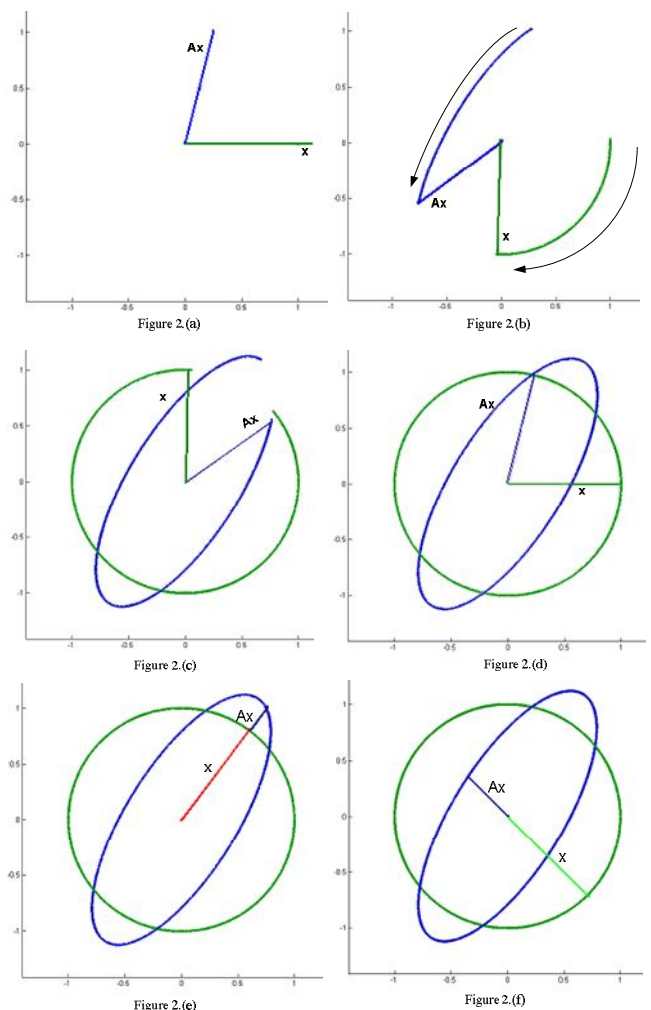
Fig. 2. Plot of eigenvalues.
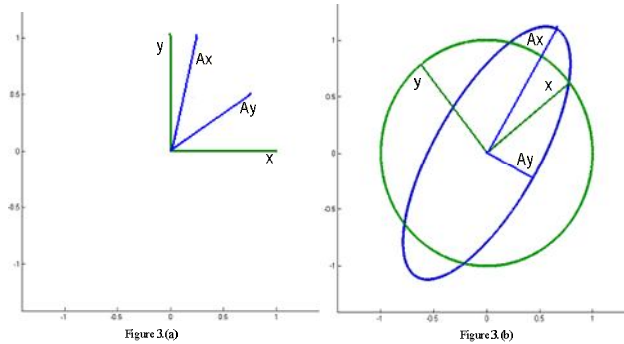
Fig. 3. Plot of singular values.

6 GRAPHICAL PRESENTATION OF EIGENVALUE AND SINGULAR VALUE

Here we demonstrate a graphical view of eigenvalues and singular values by Mat Lab with a numerical example. We use the transformation matrix

$$A = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ 1 & \frac{1}{2} \end{bmatrix}$$

of order 2 and a unit vector $x = \begin{bmatrix} 1 & 0 \end{bmatrix}'$ to facilitate of drawing a unit circle.

The resulting trajectory of **Ax** is plotted. In Figure-2, the first four subplots, Figure-2.(a), 2.(b), 2.(c) and 2.(d), are the intermediate steps of their traversed orbits. The goal is to find the vectors **x** so that **Ax** is parallel to **x**. Generally, for such a direction **x**, the transformation matrix A is simply a stretching or shrinking by a factor of λ. Each such **x** is an eigenvector and the length of **Ax** corresponding eigenvalue. From the last two subplots Figure-2.(e), 2.(d) of Figure-2 , the first eigenvalue is positive, so **Ax** lies on top of the eigenvector **x** and the second eigenvalue is negative and **Ax** is parallel to **x** but points in the opposite direction. We have plotted the SVD orbit in Figure-3 with subplots 3.(a) and 3.(b) respectively. The vectors **x** and **y** are perpendicular each other and the resulted **Ax** and **Ay** are plotted. In SVD mode, the axes of the ellipse do play a key role. The resulting **Ax** and **Ay** traverse on the ellipse, but are not perpendicular to each other though **x** and **y** are perpendicular. When **Ax** and **Ay** are perpendicular, they form the axes of the ellipse. In this case, **x** and **y** are right singular vectors as well as the columns of U in the SVD, the vectors **Ax** and **Ay** are multiples of the columns of V and the lengths of the axes of ellipse are the singular values.

7 CONCLUSIONS

In this paper, we have discussed several aspects of eigenvalues and singular values from the point of view of their underlying relations. To catch a better understanding, a pictorial visualization and elucidation, are presented. Here,

we have characterized and discussed the effect of the traverse orbit both of eigenvalues and singular values. It puts on display of a mapping from sphere onto an ellipsoid. When the vector **x** moves on a circle, **Ax** moves on an ellipsoid and the length of **Ax** is the eigenvalue. In singular value mode, non-zero singular values are the simply the lengths of the semi axes of the ellipsoid.

REFERENCES

[1] G.W. Stewart, "Matrix Algorithms: Basic Decompositions," SIAM, Philadelphia, 1998.

[2] H. Anton, C. Rorres, "Elementary Linear Algebra with Applications," 9-th Edition, John Wiley and Sons, 2005.

[3] V.C. Klema, "The Singular Value Decomposition: Its Computation and Some Applications," IEEE Trans. Automatic Control, Vol. 25, pp.164-176, 1980.

[4] W. Givens, "Numerical Computation of the Characteristic Values of a Real Symmetric Matrix," ORNL Rep. 1574.

[5] G.H. Golub and W. Kahan, "Calculating the Singular Values and Pseudo-Inverse of a Matrix," SIAMJ. Nwner. Anal., vol. 2, pp. 205-224, 1965.

[6] R.B. Lehoucq, D.C. Sorensen and C. Yang, "ARPACK User's Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Menhods," SIAM, Philadelphia, 1998.

[7] D. Jonsson, "Some Limit Theorems for the Eigenvalues of a Sample Covariance Matrix," J. Multivariate Analysis 12, pp.1-38, 1982.

[8] Y.Q. Yin, Z.D. Bai and P.R. Krishnajah, "On the Limit of the Largest Eigenvalue of the Large Dimension Sample Covariance Matrix,"Probability Theory and Related Fields, Volume 78, Number 4, 509-521, 1998.

[9] C.B. Moler, "Numerical Computing with MATLAB," February 15, 2008.

[10] Y. Jiao, B. Yang, H. Wang, X. Niu, "SVD Based Robust Image Content Retrieval," International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 351-354, Dec. 2006.

[11] Y.H. Wang, T.N. Tan and Y. Zhu, "Face Verification Based on Singular Value Decomposition and Radial Basis Function Neural Network," (NLPR), Institute of Automation, Chinese Academy of Sciences, 2005.

[12] C.L., Sun, J. Hahn, "Parameter Reduction for Stable Dynamical Systems Based on Hankel Singular Values and Sensitivity Analysis," Chem. Eng.Sci., 61(16), pp. 5393–5403, 2006.

[13] S. Lall, J.E. Marsden, S. Glavaski, "Empirical Model Reduction of Controlled Nonlinear Systems," 14th IFAC World Congress, Beijing, 1999.

[14] C. Sun, J. Hahn, "Reduction of Differential-Algebraic Equation Systems Via Projections and System Identification," Journal of Process Control, vol. 15, pp. 639-650, 2005.

[15] G.H. Golub and C. Reinsch, "Singualr Value Decomposition and Least Squares Solutions," Numer. Math., vol. 14, pp. 403-420, 1970.

[16] M.H.A. Biswas, M.H. Azmol, Munnujahan Ara, M.A. Rahman, "Optimal Control Theory and its Applications in Aerospace Engineering," International Journal on Academic Research, Vol. 3, No. 2, March 2011.

Fuzzy Logic Based Hydro-Electric Power Dam Control System

M. Abbas, M. Saleem Khan, Nasir Ali

Abstract — This research paper presents the construction design of Hydro-Electric Power Dam Control System using Fuzzy Logic. In this design two input parameters: water level and flow rate and two output parameters: release valve control and drain valve control are used. This proposed system uses a simplified algorithmic design approach with wide range of input and output membership functions. The hardware of control system for fuzzifiers and defuzzifiers is designed according to the need of system. The proposed simplified algorithmic design is verified using MATLAB simulation and results are found in agreement to the calculated values according to the Mamdani Model of the Fuzzy Logic Control System.

Index Terms—Fuzzy Logic Control, Hydro-Electric Power Plant, Inference Engine, Rule Selection.

————————————————◆————————————————

1 INTRODUCTION

THE modern-day technologies in the areas of information storage and retrieval, web search, image processing, control, pattern recognition, bio-information and computational biology, e-markets, autonomous navigation, and guidance are benefited using fuzzy sets. An integrated framework sustaining a variety of facets of human-centric computing is developed by means of fuzzy sets. The current trends of information technology have proved that the increasing level of intelligence, autonomy and required flexibility comes true with the increased human centricity of resulting results. The holistic view covers concepts, design methodologies, and algorithms with interpretation, analysis, and engineering knowledge. The computing systems are based on predefined models of two-valued logic and human information processing, concerned with two distinct words. In order to communicate between these two words we need to develop an interface. This is the key motivation behind the emergence of human-centric systems and human-centric computing [1].

The construction of a dam is necessary for the electric power generation, flood control, irrigation system, metropolitan and industrial water supply. Different kind of methods have been introduced and implemented to control the hydro-electric power dam due to non-deterministic behavior of water parameters such as flow rate and release etc. [2].

Fuzzy Set Theory along with its membership functions was implemented to the Fairbairn reservoir in Emerald,

Central Queensland, Australia where fuzzy rules were generated with explicit recognition of storage volume non-specificity in the discrete Stochastic Dynamic Programming (SPD) [3].

Fuzzy dynamic programming model was used for Hirakud dam in the State of Orissa in India in which irrigation; hydropower generation and flood control were considered as fuzzy variables [4].

The neural network and fuzzy systems were also adopted for dam control in which a comparison was made between reservoir operations using the fuzzy and neural network systems and actual one by operator, using examples of floods during flood and non-flood seasons [5].

Reports show that hydroelectric dams produce 20 percent of the world's total production of electrical energy. The development of a hydro-electric power dam control system based on fuzzy logic with two inputs and two outputs. Using water level and flow rate measuring devices for feedback control, and two control elements for draining and valve controlling (release), and formulated fuzzy rules for water level and flow rate has been achieved.

To control the water release, the controller reads the water level and flow rate after every sampling period. This proposed design work of Hydro-Electric Power Dam System is the application of fuzzy logic control system consisting of two input variables: water level and flow rate, and two output variables: Drain valve and (Releasing) Valve control used in a reservoir plant of Hydro-Electric Power Dam to monitor the system of Dam.

The basic structure of the proposed model is described in Section 2. Section 3 gives the simplified design algorithm

————————————————
- *M. Abbas is working as Visiting Faculty in the Department of Electrical Engineering in National University of FAST Lahore and also Lecturer Internee in Physics, Government Islamia College, Lahore, Pakistan (e-mail: shairabbas@yahoo.com)*
- *Dr. M. Saleem Khan is with the GC University Lahore Pakistan, working as Director Computer Science Department (e-mail: mskgcu@yahoo.com, s.khan@ed.ac.uk)*
- *Nasir Ali is an Assistant Professor of Physics in Govt. DS College, Lahore, Pakistan (e-mail nasirali8787@yahoo.com)*

of fuzzy logic for Hydro-Electric Power Dam System. Section 4 describes the simulation results of this system. Conclusion and future work is given in Section 5.

2. BASIC STRUCTURE OF THE PROPOSED HYDRO-ELECTRIC POWER DAM
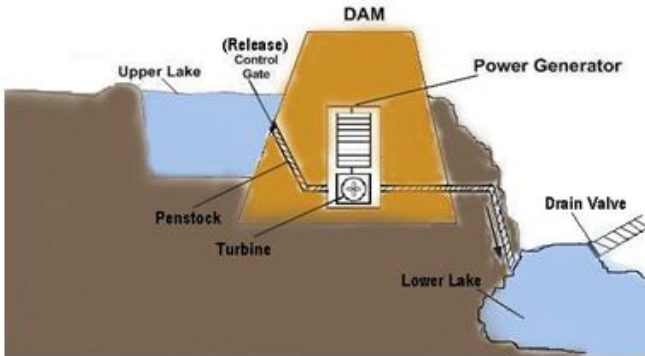
Fig.1 Arrangement of proposed hydro–electric power system

The main parts of the proposed hydro-electric power plant are shown in Fig. 1. Upper lake where water is stored presents the water level. The greater the vertical distance b/w the upper and lower lakes, the more is the generation of electricity. In order to release or block water, a control valve is used according to the need. Water on releasing from the dam gets to the blades of the turbine all the way through the penstock. Its slope and thickness determines the efficiency of the dam. Turbine produces electrical energy and water released from the turbine is released to lower lake where the drainage system is brought into action according to the requirements [6].

The schematic diagram of the proposed hydro-electric power plant is shown in Fig. 2. Water level and flow rate devices are used to monitor the status of water in the plant which is connected with the two fuzzifiers of the fuzzy logic control system after suitable amplification and voltage adjustment unit. Two outputs of defuzzifiers are the releasing control valve and drainage valve.
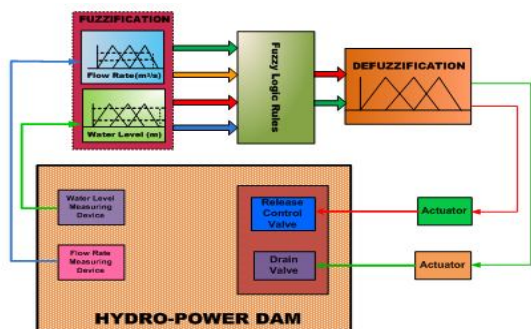
Fig. 2 Block Diagram of Hydro-Electric Power Dam fuzzy control system

3. DESIGN ALGORITHM

The algorithm designed for this system consists of two fuzzy input variables. Five triangular membership functions are equally determined over a scale range of 0 m to 20 m for the water level and $0(m^3s^{-1})$ to$100000(m^3s^{-1})$ for flow rate inputs. The five fuzzy membership functions for water level input are termed as: very low 0-5 m, low 0-10 m, below danger5-15 m, danger10-20 m and above danger15-20 m. The five fuzzy membership functions for flow rate input are: very slow $0m^3s^{-1}$ -$25000m^3s^{-1}$, slow $0m^3s^{-1}$-$50000m^3s^{-1}$, normal $25000m^3s^{-1}$-$75000m^3s^{-1}$, fast $50000m^3s^{-1}$-$100000m^3s^{-1}$, and very fast $75000m^3s^{-1}$-$100000m^3s^{-1}$.Two outputs of this proposed system are: (release) control valve and drainage valve. The control valves for release and drainage output variables consist of five membership functions: fully closed 0-5, 25% Opened 0-50, 50% Opened 40-60, 75% Opened 50-90 and Fully Opened 70-100.

3.1. Fuzzifier

The input crisp values are compared by the fuzzifier with certain levels and generate linguistic values of each input variable for inference engine. The inference engine simulates human decision with fuzzy concepts, implication and rules of inference in fuzzy logic [7]. The occupied region description, membership functions and range for two input variables are given in Table 1 and Table 2.

TABLE 1
MEMBERSHIP FUNCTIONS AND RANGES OF INPUT VARIABLE WATER LEVEL (m)

Membership Function (MF)	Ranges	Region Occupied
Very Low	0-5	1
Low	0-10	1-2
Below Danger	5-15	2-3
Danger	10-20	3-4
Above Danger	15-20	4

TABLE 2
MEMBERSHIP FUNCTIONS AND RANGES OF INPUT VARIABLE FLOW RATE (m^3/s)

Membership Function (MF)	Ranges	Region Occupied
Very Slow	0 -25000	1
Slow	0-50000	1-2
Normal	25000-75000	2-3
Fast	50000-100000	3-4
Very Fast	75000-100000	4

For each input variable, five membership functions are used as shown in Fig. 2 and in Fig. 3.
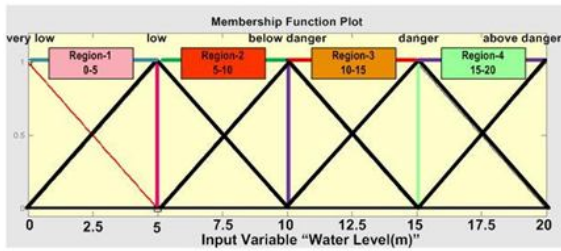
Fig.3 Plot of membership functions for input variable,
"WATER LEVEL"

The five membership functions, "very low", "low", "below danger", "danger" and "above danger" are used to show the various ranges of input fuzzy variable "WATER LEVEL" in a plot consisting of four regions as shown in Fig. 2.

The five membership functions, "very slow", "slow", "normal", "fast" and "very fast" are used to show the various ranges of input fuzzy variable "FLOW RATE" in a plot also consisting of four regions as shown in Fig. 3.
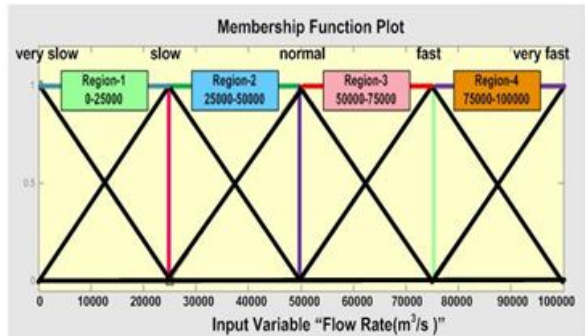
Fig. 4 Plot of membership functions for input variable, "FLOW RATE"

The linguistic values are the mapping values of the fuzzy input variables with the membership functions occupied in the regions. As we are using two variables, therefore four linguistic values are shown in Fig.4. The mapping of input fuzzy variables with the functions in four regions is listed in Table 3.

TABLE 3
LINGUISTIC VALUES OF FUZZIFIERS OUTPUTS IN ALL REGIONS

Input Variables	Linguistic Fuzzifiers Outputs	Region 1	Region 2	Region 3	Region 4
Water Level	f_1	$f_1[1]$	$f_1[2]$	$f_1[3]$	$f_1[4]$
	f_2	$f_1[2]$	$f_1[3]$	$f_1[4]$	$f_1[5]$
Flow Rate	f_3	$f_2[1]$	$f_2[2]$	$f_2[3]$	$f_2[4]$
	f_4	$f_2[2]$	$f_2[3]$	$f_2[4]$	$f_2[5]$

TABLE 4
RULE MAPPING FOR REGIONS OCCUPIED

Case No.	Regions Occupied		Rules $f_n[m]$= Membership value, where n=No. of input variable, m=No. of membership function MF occupied
	Water Level Input variable 1	Flow Rate Input variable 2	
1.	1	1	$R_1 = f_1 \wedge f_3 = f_1[1] \wedge f_2[1]$ $R_2 = f_1 \wedge f_4 = f_1[1] \wedge f_2[2]$ $R_3 = f_2 \wedge f_3 = f_1[2] \wedge f_2[1]$ $R_4 = f_2 \wedge f_4 = f_1[2] \wedge f_2[2]$
2.	1	2	$R_1 = f_1 \wedge f_3 = f_1[1] \wedge f_2[2]$ $R_2 = f_1 \wedge f_4 = f_1[1] \wedge f_2[3]$ $R_3 = f_2 \wedge f_3 = f_1[2] \wedge f_2[2]$ $R_4 = f_2 \wedge f_4 = f_1[2] \wedge f_2[3]$
3.	1	3	$R_1 = f_1 \wedge f_3 = f_1[1] \wedge f_2[3]$ $R_2 = f_1 \wedge f_4 = f_1[1] \wedge f_2[4]$ $R_3 = f_2 \wedge f_3 = f_1[2] \wedge f_2[3]$ $R_4 = f_2 \wedge f_4 = f_1[2] \wedge f_2[4]$
4.	1	4	$R_1 = f_1 \wedge f_3 = f_1[1] \wedge f_2[4]$ $R_2 = f_1 \wedge f_4 = f_1[1] \wedge f_2[5]$ $R_3 = f_2 \wedge f_3 = f_1[2] \wedge f_2[4]$ $R_4 = f_2 \wedge f_4 = f_1[2] \wedge f_2[5]$
5.	2	1	$R_1 = f_1 \wedge f_3 = f_1[2] \wedge f_2[1]$ $R_2 = f_1 \wedge f_4 = f_1[2] \wedge f_2[2]$ $R_3 = f_2 \wedge f_3 = f_1[3] \wedge f_2[1]$ $R_4 = f_2 \wedge f_4 = f_1[3] \wedge f_2[2]$
6.	2	2	$R_1 = f_1 \wedge f_3 = f_1[2] \wedge f_2[2]$ $R_2 = f_1 \wedge f_4 = f_1[2] \wedge f_2[3]$ $R_3 = f_2 \wedge f_3 = f_1[3] \wedge f_2[2]$ $R_4 = f_2 \wedge f_4 = f_1[3] \wedge f_2[3]$
7.	2	3	$R_1 = f_1 \wedge f_3 = f_1[2] \wedge f_2[3]$ $R_2 = f_1 \wedge f_4 = f_1[2] \wedge f_2[4]$ $R_3 = f_2 \wedge f_3 = f_1[3] \wedge f_2[3]$ $R_4 = f_2 \wedge f_4 = f_1[3] \wedge f_2[4]$
8.	2	4	$R_1 = f_1 \wedge f_3 = f_1[2] \wedge f_2[4]$ $R_2 = f_1 \wedge f_4 = f_1[2] \wedge f_2[5]$ $R_3 = f_2 \wedge f_3 = f_1[3] \wedge f_2[4]$ $R_4 = f_2 \wedge f_4 = f_1[3] \wedge f_2[5]$
9.	3	1	$R_1 = f_1 \wedge f_3 = f_1[3] \wedge f_2[1]$ $R_2 = f_1 \wedge f_4 = f_1[3] \wedge f_2[2]$ $R_3 = f_2 \wedge f_3 = f_1[4] \wedge f_2[1]$ $R_4 = f_2 \wedge f_4 = f_1[4] \wedge f_2[2]$
10.	3	2	$R_1 = f_1 \wedge f_3 = f_1[3] \wedge f_2[2]$ $R_2 = f_1 \wedge f_4 = f_1[3] \wedge f_2[3]$ $R_3 = f_2 \wedge f_3 = f_1[4] \wedge f_2[2]$ $R_4 = f_2 \wedge f_4 = f_1[4] \wedge f_2[3]$
11.	3	3	$R_1 = f_1 \wedge f_3 = f_1[3] \wedge f_2[3]$ $R_2 = f_1 \wedge f_4 = f_1[3] \wedge f_2[4]$ $R_3 = f_2 \wedge f_3 = f_1[4] \wedge f_2[3]$ $R_4 = f_2 \wedge f_4 = f_1[4] \wedge f_2[4]$

| Case No. | Regions Occupied | | Rules $f_n[m]$= Membership value, where n=No. of input variable, m=No. of membership function MF occupied |
	Water Level Input variable 1	Flow Rate Input variable 2	
12.	3	4	$R_1 = f_1 \wedge f_3 = f_1[3] \wedge f_2[4]$ $R_2 = f_1 \wedge f_4 = f_1[3] \wedge f_2[5]$ $R_3 = f_2 \wedge f_3 = f_1[4] \wedge f_2[4]$ $R_4 = f_2 \wedge f_4 = f_1[4] \wedge f_2[5]$
13.	4	1	$R_1 = f_1 \wedge f_3 = f_1[4] \wedge f_2[1]$ $R_2 = f_1 \wedge f_4 = f_1[4] \wedge f_2[2]$ $R_3 = f_2 \wedge f_3 = f_1[5] \wedge f_2[1]$ $R_4 = f_2 \wedge f_4 = f_1[5] \wedge f_2[2]$
14.	4	2	$R_1 = f_1 \wedge f_3 = f_1[4] \wedge f_2[2]$ $R_2 = f_1 \wedge f_4 = f_1[4] \wedge f_2[3]$ $R_3 = f_2 \wedge f_3 = f_1[5] \wedge f_2[2]$ $R_4 = f_2 \wedge f_4 = f_1[5] \wedge f_2[3]$
15.	4	3	$R_1 = f_1 \wedge f_3 = f_1[4] \wedge f_2[3]$ $R_2 = f_1 \wedge f_4 = f_1[4] \wedge f_2[4]$ $R_3 = f_2 \wedge f_3 = f_1[5] \wedge f_2[3]$ $R_4 = f_2 \wedge f_4 = f_1[5] \wedge f_2[4]$
16.	4	4	$R_1 = f_1 \wedge f_3 = f_1[4] \wedge f_2[4]$ $R_2 = f_1 \wedge f_4 = f_1[4] \wedge f_2[5]$ $R_3 = f_2 \wedge f_3 = f_1[5] \wedge f_2[4]$ $R_4 = f_2 \wedge f_4 = f_1[5] \wedge f_2[5]$

Fuzzification process for two input variables need two separate fuzzifiers. Each fuzzifier consists of: input voltage to crisp value converter, operational region for a crisp value detector, fuzzy set membership value mapping and selection arrangements [8]. The design of such a fuzzifier is shown in Fig. 5.
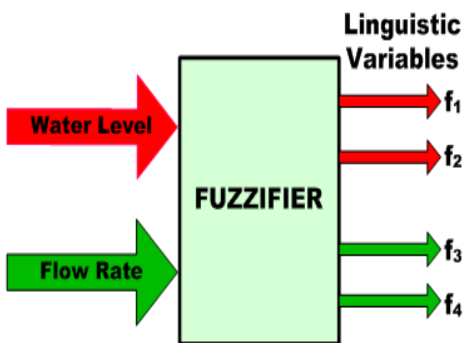
Fig. 5 Fuzzifier Block

Table 5 gives the working results of two fuzzifiers using the given values of input variables. These results are achieved using the fuzzifier design for water level and flow rate inputs shown in Fig. 6 (a) and Fig. 6 (b) [9].
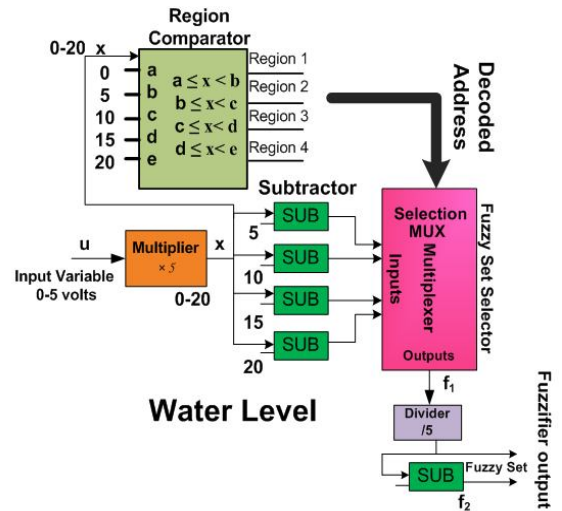
Fig. 6 (a) Design of fuzzifier for Water Level Input
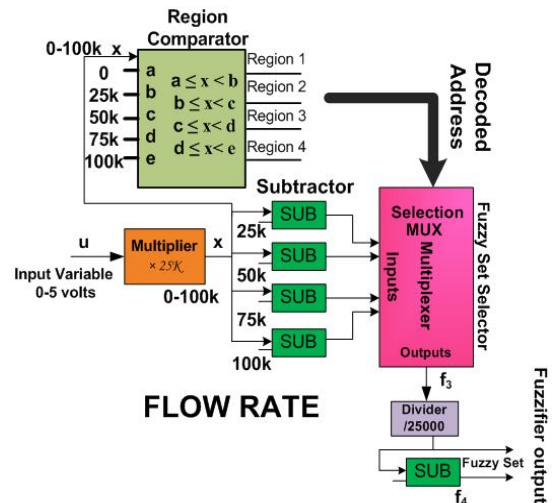
Fig. 6 (b) Design of fuzzifier for Flow Rate Input

TABLE 5
RESULTS OF FUZZIFICATION

Input Variables	Values	Region Selection	Fuzzy Set Calculation
Water Level	x=13	$10 \leq x < 15$ Region-3	$f_1=(15-13)/5=0.4$ $f_2=1-f_1=1-0.4=0.6$
Flow Rate	x=95000	$75000 \leq x < 100000$ Region-4	$f_3=(100000-95000)/25000=0.2$ $f_4=1-f_3=1-0.2=0.8$

3.2. Inference Engine

The inference engine contains four AND operators that select minimum value input for the output. This inference engine accepts four inputs from fuzzifier and applies the min-max composition to obtain the output R values. The min-max inference method uses min-AND operation

between the four inputs. Fig. 7 shows this type of inference process.

Number of active rules = **m^n**, where m = maximum number of overlapped fuzzy sets and n= number of inputs. For this design, m = 5 and n = 2, so the total number of active rules are 25.The total number of rules is equal to the product of number of functions accompanied by the input variables in their working range [10]. The two input variables described here consisted of five membership functions. Thus, 5 x 5 = 25 rules were required which are shown in Table 6.

TABLE 6
TOTAL NUMBER OF RULES

INPUTS		OUTPUTS	
Water Level (m)	Flow Rate (m³/s)	Out-flow (Valve)	Drain Valve
Very Low	Very Slow	Fully closed	Fully closed
Very Low	Slow	Fully closed	Fully closed
Very Low	Normal	Fully closed	Fully closed
Very Low	Fast	Fully closed	Fully closed
Very Low	Very Fast	Fully closed	Fully closed
Low	Very Slow	Fully closed	Fully closed
Low	Slow	Fully closed	Fully closed
Low	Normal	Fully closed	Fully closed
Low	Fast	Fully closed	Fully closed
Low	Very Fast	Fully closed	Fully closed
Below Danger	Very Slow	Fully closed	25% opened
Below Danger	Slow	25% opened	25% opened
Below Danger	Normal	25% opened	25% opened
Below Danger	Fast	50% opened	50% opened
Below Danger	Very Fast	50% opened	75% opened
Danger	Very Slow	50% opened	50% opened
Danger	Slow	50% opened	50% opened
Danger	Normal	75% opened	Fully opened
Danger	Fast	75% opened	Fully opened
Danger	Very Fast	75% opened	Fully opened
Above Danger	Very Slow	75% opened	75% opened
Above Danger	Slow	75% opened	Fully opened
Above Danger	Normal	Fully opened	Fully opened
Above Danger	Fast	Fully opened	Fully opened
Above Danger	Very Fast	Fully opened	Fully opened

In this case only 4 rules are required for the particular values of two variables because each value of two variables in a region corresponds to mapping of two functions. The corresponding mapping values of f_1 [3], f_1 [4], f_2[2], f_2[3] were used to establish the 4 rules. Here f_1 [3] means the corresponding mapping value of membership function "Below Danger" of water level in region-3 and the similar definitions are for the others.

$R_1 = f_1 \wedge f_3 = f_1[3] \wedge f_2[4] = 0.4 \wedge 0.2 = 0.2$
$R_2 = f_1 \wedge f_4 = f_1[3] \wedge f_2[5] = 0.4 \wedge 0.8 = 0.4$
$R_3 = f_2 \wedge f_3 = f_1[4] \wedge f_2[4] = 0.6 \wedge 0.2 = 0.2$
$R_4 = f_2 \wedge f_4 = f_1[4] \wedge f_2[5] = 0.6 \wedge 0.8 = 0.6$
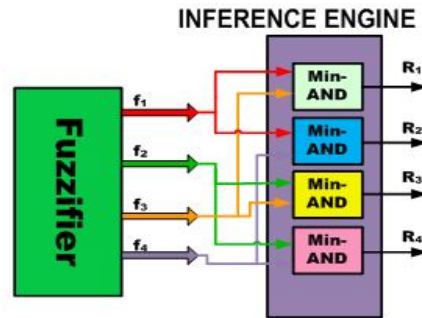
Fig. 7 Block Diagram of Inference Engine

3.3. Rule Selector

The rule selector receives two crisp values of water level and flow rate. It gives singleton values of output functions under algorithm rules applied on design model. For two variables, four rules are needed to find the corresponding singleton values S_1, S_2, S_3 and S_4 for each variable according to these rules are listed in Table 7.

TABLE 7
ILLUSTRATION OF RULES APPLIED MODEL

Rule No.	INPUTS		SINGLETON VALUES OF OUTPUTS		Singleton Values
	Water Level	Flow Rate	Release Control Valve	Drainage Valve	
1	Below Danger	Fast	0.50 =50% opened	50% opened =0.50	S_1
2	Below Danger	Very Fast	0.50 =50% opened	75% opened =0.75	S_2
3	Danger	Fast	0.75 =75% opened	Fully Opened =1.0	S_3
4	Danger	Very Fast	0.75 =75% opened	Fully Opened =1.0	S_4

The rule base accepts two crisp input values, distributes the universe of discourse into regions with each region containing two fuzzy variables, fires the rules, and gives the output singleton values corresponding to each output variable. Fig. 8 shows the main block diagram of the Rule Base.
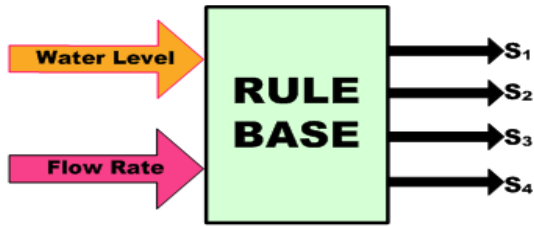
Fig. 8 Rule Base

3.4. Deffuzifier

In this system, two defuzzifiers control the actuators; Release (Valve Control and Drainage Valve. The membership functions of the two output variables are shown in Fig. 7 to Fig.9, and the detail of each plot is given in Table 8.

TABLE 8
OUTPUT VARIABLES MEMBERSHIP FUNCTIONS

MFs	Range	Release(Valve)	Drain Valve
MF1	0-5	Fully Closed	Fully Closed
MF2	0-50	25% Opened	25% Opened
MF3	40-60	50% Opened	50% Opened
MF4	50-90	75% Opened	75% Opened
MF5	70-100	Fully Opened	Fully Opened
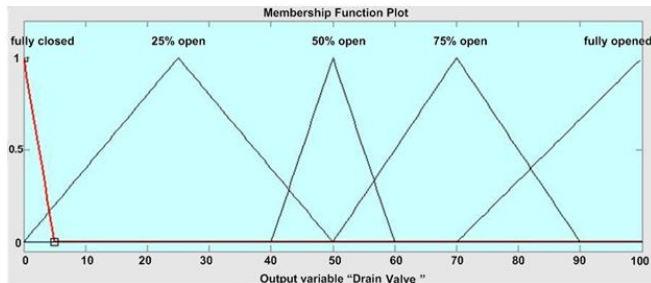
Fig. 9 Plot of Membership Functions for Output Variable, "Drain Valve"
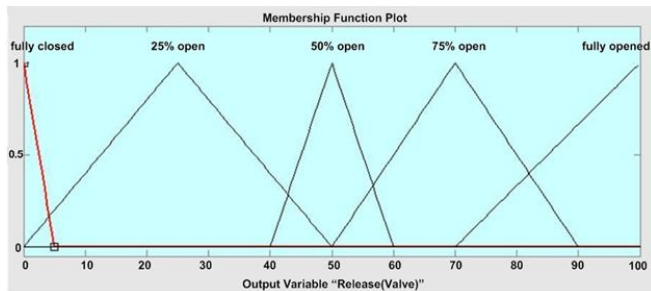
Fig. 10 Plot of Membership Functions for Output Variable, "Release (Valve Control)"

The defuzzification process provides the crisp value outputs after estimating its inputs [11]. In this system 8 inputs are given to each of the two defuzzifiers. Four values of R_1, R_2, R_3, R_4 from the outputs of inference engine and four values S_1, S_2, S_3, and S4 from the rule selector are shown in Fig. 10. Each defuzzifier estimates the crisp value output according to the center of average (C.O.A) method using the mathematical expression, $\Sigma S_i * R_i / \Sigma R_i$, where i = 1 to 4. Each output variable membership function plot consists of five functions with the same range values for simplification.
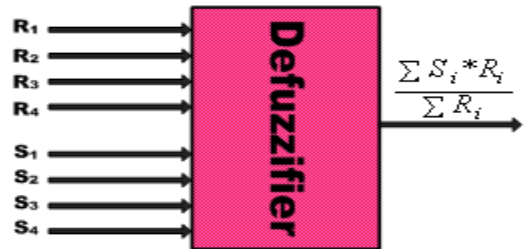
Fig. 11 Defuzzifier Block

Fig. 11 shows the design arrangement of a defuzzifier. One defuzzifier consists of: one adder for ΣR_i, four multipliers for the product of S_i*R_i, one adder for $\Sigma S_i * R_i$, and one divider for $\Sigma S_i * R_i / \Sigma R_i$. Finally a defuzzifier gives the estimated crisp value output.

4. RESULTS AND DISCUSSION

The designed values for two outputs; Release (Valve Control) and Drain Valve are given in the Table 9 and Table 10. According to the results of inference engine $\Sigma R_i = R_1+R_2+R_3+R_4 = 0.2+0.4 +0.2 +0.6 = 1.4$

TABLE 9
DESIGNED VALUE FOR CONTROL VALVE (RELEASE)

i	R_i	S_i	$R_i * S_i$
1	0.2	0.50	0.10
2	0.4	0.50	0.20
3	0.2	0.75	0.15
4	0.6	0.75	0.45

$\Sigma S_i * R_i = 0.90$

$\Sigma S_i * R_i / \Sigma R_i = 0.90/1.4 = 0.6428 = 64.28\%$
of the Valve will be Opened.

TABLE 10
DESIGNED VALUE FOR DRAIN VALVE

i	R_i	S_i	$R_i * S_i$
1	0.2	0.50	0.1
2	0.4	0.75	0.3
3	0.2	1.0	0.2
4	0.6	1.0	0.6

$\Sigma S_i * R_i = 1.2$

$\Sigma S_i * R_i / \Sigma R_i = 1.2/1.4 = 0.8571 = 85.71 \%$ of Drain Valve.

Using mathematical expression $\Sigma S_i * R_i / \Sigma R_i$ the crisp values for output variables were determined and the results were found according to the MATLAB simulation as shown in Fig. 12. These results are compared in Table 12 and found correct according to the design model. MATLAB simulation was adapted according to the arrangement of membership functions for four rules as given in Table 11.

TABLE 11
ARRANGEMENT OF MEMBERSHIP FUNCTIONS FOR-SIMULATION

Rule No.	INPUTS		OUTPUTS	
	Water Level	Flow Rate	Release(Valve)	Drain Valve
1	Below Danger	Fast	50% opened	50% opened
2	Below Danger	Very Fast	50% opened	75% opened
3	Danger	Fast	75% opened	Fully Opened
4	Danger	Very Fast	75% opened	Fully Opened

In Fig. 12 the same values of input variables, Water Level = 13, and Flow Rate = 95000 are shown. Various values of input and output variables match the dependency scheme of the system design. The simulated values were checked using MATLAB-Rule viewer as shown in Fig. 12.

The correctness of results shows the validity of the simplified design work for processing system using control system.
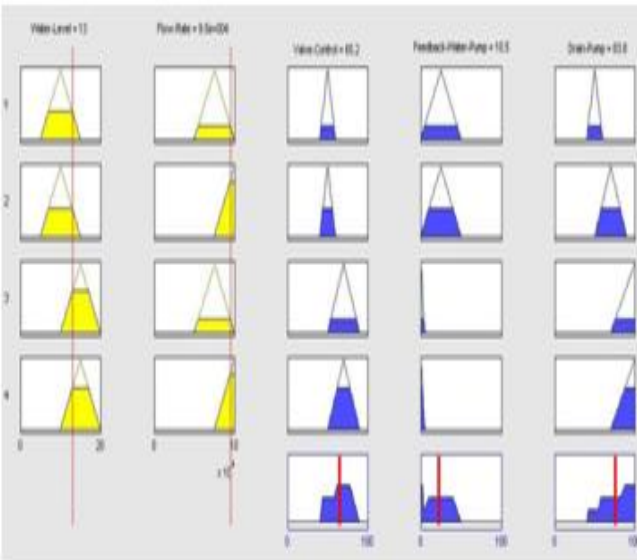
Fig. 12 MATLAB-Rule Viewer

TABLE 12
COMPARISON OF SIMULATED AND CALCULATED RESULT

Result	Release (Control Valve)	Drain Valve
Design Values	64.2	85.7
MATLAB Simulation	65.2	83.8
% error	1.55	2.21

4.1. Simulation Graphs Discussion

This system was simulated for the given range of input variables. The given value of: Water Level = 13 lies in region 3 of the range 10-15 and Flow Rate = 95000 lies in region 4 of the range 75000-100000. The four rules were applied for MATLAB simulation according to this range scheme. In this design model, the release and drain control valves depends upon the selected values of water level and flow rate. The simulated and calculated results are according to the dependence scheme.

Fig. 13(a) shows that the control valve is directly proportional to water level and it does not depend upon the flow rate. Fig. 13(b) shows that the drain valve system is directly proportional to flow rate.
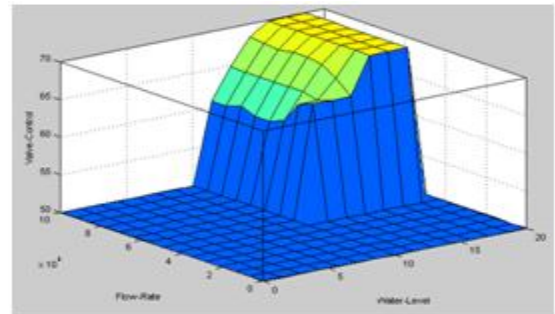
Fig. 13(a) Plot between Water Level - Flow Rate Release Valve Opened/Closed
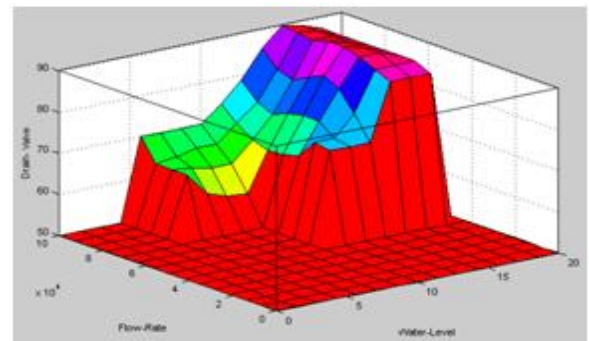
Fig. 14 Plot between Water Level and Flow Rate Drain Valve

5. CONCLUSION

Both the design model and simulation results are same. The designed system can be extended for any number of inputs and outputs. The drain valve control output can be utilized further for land irrigation according to the need and water release control valve for electric generation to fulfill the dire need of this system in automation.

The design work is being carried out to design state of the art fuzzy logic Hydro-Electric control system in future using FPGAs.

REFERENCES

[1] E. Frias-Martinez, G. Magoulas, S. Chen and R. Macredie, "Modeling human behavior in user-adaptive systems: Recent advances using soft computing techniques" *Expert Syst. Appl.***29**:2 (2005) 320-329.

[2] V. Ramani and M. Rom, "Fuzzy Logic Model on Operation and Control of Hydro-Power Dams in Malaysia" ICCES vol.4 2007, p.p 31-39.

[3] Suharyanto and Goulter, "Use of fuzzy set theory for consideration of storage non-specificity in stochastic dynamic programming for reservoir operation "*Proceeding of Seventh IAHR International Symposium on Stochastic Hydraulics*, Mackay/Queensland/Australia, 129-36.

[4] Umamahesh and Chandramouli. Fuzzy dynamic programming model for optimal operation of a multipurpose reservoir, *Journal of Water Resources Planning and Management*, 18(2): 14-20.

[5] Hasebe and Nagayama (2002). Reservoir operation using the neural network and fuzzy systems for dam control and operation support, *Advanced in Engineering Software*, 33(5): (1996) 245-260.

[6] Suvo, "How does a Hydroelectric Dam work" http://www.brighthub.com/engineering/civil/articles/42794.aspx 2009.

[7] Shabiul Islam, Shakowat,"Development of a Fuzzy Logic Controller Algorithm for Air-conditioning System", *ICSE 2006 Proc 2006 IEEE*.

[8] M. Saleem Khan and Khaled Benkrid, "A proposed Grinding admixing System using Fuzzy Time Control Discrete Event Model for Industrial Application", Lecture Notes in Engineering and Computer Science vol. 2175 2009, p.p. 1231-1236, *Directory of Open Access Journals (DOAJ)*

[9] M. Saleem Khan and Khaled Benkrid, "Multi-Dimensional Supervisory Fuzzy Logic Control DEV, Processing System for Industrial Applications" Lecture Notes in Engineering and Computer Science vol. 2175, p.p. 1208-1217, *Directory of Open Access Journals (DOAJ)*

[10] M. Berthold, D. Hand, "Intelligent Data Analysis, 2nd ed.", *Springer-Verlag, Heidelberg* (2006).

[11] M. Saleem Khan and Khaled Benkrid, "Design of Liquids Mixing Control System using Fuzzy Time Control Discrete Event Model for Industrial Applications", World Academy of Science, Engineering and Technology vol. 72 2010, p.p. 545-553, *Directory of Open Access Journals (DOAJ)*.

Mr. M. Abbas is working as Visiting Faculty in the Department of Electrical Engineering in National University of FAST Lahore and also Lecturer *Internee in Physics, Government Islamia College Lahore, Pakistan*. He completed his Master Degree in Physics with Specialization in Advanced Electronics from Government Islamia College Civil Lines, Lahore. His research interests include FPGAs and Fuzzy Logic Control System and Industrial Applications.

(e-mail: shairabbas@yahoo.com)

Dr. M. Saleem Khan is an Assistant Professor at GC University Lahore Pakistan. He availed research fellowship from The School of Electronics & Engineering, University of Edinburgh, UK and completed his Ph.D thesis in the field of control systems: design, simulation and analysis in local and distributed environment. He contributed his services on various projects in the field of Advanced Electronics and Communication. His research interests include control systems design and industrial applications. He promoted a large team of Electronics researchers and organized this field in his country. Mr. Khan had also been served as a Senior Scientific officer in a classified defense research organisation in his country.
(e-mail mskgcu@yahoo.com)

Mr. Nasir Ali is an Assistant Professor in DS College Lahore. Mr.Nasir also served as a Scientific Officer in a classified defense research organisation in his country. He has been involved in the various research projects of Electronics.

(e-mail nasirali8787@yahoo.com

Analysis and Simulation of DVR by Using AC Chopper for Distribution System

D.R.K.Praveen, T.Vijay Muni, V V Satyanarayana Rao.R

Abstract— This paper deals with modelling and digital simulation of AC chopper based Dynamic Voltage Restorer (DVR). The control of DVR that injects a voltage in series with a distribution feeder is presented. DVR is a power electronic controller that can protect sensitive loads from disturbances of voltage fluctuations in supply system. It is observed that DVR can regulate the voltage at the load. Circuit model is developed for the single phase model and the same is used for simulation studies. It is tested for different capacity loads.

Index Terms— AC chopper, Dynamic Voltage Restorer (DVR), Sensitive loads, Series Compensation, Simulink, Voltage sags, Voltage swells.

——————————————— ◆ ———————————————

1 INTRODUCTION

Power Quality problems encompass a wide range of disturbances such as voltage sags/swells, flicker, harmonics distortion, impulse transient, and interruptions.

Voltage sags can occur at any instant of time, with amplitudes ranging from 10 – 90% and a duration lasting for half a cycle to one minute. Voltage swell, on the other hand, is defined as a *swell* is defined as an increase in rms voltage or current at the power frequency for durations from 0.5 cycles to 1 min. typical magnitudes are between 1.1 and 1.8 up. Swell magnitude is also described by its remaining voltage, in this case, always greater than 1.0.

Voltage swells are not as important as voltage sags because they are less common in distribution systems. Voltage sag and swell can cause sensitive equipment (such as found in semiconductor or chemical plants) to fail, or shutdown, as well as create a large current unbalance that could blow fuses or trip breakers.

These effects can be very expensive for the customer, ranging from minor quality variations to production downtime and equipment damage. There are many different methods to mitigate voltage sags and swells, but the use of a custom Power device is considered to be the most efficient method. Switching off a large inductive load or Energizing a large capacitor bank is a typical system event that causes swells.

This paper introduces a power electronic converter based series compensator that can protect critical loads from all supply side voltage disturbances other than outages is called a dynamic voltage restorer (DVR). The restorer is capable of generating or absorbing independently controllable real and reactive power at its AC output terminal. This device employs AC chopper or AC voltage controller structure. It injects output voltages in series with the distribution feeder at the given load location. In 1996, Westinghouse Electric Corporation installed world's first dynamic voltage restorer in 12.47kV substations in Anderson, USA. This was installed to provide voltage correction to an automated rug manufacturing plant [1, 2].

This paper extends the concept of dynamic voltage restorer to maintain the load voltage near the rated value. It can also perform the primary functions of the restorer, i.e., to protect the load from temporary sag and swell. This device is a DVR. The output of the chopper voltage will inject to the distribution feeder through the injecting transformer in order to maintain the constant voltage. For the control operation of DVR

————————————————————

- *D.R.K.Praveen received Bachelors Degree from Gulbarga University. He has 11 years experience in various positions. Presently he working as Associate Professor & Head in Electrical & Electronics Department, Sri Sarathi Institute of Engineering & Technology, Nuzvid, India. PH-09440875693. E-mail: drkpraveenatnet@yahoo.com*
- *T.Vijay Muni received Masters Degree in Power & Industrial Drives from JNT University, Kakinada, India.Presently he working as Assistant Professor in Electrical Department , Sri Sarathi Institute of Engineering & Technology, Nuzvid, India.PH-09000055144. E-mail: www.vijaymuni@gmail.com*
- *V V Satyanarayana Rao.R received Masters Degree in Power Engineering from JNT University, Hyderabad, India in 2007. He is pursing the Ph.D in Power Systems at SCSVMV University, Kanchipuram, Chennai, India. Presently he is working as Senior Assistant Professor in Sri Sarathi Institute of Engineering & technology, Nuzvid, India. His areas of interest are FACTS, Power Electronics and Deregulated Power Systems. PH-09396677506. E-mail: rvvs_ssiet@yahoo.com*

designs the on-off control of ac voltage controller with different duty cycles (from 10% to 99.99%).

The actual implementation of the DVR using chopper raises additional issues of harmonics. The filter is used to eliminate these harmonics. The operation of ac chopper is given by [3].

The above literature does not deal with the modelling of DVR system using simulink. In this work, an attempt is made to model and simulate the DVR system using Simulink.
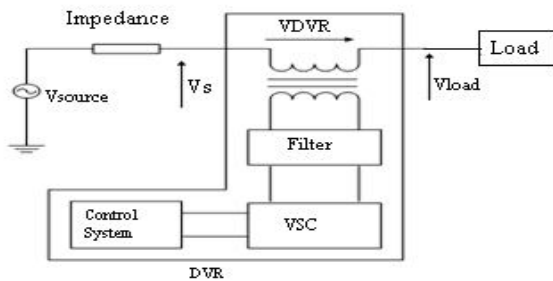
Fig 1: Schematic Diagram of DVR

2 PROBLEM FORMULATION

A DVR is a recently proposed series connected solid state device that injects the voltage in to the system in order to regulate the load side voltage. It is normally installed in a distribution system between supply in feeder at critical load location [4].
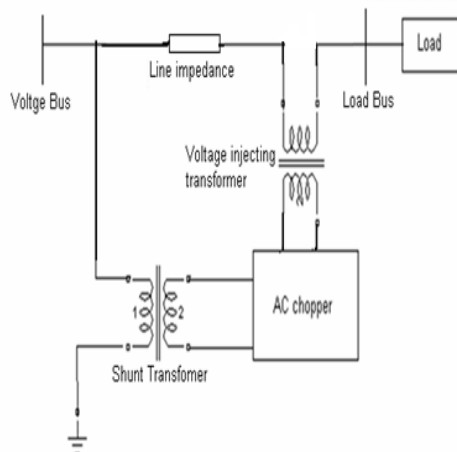
Fig 2: Block Diagram of DVR using chopper

The general configuration of a DVR consists of a shunt transformer, AC chopper, injection/booster transformer [5] and a harmonic filter as shown in fig2.

The voltage at load terminal (V_L) is equal to sum of supply voltage (V_{Supply}) and the output voltage of the chopper ($V_{injected}$).

$$V_L = V_{Supply} + V_{injected} \qquad (1)$$

The basic function of the injection transformer is to increase the voltage by injecting a series voltage derived as output of the chopper while isolating the DVR circuit from the distribution networks.

The rating of the injection transformer is an important factor while deciding the DVR performance, since it limits the maximum compensation ability of the DVR. Further the leakage inductance of the transformer brings to a low value to reduce the voltage drop across the transformer. In order to reduce the saturation of the injection transformer under normal operating conditions it is designed to handle a flux, which is higher than the normal maximum flux requirement. The winding configuration of the injection transformer mainly depends upon the upstream distribution transformer.

For any type of load (i.e. R, RL and RC) the supply voltage is same because there is no voltage drop due to source impedance except the drop in line impedance.

3 SIMULATION RESULTS

A typical system data is given in table1 for a radial feeder connected to a load along with two transformers and filter. Digital simulation is done by using Simulink and the results are presented here.

Table1. System Data

Supply Voltage	6350V
Line Impedance	1+j0.001Ω
Shunt Transformer Turns Ratio	1:5
Injecting Transformer Turns Ratio	1:1
Filter Inducatnce	7μH
Filter Capacitance	0.09046F
Load Resistance	635Ω

3.1 DVR Using AC Chopper without Filter:

DVR using AC chopper without filter is shown in Figure 3a. The output of the Chopper is injected into line through a series transformer. The voltage across load for different duty cycles is shown in Figure 3b. FFT analysis for the output voltage is shown in Figure 3c. THD is 30.17% for 90% duty cycle.
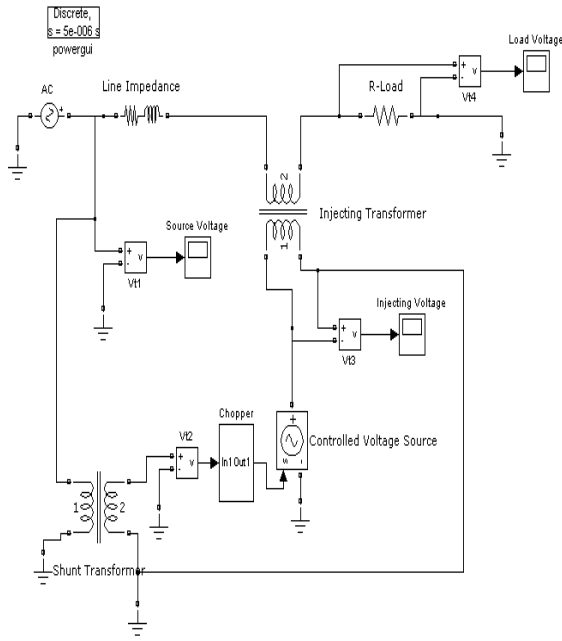
Fig 3a: DVR Circuit without Filter

Fig 3b: Waveform of load voltage with out filter

Fig 3c: FFT Analysis for harmonics load voltage without filter

The characteristics of on time and off time versus rms voltage as shown in fig 3d and 3e.
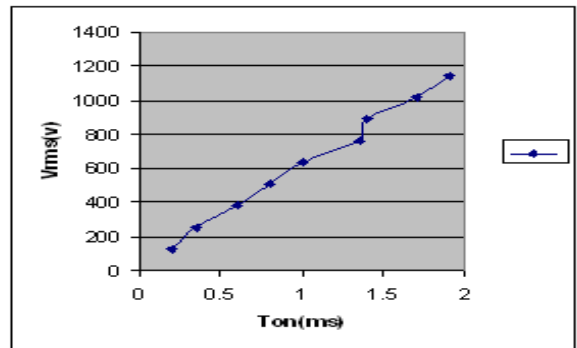
Fig.3d: Ton versus Vrms characteristics
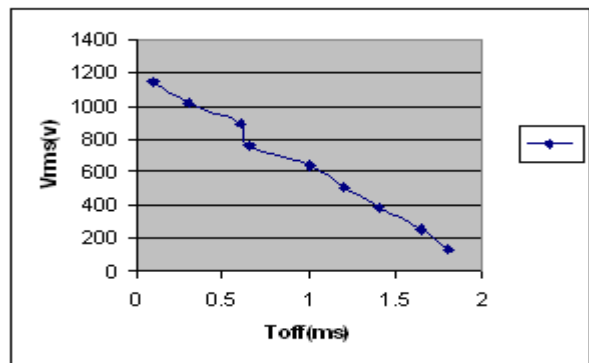
Fig.3e. Toff versus Vrms characteristics
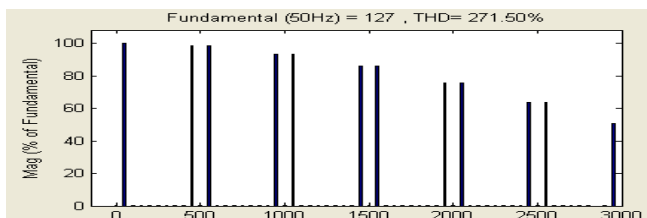
3.2 DVR Using AC Chopper with Filter:

DVR system with LC filter is shown in Figure 3f. Injected voltage and the voltage across loads for ±25% of the source voltages for different duty cycles are shown in Figure 3g and 3h. FFT analysis for the output voltages is shown in Figure 3i and 3j. THD Reduces to 0.01%. Thus the harmonics are reduced from 30.17% to 0.01%.
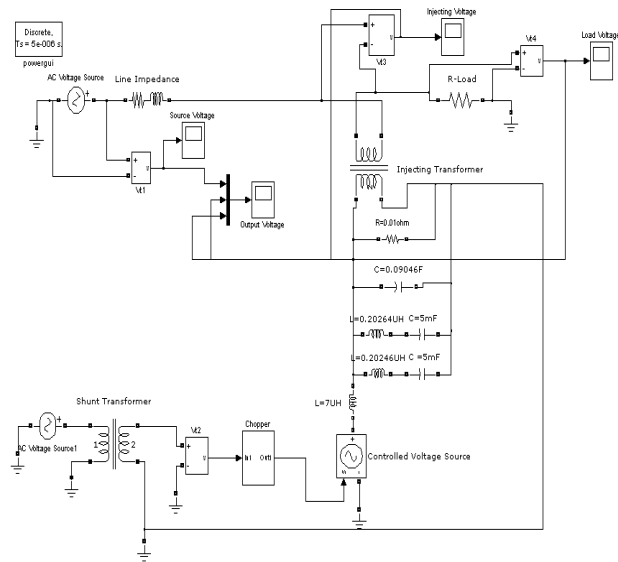
Fig.3f: DVR Circuit with filter
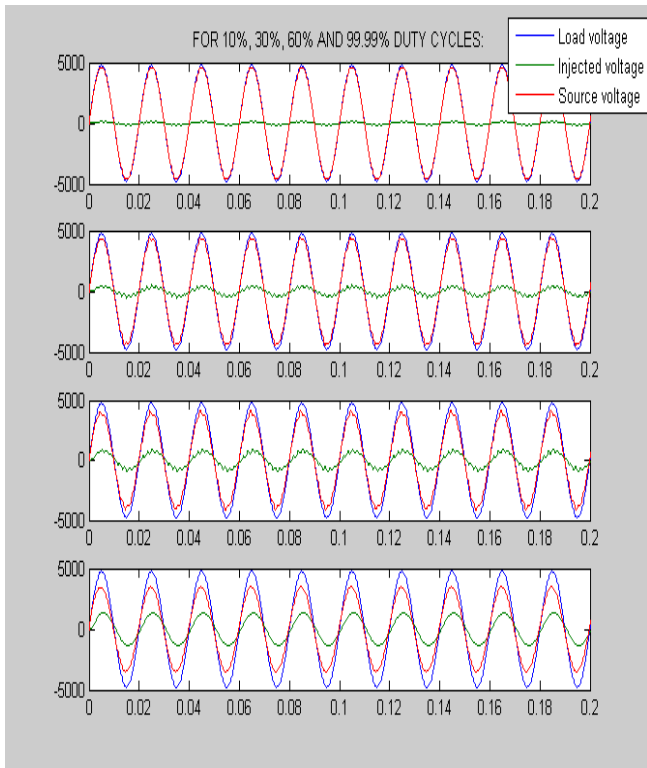
Fig.3g.Waveform for Input voltage is -25% (4762.5V)
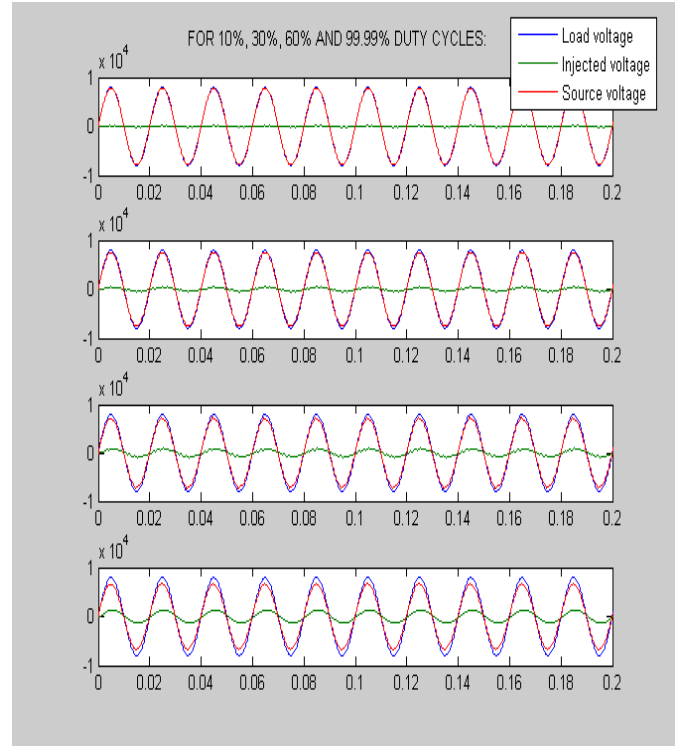
Fig.3h.Waveform for input voltage is +25% (7937.5V)
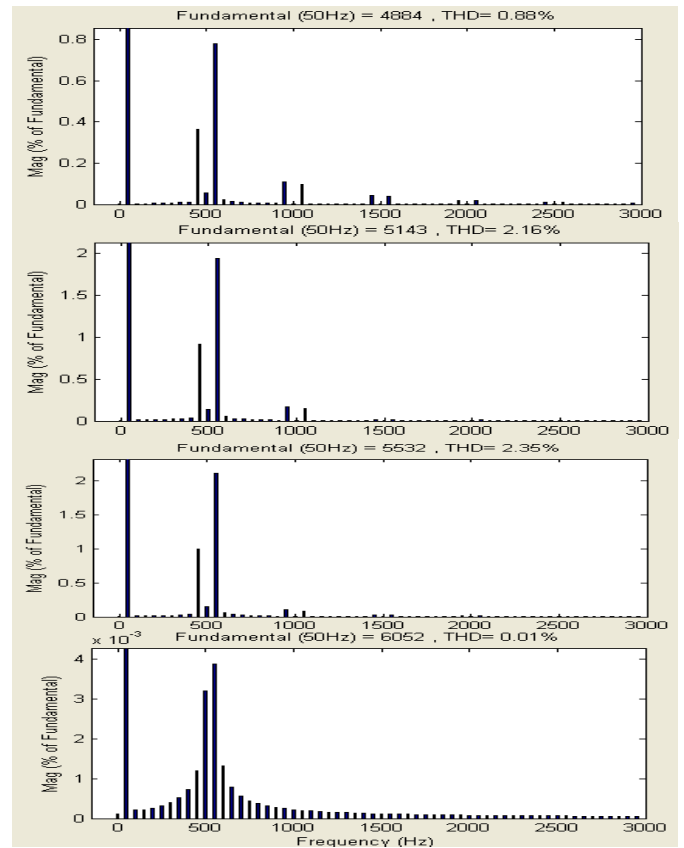
Fig.3i: FFT Analysis -25% of Input voltage
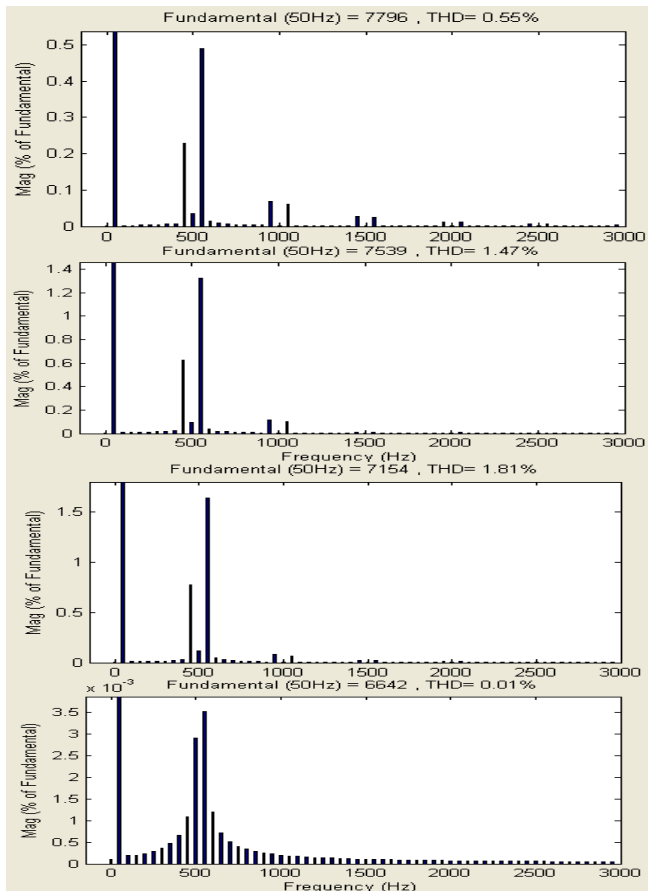
Fig.3j: FFT Analysis +25% of Input voltage

The results are given in Table 2 for supply voltage variation from +25% to -25% above the rated voltage in steps of ±5%. The corrected voltage across the load, variation and the THDs are shown for each case.

4 CONCLUSION

The load voltage is controlled to within a maximum value of 4.69% the THD within 2.05%.The use of AC chopper for voltage restoration is much simpler then to previous used DVR schemes. The authors feel that this method is for the first time used in DVR schemes .It is the authors' original contribution. The work will be extended to 3-phase system and reported later.

ACKNOWLEDGEMENT

REFERENCES

[1] N.H. Woodley, L. Morgan and A. Sundaram, *"Experience with an inverter-base dynamic voltage restorer,"* IEEE Trans. Power De-livery, Vol. 14, No.3, pp.1181-1185, 1999.

[2] Lei B.H., Choi S. S., and Vilathgamuwa D.M.: *"Design considerations on the line-side filter used in the dynamic voltage restorer"*.IEEE Proceedings — Generation, Transmission, and Distribution, vol.148, pp. 1–7, Jan. 2001.

[3] Mozder, A. JR. – Bose, B. K., "*Three-Phase AC Power Control Using Power Transistors,*" IEEE Trans. Ind. Appl., IA-12 (1976), pp. 499–505.

[4] Chan, K., 1998. "*Technical and performance aspects of a dynamic voltage restorer*". In IEE Half DayColloquium on Dynamic Voltage Restorers- Replacing Those Missing Cycles, pp: 5/1-525.

[5] Kularatna N.: "*Power Electronics Design Handbook: Low-Power Components and Applications*". Boston: Newnes, 1998.

Table 2. Look up Table

Duty Cycle (%)	Supply Voltage (V)		Load Voltage (V)		THD (%)
	Total voltage at supply	Voltage injected in Percentage	Total voltage at load	Load variation in Percentage	
99.99	7937.5	+25	6642	+4.6	0.01
80	7620	+20	6581	+3.6	1.23
50	7302.5	+15	6648	+4.69	2.05
30	6985	+10	6588	+3.74	1.68
5	6667.5	+5	6528	+3.8	0.33
99.99	4762.5	-25	6052	-4.6	0.01
80	5080	-20	6108	-3.8	1.85
50	5397.5	-15	6036	-4.9	2.26
30	5715	-10	6094	-4.0	1.82
5	6032.5	-5	6087	-4.1	0.36

Advanced Coma Patient Monitoring System

Naveen Kansal, Hardeep Singh Dhillon

Abstract- Embedded system is becoming an integral part of Engineering design process for efficient analysis and effective operation. From data analysis to hardware work, everywhere embedded products are the main interest because of its reliability and time bound perfection. There is not much time with anyone now a day to give enough in all aspects, so demand of embedded products which serve as we want is high on demand. Further it describes the design of an embedded system for the control of Temperature & Light intensity with continuous monitoring in a single system using sensors, microcontroller and LCD. It describes the controlling action incorporated in the hardware to control any device connected when specific conditions are met. Further set up is made such that data can be stored for future offline analysis.
The hardware developed namely "Advanced Coma Patient Monitoring System" is very advanced product related to physical changes in body movement of the patient and gives Warning in form of alarm and display on the LCD in less than one second time. It also passes a sms to a person sitting at the distant place if there exists any movement in any body part of the patient.

Index Terms── Coma, Patient Monitoring, Advanced Coma Monitoring, Coma Patient, Non intrusive system, image processing based sytem, Future coma product.

———————————— ◆ ————————————

1 INTRODUCTION

A coma, sometimes also called persistent vegetative state, is a profound or deep state of unconsciousness. Persistent vegetative state is not brain-death. An individual in a state of coma is alive but unable to move or respond to his or her environment. Coma may occur as a complication of an underlying illness, or as a result of injuries, such as head trauma. Individuals in such a state have lost their thinking abilities and awareness of their surroundings, but retain non-cognitive function and normal sleep patterns. Even though those in a persistent vegetative state lose their higher brain functions, other key functions such as breathing and circulation remain relatively intact. Spontaneous movements may occur, and the eyes may open in response to external stimuli. Individuals may even occasionally grimace, cry, or laugh. Although individuals in a persistent vegetative state may appear somewhat normal, they do not speak and they are unable to respond to commands. So due to very rare change in physical movement of the coma patient we have to monitor him regularly. But now with the help of this system we can monitor the patient regularly and if there exist any physical change in body of the coma patient then it will alarm the signal and send a sms through a CDMA mobile to a person at a distant place.

2 BLOCK DIAGRAM OF THE SYSTEM

Fig. 1 Block diagram of Advanced coma patient monitoring

―――――――――――

- **Naveen Kansal is currently pursuing his PHD (Electronics) from Thapar University, Patiala** and has completed his Masters of Engg (batch 2008-10), E&EC Deptt, from P.E.C University of Technology (formally Punjab Engg College), Chandigarh Sector-12, Chandigarh-160012, INDIA.

- **Hardeep Singh Dhillon** graduated from Sri Sukhmani Institute of Engg and Tech, Dera Bassi, India, and completed his Masters of Engg (batch 2008-10), E&EC Deptt, from **P.E.C University of Technology** (formally Punjab Engg College), Chandigarh Sector-12, Chandigarh-160012, INDIA.

system.

3 IMAGE PROCESSING

As The Pattern recognition and pattern rejection algorithms use MATLAB for use in geographic information system images and map. Methods of algorithm selection in this research will be based on critical review literature on image preprocessing, pattern recognition using geometric algo-

rithm, line detection, extraction of curve lines, semantic retrieval by spatial relationships, and structural object recognition using shape-form shading. The results of this research will give a user an complete knowledge of which pattern recognition algorithm will best fit in analyzing geometric and structural pattern from a given image. The conclusion of the research will show which among the pattern recognition and rejection algorithms using MATLAB will produce the best result when looking for a specific pattern.

4 INPUT AND OUTPUT UNITS

This system will consist of an input unit i.e Digital camera to study the movement in any part of the body of patient.

For indication of warning we will use three approaches i.e. one by blowing alarm, by displaying message on the LCD and third by sending sms through the CDMA mobile to a distant person. Moreover the warning will be deactivated manually rather than automatically. So for this purpose a deactivation switch will be used to deactivate warning.

5 SOFTWARE RESULTS

5.1 This window shows Original Image and changes in Original Image

Fig. 2 This window is used for noting the changes in the original image of the patient and after noting it a sms will be automatically passed to a person.

5.2 This window shows the physical changes in the body of the patient
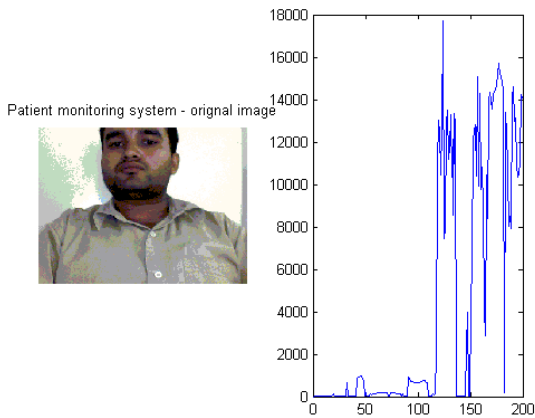
Fig. 3 This window is used to show physical changes in body.
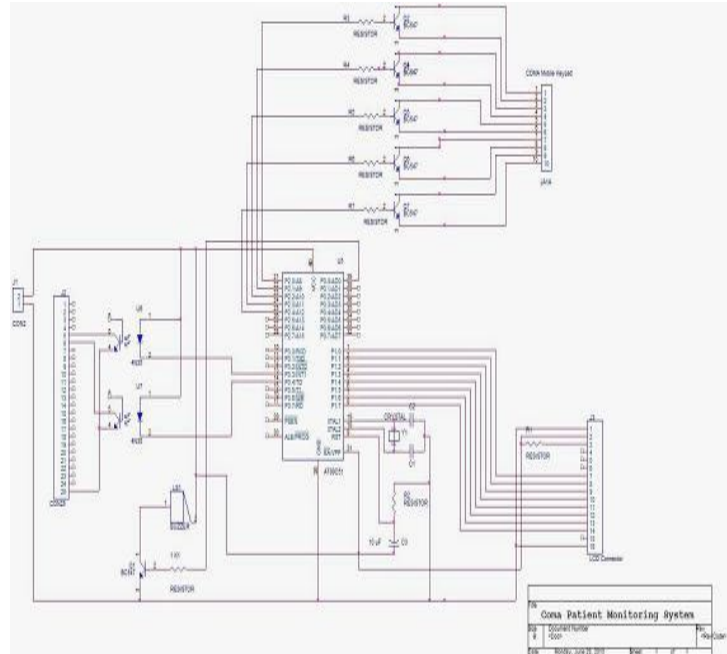
6 CIRCUIT DIAGRAM

Fig. 4 Circuit diagram of Advanced coma patient monitoring system.

7 HARDWARE DEVELOPED

Fig. 5 Complete hardware of Advanced coma patient

monitoring system.

8 CONCLUSION

With the help of this system we can monitor the whole body of a coma patient and if there exists a change in any part of the body of a patient then with the help of image processing using MATLAB software, it will automatically generates an alarm or send a sms via CDMA mobile to a person sitting at a distant place.

REFERENCES

[1] A. Goswami, T.Bezboruah and K.C. Sarma, **"Design of an embedded system for monitoring and controlling temperature and light"**, proceedings. of IJEER, Vol 1, No. 1, pp. 27-36, 2009

[2] B.Yuksekkaya et al., **"A GSM, Internet and Speech Controlled Wireless Interactive Home Automation System"**, IEEE Transactions on Consumer Electronics, Vol. 52, No. 3, August, 2006.

[3] Carlos Sandoval Zuria et al., **"MATLAB based image processing lab experiments"**, Computer Graphics proceedings, Annual Conference Series, pp. 297-305, 1995.

[4] Lourdes A. Funtanilla, **"GIS pattern recognition and rejection analysis using MATLAB"**, Proc. of the 2nd IEEE/EURASIP Conference on Pattern Recognition, Southampton, UK, 19-20 Sept. 1996.

[5] Muhammad Ali Mazidi and Janice Gillispe Mazidi, **"The 8051microcontroller and embedded systems"**, Pearson education ltd., India, pp. 35-105, 2006.

[6] Nassir H. Salman, **"New Image Processing Toolbox using MATLAB codes"**, Springer-Verlag Berlin, Hidelberg, pp. 126-130, 1991.

[7] Nahed H. Solouma et al., **"Robust computer-assisted laser treatment using real time retinal tracking"** Proceedings of the 23rd annual EMBS international conference, vol. 45, no.1, pp.105-117, 2001.

[8] Yan-Fang Li et al., **"A low-power 20GHz static frequency divider with programmable input sensitivity,"** proc. RFIC/IMS Symposium, pp. 235-238, 2003.

BIOBRAPHIES

Hardeep Singh Dhillon S/O Late. S. Jaspal Singh Dhillon was born in Village Sundran, P.O Mubarikpur, Tehsil Dera Bassi (Punjab) in India, on April 6, 1986. He graduated from Sri Sukhmani Institute of Engg and Tech, Dera Bassi, India, and completed his Masters of Engg (batch 2008-10), E&EC Deptt, from **P.E.C University of Technology** (formally Punjab Engg College), Chandigarh Sector-12, Chandigarh-160012, INDIA.

Actively Participated at various extracurricular events including participation in **8th National Convention of ISTE Students** from (21 Oct 2005 to 23 Oct 2005) at from Baba Banda Singh Bahadar Engineering college Fatehgarh Sahib (Punjab) INDIA. He has successfully completed his Six Weeks Industrial training using MODEL SIM software from (12 June 2006 to 22 July 2006) at **Semiconductor Complex Ltd S.A.S. Nagar Mohali, India** (*A Government of India Enterprise*). Also during his Six Month Industrial Project he developed Infra Red based Velocity Measurement system for Split Hopkinson pressure bar from (2 July2007 to 31 Dec 2007) at **Terminal Ballistic Research Laboratory Range, Ramgarh, District-Panchkula India** (*Ministry of defense*).

Presented research paper *"Eye Tracking based driver fatigue monitoring and warning system"* in ***IEEE explorer*** with ***ISBN No. 978-1-4244-7883-5*** *at Indian International conference on Power electronics held on 28-30 Jan 2011 at NSIT Dwarka Sector-3 Dwarka.*

Naveen Kansal S/O Sh. Manmohan Kansal was born in (Punjab) in India, on October 28, 1987. He graduated from RIMT Institute of Engg and Tech, Mandi Gobindgarh, India, and completed his Masters of Engg (batch 2008-10), E&EC Deptt, from P.E.C University of Technology (formally Punjab Engg College), Chandigarh Sector-12, Chandigarh-160012, INDIA.

He has successfully completed his Six Weeks Industrial training from 15 June 2005 to 31 July 2005 at **Punjab State Electricity Board, Ablowal, Patiala India** (*A Government of India Enterprise*). Also during his Six Month Industrial Project he worked on project named Head Up Display (HUD), (Ministry of Defence project) (Jan,2008 to June 2008) at **Central Scientific Instruments Organisation, Chandigarh.**

Spray Deposition Process of Hypereutectic Al – Si alloys: An overview

Dayanand .M. Goudar., Rudrakshi G.B., Srivastav V.C., Jagannath reddy, Ajith G. Joshi

Abstract—Modified Al alloys are gaining emphasis in the past few years due to wide range of properties available for the engineering applications. These alloys are manufactured by spray forming process have superior properties over the conventional cast alloys. Hence in this article an attempt has been made to summarize the work that has been carried out in the field of spray forming of hypereutectic Al – Si alloys and with other alloying elements such as Fe, Cu, Mg and Mn. As hypereutectic Al – Si alloys are attractive for the industrial applications due to their low thermal expansion co – efficient and high tribological properties. In this study attention has been paid towards the microstructure, mechanical and tribological properties of the studied alloys. The study of process parameters over the size and shape of the powder in spray forming is other part of this work. The change in properties of spray formed alloys by the working of secondary process is also been considered.

Index Terms— Hypereutectic Al – Si alloys, Spray deposition, Microstructure, Secondary process.

———————————————— ◆ ————————————————

1 INTRODUCTION

IN document is a template for Microsoft Word versions the present scenario Al alloys are gaining more importance in engineering applications like aerospace, automobile, etc., due to its wide range of physical, mechanical and tribological properties [1]. However conventional cast materials exhibits poor properties compared with spray formed materials [2]. Previous literatures have shown that the particle size, microstructure and mechanical properties are good when Al based alloys are manufactured by spray deposition process [3 – 6]. Gupta and Lavernia [7] studied the microstructure of hypereutectic Al – 17Si – 4.5%Cu alloy and their analysis revealed that spray atomization and spray deposition processing leads to significant refinement of microstructure as compared to its as cast counter part. They observed the microstructure characteristics as follows: (a) an equiaxed grain size (b) unconverted porosity (c) nearly spherodized secondary phases.

2 SPRAY DEPOSITION PROCESS AND PROCESS VARIABLES

Spray deposition process is one of the rapid solidification processes to produce new material by controlling the microstructure. The schematic representation of the spray deposition process is shown in Fig. 1. During the process, melt is atomized by transfer of high kinetic energy from gas to melt and melt droplets are collected on substrate, process variables like gas to melt ratio (G/M), melt superheat and atomizer design affects the size and size distribution of droplets during atomization [8 - 9]. Thus it finds necessary to understand the effect of process variables.

Jeyakumar et al. [10] studied the influence of process parameters and characteristics of Al alloys spray deposition process. They reported that powder size decreases with the increase in the applied gas pressure and also the yield, the result is also similar for decrease in the flight distance. A number of models were proposed by many literatures which relates the particles size and standard deviation which is specific for the particular nozzle designs [11 - 13].
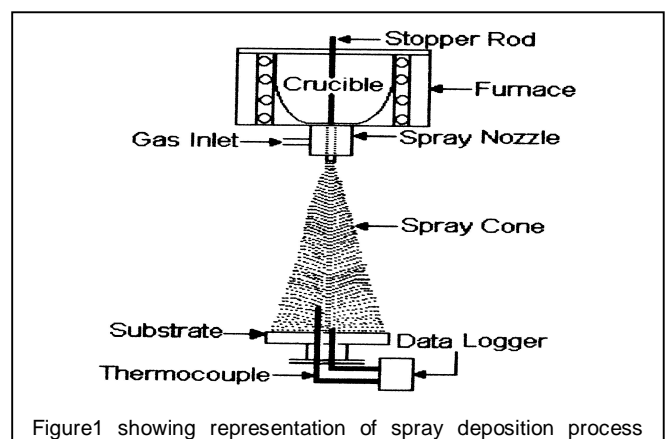
Figure1 showing representation of spray deposition process

————————————————————

- *Dayanand M Goudar is currently Asst. Prof. in Tondarya College of Engineering, Gadag, India, E-mail: dayanand_goudar@yahoo.co.in*
- *Rudrakshi G.B. is currently Prof. in Basaveshwara Engineering College, Bagalkot, India*
- *Srivastav V.C. is currently is scientist Metal Extraction & Forming Division, National Metallurgical Laboratory, Jamshedpur, India*
- *G .Jagannath Reddy is currently Asst. Prof. in vijaynagar engineering college, bellary,India, India,*
- *Ajith G Joshi is currently Lecturer in Canara Engineering College, Bejananapadavu, Bantwal, Mangalore, India.*
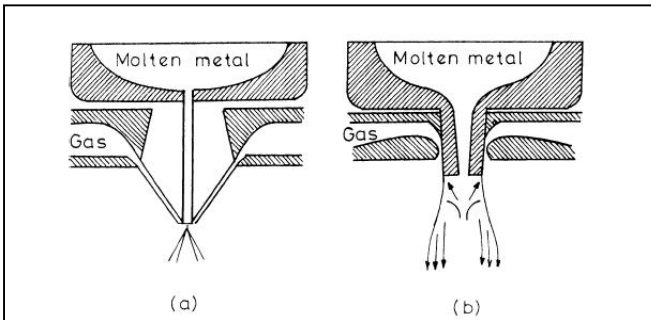
Fig. 2 Design of atomizers (a) free fall nozzle (b) closed couple nozzle [15]

Srivastav and Ojha [14] reported that G/M plays an important role in determining the particle size, standard deviation and flake formation. The productivity also increases with decrease in standard deviation for a given range of particle size. Singh et al. [15] found that the pressure supply by the given amount of gas for atomisation of freely falling molten metal stream based on an empirical formulae. They considered limiting velocity and plenum pressure and proposed models as equations (1) and (2) for the free fall atomizer designs (Fig. 2).

$$P_L = 9.24 \times 10^6 \rho^{0.3} F^{1.43} \alpha^{-0.97} \qquad \ldots\ldots\ldots\ldots(1)$$

$$V_L = 2.7 \times 10^3 \rho^{0.22} L^{0.62} \alpha^{-0.93} \qquad \ldots\ldots\ldots\ldots(2)$$

The benefits of spray deposition process are fine microstructure, enhanced mechanical, physical and tribological properties and also facilitate to produce composites with greater extent of uniform distribution. Srivastav et. al. [16] reported that there was considerable fine refinement of microstructure produced by spray forming compared to as cast alloys (Fig. 3 and Fig. 4). However the low efficiency, porosity in the material and material loss during the process are its limitations. The applications of spray deposited alloys and composites include automotive, aerospace, electrical and electronics applications. The bearing materials are also being manufactured by spray formation as it provides better tribological properties.

These alloys are attracted the researchers and industries as it provides excellent properties such as specific strength and modulus with low co-efficient of thermal expansion and good wear resistance. It is difficult to produce such alloys by conventional casting process [17 - 18] and maximum possible addition of silicon by casting is 22 - 24% [19]. However it is possible to produce hypereutectic Al - Si with greater than 25% Si by spray forming and microstructure evolution of hypereutectic Al-Si alloys by spray forming is shown in Fig. 5 [20].

In spray forming process of hypereutectic Al - Si alloys, when the diameters of nozzle increases as G/M ratio decreases and vice versa. The increase in G/M ratio leads to the fall in the size of the primary Si phase from 5 - 10 to 2 -

5μm, because the cooling rate is being enhanced during the flight time. Primary Si phase is recrystallised from the melt and develops a variety of regular polygon shape and spherical shape particles. The size of the primary Si phase also increases with the increase in preheating temperature [21]. Ha et al. [23] produced Al - 25%Si alloy by spray forming, the Si particles are uniform size and distribution. This study revealed that strain rate sensitivity factor is very low 0.1 below 300^0C and reaches maximum value to 0.2 at 500^0C from a series of load relaxation and compression test. The fine Si particle size production is possible in spray forming due to its higher cooling rate, which enhances the toughness of the alloy [24].
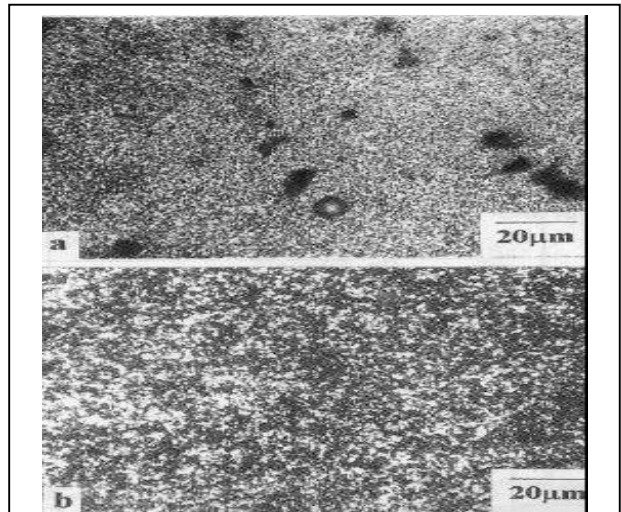
Fig. 3 Microstructure of spray-deposit near the substrate showing a. porosity and b. fine scale microstructure of second phase particles [16]
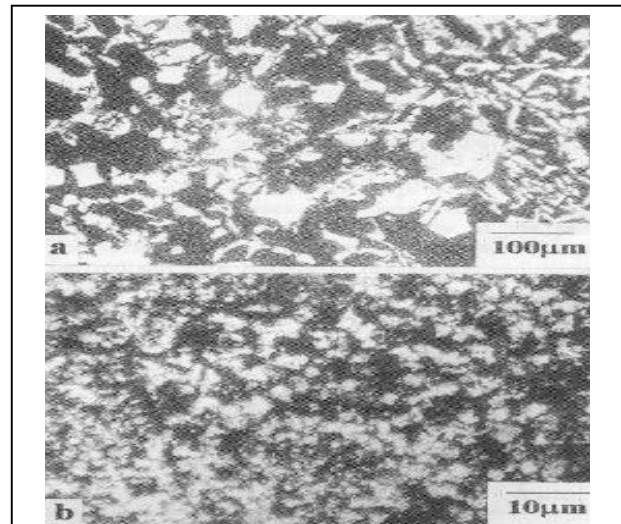
Fig. 4 Microstructure of the alloy showing a. coarse particles in the as cast alloy and b. fine second phase particles in spray-deposit [16]
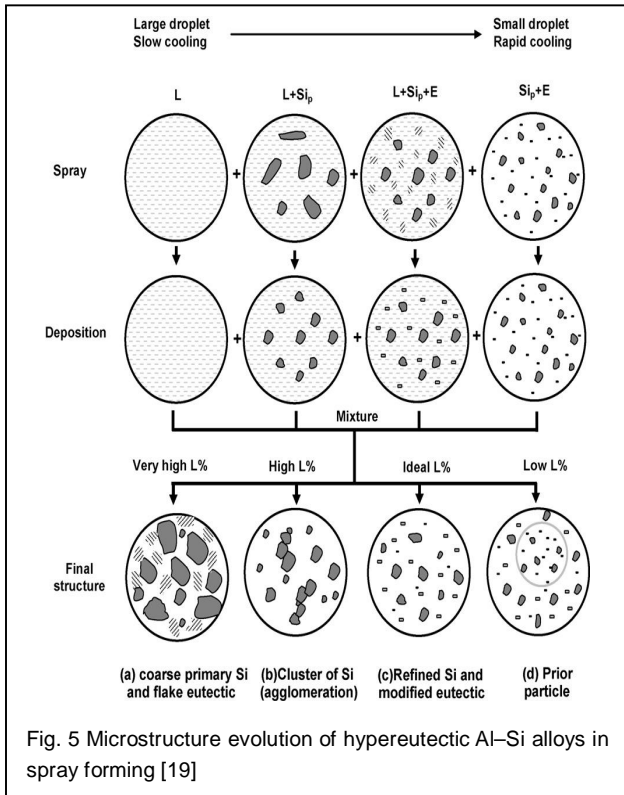
Fig. 5 Microstructure evolution of hypereutectic Al–Si alloys in spray forming [19]

Fig. 6 Variation of (a) wear rate and (b) coefficient of friction with applied pressure for spray formed Al-Si alloys (constant sliding velocity of 1.44m/s) [28]
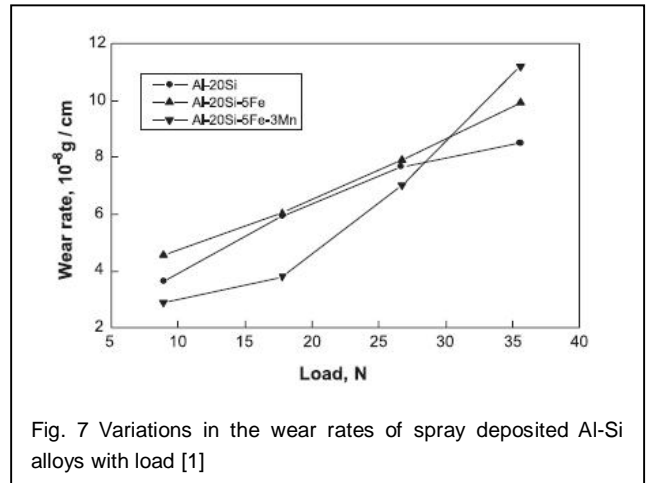
Fig. 7 Variations in the wear rates of spray deposited Al-Si alloys with load [1]

A comparative study shown that Al - 70%Si and Al - 60%Si exhibits the average density of 2421 kgm^{-3} and 2465 kgm^{-3}, average ultimate flexural strength of 180Mpa and 220Mpa and with the hardness 261 and 162 BHN respectively [25]. An another study has shown that Al - 25%Si exhibited 164MPa ultimate tensile strength (UTS) with 14.8% elongation, elastic modulus as 88.4 GPA and specific modulus as 3.4 x 10^9 mm [26]. Kim et al. [27] demonstrated that Al - 25%Si alloy almost behaves like superplastic Al - matrix at elevated temperature and has obtained 250% maximum tensile elongation, at 5 x 10^{-4} S^{-1} and at 500^0C. Extensive viscous formation was another observation of this study. They have concluded that liquidation of grain boundary material during grain boundary sliding may be responsible for the result.

The tribological characteristics of hypereutectic Al – Si alloys is better than the hypoeutectic alloys. Al – 18%Si shown very good wear behaviour than the Al – 6.5%Si. As the applied pressure increases, the wear rate increases. However, the hypereutectic alloy has shown less wear rate in all condition of applied pressure compared with hypoeutectic alloys (Fig. 6) [28]. It has been shown that as the applied load increases, the wear rate increases rapidly (Fig. 7) [1]. As the sliding distance increases, the wear volume increases and was very high at sliding distance greater than 1000m [29]. Elmadagli et al. [30] studied the microstructure and wear resistance of Al – Si alloys. They compared wear characteristics of as cast 383 alloy, A390 with spray formed Al-25Si alloys and reported that there was an increase in wear resistance and transition load considerable with increase in the Si content (Fig. 8-10). It can be also soldered without any difficulty [23]. Thus it provides a wide range of properties for a numerous engineering applications
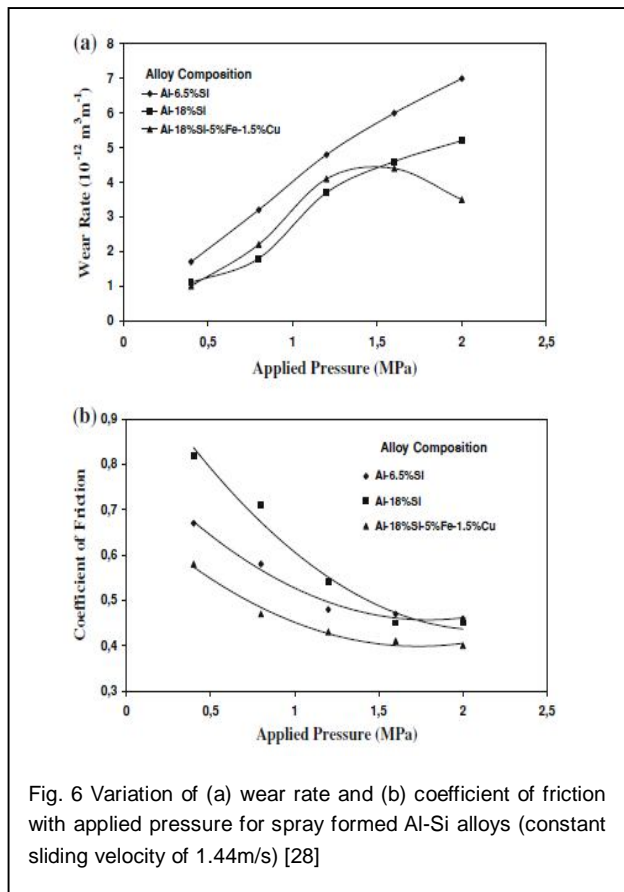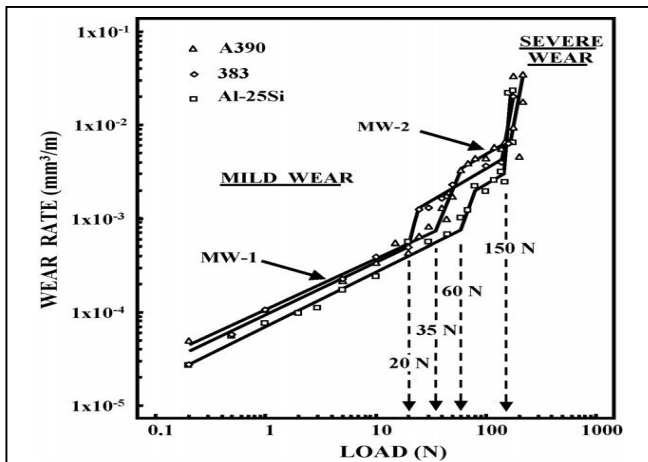
Fig. 8 Wear rates of the as-received A390, 383 and Al–25Si alloys plotted against load on a logarithmic scale. The two sub-regimes of mild wear MW-1 and MW-2 are separated by a transition regime. The loads 20, 35, and 60N correspond to the highest load inMW-1. The transition fromMW-2 to the severe wear regime occurs at 150N for all the alloys [30]
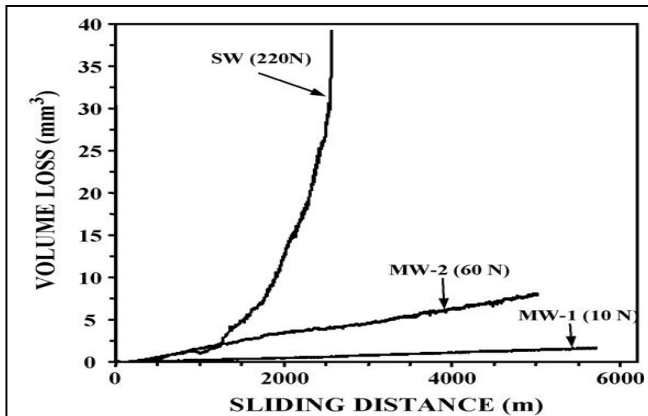
Fig. 9 Volume loss vs. sliding distance plots of the A390 in mild wear regime (MW-1:10 N,MW-2: 60 N) and severe wear regime (SW: 220 N) computed from height changes of the samples [30]
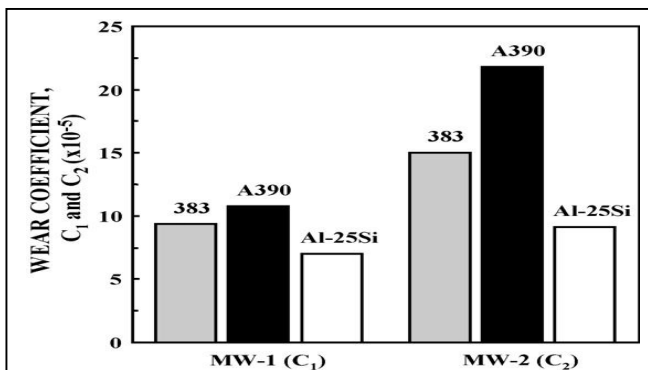
Fig. 10 Comparison of the wear coefficients of the as-received alloys. The wear coefficients (C1) of MW-1 and (C2) of MW-2 reflect the differences in the wear rates of the alloys [30]

3 HYPEREUTECTIC AL-SI TERNARY ALLOYS

3.1 Effect of Fe addition on the properties of Al – Si alloys

Eventhough Al – Si alloys is providing a wide range of properties, a new emphasis was aroused in the fast few years that to give an additional requirement at high temperature stability and microstructure. Accordingly Fe is one of the additive to the Al – Si alloys which results in stable intermetallic phase but the limitation is, in conventional casting due to slow cooling rate segregation of microstructure and hence porosity forms [31 – 33]. A study has revealed that Al – 5%Fe alloys exhibits drastic refinement and modification in the microstructure with a uniform distribution of secondary phase particles in the matrix [29]. In another phenomenon, during solidification at high undercooling it may form several dispersoids of Al_xFe [34].

Thus phase and microstructure of alloys are highly influenced by the duration of undercooling and concentration of Fe [34 – 36]. Islas et al. [37]studied the solidification of microstructure of Al – Fe binary alloys and reported that the content of Fe in the α – Al phase alloys increases with decrease in powder size and undercooling increases with increasing cooling rate and decreasing with the size of atomized powders. The addition of Fe to the Al – Si alloys reveals better development in the properties of Al – Si alloys. But it is difficult to produce by conventional casting, because of the formation of large size Si phase and vary plates of Al – Fe. However it is possible to cast Al – Si – Fe alloys adequately by spray deposition [38 – 45]. Srivastav et al. [3] produced Al – Fe alloys by spray deposition with uniform distribution Al_3Fe phase and Al – Fe alloy exhibited ductile mode of fracture compared to mixed mode of fracture of Al – 18Si alloys under tensile loading and was found to be 19% elongation of Al – 5Fe alloy.

3.2 Effect of non ferrous alloying elements addition on the properties of Al – Si alloys

Blockford et al. [46] produced Iron Aluminide preform by spray forming and found that spray deposited layer exhibited some oxide and porosity. However porosity decreased with heat treatment. A layer of Fe_2O_3 was noticed in this study instead of Al_2O_3 layer, which is usually found in Al binary alloys. Wang et al. [1] and Srivastav et al. [4] reported that addition of 5% Fe to Al – 20Si results in the formation of needle shaped intermetallic phase (Fig. 11-12). Wang et al. [11] also studied the wear characteristics as follows. At lower loads Fe containing alloy showed higher wear rate compared with Al – 20Si alloy due to tendency for embrittlement because of needle shaped intermetallic and even at higher loads Fe containing alloy exhibited inferior wear resistance than the binary alloy due to fragmentation at intermetallics, result-

ing in crack nucleation and propagation. Wear process is controlled by subsurfaced cracking assisted by sliding mechanism.
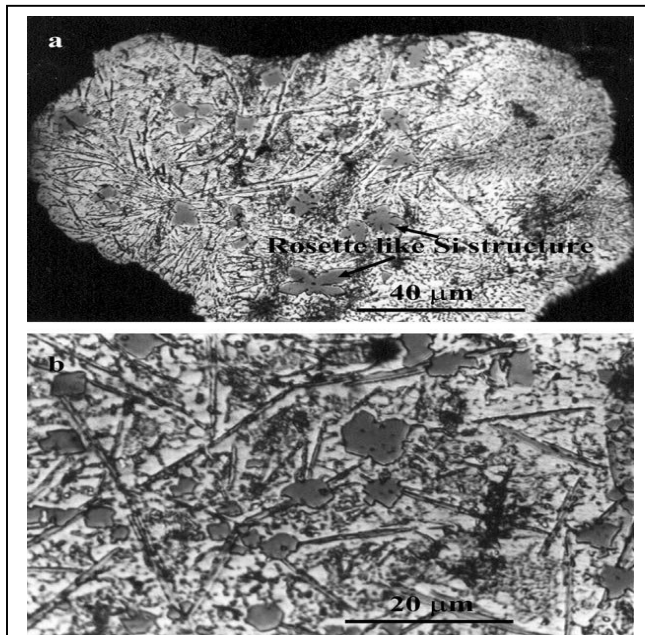
Fig. 11 Optical micrographs of a section of oversprayed powder particles showing: (a) facetted and needle-shaped morphologies of primary Si and intermetallics and (b) rosette-shaped morphology of Si. [4]
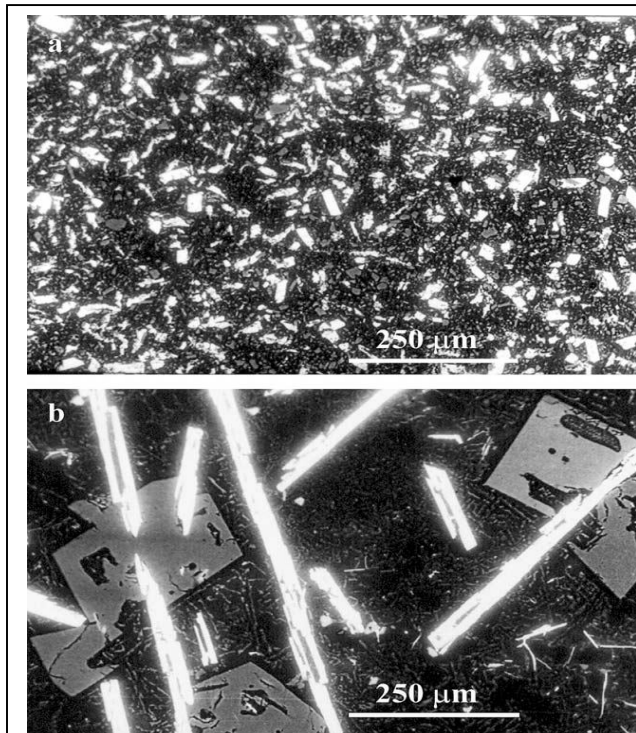
Fig. 12 SEM backscattered images showing: (a) microstructure of spray deposited alloy showing primary Si phase along with refined intermetallics and (b) microstructure of conventional as-cast alloy. [4]

Leo et al. [47] reported that there was an internal friction in the temperature range of 50 – 250⁰. In spray deposited Al – 3.3Fe – 10%Si alloy. Perhaps due to influence of $FeAl_2$ boundary relaxation and also affected by Al – Fe – Si intermetallic particle at the grain boundary. Mocellian et al. [48] reported that there was damage in the tensile loading of T_4 temperature alloy due to cracking of Si primary particles and is associated with the loss of elastic modulus. As the applied stress increased, the silicon particles are broken because a critical tensile stress level is reached within the critical applied stress value, the stress value at which silicon particles are broken depends on the parameters such as internally induced stresses and also stress concentration effect due to plastic deformation in the Al – alloy matrix around the silicon particles, the stress concentration factor is generally reduced by decrease in grain size and internally induced stresses, corresponding to a compressive stress tied in one silicon and particles delays the beginning of damage.

Addition of Fe provides high diffusivity in liquid state and low diffusivity in solid state [49 – 51], which ensures chemical homogeneity in one α – Al matrix and its thermal stability, in both processing service. Iron has an extremely low equilibrium solid solubility in α – Al at room temperature under the condition of rapid solidification [51]. Enhancement of the mechanical properties of the Al base alloys can also be done by the addition of Cu, Mg. The addition of Cu and Mg to the Al – Si alloys may form the precipitation producing phases such as Al_2Cu, Mg_2Si, Al_2CuMg [51 – 55]. Both Cu and Mg act as precipitation hardening elements and their strengthening effect is dimished at above 150⁰C because of low thermal stability of the precipitates [49].

Kai et al. [56] produced Al2024 alloy by spray deposition and reported that there was an increase in the porosity of the alloy. Srivastav et al. [25] reported that the microstructure of Al – 305Cu – 10Si alloys exhibited equiaxed grain morphology of one primary α – phase with particulate morphology of Si particles located at the grain boundaries and finer Si particles in the grain interior. Addition of 1.5%Cu to Al - 18Si – 5Fe reveals the microstructure with uniform distribution of primary Si phase and intermetallic compounds such as β – Al – Fe – Si and δ – Al_4Si_2Fe phases [25, 57]. Santos et al. [58] studied the corrosion performance of Al – Si – Cu hypereutectic alloys in synthetic condensed automotive solution and their results revealed that the preferential attack of Al matrix phase in all their studied alloys. The alloy with higher Cu content and prepared by spray forming was more susceptible to pitting compared to other alloy.

Jodoin et al. [59] reported that conventional nanocrystalline Al2618 alloy coatings sprayed on to Al substrate using cold gas dynamic spraying process exhibited negligible porosity and excellent interface with the substrate material in the range of 5 to 10% and morphology of ob-

tained powder particles is shown in Fig. 13. Ikram [60] studied that microstructure of Al – 6Si – 1.9Cu – Mg alloys and found that harndness was increased by increasing Mg content except for 1.13% Mg alloy aged for 10hrs and Q phase particles are perhaps responsible for the trend of hardness. The microstructure and mechanical properties of spray formed and squeeze cast Al – 25Si – 0.8Cu – 0.84Mg alloys was studied and compared by Chiang and Tsao [2] observed that spray formed alloy has higher degree of super saturation than that of squeeze cast alloy and peak aged spray formed alloys exhibited ultimate strength of 366MPa yield strength of 235.3MPa and elongation of 3.4%. The study of Wang et al. [1] revealed that addition of Mn to Al – 20Si – 5Fe alloy reduced the wear rate at lower loads of about 25N and has got increased than that at higher loads of about 35N.

Estrada and Duszczyk [61] studied the characteristics of rapidly solidified spray deposited Al – Si – X preforms and concluded that Al alloy can be spray deposited without oxide films and with fine precipitates. The deposited matrix exhibited an average porosity level as low as 1.3% and the porosity was uniform throughout preform and microstructure does not depends on the small local porosity variation. Thus addition of non ferrous alloying elements with Fe to the hypereutectic Al – Si alloys enhances the microstructure, mechanical and tribological properties.
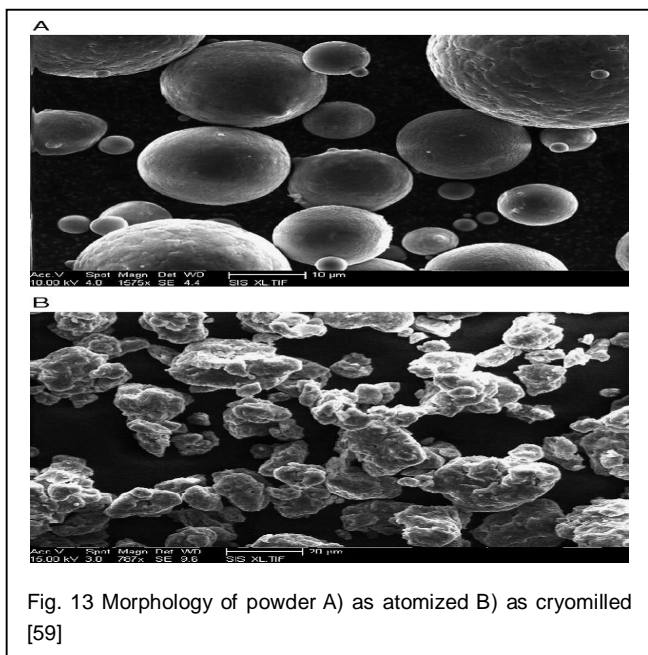
Fig. 13 Morphology of powder A) as atomized B) as cryomilled [59]

4 EFFECT OF SECONDARY PROCESSING ON THE PROPERTIES OF AL – SI ALLOYS

Eventhough spray atomization provides better microstructure and enhanced physical properties than as cast alloys. There is a problem with the formation of intermetallic phases with the needle shaped and having relatively large size in α matrix. This may also induce detrimental effect on the mechanical and tribological properties on the final product [49]. It also affects on the effectiveness at stabilizing the microstructure, hence it is desired always to do the post atomization process. It leads to the decrease in their size and interspacing often heat treatment can be done but it cannot complete the necessary requirement. Since intermetallic cannot be resolved in the solid matrix. The principle advantages of secondary process are refinement of the microstructure and reduce the porosity of the deposit.
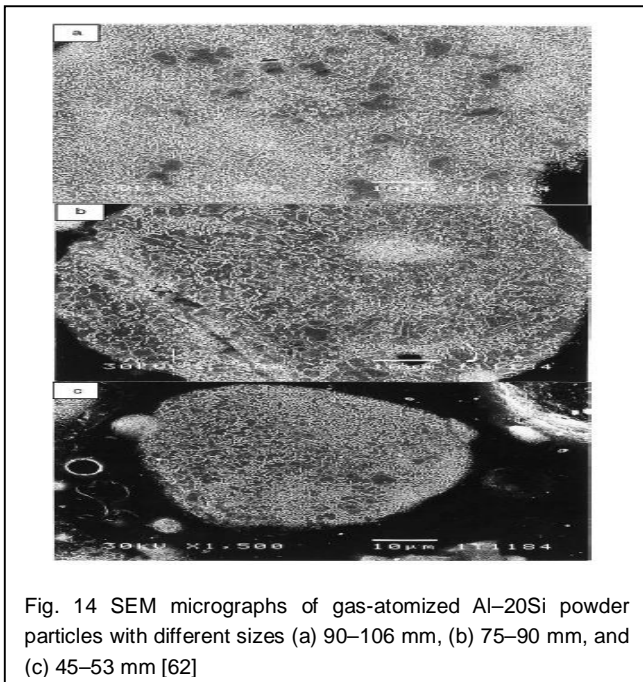
Fig. 14 SEM micrographs of gas-atomized Al–20Si powder particles with different sizes (a) 90–106 mm, (b) 75–90 mm, and (c) 45–53 mm [62]

Extrusion itself is a complicated thermomechanical process [7]. Hong et al. [62] observed that due to severe shear deformation takes place during hot extrusion. The large eutectic structure and primary Si phase and were fragmented into smaller particle and distributed homogenously in the Al – matrix (Fig. 14). Seok et al. [25] found that spray formed and extruded Al – 25Si has properties as follows UTS – 164MPa, elongation – 14.3%, elastic modulus – 88.4GPa and specific modulus 3.4 x 10^9mm. Cui et. al. [63] studied the characterization of Si phases in spray formed and extruded hypereutectic Al – Si alloys and concluded that extrusion plays a role on the interparticles spacing according to the particle grooves or fragmentation mechanism. The Si particles in extruded Al – Si alloys appear more homogenous and regular than as deposited but with a certain amount of anisotrophy and tendency to preferred orientation (Fig. 15 and 16). Ha et. al. [22] has reported that there was refinement of microstructure after extrusion (Fig. 17). Baiqing et al. [64] concluded that high extrusion ratio refines the primary Si phase in microstructure simultaneously, it refills microcracks (Fig. 18). Thus in almost previous literatures secondary processing has been done to refine the grain structure and enhance their characteristics. The Table 1 gives

the brief comparison of properties of different alloys considered in previous studies.
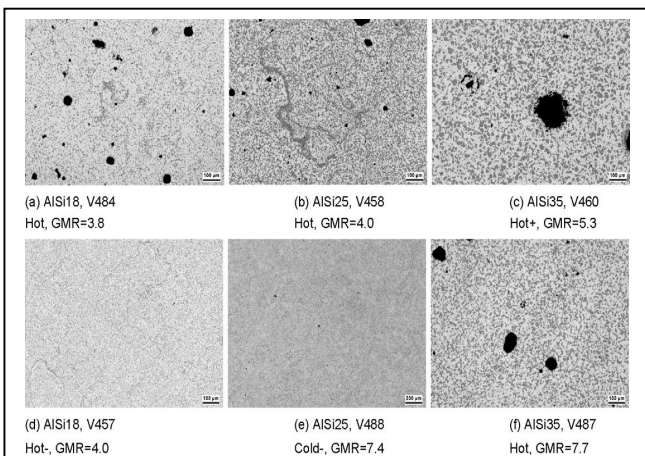
Fig. 15 Porosity in spray formed hypereutectic Al–Si alloy billets. (a) AlSi18, V484 Hot, GMR= 3.8; (b) AlSi25, V458 Hot, GMR= 4.0; (c) AlSi35, V460 Hot+, GMR= 5.3; (d) AlSi18, V457 Hot−, GMR= 4.0; (e) AlSi25, V488 Cold−, GMR= 7.4 and (f) AlSi35, V487 Hot, GMR= 7.7 [63]
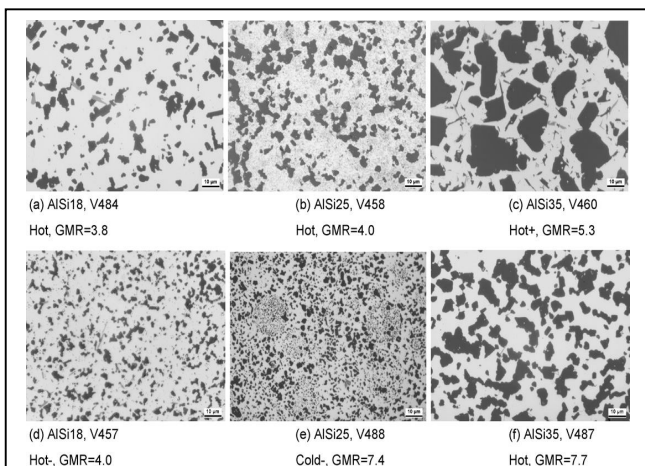
Fig. 16 Silicon precipitations in spray formed hypereutectic Al–Si alloy billets. (a) AlSi18, V484 Hot,GMR= 3.8; (b) AlSi25, V458 Hot,GMR= 4.0; (c) AlSi35, V460 Hot+,GMR= 5.3;(d) AlSi18, V457 Hot−, GMR= 4.0; (e) AlSi25, V488 Cold−, GMR= 7.4 and (f) AlSi35, V487 Hot, GMR= 7.7 [63]
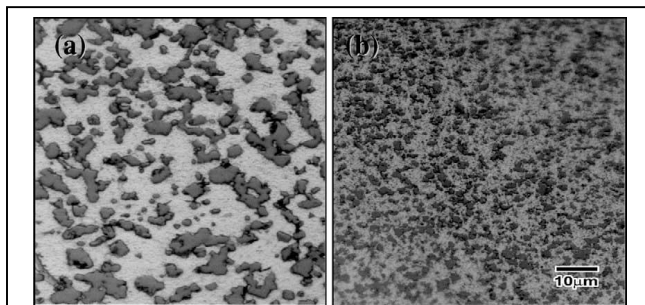
Fig. 17. Typical microstructures taken (a) before and (b) after the extrusion [22]
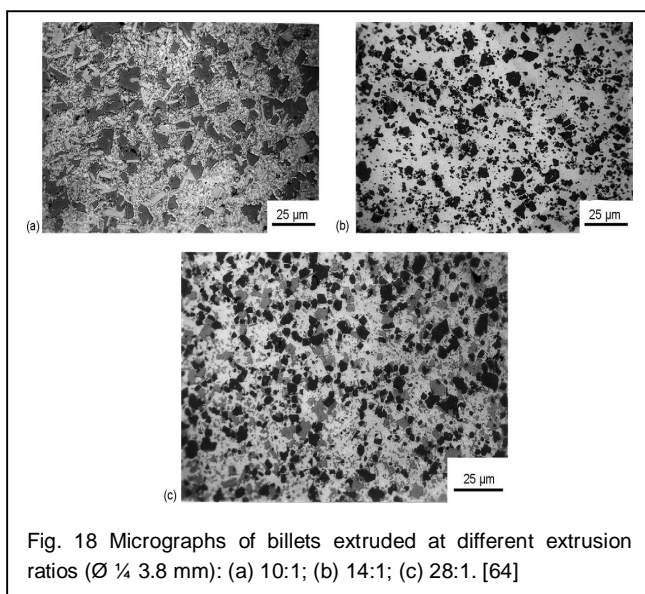
Fig. 18 Micrographs of billets extruded at different extrusion ratios (Ø ¼ 3.8 mm): (a) 10:1; (b) 14:1; (c) 28:1. [64]

5 CONCLUSIONS

In the present study, the previous works on the microstructure, mechanical and tribological properties of spray deposited hypereutectic Al – Si alloys are summarized. On this basis following conclusions were drawn.

- Good amount of work has been carried out related to the study of particle size and morphology of powder produced during spray deposition process. The works has paid concentration on the process variables and nozzle design. There is necessity of relation which provides the relation between the process variables and particle size and morphology for the general nozzle designs. Inspite optimization of process parameters is also vital.
- Hypereutectic Al – Si alloys produced by spray deposited is well studied in consideration of their properties. Yet study is going on corresponding to grain refinement and development of properties with high Si content.
- The development of fine grain and with high wear resistance would be lucrative to aerospace and automotive applications.
- In recent years researchers are concentrating towards the addition of ferrous and non ferrous alloying elements to hypereutectic Al – Si alloys which intensifies alloy properties.
- Secondary processes are also being carried out in supplement with spray deposition process. Since the probability of the presence of porosity in the microstructure of spray deposited alloys, resulting in deteriorated effect on its characteristics. Thus to augments the mechanical properties by decreasing the porosity, generally hot extrusion is processed in

TABLE 1

COMPARISON OF THE PROPERTIES OF CAST AND SPRAY DEPOSITED ALLOYS STUDIED IN PREVIOUS LITERATURES

Alloy	Processing route	Yield Strength (MPa)	UTS (MPa)	Elongation (%)	Hardness	Elastic Modulus (GPa)	Reference
Al – 18Si	AC		50	<3	50HV		[3]
Al – 18Si	ACE	97	129	7	41HV		[3]
Al – 18Si	SDE	120	158	19	59HV		[3]
Al – 5Fe	AC		48		17HV		[3]
Al – 5Fe	SDE	176	201	23	59HV		[3]
Al – 18Si – 5Fe – 1.5Cu	AC		55	<3	53HV		[3]
Al – 18Si – 5Fe – 1.5Cu	SDE	221	262	8	89HV		[3]
Al – 25Si – 0.89Cu – 1Ni – 0.84Mg	SDE	108.9 (Proof stress)	262.9	7.2			[2]
Al – 25Si	SDE		164	14.8		88.4	[22]
Al – 25Si – 3.67Cu – 1.11Mg – 0.4Fe	SDE		420	0.8		88.6	[22]
Al – 17.48Si – 3.67Cu – 0.42Fe – 1.11Mg – 0.02Ni	SDE		305.8	3.5			[20]
Al – 17Si – 3Fe	SDE	400	470	13			[6]
Al – 21Si – 3Fe	SDE	410	473	0.8			[6]
Al – 25Si – 3Fe	SDE	440	503	0.6			[6]
DISPAL S220*	ACE	95	165	2.5	65HV	85	[65]
DISPAL S250*	ACE	205	334	2.7	105HV	95	[65]
DISPAL S260*	ACE	180	265	1.0	110HV	85	[65]
DISPAL 220*	SD	135	230	2.5		86	[65]
DISPAL 240*	SD	425	490	1.5		99	[65]
DISPAL 250*	SD	438	490	1.2		97	[65]
DISPAL 260*	SD	240	360	2.1		98	[65]
DISPAL 260*	SD	180	245	1.2			[65]

*DISPAL is a trade name of PEAK Werkstoff GmbH, Velbert, Germany

the previous literatures

REFERENCES

[1] Feng Wang, Zhengye Zhang, Yajun Ma, Yuansheng Jin, "Effect of Fe and Mn additions on microstructure and wear properties of spray deposited Al – 20Si alloy" Materials Letters, 58 (2004) 2442 -2446.

[2] C.H. Chiang, Chi Y.A. Tsao, "Microstructures and mechanical properties of spray formed and squeeze cast Al – 25Si – 0.89Cu – 1.0Ni – 0.84Mg alloys in solutionized and aged conditions", Materials Science and Engineering A 417 (2006) 90 – 98.

[3] V.C. Srivastav, R.K. Mandal, S.N. Ojha, K. Venkateshwaralu, "Microstructural modifications induced during spray deposition of Al – Si – Fe alloys and their mechanical properties", Materials Science and Engineering A 471 (2007) 33 – 49.

[4] A.K. Srivastav, V.C. Srivastav, A. Gloter, S.J. Ojha, "Microstructural features induced by spray processing and hot extrusion of an Al – 18% Si – 5% Fe - 1.5% Cu alloy", Acta Materialia 54 (2006) 1741 – 1748.

[5] L.A. Bereta, C.F. Ferrarini, C.S. Kiminami, W.J.F. Botta, C. Bolfarini, "Microstructure and mechanical properties of spray deposited and extruded/heat treated hypoeutectic Al – Si alloy", Materials Science and Engineering A 449 - 451 (2007) 850 – 853.

[6] M. Courbiere and A. Mocellin, "Spray cast Al – Si base alloys for stiffness and fatigue strength requirements", Journal De Physique IV, 3 (1993) 207 – 213.

[7] M. Gupta, E.J. Lavernia, "Effect of processing on the microstructural variation and heat treatment response of a hypereutectic Al – Si alloy", Journal of Material Processing Technology, 54 (1995) 261 – 270.

[8] P.S. Grant, "Spray forming", Progress in Materials Science, 39 (1995) 497-545.

[9] E.J. Lavernia and Yue Wu, "Spray Atomization and Deposition", 1996, John Wiley and Sons Ltd., West Sussex, England.

[10] M. Jeyakumar, S. Kumar, G.S. Gupta, "The influence of processing parameters on characteristics of an aluminium alloy spray deposition", Materials and Manufacturing processes 24 (2009) 693 – 699.

[11] H. Lubanska, "Correlation of spray ring data for gas atomization of liquid metals", Journal of metals, 22 (2) 45 – 49

[12] K.Y. Kim and W.R. Marshall, "Drop-size distributions from pneumatic atomizers", AIChE Journal, 17 (1971) 575 – 584.

[13] S. Nukiyama and Y. Tanasawa, "Experiments on the atomization of liquids in an air stream", Transactions of society of mechanical engineers of Japan, 6 (23) (1940), 18 - 28

[14] V.C. Srivastav and S.N. Ojha, "Effect of aspiration and gas – melt configuration in close coupled nozzle on powder productivity," Powder metallurgy 49 (2006) 213 – 218.

[15] D. Singh, S.C. Koria, R.K. Dube, "Development of operating pressure diagram for free fall gas atomization of liquid metals" ISIJ International, 43 (12) (2003) 2067 – 2069.

[16] V.C. Srivastav, Anish Upadhaya and S.N. Ojha, "Microsturctural features induced by spray forming of a ternary Pb-Sn-Sb alloy", Bulletin of Material Science, 23 (2000) 73 – 78.

[17] K. Raju, S.N. Ojha, A.P. Harsha, "Spray forming of an aluminium alloys and its composites: An overview", Journal of Material Science, 43 (2008) 2509-2521.

[18] Hatch J.E., "Aluminium propeties and physical metallurgy", American Society for Metals, Metals Park, OH. pp. 214

[19] Key to metals: articles – Al – Si alloys (www.keytometals.com/Articles.htm)

[20] H. Hu, Z.H. Lee, D.R. White, E.J. Lavernia, "On the evolution of porosity in spray deposited tool steels", Metallurgy Materials Transactions, 31 (2000) 725 – 735.

[21] C. Cui, A. Schulz, K. Schimanski, H. W. Zoch, "Spray forming of hypereutectic Al – Si alloys", Journal of Materials Processing Technology, 209 (2009) 5220 – 5228.

[22] Xiong Baiqing, Zhang Yongan, Wei Qiang, Shi Likai, Xie Changan, Shang Chengjia, He Xinlai, "The study of primary Si phase in spray forming hypereutectic Al – Si alloy", Journal of Materials Processing Technology 137 (2003) 183 – 186.

[23] Tae Kwon Ha, Woo-Jin Park, Sangho Ahn, Young Won Chang, "Fabrication of spray-formed hypereutectic Al–25Si alloy and its deformation behavior", Journal of Materials Processing Technology 130–131 (2002) 691–695.

[24] Min Ryou, Chul Hyun Kim, Myung Ho Kim, "Microstructure and mechanical properties of hypereutectic Al – Si alloys fabricated by spray casting", Journal of Materials Science Technology, 24 (2008) 48 – 50.

[25] Wei Yan guang, Xiong Bai qing, Zhang Yong an, Liu Hong wei, Wang Feng, Zhu Bao hong, "Property measurements on spray formed Si – Al alloys", Transactions nonferrous materials society china 17 (2007) 368 – 372.

[26] Hyun Kwang Seok, Jae Chul Lee, Ho In Lee, "Extrusion of spray formed Al – 25Si – X composites and their evaluation", Journal of Materials Processing Technology 160 (2005) 354 – 360.

[27] Woo Jin Kim, J.H. Yeon, J.C. Lee, "Superplastic deformation behaviour of spray deposited hyper – eutectic Al – 25Si alloy", Journal of alloys and compounds 308 (2000) 237 – 243.

[28] V.C. Srivastav, G.B. Rudrakshi, V. Uhlenwinkel, S.N. Ojha, "Wear Characteristics of spray formed aluminium alloys and their composites", Journal of Material Science, 44 (2009) 2288 – 2299.

[29] Ph. D thesis of V.C. Srivastav, "Spray forming and characterization of Al – Si – Fe alloys", Banaras Hindu University, Varanasi, 2003.

[30] M. Elmadagli a, T. Perry b, A.T. Alpas, "A parametric study of the relationship between microstructure and wear resistance of Al–Si alloys", Wear 262 (2007) 79–92

[31] S. Nafisi, D. Emadi, M.T. Shehata, R. Ghomashchi, "Effects of electromagnetic stirring and superheat on the microstructural characteristics of Al – Si – Fe alloy", Material Science Engineering A 432 (1-2) 2006 71 – 83.

[32] P.N. Crepeau, "Effect of Iron in Al-Si Casting Alloys: A Critical Review" AFS Transcations 103 (1995) 361 – 366.

[33] C.M. Dinnis, J.A. Taylor, A.K. Dahle, " Iron – related porosity in Al – Si – Cu foundry alloys", Material Science Engineering A 425 (1-2) (2006) 286 – 296

[34] Wolfgang G.J. Bunk "Aluminium RS metallurgy" Material Science Engineering A134 (1991) 1087 - 1097

[35] W. J. Boettinger, L. Bendersky, J.G. Early, "An analysis of the microstructure of rapidly solidified Al-8 wt pct Fe powder", Metallurgical and Materials Transactions, 17A (1986) 781 – 790.

[36] R.F. Cochrane, P.V. Evans, A.L. Greer, "Competitive growth analysis of phase formation in Al-8wt.%Fe", Materials Science Engineering A 133 (1991) 803 – 806.

[37] J. Juarez – Islas, Y. Zhou, E.J. Lavernia, "Spray atomization of two Al – Fe binary alloys: Solidification and microsturcture characterization", Journal of Material Science 34 (1999) 1211 – 1218.

[38] R.F. Cochrane, S.B. Newcomb, P.V. Evans, A.L. Greer, "Microstructural Development in Drop-Tube Processed Al-8wt.%Fe", Key Engineering Materials 38 & 39 (1989) 21.

[39] J.D. Cotton, M.J. Kaufman, "Microstructural evolution in rapidly solidified Al-Fe alloys: An alternative explanation", Metallurgical and Materials Transactions A, 22A (1991) 927 - 934.

[40] B. Yang, F. Wang, J.S. Zhang, B.Q. Xiong, X.J. Duan, "The effect of Mn on the microstructure of spray-deposited Al–20Si–5Fe–3Cu–1Mg alloy" Scripta Materialia, 45 (5) (2001) 509 – 515.

[41] V.C. Srivastav, R.K. Mandal, S.N. Ojha, "Microstructure and mechanical properties of Al–Si alloys produced by spray forming process", Material Science Engineering A 304 – 306 (2001) 555 – 558.

[42] V.C. Srivastav, R.K. Mandal, C. Ramachandra, B. Chatterjee, S.N. Ojha, "Microstructural and Wear Characteristics of spray deposited hypereutectic Al – Si alloy", Transactions IIM, 51 (1) (1999) 29 – 40.

[43] Y. Wu, W.A. Cassada, E.J. Lavernia, "Microstructure and Mechanical Properties of Spray Deposited Al-17Si-4.5Cu-0.6Mg Wrought Alloy", Metallurgical Transactions A 26 (1995) 1235 – 1247.

[44] C.H. Chiang, C.Y.A. Tsao, "Si Coarsening of spray formed high loading hypereutectic Al – Si alloys in the semisolid state", Materials Science Engineering A 396 (1-2) (2005) 263 – 270.

[45] C.F. Ferrarini, C. Bolfarini, C.S. Kiminami, W.J. Botta, "Microstructure and mechanical properties of spray deposited hypoeutectic Al – Si alloy", Materials Science Engineering A 375 – 377 (2004) 577 – 580.

[46] J.R. Blackford, R.A. Buckley, H. Jones, C.M. Sellars, D.G. McCartney, A.J. Horlock, "Spray deposition of an iron aluminide", Journal of Materials Science, 33 (1998) 4417- 4421.

[47] B.H. Luo, F.J. Xu, Y.Q. Xie, "Microstructure and internal friction of spray deposited Al – 3.3Fe – 10.7Si alloy", Journal of materials science 35 (2000) 109 – 113.

[48] A. Mocellin, R. Fougeres, P.F. Gobin, "A study of damage under tensile loading in a new Al – Si – Fe alloy processed by the osprey route", Journal of materials science 28 (1993) 4855 – 4861.

[49] J. Zhou, J. Duszczyk, B.M. Korevaar, "Structural development during the extrusion of rapidly solidified Al – 20Si – 5Fe – 3Cu – 1Mg alloy", Journal of Material Science, 26 (1991) 824 – 834.

[50] G. Thursfield, M.J. Stowell, "Mechanical properties of Al-8 wt% Fe-based alloys prepared by rapid quenching from the liquid state", Journal of Material Science 9 (1974) 1644

[51] H. Jones, "Observations on a structural transition in aluminium alloys hardened by rapid solidification", Materials Science and Engineering 5 (1969) 1 - 18

[52] F. Wang, B. Wang, X.J. Duan, B.Q. Xiong, J.S. Zhang, " The microstructure and mechanical properties of spray-deposited hypereutectic Al–Si–Fe alloy" Journal of Material Processing Technology, 137 (2003) 191 – 194.

[53] Z. Lee, A.M. Samuel, F.H. Samuel, C. Ravindra, S. Valtierra, "Effect of

alloying elements on the segregation and dissolution of CuAl2 phase in Al-Si-Cu 319 alloys" Journal of Materials Science 38 (2003) 1203 – 1219.

[54]   R.M. Gomes, T. Sato, H. Tezuka, A. Kamio, "Precipitation behavior of P/M hypereutectic Al-Si-Cu-Mg alloys containing Fe and Ni" Materials Transactions, JIM, 39 (1998) 357 – 364.

[55]   A.K. Gupta, M.C. Chaturvedi, A.K. Jena, "Effects of silicon additions on aging behaviour of Al-1.52 Cu-0.75 Mg alloy" Materials Science and Technology 5 (1989) 50-55.

[56]   B.C. Moon, Z.H. Lee, "Damping behavior of Al---Zn alloys produced by spray forming process" Scripta Materialia 38 (1998) 207 - 213.

[57]   V.C. Srivastav, P. Ghoshal, S.N. Ojha, "Microstructure and phase formation in spray deposited Al – 18%Si – 5%Fe – 1.5%Cu alloy", Materials Letters 56 (2002) 797 – 801.

[58]   Hamilta de Oliveira Santos, Fernando Morais dos Reis, Clarice Terui Kunioshi, Jesualdo Luiz Rossi, Isolda Costa, " Corrosion performance of Al – Si – Cu hypereutectic alloy in a synthetic condensed automotive solution", Materials Research 8 (2005) 155 – 159.

[59]   B. Jodoin, L. Ajdenlsztajn, E. Sansoucy, A. Zuniga, P. Richer, E.J. Lavernia, "Effect of particle size, morphology and hardness on cold gas dynamic sprayed aluminium alloy coatings",   Surface and Coatings Technology, 201 (2006) 3422 – 3429.

[60]   S. Zafar, Nazma Ikram, M.A. Shaikh, K.A. Shoaib, "Microstructure studies in Al – 6%Si – 1.9%Cu – X%Mg alloys", Journal of Material Science, 25 (1990) 2595 – 2597.

[61]   J.L. Estrada, J. Duszczyk, "Characteristics of rapidly solidified Al – Si – X preforms produced by the osprey process", Journal of Materials Science 25 (1990) 1381 – 1391.

[62]   Soon-Jik Hong, C. Suryanarayana, Byong-Sun Chun, "Section – dependent microstructure and mechanical properties of rapidly solidified and extruded Al – 20Si alloy", Materials Research Bulletin, 39 (2004) 465 – 474.

[63]   Chengsong Cui, Alwin Schulz, Ellen Matthaei Schulz, Hans – Werner Zoch, "Characterization of silicon phases in spray formed and extruded hypereutectic Al – Si alloys by image analysis", Journal of Materials Science, 44 (2009) 4814 – 4826.

[64]   Xiong Baiqing, Zhang Yongan, Wei Qiang, Shi Likai, Xie Changan, Shang Chengjia, He Xinlai, "The study of primary Si phase in spray forming hypereutectic Al–Si alloy", Journal of Materials Processing Technology 137 (2003) 183–186.

[65]   Data sheet PEAK Werkstoff GmbH, Velbert, Germany, without year.

A Fuzzy Inference System for Synergy Estimation of Simultaneous Emotion Dynamics in Agents

Atifa Athar, M. Saleem Khan, Khalil Ahmed, Aiesha Ahmed and Nida Anwar

Abstract— This paper presents that emotions manifest the information processing mechanism of human mind that infers the synergic effect of simultaneous emotions to achieve focused communication and decision making. This proposed work considers integration mechanism of complex emotional dynamics for agents to communicate reason and decide in conflicting situations like humans. The proposed inference system is used to estimate the blended effect of simultaneously activated emotions in agents using fuzzy logic as it is an unsurpassed choice to deal with uncertain information and classification of non-deterministic events.

Index Terms— Fuzzy inference, Simultaneous emotion dynamics, Synergy estimation, blended emotions, PAD, Wheel of emotions, social cohorts.

————————————————   ◆   ————————————————

1 INTRODUCTION

Emotional intelligence corresponds to the awareness and usage of emotions to make smart decisions in different situations. The segregation between an intelligent and non-intelligent mind is based on awareness that is an outcome of inference and regarding psychological viewpoint emotions are well thought out as inferential shortcuts [1].

An emotion is not an isolated feature of human mind; it is to be considered as a variable in nature. It grows or decays temporally due to the change in the environmental information absorbed by mind. Since this information is not always discrete and certain thus it may cause simultaneous emotion activation in mind. The dynamic mechanism of human mind can have a variety of emotional states simultaneously and can convey their shared or diverse effect [2]. To infer these effects it is required to blend these simultaneously active emotions. There could be one of the four varieties i.e. quick succession, superposition, masking and suppression that are plausible to the phenomenon of blended emotions [3].

To measure the blended effect of simultaneous emotions it is important to capture their properties. Each emotion is unique and can be distinguished in terms of their

dynamic properties i.e. intensity, valence and dominance.

Therefore, fear, anger, sadness, and disgust are negative primary emotions while happiness and surprise are positive ones. Scherer modeled emotions as a continuous progression in three-dimensional space that consists of Pleasure/Valence (P), representing the overall valence information, Arousal (A), accounting for the degree of activeness of an emotion, and Dominance/ Power (D), describing the experienced "control" over the emotion itself or the situational context [4].

The association between primary emotions caters the possibility of complex or secondary emotions. The dynamics of secondary emotions emerged from the association of primary emotions and experience [5], which are possible to be mapped along the PAD coordinates for example, smugness might be considered a blend of the two elemental emotions: happiness, and contempt [6].

The fuzzy theory has been used recently by different researchers in emotion modeling for artificial agents. Arief has discussed visualization of emotional facial expression using Naive Bayes and Fuzzy logic [7]. Elisabetta presented an embodied conversational agent to show complex emotional facial expression[8].

Emotions are an important feature of non-verbal HCI. Ayesha has presented her research work in this area using rough fuzzy sets to resolve the complexity [9]. A fuzzy emotion model VISBER is presented by Natascha that is applicable for the real time facial emotion recognition in agents [10]. One of the advantages of Fuzzy logic includes that it captures the changes smoothly from the environment and generate even output instead of crisp values [11].

This research paper proposes an inference system using the fuzzy logic to process the dynamics of simultaneously active emotions and estimates their mixed effects

————————————————

- Ms. Atifa Athar is working as a faculty member at the School of Computer Science, NCBA&E Lahore, Pakistan. She is a research fellow at SCS, NCBA&E (e-mail: atifaathar@yahoo.com)
- Dr. M. Saleem Khan is with the Computer Science Department as Director in GC University Lahore, Pakistan (e-mail: mskgcu@yahoo.com).
- Dr. Khalil Ahmed is with the School of Computer Science at NCBA&E Lahore, Pakistan, He is an expert academician and passionately engaged in research. (e-mail: drk@ncbae.edu.pk).
- Ms. Aiesha Ahmad is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan (e-mail: aiesha@ncbae.edu.pk).
- Ms. Nida Anwar is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan. She is currently working as faculty member in VU, Lahore, Pakistan (e-mail: nidaanwar85@gmail.com)

or transitional succession. The output of the system could be used in artificial agency for the decision making etc.

The emotionless agents are simply considered as machines and not trustworthy for humans [12]. Therefore, in accordance with human psychology, the role of agents as our future social cohorts suggests that these are required to be equipped with integrated mechanism of complex emotional dynamics to communicate reason and decide in conflicting situations like humans.

This paper is structured as follows: Section 2 elaborates the approach adopted for emotion blends. Section 3 explains the fuzzy logic scheme. Section 4 argues about the fuzzy rule base scheme to infer the blended effect of simultaneous emotions. Section 5 provides the comparison of Matlab simulated and calculated results. Section 6 concludes this work and suggests some possible future directions.

2 APROACH FOR EMOTION BLEND IN PSYCHOLOGY

According to evolutionary and developmental psychology it is known that a human mind may not experience just one basic emotion at a time rather it feels simultaneous or complex emotions while sensing the environmental events. The term "basic" has been used to describe elements that combine to form more complex or compound emotions for example, smugness might be considered a blend of the two elemental emotions, happiness, and contempt.[6] Considering psycho-evolutionary theory of emotions articulated by Plutchik there are eight primary emotions in humans i.e. anger, anticipation, joy, trust, fear, surprise, sadness and disgust,. Fig 1 presents the Plutchik's complete scheme of human emotions having eight basic emotions as "Wheel of Emotions". [13]
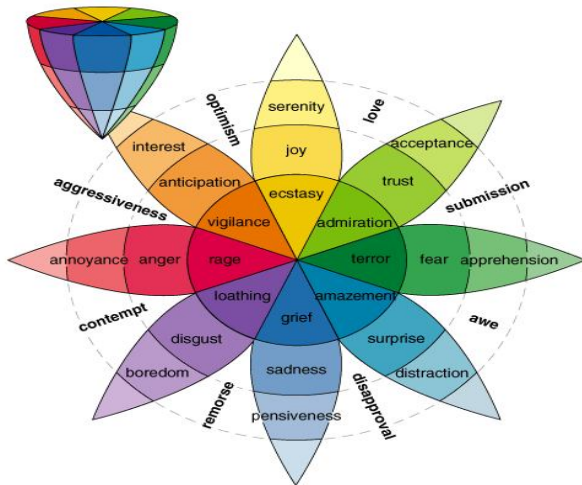
Figure 1: Wheel of Emotions

The blend of emotions at three different levels could be achieved through primary, secondary and tertiary dyads on the Wheel. [14] Each primary dyad shows the combi-

nations of neighboring pairs of emotion on the wheel e.g. trust and fear leads to submission. The primary blend of eight basic emotions is presented in table I.

Table I: Primary Level Blended Emotions

	Primary Blend	Emotion Generated from Primary Blend
1	Anger + Anticipation	Aggressiveness
2	Anticipation + Joy	Optimism
3	Joy + Trust	Love
4	Trust + Fear	Submission
5	Fear + Surprise	Alarm / Awe
6	Surprise + Sadness	Disappointment
7	Sadness + Disgust	Remorse
8	Disgust + Anger	Contempt

Likewise secondary dyad combines two emotions with the gap of one emotion on the wheel and produces the complex emotions from primary ones as presented in table II.

Table II: Secondary Level Blended Emotions

	Secondary Blend	Emotion Generated from Secondary Blend
1	Anger + Joy	Pride
2	Anticipation + Trust	Hope
3	Joy + Fear	Guilt
4	Trust + Surprise	Curiosity
5	Fear + Sadness	Despair
6	Surprise + Disgust	Unbelief
7	Sadness + Anger	Envy
8	Disgust + Anticipation	Cynicism

The tertiary dyad combines emotions with the gap of two emotions on the Wheel, presented in table III.

Table III: Tertiary Level Blended Emotions

	Tertiary Blend	Emotion Generated from Tertiary Blend
1	Anger + Trust	Dominance
2	Anticipation + Fear	Anxiety
3	Joy + Surprise	Delight
4	Trust + Sadness	Sentimentality
5	Fear + Disgust	Shame
6	Surprise + Anger	Outrage
7	Sadness + Anticipation	Pessimism
8	Disgust + Joy	Morbidness

3 FUZZY LOGIC SCHEME

Fuzzy logic is based on Boolean logic and works with partially true or false values. The fuzzy systems deals with the truth values in Fuzzy logic or membership values in fuzzy sets that are indicated by a value on the range [0.0, 1.0], with 0.0 representing absolute Falseness and 1.0 representing absolute Truth.[15]

4 FUZZY INFERENCE SYSTEM FOR SYNERGY ESTIMATION OF SIMULTANEOUS EMOTIONS

From the literature it is clear that an emotion's attributes are usually not represented by crisp values rather these have fuzzy boundaries. Therefore, we propose a fuzzy inference system with the classification of primary emotional states that blends their intensities together and constitute the secondary emotion. The design of proposed system is represented in Fig. 2.
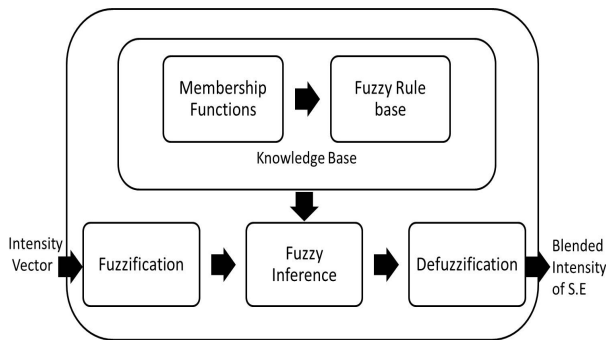
Figure 2: System Design

4.1 Fuzzification of Emotion Intensities

The fuzzification process is used to transform the crisp values of emotion intensities into degrees of membership for linguistic terms of fuzzy sets. We are considering the primary emotion identified by Plutchik represented in table IV as input to the proposed fuzzy inference system.

In the proposed scheme, we are assuming that the intensity of an emotion can be mapped to the interval [0, 1]. The following linguistic variables have been chosen for fuzzification of input intensities and divided into three types based on probability values of intensities of all eight primary emotions.

a. Low: 0.0 – 0.5
b. Medium: 0.3 – 0.7
c. High: 0.5 – 1.0

These linguistic variables are used to decide the degree of membership in fuzzification set. The plot of membership functions for one of the fuzzy input variables is represented in Fig. 3.
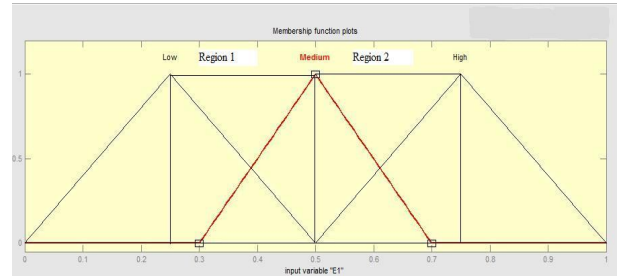
Figure 3: Membership Function Plot for Input Variable E1

The three membership functions, f1 [1] for Low, f1 [2] for Medium, and f1 [3] for High are used to show the various ranges of input fuzzy variable "E1" in a plot consisting of two regions as shown in Fig. 3. The number of membership functions and range values for each of the fuzzy input variables are taken same as E1,……..,E8 are representing the intensity of primary emotions.
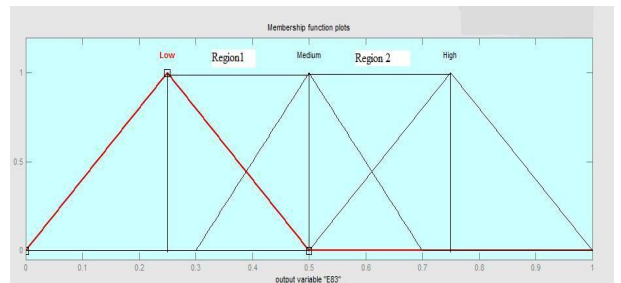
Figure 4: Membership Function Plot for Output Variable E83

Table IV: Fuzzy Input and Output Variables

Input Variables		Output Variables					
E1	Anger	E12	Aggressiveness	E13	Pride	E14	Dominance
E2	Anticipation	E23	Optimism	E24	Hope	E25	Anxiety
E3	Joy	E34	Love	E35	Guilt	E36	Delight
E4	Trust	E45	Submission	E46	Curiosity	E47	Sentimentality
E5	Fear	E56	Alarm / Awe	E57	Despair	E58	Shame
E6	Surprise	E67	Disappointment	E68	Unbelief	E61	Outrage
E7	Sadness	E78	Remorse	E71	Envy	E72	Pessimism
E8	Disgust	E81	Contempt	E82	Cynicism	E83	Morbidness

For generalization, the range values of each output membership function plot are taken same. Therefore, the shape of the plot for each output variable, used in design is the same and shown in Fig. 4.The proposed system consists of eight input variables and their values may lie in any one of the two regions as listed in Table V.

The linguistic values are the plotting values of the fuzzy input variables with the membership functions employed in different regions. As eight variables are used, therefore sixteen linguistic values are represented in Fig. 5.
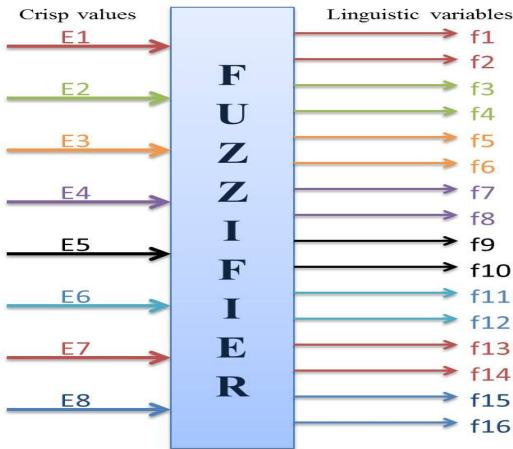
Figure 5: Block Diagram of Fuzzifier

The mapping of input fuzzy variables with the membership functions in two regions is listed in Table V.

Table V: Linguistic Values of Fuzzifier outputs in Two Regions

Input Variables	Linguistic Fuzzifier Outputs	Region 1	Region 2
E1	f_1	$f_1[1]$	$f_1[2]$
	f_2	$f_1[2]$	$f_1[3]$
E2	f_3	$f_2[1]$	$f_2[2]$
	f_4	$f_2[2]$	$f_2[3]$
E3	f_5	$f_3[1]$	$f_3[2]$
	f_6	$f_3[2]$	$f_3[3]$
E4	f_7	$f_4[1]$	$f_4[2]$
	f_8	$f_4[2]$	$f_4[3]$
E5	f_9	$f_5[1]$	$f_5[2]$
	f_{10}	$f_5[2]$	$f_5[3]$
E6	f_{11}	$f_6[1]$	$f_6[2]$
	f_{12}	$f_6[2]$	$f_6[3]$
E7	f_{13}	$f_7[1]$	$f_7[2]$
	f_{14}	$f_7[2]$	$f_7[3]$
E8	f_{15}	$f_8[1]$	$f_8[2]$
	f_{16}	$f_8[2]$	$f_8[3]$

Both of the regions are divided in two halves and each region consists of two membership functions at a time.

4.2 Fuzzy Inference Engine

After determining the degree of membership in the fuzzification process the subsequent step is to make linguistic rules to decide that which secondary emotion will be generated in response to the inputs provided to the system.

In inference process the value of intensity given to each emotion synergies and estimates the intensity of the secondary emotion produced with the help of inference rules. On the basis of number of inputs and linguistic variables, the number of fuzzy rules is determined.

In reference to the "Wheel of emotions" it is observed that secondary emotions cannot be highly intense rather these could be produced with medium intensity by blending two highly intense primary emotions. And if primary emotions are blended with medium intensity then the secondary emotion produced is comparatively less intense.
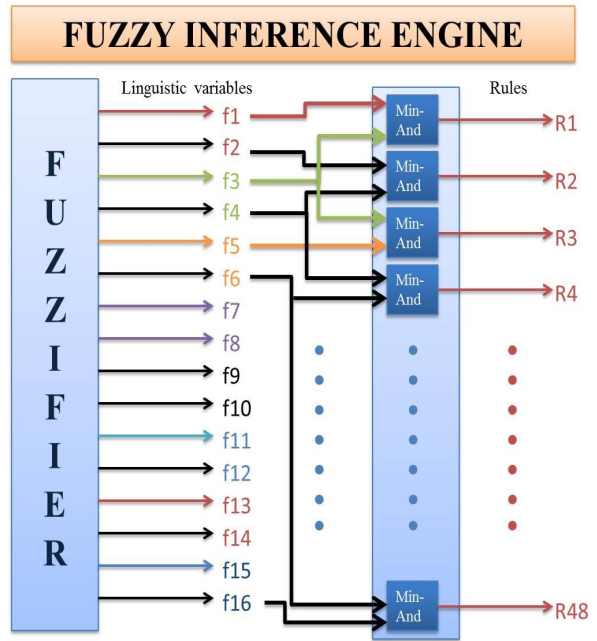
Figure 6: Block Diagram of Fuzzy Inference Process

In our proposed system there are 8 input variables i.e. E1, E2, ..., E8 and 3 linguistic variables therefore by using AND connector it is possible to have 6561 rules, but in reference to the combinations of emotions presented in table I, II, III only 24 combinations i.e. E12 . . . . E83 presented in table IV have been chosen to design the inference rules. These combinations cater the blend of only two primary emotions at three different levels.

The fuzzified intensity of emotions E1 ...... E8 is provided as input to each rule and the intensity of one of the relevant secondary emotion provided in table 5 is produced.

For example the inference rules could be

- If E1 is Medium AND E2 is Medium THEN E12 is Low
- If E6 is High AND E7 is High THEN E67 is Medium

There are 48 rules that has been used for 24 combinations because of the usage of AND connector that caters the minimum value of variables.

The inference engine contains 48 AND gates, it accepts sixteen inputs from the fuzzifier and produce R values by applying min-max composition. Each rule takes two input values, synergies these and inferences the final output using min-AND operation. Following is the calculation of R values from some of the selected rules out of 48 rules with the values E1 = 0.65, E2 = 0.7, E8 = 0.7, E3 = 0.6 in region2.

$$R_1 = f_1 \wedge f_3 = f_1[2] \wedge f_2[3] = 0.6 \wedge 0.8 = 0.6$$
$$R_2 = f_2 \wedge f_4 = f_1[2] \wedge f_2[3] = 0.4 \wedge 0.2 = 0.2$$

$$R_{47} = f_5 \wedge f_{15} = f_3[2] \wedge f_8[2] = 0.4 \wedge 0.8 = 0.4$$
$$R_{48} = f_6 \wedge f_{16} = f_3[3] \wedge f_8[3] = 0.6 \wedge 0.2 = 0.2$$

By applying the Mamdani-min process we get the minimum value from the membership function values using AND operation. The sign ^ between the membership function values represents the Min-ANDing process.

4.3 Rule Selector

The rule base works with eight crisp input values, dividing the universe of discourse into two regions, each containing two fuzzy variables, fires the rules, and gives the output singleton values corresponding to each output variable as presented in Fig. 7.
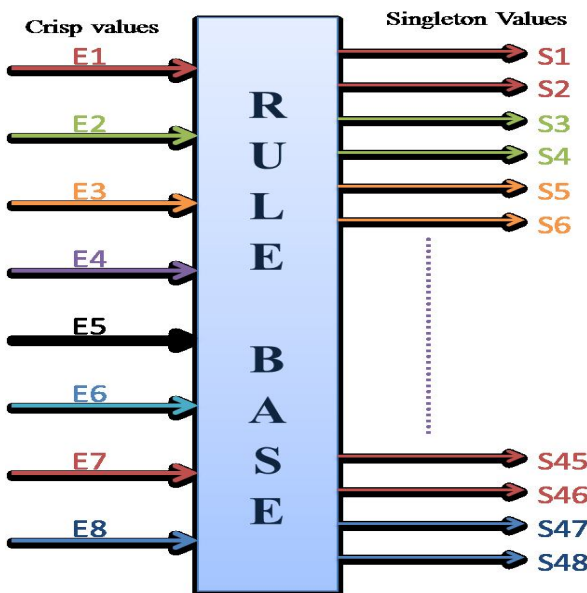
Figure 7: Block Diagram of Rule Base

The rule selector for the proposed system receives eight crisp values from the input variables. It provides singleton values of output functions according to the forty eight rules. According to the division of the regions for each output variables there are forty eight singleton values S1, S2, S3, ……., S48.

4.4 Defuzzification

After the estimation of inputs the defuzzification process generates the crisp values for output variables. In the proposed system there are 96 inputs that are provided to each of 48 defuzzifiers, forty Eight values of R1,R2,……..,R48 from the outputs of inference engine and forty eight singleton values S1, S2, ………,S48 from the rule selector as shown in Fig.7. Each defuzzifier estimates the crisp value output according to the center of average (C.O.A) method using the mathematical expression, $\Sigma S_i * R_i / \Sigma R_i$, where i = 1 to 48.

S1 = 0.25
S2 = 0.5
S3 = 0.75
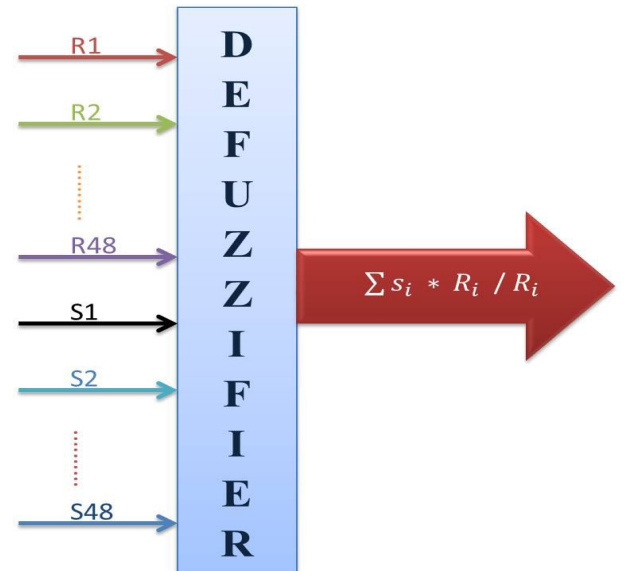
Figure 8: Block Diagram of Defuzzifier

The design formation of defuzzifier is represented in Fig. 9. Each defuzzifier consists of 48 multipliers for $S_i * R_i$ one adder to sum up, one adder for $\Sigma S_i * R_i$ and one divider for $\Sigma S_i * R_i / \Sigma R_i$ to provide the final estimated output crisp value.

MULTIPLIER
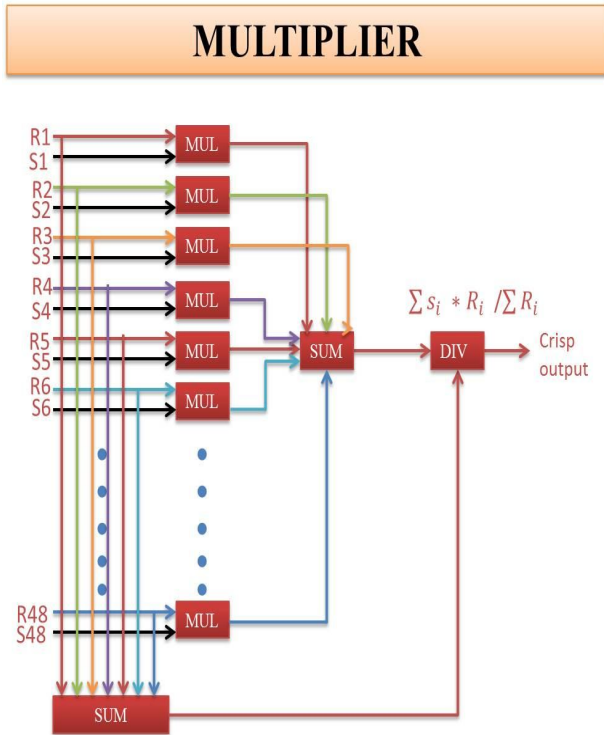
Figure 9: Defuzzifier Design

Figure 10: MATLAB Rule Viewer for Input Variables

5. COMPARISON OF SIMULATED AND CALCULATED RESULTS

The results calculated from the fuzzy inference engine

$$\sum R_i = R_{1} + R_2 + \ldots + R_{48}$$

Two output variables E12 and E83 have been chosen for the comparisons. The crisp values for E12 and E83 are determined using mathematical expression $\sum S_i * R_i / \sum R_i$ from the values of the input variables that are provided in the fuzzy inference process.

These calculated results according to the system design are compared with the results according to the MATLAB simulation in Table VI and found correct. Fig. 10 represents the input variables & Fig. 11 represents the output variables from the Rule Viewer in MATLAB.

Table VI: Comparison of Simulated and Calculated Results

Results	E12	E83
MATLAB Simulation	0.5	0.5
Calculated Values	0.56	0.58

Figure 11: MATLAB Rule Viewer for Output Variables

5. SIMULATED GRAPHS

Following is one of the simulated graphs for output E12 for the proposed inference system according to the values of E1 and E2 provided for inference process mentioned above.
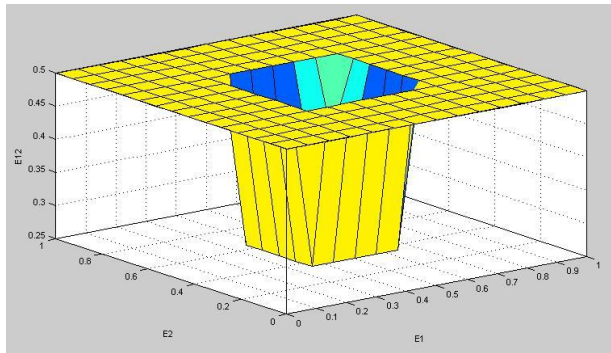
Figure 12: Plot between E1 and E2

6. CONCLUSION

There are several emotion inference systems presented by researchers but these do not infer and verify the coexistence of numerous emotions and their blends according to Plutchik theory of emotional blend. However our proposed fuzzy inference system estimates the synergy of two emotions that activates simultaneously quite efficiently.

The comparisons between calculated and simulated values afford it a role in estimation authenticity for blended emotions.

The proposed system design and simulation work could also lead to the new avenues in the field of modeling complex emotion dynamics for agency.

REFERENCES

[1] P. Scaruffi. *The Nature of Consciouss.* 2006, ISBN 0976553112

[2] G-Y. Park, S-I. Lee W-Y. Kwon and J-B. Kim. "Neurocognitive Affective System for an Emotive Robot". *Proceedings of IEEE/RSJ International Conference on Intelligent Robot and Systems* (pp. 2595 - 2600). Beijing: IEEE, China, 2006.

[3] S. Buisine, S. Abrilian, R. Niewiadomski, J. Martin, L. Devillers, and C. Pelachaud, "Perception of Blended Emotions: from Video Corpus to Expressive Agent". *Proceedings of 6th International Conference on Intelligent Virtual Agents (IVA 2006).* 4133/2006, pp. 93-106. Marina Del Rey, CA, USA: Springer Berlin / Heidelberg, August 10, 2006.

[4] C. Becker-Asano, I. Wachsmuth. "Virtual Humans Growing up: From primary toward Secondary Emotion". Autonomous Agents and Multi-Agent Systems, 20(1), pp. 32 – 49, 2009.

[5] L.Custodio, R.Ventura, and C. P.Ferreira. "Artificial Emotions and Emotion-Based Control Systems." *Proceedings of 7th IEEE International Conference on Emerging Technologies and Factory Automation .(ETFA),* 2, pp. 1415 - 1420. Barcelona 1999.

[6] P. Ekman. *Basic Emotions. In Handbook of Cognition and Emotion.* Sussex, U.K.: John Wiley & Sons, 1999.

[7] M. Arief, M. I. Wardhana, S. Sumpeno, and M. Hariad. "Emotion Expression of Three Dimensional Face Model Using Naive Bayes and Fuzzy Logic." International Journal of Computer Science and Network Security (IJCSNS), 10(5), May 2010.

[8] E. Bevacqua, M. Mancini, R. Niewiadomski and C. Pelachaud. "An expressive ECA showing complex emotions." *Proceedings of Artificial and Ambient Intelligence.(AISB, 07).* Newcastle University, Newcastle upon Tyne, UK, 2007.

[9] A. Butalia, A. R Ramani, and P. Kulkarni. "Emotional Recognition and towards Context based Decision." International Journal of Computer Applications, 9(3), pp. 42 – 54, November 2010.

[10] N. Esau, E. Wetzel, L. Kleinjohann, B. Kleinjohann, "Real-Time Facial Expression Recognition Using a Fuzzy Emotion Model." *IEEE International Conference In Fuzzy Systems,* (pp. 1 - 6), 2007.

[11] H. Mobahi, S. Ansari. "Fuzzy Perception, Emotion and Expression for Interactive Robots." *IEEE International Conference on Systems, Man and Cybernetics,* 2003 .

[12] L .PourMohammadBagher. "Intelligent Agent System Simulation Using Fear Emotion." World Academy of Science, Engineering and Technology 48, 2008.

[13] R. Plutchik and H. Kellerman. *Emotion: Theory, research, and experience.* Newyork, USA: Academic Press, 1983.

[14] J. Turner. *Origin of Human Emotions: A Sociological Inquiry into the Evolution of Human Affect,* Stanford University Press, 2000.

[15] M. Hellmann. Fuzzy Logic Introduction. Epsilon Nought Radar Remote Sensing Tutorials, 2001.

Performance Analysis of Routing Protocols Using Different Environment in MANET

Rakesh kumar Yadav, Manish Bhardwaj, Sachi Pandey

Abstract— Mobile Ad hoc networks (MANET) represents complex distributed systems that comprise wireless ad hoc network that does not rely on any fixed infrastructure and has dynamic topology also called short lived networks. MANET is self-organized and self-configurable network where the mobile nodes move arbitrarily. The mobile nodes can receive and forward packets as a router. Routing is a critical issue in MANET; efficient routing protocols will make routing reliable. The most popular ones are AODV, DSDV and DSR. The performance measurements are based on the various performance metrics such as packet delivery fraction, average end to end delay. This paper also analysis the performance of using execution time by varying different MANET simulation parameters.

Index Terms— MANET, AODV, DSDV, DSR, GloMoSim 2.03, Transmission range, Topologies, Bandwidth.

——————————— ◆ ———————————

1 INTRODUCTION

Mobile Ad Hoc Network (MANET) is a collection of communication devices or nodes that wish to communicate without any fixed infrastructure and pre-determined organization of available links. The nodes in MANET themselves are responsible for dynamically discovering other nodes to communicate. It is a self-configuring network of mobile nodes connected by wireless links the union of which forms an arbitrary topology. The nodes are free to move randomly and organize themselves arbitrarily thus, the network's wireless topology may change rapidly and unpredictably. Routing is a core problem in networks for sending data from one node to another. Such networks are aimed to provide communication capabilities to areas where limited or no communication infrastructures exist. MANET's can also be deployed to allow the communication devices to form a dynamic and temporary network among them. A mobile Ad Hoc network (MANET) is receiving attention due to many potential military and civilian applications. MANET's have several salient characteristics: 1) Dynamic topologies 2) Bandwidth-constrained links 3) Energy constrained operation 4) limited physical security. Therefore the routing protocols for wired networks cannot be directly used for wireless networks.

A MANET uses multi-hop routing instead of a static network infrastructure to provide network connectivity. Several routing protocols have been proposed for mobile Ad Hoc networks. In this paper represents the execution time analysis by varying different MANET simulation parameters, because execution time is a main parameter

————————————————

- *Rakesh Kumar Yadav, M.E, NITTTR, Chandigarh, India,E-mail: rky019@gmail.com*
- *Manish bhardwaj, Pursuing M.Tech. in Computer Science & Engg, SRM University, India, E-mail: simenscomputer4u@yahoo.com*
- *Sachi Pandey Pursuing M.Tech. in Computer Science & Engg, SRM University, India, E-mail: sachipandey_08@yahoo.co.in*

for analyzing the performance of mobile ad-hoc network. Here we are using the simulation tool GloMoSim for large scale scenarios to obtain execution time analysis by varying different simulation parameters. Because it is giving more accuracy while constructing large scale scenarios.

2 ADHOC ON DEMAND DISTANCE VECTOR ROUTING

The AODV routing protocol shares features of both DSDV and DSR algorithms. AODV shares DSR's on-demand characteristics in that it also discovers route as and when needed by initiating a route discovery process. It maintains one entry per destination in its routing tables unlike in DSR, which maintains multiple route entries for each destination in its route cache. In AODV, the packets carry the destination address and sequence number. In AODV, when a source requires a path to the destination, a route request (RREQ) message is flooded in the network.

When an intermediate node receives such a RREQ, it examines its local route cache to check whether a fresh route to the required destination is available or not. If a fresh route exists, then the node unicasts a route reply (RREP) message immediately back to the source. As an optimization, AODV uses an "expanding ring" flooding technique, where a RREQ is issued with a limited TTL only. If no RREP message is received within a certain time by the source node, then another RREQ is issued with a larger TTL value. If still no reply, the TTL is increased in steps, until a certain maximum value is reached. During route discovery process, all IP-Packets generated by the application for destination are buffered in the source node itself. When a route is established, then the packets are transmitted. An important feature of AODV is the maintenance of timer-based states in each node, regarding utilization of individual routing table entries. A routing table entry is said to be expired if not used within certain duration. These nodes are notified with route error (RERR) packets when the next-hop link breaks. In the

situation of link break, each predecessor node, forwards the RERR to its own set of predecessors. In this way all routes, which contain the broken link, are removed.

3 DESTINATION SEQUENCED DISTANCE VECTOR

DSDV is a hop-by-hop distance vector routing protocol. It is proactive; each network node maintains a routing table that contains the next-hop for, and number of hops to, all reachable destinations. Periodical broadcasts of routing updates attempt to keep the routing table completely updated at all times. To guarantee loop-freedom DSDV uses a concept of sequence numbers to indicate the freshness of a route. A route R is considered more favorable than R' if R has a greater sequence number or, if the routes have the same sequence number, R has lower hop-count.

The sequence number for a route is set by the destination node and increased by one for every new originating route advertisement. When a node along a path detects a broken route to a destination D, it advertises its route to D with an infinite hop-count and a sequence number increased by one. Route loops can occur when incorrect routing information is present in the network after a change in the network topology, e.g., a broken link. In this context the use of sequence numbers adapts DSDV to a dynamic network topology such as in an ad-hoc network. DSDV uses triggered route updates when the topology changes. The transmission of updates is delayed to introduce a damping effect when the topology is changing rapidly. This gives an additional adaptation of DSDV to ad-hoc networks.

4 DYNAMIC SOURCE ROUTING

It is one of the most well known routing algorithms for ad hoc wireless networks. It was originally developed by Johnson, Maltz, and Broch. DSR is on demand, which reduces the bandwidth use especially in situations where the mobility is low. It is a simple and efficient routing protocol for use in ad hoc networks. It has two important phases, route discovery and route maintenance. The main algorithm works in the following manner. A node that desires communication with another node first searches its route cache to see if it already has a route to the destination. If it does not, it then initiates a route discovery mechanism. This is done by sending a Route Request message. When the node gets this route request message, it searches its own cache to see if it has a route to the destination. If it does not, it then appends its id to the packet and forwards the packet to the next node; this continuous until either a node with a route to the destination is encountered (i.e. has a route in its own cache) or the destination receives the packet. In that case, the node sends a route reply packet which has a list of all of the nodes that forwarded the packet to reach the destination. This constitutes the routing information needed by the source, which can then send its data packets to the destination

using this newly discovered route. DSR can support relatively rapid rates of mobility.

5 EXECUTION TIME

In Mobile adhoc network (MANET) consist of mobile hosts without any infrastructure. Here the Execution time is the essential parameter in performance analysis for the research peoples. Execution time is the time for executing a particular scenario. Here in GloMoSim execution time is measured by

```
Execution time :   11.5000 sec
Number of messages processed : 4
Number of context switches occurred : 12
Number of Local NULL messages sent : 0
Number of Remote NULL messages sent : 0
Total Number of NULL messages sent : 0
NULL messages / Regular messages : 0.000
```

6 GLOMOSIM

Global mobile information system simulator (GloMoSim) is a scalable simulation environment for large wireless and wire line communication networks. GloMoSim is a scalable simulation environment for mobile and wireless networks developed at UCLA Parallel Computing Laboratory. GloMoSim is a discrete event simulator built using PARSEC a C based environment designed for parallel simulations also developed at UCLA Parallel Computing Laboratory. GloMoSim is built using a layered approach similar to the OSI seven layers network architecture, with standard APIs between layers. This makes it easy to implement and integrate new protocols and models at different. GloMoSim is written in C and utilizes PARSEC (Parallel Simulation Environment for Complex Systems) which is a simulation environment developed for parallel and sequential execution of discrete event simulations. By discrete event it means that the execution is mainly based on event handling, i.e., the execution consists of set of events and as an event occurs, the appropriate action is taken in its response. An event is defined as an incident which results in the change of state of the system. A certain event or combination of events may invoke other events and so on, and this is how the execution proceeds. Some examples are arrival of packet at a particular layer, expiry of a timer, etc. Events only occur at discrete units of time are not permitted to occur in between these units of time. GloMoSim is freely available for educational purposes. However the free version only supports the sequential execution of simulations.

GloMoSim simulates networks with up to thousand nodes linked by a heterogeneous communications capability that includes multicast, asymmetric communications using direct satellite broadcasts, multi-hop wireless communications using ad-hoc networking and traditional Internet protocols. The following table lists the GloMoSim

models currently available at each of the major layers mention in table 1.

TABLE 1

GLOMOSIM MODELS AT DIFFERENT LAYER

LAYER	MODELS
Physical (Radio Propagation)	Free space, Rayliegh,Two ray
Data Link	CSMA, MACA, TCMA, 802.11
Network (Routing)	Flooding, AODV, DSR, Fisheye
Transport	TCP, UDP
Application	TCPlib, Telnet, FTP

The node aggregation technique is introduced into GloMoSim to give significant benefits to the simulation performance. Initializing each node as a separate entity inherently limits the the scalability because the memory requirements increase dramatically for a model with large number of nodes. With node aggregation, a single entity can simulate several network nodes in the system. Node aggregation technique implies that the number of nodes in the system can be increased while maintaining the same number of entities in the simulation. In GloMoSim, each entity represents a geographical area of the simulation. Hence the network nodes which a particular entity represents are determined by the physical position of the nodes.

7 EXPERIMENT AND SIMULATION SETUP

Below we mention parameters values in table 2 for routing protocols simulation.

TABLE 2

PARAMETER EVELUATION

PARAMETERS	VALUES
Simulator	GlomoSim-2.03
Protocols Studied	AODV, DSDV, DSR
Traffic Type	CBR, Data stream
Movement of Node	Random Way Point
Terrain Dimension	Initial 15(not fixed)
No of Nodes	5,10,15,20,25
Sending Rate	10 Packets/seconds
Packet Size	512 Bytes

In order to evaluate the performance of ad hoc network routing protocols, the following metrics were considered:

7.1 Packet Delivery Fraction

Also known as the ratio of the data packets delivered to the destinations to those generated by the CBR sources. The PDF shows how successful a protocol performs deli-

vering packets from source to destination. The higher for the value give use the better results. This metric characterizes both the completeness and correctness of the routing protocol also reliability of routing protocol by giving its effectiveness.
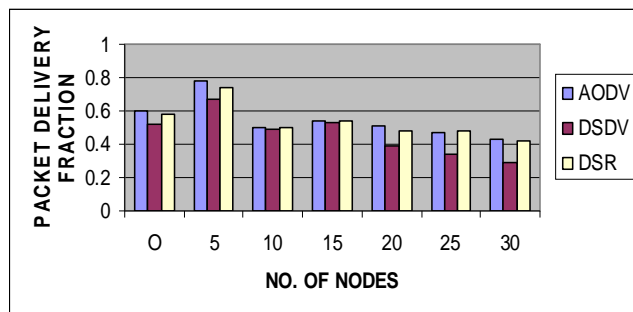
Fig. 1. Packet delivery fraction vs No. of node

7.2 Average End to End Delay

Average end-to-end delay is an average end-to-end delay of data packets. It also caused by queuing for transmition at the node and buffering data for detouring. Once the time difference between every CBR packet sent and received was recorded, dividing the total time difference over the total number of CBR packets received gave the average end-to-end delay for the received packets.
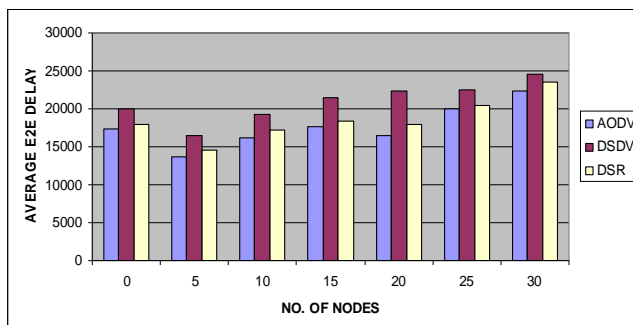
Fig. 2. Average E2E delay vs No. of nodes

This graph shows that the com,parison between the three routing protocol like DSDV, AODV,and DSR.According to this graph the average end to end delay of two routing protocol like AODV and DSR are very similar or say that it contain very minute difference.

7.3 Execution Time

7.3.1 AODV, DSR, DSDV VS EXECUTION TIME

Here we compared the execution time analysis by using AODV, DSR and DSDV routing protocols. And we noticed the execution time varies as follows
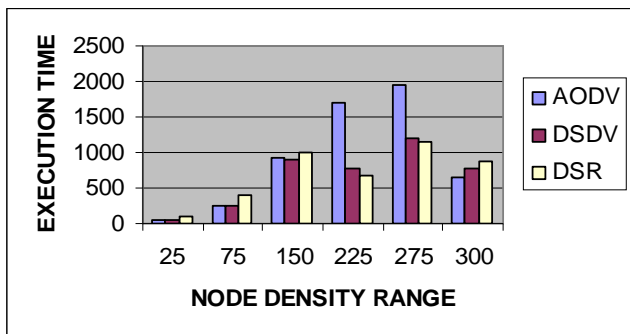
Fig. 3. Execution time vs Node density range

7.3.2 TERRAIN RANGE VS EXECUTION TIME

In the same scenario we changed the terrain dimensions, the execution time varies or high for increasing the terrain dimension. The result is following in the figure
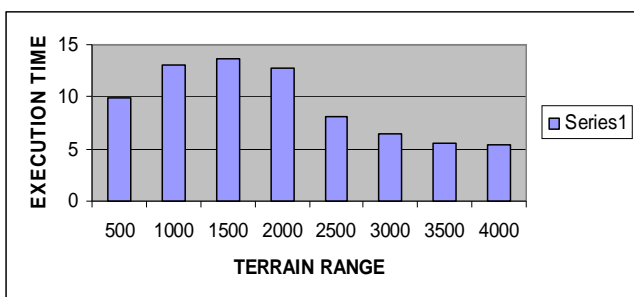
Fig. 4. Execution time vs Terrain range

7.3.3 TRANSMISSION RANGE VS EXECUTON TIME

In Mobile Ad hoc Network the mobile nodes are battery constraint.so the transmission range is very essential parameter to analyse the Execution time. The following graph has shown the analysis of execution time with varying transmission range.
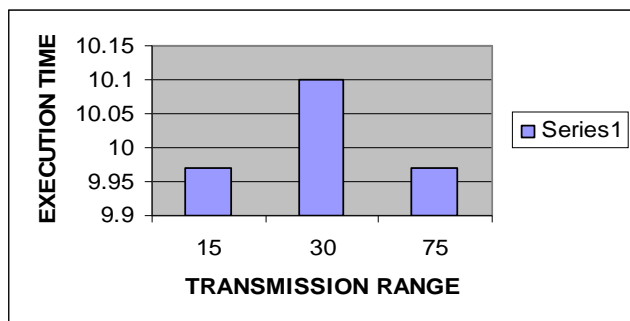
Fig. 5. Execution time vs Transmission range

8 CONCLUSION

The work was initiated with an intention of carrying out exhaustive study of the performance issues of large scale MANETs. We obtained convincing results for the execu-

tion times and it's affected while changing network parameters vs. real times in large scale scenario.

REFERENCES

[1] R.Balakrishna, M.V.Panduranga Rao, Dr.K.C.Shet," Development of Scheduler for Real Time and Embedded System Domain", in digital library at IEEE AINA-2008 International Conference, JAPAN.The details about the conference is available at http://www.aina-conference.org/2008

[2] A. K. Gupta "Performance analysis of AODV, DSR and TORA Routing Protocols" IACSIT international journal of Engineering and Technology, Vol. 2, No. 2, April 2010 ISSN: 1793-8236.

[3] C. E. Perkins and E. M. Royer, Ad hoc Networking, chapter Ad hoc On-Demand Distance Vector Routing, Addison Wesley, 2009.

[4] Ahmed Al- Maashri and Mohamed Ould-Khaoua."Performance Analysis of MANET Routing Protocols in the Presence of Self- Similar Traffic". IEEE, ISSN- 0742-1303 . First published in Proc. of the 31st IEEE Conference on Local Computer Networks, 2006.

[5] Khan, Khaleel Ur Rahman; Zaman, Rafi U.; Reddy, A. Venugopal," Performance Comparison of On-Demand and Table Driven Ad Hoc Routing protocols Using NCTUns",, UKSIM: IEEE Tenth International Conference on Computer Modeling and Simulation Volume , Issue , 1-3 , Page(s):336 – 341, April 2008.

[6] A.E. Mahmoud, R. Khalaf & A, Kayssi, Performance Comparison of the AODV and DSDV Routing Protocols in Mobile Ad-Hoc Networks, Lebanon, 2007.

[7] C. E. Perkins and E. M. Royer, "Ad-hoc On-Demand Distance Vector Routing," Proc. 2nd IEEE Wksp. Mobile Comp. Sys. and Apps., Feb. 1999, pp. 90–100.

[8] GloMoSim: Global Mobile Information Systems Simulation Library. http://pcl.cs.ucla.edu/projects/glomosim/

[9] IEEE 802.11: part 11: Wireless LAN Medium Access control (MAC) and Physical Layer (PHY) specification, Aug.1999

[10] S. Bhatt, R. Fujimoto, A. Ogieski and K. Permalla. Parallel Simulation Techniques for Large Scale Networks. IEEE Communication Magazine, 98, pp. 42-47.

[11] Stuart Kurkowski, Tracy Camp, Michael Cologrosso, "MANET Simulation Studies: The Incredible", Mobile Computing and Communications Review, Volume 9, Number 4, ACM, 2005.

[12] Tope. M.A, McEachen, J.C, and Kinney. A.C, "Ad-hoc net work routing using co-operative diversity", Advanced Information Networking and Applications, IEEEConference, 18-20 April 2006.

[13] A.Boomarani Malany, V.R.Sarma Dhulipala, and RM.Chandrasekaran "Throughput and Delay Comparison of MANET Routing Protocols" Int. J. Open Problems Compt. Math., Vol. 2, No. 3, September 2009.

[14] V.R. Sarma Dhulipala, RM.Chandrasekaran, R.Prabakaran " Timing Analysis and Repeatability Issues of Mobile Ad-Hoc Networking Application traffics in Large Scale Scenarios " International Journal of Recent Trends in Engineering, Vol. 1, No. 1, May ,2009.

A semi-implicit finite-difference approach for two-dimensional coupled Burgers' equations

Mohammad Tamsir, Vineet Kumar Srivastava

Abstract— In this paper, a semi-implicit finite-difference method is used to find the numerical solution of two-dimensional Coupled Burgers' equation. The proposed scheme forms a system of linear algebraic difference equations to be solved at each time-step. The linear system is solved by direct method. Numerical results are compared with those of exact solutions and other available results. The present method performs well. The proposed scheme can be extended for solving non-linear problems arising in mechanics and other areas of engineering and science.

Index Terms— *Burgers'* equations; finite- difference; semi-implicit scheme; Reynolds number.

————————————— ◆ —————————————

1 INTRODUCTION

THE two-dimensional Burgers' equation is a mathematical model which is widely used for various physical applications, such as modeling of gas dynamics and traffic flow, shock waves [1], investigating the shallow water waves[2,3], in examining the chemical reaction-diffusion model of Brusselator[4] etc. It is also used for testing several numerical algorithms. The first attempt to solve Burgers' equation analytically was given by Bateman [5], who derived the steady solution for a simple one-dimensional Burgers' equation, which was used by J.M. Burger in [6] to model turbulence. In the past several years, numerical solution to one-dimensional Burgers' equation and system of multidimensional Burgers' equations have attracted a lot of attention from both scientists and engineers and which has resulted in various finite-difference, finite-element and boundary element methods. Since in this paper the focus is numerical solution of the two-dimensional Burgers' equations, a detailed survey of the numerical schemes for solving the one-dimensional Burgers' equation is not necessary. Interested readers can refer to [7-13] for more details.

Consider two-dimensional coupled nonlinear viscous Burgers' equations:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} - \frac{1}{\text{Re}}(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}) = 0, \qquad (1)$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} - \frac{1}{\text{Re}}(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}) = 0, \qquad (2)$$

subject to the initial conditions:

$$u(x, y, 0) = \psi_1(x, y); \ (x, y) \in \Omega,$$
$$v(x, y, 0) = \psi_2(x, y); \ (x, y) \in \Omega,$$

and boundary conditions:

$$u(x, y, t) = \xi(x, y, t); \ (x, y) \in \partial\Omega, \ t > 0,$$
$$v(x, y, t) = \zeta(x, y, t); \ (x, y) \in \partial\Omega, \ t > 0,$$

where $\Omega = \{(x, y) : a \le x \le b, \ c \le y \le d\}$ and $\partial\Omega$ is its boundary; $u(x, y, t)$ and $v(x, y, t)$ are the velocity components to be determined, $\psi_1, \psi_2, \ \xi$ and ζ are known functions and Re is the Reynolds number.

The analytic solution of eqns. (1) and (2) was proposed by Fletcher using the Hopf-Cole transformation [14]. The numerical solutions of this system of equations have been solved by many researchers. Jain and Holla [15] developed two algorithms based on cubic spline method. Fletcher [16] has discussed the comparison of a number of different numerical approaches. Wubs and Goede [17] have applied an explicit–implicit method. Goyon [18] used several multilevel schemes with ADI. Recently A. R. Bahadır [19] has applied a fully implicit method. Vineet etl.[20] have used Crank-Nicolson scheme for numerical solutions of two dimensional coupled Burgers' eqations. The usual implicit schemes are obviously unconditionally stable with higher order truncation error $O((\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2)$. However, they involve solving a nonlinear algebraic system of equations which makes it inefficient in practice. In this paper, to resolve the above issue, the semi-implicit scheme proposed by

_____

- *Mohammad Tamsir, Faculty member, Department of Mathematics, The ICFAI University, Dehradun, India, Email: tamsiriitm@gmail.com.*
- *Vineet Kumar Srivastava, Faculty member, Department of Mathematics, The ICFAI University, Dehradun, India, E-mail: vineetsriiitm@gmail.com.*

Ozis [7] is used for solving two-dimensional Burgers' equations which has a truncation error $O((\Delta t) + (\Delta x)^2 + (\Delta y)^2)$. Three numerical experiments have been carried out and their results are presented to illustrate the efficiency of the proposed method.

2 THE SOLUTION PROCEDURE

The computational domain Ω is discretized with uniform grid. Denote the discrete approximation of $u(x, y, t)$ and $v(x, y, t)$ at the grid point $(i\Delta x, j\Delta y, n\Delta t)$ by $u_{i,j}^n$ and $v_{i,j}^n$ respectively $(i = 0, 1, 2......, n_x; j = 0, 1, 2....., n_y; n = 0, 1, 2......)$, where $\Delta x = 1/n_x$ is the grid size in x-direction, $\Delta y = 1/n_y$ is the grid size in y-direction, and Δt represents the increment in time.

Semi-implicit finite-difference approximation to (1) and (2) are given by:

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} + u_{i,j}^n (\frac{u_{i+1,j}^{n+1} - u_{i-1,j}^{n+1}}{2\Delta x}) + v_{i,j}^n (\frac{u_{i,j+1}^{n+1} - u_{i,j-1}^{n+1}}{2\Delta y})$$

$$-\frac{1}{\text{Re}}[(\frac{u_{i+1,j}^{n+1} - 2u_{i,j}^{n+1} + u_{i-1,j}^{n+1}}{(\Delta x)^2}) + (\frac{u_{i,j+1}^{n+1} - 2u_{i,j}^{n+1} + u_{i,j-1}^{n+1}}{(\Delta y)^2})] = 0$$

$$\frac{v_{i,j}^{n+1} - v_{i,j}^n}{\Delta t} + u_{i,j}^n (\frac{v_{i+1,j}^{n+1} - v_{i-1,j}^{n+1}}{2\Delta x}) + v_{i,j}^n (\frac{v_{i,j+1}^{n+1} - v_{i,j-1}^{n+1}}{2\Delta y})$$

$$-\frac{1}{\text{Re}}[(\frac{v_{i+1,j}^{n+1} - 2v_{i,j}^{n+1} + v_{i-1,j}^{n+1}}{(\Delta x)^2}) + (\frac{v_{i,j+1}^{n+1} - 2v_{i,j}^{n+1} + v_{i,j-1}^{n+1}}{(\Delta y)^2})] = 0$$

The above linear system of equations is solved by direct method.

3 NUMERICAL EXAMPLES AND DISCUSSION

3.1 Problem 1

The exact solutions of Burgers' equations (1) and (2) can be generated by using the Hopf–Cole transformation [3] which is:

$$u(x, y, t) = \frac{3}{4} - \frac{1}{4[1 + \exp((-4x + 4y - t)\,\text{Re}/32\,)]},$$

$$v(x, y, t) = \frac{3}{4} + \frac{1}{4[1 + \exp((-4x + 4y - t)\,\text{Re}/32\,)]},$$

Here the computational domain is taken as a square domain $\Omega = \{(x, y) : 0 \le x \le 1, 0 \le y \le 1\}$. The initial and boundary conditions for $u(x, y, t)$ and $v(x, y, t)$ are tak-

en from the analytical solutions. The numerical computations are performed using uniform grid, with a mesh width $\Delta x = \Delta y = 0.05$. From Tables 1-4, it is clear that the results from the present study are in good agreement with the exact solution for different values of Reynolds number. Comparison of numerical and exact solutions for u and v for $\text{Re}=100$ at $t = 0.5$ with $\Delta t = 0.001$ are shown in Figs. 1-4.

3.2 Problem 2.

Here the computational domain is taken as $\Omega = \{(x, y) : 0 \le x \le 0.5, 0 \le y \le 0.5\}$ and Burgers' equations (1) and (2) are taken with the initial conditions:

$$u(x, y, 0) = \sin(\pi x) + \cos(\pi y)$$

$$v(x, y, 0) = x + y$$

and boundary conditions:

$$\left. \begin{array}{ll} u(0, y, t) = \cos(\pi y), & u(0.5, y, t) = 1 + \cos(\pi y) \\ v(0, y, t) = y, & v(0.5, y, t) = 0.5 + y \end{array} \right\} 0 \le y \le 0.5,\ t \ge 0,$$

$$\left. \begin{array}{ll} u(x, 0, t) = 1 + \sin(\pi x), & u(x, 0.5, t) = \sin(\pi x) \\ v(x, 0, t) = x, & v(x, 0.5, t) = x + 0.5 \end{array} \right\} 0 \le x \le 0.5,\ t \ge 0,$$

The numerical computations are performed using 20×20 grids and $\Delta t = 0.0001$. The steady state solutions for $\text{Re} = 50$ and $\text{Re} = 500$ are obtained at $t = 0.625$. Perspective views of u and v for $\text{Re} = 50$ at $\Delta t = 0.0001$ are given in Figs. 5 and 6 respectively. The results given in Tables 5-8 at some typical mesh points (x, y) demonstrate that the proposed scheme achieves similar results given by [15, 19].

3.3 Problem 3.

In this problem the computational domain is $\Omega = \{(x, y) : 0 \le x \le 1, 0 \le y \le 1\}$ and Burgers' equations (1) and (2) are taken with the initial conditions:

$$u(x, y, 0) = \frac{-4\pi \cos(2\pi x)\sin(\pi y)}{\text{Re}(2 + \sin(2\pi x)\sin(\pi y))}, \qquad (x,y) \in \Omega$$

$$v(x, y, 0) = \frac{-2\pi \sin(2\pi x)\cos(\pi y)}{\text{Re}(2 + \sin(2\pi x)\sin(\pi y))}, \qquad (x,y) \in \Omega$$

with boundary conditions:

$$u(0, y, t) = -\frac{2\pi e^{-\frac{5\pi^2 t}{\text{Re}}}\sin(\pi y)}{\text{Re}}, \quad t \ge 0;$$

$$u(1, y, t) = - \frac{2\pi e^{-\frac{5\pi^2 t}{\text{Re}}} \sin(\pi y)}{\text{Re}}, \quad t \geq 0$$

$$u(x, 0, t) = 0, \quad t \geq 0; \quad u(x, 1, t) = 0, \quad t \geq 0$$

$$v(0, y, t) = 0, \quad t \geq 0; \quad v(1, y, t) = 0, \quad t \geq 0$$

$$v(x, 0, t) = - \frac{\pi e^{-\frac{5\pi^2 t}{\text{Re}}} \sin(2\pi x)}{\text{Re}}, \quad t \geq 0;$$

$$v(x, 1, t) = \frac{\pi e^{-\frac{5\pi^2 t}{\text{Re}}} \sin(2\pi x)}{\text{Re}}, \quad t \geq 0$$

for which the exact solutions are:

$$u(x, y, t) = - \frac{4\pi e^{-\frac{5\pi^2 t}{\text{Re}}} \cos(2\pi x)\sin(\pi y)}{\text{Re}(2 + e^{-\frac{5\pi^2 t}{\text{Re}}} \sin(2\pi x)\sin(\pi y))}$$

$$v(x, y, t) = - \frac{2\pi e^{-\frac{5\pi^2 t}{\text{Re}}} \sin(2\pi x)\cos(\pi y)}{\text{Re}(2 + e^{-\frac{5\pi^2 t}{\text{Re}}} \sin(2\pi x)\sin(\pi y))}$$

The computed solutions for u and v are ploted in Figs 7 and 9 respestively while the analytical solutions for u and v are shown in Figs 8 and 10 at 20×20 grids and at time level $t = 1.0$ with $\Delta t = 0.001$ for $\text{Re} = 1000$. From these figures it is obvious that numerical solutions are in excellent agreement with the corresponding analytical solutions.

Table 1.
The numerical results for u in comparison with the exact solution at $t = 0.01$ and $t = 1$ with $\Delta t = 0.0001$, and $\text{Re} = 10$.

(x, y)	t=0.01		t=1.0	
	Numerical	Exact	Numerical	Exact
(0.1, 0.1)	0.624805	0.624805	0.605626	0.605626
(0.5, 0.1)	0.594202	0.594202	0.576840	0.576840
(0.9, 0.1)	0.567082	0.567082	0.553017	0.553017
(0.3, 0.3)	0.624805	0.624805	0.605627	0.605626
(0.7, 0.3)	0.594202	0.594202	0.576840	0.576840
(0.1, 0.5)	0.655431	0.655431	0.636685	0.636685
(0.5, 0.5)	0.624805	0.624805	0.605628	0.605626
(0.9, 0.5)	0.594202	0.594202	0.576840	0.576840
(0.3, 0.7)	0.655431	0.655431	0.636687	0.636685
(0.7, 0.7)	0.624805	0.624805	0.605629	0.605626
(0.1, 0.9)	0.682611	0.682611	0.666353	0.666353
(0.5, 0.9)	0.655431	0.655431	0.636687	0.636685
(0.9, 0.9)	0.624805	0.624805	0.605628	0.605626

Table 2.
The numerical results for v in comparison with the exact solution at $t = 0.01$ and $t = 1.0$ with $\Delta t = 0.0001$, and $\text{Re} = 10$.

(x, y)	t=0.01		t=1.0	
	Numerical	Exact	Numerical	Exact
(0.1, 0.1)	0.875195	0.875195	0.894374	0.894374
(0.5, 0.1)	0.905798	0.905798	0.923160	0.923160
(0.9, 0.1)	0.932918	0.932918	0.946983	0.946983
(0.3, 0.3)	0.875195	0.875195	0.894373	0.894374
(0.7, 0.3)	0.905798	0.905798	0.923160	0.923160
(0.1, 0.5)	0.844569	0.844569	0.863315	0.863315
(0.5, 0.5)	0.875195	0.875195	0.894372	0.894374
(0.9, 0.5)	0.905798	0.905798	0.923160	0.923160
(0.3, 0.7)	0.844569	0.844569	0.863313	0.863315
(0.7, 0.7)	0.875195	0.875195	0.894371	0.894374
(0.1, 0.9)	0.817389	0.817389	0.833647	0.833647
(0.5, 0.9)	0.844569	0.844569	0.863313	0.863315
(0.9, 0.9)	0.875195	0.875195	0.894372	0.894374

Table 3
The numerical results for u in comparison with the exact solution at $t = 0.01$ and $t = 1.0$ with $\Delta t = 0.0001$, and $\text{Re} = 100$.

(x, y)	t=0.01		t=1.0	
	Numerical	Exact	Numerical	Exact
(0.1, 0.1)	0.623106	0.623047	0.510307	0.510522
(0.5, 0.1)	0.501617	0.501622	0.500072	0.500074
(0.9, 0.1)	0.500011	0.500011	0.500000	0.500000
(0.3, 0.3)	0.623106	0.623040	0.509824	0.510522
(0.7, 0.3)	0.501617	0.501622	0.500067	0.500074
(0.1, 0.5)	0.748272	0.748274	0.716947	0.716759
(0.5, 0.5)	0.623106	0.623047	0.509499	0.510522
(0.9, 0.5)	0.501617	0.501622	0.500063	0.500074
(0.3, 0.7)	0.748272	0.748274	0.717266	0.716759
(0.7, 0.7)	0.623106	0.623047	0.509314	0.510522
(0.1, 0.9)	0.749988	0.749988	0.749738	0.749742
(0.5, 0.9)	0.748272	0.748274	0.717530	0.716759
(0.9, 0.9)	0.623106	0.623047	0.509172	0.510522

Table 4
The numerical results for v in comparison with the exact solution at $t = 0.01$ and $t = 1.0$ with $\Delta t = 0.0001$, and $\text{Re} = 100$.

(x, y)	t=0.01		t=1.0	
	Numerical	Exact	Numerical	Exact
(0.1, 0.1)	0.876894	0.876953	0.989693	0.989478
(0.5, 0.1)	0.998383	0.998378	0.999928	0.999926
(0.9, 0.1)	0.999989	0.999989	1.000000	1.000000
(0.3, 0.3)	0.876894	0.876953	0.990176	0.989478
(0.7, 0.3)	0.998383	0.998378	0.999933	0.999926
(0.1, 0.5)	0.751728	0.751726	0.783053	0.783241
(0.5, 0.5)	0.876894	0.876953	0.990501	0.989478
(0.9, 0.5)	0.998383	0.998378	0.999937	0.999926
(0.3, 0.7)	0.751728	0.751726	0.782734	0.783241
(0.7, 0.7)	0.876894	0.876953	0.990686	0.989478
(0.1, 0.9)	0.750012	0.750012	0.750262	0.750258
(0.5, 0.9)	0.751728	0.751726	0.782470	0.783241
(0.9, 0.9)	0.876894	0.876953	0.990828	0.989478

Table 5.

Comparison of computed values of u for $\mathrm{Re} = 50$ at $t = 0.625$.

(x, y)	Present work	A.R.Bahadir	Jain and Holla
(0.1, 0.1)	0.97146	0.96688	0.97258
(0.3, 0.1)	1.15280	1.14827	1.16214
(0.2, 0.2)	0.86308	0.85911	0.86281
(0.4, 0.2)	0.97984	0.97637	0.96483
(0.1, 0.3)	0.66316	0.66019	0.66318
(0.3, 0.3)	0.77232	0.76932	0.77030
(0.2, 0.4)	0.58181	0.57966	0.58070
(0.4, 0.4)	0.75860	0.75678	0.74435

Table 6.

Comparison of computed values of v for $\mathrm{Re} = 50$ at $t = 0.625$.

(x, y)	Present work	A.R.Bahadir	Jain and Holla
(0.1, 0.1)	0.09869	0.09824	0.09773
(0.3, 0.1)	0.14158	0.14112	0.14039
(0.2, 0.2)	0.16754	0.16681	0.16660
(0.4, 0.2)	0.17110	0.17065	0.17397
(0.1, 0.3)	0.26378	0.26261	0.26294
(0.3, 0.3)	0.22655	0.22576	0.22463
(0.2, 0.4)	0.32851	0.32745	0.32402
(0.4, 0.4)	0.32501	0.32441	0.31822

Table 7.

Comparison of computed values of u for $\mathrm{Re} = 500$ at $t = 0.625$.

(x, y)	Present work	A.R.Bahadir	Jain and Holla
(0.15, 0.1)	0.96870	0.96650	0.95691
(0.3, 0.1)	1.03200	1.02970	0.95616
(0.1, 0.2)	0.86178	0.84449	0.84257
(0.2, 0.2)	0.87813	0.87631	0.86399
(0.1, 0.3)	0.67920	0.67809	0.67667
(0.3, 0.3)	0.79945	0.79792	0.76876
(0.15, 0.4)	0.66039	0.54601	0.54408
(0.2, 0.4)	0.58958	0.58874	0.58778

Table 8.

Comparison of computed values of v for $\mathrm{Re} = 500$ at $t = 0.625$.

(x, y)	Present work	A.R.Bahadir	Jain and Holla
(0.15, 0.1)	0.09043	0.09020	0.10177
(0.3, 0.1)	0.10728	0.10690	0.13287
(0.1, 0.2)	0.17295	0.17972	0.18503
(0.2, 0.2)	0.16816	0.16777	0.18169
(0.1, 0.3)	0.26268	0.26222	0.26560
(0.3, 0.3)	0.23550	0.23497	0.25142
(0.15, 0.4)	0.29022	0.31753	0.32084
(0.2, 0.4)	0.30418	0.30371	0.30927
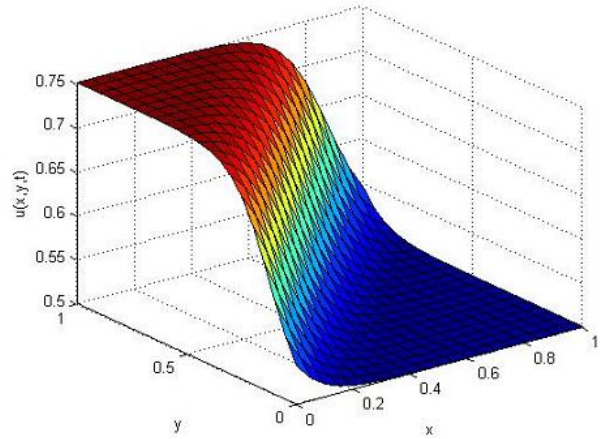
Fig.1. The numerical value of u for $\mathrm{Re} = 100$ at time level $t = 0.5$ with $\Delta t = 0.0001$.
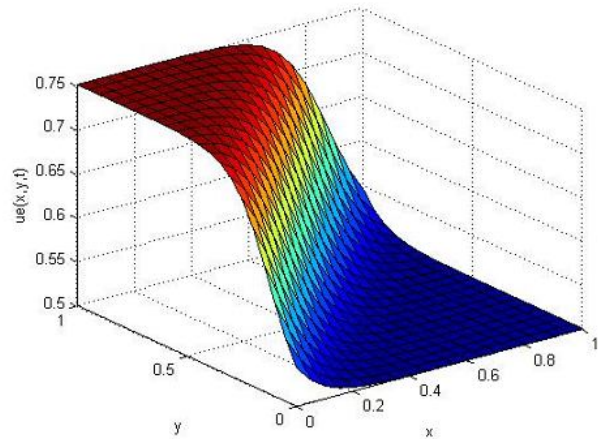
Fig.2. The exact value of u for $\mathrm{Re} = 100$ at time level $t = 0.5$ with $\Delta t = 0.0001$.
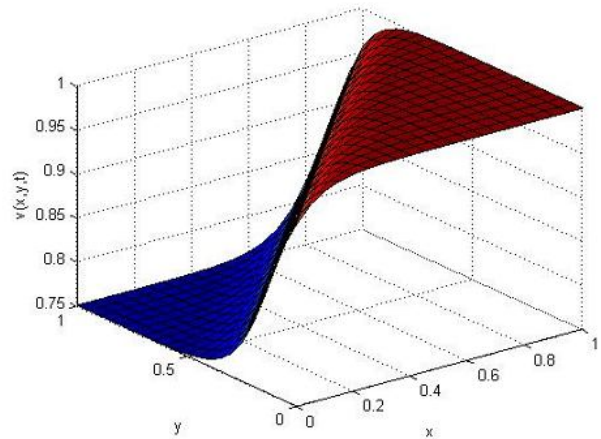
Fig.3. The numerical value of v for $\mathrm{Re} = 100$ at time level $t = 0.5$ with $\Delta t = 0.0001$.
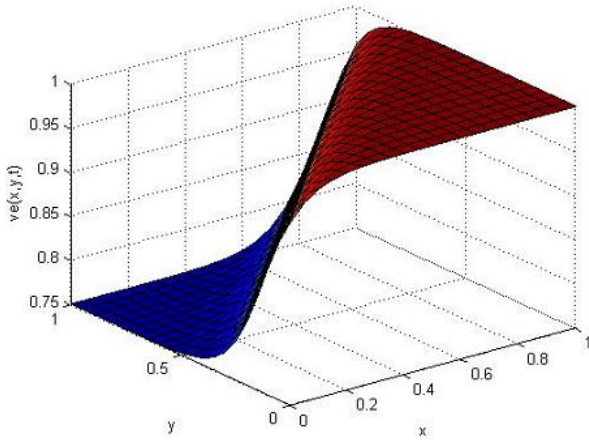
Fig.4. The exact value of v for $\mathrm{Re} = 100$ at time level $t = 0.5$ with $\Delta t = 0.0001$.
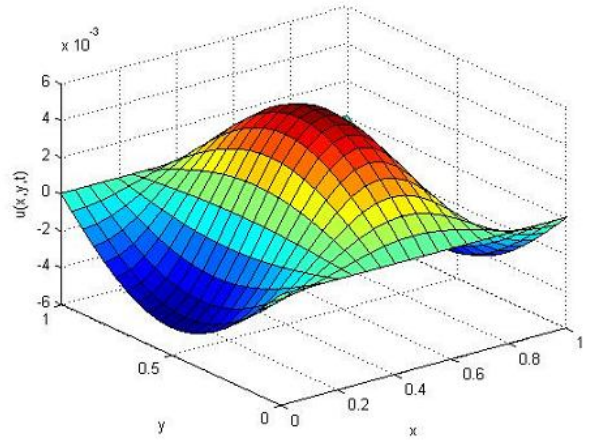
Fig.7. The numerical value of u at 20×20 grids for $\mathrm{Re} = 1000$ and at time level $t = 1.0$ with $\Delta t = 0.001$.
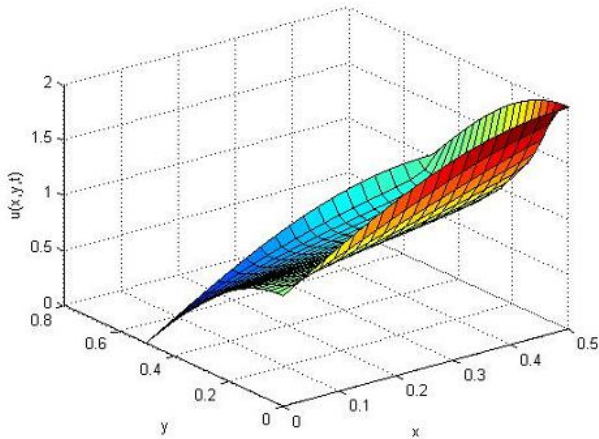
Fig.5. The computed value of u for $\mathrm{Re} = 50$ at time level $t = 0.625$.
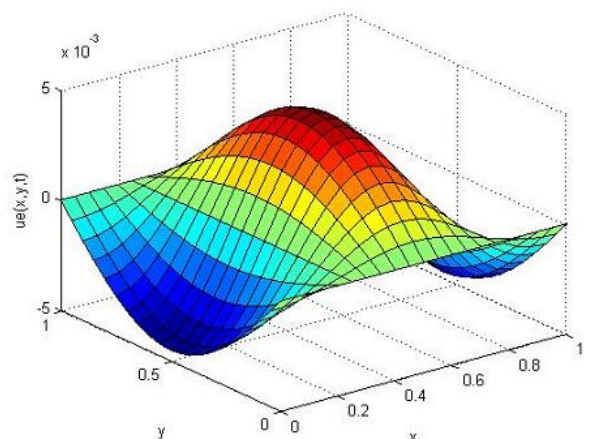
Fig.8. The exact value of u at 20×20 grids for $\mathrm{Re} = 1000$ and at time level $t = 1.0$ with $\Delta t = 0.001$.
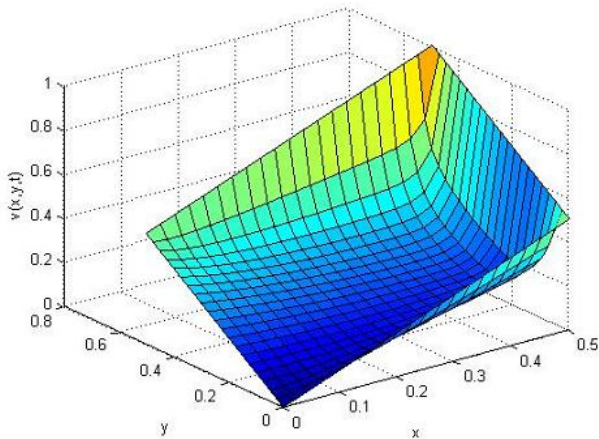
Fig.6. The computed value of v for $\mathrm{Re} = 50$ at time level $t = 0.625$.
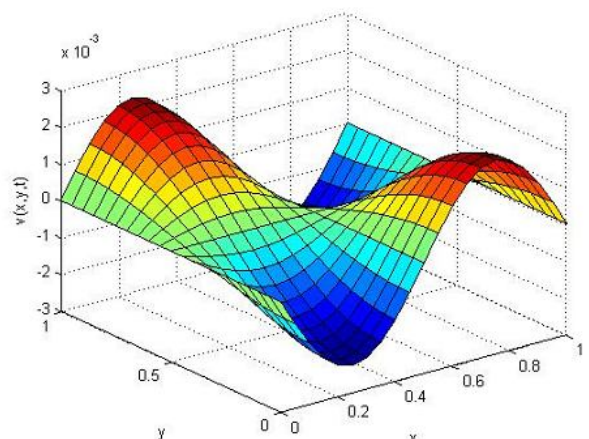
Fig.9. The numerical value of v at 20×20 grids for $\mathrm{Re} = 1000$ and at time level $t = 1.0$ with $\Delta t = 0.001$.
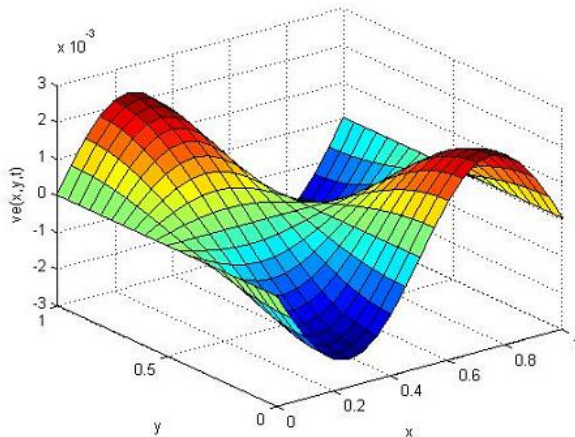
Fig.10. The exact value of v at 20×20 grids for $\mathrm{Re} = 1000$ and at time level $t = 1.0$ with $\Delta t = 0.001$.

4 CONCLUSION

A semi-implicit finite-difference method based on Ozis [7] has been presented for solving two-dimensional coupled nonlinear viscous Burgers' equations. The efficiency and numerical accuracy of the present scheme are validated through three numerical examples. Numerical results are compared well with those from the exact solutions and previous available results.

REFERENCES

[1] Cole JD, "On a quasilinear parabolic equations occurring in aerodynamics," Quart Appl Math 1951; 9:225-36.

[2] J.D. Logan, "An introduction to nonlinear partial differential equations, " Wily-Interscience, New York, 1994.

[3] L. Debtnath, "Nonlinear partial differential equations for scientist and engineers," Birkhauser, Boston, 1997.

[4] G. Adomian, "The diffusion-Brusselator equation," Comput. Math. Appl. 29(1995) 1-3.

[5] Bateman H., "some recent researches on the motion of fluids," Monthly Weather Review 1915; 43:163-170

[6] Burger JM, "A Mathematical Model Illustrating the Theory of Turbulence, "Advances in Applied mathematics 1950; 3:201-230

[7] T.Ozis, Y.Aslan, "The semi-approximate approach for solving Burgers' equation with high Reynolds number," Appl. Math. Comput. 163 (2005) 131-145.

[8] Hon YC, Mao XZ,"An efficient numerical scheme for Burgers-like equations,"Applied Mathematics and Computation 1998; 95:37-50.

[9] Aksan EN, Ozdes A, "A numerical solution of Burgers' equation,"Applied Mathematics and Computation 2004; 156:395-402.

[10] Mickens R, "Exact solutions to difference equation models of Burgers' equation,"Numerical Methods for Partial Differential Equations 1986; 2(2):123-129.

[11] Kutluay S, Bahadir AR, Ozdes A,"Numerical solution of one-dimensional Burgers' equation: explicit and exact-explicit finite difference methods," Journal of Computational and Applied Mathematics 1999; 103:251-261.

[12] Kutluay S, Rsen A,"A linearized numerical scheme for Burgers-like equations," Applied Mathematics and Computation 2004; 156:295-305.

[13] Wenyuan Liao," An implicit fourth-order compact finite difference scheme for one-dimensional Burgers' equation," Applied Mathematics and Computation 2008; 206:755-764.

[14] C.A.J. Fletcher, "Generating exact solutions of the two-dimensional Burgers' equation", Int. J. Numer. Meth. Fluids 3 (1983) 213–216.

[15] P.C. Jain, D.N. Holla, "Numerical solution of coupled Burgers_ equations, "Int. J. Numer. Meth.Eng. 12 (1978) 213–222.

[16] C.A.J. Fletcher," A comparison of finite element and finite difference of the one- and two-dimensional Burgers' equations," J. Comput. Phys., Vol. 51, (1983), 159-188.

[17] F.W. Wubs, E.D. de Goede, "An explicit–implicit method for a class of time-dependent partial differential equations, "Appl. Numer. Math. 9 (1992) 157–181.

[18] O. Goyon," Multilevel schemes for solving unsteady equations," Int. J. Numer. Meth. Fluids 22(1996) 937–959.

[19] Bahadir AR. "A fully implicit finite-difference scheme for two-dimensional Burgers' equation," Appl Math Comput 2003; 137:131-7.

[20] Vineet Kumar Srivastava, Mohammad Tamsir, Utkarsh Bhardwaj, YVSS Sanyasiraju, "Crank-Nicolson scheme for numerical solutions of two dimensional coupled Burgers' equations," International Journal of Scientific & Engineering Research Volume 2, Issue 5, May-2011.

Audio Streaming on Mobile Phones

Ajinkya Patil, Apurva Mayekar, Shruti Gurye, Varun Karandikar, Pramila Chavan

Abstract— The development of digital communication with increased bandwidths and support of various protocols that allow streaming of multimedia data on mobile devices have created the need to manage and access the multimedia content from mobile devices. A few mobile applications have been developed which readily enable streaming of multimedia files on mobile phones. There are several issues that need to be handled in the process. The paper talks about the critical issues and presents essential concepts with respect to application development for streaming content to a mobile device. It provides a comparison between protocols and audio formats. A system for streaming audio has been proposed exploring various aspects of streaming.

Index Terms— Audio Formats, Interleaving, Mobile, Multimedia, Progressive Streaming, Protocols, Streaming Server.

———————————————— ◆ ————————————————

1 INTRODUCTION

Today, the mobile devices do not depend only on network service providers for wireless network access. With support for secondary wireless channels such as IEEE 802.11 or Bluetooth, the devices are able to exchange data with a larger bandwidth and speed. The availability of the secondary channels reduces the cost the user may have to pay compared to the cost when the device is connected to a network over 3G or GPRS. Not only that but the processing power, the battery life and the memory capacities have been improved to a great extent, making the mobile device and PDAs capable of handling continuous streams.

Understanding Downloading and Streaming

In general, audio content can be delivered over a network in two ways:

1. The audio file can be downloaded and then played from the local hard disk of the client.
2. The audio content can be streamed from the server to the client who decodes the received packets in real time, displays the content immediately.

We will discuss the advantages and disadvantages of both the methods in brief.

Advantages of downloading:

- Downloading works with any data rate and allows any audio quality one wants to offer.
- The file is transmitted error-free thus no quality reduction occurs during transmission.
- The download results in permanent storage of the data. This can be reused as per the demand.

Disadvantages of downloading:

- Download times are extremely long if high audio quality is provided over a low-bandwidth network. For example, a five-minute-long music clip encoded with 128kbps may take more than 20 minutes to download over a typical private Internet access with an effective long-term bandwidth of 20kbps. To reduce download times, audio quality must be reduced.

- The user has to load the complete file before he can listen to any part of it. He cannot preview the file to decide whether he is interested in the content.
- There is no possibility to provide a "live" service as re-broadcasting a radio program into the network simultaneously with e.g. the terrestrial broadcast.

Advantages of streaming

- The user can listen to the content immediately after he has demanded for it.
- He can seek to any time frame and listen to that part of the audio content without waiting for the whole file to download.
- It is possible to provide ''live'' services by encoding the audio signal in real time and sending the resulting audio data stream immediately to the client.

Disadvantages of streaming:

- To stream real time audio data, the transmission line has to provide the full bandwidth of the stream during the whole transmission period. This imposes a limit on the bandwidth and hence the quality of the audio stream.
- The transmitted stream is very sensitive to network load which may cause lost or delayed data packets. This leads to drop-outs in the client audio output.
- To process the required flow control, additional server software (a "streaming server") is needed.

In particular, if the client on the network has limited memory, eg a mobile device, then the choosing one of the above two methods make a lot of difference with respect to cost and user experience. For such devices, due to their limited memory downloading the entire file may not be feasible. Streaming on the other hand, forms a better solution, which can be implemented by flushing the unused memory time to time.   It has to be considered that the main advantage of downloading, the possible high audio quality even over low bandwidth networks, is

compensated in practical use by the required high download times.

The drawbacks of streaming, however, can be reduced by dynamically adapting the amount of transmitted data to the actual capacity of the network connection. To do this, the audio data must either be stored in a format that allows dropping parts of the data resulting in a "graceful degradation" or they must be stored in different formats that the server can choose the format best suited for the bandwidth of a given connection.

As far as audio applications on mobile phones are concerned downloading is not a viable option because mobile phones do not have enough memory to download audio data and then play it. Therefore streaming is the only option for mobile phones.

2 COMPARISON OF PROTOCOLS

2.1 HTTP (Windows Media Services HTTP Streaming Protocol Extensions)

HTTP Streaming Protocol Extensions are Microsoft Media Server (MMS) streaming which is tunneled through HTTP. HTTP has following features:

HTTP supports TCP because it uses HTTP for all data transmission. It never loses packets. If the network is unreliable, your client will stop and start buffering all over again and fall behind. This means a video being streamed will suffer unwanted delays if it is streamed using HTTP. There will be no loss of frames but this losslessness inot intended in streaming.

The advantage is that error correction is already taken care of. So, there is no additional overhead. Also, the firewalls do not block HTTP. Again, the number of messages that are sent by HTTP are lesser than that of RTSP. E.g. HTTP streaming can be started with two GET requests whereas, RTSP will require at least six messages to do the same task.

RTSP requests are sent over the single TCP connection but HTTP requests require different TCP connections. This means more three way handshakes.

Modes:

Simple mode:

It requires two HTTP connections when the streaming is started. Again, whenever the client pauses, resumes or seeks, one more connection is needed. Load balancing in such a scenario is a problem as each of these connections might be sent to a different server. Streaming protocol is considered to be stateful. It must have a consistent connection to one server. This mode can cause erros while streaming when content is received through a load balancer.

Pipelined mode:

This mode uses the HTTP 1.1 pipelining feature. It send all requests on a single connection. This can reduce the problem of load balancer that occurs with simple mode but cannot totally eliminate it. But, most of the HTTP proxy servers do not support this kind of pipelining, in which case, the simple mode is used by the application that does not involve pipelining.

2.2 Real Time Streaming Protocol

RTSP has the following features:

Packets in RTSP have a choice between TCP and UDP. If packet loss is not an issue, UDP can be used for streaming as it has less overhead compared to TCP.

The encapsulation of Advanced Streaming Format (ASF) packets in RTP is proprietary. The description of the ASF file, called ASF encapsulated in SDP, is proprietary.

If an RTP packet sent over UDP is lost it can be retransmitted in WMS. Thus, the client will not wait for expired RTP packets.As a result, it will not fall behind in streaming if it loses some packets.

A forward error correction (FEC) scheme for RTP packets is also supported for WMS.

A firewall can block ports and protocols that are used by RTSP. This will result into failure of streaming if the firewall exists between the client and server. This is genetally found in home Internet gateways. In case of presence of a built-in RTSP NAT, streaming is ought to fail.

RTSP needs to send multiple requests before the streamed content can actually start playing. But, the client can aggregate together these requests and pipeline them over a single TCP connection.

2.3 HTTP v/s RTSP

If the content to be streamed is pre-encoded and end-to-end delay is not a concern HTTP protocol can be used.

If the client runs a desktop OS and the stream is a live broadcast, the end-to-end delay must be reduced. Here, RTSP should be preferred over HTTP.

3 PROGRESSIVE STREAMING

The main problem with the mobile phones is that it does not support progressive streaming. The player will wait until the entire audio/video is streamed and then only it starts playing. Normally mobile phones do not have enough memory to save audio/video therefore the audio/video that the mobile phone can stream and play depends on the memory of the phone.
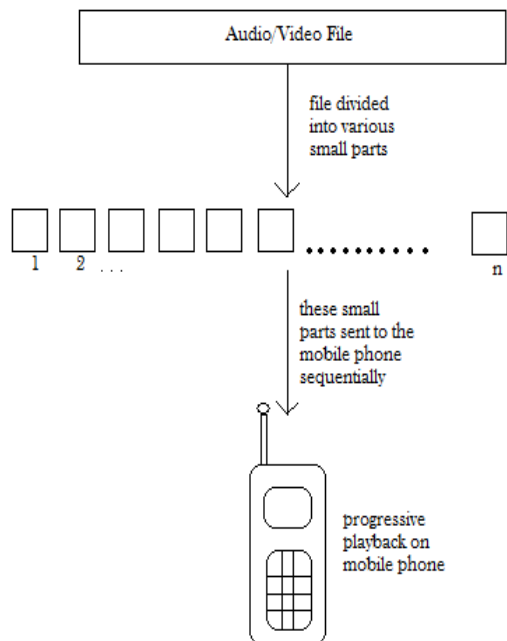
Overview of Progressive Playback

One way to work around it is to divide the audio/video into many small parts such that they can be easily saved and then played on the mobile phones. This will also allow progressive streaming by sending parts of the audio/video sequentially. For example consider an audio file which is divided into n parts. The server will send these parts sequentially to the mobile client. When the mobile phone receives the first part it will play it then when it receives the second part it will play only after the first part is done playing. Therefore the player will not wait for the entire audio/video to be streamed.

The size of each part can be decided depending on the mobile client, the memory available and the bandwidth of the network. The file may be divided into more number of parts, if serving to a client with low memory. The system can be designed such that to download the initial part, it will spend some time and start playing as soon as done. While it is playing the first part, the consecutive part is downloaded too and played at the right time.

4 AUDIO FORMATS

There are a number of audio formats available. The choice of audio format depends upon two factors: Quality of service and Bandwidth required for sending audio.

These two factors are inversely proportional. Thus, there is a trade off involved. For streaming over internet, to avoid unwanted delays, the bandwidth required should be kept minimum. But, reducing the bandwidth will lower the quality of audio. Thus, the aim is to obtain a good quality stream that will take minimum bandwidth.

AMR codec:

AMR, or Adaptive Multi-Rate, is an audio data compression scheme used in speech coding. This open standard format involves compressing of audio data to allow more storage on voice files. Originally developed for GSM (Global System for Mobile), a circuit-switched telecommunication system, it has now been adopted by most cellular companies all over the world.

AMR operates on narrow-band signals (200-3400 Hz) at eight different bit rates - 12.2, 10.2, 7.95, 7.40, 6.70, 5.90, 5.15 and 4.75 kb/s – which are based on frames containing 160 samples and are 20 milliseconds long. AMR is optimized for link adaptation, which enables it to select the best coding technique for better reproduction. The coding techniques that AMR uses are ACELP (Algebraic Code Excited Linear Prediction), DTX (Discontinuous Transmission), VAD (Voice Activity Detection), and CNG (Comfort Noise Generation). Hence the name Adaptive Multi-Rate.

As AMR offers eight different bit rates, it can help during network congestion improving the Quality of Service. It is known for its robustness to loss of packets, bit errors and is more immune to background noise

Advantages of AMR:

-AMR provides the best end-to-end solution to various applications and its flexible bit rate enables error correction

-AMR can be used in multi-media applications after being encapsulated in 3GP or MPEG-4 file formats

-AMR can easily be applied in low bit rate speech coding

-AMR has the most consistent performance across all major languages

5 ISSUES IN PROGRESSIVE STREAMING

Following are the issues that are encounter by using the progressive streaming algorithm that we suggested:
1. Loss of data during transmission
2. Size of the chunk in which the audio is cut
3. Static slicing of the audio vs. dynamic slicing of the audio

First we will discuss the issue of loss of data during transmission. Suppose the audio is cut into 10 parts and is transmitted and while transmission some of the parts are lost. Since this is progressive streaming we cannot request for the lost part because then we will have to wait for the audio part because that will take a lot of time and will create a gap when the audio is played. Now we need to compensate in some way for the last part or else there will be an abrupt gap when the audio is played. We can use interleaving along with an error concealment algorithm.
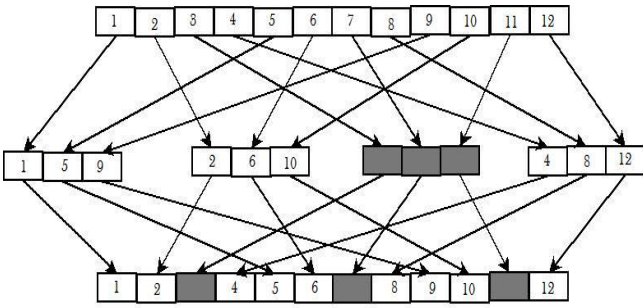
What happens in interleaving is that the audio is divided into very small chunks and then they are interleaved as follows:

Therefore if one of the parts is lost there will not be a long gap in the audio that is played. These small gaps can be compensated by using an error concealment algorithm.

Another issue of upmost importance is the size of the chunk in which the audio is cut. If the size is too small it will create a lot of network traffic by sending an individual part per packet. Alternatively if the chunk size is too large and if a part is lost during transmission then there will be a long abrupt gap in when the audio is played. Also the mobile phone may not have that much memory to temporarily save an individual part of the audio file. Therefore the size of the chunk should not be very small neither should it be very long.

The last topic of concern is whether to adapt static slicing of audio or dynamic slicing of audio. In static slicing the audio will be cut beforehand and then it will be saved in the database. Hence in static slicing when a request comes in the chunks will be retrieved from the database and then it will be streamed. Disadvantage of static slicing is that the time required to retrieve each part may exceed when the audio file is very large and is divided into many small chunks. Another disadvantage is that later if we want to change the chunk size then we will have to retrieve all the chunks from the database and then merge it and then again cut it according to the new size and save it in the database. This will be a very time consuming task if the database contains a large amount of audios saved in it. In dynamic slicing the audio file is saved as a whole unit. So when a request for an audio comes in the cutting of the audio is done dynamically.

6 MEDIA STREAMING PROTOCOL FOR REDUCING FRAME ERRORS

The following three methods are used in media streaming protocol to reduce frame errors:

1. Link-Layer Retransmission
In EGPRS there is an option of using Automatic Repeat

request (ARP). This ARP is used at the data link layer. Therefore it enables us to create a reliable pipeline for transmission of audio data over radio-link. The network has the capability of choosing the maximum number of retransmissions for each radio block. This number in many practical applications is set to 3.

2. Dynamic Packet Assignment
EGPRS wanted to efficiently support bursty data therefore it provides a facility of dynamic packet assignment (DPA). In DPA a link is assigned different channel dynamically which depends on the channel quality at the receiver. Initially each mobile is assigned to a channel. It continues to use the same channel till M consecutive blocks error occurs. Here M is a constant which is assigned to 2 in most of the cases. When M consecutive block errors occur the mobile is reassigned to another channel. DPA increases the capacity of the system therefore it offers better coverage.

3. Packet Shuffling
Shuffling or interleaving the frame and then transmitting increases the effectiveness of the error concealment algorithm. Shuffling and interleaving is done at the server side or it can be done at some other intermediate node. Reassembling of the frames takes place at the receiver and after reassembling the frames are placed in the play out buffer. This method helps in dispersing the effect of frame burst error. The frame number should be included in the frame so that we can hide the shuffling and interleaving from the mobile device. This hiding is required so that packet shuffling feature can be activated without changing the mobile. The shuffling technique is based on convolution interleaver which is defined by two interger parameters, N and B.

Let us take an example where N = 7 and B = 1. If T (k) denotes the position at which frame k is transmitted, then we have:
T (k) = k + (k mod N) NB
The order of transmission of the frames is 0, 8, 16, 24, 32, 40, 48, 7, 15, 23, 31, 39 etc .Using this technique two consecutive frames are separated by at least NB frames after the shuffling.

k	0	1	2	3	4	5	6	7
T(k)	0	8	16	24	32	40	48	7

Convolutional interleaving for N = 7, B = 1.

This technique is very effective in shortening gaps in the received frame sequence that result from burst trans-

mission errors. This technique also introduces a delay in the transmission of some frames. The choice of interleaver should provide a suitable trade-off between the frame error spreading and the reduced buffer length.

7 STREAMING SERVER

The process of streaming can also be achieved by the use of server side processing. Streaming servers such as Darwin Streaming Server, IceCast, Ampache process the audio data available on the server to create a stream of data.
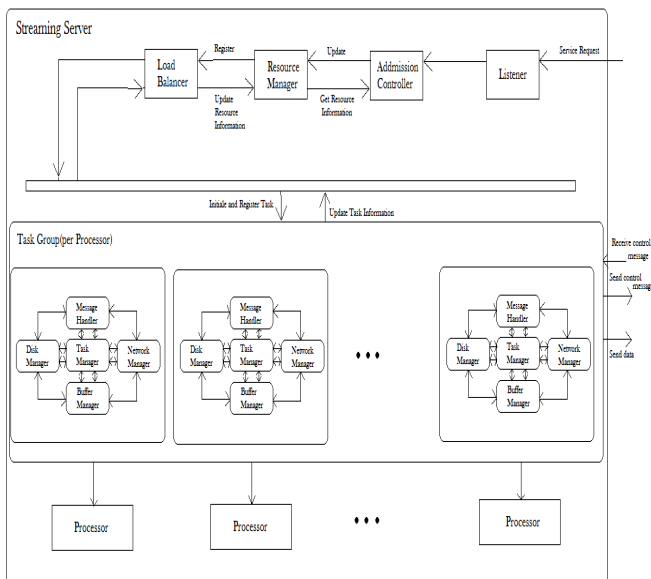
On request, this data can be made available to a mobile client via different protocols including rtsp and http. The client hardware needs to support rtsp protocol suite in order to get served by an RTSP server.

8 STREAMING SERVER ARCHITECTURE

A streaming server is basically structured for a multiple CPU platform that has the capability to provide support to many concurrent media streams.
The design of the server is based on:
1. Efficiency
2. Reliability
3. Scalability

Figure above illustrates the server structure. The main parts of the server are as follows:
1. Admission controller
2. Resource manager
3. Load balancer
4. Task pool
 There exists a task pool for each processor. Actual media streaming is carried out by the task which is allocated to the processor.

The main components of a task pool are as follows:
1. Disk manager
2. Network manager
3. Buffer manager
4. Message handler
5. Task manager

A service request by any user is handled as follows:
1. A listener basically listens to new service request. So when it receives a new service request it passes it on to the admission controller.
2. The job of the admission controller is to decide whether to service the request or reject it. The decision is based on the availability of system resources. A request will be serviced only when there is enough resources available to service it. The server has to base its decision on resource availability because it has to guarantee quality of service (QoS) to the existing user sessions.
3. The admission controller gets the status of the system resources from the resource manager. The resource manager maintains various data related to CPU usage, disk bandwidth, network bandwidth and memory size. It is the job of the resource manager to maintain latest information for various system resources.
4. If the service request is admitted resource usage is updated and the load balancer selects a processor to execute the newly admitted request. The main job of the load balancer is to intelligently balance the load between different processors.
5. The job of the task manager is to schedule the stream request so that it meets its timing constraints.
6. When the server gets a request from the client it transfers a fixed size media from the disk to the buffer and then analyses the stream and transfers the stream to the network depending on the media rate.
7. Each packet in the server is associated by a timestamp that denotes the time before which it must be transmitted to the network. The local timestamp which is there in each packet is added to the system which then it is added to the time-lined job queue where requests which are sorted according to the final timestamps are kept.
8. The job of the scheduler is to examine the queue constantly after a specific small time interval and to send the media packets whose timestamps are within the current scheduling period.

When the system capacity is exceeded by allowing excessive number of system requests then the quality of service may not be guaranteed. The admission control mechanism in the streaming server should be very strict so that it does not allow users in exceeding numbers which the server cannot handle. The decision should be based on the current resource availability at the server side. The admission decision should be such that it does not deteriorate the QoS requirements of the clients which are already being serviced by the server. The admission control mechanism does not only depend on the current system

resource availability but it also depends on the resource of the new request. The processor usage and disk access bandwidth vary dynamically depending on the way that the client behaves. It also depends on the underlying operating system. But the resource requirements of a new stream are known beforehand. Since the resource usage is dynamic and depends on the operating system and the user behaviors the resource usage table must also be dynamically updated. The updating of the resource tables is done by the server. The information from this table is used by the admission controller to admit new requests to the system.

The criteria that are used for the resource availability are:

1. Processor usage
2. Memory capacity
3. Disk bandwidth
4. Network bandwidth

4 CONCLUSION

Playing audio files on mobile device should be achieved through streaming as it will reduce the user waiting time involved in downloading. The effect of streaming can be achieved through progressive streaming that does not involve a streaming server, which can also be used alternatively. AMR turns out to be the most suitable audio format for streaming as it will require minimum bandwidth with a good audio quality compared to other audio formats. HTTP protocol can be preferred over RTSP as it incurs less overhead. In case a streaming server is used, various methods like link layer retransmission, dynamic packet assignment, packet shuffling etc. should be used.

References

[1] Anna Kyriakidou, N. K. (n.d.). Video Streaming for fast moving users in 3G mobile networks. *Univ. of Athens, Greece*, ACM.

[2]Christopher E.Hess, R. H. (n.d.). Media Streaming Protocol: An Adaptive Protocol For The Delivery Of Audio and Video Over the Internet. *Department of Computer Science, University of Illinois at Urbana-Champaign,Urbana, IL 61801*, IEEE.

[3]HOJUNG CHA, J. L. (2005). A Video Streaming System for Mobile Phones: Practice and Experience. *Department of Computer Science, Yonsei University*, ACM.

[4]Kapil Chawla, P. F. (n.d.). Transmission of Streaming Data over an EGPRS Wireless Network. *AT&T Laboratories - Research*, ACM.

[5]Karl Jonas, P. K. (n.d.). Audio Streaming on the Internet Experiences with Real-Time Streaming of Audio Streams. *German National Research Center for Information Technology*, ACM.

[6]Samjani, A. (2002). General Packet Radio Service, IEEE.

[7]Samjani, A. (2001). Mobile Internet Protocol, IEEE.

[8]Sumit Roy, M. C. (n.d.). A System Architecture for Managing Mobile Streaming Media Services. *Hewlett-Packard Laboratories*, IEEE.

[9]VAZQUEZ, M. (n.d.). A Mobile Audio Messages Streaming System. *National Institut of Telecommunications*, ACM.

[10]YE WANG, W. H. (N.D.). A FRAMEWORK FOR ROBUST AND SCALABLE AUDIO STREAMING. *SCHOOL OF COMPUTING, NATIONAL UNIVERSITY OF SINGAPORE*, IEEE

Performance Analysis of Backoff Algorithm in IEEE 802.11 Networks

Sakshi Suhane, Dr. Sanjeev Sharma, Prof. Varsha Sharma

Abstract— The primary Medium Access Control (MAC) technique of IEEE 802.11 is called Distributed Coordination Function (DCF). This protocol adopts a Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) with a binary exponential backoff (BEB) algorithm to access the channel. The protocol performance mainly depends on backoff procedure which reduces the probability of collision. With BEB, waiting time of a node gets doubled after every unsuccessful transmission. This introduces fast-growing retransmission delays for the backlog traffic. In a mobile ad hoc network (MANET), it would be worthwhile to slow down the growth-rate of waiting time because the nodes communicating in a MANET might move out of collision range while waiting for retransmission. Moreover, DCF reduces the Contention Window to the initial value after each successful transmission which essentially assumes that each successful transmission is an indication that the system is under low traffic loading. In this paper analyzed the problem with existing Backoff Algorithm.

Index Terms— DCF, MAC, CSMA/CA, BEB, MANET.

———————————————— ◆ ————————————————

1 INTRODUCTION

In recent years, much interest has been involved in the design of wireless networks for local area communication. Study group 802.11 was formed under IEEE Project 802 to recommend an international standard for Wireless Local Area Networks (WLAN's). The final version of the standard was released in 1999, and provides detailed medium access control (MAC) and physical layer (PHY) specification for WLAN's [1].

WLANs can operate in two modes namely infrastructure based and infrastructure-less mode or ad-hoc mode. In infrastructure based mode, a central coordinator or an Access Point (AP) is needed for operation of network. The AP resolves issues related to channel access and transfer of information between stations. AP based networks are also called as a single-hop networks where the all the information from a source to destination is transferred via the AP. Stations cannot communicate with each other directly. In the other mode of operation, known as the Mobile Ad-hoc Network (MANET) nodes communicates directly with each other without any central coordinator. This requires that all nodes must act as packet forwards to relay packets between two stations that are outside the radio coverage of each other. This provides greater flexibility and robustness.

To transmit packets to a node outside its range, the network uses multi-hop store-and-forward routing.

WLANs have great potential for both military and commercial applications. In a WLAN, nodes transmit packets in an unsynchronized fashion. The protocol employed in the medium access control (MAC) layer is responsible for coordinating access to the shared channel while minimizing conflicts. Hence it is important to design an efficient and effective MAC protocol.

In the 802.11 protocol, the fundamental mechanism to access the medium is called distributed coordination function (DCF). This is a random access scheme, based on the carrier sense multiple access with collision avoidance (CSMA/CA) protocol. Retransmission of collided packets is managed according to binary exponential backoff rules. The standard also defines an optional point coordination function (PCF), which is a centralized MAC protocol able to support collision free and time bounded services.

2. 802.11 Medium Access Control Layer

The MAC layer has to fulfill several tasks. It has to control medium access but it can also offer support for roaming, authentication, and power conservation. The basic services provided by the MAC layer are the mandatory asynchronous data service and an optional time bounded service. The following three basic access mechanisms have been defined for IEEE 802.11: the mandatory basic access method based on CSMA/CA, an optional method avoiding hidden terminal problem, and finally a contention-free polling method for time bounded services. The first two methods are also summarized as Distributed Coordination Function (DCF), the third method is called Point Coordination Function (PCF). The MAC mechanisms are also called as distributed foundation wireless medium access control (DFWMAC).

————————————————

- *Sakshi Suhane is currently pursuing masters degree (M.Tech-IT) in Ad-hoc Networking in SOIT RGPV ,BHOPAL(MP),INDIA, PH-9424958241. E-mail: sakshi.suhane@gmail.com,sakshi_srit@yahoo.com*
- *Dr. Sanjeev Sharma is currently working as HOD (SOIT) in SOIT RGPV,BHOPAL(MP),INDIA. E-mail: **sanjeev@rgtu.net**.*
- *Prof. Varsha Sharma is currently working as Professor in SOIT RGPV,BHOPAL(MP),INDIA. E-mail: **varshasharma@rgtu.net***

Following figure depicts the architecture of 802.11 MAC layer –
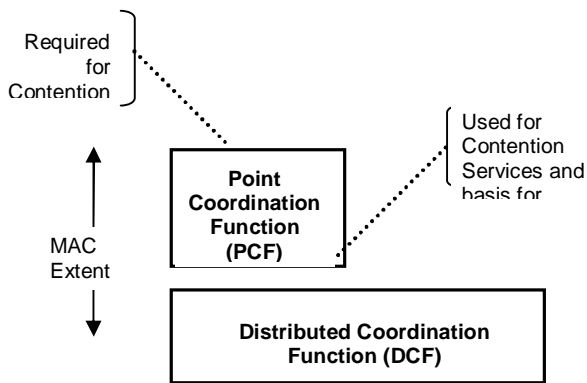
Figure 1: MAC Architecture

3. Distributed Coordination Function

The fundamental access method of the IEEE 802.11 MAC is a DCF known as carrier sense multiple access with collision avoidance (CSMA/CA). The DCF must be implemented in all stations, for use within both ad-hoc and infrastructure network configurations. For a station to transmit, it shall sense the medium to determine if another station is transmitting. If the medium is not determined to be busy, the transmission may proceed. The CSMA/CA distributed algorithm mandates that a gap of a minimum specified duration exist between contiguous frame sequences. A transmitting station must ensure that the medium is idle for this required duration before attempting to transmit. If the medium is determined to be busy, the station shall defer until the end of the current transmission. After deferral, or prior to attempting to transmit again immediately after a successful transmission, the station shall select a random backoff interval and should decrement the backoff interval counter while the medium is idle.

4. Operation mode of Conventional DCF

In 802.11, the DCF is the fundamental access method used to support asynchronous data transfer on a best effort basis. As specified in the standards, the DCF must be supported by all the stations in a basic service set(BSS). The DCF is based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). CSMA/CD is not used because a station is unable to listen to the channel while transmitting. In 802.11 CS is performed both at physical layer, which is also referred to as physical carrier sensing, and at the MAC layer, which is known as virtual carrier sensing. The PCF in the 802.11 is a polling-based protocol, which is designed to support collision free and real time services.

There are two techniques used for packet transmitting in DCF. The default one is a two-way handshaking mechanism, also known as basic access method. A positive MAC acknowledgement (ACK) is transmitted by the destination station to signal the successful packet transmission. The other optional one is a four-way handshaking mechanism, which uses request-to-send/clear-to-send(RTS/CTS) technique to reserve the channel before data transmission. Before transmitting a packet, a station operating in RTS/CTS mode "reserves" the channel by sending a special Request-To-Send short frame. The destination station acknowledges the receipt of an RTS frame by sending back a Clear-To-Send frame, after which normal packet transmission and ACK response occurs. Since collision may occur only on the RTS frame, and it is detected by the lack of CTS response, the RTS/CTS mechanism allows to increase the system performance by reducing the duration of a collision when long messages are transmitted. As an important side effect, the RTS/CTS scheme designed in the 802.11 protocol is suited to combat the so-called problem of Hidden Terminals, which occurs when pairs of mobile stations result to be unable to hear each other. However, the drawback of RTS/CTS mechanism is increased overhead for short data frames.

A station with a new packet to transmit monitors the channel activity. If the channel is idle for a period of time equal to a distributed inter-frame space (DIFS), the station transmits. Otherwise, if the channel is sensed busy (either immediately or during the DIFS), the station persists to monitor the channel until it is measured idle for a DIFS. At this point, the station generates a random backoff interval before transmitting (this is the Collision Avoidance feature of the protocol), to minimize the probability of collision with packets being transmitted by other stations. In addition, to avoid channel capture, a station must wait a random backoff time between two consecutive new packet transmissions, even if the medium is sensed idle in the DIFS time.

For efficiency reasons, DCF employs a discrete-time backoff scale. The time immediately following an idle DIFS is slotted, and a station is allowed to transmit only at the beginning of each slot time. The slot time size is set equal to the time needed at any station to detect the transmission of a packet from any other station.It depends on the physical layer, and it accounts for the propagation delay, for the time needed to switch from the receiving to the transmitting state, and for the time to signal to the MAC layer the state of the channel (busy detect time).

5. Backoff procedure

The backoff procedure shall be invoked for a STA to transfer a frame when finding the medium busy as indicated by either the physical or virtual carrier-sense mechanism (see Figure). The backoff procedure shall also be invoked when a transmitting STA infers a failed transmission.To begin the backoff procedure, the STA

shall set its Backoff Timer to a random backoff time. All backoff slots occur following a DIFS period during which the medium is determined to be idle for the duration of the DIFS period, or following an EIFS period during which the medium is determined to be idle for the duration of the EIFS period following detection of a frame that was not received correctly. A STA performing the backoff procedure shall use the carrier-sense mechanism to determine whether there is activity during each backoff slot. If no medium activity is indicated for the duration of a particular backoff slot, then the backoff procedure shall decrement its backoff time by slotTime.If the medium is determined to be busy at any time during a backoff slot, then the backoff procedure is suspended; that is, the backoff timer shall not decrement for that slot. The medium shall be determined to be idle for the duration of a DIFS period or EIFS, before the backoff procedure is allowed to resume. Transmission shall ommence whenever the Backoff Timer reaches zero.A backoff procedure shall be performed immediately after the end of every transmission.In the case of successful acknowledged transmissions, this backoff procedure shall begin at the end of the received ACK frame. In the case of unsuccessful transmissions requiring acknowledgment, this backoff procedure shall begin at the end of the ACK timeout interval. If the transmission is successful, the CW value reverts to aCWmin before the random backoff interval is chosen, and the STA short retry count and/or STA long retry count are updated . This assures that transmitted frames from a STA are always separated by at least one backoff interval. The effect of this procedure is that when multiple STAs are deferring and go into

random backoff, then the STA selecting the smallest backoff time using the random function will win the contention.

6. Problems with existing Backoff Algorithm

From the above discussion we can see that DCF resolves collision through Contention Window and backoff time. As specified in the original standard, after each successful transmission, the backoff stage will resume to the initial stage 0, and the contention window will be set to CWmin regardless of network conditions such as number of competing nodes. This method, referred to as 'heavy decrease' tends to work well when the number of competing nodes is less. When the number of competing nodes increases, it will be shown to be ineffective, since the new collisions can potentially occur and cause significant performance degradation.

For example, let us assume that the current backoff stage is 'i' with contention window $CW(i) = 2^i *$ CWmin , and there is a successful transmission, the next backoff stage will be stage 0 with contention window $CW(0) = 31$ according to the specification. But if the number of competing nodes is large enough (>>31), the new collision will likely occur at the backoff stage 0. The main argument is that since the current backoff stage is 'i' some collision must have occurred recently at the previous stage. Now if the number of current competing nodes is larger than or close to CW(i), and if the backoff stage is set to 0, there is a high probability that new collisions will happen. So resetting the contention
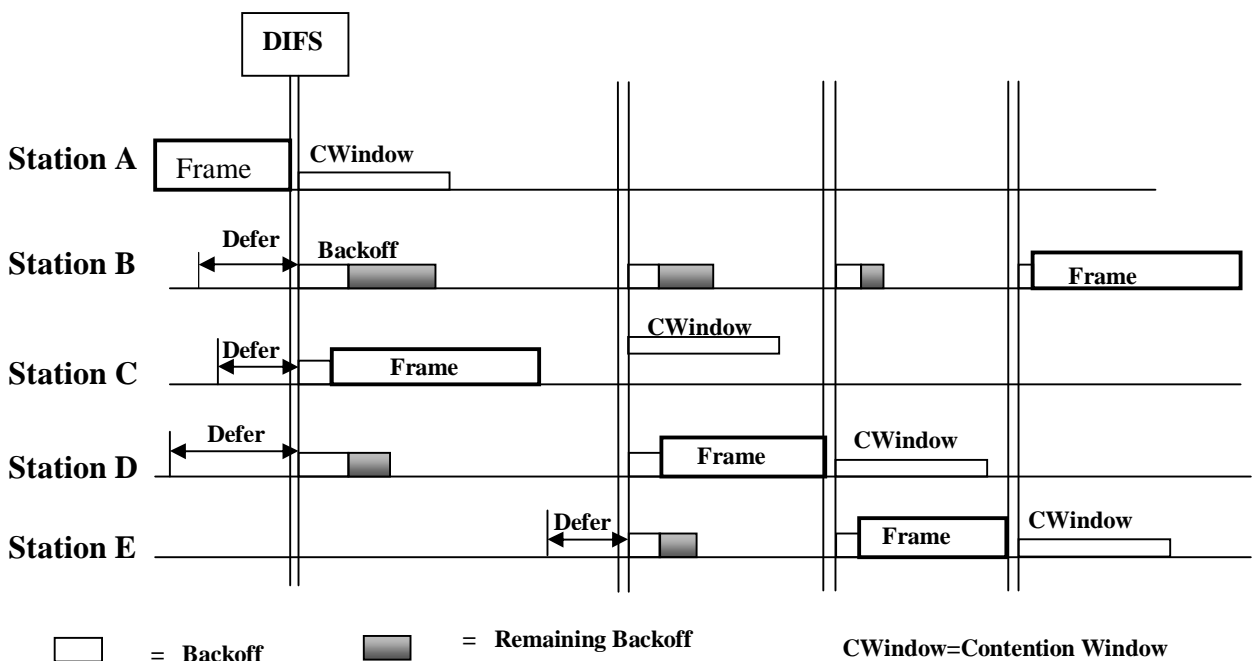
Figure 2: Backoff Procedure

window after every successful transmission is an inefficient approach if the number of nodes is large.
Thus, the operation of the BEB algorithm can be summarized as follows:

CW = min [2*CW, CWmax], upon collision (1)
CW = CWmin , upon success (2)

The operation of existing DCF protocol can be summarized from the following figure –
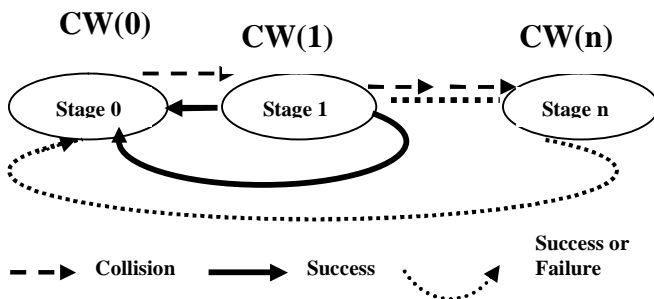
Figure 3. Operation of 802.11 DCF with BEB Algorithm

We also observe that The BEB algorithm causes a fast build-up (i.e., growth-rate) of waiting times spreading the backlog traffic over a larger time frame. However, this fast build-up of waiting time with increasing number of occurrence of collisions might not be appropriate for a MANET, wherein the contending nodes might leave the geographical location of contention itself after a short while due to their mobility. In view of this, we conjecture that it may not be necessary to make a node wait for a duration that builds up exponentially with a binary base.

REFERENCES

[1]   IEEE 802.11 standard, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE Std. June 1999.

[2]   C. Rama Krishna, Saswat Chakrabarti, and Debasish Dutta, "A modified backoff algorithm for IEEE 802.11 DCF based MAC protocol in a Mobile Ad-Hoc Network", TENCON 2004.

[3] Chonggang Wang, Weiwen Tang, Kazem Sohraby, Bo Li, "A simple mechanism on MAC layer to improve the performance of IEEE 802.11 DCF", Broadband Networks, 2004. BroadNets 2004. Proceedings of First International Conference on Broadband networks 2004

[4] Guanghong Wang, Yantai Shu, Liang Zhang, Oliver W.W. Yang,"Improving DCF in IEEE 802.11 Wireless LAN", IEEE CCECE 2003. CanadianConference on Electrical and Computer Engineering, 2003.

[5] Yunli Chen, Qing-An Zeng, Dharma P. Agrawal, " Performance analysis and enhancement for IEEE 802.11 MAC protocol", ICT 2003. 10th International Conference on Telecommunications, 2003, Volume 1, 23 Feb.- 1 March 2003 Page(s):860 – 867

[6] Manshaei M.H., Cantieni G.R., Barakat C., Turletti T, "Performance analysis of the IEEE 802.11 MAC and physical layer protocol", Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. 13-16 June 2005 Page(s):88 – 97

[7] Haitao Wu, Yong Peng, Keping Long, Shiduan Cheng, "A simple model of IEEE 802.11 wireless LAN", International Conferences on Info-tech and Info-net, 2001. Proceedings. ICII 2001 - Beijing. 2001 Volume 2, 29 Oct.-1 Nov. 2001 Page(s):514 – 519

A Mechanism for Booster Approach in Mobile Ad Hoc Network

Ahamd Anzar, Husain Shanawaz, Dr S.C Gupta

Abstract— In Mobile Adhoc network Nodes are highly mobile. They move around the Network. Due to this network topology and number of neighboring nodes in each node frequently change. Movement of nodes from one to another network also affect to the communication between them. As we know if nodes are within the range of each other they will work properly. But any of one node is not in the range of other node communication will Break. So we have to Design and develop that kind of system/approach which can handle such types of situation and prevent frequent link failure. To overcome this Link failure we introduced Booster Approach. In this Approach, we are dealing with the Power levels at both the ends (Tx as well as Rx) in order to make healthy communication between nodes in Adhoc Network. Suppose if our one node is moving away from the source node the moving node will measure its Rx power with respect to the distance and after a time duration when its current power level reaches the Threshold level (mini. Power level require for communication between nodes) it "ON its Booster" and at the same time it send one packet message MEP (Member Packet) to source node which contains (Received power level of moving node) due to this source node also "ON its Booster" enhance the Power level and trying to connect together to prevent the data lost during node mobility. In this paper we proposed a mechanism based on the Mathematical parameters . like Tx power, Rx power (Threshold level- minimum power require for communication between nodes), Distance between Tx and Rx, which are making an important role in Booster approach mechanism and Prevent link break during Nodes mobility at data transaction time.

Index Terms— MANET, Mechanism, Node, CHG,Booster, Tx/Rx Power.

———————————————— ◆ ————————————————

1   INTRODUCTION

Adhoc Network is self organizing, Multi-Hop based future technology, But in MANET we have some various challenges like Battery Back up, Node Mobility, Bandwidth constraint, Routing Protocols & Security problems. As we know MANET is used where no infrastructure is available for communication, such like disastrous area, military application. A mobile Transaction is structured as a Distributed transaction [2, 3]. In which the transaction is completed by the help of mobile nodes, providing different services. The mobile environment produces the significant challenges to transaction processing. The wireless network provides limited bandwidth so network bandwidth is a scarce resource. Battery power drains with data transmission and transaction processing. Due to the Dynamic mobility of nodes it affect to the communication link which should be maintain properly during mobility. Here we are trying to prevent data loss during node mobility. For this we develop a Booster Technique. Booster is a kind of Device and Algo's combination attached within the equipment as

————————————————————

- *Anzar Ahamd is currently pursuing Phd in Mobile Ad Hoc Network in Graphic Era University, India, PH-09634321293. E-mail: anz.hmd@gmail.com*
- *Shahnawaz Husain is currently pursuing Phd in Mobile AdHoc Network in Graphic Era University, India, PH-097585804269. E-mail: shahnawaz.husain@gmail.com*
-

laptop, mobile etc. and as per the requirement Booster operate automatically need not require manually. Before operate 1st it will check the device parameters which are

require for operating this like (Tx power, Rx power-Threshold level, Distance between Tx and Rx). Once Booster activated the mobile node can cover to each other without any difficulty. Primary applications of MANET are military tactical application, oceanography and the situation where a human being is unable to go and we have the condition to access the information about the climate and environment of that location. Various approaches have been given to manage a MANET. The Booster Approach is Newest and Latest in this direction which can be helpful for Mobile Ad hoc Network.

2 Related Work

In this model assuming that each node in MANET is booster enabled, a threshold value is used to activate and deactivate the booster. The threshold power level (T_{MIN}) value is related to strength of signal. All the nodes in the cluster send the performance table to the cluster head (CH). According to efficiency of the node CH maintains a Performance table. On the basis of efficiency the CH assign the task to the nodes in the cluster. If any node goes out from the cluster and receiving a low signal (equivalent to link failed), then outgoing node checks the value of signal strength, if signal strength is less than threshold power level (T_{MIN}) then booster becomes active, the booster amplifies the signal, the node which is somehow connected to outgoing node also activate its booster when it realize the equivalence of link failure.

    Here, The provided approach showing that the cluster head always try to tie up with the efficient node with-

in its cluster, through the use of signal booster technology.
. Using this technique we can increase strength to the various QoS models in MANET, & designing of low weighted and less battery consumption enabled boosters.

3 Proposed Model

In MANET we can not restrict over the speed of node i.e they are free to move any where. So provided approach showing that the cluster Head(CH) always to tie up with the efficient node within its cluster, through the use of signal booster Technology. So in the proposed model we developed a mechanism for booster RF approach..
In proposed mechanism we are taking three major points into consideration.. [7]

- Tx Power level.
- Rx Power level (Threshold level- minimum power level require for communication between nodes).
- Distance between Tx and Rx.
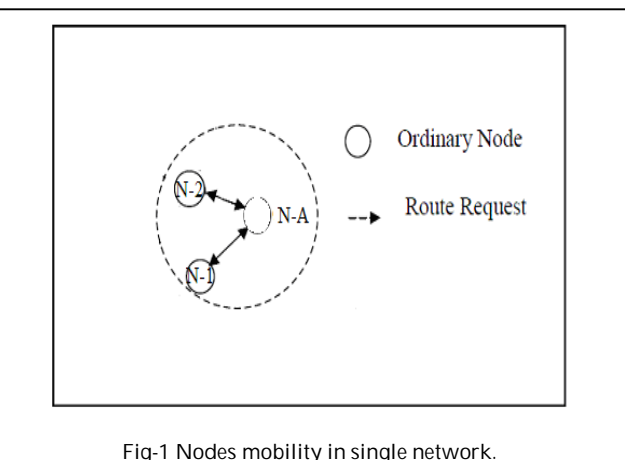
Fig-1 Nodes mobility in single network.

As we can see in the fig.1 we are focusing on Node N-A and node N-1. Here node N-A wants to communicate to node N-1 But both nodes are far from each other due to this communication link can be Break and the data can be lost. To prevent this we have to increase the Power (Strength) level of our signal. To increase the Power level we are introducing "Booster Mechanism".

In this Mechanism, we are dealing with the Power levels at both the ends (Tx as well as Rx) in order to make healthy communication between nodes in Adhoc N/w.

Before Booster activated. 1st Both of the node (N-A & N-1) should know the Rx power level of each other. Here we introduce MEP packets (Message packet). MEP packets are responsible for Exchanging of data including Rx power level between nodes in Adhoc network.

Once both the nodes know the Rx power level of each other they compare it with Threshold power level (mini. Power level require for communication between nodes) and if current Rx power level reaches the Threshold level they will activate their Booster respectively. Suppose in the given fig.1 node N-1 is moving far from

node N-A so as per the formula if distance increases Rx power level will decreases. After a certain time of period if node N-1 has Rx power level will reach the Threshold power level its Booster will activated automatically and at the same time node N-1 send MEP packets to node N-A. Informing its Rx power level to node N-A.

As node N-A get this MEP packet its Booster will be automatically activated and both the nodes try to reconnect in order to enhance the power level and make proper Communication.

Before doing all these things we have to consider some points-

1. Tx power of mobile.
2. Rx power of mobile. (i.e. Threshold Power level)
3. Distance between Tx and Rx.
4. Formula

$$Pr=Pt*Gt*Gr/(4r\pi/\lambda)^2$$

1 Step – Tx Power level—
In this step we are dealing with the Tx power of nodes in Adhoc network. As we know more Tx power covers more area and increase vicinity and the data loss chances become less.

TABLE 1

Transmission power level and its corresponding transmission range of a node

Level	Transmission Power (db)	Distance (m)
4	26.82	250
3	24.01	200
2	21.51	150
1	17.99	100

After analyzing these results we can say our network support up to 250 m. single hop communication (Maxi. Communication distance between one node to another node directly) without facing any difficulty and for 1 m distance we require approximate 0.1 db Tx power in Ad hoc network.

As per given formula we can say as distance increase Tx power level increase while Rx power level decrease.

2 & 3 Step- Rx Power level with respect to distance—
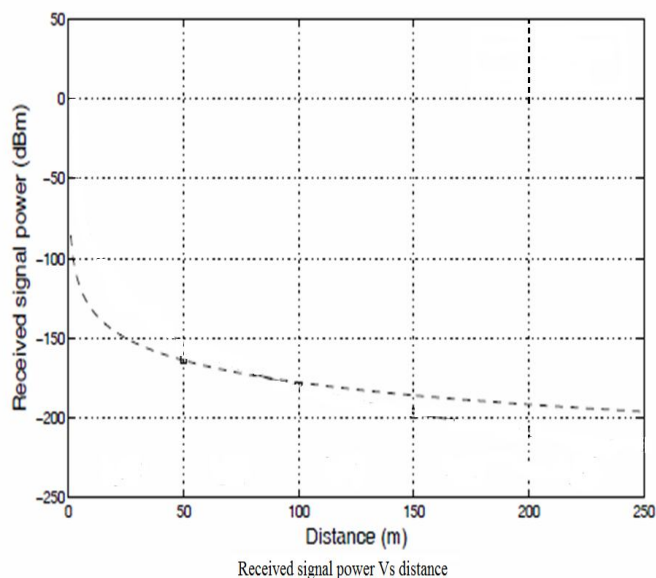
Figure-Received power Vs Distance

TABLE 2

Receive power level and its corresponding Receive range of a
node

Level	Receive Power (Threshold Level) (dbm)	Distance (m)
4	-199	250
3	-192	200
2	-185	150
1	-178	100

So, here as distance increases Rx power level decreases.
With the help of given curve Rx power level vs. distance
we can say our Threshold Rx power level is = -199 dbm.
At 250 m. (for ad-hoc wireless network).
Where Tx power = 26.82 db.
Rx power level (Maxi. Threshold level) approximate = -
199 dbm.
Distance between Tx and Rx = 250m.

4 CONCLUSION

Mobile Adhoc Network is the future technology; we have
provided the mechanism for Booster approach .By the
proposed mechanism we can save the efficient node in
the cluster which may be responsible to maintain com-
munication in our cluster.By measuring Tx power level,
Rx power level (Threshold level) and Maximum. Dis
tances between nodes. This paper has proposed the ap-
proach in which cluster head always try to tie up with the
efficient node within its cluster through the use of Booster
Technology. Using this this approach we can increase
network efficiency and also increase the QOS of network.

REFRENCES

1. Husain, Shahnawaz, Ahmad, Anzaar, Chand, Mukesh "A Fair Load Distribution Using Greedy Booster Approach in MANET " 3rd International Conference on Data Mining (ICDM 2010), joint-ly organized by University of Saskatchewan Canada, IMT Ga-ziabad, Nanyang Technological University, Singapore.11-12 March 2010, India.
2. W. Navidi and T. Camp. "Stationary distributions for the ran-dom waypoint mobility model", IEEE Transactions on Mobile Computing, pages 99–108, 2004.
3. Leslie D. Fife, Le Gruenwald , "Research Issues for Data Com-munication in Mobile Ad-Hoc Network Database Systems", SIGMOD Record, Vol. 32, No. 2, June 2003.
4. Dunham, M. H, Helal, A., & Balakrishnan, S. "A mobile transac-tion that captures both data and movement behavior", In ACM-Baltzer journal on Mobile Networks and Application, VOL. 2 (1997). pp 149-162.
5. Gruenwald, L., Javed, M., and Gu, M. "Energy- Efficient Data Broadcasting in Mobile Ad-Hoc Networks", In Proc. Internation-al Database Engineering and Applications Symposium (IDEAS '02), July, 2002.
6. G. Pei, M. Gerla, X. Hong, and C. Chiang. "A wireless hierar-chical routing protocol with group mobility", In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), pp 1538-1542, 1999.
7. Pitoura, E., & Bhargava, B., "Revising Transaction concept for mobile computing", In first IEEE workshop on mobile compu-ting systems and applications. (June 1995). pp 164-168.
8. Pitoura E.,and Bhargava B., "maintaining consistency of data in mobile distributed environments", 15th International Conference of distributed computing system (1996), pp 404-413.
9. Walborn G., & Chrysanthis P, "Supporting semantics based transaction processing in mobile database application", 14th IEEE symposium on reliable distributed system (1995), pp 31-40.
10. Yeo L. & Zaslavksy A. "submission of transaction from worksta-tion in a cooperative multidatabase processing environment", 14th ICDCS-1994.
11. P. Spentzouris, J. Amundson, "FNAL Booster Experiment and Modeling", proc. of the 2003 particle accelerator conference, IEEE computer society.
12. CHRYSANTHIS, P. Transaction Processing in Mobile Compu-ting Environments in IEEE Workshop on Advances in Parallel and Distributed Systems (1993).
13. Mohan, C., Harderle, D., Lindsat, B., Pirahesh, H., Schwarz, P. "Aries: A Transaction Recovery Method supporting fine granu-larity locking and partial rollback using write ahead logging", In ACM Transactions on Database Systems, VOL. 17 No. 1 (March 1992). pp 94-162.
14. Arup Acharya, B. R. Badrinath, and T. Imielinski. Checkpointing Distributed Applications on Mobile Computing. In Proceedings of the Third International Conference on Parallel and Distributed Information Systems, September 1994.
15. Evaggelia Pitoura and Bharat Bhargava. Maintaining Consisten-cy of Data in Mobile Distributed Environments. Technical Re-port TR-94-025, Purdue University, Dept. of Comp. Sciences,1994.
16. M.Q.Rafiq, M.Chauhan, S.Kumar, S.Husain "Mathematical si-mulation of Cochannel Interference Ratio for the Omni Direc-tional Antenna in Mobile Computing" NCAICT, CSI Allahabad chapter, India,15 -16 March,2008.
17. NFS Wireless & Mobile Communications Workshop, Northern Virgenia,March 1997.
18. Dharma Prakash Agrawal and Qing- An Zeng," Introduction to Wireless and Mobile System" University of Cincinnati.
19. P.Krishna, M.Chatterjee, N.H Vaidya,& D.K Pardhan," A Clus-ter based Approach for routing in adhoc networks", Proceeding

of 2nd USENIX Symposium on Mobile and Location Independent Computing.P.1 1995

20. B.Das,V Bhargavan,"Routing in adhoc network using a spine", IEEE International Conference on computer & communication network (ICC'97),1997.

21. Vincent D. Park and M. Scott Corson, "A performance comparison of the Temporally-Ordered Routing Algorithm and Ideal Link-state routing". In *Proceedings of IEEE Symposium on Computers and Communication '98* , June 1998.

22. David B. Johnson and David A.Maltz, "Dynamic source routing in adhoc wireless networks". In *MobileComputing* , edited by Tomasz Imielinski and Hank Korth, chapter 5, pages 153-181. Kluwer Academic Publishers.

23. Charles E. Perkins, "Ad Hoc On Demand Distance Vector (AODV) Routing". Internet draft, draft-ietf-manet-aodv-01.txt, August 1998.

24. Charles E. Perkins, "Ad Hoc On Demand Distance Vector (AODV).

Efficient Algorithm for ECG Coding

Ms. Manjari Sharma, Dr. A. K. Wadhwani

Abstract— Electrocardiogram (ECG) data compression algorithm is needed to reduce the amount of data to be transmitted, stored and analyzed, without losing the clinical information content. This work investigates a set of ECG data compression schemes in frequency domain to compare their performances in compressing ECG signals. These schemes are based on transform methods such as discrete cosine transform (DCT), fast fourier transform (FFT), discrete sine transform (DST), and their improvements. An improvement of a discrete cosine transform (DCT)-based method for electrocardiogram (ECG) compression is also presented as DCT-II. A comparative study of performance of different transforms is made in terms of Compression Ratio (CR) and Percent root mean square difference (PRD).The appropriate use of a block based DCT associated to a uniform scalar dead zone quantiser and arithmetic coding show very good results, confirming that the proposed strategy exhibits competitive performances compared with the most popular compressors used for ECG compression. Each specific transform is applied to a pre-selected data segment from the MIT-BIH database and then compression is performed.

Index Terms—Compression, Compression ratio, Cosine transform, ECG, Fourier transform, Frequency domain techniques, PRD, Time domain techniques.

———————————— ◆ ————————————

1 INTRODUCTION

CODING is useful because it helps reduce the consumption of expensive resources, such as hard disk space or transmission bandwidth. On the downside, compressed data must be decompressed to be used and this extra processing may be detrimental to some applications. Basically, a data coding algorithm seeks to minimize the number of code bits stored by reducing the redundancy present in the original signal. The design of data compression schemes therefore involves trade-offs among various factors including the degree of compression, the amount of distortion introduced (if using a lossy compression scheme) and the computational resources required to compress and uncompress the data. To deal with the huge amount of electrocardiogram (ECG) data for analysis, storage and transmission; an efficient ECG compression technique is needed to reduce the amount of data as much as possible while pre-serving the clinical significant signal for cardiac diagnosis, for analysis of ECG signal for various parameters such as heart rate, QRS-width, etc. Then the various parameters and the compressed signal can be transmitted with less channel capacity**.** Compression connotes the process of starting with a source of data in digital form (usually either a data stream or a stored file) and creating a representation that uses fewer bits than the original [1]. An effective data compression scheme for ECG signal is required in many practical applications such as ECG data storage, ambulatory recording systems and ECG data transmission over telephone line or digital telecommunication network for telemedicine. The main goal of any compression technique is to achieve maximum data volume reduction while preserving the significant features [2] and also detecting and eliminating redundancies in a given data set. Data compression methods can be classified into two categories: 1) lossless and 2) lossy coding methods. Lossy compression is useful where a certain amount of error is acceptable for increased compression performance. Loss less or information preserving compression is used in the storage of medical or legal records. In lossless data compression, the signal samples are considered to be realizations of a random variable or a random process and the entropy of the source signal determines the lowest compression ratio that can be achieved. In lossless coding the original signal can be perfectly reconstructed. For typical biomedical signals lossless (reversible) compression methods can only achieve Compression Ratios (CR) in the order of 2 to 1. On the other hand lossy (irreversible) techniques may produce CR results in the order of 10 to 1. In lossy methods, there is some kind of quantization of the input data which leads to higher CR results at the expense of reversibility. But this may be acceptable as long as no clinically significant degradation is introduced to the encoded signal. The CR levels of 2 to 1 are too low for most practical applications. Therefore, lossy coding methods which introduce small reconstruction errors are preferred in practice. In this paper we review the lossy biomedical data compression methods.

Biomedical signals can be compressed in time domain, frequency domain, or time-frequency domain. ECG data compression algorithms have been mainly classified into three major categories [3]: 1) Direct time-domain techniques, e.g., turning point (TP), amplitude-zone-time epoch coding (AZTEC) [4], coordinate reduction time encoding system (CORTES) and Fan algorithm. 2) Transformational approaches [3], e.g., discrete cosines transformation (DCT), fast fourier transform (FFT), discrete

————————————————

- *Manjari Sharma is currently pursuing masters degree program in Measurement and Control in Electrical Engineering Department, MITS, Gwalior, India, PH-08103780811. E-mail: manjari.sharma1@gmail.com*

- *Dr. A. K. Wadhwani is currently Professor in Electrical Engineering Department, MITS, Gwalior, India, PH-09425308846. E-mail: wadhwani_arun@rediffmail.com*

sine transform (DST), wavelet transform (WT) etc. 3) Parameter extraction techniques, e.g., Prediction and Vector Quantization (VQ) methods [2]. The time domain techniques which are based on direct methods were the earlier approaches to biomedical signal compression. Transform Coding (TC) is the most important frequency-domain digital waveform compression method.

When we compare these methods we find that direct data compression is a time domain compression algorithm which directly analyses samples where inter-beat and, intra-beat correlation is exploited. These algorithms suffer from sensitiveness to sampling rate, quantization levels and high frequency interference. It fails to achieve high data rate along with preservation of clinical information [5]. In Transform based technique [6] compressions are accomplished by applying an invertible orthogonal transform to the signal. Due to its de-correlation and energy compaction properties the transform based methods achieve better compression ratios [7]. In transform coding, knowledge of the application is used to choose information to discard, thereby lowering its bandwidth. The remaining information can then be compressed via a variety of methods. When the output is decoded, the result may not be identical to the original input, but is expected to be close enough for the purpose of the application. In parameter extraction methods a set of model parameters/features are extracted from the original signal (model based) which involves methods like Linear term prediction (LTP) and analysis by synthesis.

In this paper different transforms like DCT, FFT, DST and are studied for ECG signal compression. A threshold based algorithm is proposed to achieve better compression as DCT-II.

2 NECESSITY FOR ECG COMPRESSION

Compression of electrocardiography (ECG) is necessary for efficient storage and transmission of the digitized ECG signals. A typical ECG monitoring device generates a large amount of data in the continuous long-term (24-48 hours) ambulatory monitoring tasks. For good diagnostic quality, up to 12 different streams of data may be obtained from various sensors placed on the patient's body. The sampling rates of ECG signals are from 125Hz to 500Hz, and each data sample may be digitized into 8 to 12 bits binary number. Even with one sensor at the lowest sampling rate of 125 Hz and 8-bit encoding, it generates data at a rate of 7.5KB per minute and 450KB per hour. For a sampling rate of 500Hz and 12-bit encoding recording, it generates data at a rate of 540KB per minute and 30MB per hour. The data rate from 12 different sensors totally will generate 12 times amount of data and it is enormously big. Besides, recording over a period of time as long as 24 hours maybe needed for a patient with irregular heart rhythms . The monitor device such as Holter must have a memory capacity of about 400-800 MB for a 12-lead recording, but such a big memory cost may render a solid-state commercial Holter device impossible. Thus, efficient ECG data compression to dramatically reduce the data storage capacity is a necessary solution. On the other hand, it makes possible to transmit ECG data over a telephone line from one cardiac doctor to another cardiac doctor to get opinions.

3 EVALUATION CRITERIA

Before studying the various algorithms it is important to understand the criteria on the basis of which results are compared. The evaluation of performance for testing ECG compression algorithms includes three components: compression efficiency, reconstruction error and computational complexity. The compression efficiency is given by compression ratio (CR). The compression ratio and the reconstruction error are usually dependent on each other. The computational complexity component is part of the practical implementation consideration [8], [9]. In present work we have tested the data on the basis of Compression ratio CR and Percent root mean square difference PRD.

3.1 Compression Ratio

This is one of the most important parameters in data compression algorithms which specifies the amount of compression. A large value of this ratio shows success of any algorithm. Compression Ratio (CR) is defined as the ratio between the rates of the compressed signal (in terms of bits per second) to the rate of the original signal. A data compression algorithm must also represent the data with acceptable fidelity while achieving high CR, given by (1).

$$CR = \frac{Number\ of\ samples\ before\ compression}{Number\ of\ samples\ after\ compression} \qquad (1)$$

3.2 Distortion Measurement

One of the most difficult problems in ECG compression applications and reconstruction is defining the error criterion. The purpose of the compression system is to remove redundancy and irrelevant information. Consequently the error criterion has to be defined so that it will measure the ability of the reconstructed signal to preserve the relevant information. Since ECG signals generally are compressed with lossy compression algorithms, we have to have a way of quantifying the difference between the original and the reconstructed signal, often called distortion. Different objective error measures namely; root mean square error (RMSE), percentage root mean difference (PRD), signal to noise ratio (SNR) are used for calculation of reconstruction error. The most prominently used distortion measure is the Percent Root mean square Difference (PRD) [10] that is given by (2)

$$PRD = \sqrt{\frac{\sum_{n=1}^{L_b}[x(n)-x\prime(n)]^2}{\sum_{n=1}^{L_b}[x(n)]^2}} \qquad (2)$$

where $x(n)$ is the original signal, $x\prime(n)$ is the reconstructed signal and L_b is the length of the block or sequence over which PRD is calculated. PRD provides a numerical measure of the residual root mean square (rms) error.

4 THE FREQUENCY DOMAIN TECHNIQUES

In transform coding, knowledge of the application is used to choose information to discard, thereby lowering its bandwidth. The remaining information can then be compressed via a variety of methods. When the output is decoded, the result may not be identical to the original input, but is expected to be close enough for the purpose of the application. Transform Coding (TC) is the most important frequency-domain digital waveform compression method. The key is to divide the signal into frequency components and judiciously allocate bits in the frequency domain. In most TC methods, the input signal is first divided into blocks of data and each block is linearly transformed into the frequency domain. Transformation methods involve processing of the input signal by a linear orthogonal transformation and encoding of output using an appropriate error criterion. For signal reconstruction an inverse transformation is carried out and the signal is recovered with some error. Thus transformation techniques involve preprocessing the input signal by means of linear orthogonal transformation and properly encoding the transformed output (expansion coefficients) and reducing the amount of data needed to adequately represent the original signal. Various orthogonal transformations include Karhunen-Loeve Transform (KLT), Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT), Discrete Sine Transform (DST), and WAVELET transforms etc.

In this work we have compared the performance of four different frequency domain transformation methods for ECG compression and then their performance is evaluated. The various compression techniques have been discussed below:

4.1 Discrete Cosine Transform (DCT)

The Discrete Cosine Transform (DCT) was developed to approximate Karhunen-Loeve Transform (KLT) when there is high correlation among the input samples, which is the case in many digital waveforms including speech, music, and biomedical signals.

The DCT $\mathbf{v} = [v_0 \; v_1 \ldots \ldots \ldots v_{N-1}]^T$ of the vector $\mathbf{x}$ is defined as follows:

$$v_{0 = \frac{1}{\sqrt{N}}} \sum_{n=0}^{N-1} x_n \qquad (3)$$

$$v_k = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x_n \cos \frac{(2n+1)k\pi}{2N} \quad , \; k=1, 2,\ldots\ldots, (N\text{-}1) \quad (4)$$

where v_k is the *kth* DCT coefficient. The inverse discrete cosine transform (IDCT) of $\mathbf{v}$ is given by (5).

$$x_n = \frac{1}{\sqrt{N}} v_0 + \sqrt{\frac{2}{N}} \sum_{k=1}^{N-1} v_k \cos \frac{(2n+1)k\pi}{2N} \quad , \; n=0, 1, 2,\ldots\ldots(N\text{-}1)$$

$$(5)$$

There exist fast algorithms, Order (*N* log *N*), to compute the DCT .Thus, DCT can be implemented in a computationally efficient manner. Two recent image and video coding standards, JPEG and MPEG, use DCT as the main building block.

A discrete cosine transform (DCT) expresses a sequence of finitely many data points in terms of a sum of cosine functions oscillating at different frequencies [11]. DCTs are important to numerous applications in science and engineering, from lossy compression of audio (e.g. MP3) and images (e.g. JPEG) (where small high-frequency components can be discarded), to spectral methods for the numerical solution of partial differential equations. The use of cosine rather than sine functions is critical in these applications. For compression, it turns out that cosine functions are much more efficient whereas for differential equations the cosines express a particular choice of boundary conditions [12]. In particular, a DCT is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using only real numbers. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real and even function is real and even), where in some variants the input and/or output data are shifted by half a sample.

Discrete Cosine Transform is a basis for many signal and image compression algorithms due to its high decorrelation and energy compaction property [7]. A discrete Cosine Transform of *N* sample is defined as in (6):

$$F(u) = \sqrt{\frac{2}{N}} C(u) \sum_{x=0}^{N-1} f(x) \cos \left[\frac{\pi(2x+1)u}{2N}\right] \qquad (6)$$

$$u=0,1,\ldots\ldots,N\text{-}1,$$

where $C(u) = \frac{1}{\sqrt{2}}$ for $u = 0$

$= 1$ otherwise.

The function *f(x)* represents the value of *xth* samples of input signals [7]. *F(u)* represents a DCT coefficients. The inverse DCT is defined in similar fashion as in (7):

$$f(x) = \sqrt{\frac{2}{N}} \sum_{u=0}^{N-1} C(u) F(u) \cos \left[\frac{\pi(2x+1)u}{2N}\right] \qquad (7)$$

$$x=0, 1,\ldots N\text{-}1.$$

4.2 Fast Fourier Transform (FFT)

A fast Fourier transform (FFT) [13] is an efficient algorithm to compute the discrete fourier transform (DFT) and it's inverse [14]. There are many distinct FFT algorithms involving a wide range of mathematics, from simple complex-number arithmetic to group theory and number theory. A DFT decomposes a sequence of values into components of different frequencies but computing it directly from the definition is often too slow to be practical. An FFT is a way to compute the same result more quickly. Computing a DFT of N points in the naive way, using the definition, takes $O(N^2)$ arithmetical operations [15], while an FFT can compute the same result in only $O(N \log N)$ operations. The difference in speed can be substantial, especially for long data sets where N may be in the thousands or millions—in practice, the computation time can be reduced by several orders of magnitude in such cases, and the improvement is roughly proportional to $N / \log (N)$. This huge improvement made many DFT-based algorithms practical; FFTs are of great importance to a wide variety of applications, from digital signal processing and solving partial differential equations to algorithms for quick multiplication of large integers. The most well known FFT algorithms depend upon the factorization of N, but there are FFTs with $O(N \log N)$ complexity for all N, even for prime N. Many FFT algorithms only depend on the fact that is an Nth primitive root of unity, and thus can be applied to analogous transforms over any finite field, such as number-theoretic transforms.

Fast Fourier Transform is a fundamental transform in digital signal processing with applications in frequency analysis, signal processing etc [7]. The periodicity and symmetry properties of DFT are useful for compression. The uth FFT coefficient of length N sequence $\{f(x)\}$ is defined as in (8):

$$F(u) = \sum_{x=0}^{N-1} f(x) e^{-j2\pi ux/N} \qquad (8)$$

$$u=0,1,\dots.N-1.$$

And its inverse transform is calculated from (9):

$$f(x) = \frac{1}{N} \sum_{u=0}^{N-1} f(x) e^{j2\pi ux/N} \qquad (9)$$

$$x=0,1,\dots.N-1$$

4.3 Discrete Sine Transform (DST)

Discrete sine transform (DST) [16] is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using a purely real matrix. It is equivalent to the imaginary parts of a DFT of roughly twice the length, operating on real data with odd symmetry (since the Fourier transform of a real and odd function is imaginary and odd), where in some variants the input and/or output data are shifted by half a sample. Like any Fourier-related transform, discrete sine transforms (DSTs) express a function or a signal in terms of a sum of sinusoids with differ-

ent frequencies and amplitudes. Like the discrete Fourier transform (DFT), a DST operates on a function at a finite number of discrete data points. The obvious distinction between a DST and a DFT is that the former uses only sine functions, while the latter uses both cosines and sines (in the form of complex exponentials). However, this visible difference is merely a consequence of a deeper distinction: a DST implies different boundary conditions than the DFT or other related transforms [17].

Formally, the discrete sine transform is a linear, invertible function $F: R^N \rightarrow R^N$ (where R denotes the set of real numbers), or equivalently an $N \times N$ square matrix. There are several variants of the DST with slightly modified definitions. The N real numbers $x_0,\dots, x_{N-1}$ are transformed into the N real numbers $X_0, \dots, X_{N-1}$ according to (10):

$$X_k = \sum_{n=0}^{N-1} x_n \sin\left[\frac{\pi}{N+1}(n+1)(k+1)\right] \qquad (10)$$

$$k=0,\dots\dots N-1$$

4.4 Discret Cosine Transform-II (DCT-II)

The most common variant of discrete cosine transform is the type-II DCT [18]. The DCT-II is typically defined as a real, orthogonal (unitary), linear transformation by the formula in (11):

$$C_k^{II} = \sqrt{\frac{2-\delta_{k,0}}{N}} \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right], \qquad (11)$$

for N inputs x_n and N outputs C_k^{II}, where $\delta_{k,0}$ is the Kronecker delta (= 1 for k = 0 and = 0 otherwise).

DCT-II can be viewed as special case of the discrete fourier transform (DFT) with real inputs of certain symmetry [19]. This viewpoint is fruitful because it means that any FFT algorithm for the DFT leads immediately to a corresponding fast algorithm for the DCT-II simply by discarding the redundant operations.

The discrete Fourier transform of size N is defined by (12):

$$X_k = \sum_{n=0}^{N-1} x_n \qquad (12)$$

where $\omega_N = e^{-2\pi i/N}$ is an Nth primitive root of unity. In order to relate this to the DCT-II, it is convenient to choose a different normalization for the latter transform [19] as in (13):

$$C_k = 2 \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right] \qquad (13)$$

This normalization is not unitary, but it is more directly related to the DFT and therefore more convenient for the development of algorithms. Of course, any fast algorithm for Ck trivially yields a fast algorithm for C_k^{II} although the exact count of required multiplications depends on the normalization. In order to derive C_k from the DFT formula, we can use the identity $2\cos(\pi l/N) = \omega_{4N}^{2l} + \omega_{4N}^{4N-2l}$ to write (14):

$$C_k = 2\sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right]$$

$$= \sum_{n=0}^{N-1} x_n \,\omega_{4N}^{(2n+1)k} + \sum_{n=0}^{N-1} x_n \,\omega_{4N}^{(4N-2n-1)k}$$

$$= \sum_{n=0}^{4N-1} \widehat{x_n} \,\omega_{4N}^{nk} \qquad (14)$$

where $\widehat{x_n}$ is a real-even sequence of length $4N$, defined as follows for $0 \le n < N$:

$$\widehat{x_{2n}} = \widehat{x_{4N-(2n+1)}} = x_n \qquad (15)$$

Thus, the DCT-II of size N is precisely a DFT of size $4N$, of real-even inputs, where the even-indexed inputs are zero.

5 THE CODING ALGORITHMS

The various compression techniques DCT, FFT, DST algorithms are compared with PRD and Compression ratio CR and best suitable is considered. A threshold based algorithm is proposed to achieve better compression as DCT-II. The algorithms are performed on MIT-BIH database shown in Fig. 1.

The plot for compressed ECG by DCT is shown in Fig. 2.

Fig 2. The ECG signal after DCT coding
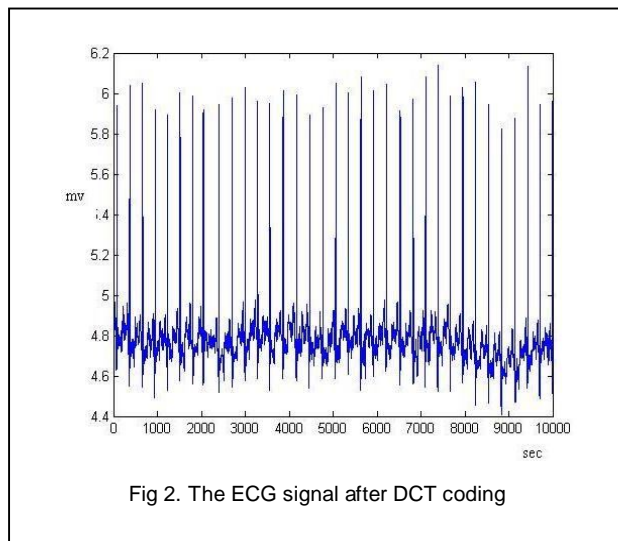
The error signal after DCT coding is shown in Fig. 3.

Fig.3. The Error signal after DCT coding

Fig. 1.The ECG signal before coding

5.1 The DCT Coding Algorithm

1. Separate the ECG components into three components x, y, z.
2. Find the frequency and time between two samples.
3. Find the DCT of ECG signal and check for DCT coefficients (before compression) =0, increment the counter A if it is between +0.22 to -0.22 and assign to Index=0.
4. Check for DCT coefficients (after compression) =0, increment the Counter B.
5. Calculate inverse DCT and plot decompression, error.
6. Calculate the compression ratio CR and PRD.

5.2 The FFT Coding Algorithm

1. Separate the ECG components into three components x, y, z.
2. Find the frequency and time between two samples.
3. Find the FFT of ECG signal and check for FFT coefficients (before compression) =0, increment the counter A if it is between +25 to-25 and assign to Index=0.
4. Check for FFT coefficients (after compression) =0, increment the Counter B.
5. Calculate inverse FFT and plot decompression, error.
6. Calculate the compression ratio CR and PRD.

The plot for compressed ECG by FFT is shown in Fig. 4.

The plot for compressed ECG by DST is shown in Fig. 6.
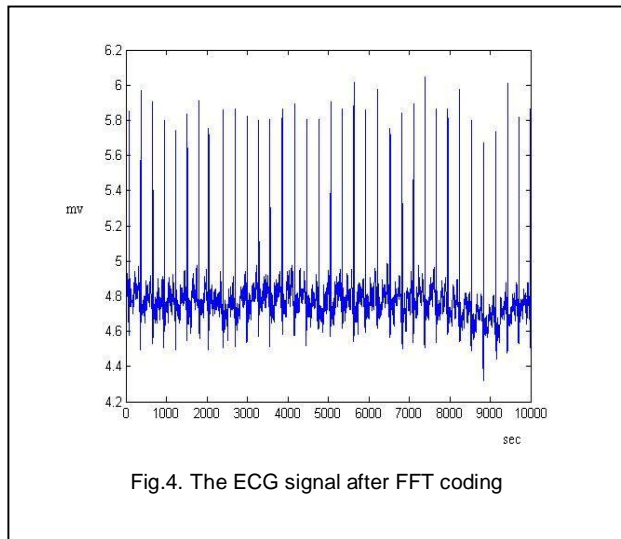
Fig.4. The ECG signal after FFT coding
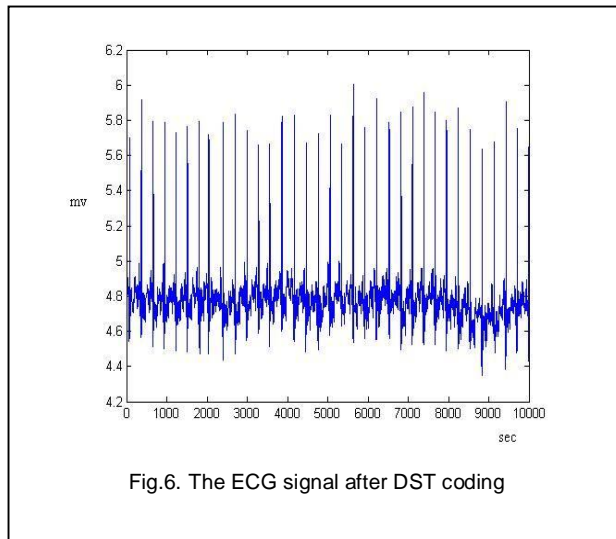
Fig.6. The ECG signal after DST coding

The error signal after FFT coding is shown in Fig. 5.
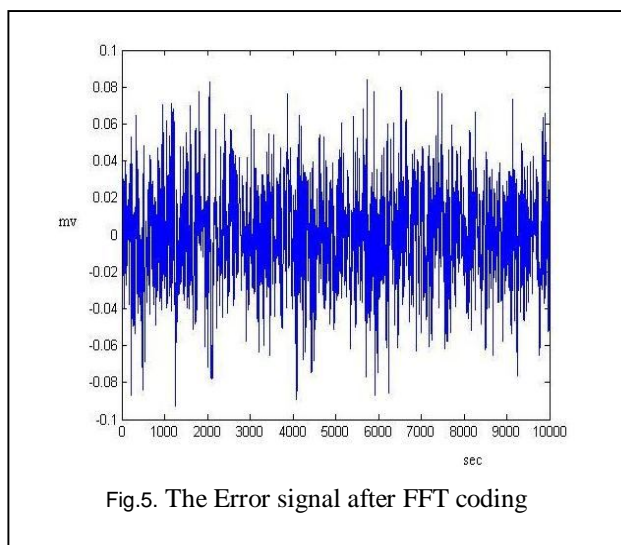
The error signal after DST is shown in Fig. 7.

Fig.5. The Error signal after FFT coding
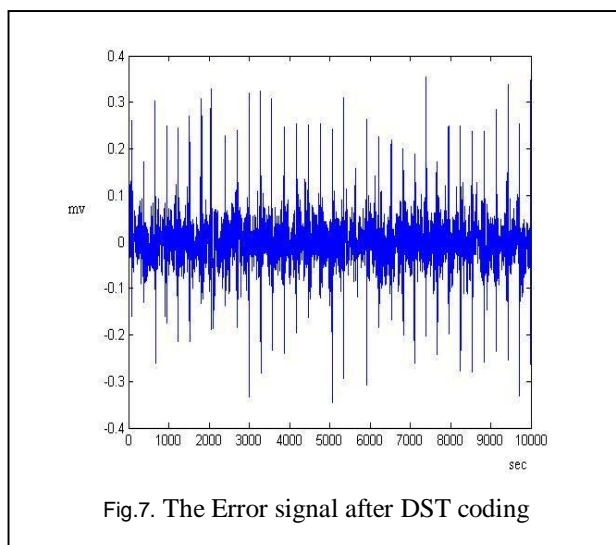
Fig.7. The Error signal after DST coding

5.3 The DST Coding Algorithm

1. Separate the ECG components into three components x, y, z.
2. Find the frequency and time between two samples.
3. Find the DST of ECG signal and check for DST coefficients (before compression) =0, increment the counter A if it is between +15 to-15 and assign to Index=0.
4. Check for DST coefficients (after compression) =0, increment the Counter B.
5. Calculate inverse DST and plot decompression, error.
6. Calculate the compression ratio CR and PRD.

5.4 The DCT-II Coding Algorithm

1. Partition of data sequence x in N_b consecutive blocks bi, i =0, 1,…….., N_b-1, each one with L_b samples.
2. DCT computation for each block.
3. Quantization of the DCT coefficients.
4. Lossless encoding of the quantized DCT coefficients.

Increasing the block size increases the CR and the DCT computing time. Various results show that increasing the block size above a certain point results in a very modest CR gain, while the processing time increases. The type II DCT is commonly used for data compression due to its greater capacity to concentrate the signal energy in few transform coefficients.

The algorithm is explained in detail as:

1. Let $b_i[n]$, n=0,1,...,L_b-1,represent the L_b values in block b_i.

2. The one-dimensional DCT-II of this block generates a transformed block B_i constituted by a sequence of L_b coefficients $B_i[m]$, m=0,1,.....L_b-1,given by (16):

$$B_i[m] = \sqrt{\left(\frac{2}{Lb}\right)} c_m \sum_{n=0}^{L_b-1} bi[n]cos\left[\frac{(2n+1)m\pi}{2L_b}\right] \quad (16)$$

$$m = 0,1,...L_b - 1$$

where c_m=1 for 1≤ m ≤ L_b-1 and c_o=(1/2)$^{(1/2)}$.

The DCT can be seen as a one-to-one mapping for N point vectors between the time and the frequency domains. The coefficient $B_i[0]$,which is directly related to the average value of the time-domain block is called DC coefficient and the remaining coefficients of a block are called AC coefficients.

Given B_i, b_i can be recovered by applying the inverse DCT-II:

$$b_i[n] = \sqrt{\left(\frac{2}{Lb}\right)} \sum_{m=0}^{L_b-1} c_m Bi[m]cos\left[\frac{(2n+1)m\pi}{2L_b}\right] \quad (17)$$

$$n=0,1,....L_b-1.$$

3. To quantize B_i we use quantization vector q. Each element $q[n]$, n=0,1,....L_b-1,of q is a positive integer in a specified interval and represents the quantization step size for the coefficient $B_i[n]$.The elements $\widehat{B_i}[n]$of the quantized DCT block $\widehat{B_i}$obtained by the following operation:

$$\widehat{B_i}[n] = B_i[n]//q[n] \quad , n=0,1,....L_b-1 \quad (18)$$

$$i=0,1,.....N_b-1$$

where // represents division followed by rounding to the nearest integer.

4. The lossless encoding of the quantized DCT coefficients involves run-length encoding, because the quantization normally generates many null values followed by an entropy encoder.

The plot for compressed ECG by DCT-II is shown in Fig. 8.
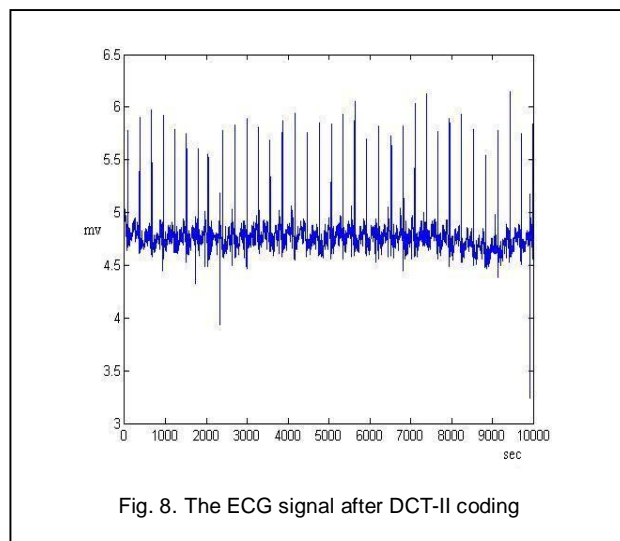
Fig. 8. The ECG signal after DCT-II coding

The error signal after DCT-II coding is shown in Fig. 9.
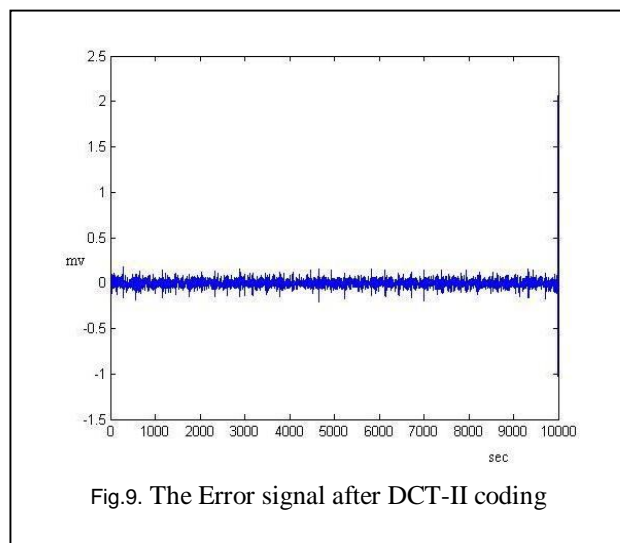
Fig.9. The Error signal after DCT-II coding

6 RESULTS AND DISCUSSION

Since signal recording conditions and noise levels vary from study to study, a thorough comparison of the coding methods is very difficult to make. But, frequency-domain coding methods produce higher coding results than time-domain coding methods. We used data in the MIT-BIH database to test the performance of the four coding techniques. The ECG data is sampled at 360Hz and the resolution of each sample is 11 bits/samples that total bit rate is 3960bps. The amount of compression is measured by CR and the distortion between the original and reconstructed signal is measured by PRD. The comparison table shown in Table 1, details the resultant compression techniques. This gives the choice to select the best suitable compression method. A data compression algorithm must represent the data with acceptable fidelity while achiev-

ing high CR. As the PRD indicates reconstruction fidelity; the increase in its value is actually undesirable. Although DCT-II provides maximum CR, but distortion is more. So a compromise is made between CR and PRD.

TABLE 1

COMPARISON OF COMPRESSION TECHNIQUES

Method	Compression Ratio	PRD
DCT	91.6800	0.8392
FFT	89.5723	1.0237
DST	70.4073	1.1967
DCT-II	94.28	1.5729

Considering that the number of electrocardiogram records annually numbers in the millions and the use of sending electrocardiogram records over telephone lines for remote analysis is increasing, the need for effective electrocardiogram compression techniques is great. Many existing compression algorithms have shown some success in electrocardiogram compression; however, algorithms that produce better compression ratios and less loss of data in the reconstructed data are needed. This paper has provided an overview of several compression techniques and has formulated new algorithms that should improve compression ratios and lessen error in the reconstructed data. This has allowed comparing various compression schemes and analyzing reconstructed electrocardiogram records through a graphic interface, without detailed knowledge of the mathematics behind the compression algorithm.

6 CONCLUSIONS

Transform based techniques because of their high compression ability have gained popularity. In this paper the preprocessed signal is transformed to get the decorrelated coefficients. The thresholding or quantization of transformed coefficients gives the actual compression, which is lossy one. But it has good performance and low computational cost. Among the four techniques presented, DST provides lowest CR and distortion is also high. FFT improves CR and lowers PRD. So FFT is better choice than DST. Next is DCT which gives higher CR upto 91.68 with PRD as 0.8392. But DCT-II provides an improvement in terms of CR of 94.28 but PRD increases up to 1.5729. Thus an improvement of a discrete cosine transform (DCT)-based method for electrocardiogram (ECG) compression

is presented as DCT-II. The appropriate use of a block based DCT-II associated to a uniform scalar dead zone quantiser and arithmetic coding show very good results, confirming that the proposed strategy exhibits competitive performances compared with the most popular compressors used for ECG compression.

REFERENCES

[1] Pranob K. Charles and Rajendra Prasad K. (2011): A Contemporary Approach For ECG Signal Compression Using Wavelet Transforms. Signal and Image Processing: An International Journal (SIPIJ). Vol. **2**, No. 1, 178-183.

[2] Yucel Kocyigit, Mehmet Korurek and Bekir Karlik (1999): ECG Data Compression by Artificial Neural Networks. In *ELEC0'99 International Conference On Electrical And Electronics Enginering.* E01.44/D-10, 338-339.

[3] S. Jalaleddine, C. Hutchens, R. Stratan, and W. A. Coberly (1990): ECG data compression techniques-A unified approach. *IEEE Trans. Biomed. Eng.,* **37,** 329-343.

[4] J. R. Cox, F. M. Nolle, H. A. Fozzard and G. C. Oliver (1968), AZTEC: A preprocessing scheme for real time ECG rhythm analysis. *IEEE Tran. Biomed. Eng.,* vol-BME-**15**, 128-129.

[5] M. Sabarimalai Manikandan and S. Danpat (2006): Wavelet Threshold based ECG Compression using USZZQ and Huffman Coding of DSM. In *Science Direct Biomedical Signal Processing and Control.* 261-270.

[6] B. R. S. Reddy and I. S. N. Murthy (1986): ECG data compression using Fourier descriptors, IEEE Trans. Bio-med. Eng., BME-**33**, 428-433.

[7] Mrs. S. O. Rajankar and Dr. S. N. Talbar (2010): An Optimized Transform for ECG Signal Compression. In *Proc. Of Int .Conf. on Advances in Computer Science,* 94-96.

[8] Shang-Gang Miaou, Heng-Lin Yen, Chih-Lung Lin (2002): Wavelet based ECG compression using Dynamic vector Quantization with Tree Code vectors in single codebook. In IEEE Transaction on Biomedical Engineering, vol. 49, no. 7, pp. 671-680.

[9] R.shanta selva Kumari, V Sadasivam (2007): A novel algorithm for wavelet based ECG signal coding. S*cience Direct Computers and Electrical Engineering,* 33, pp. 186-194.

[10] J. Abenstein and W. Tompkins (1982): A new data-reduction algorithm for real time ECG analysis. IEEE Tran. On Biomed. Engg., 29(BME-1):4, 3-8.

[11] N. Ahmed, T. Natarajan and K. R. Rao (1974): Discrete Cosine Transform. *IEEE Trans. Trans. On Computers.* C-23, 90-93.

[12] K. R. Rao and P. Yip (1990): Discrete cosine transform – algorithms, advantages, applications, San Diego: *Academic Press.*

[13] M. Clausen and U. Baum (1993): Fast Fourier Transforms. *BI-Wiss.-Verl.*

[14] L. Auslander, E. Feig and S. Winograd (1984): Abelian Semi-simple Algebras and Algorithms for the Discrete Fourier Transform. In *Advances in Applied Mathematics.* 5, 31-55.

[15] Tinku Acharya and Ajoy K. Roy. Image Processing Principles and Applications. *John Wiley.*

[16] S. Chan and K. Ho (1990): Direct Methods for computing discrete sinu-

soidal transforms. *IEEE Proceedings,* 137, 433-442.

[17] G. Steidl and M. Tasche (1991): A Polynomial approach to Fast algorithms for Discrete Fourier –cosine and Fourier-sine Transforms. In Mathematics in Computation, 56 (193), 281-296.

[18] E. Feig and S. Winograd (1992): Fast Algorithms for Discrete Cosine Trnsforms. *IEEE Tran. On Signal Processing.*vol-40(9), pp 2174-2193.

[19] Xuancheng Shao and Steven G. Johnson (May 10, 2007): Type-II/III DCT/DST algorithms with reduced number of arithmetic operations. *Preprint submitted to Elsevier.*

A Study of Finding Similarities in Web Service Using Metrics

D.Chandramohan, Shailesh Khapre, S.Ashokkumar

Abstract— Newest mounting and escalated recognition of services and its equivalent building of significant Web service mature to be a considerable rally. Web service expertise for permitting right to entry for promoting services regardless of locality and execution policies exist already. On the other hand, a huge discrepancy on structural, semantically and technological intensity all along with the emergent number of offered web services formulate their discovery a significant challenge. Our proposal progresses to classify an ultimate approach for identifying web service resemblance with the help of metrics. In particular, we analyzed the intrinsic worth of using metrics based Models, WordNet metrics and semantic similarity metrics for this evaluation purpose.

Index Terms—Web Service, metrics, Web Service Architecture, Simple Object Access Protocol (SOAP), (SOAP) based multicast protocol (SMP). (Web Service Manipulation Language) WSDL, Universal Description, Discovery and Integration (UDDI), XML, Web Ontology Language (OWL) ontology

———————————————— ◆ ————————————————

1 INTRODUCTION

Web service allows different application from different sources to communicate with each other without time consuming and custom coding. It is a standardized way of integrating web based applications. It allows organizations and users to communicate data without intimate knowledge of each other and share business logic, data and process through a programmatic interface across a network. It has been employed in a wide range of applications and has become a key technology in developing business operations on the web. In order to leverage on the use of web services, web service discovery and composition need to be fully supported. Several systems have been proposed to meet this need. The search and concepts lets to turn out the useful information retrieval and matching scenarios. Determination of conceptual overlap simplifies phrasing an adequate search concept. These results are ordered by degree of similarity to the searched and compared to the concepts. Delivering a flexible degree of conceptual overlap to a searched concepts gives the similarity measurements and it would additionally deliver the concepts of whose instances are located inside and adjacent to the search result. These concepts are close and identical to the user intended concepts. Identifying the attribute filler or relations for identifying the similarity in WSML (Web Service Manipulation Language) and the service integration scenarios the similarities between the concepts are based on

---

- *D.Chandramohan* and Shailesh Khapre are currently pursuing masters in technology program in Computer Science & engineering in Pondicherry Central University,Pondicherry, India, pdchandramohan@gmail.com , shaileshkhaprerkl@gmail.com*

- *S.Ashok kumar is currently pursuing masters in Engineering program in Computer Science and engineering in Veltech DR.RR & DR.SR Technical University,Chennai, India, ashokkumar782@gmail.com*

their specification and their chosen representation languages. This proposal leads to develop metrics for identifying the service similarities in web services that can help the clients / users to get serviced by

i) Efficiently finding the web service on web by ranking them according to their category with the help of metrics and proposing their degree of similarity
ii) Implicit and explicit data flow identification in services and computing the service to the best for the requesters.
iii) Eliminating fuzzy decision to identify the web service through improved fuzzy similarity algorithms.

We are proposing an approach enabling the implicit representation of similarities across distinct services which will help the requestors for identifying the exact service in which they are looking for. Similarity between words is becoming a generic problem for many applications of computational linguistics and it explores the determination of similarity by a number of information sources, which consists of structural information and information content from a huge quires.

2 RELATED WORK AND BACKGROUND

Services are provided by logical grouping of operations and its functions, if we consider the relevant service in business process as an activity with specific business goals and new products going to introduces in the near feature, current products to sales and ready for service and fulfilled order and business process consists of service and its operations which are executed in an ordered sequence according to a set of business rules. A service similarity matching algorithm is proposed to address the various degree of similarity in the qualitative matching

level. A multi-level matching frame work for modeling web services are achieved in five different levels syntactic, static semantic, dynamic semantic, qualitative services, dependable services and collaborative design for demonstrating these approaches. The appropriate measuring of distance of different services and their similarity relations between the sequences are based on the non computable notations and shows all its computable similarities globally and this computable similarity measure is similarity metrics as discussed [5], [7], [10], [12], & [15]. The UDDI registry by an approach for web service retrieval based on the evaluation of similarity between web services interfaces are defined with web service description language (WSDL). The higher the similarity and the less are the difference among their interfaces. A semantic oriented variant approach was proposed and discussed in [1], [8], [9], [11], & [13] as a language to annotate a WSDL description. Effective web service retrieval algorithm and to evaluate the similarity degree between two web services by comparing the related WSDL descriptions and the relation between the main elements composing a WSDL description by means of port Type, operation, message, and part Type. The automatic semantic categorization of web services are enabled by categorizing the pre-defined terms to calculate the sum of measure of semantic relatedness, nearest similarity score which are extracted from WSDL of a web service discussed in [2], [14], [17], & [19]. The measures of semantic relatedness are statistical methods for extracting word associations of text from WSDL. Vector based and probability based measure of semantic relatedness to find the point of mutual information and the normalized service distances are easily identifiable. A scalable similarity search for learned metrics and their pair wise similarity was measured using the Mahalanobis distance function enables efficient indexing with learned distances with a very large database as discussed in [3], [16], [18], & [21]. This approach to evaluate the impact on accuracy a learned metric has relative to both standard baseline metrics and state of the art methods and to test how reliably our semi-supervised hash functions preserve the learned metrics in practice when performing sub linear time database searches. Simple Object Access Protocol (SOAP) [1], [2], [4], [20] based multicast protocol (SMP). The SMP reduces the network traffic by aggregating syntactically similar SOAP messages to form a compact SMP message to the service requesters. The similarity of SOAP message is measured in pairs and both based on the message template and on the values of each XML tag in the message as discussed [20], [22], [24]. For measuring the similarity of semantic services and their compatibility by annotated with Web Ontology Language (OWL) ontology. This technique to OWL-S an emerging standard based on the inference as discussed [6] and this ontology based techniques and metrics similarities are under development by the research teams. The approach of UDDI registry for web service retrieval based on the evaluation of similarity between web services interfaces

are defined with the WSDL [1] and the clarification by semantically enriching the WSDL specifications in a semantic explanation. The method of enabling the scalable similarity search for learned metrics by finding the Mahalanobis distance function by construction of enabling the hashing construction indexing [23] in retrieved services.

The service categorization and discovery is based on the keywords [2] which are extracted from WSDL and measuring the semantic relatedness of each word with predefined category then finding the nearest similarity score. The similarity based SOAP multicast protocol to address the issue of latency service, the similarity of SOAP message is measured in pairs and is based on the message template [4] and on the value of each XML tag in the message. Finding valuable and attractive web services are becoming difficult due to massive number of web services, mining web service frame work in which the fuzzy logic, fuzzy set theory and fuzzy matching for those services into composite web services [5], [6] and finding the service similarity metrics.

3 OBJECTIVES AND RESEARCH ISSUES

This proposal states the web service retrieval based on the evaluation of similarity between the services in which the service descriptions are adopted from Web service Description Language (WSDL), Simple Object Access Protocol (SOAP), Universal Description, Discovery and Integration (UDDI) and XML schema.

i. To identify the similarity in services by finding the replica of port type, operations, message, functions, prototypes, in and out parameters, parts etc from SOAP, UDDI, WSDL, and XML schema registries.

ii. Reducing the complexity of evaluating the words and phrases extracted from WSDL by introducing an improved clustering algorithms.

iii. To identify the semantic relatedness of keyword to find the similarity, machine learning techniques can reduce the complexity of extracting from WSDL.

iv. To identify the ontology based service similarity to match the degree of service recommendation by hybrid match making algorithms and techniques in schemas.

v. To identify the web ontology language OWL similarity functions by targeting inference based matching with generic clustering algorithms.

vi. To evaluate the scalable similarity of Mahalanobis distance from very large database by means of introducing the advance hashing algorithm for implicit and explicit parameters existing in service.

vii. To explore a generic similarity service view in automated discovery of mediation by identifying the means of scalability.

viii. A statistical approach for service value decomposition to find the balance point accuracy and ranking of web services based on QOS attributes and cosine values.

ix. Exploiting a linguistic semantics search instead of plain key word based search will make the similarity selection of web service more appropriate.

x. Introducing a retrieval frame work by identify the similarity in service to improve the sophisticated similarity measures in XML service retrieval.

xi. To find the similarity in services an improved fuzzy similarity algorithm can improve the scalability among the services.

Heuristic and complex schema matching algorithms for both functional and non functional semantic distance for finding the web service similarity and the process model search.

4 PROPOSED APPROACH AND WORK PACKAGES

The proposed approach can be identified by means of extracting details from different services of UDDI registry, SOAP, WSDL, and XML schema applies the proposed model and it's easy to identifying the similarity metrics from any service. To realize the strategies mentioned in this proposed project, the following metrics are used in different phase to identify the similarity in web service.

Phase1

The different similarity functions from web services are identified on certain parameters like port type, operation name, input message, data type similarity, information loosed maximization function and similarity functions from the service.

$$\text{MaxSim}(f, Q, P) = \frac{1}{|Q|}.\max \sum_{i \in I}^{j \in J} f(qi, pj).xi, j, \quad (1)$$

$$\sum_{j \in J} x\,i, j \leq 1, \quad \forall i \in I,$$

$$\sum_{i \in I} xi, j \leq 1, \quad \forall j \in J \quad I = [1.. ||Q|], \; J = [1 ... |P|]$$

$$P(\sigma_q) = \frac{|\{\sigma i \in \Sigma_{\sigma q} | \sigma i \in R\sigma q\}|}{|\Sigma_{\sigma q}|}$$

$$R(\sigma_q) = \frac{|\{\sigma i \in \Sigma_{\sigma q} | \sigma i \in R\sigma q\}|}{|\Sigma_{\sigma q}|}$$

Where σq is the query, $\sum_{\sigma q}$ is the returned services after submitting the query, and Rσq is the relevant services for the given query. Precision and recall are measures for the entire result set without considering the ranking order. Thus, R-Precision and AP (Average Precision) parameters are also considered. Both of them depend on the precision at a given cutoff point (Pn). Thus, assuming $\sum_{\sigma q}$ as the set including the first n returned services

Phase2

Similarity in words in different services are denoted by lexical meaning between two words for example W_1 and W_2 are two words in WSDL document then the SimSet (Set1, Set2). The set denotes by S_1 and S_2 and their functional similarity value can be calculated as Sim (S_1, S2),

$$\text{Sim}(S_1, S_2) = \alpha * \text{SimSet}(S_1.T, S_2.T) +$$
$$B * \text{SimSet}(S_1.B, S_2.B) +$$
$$\Gamma * \text{SimSet}(S_1.A, S_2.A)$$

$$\text{Sim}(S1, S2) = \frac{\sum_{w \in S1} \text{Sim } m(w, S2) + \sum_{w \in S2} \text{Sim } m(w, S1)}{|S1| + |S2|} \quad (2)$$

S, T collection of words in service title information B denotes the service body information, A denotes the service additional information, and T, B and A represents the weights in similarity computation, α, β, γ are coefficients of different parts.

Phase3.

The Web Services are initially categorized under seven different categories Zip Code, Country Information Stock Market, Temperature, Weather, Fax and Currency. The extracted words of a particular Web Service are compared with each category from the WSDL of a web service (pressure, humidity, rainfall etc). The NSS (Normalized Similarity Score) of each word was calculated with category. We use the probability-based MSR – Normalized Similarity Score (NSS). NSS is an MSR that is derived from NGD. To be more precise, the relatedness between two words x and y is derived by NSS(x, y) =1-NGD(x, y)

Where NGD is a formula derived by cilibrasi and derived as

$$\text{NGD}(x, y) = \frac{\max\{\log f(x), \log f(y)\} \log f(x,y)}{\log M \min\{\log f(x), \log f(y)\}} \quad (3)$$

Phase4.

Consider two concepts that belong to a same ontology, the similarity between them are calculated by the following equation

$$\text{Sim}(C_1, C_2) = e^{-ch} \cdot \frac{e^{\beta h} - e^{-\beta h}}{e^{\beta h} + e^{-\beta h}} \quad (4)$$

Where l stands for the shortest path length between the two concepts in the ontology, and h stands for the depth of the closest common ancestor of the two concepts. The explanation of this measure is apparent. f_1 takes into account the influence of the path length. The similarity between the two concepts decreases at an exponential rate as the path length increases. f_2 reveals the influence of the concepts depth. Deeper a concept in the ontology, more concrete concept that means, a pair of deeper concepts will share more common information comparing with a pair of shallower concepts, when the path lengths of the two pairs are equal.

Phase5.

Distance metric for semantic nets and similarity for ontology framework and the semantic similarity was calculated by the similarity of OWL objects a and b is formally defined as

$$\text{Sim}(a, b) = \frac{f \text{ common }(a, b)}{f \text{ desc }(a, b)}$$

$$0 \leq \text{sim} \leq 1$$
$$\forall a : sim\,(a,a) = 1$$
$$\forall a, b : sim\,(a,b) = sim\,(b,a)$$
$$f\,common\,(a,b) = f\,desc\,(a,b)$$

We have the similarity metric by inference based information value, for any OWL object a and b

$$Sim\,(a,\,b) = \frac{\sum_{(x,y)\in com\,(a,b)}(IBIV\,(x) + IBIV\,(y))}{\sum_{s \in code\,sc\,m,n\,(a,b)}IBIV\,(s)} \qquad (5)$$

fcommon is the common function measuring the information value of the description that is shared between a, b. fdesc is the description function giving the value of the total information content of a and b. The similarity is defined as the ratio of the shared information between the objects to the total information about both objects. From the definition of the similarity function, sim, we can immediately obtain some properties that confirm some of the common intuition regarding similarity.

Phase6.

Metrics learning by similarity search and locality sensitive hashing in web service environments main idea of our approach is to learn a parameterization of a Mahalanobis metric based on provided labels or paired constraints while simultaneously encoding the learned information into randomized hash functions. These functions will guarantee that the more similar inputs are under the learned metric, the more likely they are to collide in a hash table. After constructing hash tables containing the entire database examples, those examples similar to a new instance are found in sub linear time in the size of the database by evaluating the learned metric between the new example and any example with which it shares a hash bucket. Parameterized Mahalanobis Metrics was proposed by means of

$$d_A\,(x_i,\,x_j) = (x_i\text{-}\,x_j)^t A\,(x_i\text{-}\,x_j)$$
$$S_A\,(x_i,\,x_j) = x_i^T A\,x_j$$

By parameter zing the hash functions instead by G (which is computable since A is p.d.), the following relationship express the hash functions by means of described sample inputs

$$P_r\left[h_r.A(x_i) = h_r.A(x_j)\right] = 1 -$$
$$\left\{1 - \frac{1}{n}\,cos^{-1}\left(\frac{x_i^T A x_j}{\sqrt{|Gx_i||Gx_j|}}\right)\right\} \qquad (6)$$

Phase7.

Semantic web service mediation as used throughout the remainder of SWS either the description of the web service or the description of service request and it is formally represented with in a particular ontology that complies with a certain SWS reference model and defines the semantic similarity between two service members of a

space as a function of the ECD (Euclidean Distance). Euclidean distance between the points representing each of the members, the different distance metrics could be considered based on the nature and purpose of the MS and the definitions of MS given by V, U expressed by vectors $V_0, V_1, \ldots V_n$ and $U_1, U_2, \ldots U_n$ within the MS and their distance can be calculated as

$$\text{Dist}\,(u,\,v) = \sqrt{\sum_{i=1}^{n} P_i\left(\left(\frac{u_i - \bar{u}}{su}\right) - \left(\frac{v_i - \bar{v}}{sv}\right)\right)_2}$$

A similarity based mediation service can be provided as a general purpose mediator and can be implemented as a particular mediation service in which (MWS_1, MWS_2) are the standard web services and denotes the ontological refinement of SWS_i as a service request description. Consequently mediation services MWS computes a set of x sets of distance Dist (SWS_i) = {Dist (SWS_i, SWS_j), Dist (SWS_i, SWS_j)...Dist (SWS_i, SWS_x)}

Where each Dist (SWS_i, SWS_j) contains a set of distance {dist_1...dist_n}

$$\text{Sim}\,(SWS_i,\,SWS_j) = \left(\overline{Dist\,(SWS_i, SWS_j)}\right)^{-1} = \left[\frac{\sum_{k=1}^{n} dist_k}{n}\right]^{-1} \qquad (7)$$

Phase8.

Improving the semantic web services discovery through similarity search in services by means of Metric Space M=<D, d> where domain D is a collection of web services and distance function d is the semantic similarity between requester and web services. A semantic web service modeled in metric space is a triple SWS=<FS, NFS, FW> such that FS is the functional semantics, NFS is the non-functional semantics and FW is the description framework. The functional semantics of SWS is defined as a quadruple of functional features FS=<EX, RE, PR, OU> such that EX describes what a web service expects for enabling it to provide its service, RE describes what a web service returns in response to its input, PR describes conditions over the world state to met before service execution, and OU describes the state of the world after the execution of the service. The non-functional semantics of SWS is defined as a triple of non-functional features NFS = <Q, SP,CP> such that Q is a set of QOS parameters offered by the service, SP is a set of service specific parameters, and CP is context policy Distance measure is defined as semantic similarity based on functional semantics

$$d_M\,(SWS_R,\,SWS_A) = \sqrt{\sum_{i=1}^{n} Wi\,(SWS_{Ri} - SWS_{Ai})2} \qquad (8)$$

Phase9.

Similarity measures for business process models such as label similarity is one way of measuring the similarity between a pair of process models and relation between elements in one model and elements in other models

$$\frac{\sum_{(n,m)\in A} Sim\,l\,(n,m)}{|P_1| + |P_2|} \qquad (9)$$

Where Siml is a similarity measure between pairs of model elements the similarity between model elements

can be computed from their labels using syntactic similarity measures, semantic measures, or a combination of both. Syntactic measures are based on string-edit distance, n-gram, morphological analysis (stemming), and stop-word elimination techniques, whereas semantic techniques are based on synonym and other semantic relations captured. The structural similarity measures are processed as per distance.

Phase10.

Heuristic search and similarity flooding search algorithm A* algorithm was adopted here to find the distance between the services in N-M mapping O (mn) is the worst case in this condition. Fuzzy similarity clustering for the consumer centric for selection of web services, the similarity between WSA_{ai}^{jk} can be obtained via the equation Sim_{ai}^{jk}

$$Sim_{ai}^{jk} = \frac{\int \left(min \left[\hat{m} \left(wsa_{ai}^{j} \right), \hat{m} \left(wsa_{ai}^{k} \right) \right] \right) dx}{\int \left(max \left[\hat{m} \left(wsa_{ai}^{j} \right), \hat{m} \left(wsa_{ai}^{k} \right) \right] \right) dx} \quad (10)$$

We can find the average agreement degree (AAD), denoted as A (WSA_{ai}^{k}) in the form

$$A(wsa_{ai}^{k}) = \frac{1}{n-1} \sum_{\substack{j=1 \\ k \neq j}} Sim\ kj$$

The RAD (Relative Agreement Degree) for each individual opinion uses the RAD equation as given below

$$RAD(wsa_{ai}^{k}) = \frac{A(wsa_{ai}^{k})}{\sum_{j=1}^{n} A(wsa_{ai}^{k})}$$

Similarly to obtain the CDC (Consensus Degree Coefficient) for each participant can be obtained by the equation

$$CDC(wsa_{ai}^{k}) = \beta \times w_k + (1 - \beta) X\ RAD(wsa_{ai}^{k})$$

The aggregate fuzzy opinions by the CDC can be determined as

$$\hat{R}_{ai} = \sum_{k=1}^{n} CDC(wsa_{ai}^{k}) \cdot (wsa_{ai}^{k})$$

Thus the AAD, RAD and CDC equations helps to find the similarity in different services.

Phase11.

Service discovery and integration of similar service and ranking it Current web service description model fails to provide sufficient information to enable ranking mechanism for service web, but at the same time, due to the fast growth of web services, ranking becomes more and more important for a web user to easily find an appropriate service. Thus, to enrich semantic aspect of a service description for ranking, Anchor Semantic Description Model (ASDM) is proposed to incorporate request semantics and reputation as anchors into a service description. ASDM provides an easy and automatic approach to enhance service descriptions on semantic aspect. It is simpler than the complex OWL-S but can include more infor-

mation than the document element of WSDL. Selection usage level indicates that the retrieved service operation is selected by the service requestor to view its detailed information. When a service selection action is captured, the use Frequency of s-anchor element of the corresponding service operation will increase 1. Composition usage level indicates that a retrieved service operation is selected and composed into a service application. When a service composition action is captured, the use Frequency of c-anchor element will increase 1 while the use Frequency of s-anchor element will decrease 1 for the corresponding service operation

$$DScore = \cos(D_r, D_p) = \frac{\sum_{k=1}^{n} W_{rk} \times W_{pk}}{\sqrt{(\sum_{k=1}^{n} W_{rk}^2)(\sum_{k=1}^{n} W_{pk}^2)}} \quad (11)$$

Ranking by means of DScore and to compute the similarity between request description and anchor semantic elements it has a property use frequency and a term list with VSM between request description and anchor semantic elements of a service operation should firstly be represented as a vector.

$$W_{ak} = sr \times T_s + cr \times T_c + dr \times T_d + rr \times T_r$$

Where sr, cr, dr, and rr are the reputation values of the four different usage levels; and Ts, Tc, Td and Tr are the occurring times of the given term in each anchor semantic element. The AScore ranking also found by using the cos (Dr, Da)

$$AScore = \cos(D_r, D_a) = \frac{\sum_{k=1}^{n} W_{rk} \times W_{ak}}{\sqrt{(\sum_{k=1}^{n} W_{rk}^2)(\sum_{k=1}^{n} W_{ak}^2)}} \quad (12)$$

high frequency of service using leads to high quality by RScore in the mean of sr U, F, cr, dr, rr where UFs, UFe, UFd, UFr are the frequency of the four different anchor semantic elements and it is demanded by,

RScore=sr x UFs + cr x UFe + dr x UFd + rr x UFr

Similarity in query retrieval with ranking for XML services and their similarity string and a set of services descriptions

$$Sim = sim(Q, D) = \frac{\sum_{j=1}^{t} W_{aj} \cdot W_{id}}{\sqrt{\sum_{j=1}^{t}(W_{aj})^2 \cdot \sum_{j=1}^{t}(W_{id})^2}} \quad (13)$$

In this equation, wqj is a weighted vector for a word in a query vector, and wid is a weighted vector for a word in a service descriptions. We also use the following method to apply similarity of child entries in the hierarchical structure of service descriptions for measuring upper description similarity. Current Web service discovery mechanisms can be classified into three categories. Registry-based discovery like UDDI, Semantic annotation and discovery, Similarity-based search calculate term frequencies according to their position. By evaluating the different phase (i.e. from phase 1 to Phase 11) and by using equations 1 to 11 the calculation of similarity by means of metrics are described in this work. Identifying the related service and their functionality in these competitive global service providers marked as a robust task in web service

world. Fig.1 demonstrates an overview of finding similarity in the web environment.

   i.   Client request
   ii.   Query Handler
   iii. Query Processor
   iv.   Request Transporting
   v.   Registry Handler
   vi.   Response by XML
   vii. Request Execution
   viii. Save and Retrieve
   ix.   Registry Data Storage

Client Request

End-user and customers demand get processed with high demand and priority as per the request.
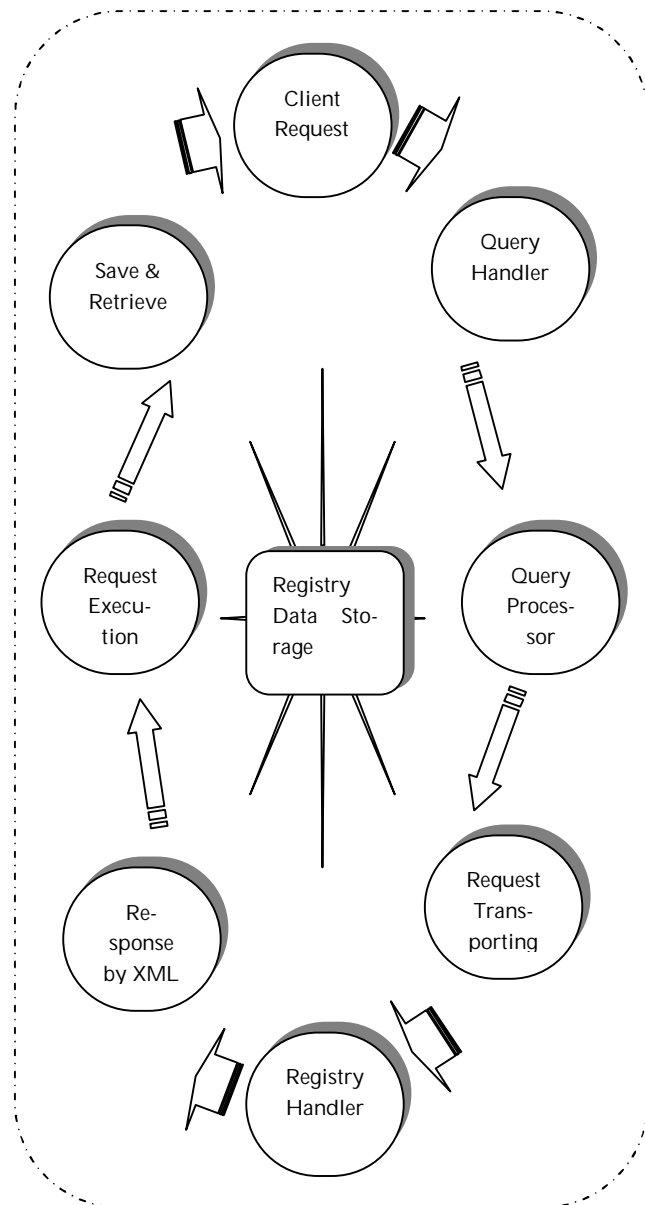
Fig.1. Identifying similarity's in different slice.

Query Handler

The demand and request of different users across globally routed in correct direction with the help of query handler.

Query Processor

As per the request and demand of query handler and end-users, their overall workstation get activated and compared with huge data sets presented in their ware house.

Request Transporting

The communication among various components and elements are by passed with necessary commands and triggers via proper channel and optimized path.

Registry Handler

Maintenance of huge database and ware house of different set of service requests are organized in the form of registry by the help of Registry Handler.

Response by XML

To achieve interoperability as a key feature across the web world the supportive language XML was used to obtain it more consistently.

Request Execution

Client request, Query Handler, Query Processor, Request Transporting, Registry Handler, and Response by XML stands in queue for execution in various environment and critical situations etc.

Save and Retrieve

Requesting services are collected by Registry Handler and processed by request executor in various aspects with respect to its own database registry.

Registry Data Storage

Bulk and huge data set information are collected from various service providers to maintain registry set up as per the conditional requests. All datasets get shuffled periodically based on its updating level.

5 CONCLUSIONS

Verdict various related issues in web environment and solving it in a matter of time, our proposed phases pin points stipulated support to web environmental world. A bulky Web services will be offered in the future, which may have the related or the similar utility, situated in unusual places and furnish by different providers across the internet. The optimization of similar service from different web services are grouped and alerted on nonfunctional requirements such as routine, expenditure and confidence etc. The proposed approach can assure the Upcoming research proposals contemplates to find the simi-

larity in different networks with infinite ranking and un-usual query retrieval in universal web world.

REFERENCES

[1] Pierluigi Plebani and Barbara Pernici,"URBE: Web Service Retrieval Based on Similarity Evaluation", IEEE Transaction on Knowledge and Data Engineering, Vol.21, No.11, Nov-2009.

[2] Shalini Batra, Seema Bawa, Thapar University, "Web Service Categorization Using Normalized Similarity Score", "International Journal of Computer Theory and Engineering", Vol.2, No.1, Feb-2010.

[3] Brian Kulis, Prateek Jain, and Kristen Grauman, "Fast Similarity Search for Learned Metrics", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.31, No.12, Dec-2009.

[4] Khoi Anh Phan, Zahir Tari, Peter Bertok, "Similarity-Based SOAP Multicast Protocol to reduce Bandwidth and latency in Web Services", "IEEE Transaction on Services Computing", Vol.1, No.2, Jun-2008

[5] Ming Li, Xin Chen, Xin Li, Bin Ma, Paul M.B.vitanyi, "The Similarity Metric", "IEEE Transaction on Information Theory", Vol.50, No.12, Dec-2004

[6] Khurram Shehzad, Muhammad Younus Javed, "Multithreaded Fuzzy Logic Based Web Services Mining Framework", "European Journal of Scientific Research", Vol.41, No.4, 633-645, 2010.

[7] Jeffrey Hau, William Lee, John Darlington, "A Semantic Similarity Measure for Semantic Web Services", "IEEE Conference", May-2005

[8] Min Liu, Weiming Shen, Qi Hao, Junwei Yan, "A Multilevel Match Making Frame Work for Semantic Web Service in Collaborative Design", "International Conference on Computer Supported Cooperative Work in Design", April 2008, pp. 392-398

[9] Xianyang Qu, Hailong Sun, Xiang LI, Xudong, Wei Lin, "A WordNet-Based Web Service Similarity Mining Mechanism", IEEE-Conference-JAN-2009.

[10] Shalini Batra, Seema Bawa, "Web Service Categorization Using Normalized Similarity Score", International Journal of Computer Theory and Engineering, Vol. 2, No. 1 February, 2010 1793-8201.

[11] Haoxiang Xia, Taketoshi Yoshida, "Web Service Recommendation with Ontology-based Similarity Measure", IEEE Conference-2009.

[12] Jeffrey Hau, William Lee, John Darlington, "A Semantic Similarity Measure for Semantic Web Services", IEEE-Conference-2007.

[13] Ricardo Mendoza Gonzalez, "Web Service-Security Specification based on Usability Criteria and Pattern Approach", Journal of Computers, Vol. 4, No. 8, AUGUST 2009.

[14] Stefan Dietze, Alessio Gugliotta, John Domingue, "Exploiting Metrics for Similarity-Based Semantic Web Service Discovery", ICWS, pp.327-334, IEEE International Conference on Web Services, 2009.

[15] Minghui Wu, Fanwei Zhu, Jia Lv, Tao Jiang, Jing Ying, tase, "Improve Semantic Web Services Discovery through Similarity Search in Metric Space", Third IEEE International Symposium on Theoretical Aspects of Software Engineering, pp.297-298, 2009.

[16] Marlon Dumas, Luciano Garcia-Ba nuelos, Remco Dijkman, "Similarity Search of Business Process Models", IEEE Conference -2009

[17] Wei-Li Lin, Chi-Chun Lo, Kuo-Ming Chao, Nick Godwin, "Fuzzy Similarity Clustering for Consumer-Centric QOS-aware Selection of Web Services", International Conference on Complex, Intelligent and Software Intensive Systems-2009.

[18] Jing Luo, Ying Li, "Anchor Semantics Enabled Ranking Method for Service Discovery and Integration", 2008 IEEE International Conference on Web Services.

[19] Shirin Sohrabi and Sheila A. McIlraith, ISWC 2009, LNCS 5823, "Optimizing Web Service Composition While Enforcing Regulations", pp. 601–617, 2009, Springer

[20] Kyong-Ha Lee, Mi-young Lee, Yun-Young Hwang, Kyu-Chul Lee, "A Framework for XML Web Services Retrieval with Ranking", 2007 International Conference on Multimedia and Ubiquitous Engineering IEEE.

[21] Jing Luo, Ying Li, "Anchor Semantics Enabled Ranking Method for Service Discovery and Integration", 2008 IEEE International Conference on Web Services.

[22] Shrabani Mallick, D. S Kushwaha, "An Efficient Web Service Discovery Architecture", International Journal of Computer Applications Vol.3 – No.12, July 2010.

[23] Hoi Chan, Trieu Chieu and Thomas Kwok, "Autonomic Ranking and Selection of Web Services by Using Single Value Decomposition Technique", 2008 IEEE International Conference on Web Services.

[24] Manish Godse, Rajendra Sonar, "The Analytical Hierarchy Process Approach for Prioritizing Features in the Selection of Web Service", Indian Institute of Technology Bombay, India, Shrikant Mulik, L& T Infotech Mumbai, India.

Selecting of Slotted AFPM Motors with High Torque Density for Electric Vehicles

S.Asghar Gholamian, M. Ardebili, K. Abbaszadeh, Seyed Akbar Gholamian

Abstract— Double-sided axial flux PM motors (AFPM) are the most promising and widely used types. There are two topologies for slotted double-sided AFPM motors. Selecting an AFPM motors with high torque density is an important parameter in applications. So, comparison of torque density between different topologies of double-sided AFPM motors seems to be necessary. In this paper, the sizing equations of axial flux slotted one-stator-two-rotor (TORUS) and two-stator-one-rotor (AFIR) type PM motors is presented and comparison of the TORUS and AFIR topologies in terms of torque density is illustrated. Field analysis of both Topologies of slotted motors is investigated using Finite Element method (FEM) software. Finally a high torque double-sided slotted AFPM motor is introduced in the paper.

Index Terms— Axial flux PM motors (AFPM), torque density, electrical loading and Finite Element method (FEM).

——————————————— ◆ ———————————————

1 INTRODUCTION

IN conventional machines, the air gap flux density has normally radial direction. In AFPMs, the air gap flux density presents mainly axial direction. In general, AFPMs exhibit an axial length much smaller than the length of a conventional motor of the same rating [1-3].

AFPM motors are particularly suitable for electrical vehicles and industrial equipment. The large diameter rotor with its high moment of inertia can be utilized as a flywheel. These machines are ideal for low speed applications, as for example, electromechanical traction drives hoists. Traction electric motors for EVs should meet high power density, high torque at low speed for starting and climbing, high speed at low torque for cruising, high efficiency over wide speed and torque ranges [1,6].

Selecting of double-sided AFPM motors with high torque density is an important parameter, especially in electrical vehicle applications. So, comparison of torque density between different topologies of double-sided AFPM motors seems to be necessary [3].

The AFPM machine, also called the disc-type machine, is an attractive alternative to the cylindrical RFPM machine due to its pancake shape, high efficiency, compact Construction and high power density [1,3].

There are two topologies for slotted double-sided AFPM motors. These topologies are axial flux slotted one-stator-two-rotor (TORUS) and two-stator-one-rotor (AFIR) type PM motors. Two AFPM motors and their acronyms are selected TORUS-S (Axial flux slotted external rotor internal stator PM stator) and AFIR-S (Axial flux slotted internal rotor external stator PM motor) for detailed analysis. The stator cores of the machine is formed by tape wound core with a lap and short-pitched polyphase AC winding located in punched stator slots. The rotor structure is formed by the axially magnetized NdFeB magnets [4,5].

The topologies used in the study are illustrated in

Fig.1.

(a)

(b)
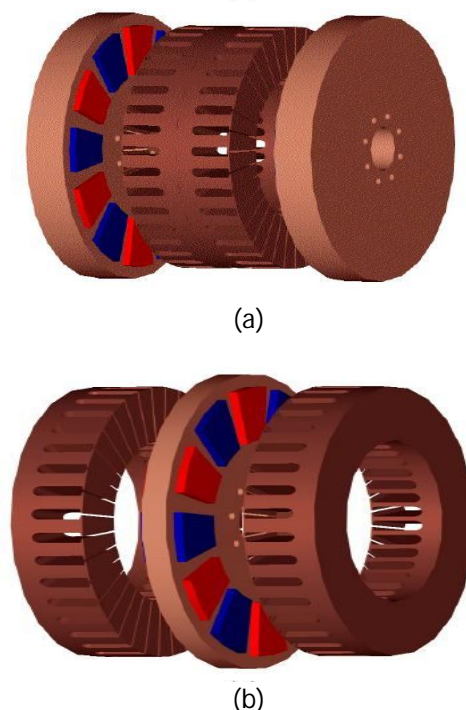
Fig.1. Axial flux slotted (a) one-stator-two-rotor TORUS-S type (b) two-stator-one-rotor AFIR-S type

Flux directions of both AFIR and TORUS slotted topologies at the average diameter in 2D are also shown in Fig. 2a and 2b.

(a)

(b)
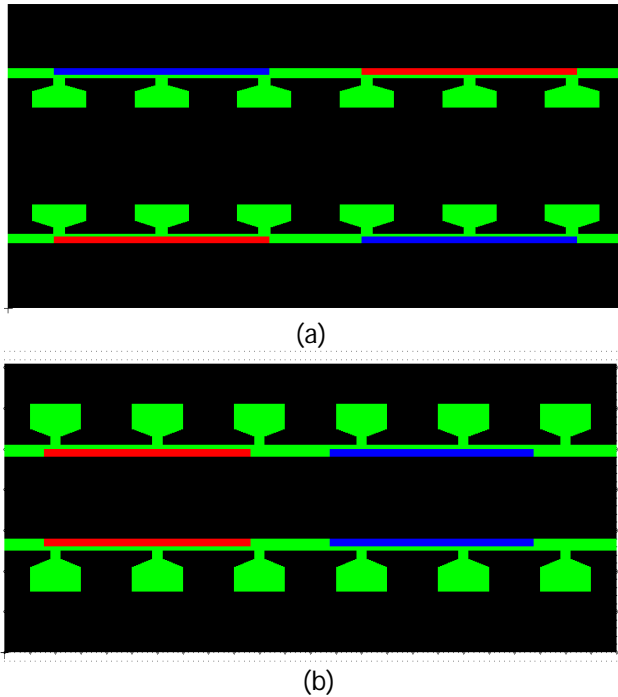
Fig.2. One pole pair of the (a) TORUS-S (b) AFIR-S

Increasing the air gap length, maximum torque density will change in AFPM motors. These changes are not the same in different topologies. Maximum torque density of TORUS-S is higher than AFIR-S in large air gap length.

In Section2, the generalized sizing approach for TO-RUS-S and AFIR-S types PM motors is briefly discussed. Then, some results of comparisons of the TO-RUS-S and AFIR-S topologies in terms of torque density are illustrated in Section3. In Section 4, Field analyses of both Topologies of slotted motors are investigated using Finite Element method (FEM) by MAX-WELL10 software. In Section 5, effect of electrical loading and current density are illustrated. The conclusions are given in Section6.

2. Sizing equations of AFPM Motors

In general, if stator leakage inductance and resistance are neglected, the output power for any electrical machine can be expressed as

$$P_{out} = \eta \frac{m}{T} \int_0^T e(t).i(t)\,dt = m K_p \eta \, E_{pk} I_{pk} \quad (1)$$

where

$e(t)$ and E_{pk} are phase air gap EMF and its peak value, $i(t)$ and I_{pk} are phase current and the peak phase current, η is machine efficiency, m is number of phases of the machine and T is period of one cycle of the EMF[4,5,7].

The quantity K_p is termed the electrical power waveform factor and defined as

$$K_p = \frac{1}{T}\int_0^T \frac{e(t)\times i(t)}{E_{pk}\times I_{pk}}\,dt = \frac{1}{T}\int_0^T f_e(t).f_i(t)dt \quad (2)$$

where

$f_e(t)=e(t)/E_{pk}$ and $f_i(t)=i(t)/I_{pk}$ are the expressions for the normalized EMF and current waveforms. In order to indicate the effect of the current waveform, a definition for current waveform factor, K_i, is also useful,

$$K_i = \frac{I_{pk}}{I_{rms}} = \left[\frac{1}{T}\int_0^T \left(\frac{i(t)}{I_{pk}}\right)^2 dt\right]^{-0.5} \quad (3)$$

where

I_{rms} is the rms value of the phase current. The peak value of the phase air gap EMF for AFPM in (1) is given by:

$$E_{pk} = K_e\, N_{ph}\, B_g \cdot \frac{f}{p}.(1-\lambda^2)\,D_o^2 \quad (4)$$

where

K_e is the EMF factor which incorporates the winding distribution factor K_w and the per unit portion of the total air gap area spanned by the salient poles of the machine (if any), N_{ph} is the number of turn per phase, B_g is the flux density in the air gap, f is the converter frequency, p is the machine pole pairs, λ is the diameter ratio for AFPM defined as D_i/D_o, D_o is the diameter of the machine outer surface, D_i is the diameter of the machine inner surface. The peak phase current in (1) is given by:

$$I_{pk} = A\,\pi\,K_i\,\frac{1+\lambda}{2}\cdot\frac{D_o}{2m_1\,N_{ph}} \quad (5)$$

where

m_1 is number of phases of each stator and A is the electrical loading.
Combining (1) through (5), the general purpose sizing equations take the following form for AFPM.

$$P_{out} = \frac{m}{m_1}\frac{\pi}{2}K_e K_p K_i A B_g\,\eta\frac{f}{p}(1-\lambda^2)(\frac{1+\lambda}{2})D_o^3 \quad (6)$$

The machine torque density for the total volume can be defined as

$$T_{den} = \frac{P_{out}}{\omega_m\frac{\pi}{4}D_{tot}^2 L_{tot}} \quad (7)$$

where

ω_m is the rotor angular speed, D_{tot} is the total machine outer diameter including the stack outer diameter and the protrusion of the end winding from the iron stack in the radial direction, L_{tot} is the total length of the machine including the stack length and the protrusion of the end winding from the iron stack in the axial direc-

tion [4,5,7].

1.1. Sizing Equations for the TORUS-S

The generalized sizing equation approach can easily be applied to axial flux permanent magnet TORUS type motor [5].

The outer surface diameter D_o can be written as

$$D_o = \left(P_{out} \Big/ \frac{\pi m}{2m_1} K_e K_p K_i A B_g \eta \frac{f}{p} (1-\lambda^2)(\frac{1+\lambda}{2}) \right)^{1/3} \quad (8)$$

The machine total outer diameter D_{tot} for the TORUS-S motor is given by

$$D_{tot} = D_o + 2W_{cu} \quad (9)$$

where

W_{cu} is the protrusion of the end winding from the iron stack in the radial direction. For the back-to-back wrapped winding, protrusions exist toward the axis of the machine as well as towards the outsides and can be calculated as

$$W_{cu} = \frac{D_i - \sqrt{D_i^2 - \left(\frac{2AD_g}{K_{cu}J_s}\right)}}{2} \quad (10)$$

where

D_g is the average diameter of the machine, J_s is the current density and K_{cu} is the copper fill factor.

Note for the slotted topology machines the depth of the stator slot for slotted motors is $L_{ss}=W_{cu}$.

The axial length of the machine L_e is given by

$$L_e = L_s + 2L_r + 2g \quad (11)$$

where

L_s is axial length of the stator, L_r is axial length of the rotor and g is the air gap length. The axial length of the stator L_s is

$$L_s = L_{cs} + 2L_{ss} \quad (12)$$

The axial length of the stator core L_{cs} can be written as

$$L_{cs} = \frac{B_g \pi \alpha_p D_o (1+\lambda)}{4p B_{cs}} \quad (13)$$

where

B_{cs} is the flux density in the stator core and α_p is the ratio of average air gap flux density to peak air gap flux density. The axial length of rotor L_r becomes

$$L_r = L_{cr} + L_{PM} \quad (14)$$

Also, the axial length of the rotor core L_{cr} is

$$L_{cr} = \frac{B_u \pi D_o (1+\lambda)}{8p B_{cr}} \quad (15)$$

where

B_{cr} is the flux density in the rotor disc core, and B_u is the attainable flux density on the surface of the PM. The PM length L_{PM} can be calculated as

$$L_{PM} = \frac{\mu_r B_g}{B_r - \left(\frac{K_f}{K_d} B_g\right)} K_c g \quad (16)$$

where

μ_r is the recoil relative permeability of the magnet, B_r is the residual flux density of the PM material, K_d is the leakage flux factor, K_c is the Carter factor, $K_f = B_{gpk}/B_g$ is the peak value corrected factor of air gap flux density in radial direction of the AFPM motor. These factors can be obtained using FEM analysis [5].

1.2. Sizing Equations for the AFIR-S

The concept of Double-sided Axial Flux two-stator-one-rotor (AFIR) type PM motors was presented in [4].

The outer surface diameter D_o is obtained from (6).

$$D_o = \left(2P_{out} \Big/ \frac{\pi m}{2m_1} K_e K_p K_i A B_g \eta \frac{f}{p} (1-\lambda^2)(\frac{1+\lambda}{2}) \right)^{1/3} \quad (17)$$

The machine total outer diameter D_{tot} for the AFIR type machines is given as

$$D_{tot} = D_o + 2W_{cu} \quad (18)$$

where

W_{cu} is the protrusion of the end winding from the iron stack in the radial direction and can be calculated as

$$W_{cu} = \frac{(0.46 - 0.62)D_o}{p} \quad (19)$$

The axial length of the machine L_e is

$$L_e = L_r + 2L_s + 2g \quad (20)$$

where

L_s is axial length of the stator, L_r is axial length of the rotor and g is the air gap length. The axial length of a stator L_s is

$$L_s = L_{cs} + d_{ss} \quad (21)$$

where

L_{cs} is the axial length of the stator core, and the depth of the stator slot for slotted machines d_{ss} is

$$d_{ss} = \frac{D_i - \sqrt{D_i^2 - \left(\frac{2AD_g}{\alpha_s K_{cu} J_s}\right)}}{2} \quad (22)$$

where

α_s is the ratio of stator teeth portion to the stator pole.

The axial length of the stator core L_{cs} can be written as

$$L_{cs} = \frac{B_g \pi \alpha_p D_o (1+\lambda)}{8p B_{cr}} \quad (23)$$

Since there is no rotor core in rotor PM topologies, the axial length of rotor Lr is

$$L_r = L_{PM} \quad (24)$$

The PM length L_{PM} can be calculated as

$$L_{PM} = \frac{2\mu_r B_g}{B_r - \left(\frac{K_f}{K_d} B_g\right)} K_c g \qquad (24)$$

3. Comparison of slotted TORUS and AFIR

Comparison of two different Double-sided axial flux slotted PM motors in terms of torque density is accomplished for 10KW output power, 4 poles and 60Hz drive. In this comparison, other constant parameters of motors are tabulated in table1.

Table1. Constant parameters of motors in comparison

Number of phases	3
Slot fill factor	0.8
Pole arc ratio	0.75
Slot per Pole per Phase	1
flux density in stator	1.5 T
flux density in rotor	1.5 T
Efficiency	90%
Residual flux density of PM	1.1 T

In AFPM motors, the air gap flux density and diameter ratio are the two important design parameters which have significant effect on the motor characteristics. Therefore, in order to optimize the motor performance, the diameter ratio and the air gap flux density must be chosen carefully. Fig.3 shows the torque density variation as a function of air gap flux density and the diameter ratio for the AFIR-S and TORUS-S motors.

As can be seen from Fig.3b, the maximum torque density occurs at Bg=0.51 (T) and $\lambda = 0.27$ □. In various air gap length, the maximum torque density occurs in different Bg and λ.
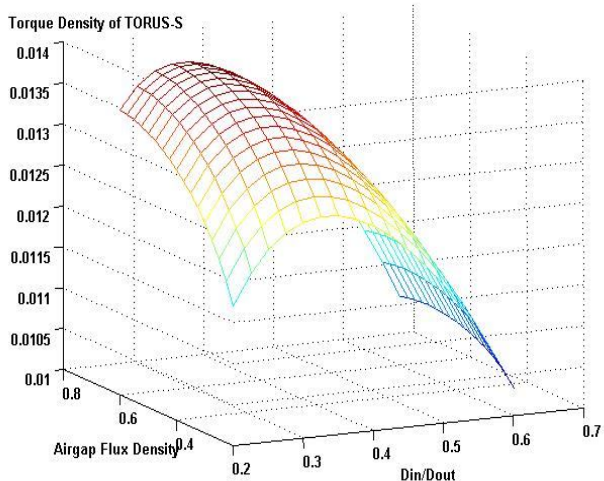
(a)
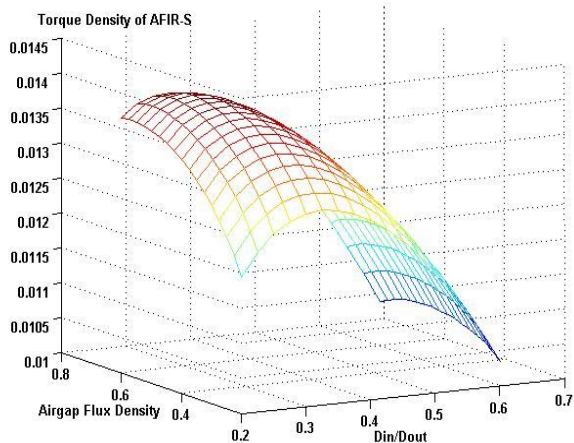
(b)

Fig.3. Torque density vs. air gap flux density and diameter ratio for A=30000 (A/m), g=1 (mm), Js=9000000 (A/m²) a) TORUS-S b) AFIR-S

Table2 shows maximum torque density with corresponding Bg and λ.

Table2. Maximum torque density with corresponding Bg and λ

Type	g (mm)	Bg (T)	λ	Maximum torque density (N.m/cm³)
TORUS-S	1	0.56	0.3	0.014
	1.5	0.57	0.3	0.0137
	2	0.58	0.27	0.0134
AFIR-S	1	0.51	0.27	0.014
	1.5	0.52	0.27	0.0136
	2	0.53	0.28	0.0133

Fig.4 shows the maximum torque density variation as a function of air gap length for the AFIR-S and TORUS-S motors for A=30000 (A/m), Js=9000000 (A/m²).

In special air gap length (this air gap length is called G_T) maximum torque density of AFIR-S and TORUS-S motors will be the same. Considering Fig.4, it can be concluded that in large air gap length, slotted TORUS motor has high power density.
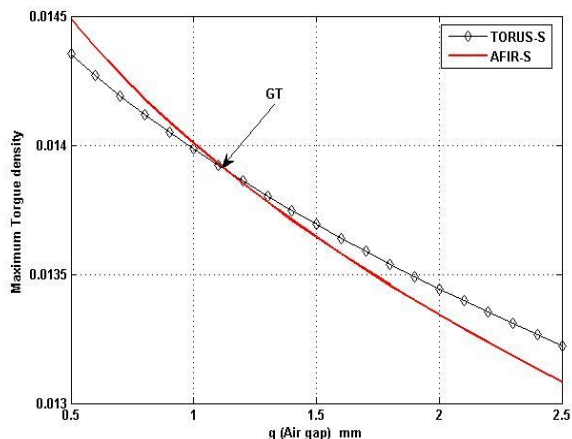
**Fig.4. Maximum torque density
AFIR-S and TORUS-S vs. air gap length**

4. 2D Finite element method Analysis of field
In order to analyze the magnetic circuit and power density, 2D Finite Element Analysis was used for both TORUS-S and AFIR-S type motors [2]. The purpose of the FEM is to get the overall picture of the saturation levels in various parts of the machine, to compare the flux densities obtained from FEM and sizing analysis.

4.1. FEM of the AFIR-S Motor
The motor parameters and important design dimensions used for the AFIR-S model are shown in Table 3. Fig.5 shows the flux distribution over one pole pair using FEM.
Fig.6 shows the air gap Flux density over one pole at the average diameter (Dg) using FEM. This curve shows that the flux density on the edge of the Slots is about 13% lower than the flux density on the center of the PM because of the magnet leakage flux.

Table3. Parameters and dimensions of slotted AFIR-S motor

Air gap length	1 mm
Slot depth	9 mm
Pole-arc-ratio	0.75
Axial length of stator core	16 mm
Axial length of rotor core	40 mm
Axial length of PM	2 mm
Outer diameter	367 mm
Inner diameter	95.5 mm
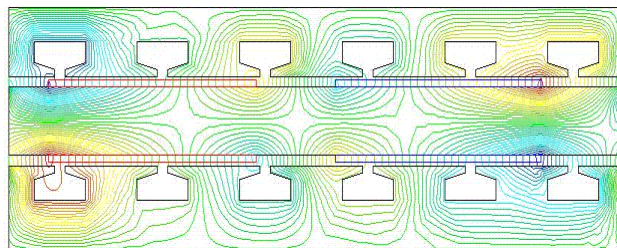
Fig.5. flux distribution over one pole pair for AFIR-S
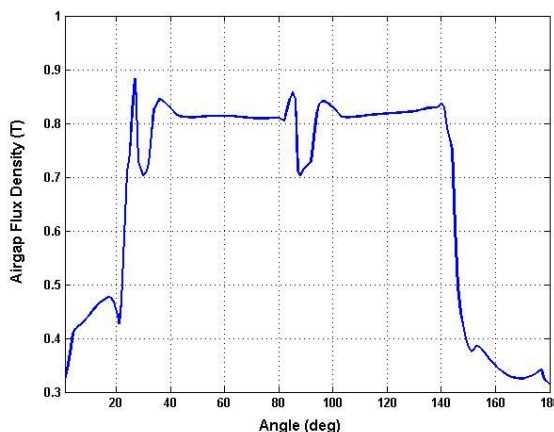
Fig.6. Air gap flux density over one pole for AFIR-S

A flux density comparison between the FEM results and sizing analysis results on various parts of the slotted AFIR motor at no load is tabulated in Table4. The comparison table shows that the FEM results are consistent with the results obtained from the sizing analysis.

Table4. Flux density comparison of slotless AFIR-S motor

	Rotor	Air gap		Stator
	B_{cr}	B_{max}	B_{avg}	B_{cs}
FEM	1.5	0.82	0.55	1.45
Sizing Eq.	1.5 T	0.8	0.53	1.5

4.2. FEM of the TORUS-S Motor

The parameters and optimized TORUS-S motor dimensions used in the design which are calculated using sizing equations are shown in Table 5.

Table5. Parameters and dimensions of slotted TORUS-S motor.

Air gap length	1 mm
Slot depth	10 mm
Pole-arc-ratio	0.75
Axial length of stator core	42 mm
Axial length of rotor core	25 mm
Axial length of PM	2 mm
Outer diameter	356 mm
Inner diameter	103 mm

Fig.7 shows the flux distribution over one pole pair using FEM. The air gap flux density at the average diameter (*Dg*) over one pole using FEM was obtained and is shown in Fig.8.
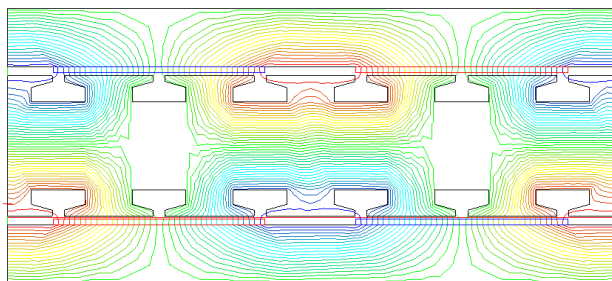
Fig.7. flux distribution over one pole pair for TORUS-S
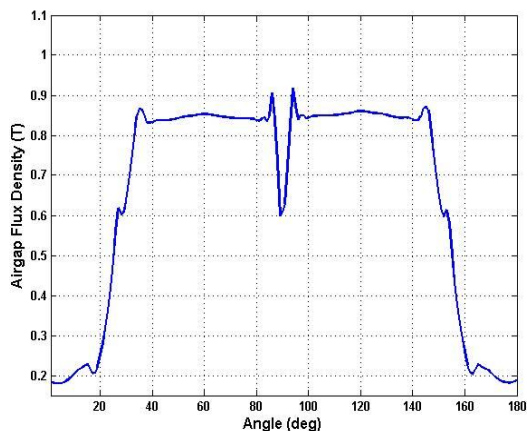
Fig.8. Air gap flux density over one pole for TORUS-S

A comparison of the flux densities between the FEM results and sizing analysis results for different parts of the machine at no load is tabulated in Table 6.

Table6. Flux density comparison of slotless TORUS-S motor

	Rotor	Air gap		Stator
	B_{cr}	B_{max}	B_{avg}	B_{cs}
FEM	1.52	0.85	0.6	1.44
Sizing Eq.	1.5 T	0.8	0.58	1.5

From the no load flux density plots, it is seen that the results are again consistent with the results obtained from the sizing analysis, the maximum flux density values on the rotor and stator came out almost the same. Also, the maximum and average air gap flux densities obtained from the FEM and sizing analysis agree well.

5. Effect of electrical loading and current density

The considerable point is that the value of G_T will vary when the electrical loading '*A*' changes.

Fig.9 shows the variation of the maximum torque density as a function of air gap length in A=25000 (A/m) for the AFIR-S and TORUS-S motors.

Fig.10 shows the variation of the maximum torque density as a function of air gap length in A=35000 (A/m) for the AFIR-S and TORUS-S motors also.

According to fig.10 it can be concluded that point G_T is shifted to larger air gaps and this means that in smaller air gaps AFIR-S motor has higher maximum torque density. According to figure 5 it can be concluded that point G_T is shifted to smaller air gaps and this means that in higher air gaps TORUS-S motor has higher maximum torque density. Other value of G_T for various A is tabulated in table 7.
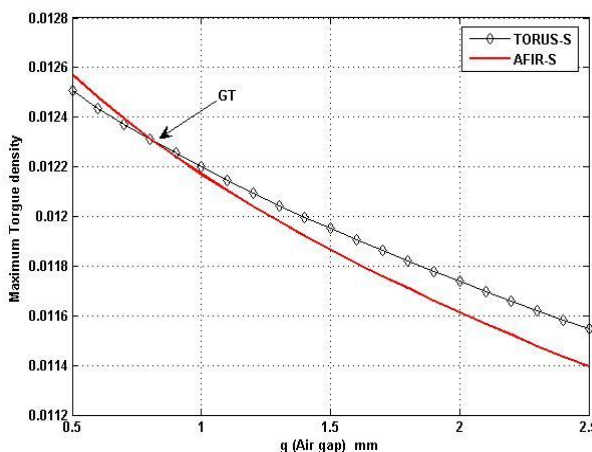
Fig.9. Maximum torque density AFIR-S and TORUS-S vs. air gap length for A=25000 (A/m)
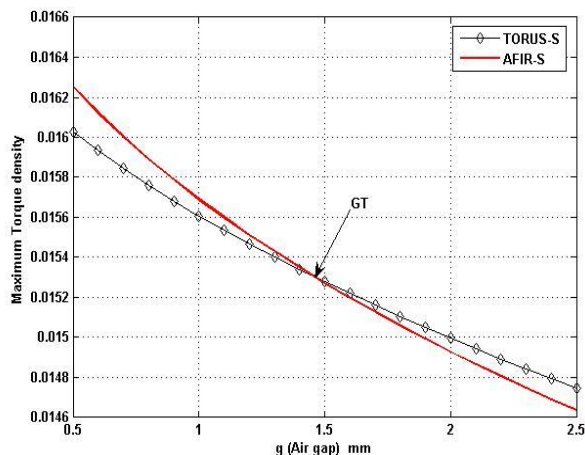
Fig.10. Maximum torque density AFIR-S and TORUS-S vs. air gap length for A=35000 (A/m)

Table7. Other value of GT for Various A

A (A/m)	G_T (mm)
15000	0.43
20000	0.58
25000	0.81
30000	1.12
35000	1.46
40000	1.93

6. CONCLUSION

Selecting an AFPM motors with higher torque density is an important parameter in applications. The main goal of this paper has been introduce to double-Sided Axial Flux Slotted PM Motors with maximum torque density. There are two topologies for slotted double-sided AFPM motors.

The maximum torque density is changed by different value of the air gap and electrical loading. TORUS-S topology has high torque density in low electrical loading. But, AFIR-S topology has high torque density in high electrical loading.

A flux density comparison between the various parts of the slotted AFIR-S and TORUS-S motors obtained from the FEM and sizing analysis at no load agree well.

REFERENCES

[1] S.A. Gholamian, M. Ardebil, K. Abbaszadeh and S. Mahmodi charati; "Optimum Design of 1kW Axial Flux Permanent Magnet Slotted TORUS Motor", European Journal of Scientific Research, ISSN 1450-216X, Vol.21 No.3 (2008), pp.488-499.

[2] S. Asghar Gholamian, M. Ardebili. K. Abbaszadeh,"Selecting and Construction of High Power Density Double-Sided Axial Flux Slotted Permanent Magnet Motors for Electric Vehicles", IREE, Vol. 4. n. 3, pp. 477-484, 2009.

[3] Jacek F. Gieras, Rong-Jie Wang and Maarten J. Kamper, "Axial Flux Permanent Magnet Brushless Machines",Publisher: Springer; 1 edition (January 4, 2005).

[4] Aydin, M.; Huang, S.; Lipo, T.A.; "Optimum design and 3D finite element analysis of nonslotted and slotted internal rotor type axial flux PM disc Machines", Power Engineering Society Summer Meeting, 2001. IEEE Volume 3, 15-19 July 2001 Page(s):1409 - 1416 vol.3.

[5] Aydin, M.; Surong Huang; Lipo, T.A.; "Design and 3D electromagnetic field analysis of non-slotted and slotted TORUS type axial flux surface mounted permanent magnet disc machines", Electric Machines and Drives Conference, 2001. IEMDC 2001. IEEE International2001 Page(s): 645 – 651.

[6] Caricchi, F.; Capponi, F.G.; Crescimbini, F.; Solero, L.; "Experimental study on reducing cogging torque and core power loss in axial-flux permanent-magnet machines with slotted winding", Industry Applications Conference, 2002. 37th IAS Annual Meeting. Conference Record of the Volume 2, 13-18 Oct. 2002 Page(s):1295 - 1302 vol.2.

[7] S. Huang, J. Luo, F. Leonardi and T. A. Lipo, "A Comparison of Power Density for Axial Flux Machines Based on the General Purpose Sizing Equation", IEEE Trans. on Energy Conversion, Vol.14, No.2 June 1999, pp. 185-192.

Addressing Challenges in Multilingual Machine Translation

Prof. Rekha Sugandhi, Sayali Charhate, Anurag Dani, Amol Kawade

Abstract- The machine translation process may be unidirectional or bidirectional between a pair of languages. Or it can be multilingual too. A number of software's are developed till date and different advancements are taking place in this field to overcome the language barriers and create borderless marketplace. Still there are many challenges involved in this field of AI which are yet to be overcome. The translation quality of MT systems may be improved by developing better methods as well as by imposing certain restrictions on the input. All sort of challenges peak in case of Multilingual Machine Translation as compared to bilingual one. Paper focuses on long term challenges like High-Quality MT for many more language pairs, training with limited data resources, robustness across domains, genres and language styles, Achieving human-level translation quality and fluency.

Index Terms— Machine Learning, Machine Translation (MT), Natural Language Processing (NLP), Word sense disambiguation (WSD), Interlingua, Source Language (SL), Target Language (TL) , Artificial Intelligence.

———————————— ◆ ————————————

1 INTRODUCTION

Machine Translation (MT) is a sub-field of Artificial Intelligence (AI) which involves automated translation of text from one natural language to another with the help of computer. At basic level, Machine translator performs simple substitution of words in one natural language for words in another, but that alone usually cannot produce a good translation of a text, because recognition of whole phrases and their closest counterparts in the target language is needed. Every natural language has got its own grammatical structure and certain set of rules. Thus, during task of translation, two things should be taken care of, pertain the meaning as of source language (SL) and output must satisfy lexical rules of target language (TL).

There are many challenges involved in this field of AI which are yet to be overcome. Multilingual MT mainly suffers from 4 types of ambiguity [13]; Lexical ambiguity, Referential ambiguity, scope ambiguity, structural ambiguity. Lexical and structural ambiguities affect quality of translator the most. Lexical ambiguity arises due to multiple meanings of same word, while structural ambiguity arises due to multiple interpretations of the same sentence. During translation, corpus is processed for a

———————————————

- *Rekha S. Sugandhi , working as Assistant Professor at the Department of Computer Engineering, MIT College of Engineering Pune is currently pursuing her Ph.D. in Computer Engineering at the SGB, Amravati University. She has completed her M.Tech in Computer Engineering and B.E. in Computer Engineering from the University of Pune. Her research area and areas of interest include Machine Learning, Natural Language Processing and Theory of Computation. PH-+91-02030273130. E-mail:rekha.sugandhi@gmail.com*
- *Sayali Charhate, is currently pursuing her bachelor degree in Computer Engineering from the University of Pune*
- *Anurag Dani, is currently pursuing his bachelor degree in Computer Engineering from the University of Pune*
- *Amol Kawade is currently pursuing his bachelor degree in Computer Engineering from the University of Pune*

number of times to perform different operations rather than directly translating it to the TL. In this paper we present overview of challenges involved in these processes and long term challenges in this field.

Section 2 focuses on challenges in preprocessing of text i.e. different types of analysis. Section 3 involves challenges during training of machine which affect overall performance and accuracy. Section 4 is about challenges in dealing with multiple languages from different origin and having different characteristics; how these aspects were considered till date. Section 5 comprises of long term challenges followed by some suggestions in section 6. Section 7 contains a proposed model.

2 CHALLENGES IN ANALYSIS

Analysis of input text from SL consists of morphological analysis, syntactic analysis and semantic analysis. Each kind of analysis poses some challenges in MT.

2.1 In Morphological Analysis

There are no generalized grammatical rules in any language [1] which we can use to reduce the size of lexicon. E.g. suffix 'er' is used to indicate a person performing an action like the one who 'Farms' is a Farmer. But this is not always the case e.g. one who 'Cooks' is known as 'Cook', one who 'Drafts' is called as 'Draftsman' and so on. Similarly different forms of a verb are obtained like 'book', 'book-s', 'book-ing', 'book-ed'. But there are many exceptions too like 'do', 'does', 'doing', 'did', 'done' or 'sing', 'sings', 'singing', 'sang', 'sung'. So if we reduce size of lexicon using this approach, it will certainly display the minimal units of grammatical analysis in a vast amount of language data. This technique may also lead to an imperfect attempt to describe something which is too complex [17].

Moreover derivational morphology will tend to change category of the word (POS) unlike the inflectional morphology [1]. As each language varies greatly, even if we find out a rule which is applicable in most of the cases then it will be based on the experience of only one language. Consequently, it will get counter exampled from other

languages. It becomes difficult to handle this issue in case of multilingual MT. Different languages have different way of expressing the same concept. When it comes to multilingual machine translation, it will have to frame different strategies for different TL.

A morphological analyzer has to be able to separate the clitic from its attached morpheme. On the other hand when it comes to agglutinative languages like Turkish, morpheme concept is invaluable [17], as morphemes are combined without fusion or morphophonemic changes.

If we think of listing all inflectional and derivational forms regardless of their type as well as learning program to learn appropriate rules, this may result into redundancy involved in typing out such a huge data and most machine algorithms are designed so as to deal with single token entities.
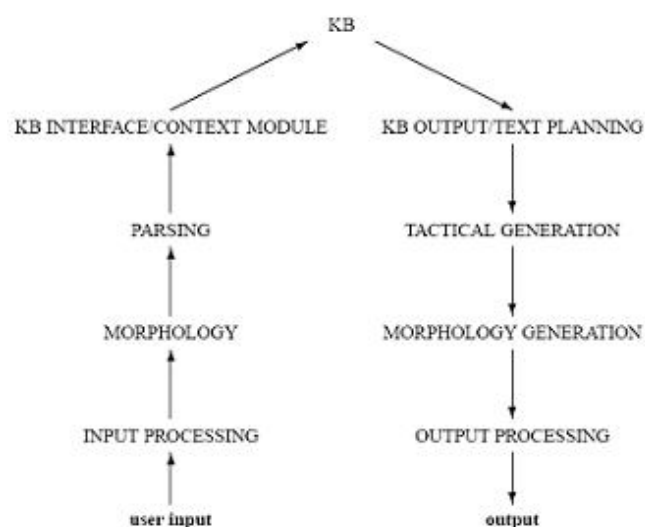
Fig 1 Steps involved in machine translation [11]

There are various methods that could be used to abate this problem, such as passing as many alternatives as possible from the morphological analyzer to the syntactic parser and hoping that the latter can resolve the ambiguities. Another method might be to attempt to operate morphological and syntactic analysis in parallel. But these can be imperfect ways of describing language and we should find a better descriptive model.

2.1.1 Stemming challenges
Challenges arise due to same problem of no generalization of grammatical rules. They work only for few commonly used suffixes [2] e.g. remove 'ed', 'ing', 'ness', 'tion', 'ly' , 's' etc.

> Actively → active (stem) + ly (suffix)
> Wolves → wolve (stem) + s (suffix)
> (2nd example does not work well.)

As all parts of speech have no such well formulated set of rules it's not always straight forward to break word to its

stem and affixes which may lead to two types or errors after stemming-
1. Understemming error- word with same stem are converted to different stems
2. Overstemming error-word with different stem converted to same stem.

Multilingual Stemming is similar to regular stemming but uses morphological rules of several languages at the same time instead of rules relative to only one language [10]. When rules of more than one language apply to a word, extra procedures are used to determine the right rule. Although these procedures cover the majority of possible conflicts, there are cases that cannot be solved. As a result, multilingual stemming is slightly less precise than monolingual stemming for a given language.

So stemming is not suitable when-
1. Precision is major concern
2.index is too large

2.1.1.1 Lemmatization
Lemmatization can be a solution for this problem. It involves detection of parts of speech (POS) followed by stemming rules as per POS [1]. It operates with the help of the dictionary. We can make use of small look up tables to store few frequent exceptions. If the word is not there in this list, then only suffix strapping or lemmatization is done. Lemmatization is more accurate than stemming requires additional resources unlike stemming. Moreover it might not be available for all the languages or it might be costly.

2.1.1.2 Challenges in information retrieval
Problems in stemming impose challenges in IR. Single query used to retrieve the information may return multiple related documents along with the desired document. In such a case, we can narrow it down by adding more words to the query. But still there will be some exceptions. It would be expensive to handle exceptions in terms of memory and speed considering their relative rareness. Exact match queries can also occasionally retrieve non expected results. For instance, the exact match string "sales reports" matches "sal- report-". Moreover it is not preferred to stem short words to favor performance. So, morphological variations of short words should be implicitly included in a query.

Problem becomes severe in case of multilingual machine translation because translating different languages is not necessarily transitive. Proper handling should be done in multilingual MT.

E.g. For example, using Google, "River Bank" in English translates to "Rive" in French [15], which translates to "Ufer" in German, which translates to "Shore" in English.

2.2. In Syntactic Analysis

It includes challenges in tagging and parsing.

2.2.1 Tagging

When data is annotated using part of speech (POS), it is used for parsing and chunking. It can also be used for word alignment. But this type of tagging is not suitable for small training data. Another approach i.e. sense tagging makes use of several predetermined tags which denote the sense of the word in current context. So, this helps in task of disambiguation which is one of the important steps in machine translation. The problem in this technique is that there are no standard tags used by all people or tagging system. This tagging is difficult as compared to POS tagging.

2.2.2 Parsing challenges

It is necessary that every sentence will have only single parse tree. Many times parsing results in multiple parse trees. Meaning or translation relationship changes with each representation. Multiple trees for a single sentence will lead to ambiguities. Moreover it becomes highly complex to represent large sentences using parse trees. Parsing takes long time for large corpora. It slows down processing too because of high computational complexity (of the order of $O(n^3)$ considering length of sentence and of the order of $O(G^2)$ where G➔grammar size). Full parsing is not very robust too. These limitations affect the performance of Syntax -Based Statistical Machine Translation. Chunking can be an option for parsing being more efficient [4].

In case of multilingual MT, parsing technique changes as per the structure of language which of course different for languages with different origin.

2.2.2.1 Chunking

Chunking divides sentences into their sub constituents, such as noun, verb, and prepositional phrases. Chunking might be a first step in full parsing or it's known as partial parsing. Chunks are non-overlapping regions of text which are non-recursive, non-exhaustive [3].

But within chunks, (syntactic) attachment ambiguities do not arise, and simple context-free parsing techniques are very effective. By having separate chunker and attacher, we can limit the use of expensive techniques for dealing with attachment ambiguities to the parts of grammar where they are really necessary—i.e., in the attacher.

Chunk parsing [4] approach can also be used in which parse trees are formed foe individual chunk rather than for whole sentence. Alignment of these chunks will differ for multilingual MT as per every TL. It should not be a limitation of MT.

2.3 In Semantic Analysis

It mainly includes the challenges in disambiguation. First sense baseline is a real challenge for all-words WSD systems. In case of supervised learning, homonym disambiguation challenges are there. Structural ambiguities also come under semantics analysis. Different languages follow different structure of sentence e.g. English follows 'Subject Verb Object' structure, Indian languages follow 'Subject Object Verb' structure while Arabic language follows 'Verb Subject Object' kind of structure. A sentence can also be in passive voice. Phrases must be properly rearranged after the translation in order to have proper grammatical structure. So in multilingual MT it becomes difficult to handle different structure of different TL. Two steps can be implemented in semantic analysis, context independent interpretation and context interpretation [13].

3 CHALLENGES IN TRAINING

Every system should undergo proper training phase which is done using training data i.e. some part of corpus. More the training data more is the system efficiency. This training data is tagged. Corpus contains at least few billions of words. If we consider hand labeling of data, then it's a very tedious job to annotate it. As human interference is there its more error prone. If we consider automated taggers, then there are challenges due to ambiguity of POS of word in current context. In such a case automated tagger need help of morphological analyzer.

3.1 Data bottlenecks

As performance of MT depends on training phase, it may suffer from the problem of data sparseness. In case of small training set, there is unreliability of data prediction. This requires wide coverage system. It must cover full lexicon of languages of interest. Problem arises even when training data is from one domain and system is applied to corpora in some other area.

Now a day, parallel corpus are becoming increasingly available for different language pairs, the magnitude of such corpora that is likely to be available for most language pairs in the foreseeable future is limited. The levels of translation performance that can be achieved using today's MT models with such limited amounts of data are rather unsatisfying. Significant progress is therefore required in developing new types of translation models that better generalize from limited amounts of available training data, in order to enable the development of MT systems for a far broader range of language pairs.

3.2 Knowledge bottleneck

All WSD systems heavily rely on knowledge bases. Lexicon has to be morphologically rich in order to guarantee proper translation. The problem comes in creating and maintaining the knowledge base. Insufficient knowledge base will obviously affect throughput of system. Knowledge base contains billions of words. So, whenever

system searches for a specific word in lexicon, it must retrieve the target meanings in negligible time and then perform the task of disambiguation if needed. It's a costly affair to find out single word each time from billions of words. So it is better to go for domain model in which translator will deal with corpus from a specific domain only. This will result into smaller granularity of knowledge base, but we will have to compromise with precision. Such type of MT is not capable of operating on a morphologically rich language.

Knowledge acquisition bottleneck is more serious kind of problem. Manual creation of knowledge is time consuming and expensive effort. Manual identification and entry of relevant information into knowledge base with change in disambiguation scenario is practically impossible. When we consider automated knowledge acquisition, it is must for every MT so as to improve its morphological knowledge. The basic concept behind this is that what the machine has used recently, which was not there in its knowledge, will be needed in future also. And when we come to a multilingual MT, this problem becomes more severe because quality of translation depends on existing multilingual corpora. A minimum of 2 million words for a specific domain and even more for general language are required. Theoretically it is possible to reach the quality threshold but many times such large amounts of existing multilingual corpora are not there.

4 LANGUAGE-CONSTRAINTS

4.1 Due to the concepts specific to language
Cultural differences also impose challenges during translation. Some words normally related to culture have no equivalent words in other languages. E.g. Indian word 'sari' (traditional dress of Indian women) has no equivalent word in other languages. Same is the case of 'kimono' in Japanese.

In Hindi, person is addressed according to his age. Elder person is addressed using 'aap' while younger one with 'tum'. In most of the Indian languages, same scenario is there. But in language like English there is no such differentiation. 'You' is the only word used for addressing. However in German, unknown person is addressed using 'du', but a known person is addressed using 'sie'.

In many Indian languages, morphological forms of verbs are depends upon gender of the subject or the object usually observed at the end of the sentence which is the not case in English.

In Spanish, appendages of hand or foot are not differentiated, 'dedo' is the only word used to refer those. But in English tow different words are used namely 'toe' and 'finger'.

4.2 Due to change in Linguistic theory

Developments in MT as well as in linguistic theory must go hand in hand considering the performance of system. But practically this is very difficult to achieve. Modification of knowledge base is not simple. Especially if we consider the colloquialism, it's not feasible to modify lexicon time to time neither it's simple to go for online knowledge base as far as MT is concerned. This results in a wide communication gap between theoretical linguistics and practical MT research. Multilingual MT, being concerned with multiple languages, suffers from this problem the most.

Care should be taken for acronyms and official words from languages of interest especially in multilingual MT. Acronyms do not translate well because of different letters used in different languages in different order. To avoid this, the acronyms should become accepted words in multiple languages, they are difficult to translate. Translations of official names as opposed to unofficial and shortened names also must be accounted (e.g., United States, vs. United States of America).

4.3 Focused only on English
USA is the heart of development of computer technology. It is one of the most thoroughly monolingual societies in the world. For this reason, problems arising from the use of the computer with languages other than English did not at first receive much attention, and even today only few languages are there in focus. Most linguistic theory of MT is based upon phenomenon observed in English [14]. This results in less impact of linguistic theory on MT. One of the reasons is that MT research is sometimes regarded as an 'engineering' task, a search for computational methods that work with the facts of language. On the other hand, the aims of many theoretical linguists are more concerned with investigations of human aptitude. Statistical MT in particular, has focused on a small number of language pairs for which vast amounts of sentence-aligned parallel text have become available or was explicitly constructed However, recent theories such as Lexical Functional Grammar or Generalized Phrase-Structure Grammar and their various derivatives have set out explicitly to cover as broad a range as possible, not only within one specific language, but also for different types of languages.

When we compare different languages with English, they vary in writing system, grammatical structure, way of expressing similar meanings and intentions. Many languages of the world use either a completely different writing system from English, or else a slightly different variant of the English writing system, involving in particular special characters. All these differences are problems for which computational solutions must be found. Languages which are not used widely will need deep study and special attention. Languages like Arabic and Hebrew are written from right to left. This needs change in processing technique of different input output

devices. Some languages do not even use alphabetical writing systems and character sets of about 3,000 or 6,000 characters [5]. Japanese in particular have achieved much advancement overcoming this problem.

4.4 Morphological complexity

Every language has different level of morphological complexity and as far as English is concerned, the morphology is quite simple. As a result of this, recent MT research work has placed only limited focus on issues of effectively handling languages with complex morphology especially as a target language. Furthermore, all methods of MT tend to retain SL structural features; however, theoretically MT methods should be more TL oriented. Bur in multilingual MT, translator will have to have different MT methods for different TL.

To address this issue, researchers may have to work upon new translation models that can effectively deal with complex morphology. Suitable language model or n-gram model should be developed [14]. The issue of data sparseness should be taken care of in this case. This can be done by developing translation models that generalize far better from their given training data which is a grand challenge.

As the field develops advanced new types of models, new adaptation techniques, most suitable for these new models will need to be developed and explored.

5 LONG TERM CHALLENGES

The aim behind this field of AI is to achieve human level translation with or without human intervention which will need minimum efforts and minimum time to develop and to process. To achieve this there are some grand challenges in this field.

5.1 Domain Model

In many scenarios, ample amounts of data for training MT systems are available in specific domains and text styles. In that case, machine should considers most common general patterns of mappings between basic syntactic structures in two languages that are likely to hold across genes and text styles. Such patterns can be acquired from training data in one domain, and yet be effective when translating text from a different domain. Word-level translation pairs for the new domain still must be acquired somehow, but this can conceivably be done using far smaller amounts of new domain-specific training data.

When very limited amounts of training data is available in targeted genres and domains, a major challenge that how to adapt or extend MT systems to such new domains and genres to utilize the most out of the limited data that is available. Very little research has been done to date on methods that can identify the differences between genres and domains and use this information for targeted learning of new models [14]. Furthermore, in some scenarios, it may

be possible to actively create small amounts of targeted new training data that are most useful for improving MT system performance.

5.2 Human level translation quality

To achieve the MT performance close to human performance level, translation models must be able to capture advanced syntax and semantic representations and their correspond across languages. The key technical challenge is identifying representation formalisms that are rich and powerful enough on the one hand, yet are simple enough to support the development of algorithms for automatic acquisition of the models from training data, appropriately annotated. The main learning task in this case is to discover the correspondences between the structures in the two languages and to model these correspondences statistically [14]. A more challenging scenario is to learn such models using parallel data where syntax and semantic structures are available for only one of the two sides. The learning task in this case is to project the structures from one language to their corresponding structures in the other language, using word-to-word correspondences. The bottleneck is obtaining the needed training resources. Only small and very limited annotated corpora of this kind currently exist. The development of such annotated corpora is a critical enabling step, without which this research direction cannot hope to even get started. In its easiest to implement multilingual form, multilingual users contribute in multilingual settings. Unfortunately, the number of multilingual users is limited and the number of languages that any one individual communicates in is limited. As a result, in some cases users can employ their own language and that language will be translated.

5.3 Inter-sentential context

During translation, the focus is on an individual sentence. Thus for the correct translation of pronominal referents inter-sentential context has to be considered to resolve specific types of ambiguity. Result should be a coherent multi sentence discourse structure in the target language.

5.4 In Interlingua approach

Interlingua is standard of representing a natural language. It has got its own rules and structure independent of any language [18][19]. Source text is represented using Interlingua. To represent source language sentences all the way into a language-independent representation requires a complex series of NLP components. It is very complex and requires experts. Generating target sentences from Interlingua is similarly challenging. Furthermore, researchers have broadly recognized the extreme challenge of devising a true Interlingua representation that is simultaneously adequate for all languages it is intended for, rich enough to represent all intended meaning, and simple

enough for humans to agree upon and for NLP algorithms to analyze and generate from.

If such a system is developed then we can have multiple source languages for a given system.

6 SUGGESTIONS

The previous section of the paper discusses various grave challenges that come in the way of accurate and efficient translated target language outcome for any source input. We therefore suggest a generalized intermediate language representation that will easily map SL and TL constructs with the help of auxiliary tags that will take care of extra information that may or may not be considered (depending on the target language constructs) in the translation. Though this would involve storing additional tags in the intermediate language document, it stills proves feasible considering the various choices of target languages that we can output.

Secondly, we can also maintain tuning parameters that can decide the granularity to be considered during the translation. For highly granular systems, the word sense disambiguation can be precisely or accurately done, while low granularity, some disambiguation steps may be speeded up by compromising between precision and translation time.

7 PROPOSED MODEL

Based on the overall study of challenges and suggestions from previous section, we are proposing a new model for machine translation. Following figure shows the block diagram of the model.
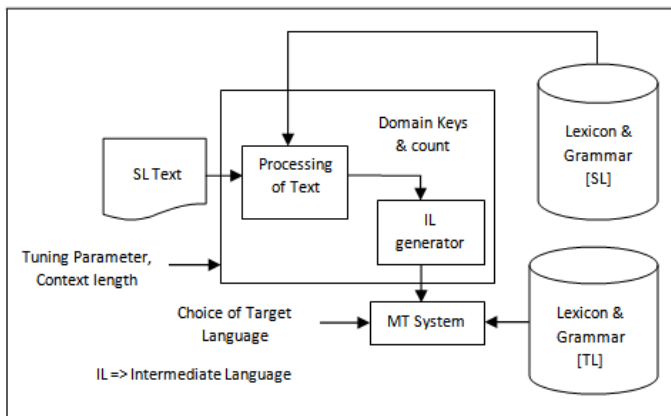
Fig2 Improved Model for Multilingual Machine Translation

Considering structural complexity of Interlingua, it is better to go for a simplistic approach like an intermediate language. It is better to consider the structural constructs for basic languages like *Sanskrit*, for the structure definition of the IL and the mapping of SL with IL and then with the TL. The reason for consideration of basic and simpler languages for the IL structure definition is because

inconsistencies in sentence formation are least in such languages. A major problem with derived languages such as *English* is that many sentences may be correctly formed in more than one ways. Hence automated syntax checks in such cases become very difficult. Also some languages are derived further into dialects for which the grammar may not be well-defined or consistent. The solution we propose to this problem is to convert the source language into its base language and generate some auxiliary tags for extra information by applying context checks on the input text. Hence, size of the context window can be varied depending on the tuning required for accuracy of meaning extraction.

Context check is very important in multi-lingual translation for various reasons explained in the previous section. One example for such case would be where in *English* a simple sentence like "Where are you going?" cannot be properly converted into *Hindi* equivalent unless the gender of the person (to whom the question is addressed) is known. This knowledge can only be extracted by mining the text surrounding the sentence to be translated.

Languages, their evolvement into other languages and their derivatives can be represented in a tree structure. Thus a language that exists at the root of this tree is the ancestral language. It is easier to map one ancestral language to another ancient language. Even we can easily map a language to its ancestral language and vice versa. By making use of this concept we can perform translation.
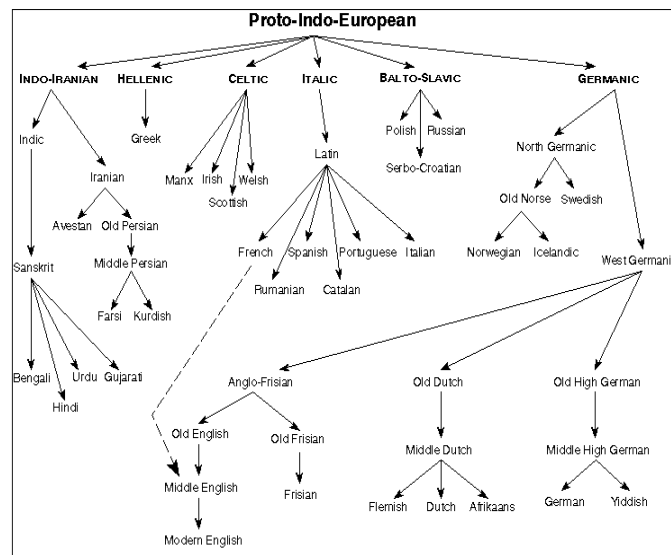
Fig 3 Language Tree [21]

In first phase, text from SL is processed using knowledge repository comprising of lexicon and grammar rules. Different domain keys are provided to processing module. Domain count will be calculated during processing and passed on to IL generator. Processed text will be

converted to intermediate language in second phase. Along with the text to be translated, other inputs can be tuning parameter and context length or context window [20] for disambiguation. This overall process will result into intermediate language file. Now the remaining task is to convert it to the desired TL text which is done in third phase by MT system. Being multilingual MT, there will be choice of TL. MT will be provided with the lexicon of corresponding TL as well as grammar rules i.e. knowledge repository for TL.

Implementation will be easy if SL and TL are having same ancestral language. For example to translate text in Marathi to Malayalam (both of which are Indian Languages), Sanskrit seems to be the obvious choice for the Intermediate language representation. Furthermore, we can even use two intermediate languages. One will be holding characteristics of SL and another that of TL.

8 CONCLUSION

Although the theory of machine translation is quite old but it still poses long term challenges. There is good scope of substantial improvement in translation quality, robustness across domains, genres and language styles. Research till date is mainly focused on English. Except very few languages like Japanese, German and Chinese, other languages had been neglected. To increase the scope of application on MT, Research on other languages should be equally emphasized. It is needed to develop high quality MT for other languages too. Multilingual MT will be effective one rather than having separate MT for each pair of language. MT quality should be so high so that it should match with human-level translator quality and fluency which is the ultimate goal of MT. The MT Models should be Trainable with Limited Data Resources so as to train MT in short span of time. It will reduce cost of training. Low cost techniques should be explored without compromising the quality of output. With the reduction of cost of Machine Translation, its area of application will be augmented.

REFERENCES

[1]   Philipp Koehn, "Advanced Natural Language Processing," Lecture 2, Morphology, 24 September 2010.

[2]   D. Grimshaw, "Artificial Intelligence Topics With Agents," Fall 2001.

[3]   W. Raynor, "International Dictionary Of Artificial Intelligence," (Amacom, 1999). (Book)

[4]   Shin-ichiro KAMEI and Kazunori MURAKI, "Interlingua Developed And Utilized In Real Multilingual MT Product Systems," 1997.

[5]   W. John Hutchins, Harold L. Somers , " An Introduction To Machine Translation," The Library,University of East Anglia, Norwich, UK, 1992.

[6]   Cyril Goutte, Nicola Cancedda,Marc Dymetman, " Learning Machine Translation," George Foster, Massachusetts Institute of Technology.

[7]   Yaser Al-Onaizan, Ulrich Germann, Ulf Hermjakob, Kevin Knight,Phillip Koehn, Daniel Marcu, Kenji Yamada, " Translation with Scarce Resources," American Association For Artificial Intelligence, 2000.

[8]   Fuji Ren, Hongchi Shi ,"A General Ontology Based Multi-Lingual Multi-Function Multi-Media Intelligent System" .

[9]   Philipp koehn, http://homepages.inf.ed.ac.uk/pkoehn (URL)

[10] Coveo knowledge base, http://www.coveo.com/en/support/articles/Information%20-%20CES4-060330-3%20-%20Understanding%20Stemming.htm (URL)

[11] http://www.rustyspigot.com/Computer_Science/Natural%20Language%20Processing2.html (URL)

[12] LIM Lian Tze," Multilingual Lexicons for Machine Translation"

[13] Shaidah Jusoh , Hejab M. Alfawareh , "Automated Translation Machines: Challenges and a Proposed Solution," Second International Conference on Intelligent Systems, Modelling and Simulation, 2011.

[14] Alon Lavie, David Yarowsky, Kevin Knight, Chris Callison-Burch, Nizar Habash, Teruko Mitamura, "MINDS Workshops Machine Translation Working Group Final Report".

[15] Daniel E. O'Leary," Multilingual Knowledge Management ".

[16] Hozumi Tanaka, "Progress in Machine Translation", ed. Sergei Nirenburg, "Multi-Lingual Machine Translation" Systems in the Future", Amsterdam: IOS Press, 1993.

[17] http://www.cs.bham.ac.uk/~pjh/sem1a5/pt2/pt2_intro_morphology.html (URL)

[18] Teruko Mitamura, Eric Nyberg, Jaime Carbonell ,"Kant knowledge based , accurate natural language translation," Center for Machine Translation

[19] Sergei Nirenburg, "New developments in knowledge-based machine translation," Carnegie Mellon University.

[20] ROBERTO NAVIGLI, "Word Sense Disambiguation: A Survey," ACM Computing Surveys, Vol. 41, No. 2, Article 10, February 2009.

[21] http://andromeda.rutgers.edu/~jlynch/language.html (URL)

Evolving Data Mining Algorithms on the Prevailing Crime Trend – An Intelligent Crime Prediction Model

A. Malathi and Dr. S. Santhosh Baboo

Abstract— Crime is a behavior deviation from normal activity of the norms giving people losses and harms. Crimes are a social nuisance and cost our society dearly in several ways. In this paper we look at use of missing value and clustering algorithm for crime data using data mining. We will look at MV algorithm and Apriori algorithm with some enhancements to aid in the process of filling the missing value and identification of crime patterns. We applied these techniques to real crime data. Crime prevention is a significant issue that people are dealing with for centuries. We also use semi-supervised learning technique in this paper for knowledge discovery from the crime records and to help increase the predictive accuracy.

Index Terms— Crime-patterns, clustering, data mining, law-enforcement, Apriori.

———————————————— ◆ ————————————————

1 INTRODUCTION

Crime is a behavior disorder that is an integrated result of social, economical and environmental factors. Crimes are social nuisance and cost our society in several ways. In the world today crime analysis is gaining significance and one of the most popular disciplines is crime prediction. Stakeholders of crime intend to forecast the place, time, number of crimes and crime types to get precautions. With respect to these intentions, in this paper a crime prediction model is generated. .

We today, security are considered to be one of the major concerns and the issue is continuing to grow in intensity and complexity. Security is an aspect that is given top priority by all political and government worldwide and are aiming to reduce crime incidence[5]**.** Reflecting to many serious situations like September 11, 2001 attack, Indian Parliament Attack, 2001, Taj Hotel Attack, 2006 and amid growing concerns about theft, arms trafficking, murders, the importance for crime analysis from previous history is growing. The law enforcement agencies are actively collecting domestic and foreign intelligence to prevent future attacks.

The model is generated by utilizing crime data for few years from the years 2006 to 2010. Methodology starts with obtaining clusters with different clustering algorithms. Then clustering methods are compared to select the most appropriate clustering algorithms.

- *A. Malathi, Assistant Professor, PG and Research Department of Computer Science, Government Arts College, Coimbatore. She is currently pursuing Docrate program in Research and Developemnt centre, Bharathiar University, India, PH-09942526000.*
  E-mail: malathi.arunachalam@yahoo.com
  .
- *Dr. S. Santhosh Baboo, Reader, Post Graduate and Research Department of Computer Science, D. G. Vaishnav College, Chennai. India. PH-0999.*
  E-mail: santhos2001@sify.com

Later crime data is divided into daily apoch, to observe spatiotemporal distribution of crime. In order to predict crime in time dimension is fitted for each week day, then the forecasted crime occurrences in time are disaggregated according to spatial crime cluster patterns. Hence the model proposed in this thesis can give crime prediction in both space and time to help police departments in tactical and planning operations.

The high volume of crime datasets and also the complexity of relationships between these kinds of data have made criminology an appropriate field for applying data mining techniques. Identifying crime characteristics is the first step for developing further analysis. The knowledge that is gained from data mining approaches is a very useful tool which can help and support police forces [8]. According to[9], solving crimes is a complex task that requires human intelligence and experience and data mining is a technique that can assist them with crime detection problems. The idea here is to try to capture years of human experience into computer models via data mining.

In the present scenario, the criminals are becoming technologically sophisticated in committing crimes [1]. Therefore, police needs such a crime analysis tool to catch criminals and to remain ahead in the eternal race between the criminals and the law enforcement. The police should use the current technologies [4] to give themselves the much-needed edge. Availability of relevant and timely information is of utmost necessity in conducting of daily business and activities by the police, particularly in crime investigation and detection of criminals. Police organizations everywhere have been handling a large amount of such information and huge volume of records. There is an urgent need to analyzing the increasing number of crimes as approximately 17 lakhs Indian Penal Code (IPC) crime, and 38 lakhs local and Special Law

crimes per year.

An ideal crime analysis tool should be able to identify crime patterns quickly and in an efficient manner for future crime pattern detection and action. However, in the present scenario, the following major challenges are encountered.

- Increase in the size of crime information that has to be stored and analyzed.
- Problem of identifying techniques that can accurately and efficiently analyze this growing volumes of crime data
- Different methods and structures used for recording crime data.
- The data available is inconsistent and are incomplete thus making the task of formal analysis a far more difficult.
- Investigation of the crime takes longer duration due to complexity of issues

All the above challenges motivated this research work to focus on providing solutions that can enhance the process of crime analysis for identifying and reducing crime in India. The main aim of this research work consist of developing analytical data mining methods that can systematically address the complex problem related to various form of crime. Thus, the main focus is to develop a crime analysis tool that assists the police in

- o Detecting crime patterns and perform crime analysis
- o Provide information to formulate strategies for crime prevention and reduction
- o Identify and analyze common crime patterns to reduce further occurrences of similar incidence

The present research work proposes the use of an amalgamation of data mining techniques that are linked with a common aim of developing such a crime analysis tool. For this purpose, the following specific objectives were formulated.

- o To develop a data cleaning algorithm that
  - ▪ cleans the crime dataset, by removing unwanted data
  - ▪ Use techniques to fill missing values in an efficient manner

- o To explore and enhance clustering algorithms to identify crime patterns from historical data
- o To explore and enhance classification algorithms to predict future crime behaviour based on previous crime trends
- o To develop anomalies detection algorithms to identify change in crime patterns

These techniques do not have a set of predefined classes for assigning items. Some researchers use the statistics-based concept space algorithm to automatically associate different objects such as persons, organizations, and vehicles in crime records [7]. Using link analysis techniques to identify similar transactions, the Financial Crimes Enforcement Network AI System [10] exploits Bank Secrecy Act data to support the detection and analysis of money laundering and other financial crimes. Clustering crime incidents can automate a major part of crime analysis but is limited by the high computational intensity typically required.

2 LITERATURE REVIEW

Data mining in the study and analysis of criminology can be categorized into main areas, crime control and crime suppression. Crime control tends to use knowledge from the analyzed data to control and prevent the occurrence of crime, while the criminal suppression tries to catch a criminal by using his/her history recorded in data mining.

constructed a software framework called ReCAP (Regional Crime Analysis Program) for mining data in order to catch professional criminals using data mining and data fusion techniques[3]. Data fusion was used to manage, fuse and interprets information from multiple sources. The main purpose was to overcome confusion from conflicting reports and cluttered or noisy backgrounds. Data mining was used to automatically discover patterns and relationships in large databases.

Crime detection and prevention techniques are applied to different applications ranging from cross-border security, Internet security to household crimes. Proposed a method to employ computer log files as history data to search some relationships by using the frequency occurrence of incidents[2]. Then, they analyzed the result to produce profiles, which can be used to perceive the behavior of criminal.

Introduced a framework for crime trends using a new distance measure for comparing all individuals based on their profiles and then clustering them accordingly[6]. This method also provided a visual clustering of criminal careers and identification of classes of criminals.

From the literature study, it could be concluded that crime data is increasing to very large quantities running into zota bytes (1024bytes). This in turn is increasing the need for advanced and efficient techniques for analysis. Data mining as an analysis and knowledge discovery tool has immense potential for crime data analysis. As is the case with any other new technology, the requirement of such tool changes, which is further augmented by the new and advanced technologies used by criminals. All these facts confirm that the field is not yet mature and needs further investigations.

3. PREPROCESSING

Data preprocessing is a process that consists of data cleaning, data integration and data transformation which is usually processed by a computer program. It intends to reduce some noises, incomplete and inconsistent data. The results from preprocessing step can be later proceeding by data mining algorithm.

The dataset used in experiment contains various items like year, state code, status of administrative unit, name of the administrative unit, number of crimes with respect to murder, dacoity, riots and Arson, area in sq. meters of the administrative unit, Estimated Mid-Year Population of the Administrative Unit in 1000s (begins in 1964), Actual Civil Police Strength (numbers of personnel), Actual Armed Police Strength (numbers of personnel) and Total Police Strength (Civil and Armed Police).

3.1 Missing value handling

The experiment concentrate on only those attributes that are related to crime data, that is year, state, administrative name, number of crimes for the years 1971 to 2006. The quality of the results of the mining process is directly proportional to the quality of the preprocessed data. Careful scrutiny revealed that the dataset have missing data in state and number of crimes attributes.

3.2 Missing value handling for number of crimes ocurred attribute

In the present research work, while considering filling missing number of crimes related murder, dacoity, riots and arson, two methods were used. Initially, all the four fields are analyzed for empty values. If all the four attributes have empty values for a particular record, then the entire record is considered as irrelevant information and is deleted.

○ While taking individual attributes into consideration, a novel KNN-based imputation method is proposed. In this method, the missing values of an instance are imputed by considering a given number of instances that are most similar to the instance of interest. The similarity of two instances is determined using a distance function.

○ The new algorithm is as follows

1. Divide the data set D into two parts. Let D_m be the set containing the instances in which at least one of the features is missing. The remaining instances will complete feature information form a set called D_c.

   For each vector x in D_m:

   a. Divide the instance vector into observed and missing parts as $x = [x_o, x_m]$.

   b. Calculate the distance between the x_o and all the instance vectors from the set D_c.

   c. Use only those features in the instance vectors from the complete set D_c, which are observed in the vector x.

   d. Use the P closest instances vectors and perform a majority voting estimate of the missing values for categorical attributes. For continuous attributes replace the missing value using the mean value of the attribute in the P (related instances)

The challenging decisions that have to be carefully chosen are:

(i) The choice of the distance function. In the present work, four distance measures, Euclidean, Manhattan, Mahalanobis and Pearson, are considered and the one that produced best result is considered.

(ii) The KNN algorithm searches through all the dataset looking for the most similar instances. This is a very time consuming process and it can be very critical in data mining where large databases are analyzed.

To speed up this process a method that combines missing value handling process with classification is proposed.

(iii) The choice of k, the number of neighbors. Experiments showed that a value of 10 produce best results in terms of accuracy and hence is used in further experimentation.

Thus, the traditional KNN Imputation method was enhanced in two manners. The first enhancement is achieved by proposing a new distance metric and the second enhancement is achieved by using LVQ (Learning Vector Quantization) methods combined with generalized relevance learning to perform the classification and missing value treatment simultaneously. Both these enhancement when combined together produces a model (E-KDD) that is efficient in terms of speed and accuracy.

3.3 Missing value handling in the prediction of the size of Population of the city

The first task is the prediction of the size of the population of a city. The calculation of per capita crime statistics helps to put crime statistics into proportion. However, some of the records were missing one or more values. Worse yet, half the time, the missing value was the "city population size", which means there was no per capita statistics for the entire record. Over some of the cities did not report any population data for any of their records. To improve the calculation of "yearly average per capita crime rates", and to ensure the detection of all "per capita outliers", it was necessary to fill in the missing values. The basic approach to do this was to cluster population sizes, create classes from the clusters, and then classify records with unknown population sizes. The justification for using clustering is as follows: Classes from clusters are more likely to represent the actual population size of the cities. The only value needed to cluster population sizes was the population size of each record. These values were clustered using EM algorithm and initially 10 clusters were chosen because it produced clusters with mean values that would produce per capita calculations close to the actual value

4   CRME PREDICTION MODEL

Given a set of objects, clustering is the process of class discovery, where the objects are grouped into clus-

ters and the classes are unknown beforehand. Two clustering techniques, K-means and DBScan (Density-Based Spatial Clustering Application with Noise) algorithm are considered for this purpose. The algorithm for k-means is given below.

The HYB algorithm is given below.

The HYB algorithm clusters the data m groups where m is predefined
Input – Crime type, Number of Clusters, Number of Iteration
Initial seeds might produce an important role in the final result
Step 1: Randomly Choose cluster centers;
Step 2: Assign instances to clusters based on their distance to the cluster centers
Step 3: centers of clusters are adjusted
Step 4: go to Step 1 until convergence
Step 5: Output C0, C1, C2, C3

From the clustering result, the city crime trend for each type of crime was identified for each year. Further, by slightly modifying the clustering seed, the various states were grouped as high crime zone, medium crime zone and low crime zone. From these homogeneous groups, the efficiencies of police administration units i.e. states can be measured and the method used is given below.

Output Function of Crime Rate = 1/Crime Rate

Here, crime rate is obtained by dividing total crime density of the state with total population of that state since the police of a state are called efficient if its crime rate is low i.e. the output function of crime rate is high.

Thus the two clustering techniques were analyzed in their efficiency in forming accurate clusters, speed of creating clusters, efficiency in identifying crime trend, identifying crime zones, crime density of a state and efficiency of a state in controlling crime rate. Experimental results showed that HYB algorithm show improved results when compared with k-means algorithm and therefore was used in further investigations.

Crime Trend Prediction

The next task is the prediction of future crime trends. This involves tracking crime rate changes from one year to the next and used data mining to project those changes into the future. The basic method involves cluster the states

having the same crime trend and then using "next year" cluster information to classify records. This is combined with the state poverty data to create a classifier that will predict future crime trends.

The Major crimes under property crime are discussed here. There are many categories of crimes like Crime against women, property crime, Road Accident.

- Murder
- Murder for Gain Dacoity
- Robbery
- Burglary
- Theft

To the clustered results, a classification algorithm was applied to predict the future crime pattern. The classification was performed to find in which category a cluster would be in the next year. This allows us to build a predictive model on predicting next year's records using this year's data. The C4.5 decision tree algorithm was used for this purpose. The generalized tree was used to predict the unknown crime trend for the next year. Experimental results proved that the technique used for prediction is accurate and fast. The following are four different clusters produced depends upon the crime nature

- C0: Crime is steady or dropping. Theft is the primary crime little increased and dropping.

- C1: Crime is rising or in flux. Dacoity is the primary crime rates changing..

- C2: Crime is generally increasing. Robbery, Murder, Murder for gain, and Burglery are the primary crime on the rise.

- C3: Few crimes are in flux. Dacoity is in flux. It has gone down and increased then once again gone down.

5. IMPLEMENTATION

Major two crimes Burglary and Murder were taken to analyse the existing crime. Crime Burglary was in increasing, In the year 2006 it got decreased, then it keeps increasing till 2010. Crime Murder kept increasing from 2006 to 2010. The sample crimes Burglary and Murder belong to the cluster C2.
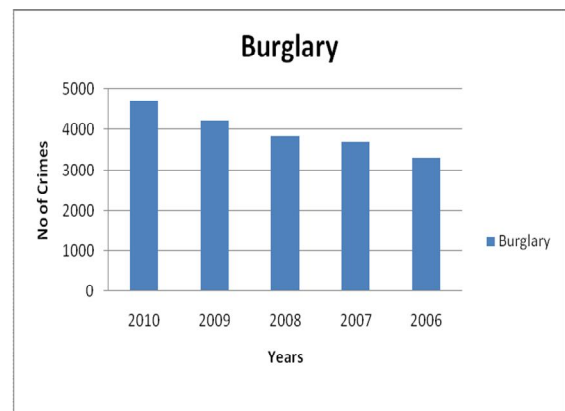
Fig. 1. Crime Burglary Analysis
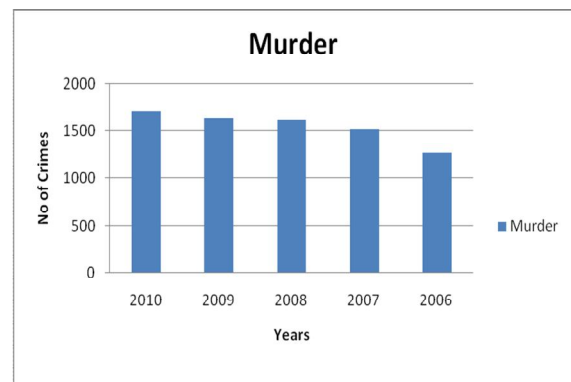
Fig. 2. Crime Murder Analysis

The Murder crime was taken to analyse the future crime prediction. This crime was analysed for the period 2006 to 2009. Both existing algorithm and the new algorithm are executed for the same data set. The existing algorithm predicted the crime as 83%. The new algorithm predicted the crime as 89%.

6 CONLUSION

A major challenge facing all law-enforcement and intelligence-gathering organizations is accurately and efficiently analyzing the growing volumes of crime data. As information science and technology progress, sophisticated data mining and artificial intelligence tools are increasingly accessible to the law enforcement community. These techniques combined with state-of-the-art Computers can process thousands of instructions in seconds, saving precious time. In addition, installing and running software often costs less than hiring and training personnel. Computers are also

less prone to errors than human investigators, especially those who work long hours.

This research work focus on developing a crime analysis tool for Indian scenario using different data mining techniques that can help law enforcement department to efficiently handle crime investigation. The proposed tool enables agencies to easily and economically clean, characterize and analyze crime data to identify actionable patterns and trends. The proposed tool, applied to crime data, can be used as a knowledge discovery tool that can be used to review extremely large datasets and incorporate a vast array of methods for accurate handling of security issues.

The development of the crime analysis tool has four steps, namely, data cleaning, clustering, classification and outlier detection. The data cleaning stage removed unwanted records and predicted missing values. The clustering technique is used to group data according to the different type of crime. From the clustered results it is easy to identify crime trend over years and can be used to design precaution methods for future. The classification of data is mainly used predict future crime trend. The last step is mainly used to identify future crimes that are emerging newly by using outlier detection on crime data.

Experimental results prove that the tool is effective in terms of analysis speed, identifying common crime patterns and future prediction. The developed tool has promising value in the current changing crime scenario and can be used as an effective tool by Indian police and enforcement of law organizations for crime detection and prevention.

REFERENCES

1. Amarnathan, L.C. (2003) Technological Advancement: Implications for Crime, The Indian Police Journal, April June.

2. Abraham, T. and de Vel, O. (2006) Investigative profiling with computer forensic log data and association rules," in Proceedings of the IEEE International Conference on Data Mining (ICDM'02), Pp. 11 – 18.

3. Brown, D.E. (1998) The regional crime analysis program (RECAP): A frame work for mining data to catch criminals," in Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Vol. 3, Pp. 2848-2853.

4. Corcoran J.J., Wilson I.D. AND Ware J.A. (2003) Predicting the geo-temporal variations of crime and disorder, International Journal of Forecasting, Vol. 19, Pp.623–634.

5. David, G. (2006) Globalization and International Security: Have the Rules of the Game Changed?, Annual meeting of the International Studies Association, California, USA, http://www.allacademic.com/meta/ p98627_index.html.

6. de Bruin, J.S. , Cocx, T.K. , Kosters, W.A. , Laros, J. and Kok, J.N. (2006) Data mining approaches to criminal career analysis," in Proceedings of the Sixth International Conference on Data Mining (ICDM'06), Pp. 171-177.

7. Hauck, R.V.Atabakhsh, H., Ongvasith, P., Gupta, H. and Chen, H. (2002) Using Coplink to Analyze Criminal-Justice Data, Computer, Volume 35 Issue 3, Pp. 30-37.

8. Keyvanpour, M.R., Javideh, M. and Ebrahimi, M.R. (2010) Detecting and investigating crime by means of data mining: a general crime matching framework, Procedia Computer Science, World Conference on Information Technology, Elsvier B.V., Vol. 3, Pp. 872-830.

9. Nath, S. (2007) Crime data mining, Advances and innovations in systems, K. Elleithy (ed.), Computing Sciences and Software Engineering, Pp. 405-409.

10. Senator, T.E., Goldberg, H.G., Wooton, J., Cottini, M.A., Khan, A.F.U., Klinger, C.D., Llamas, W.M., Marrone, M.P. and Wong, R.W.H. (1995) The FinCEN Artificial Intelligence System: Identifying Potential Money Laundering from Reports of Large Cash Transactions, AI Magazine, Vol.16, No. 4, Pp. 21-39.

Oracle Real Application Clusters

Deepali Kadam, Nandan Bhalwarkar, Rahul Neware, Rajesh Sapkale, Raunika Lamge

Abstract -This paper shows automatic failover and load balancing for Oracle real Application Clusters. rac enables you to use clustered hardware by running multiple instances against the same database. The database files are stored on disks that are either physically or logically connected to each node, so that every active instance can read from or write to them. Oracle Real Application Clusters manages data access, so that changes are coordinated between the instances and each instance sees a consistent image of the database. The Cluster Interconnect enables instances to pass coordination information and data images between each other.The architecture enables users and application to benefit from the processing power of multiple machines. Oracle RAC architecture also achieves redundancy in the case of, for example, a system crashing or becoming unavailable; the application can still access the database on any surviving instances.

Inexd Terms – Oracle, real application, clusters, rac architecture, clusterware, loadbalancing, server.

1 INTRODUCTION

In the last five years, researchers have shown that commodity clusters have the potential to provide super-computing capabilities at a fraction of the cost of traditional multiprocessor systems . At the same time, much research has been conducted in software distributed shared memory to make it as easy to program clusters as it is to program shared memory multiprocessors.

These advances in programmability and performance are making it possible to use very large clusters as a cost-effective platform for data-intensive, long-running applications. As cluster size and application running times increase, adding failover becomes critical.At the same time, to preserve performance scalable.
In this paper, we address the problem of designing Automatic failover system, specifically targeting scalable system. The novelty of our work lies in the combination of failover techniques with memory-mapped communication, such that the resulting system can be used in very large local-area clusters. Furthermore,A common approach is, Real Application Clusters manages data access, so that changes are coordinated between the instances and each instance sees a consistent image of the database. The **Cluster Interconnect** enables instances to pass coordination information and data images between each other.

Furthermore,A common approach is, Real Application Clusters manages data access, so that changes are coordinated between the instances and each instance sees a consistent image of the database. The **Cluster Interconnect** enables instances to pass coordination information and data images between each   other.
The architecture enables users and application to benefit from the processing power of multiple machines. RAC architecture also achieves redundancy in the case of, for example, a system crashing or becoming unavailable; the application can still access the database on any surviving instances.

_____

- *DEEPALI KADAM Btech computer engineer from*
  VJTI,MUMBAI,INDIA email :deepalisushpriya@gmail.com
- *NANDAN BHALWANKAR Btech computer engineer from*
  VJTI,MUMBAI,INDIA email :nandanbhalwankar123@gmail.com
- *RAHUL NEWARE Btech computer engineer from*
  VJTI,MUMBAI,INDIA email :rahulneware2007@gmail.com
- *RAJESH SAPKALE Btech computer engineer from*
  VJTI,MUMBAI,INDIA email :rajeshsapkale@hotmail.com
- *RAUNIKA LAMGE Btech computer engineer from*
  VJTI,MUMBAI,INDIA email :raunika24@gmail.com
- *PRAMILA CHAWAN professor computer department*
  VJTI,MUMBAI,INDIA email :pmchawan@vjti.org.in
- *BANDU MESHRAM head of computer department*
  VJTI,MUMBAI,INDIA email :bbmeshram@vjti.org.in
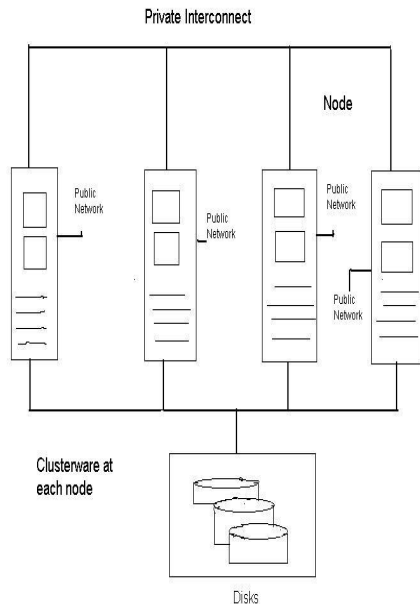
RAC ARCHITECTURE

The diagram consists of two things:

Collection of interconnected Servers, called Clusters connected together using   private interconnect.

The database files storing collection of database tables which is shared   among multiple instances among servers.

A cluster consists of two or more independent, but interconnected, servers. Several hardware vendors have provided cluster capability over the years to meet a variety of needs. Some clusters were intended only to provide high availability by allowing work to be transferred to a secondary node if the active node fails. Others were designed to provide scalability by allowing user connections or work to be distributed across the nodes.

Another common feature of a cluster is that it should appear to an application as if it were a single server. Similarly, management of several servers should be as similar to the management of a single server as possible. The cluster management software provides this transparency. For the nodes to act as if they were a single server, files must be stored in such a way that they can be found by the specific node that needs them. There are several different cluster topologies that address the data access issue, each dependent on the prim goal of the cluster designer. The interconnect is a physical network used as a means of communication between each node of the cluster.

A cluster comprises multiple interconnected computers or servers that appear as if they are one server to end users and applications. Oracle Real Application Clusters

(Oracle RAC) enables you to cluster Oracle databases. Oracle RAC uses **Oracle Clusterware** for the infrastructure to bind multiple servers so they operate as a single system.Oracle Clusterware is a portable cluster management solution that is integrated with the Oracle database. Oracle Clusterware is also a required component for using Oracle RAC. In addition, Oracle Clusterware enables both single-instance Oracle databases and Oracle RAC databases to use the Oracle high-availability infrastructure. Oracle Clusterware enables you to create a clustered pool of storage to be used by any combination of single-instance and Oracle RAC databases.

Oracle Real Application Clusters (RAC) is a software component you can add to a high-availability solution that enables users on multiple machines to access a single database with increased performance. RAC comprises two or more Oracle database instances running on two or more clustered machines and accessing a shared storage device via cluster technology. To support this architecture, the machines that host the database instances are linked by a high-speed interconnect to form the cluster. The interconnect is a physical network used as a means of communication between the nodes of the cluster. Cluster functionality is pro vided by the operating system or compatible third party clustering software.

Oracle RAC offers features in the following areas:

Scalability

Availability

Failover

Load balancing

Oracle RAC is a cluster database with a shared cache architecture that overcomes the limitations of traditional shared-nothing and shared-disk approaches to provide highly scalable and available database solutions for all your business applications. Oracle RAC is a key component of Oracle's enterprise grid architecture. Oracle RAC support is included in the Oracle Database Standard Edition for higher levels of system uptime.

SHARED DISK ARCHITECTURE FOR ORACLE RAC:

A **shared disk file system** uses a storage area network (SAN) or RAID array to provide direct disk access from multiple computers at the block level. Translation from file-level operations that applications use to block-level operations used by the SAN must take place on the client node. The most common type of clustered file system, a shared disk file system adds a mechanism for concurrency control which conventional file systems intended for local storage do not have. Clients are afforded a consistent and serializable view of the file system, avoiding corruption and unintended data loss even when multiple clients try to access the same files at the same time. Shared disk file systems also usually employ some sort of a fencing mechanism to prevent data corruption in

case of node failures.
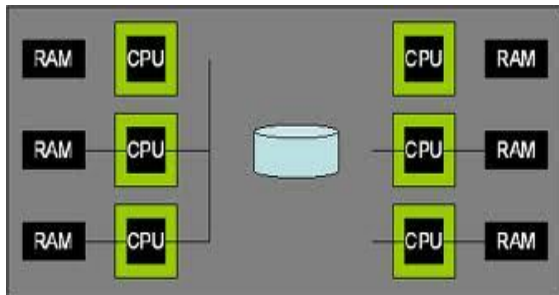
The underlying storage area network might use any of a number of block-level protocols, including SCSI, iSCSI, HyperSCSI, ATA over Ethernet (AoE), Fibre Channel, and InfiniBand.
There are different architectural approaches to a shared disk file system. Some distribute file information across all the servers in a cluster (fully distributed). Others utilize a centralized metadata server. Both achieve the same result of enabling all servers to access all the data on a shared storage device

CLUSTERWARE :
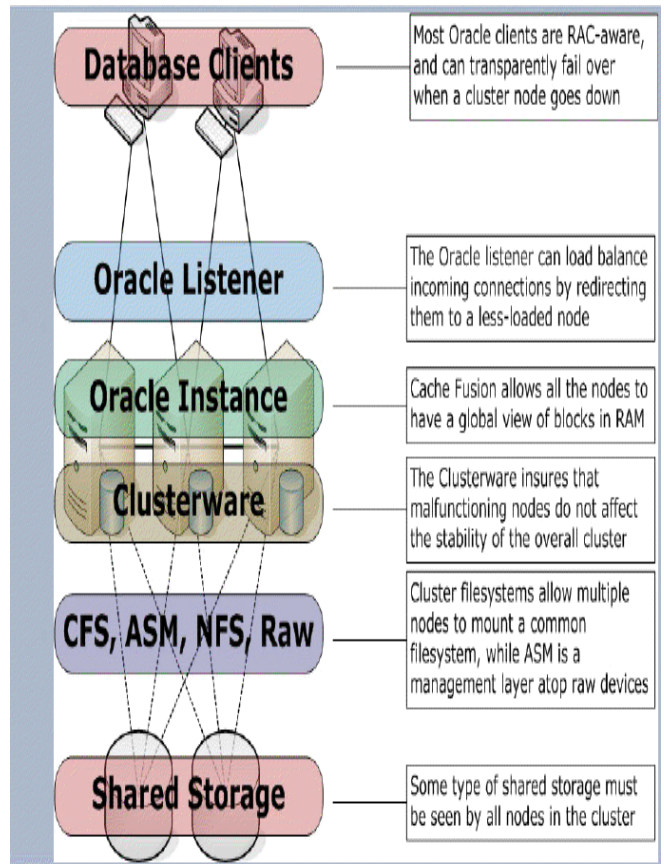
Oracle Clusterware is the software, which enables the nodes to communicate with each other, and forms the cluster and makes the nodes as single logical server. Oracle Clusterware is run by Cluster Ready Services (CRS) using two key components. They are Oracle Cluster Registry (OCR), which records and maintains the cluster and node membership information. The other component is voting disk, which acts a tiebreaker during communication failures. Consistent heartbeat information from all the nodes is sent to voting disk when the cluster is running. CRS service has four components namely OPROCd, CRS Daemon (crsd), Oracle Cluster Synchronization Service Daemon (OCSSD) and Event Volume Manager Daemon (evmd) and each handles a variety of functions. Failure or death of the CRS daemon can cause the node failure and it automatically reboots the nodes to avoid the data corruption because of the possible communication failure between the nodes. The CRS daemon runs as the super user 'root' in the UNIX platforms and runs as a service in the windows platforms.

LOAD BALANCING :

Clusters have become increasingly popular as powerful and cost-effective platforms for executing parallel applications In such systems, load-balancing schemes can improve system performance by attempting to assign work, at run time, to machines with idle or underutilized resources.
The Oracle RAC system can distribute the load over many nodes this feature called as load balancing.

There are two methods of load balancing

1.Client load balancing
2.Server load balancing

Client Load Balancing distributes new connections among Oracle RAC nodes so that no one server is over-loaded with connection requests and it is configured at net service name level by providing multiple descriptions in a description list or multiple addresses in an address list. For example, if connection fails over to another node in case of failure, the client load balancing ensures that the redirected connections are distributed among other nodes in the RAC.

Server Load Balancing distributes processing workload among Oracle RAC nodes. It divides the connection load evenly between all available listeners and distributes new user session connection requests to the least loaded listener(s) based on the total number of sessions which are already connected. Each listener communicates with the other listener(s) via each database instance's PMON process.
A farm of servers with the same function is the base of a load balancing cluster. To distribute the user requests to several nodes, a load balancer is useful. The load balancer checks the utilization of all nodes. The node with the estimated best performance will get the next user request. This algorithm ensures the best performance available at the time is given to the users.
A very important point is the quality of the load balancer. In this case, quality means the opportunity of the system to made a highly qualified forecast about which node will offer the best performance concerning an individual application. Another aspect to make the decision is to re-connect the client to an existing session. Without any kind of protection, the load balancer could be a single point of failure.
If one node is out of service, then the whole system will still work. The load balancer will recognize the failed node, and mark the crashed system. The total performance of the load balancing cluster will be reduced, but services will still be provided.

AUTOMATIC FAILOVER:

Automatic failover is supported in database for high-safety mode. In high-safety mode with automatic failover, once the database is synchronized, if the principal server becomes unavailable, an automatic failover occurs. An automatic failover causes the secondary server to take over the role of principal server and bring its copy of the database to the user. Requiring that the server be synchronized prevents loss to the user during failover, because every transaction committed on the principal server can also be committed on the secondary server.

Automatic failover requires the following conditions:

1. The secondary server must be running in high-safety mode
2. The secondary server must have access to the main database

How Automatic Failover Works

Under the preceding conditions, automatic failover initiates the following sequence of actions:

1. If the principal server fails, it changes the state of the principal server to DISCONNECTED and disconnects all clients from the principal server.
2. The secondary server register that the principal server is unavailable.
3. All clients from the principal server are shifted to the secondary server by clusterware.

CONCLUSIONS:

In this paper we have shown automatic failover and load balncing for oracle real Application Clusters. an oracle rac system can protect against computer failures caused by unexpected hardware failures and operating system or server crashes, as well as processing loss caused by planned maintenance.. When node failover occurs and a service connection is redirected to another node, users can continue to access the service, unaware that it is now provided from a different node. We have also explained the architecture of oracle real application cluster,oracle clusterware,shared disk architecture.
.

REFERENCES

[1] T.E. Anderson, D.E. Culler, D.A. Patterson. A Case
for NOW (Networks of Workstations). *IEEE Micro*, v.
15, no. 1, pp. 54-64, February 1995.
[2] R. Bianchini, L. I. Kontothanassis, R. Pinto, M. De
Maria, M. Abud, C.L. Amorim. Hiding Communication
Latency and Coherence Overhead in Software
DSMs. *Proc. 7th International Conference on Architectural
Support for Programming Languages and Operating
Systems (ASPLOS)*, October 1996.
[3] N. J. Boden, D. Cohen, R. E. Felderman, A. E. Kulawik,
C. L. Seitz, J. N. Seizovic, W. Su. Myrinet: A

Face Recognition System Based on Principal Component Analysis (PCA) with Back Propagation Neural Networks (BPNN)

Mohammod Abul Kashem, Md. Nasim Akhter, Shamim Ahmed, and Md. Mahbub Alam

Abstract— Face recognition has received substantial attention from researches in biometrics, pattern recognition field and computer vision communities. Face recognition can be applied in Security measure at Air ports, Passport verification, Criminals list verification in police department, Visa processing , Verification of Electoral identification and Card Security measure at ATM's**.** In this paper, a face recognition system for personal identification and verification using Principal Component Analysis (PCA) with Back Propagation Neural Networks (BPNN) is proposed. This system consists on three basic steps which are automatically detect human face image using BPNN, the various facial features extraction, and face recognition are performed based on Principal Component Analysis (PCA) with BPNN. The dimensionality of face image is reduced by the PCA and the recognition is done by the BPNN for efficient and robust face recognition. In this paper also focuses on the face database with different sources of variations, especially Pose, Expression, Accessories, Lighting and backgrounds would be used to advance the state-of-the-art face recognition technologies aiming at practical applications.

Index Terms— Face Detection, Facial Features Extraction, Face Database, Face Recognition, Increase Acceptance ratio and Reduce Execution Time.

————————————— ◆ —————————————

1 INTRODUCTION

WITHIN computer vision, face recognition has become increasingly relevant in today's society. The recent interest in face recognition can be attributed to the increase of commercial interest and the development of feasible technologies to support the development of face recognition. Major areas of commercial interest include biometrics, law enforcement and surveillance, smart cards, and access control. Unlike other forms of identification such as fingerprint analysis and iris scans, face recognition is user-friendly and non-intrusive. Possible scenarios of face recog-

————————————————

- **Mohammad Abul Kashem** *has been serving as an Associate Professor and Head of the Department, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. Field of interest: Speech Signal Processing. E-mail:* **drkashem11@duet.ac.bd**.
- **Md. Nasim Akhter** *has been serving as an Assistant Professor, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. Field of interest: Operating System, Data Communication. E-mail:* **nasimntu@yahoo.com**.
- **Shamim Ahmed** *has been studying as an M.Sc. in Engineering Student, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. He got B.Sc. in engineering degree in CSE in the year of 2010 from DUET, Gazipur, Bangladesh. He joined at Dhaka International University (DIU) as lecturer (part time), in the department of Computer Science and Engineering (CSE) in January 2011 after completion his B.Sc. in Engg. (CSE). Field of interest: Digital Image Processing, Artificial Neural Network, Artificial Intelligence & Visual Effects. E-mail:* **shamim.6feb@gmail.com**.
- **Md. Mahbub Alam** *has been serving as a Lecturer, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. He got B.Sc. in engineering degree in CSE in the year of 2009 from DUET, Gazipur, Bangladesh. Field of interest: Digital Image Processing, Neural Network, Artificial Intelligence. E-mail:* **emahbub.cse@gmail.com**.

nition include: identification at front door for home security, recognition at ATM or in conjunction with a smart card for authentication, video surveillance for security. With the advent of electronic medium, especially computer, society is increasingly dependent on computer for processing, storage and transmission of information. Computer plays an important role in every parts of today life and society in modern civilization. With increasing technology, man becomes involved with computer as the leader of this technological age and the technological revolution has taken place all over the world based on it. It has opened a new age for humankind to enter into a new world, commonly known as the technological world. Computer vision is a part of every day life. One of the most important goals of computer vision is to achieve visual recognition ability comparable to that of human [1],[2],[3].

Face recognition has received substantial attention from researches in biometrics, pattern recognition field and computer vision communities. In this paper we proposed a computational model of face detection and recognition, which is fast, reasonably simple, and accurate in constrained environments such as an office or a household. Face recognition using Eigen faces has been shown to be accurate and fast. When BPNN technique is combined with PCA, non-linear face images can be recognized easily.

2 OUTLINE OF THE SYSTEM

In this papers to design and implementation of the Face Recognition System (FRS) can be subdivided into three

main parts. The first part is face detection-automatically face detection can be accomplished by using neural networks back propagation. The second part is to perform various facial features extraction from face image using digital image processing and Principal Component Analysis (PCA). And the third part consists of the artificial intelligence (face recognition) which is accomplished by Back Propagation Neural Network (BPNN).

The first part is the Neural Network-based Face Detection described in [4]. The basic goal is to study, implement, train and test the Neural Network-based machine learning system. Given as input an arbitrary image, which could be a digitized video signal or a scanned photograph, determine whether or not there are any human faces in the image, and if there are, return an encoding of the location and spatial extent of each human face in the image. The first stage in face detection is to perform skin detection. Skin detection can be performed in a number of color models. To name a few are RGB, YCbCr, HSV, YIQ, YUV, CIE, XYZ, etc. An efficient skin detection algorithm is one which should be able to cover all the skin colors like black, brown, white, etc. and should account for varying lighting conditions. Experiments were performed in YIQ and YCbCr color models to find out the robust skin color model. This part consists of YIQ and YCbCr color model, skin detection, blob detection, smooth the face, image scaling.

Fig: 3. (a) skin detection, and (b) face detection.   The second part is to perform various facial features extraction from face image using digital image processing and Principal Component Analysis (PCA) and the Back Propagation Neural Network (BPNN). We separately used iris recognition for facial feature extraction. Facial feature extraction consists in localizing the most characteristic face components (eyes, nose, mouth, etc.) within images that depict human faces. This step is essential for the initialization of many face processing techniques like face tracking, facial expression recognition or face recognition. Among these, face recognition is a lively research area where it has been made a great effort in the last years to design and compare different techniques. The second part consists of face landmarks, iris recognition, fiducial points.

The third part consists of the artificial intelligence (face recognition) which is accomplished by Back Propagation Neural Network (BPNN). This paper gives a Neural and PCA based algorithm for efficient and robust face recognition. This is based on principal component-analysis (PCA) technique, which is used to simplify a dataset into lower dimension while retaining the characteristics of dataset. Pre-processing, Principal component analysis and Back Propagation Neural Algorithm are the major implementations of this paper.

This papers also focuses on the face database with different sources of variations, especially Pose, Expression, Accessories, and Lighting would be used to advance the state-of-the-art face recognition technologies aiming at practical applications especially for the oriental.

3   FACE DETECTION

The face detection can be perform by given as input an arbitrary image, which could be a digitized video signal or a scanned photograph, determine whether or not there are any human faces in the image, and if there are, return an encoding of the location and spatial extent of each human face in the image[5].

3.1 The YIQ & TCbCr color model for skin detection

The first stage in face detection is to perform skin detection. Skin detection can be performed in a number of color models. To name a few are RGB, YCbCr, HSV, YIQ, YUV, CIE, XYZ, etc. An efficient skin detection algorithm is one which should be able to cover all the skin colors like black, brown, white, etc. and should account for varying lighting conditions. Experiments were performed in YIQ and YCbCr color models to find out the robust skin color model.

a   b   c
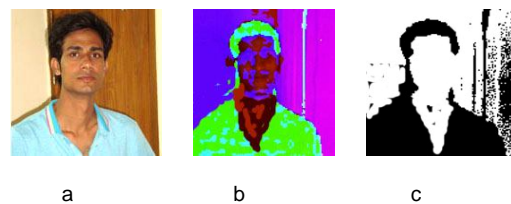
Fig. 1. (a) RGB, (b) RGB to YIQ, and (c) Skin threshold in YIQ.

a   b

Fig. 2. (a) RGB to YCbCr, and (b) Skin threshold in

a   b
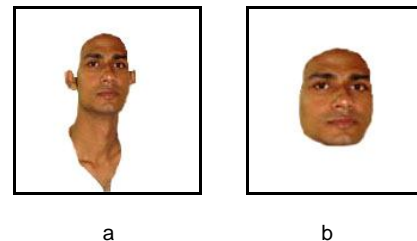
Fig. 3. (a) skin detection, and (b) face detection.

3.2 Blob detection and Smooth the face

We used an open GL blob detection library. This library designed for finding 'blobs' in an image, i.e. areas whose luminosity is above or below a particular value. It computes their edges and their bounding box. This library does not perform blob tracking; it only tries to find all blobs in each frame it was fed with. Blobs in the image which are elliptical in shape are detected as faces. The blob detection algorithm draws a rectangle around those blobs by calculating information such as position and center. After the previous steps, the above face would be a possible outcome. When the face is zoomed in, it turns out the outline of the face is

not smooth. So the next step is to smooth the outline of the face.

a   b
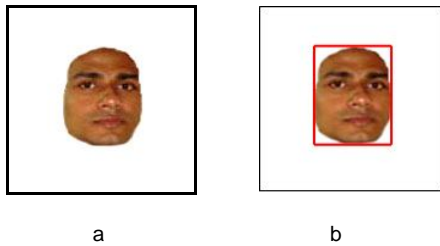
Fig. 4. (a) Blob detection, and (b) Smooth the face image.

3.3 Image Scaling.

According to the client's requirement, the image is to be scaled to the size of 80*80 pixels with the face centred. The face should contain 3350 pixels and all the rest of the pixels are white. Some edge detection algorithms cannot be applied to color images, so it is also necessary to convert the image to grey scale.

There are four steps in this stage:

1. Scaling the face to the number of pixels which is most approximate to and greater than 3350.
2. Making the number of pixels of the face exactly equal to 3350.
3. Making the size of the image 80*80 pixels.
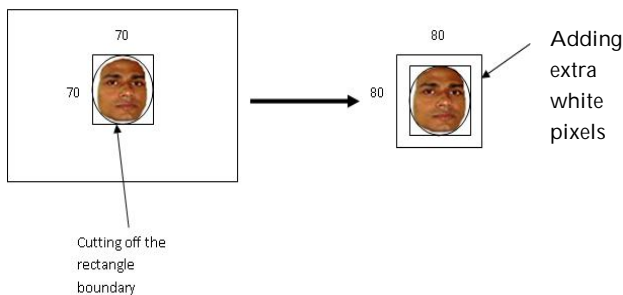
Fig. 5. Making the size of the image 80*80 pixels.

4. Converting the image to grey scale.

3.4 Face Detector Algorithms

Training Data Preparation:
- For each face and non-face image:
  o Subtract out an approximation of the shading plane to correct for single light source effects.
  o Rescale histogram so that every image has the same gray level range.
- Aggregate data into data sets.

Backpropagation Neural Network.
- Set all weight to random value range from -1.0 to 1.0.
- Set an input pattern (binary values) to the neurons of the net's input layer.
- Active each neuron of the following layer:

  o Multiply the weight values of the connections leading to this neuron with the output values of the preceding neurons.
  o Add up these values.
  o Pass the result to an activation function, which computes the output value of this neuron.
- Repeat this until the output layer is reached.
- Compare the calculated output pattern to the desired target pattern and compute a square error value.
- Change all weights values of each weight using the formula:

  Weight (old) + Learning Rate * Output Error * Output (Neuron i) * Output (Neuron i + 1) * (1 – Output (Neuron i + 1))
- Go to the first step.
- The algorithm end, if all output pattern match their target pattern.

Apply Face Detector to Image:
- Apply the 20 x 20 pixel view window at every pixel position in the input image.
- For each window region:
  o Apply linear fit function and histogram equalization function on the region.
  o Pass the region to the trained Neural Network to decide whether or not it is a face.
  o Return a face rectangle box scaled by the scale factor, if the region is detected as a face.
- Scale the image down by a factor of 1.2.
- Go to the first step, if the image is larger than the 20 x 20 pixel window.

4 FACIAL FEATURE EXTRACTION

The part is to perform various facial features extraction from face image using digital image processing and Principal Component Analysis (PCA). We separately used iris recognition for facial feature extraction. Facial feature extraction consists in localizing the most characteristic face components (eyes, nose, mouth, etc.) within images that depict human faces. This step is essential for the initialization of many face processing techniques like face tracking, facial expression recognition or face recognition. Among these, face recognition is a lively research area where it has been made a great effort in the last years to design and compare different techniques.

4.1 Face Landmarks (Nodal points)

Facial features can be extracted according to various face landmarks on human face. Every face has numerous, distinguishable **landmarks**, the different peaks and valleys that make up facial features. It defines these landmarks as **nodal points**. Each human face has approximately 80 nodal points. Some of these measured by the software are:

1. Distance between the eyes.
2. Width of the nose.
3. Depth of the eye sockets.
4. The shape of the cheekbones.
5. The length of the jaw line.
6. Height & Width of forehead and total face.
7. Lip height.

Fig. 6. Various Face Landmarks (nodal points).

8. Lip width.
9. Distance between nose & mouth.
10. Face skin marks, etc.
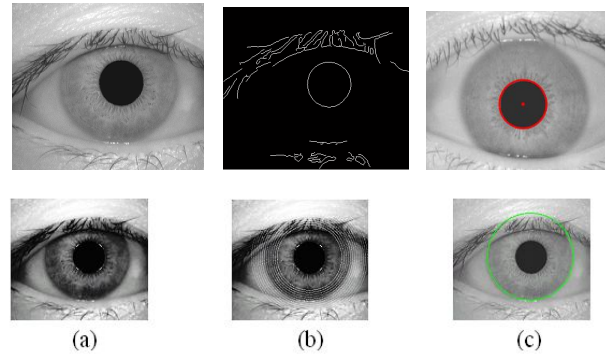
Fig. 7. Steps involved in detection of inner pupil boundary and outer iris localization.
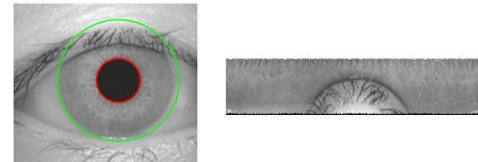
Fig. 8. Iris normalization

4.2 Iris Recognition for facial feature extraction.

The iris is an externally visible, yet protected organ whose unique epigenetic pattern remains stable throughout adult life. These characteristics make it very attractive for use as a biometric for identifying individuals.

4.3 Image Acquisition

The iris image should be rich in iris texture as the feature extraction stage depends upon the image quality. Thus, the image is acquired by 3CCD camera placed at a distance of approximately 9 cm from the user eye. The approximate distance between the user and the source of light is about 12 cm.

4.4 Iris Localization

The acquired iris image has to be preprocessed to detect the iris, which is an annular portion between the pupil (inner boundary) and the sclera (outer boundary). The first step in iris localization is to detect pupil which is the black circular part surrounded by iris tissues. The center of pupil can be used to detect the outer radius of iris patterns. The important steps involved are:

1. Pupil detection.
2. Outer iris localization.

4.5 Iris Normalization

Localizing iris from an image delineates the annular portion from the rest of the image. The concept of rubber sheet modal suggested by Daugman takes into consideration the possibility of pupil dilation and appearing of different size in different images.

4.6 Feature Extraction

Corners in the normalized iris image can be used to extract features for distinguishing two iris images. The steps involved in corner detection algorithm are as follows

S1: The normalized iris image is used to detect corners using covariance matrix

S2: The detected corners between the database and query image are used to find cross correlation coefficient

S3: If the number of correlation coefficients between the

$$M_{cv} = \begin{bmatrix} \sum D_x^2 & \sum D_x D_y \\ \sum D_x D_y & \sum D_y^2 \end{bmatrix}$$
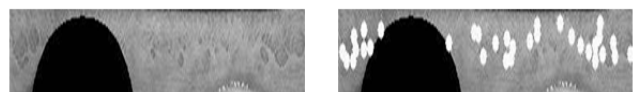
Fig. 9. Detection of corners.

detected corners of the two images is greater than a threshold value then the candidate is accepted by the system

4.7 Corner detection

Corner points can be detected from the normalized iris image using covariance matrix of change in intensity at each point. A 3x3 window centered on point p is considered to find covariance matrix M_{cv}
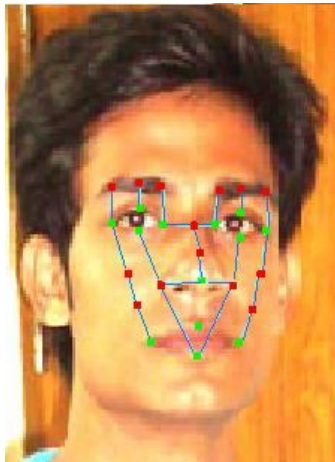
Fig. 10. A face is described by 27 fiducial points: 13 are directly extracted from the image (in green), 14 are inferred from the former ones (in red).

4.8 From eye centers to fiducial points

In this section we show how, given the eye centers, we derive a set of 27 characteristic points (*fiducial points*): three points on each eyebrow, the tip, the lateral extremes and the vertical mid-point of the nose, the eye and lip corners, their upper and lower mid-points, the midpoint between the two eyes, and four points on the cheeks (see Fig: 10).

4.9 Nose

The nose is characterized by very simple and generic prop-
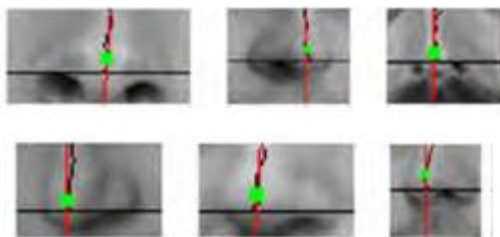
Fig. 11. Examples of nose processing. The black horizontal line indicates the nose base; the black dots along the nose are the points of maximal symmetry along each row; the red line is the vertical axis approximating those points; the green marker indicates the nose tip.

erties: the nose has a "base" the gray levels of which contrast significantly with the neighboring regions; moreover, the nose profile can be characterized as the set of points

with the highest symmetry and high luminance values; therefore we can identify the nose tip as the point that lies on the nose profile, above the nose baseline, and that corresponds to the brightest gray level. These considerations allow to localize the nose tip robustly (see Figure: 11).

4.10 Mouth

Regarding the mouth, our goal is to locate its corners and its upper and lower mid-points. To this aim, we use a snake [Hamarneh, 2000] to determine the entire contour since we verified that they can robustly describe the very different shapes that mouths can assume. To make the snake con-
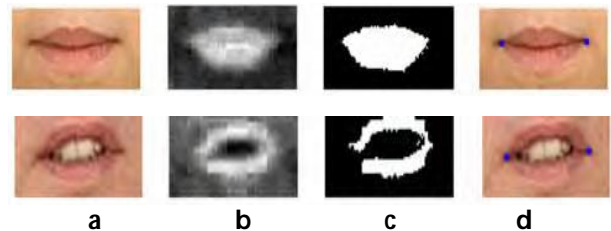
Fig. 12 Mouth corners estimation: a) mouth subimage b) mouth map c) binarized mouth map d) mouth corners.
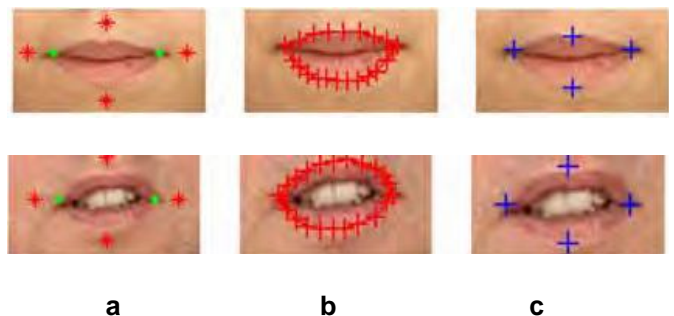
Fig. 13. Snake evolution: a) snake initialization b) final snake position c) mouth fiducial Points.

verge, its initialization is fundamental; therefore the algorithm estimates the mouth corners and anchors the snake to them: first, we represent the mouth subimage in the *YCbCr* color space, and we apply the following transformation:

$$MM = (255 - (Cr - Cb))\, Cr\, 2$$

MM is a mouth map that highlights the region corresponding to the lips; *MM* is then binarized putting to 1 the 20% of its highest values; the mouth corners are determined taking the most lateral extremes (see "Fig 12").

5 FACE DATABASE

Our ace database contains large-scale face images with different sources of variations, especially Pose, Expression, Accessories, and Lighting would be used to advance the

state-of-the-art face recognition technologies aiming at practical applications especially for the oriental. Our face database contains 99,594 images of 1040 individuals (595 males

Fig. 14. Different kinds of poses.

and 445 females) with varying Pose, Expression, Accessory, and Lighting.

5.1 Poses

In our face database we consider various kinds of poses such as front pose, left pose, right pose, left corner pose, righter corner pose, left up pose, right up pose, front up

| Normal | Happiness | Anger | Disgust |

| Sadness | Fear | Surprise | Eyes closed |

Fig. 15. Different kinds of expressions.

pose, front down pose, left down pose, and right down pose.

5.2 Facial expression

A database of facial expression images was collected. Ten expressors posed 3 or 4 examples of each of the six basic facial expressions (happiness, sadness, surprise, anger, disgust, fear) and a neutral face for a total of 219 images of facial expressions.

5.3 Varying lighting conditions

The various lighting conditions are effects the facial images. Images with varying lighting conditions are recommended

Fig. 16. Face images with varying lighting conditions.

for the purpose of image processing and face recognition under natural illumination. It is recommended to store facial images in the face database with varying lighting conditions.

Fig. 17. Face images with different kinds of accessories.

5.4 Accessories: Glasses and Caps

Several kinds of glasses and hats are prepared in the room used as accessories to further increase the diversity of the database. The glasses consisted of dark frame glasses, thin and white frame glasses, glasses without frame. The hats also have brims of different size and shape.

5.5 Backgrounds

Without special statement, we are capturing face images with a blue cloth as the default background. However, in practical applications, many cameras are working under the auto-white balance mode, which may change the face appearance much. Therefore, it is necessary to mimic this situation in the database. We just consider the cases when the background color has been changed. Concretely, five sheets of cloth with five different unicolors (blue, white, black, red and yellow) are used.

6 FACE RECOGNITION

This part consists of the artificial intelligence (face recognition) which is accomplished by Principal Component Analysis (PCA) with Back Propagation Neural Network (BPNN). This paper gives a Neural and PCA based algorithm for efficient and robust face recognition. A face recognition system [11] is a computer vision and it automatically identifies a human face from database images. The face recognition problem is challenging as it needs to account for all possible appearance variation caused by change in illumination, facial features, occlusions, etc. This is based on principal component-analysis (PCA) technique, which is used to simplify a dataset into lower dimension while retaining the characteristics of dataset. Pre-processing, Principal component analysis and Back Propagation Neural Algorithm are the major implementations of this paper. Pre-processing is done for two purposes

1. To reduce noise and possible convolute effects of interfering system,
2. To transform the image into a different space where classification may prove easier by exploitation of certain features.

PCA is a common statistical technique for finding the patterns in high dimensional data's [6]. Feature extraction, also called Dimensionality Reduction, is done by PCA for a three main purposes like

1. To reduce dimension of the data to more tractable limits
2. To capture salient class-specific features of the data,
3. To eliminate redundancy.

6.1 Experimentation and Results

When BPNN technique is combined with PCA, non linear face images can be recognized easily. One of the images as

a   b   c

Fig. 20. (a) Input Image , (b) Recognized Image by BPNN, (c) Recognized Image by PCA method.

shown in fig 20 (a) is taken as the Input image. The Recognized Image by BPNN and reconstructed output image by PCA is as shown in fig 20 (b) and 20 (c).
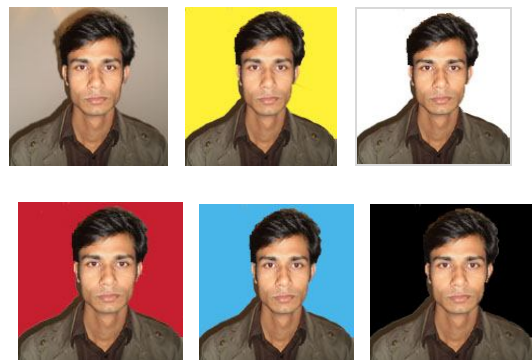
Fig. 18. Face images with different kinds backgrounds.

Fig. 19. Outline of Face Recognition System by using PCA & Back-propagation Neural Network.

Table1 shows the comparison of acceptance ratio and execu-

TABLE 1
COMPARISON OF ACCEPTANCE RATIO AND EXECUTION TIME FOR DATABASE IMAGES.

No .of Images	Acceptance ratio (%)		Execution Time (Seconds)	
	PCA	PCA with BPNN	PCA	PCA with BPNN
40	92.4	96.5	38	36
80	90.6	94.3	46	43
120	87.9	92.8	55	50
160	85.7	90.2	67	58
200	83.5	87.1	74	67

tion time values for 40, 80, 120,160 and 200 images of Yale database. Graphical analysis of the same is as shown in "Fig 22."

TABLE 2
USING BPNN THE ACCEPTANCE RATIO AND EXECUTION TIME FOR DATABASE IMAGES

No .of Images	Acceptance ratio (%)	Execution Time (Seconds)
	BPNN	BPNN
40	94.2	37
80	92.3	45
120	90.6	52
160	87.9	65
200	85.5	71
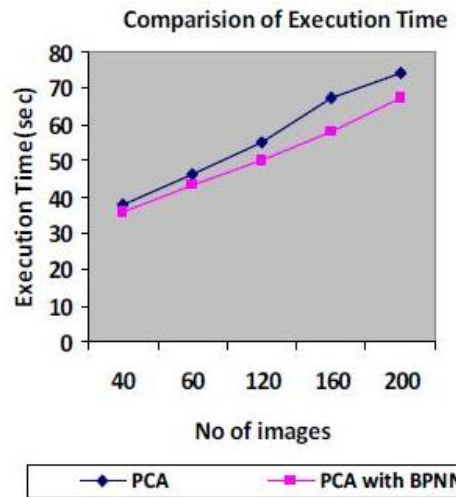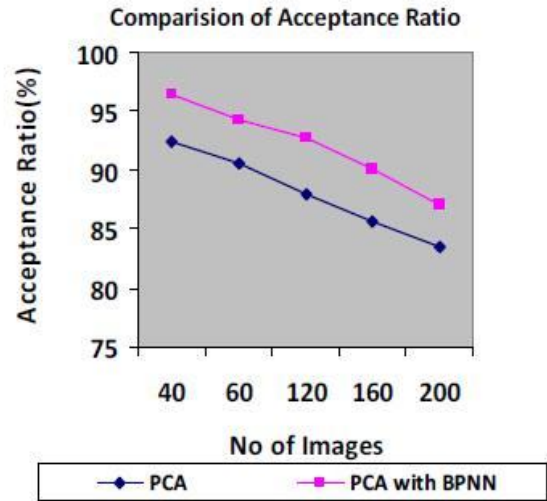
Fig. 22. comparison of Acceptance ratio and execution time.

a   b   c
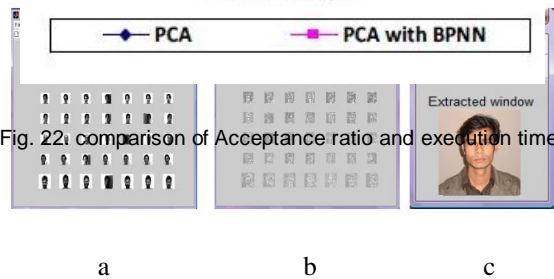
Fig. 21. (a) Training set, (b) Eigen faces , (c) Recognized Image by PCA with BPNN method.

7 CONCLUSION

In this paper, Face recognition using Eigen faces has been shown to be accurate and fast. When BPNN technique is combined with PCA, non linear face images can be recognized easily. Hence it is concluded that this method has the acceptance ratio is more than 90 % and execution time of only few seconds. Face recognition can be applied in Security measure at Air ports, Passport verification, Criminals list verification in police department, Visa processing , Verification of Electoral identification and Card Security measure at ATM's. Face recognition has received substantial attention from researches in biometrics, pattern recognition field and computer vision communities. In this paper we proposed a computational model of face detection and recognition, which is fast, reasonably simple, and accurate in constrained environments such as an office or a household.

REFERENCES

[1] Jain, *Fundamentals of Digital Image Processing, Prentice-Hall Inc., 1982.*

[2] E. Trucco, and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall Inc., 1998.

[3] L. G. Shapiro, and G. C. Stockman, *Computer Vision*, Prentice-Hall Inc., 2001.

[4] Rowley, H., Baluja, S. and Kanade, T., Neural Network-Based Face Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, January, 1998, pp. 23-38. http://www.ri.cmu.edu/pubs/pub_926_text.html

[5] Duda, R.O., Hart, P.E. and Stork, D.G. Pattern Classification. Wiley, New York, 2001.

[6] B.K.Gunturk,A.U.Batur, and Y.Altunbasak,(2003) "Eigenface-domain super-resolution for face recognition," IEEE Transactions of . Image Processing. vol.12, no.5.pp. 597-606.

[7] M.A.Turk and A.P.Petland, (1991) "Eigenfaces for Recognition," Journal of Cognitive Neuroscience. vol. 3, pp.71-86.

[8] T.Yahagi and H.Takano,(1994) "Face Recognition using neural networks with multiple
combinations of categories," International Journal of Electronics Information and Communication Engineering., vol.J77-D-II, no.11, pp.2151-2159.

[9] S.Lawrence, C.L.Giles, A.C.Tsoi, and A.d.Back, (1993) "IEEE Transactions of Neural Networks. vol.8, no.1, pp.98-113.

[10] C.M.Bishop,(1995) "Neural Networks for Pattern Recognition" London, U.K.:Oxford
University Press.

[11] Kailash J. Karande Sanjay N. Talbar "Independent Component Analysis of Edge
Information for Face Recognition" International Journal of Image Processing Volume (3) : Issue (3) pp: 120 -131.

[12] FernandoDe La Torre, Michael J.Black 2003 Internatioal Conference on Computer Vision (ICCV'2001), Vancouver, Canada, July 2001. IEEE 2001

[13] Turk and Pentland, Face Recognition Using Eigenfaces, Method Eigenfaces", IEEE CH2983-5/91, pp 586-591.

[14] Douglas Lyon. Image Processing in Java, Prentice Hall, Upper Saddle River, NJ. 1998.

[15] MATTHEW T. RUBINO, EDGE DETECTION ALGORITHMS,
<HTTP://WWW.CCS.NEU.EDU/HOME/MTRUBS/HTML/EDGEDETECTION.HTML >

[16] H. Schneiderman, "Learning Statistical Structure for Object Detection", *Computer Analysis of Images and Patterns* (CAIP), 2003, Springer-Verlag, August, 2003.

[17] Yang and Huang 1994. "Human face detection in a complex background." Pattern Recognition, Vol 27, pp53-63

[18] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In CVPR, 2001, <http://citeseer.nj.nec.com/viola01rapid.html>

[19] Angela Jarvis, < http://www.forensic-evidence.com/site/ID/facialrecog.html>

[20] Konrad Rzeszutek, <http://darnok.com/projects/face-recognition>

Mohammad Abul Kashem has been serving as an Associate Professor and Head of the Department, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. Field of interest: Speech Signal Processing. E-mail: **drkashem11@duet.ac.bd**.

Md. Nasim Akhter has been serving as an Assistant Professor, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. Field of interest: Operating System, Data Communication. E-mail: **nasimntu@yahoo.com**.

Shamim Ahmed has been studying as an M.Sc. in Engineering Student, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. He got B.Sc. in engineering degree in CSE in the year of 2010 from DUET, Gazipur, Bangladesh. He joined at Dhaka International University (DIU) as lecturer (part time), in the department of Computer Science and Engineering (CSE) in January 2011 after completion his B.Sc. in Engg. (CSE). Right now he teaches courses on Data Communication and Computer Networking, Object Oriented Programming using C++ and Java, Artificial Intelligence, Formal Language and Automata Theory. Field of interest: Digital Image Processing, Ar-

tificial Neural Network, Artificial Intelligence & Visual Effects. E-mail: **shamim.6feb@gmail.com**.

Md. Mahbub Alam has been serving as a Lecturer, Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur, Bangladesh. He got B.Sc. in engineering degree in CSE in the year of 2009 from DUET, Gazipur, Bangladesh. Field of interest: Digital Image Processing, Neural Network, Artificial Intelligence. E-mail: **emahbub.cse@gmail.com**.

Improved Password Selection Method to Prevent Data Thefts

Akash Mathur

Abstract— Password thefts have become an important data security issue in the recent times. Hackers have been intruding into the websites and personal mail accounts of the users. Prevention from this unauthorized access is the need of the hour. So here lies the major key to overcome such a situation and that is to adopt more secure and advanced password selection methods. My paper throws light on one of the most secure methods to improve password strength and security.

Index Terms— Password selection, Data security, Prevention data thefts, Cyber security, Password, Prevention from hacks.

——————————————— ◆ ———————————————

1 INTRODUCTION

THIS is the era of globalization. Every part of the world is potentially connected to the other part. The data travelling has become so easy that it takes only a few seconds for the exchange of media. But with the advancement of the worldwide internet the security of the data is an important issue. How secure is the data is an important issue. Data threats are now common, password hacks being the most common ones. Severe vulnerable attacks on websites had leaded its users to suffer a huge discomfort. Majority of steps are now employed to protect data from theft but how effective are they is an important issue. E-mail websites have been majorly affected by the improper password selection and mismanagement problem. This has resulted in unauthorized entry to the user's personal account. The password strength also is not capable enough to counter the security issues.

Since my main interest has been the web securities I decided to focus on the password selection process.

The main aim of this paper is to lay stress on the security issues related to password selection and management. The idea is to provide the users of the websites an option to choose the password along with their font color.

The passwords must not only comply with at least 6 characters scheme but also color coding pattern can prove to be beneficial. Basic colors including black, red, and white can be adopted to double protect the password. The main idea is that the user not only types the password in the password field box but also selects basic colors of the fonts too used in the password. The ASCII codes used in the password field should be merged with the color combination. The website designers should employ coding to merge the text passwords with font colors to reduce security threats and identity thefts. The credit card owners can be saved from frauds which are caused due to unauthorized access by the stealing computer programs. This unauthorized access can be prevented by using colored alphanumeric passwords. The basic colors of red, black and white won't increase the size of the website and would also not slow down the identification process on the websites.

ADVANTAGES

The advantages of using color implemented password fields are:

- With HTML coding, the password box must also support colored fonts i.e., color coding [1] for better security. There should be a blend of both characters and colors when used as a password.

- Virtual keyboard [2] can also be made into use by making it color adaptable. Many times virtual keyboard helps to protect the accounts from hacking programs. Therefore virtual keyboard should also be designed which support colored fonts when typing password in the password field.

- Hacking programs [3] which record the keyboard hits also known as key loggers [4] also fail since color combination requires mouse clicks and visualization eye of the user. Therefore once again the login process has high security especially E-mail hacking [5] will be reduced to great extent.

- Hacking issues can be minimized to a great extent. Even if multiple hacking programs run, but it is only the eye of the user which can see the color of the password font. So now users will be safer in terms of password management.

- Guessing attacks are also reduced, since guessing colored font password seems to be impossible

COMPLICATIONS

The complications accompanying color implemented password scheme are:

- The size of the website may increase. Website

size becomes heavy on using colored fonts in password fields.

- User interface may become unfriendly. Users might take some time to adapt to this pattern.

- Website hosting price increases with the increase in no. of bytes. If the websites increase in size using this technique, then the cost of hosting would definitely increase.

- One of the shortcomings can be that the identification process of the websites [6] can be slow during the login process.

MEASURES ADOPTABLE TO OVERCOME COMPLICATIONS

The following are the measures which can be adopted to overcome the complications:

- Use of only basic 3 colors red, black and white won't increase much size of the website. Since use of more than these colors would increase the size and consequently the website host prices.

- Graphically adaptable user interface can be used without increasing much of the website size. Skilled website programmers can make the user interface more convenient and user friendly without slowing down the processing of the websites.

- Website hosting prices increases but the security issue solves to a great extent. Data is now more secured with this color coding technique. Even the key loggers fail to track the typed password

- Website's identification process won't be very slow since only basic colors like red, black and white are used as colors for fonts which won't increase much the size of login page and will offer higher security to the users.

BASIC IDEA INTERPRETATION

The following figure correctly interprets the situation after the color coding pattern for password field is adopted and embedded in the websites.

Figure 1. Login format after use of technique

Assume the figure above as the login page of a website then, various color combinations can be used to set the final password.

CONCLUSION

To secure the data on the web, implementation of such a system of colored font passwords is an essence. Use of at least 6 text or alphanumeric characters with colored fonts as password in the password field makes its strength increase many folds. Protection of user accounts like that of E-mail and credit cards from hackers are some of the advantages of using such system. Prevention from identity thefts is also an added advantage. These color combinations technique provides highly secured gateway for the exchange of data on the internet. Although this might seem to be a complicated password selection procedure but is highly effective and secured method, once adopted and embedded with the websites. Even the hacking programs fail when such a system of password selection and management procedure is implemented.

ACKNOWLEDGMENT

REFERENCES

[1] L Cottrell, L M. Cottrell, "*HTML & XHTML Demystified*", USA: The McGraw Hill Companies, pp. 50-56, 2011.

[2] M Hirose, "*Human-Computer Interaction INTERACT '01,*" *Netherlands: IOS Press,* pp. 678-679, 2001.

[3] S Mclure, J Scambray, G Kurtz, "*Hacking Exposed*". 5th Ed. USA: McGraw-Hill/Osborne, 2005.

[4] R.C Newman, "*Computer Security: Protecting Digital Resources*", Jones and Bartletts publishers, pp. 58-59, 2009.

[5] A. Fadia, "*Email Hacking*". 2nd Ed. India: Vikas publishing house pvt ltd, 2009.

[6] M Kimwele, W M Wangi, S Kimani, "Strengths of a colored graphical Password Scheme," *International Journal of Reviews in Computing*, vol 4, pp. 64-65, Oct. 2010.

A DKIM based Architecture for Combating Good Word Attack in Statistical Spam Filters

Kashefa Kowser.K, Saruladha.K, Packiavathy.M

Abstract— Abuse of E-Mail by unwanted users causes an exponential increase of E-Mails in user mailboxes which is known as Spam. It is an unsolicited commercial E-mail or unsolicited bulk E-Mail produces huge economic loss to large scale organizations due to high network bandwidth consumption and heavy mail server processing overload. Statistical spam filters could be used to categorize incoming E-Mails into legitimate and spam but they are vulnerable to Good Word attack which obfuscates "good words" in spam messages to make it legitimate. This paper attempts for a counterattack strategy to eradicate insertion of good words by proposing architecture of enhanced DKIM (DomainKeys Identified Mail) as a solution. Our experimental result shows that DKIM serves to be the best as it incorporates sender evidence with random values in the E-Mail messages which is critical for the spammers to evade E-Mail filtering process. The misclassification of the spam E-Mail as legitimate E-Mail would reduce the performance of text classifiers. As the number of E-Mail increases, the misclassification percentage decreases by using DKIM

Index Terms— spam filtering, good word attack, DomainKeys Identified Mail (DKIM)

———————————————— ◆ ————————————————

1 INTRODUCTION

THE statistical spam filters use Machine Learning Techniques for automatically sorting text sets into categories from a predefined set. They are broadly classified into Reinforcement learning, supervised learning, semi-supervised learning and unsupervised learning. The learning method for each technique differs. In supervised learning method all training data are mostly labeled, unsupervised method train machines to learn by using unlabelled data, Semi-supervised learning technique uses both labeled and unlabeled data for training whereas reinforcement learning makes use of an agent to train data.

Text Categorization approach has considerable savings in labor power for organizing and handling text data than the knowledge engineering approach which requires data to be collected with the help of the domain experts either through direct interaction or through question raise with the help of the domain experts. Though Text Classification filtering Techniques is proven useful in statistical spam filters, spammers systematically modify the E-Mail messages and malicious contents enter the user's host bypassing the filters. One such type of attack is known as Good Word Attack in which spam messages are injected with enough good words which tends the text classifier system to classify a spam as a legitimate E-Mail. Spammers are explicitly trained to learn the features (keywords) which mostly occur in legitimate E-Mails and add those sets of good feature words(Most frequently occur-

ring words in legitimate E-mails) to make the spam messages legitimate.

Also they append the spam keywords with spaces and punctuation symbols so that they are not filtered by the statistical spam filters. Even though a large body of research was proposed to this good word attack, there is paucity of misclassifications of features. DKIM [8] is a defense mechanism which uses digital signatures and guarantees authenticated E-Mail service. Further Domain Keys offers end-to-end integrity from a sender to the intended recipient with randomly generated evidence values.

This paper is organized as follows. Section 2 summarizes the related work, Section 3 discusses the architecture design of the proposed work, Section 4 discusses the experimental results and Section 5 is the conclusion.

2 RELATED WORKS

Enrico Blanzieri et.al presents an overview of machine learning applications for spam filtering and compares the different filtering methods. They also discuss other branches of anti-spam protection and use of various approaches in commercial and noncommercial anti-spam software solutions [1]

Fabrizio Sebastiani et.al compares the various automated approaches of text categorization algorithms in the way the classifiers are constructed and further evaluate the above said approaches for document indexing within the general machine learning Paradigm [2].

Sirisanyalak et. al uses an E-Mail feature extraction technique that extracts a set of four features and has used those features as input for spam detection model in artificial immune spam systems [3].

Gregory Wittel et. al examines the general attack method like common word attack and dictionary attack in

• *Kashefa Kowser.K is currently pursuing masters degree program in Information security in Pondicherry Engineering College, India, PH-9597652892. E-mail: kashefais@gmail.com*
• *Saruladha. K is currently working as a Assistant Professor and pursuing PhD in Pondicherry Engineering College, India, PH-9442396080. E-mail: charusanthaprasad@yahoo.com*
• *Packiavathy. M is currently pursuing masters degree program in Information security in Pondicherry Engineering College, India, PH-9943978904. E-mail: packiavathy_m@gmail.com*

the filter's features generation through tokenization or obfuscation along with the challenges faced by developers and spammers [4].

Daniel Lowd et. al describes the naïve bayes, maximum entropy statistical spam filters and evaluates the effectiveness of active and passive good word attacks on those filters [5].

Zach Jorgensen et. al applies multiple instance logistic regression on the multiple bags of instances (segments) and an E-Mail is classified as legitimate if all the instances in it are legitimate and as spam if at least one instance in the corresponding bag is spam [6]

Allman et.al [7] defines DKIM as a digital signature domain-level authentication framework that permits potential E-mail signers to publish E-Mail signing practices information for the E-Mail receivers to make additional assessments about messages using key server technology, public-key cryptography and Mail Transport Agents (MTAs) or Mail User Agents (MUAs).

Barry Leiba et.al focuses on verifying the digital signature that creates the evidence and ensuring both the sender and the recipient about the mail origin from where it says it does [8]

Erkut Sinan Ayla Havelsan et.al discusses intra-domain E-mail security system. It keeps E-Mail messages in corresponding mailboxes as encrypted messages. Trusted Mail Gateway process keeps encrypted E-Mail messages in mail boxes and records processing results in a database as notary information [9]

Ya-Jeng Lin et.al discusses the Lightweight, Pollution-Attack Resistant Multicast authentication scheme (PARM), which generates evidence that receivers can validate on a fast, per-packet basis. Fault-tolerance coding [10] algorithm which is discussed tolerates the loss of packet and signature amortization reduces the computation and communication overhead.

3 PROPOSED DKIM BASED SOLUTION FOR GOOD WORD ATTACK

GoodWord attack is one of the most popular frequently employed attacks by spammers. The main issue in good word attack is that a spammer adds extra words or phrases to a spam message which are typically associated with legitimate E-Mail. Spam messages inflated with good words are more likely to bypass spam filters. Good word attack contains both passive and active attack. In passive attack, a word list is constructed by the attacker without any feedback from the spam filter. In active attack, text messages are sent to the filter to determine whether or not they are labeled as spam. So far, relatively little research has been done on how spam filters can be trained to account for such attacks. The misclassification of spam E-Mails as legitimate E-Mails (Good Word Attack) would reduce the performance of the text classifiers. This misclassification percentage could be reduced by the following methods.

- Frequent re-training of classifiers is an existing solution for combating good word attacks.
- Creating evidence of the sent E-mails so that the intercepted E-mails for injection of Good Words could be identified.

The first method though seen as a good solution the training of the text classifiers are to be done frequently and if the number of feature words in legitimate E-mails increases, the training time also increases.

This paper presents a novel approach for combating good word attack in statistical spam filters using DomainKeys Identified Mail (DKIM) based architecture. DKIM [8] defines a mechanism of cryptographically signing E-Mail messages permitting a signing domain to claim responsibility for the introduction of a message into an E-Mail system. Sender server publish public key in DNS (Domain Name Service) and then a sum using SHA256 [12] is calculated on selected header for sending an E-Mail. The sender generates a digital signature of the hashed message using RSA [11], a public key encryption scheme. The receiver server now looks public key up using DNS, decrypts the hash value and verifies the received sum. If the sum verifies, the sender server is verified, and the mail can be delivered. The proposed DKIM based architecture incorporates the sender evidence in the E-Mail messages to avoid the injection of good word thereby making the spam detection possible. Figure 1 shows the DKIM based architecture for combating good word attack. The following steps are to be followed for creation of evidence generation.
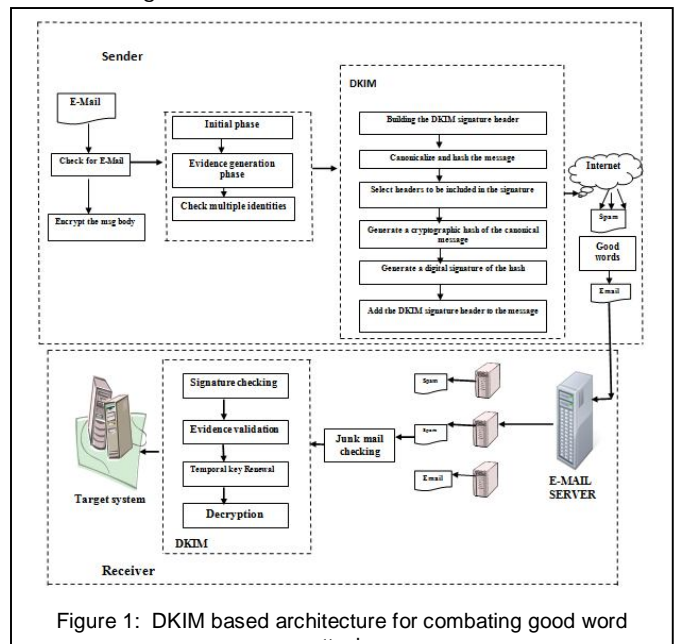
Figure 1: DKIM based architecture for combating good word

3.1 Evidence generation by the sender

In the first stage, Junk E-Mail is checked and controlled using its common spam characteristic features like discount, offer, bonus, money etc. The spam features can be identified with the help of message pre-processing used by a machine learning filter. Three main steps involved in the pre-processing of messages are

- Tokenization, decomposing the sentence into smaller units using punctuations, spaces etc., and extracting the features
- Lemmatization, reducing words to possible root word (e.g., "meeting" to "meet")
- Stop word removal, eliminating words like "to", "and", "for" that often occur in many messages. After data preprocessing if there is a spam keyword the DKIM header cannot be created otherwise the E-Mail is send for the encryption.

In the this stage, the body of the E-Mail is encrypted using Digital Signature Algorithm (DSA) [13] and a temporal key pair using one-way hash function is generated. Temporal key pair contains a Temporal Secret Key (TSK) chain and a Temporal Public Key (TPK). The sender creates the evidence of a packet from a TSK chain, and the receiver validates the evidence of a received packet with the TPK [10]. An attacker can convince receivers to accept a forged TPK if the sender not sign the TPK with the digital signature during distribution.

The sender generates the evidence which should be lightweight and fast for the receiver. The receiver should generate the evidence before sending a message to determine the validity of the packet and the sender needs to maintain a usage table for a given temporal key pair to find the number of times the column index of the TSK element array is used. The spammers may use multiple identities for good word attack in E-Mail. Hence, multiple identities of the user can be avoided by checking the user's address and time of sending the message.

Next stage is to construct the DKIM signature header in which the header will be covered by the signature. The domain and identity is to be signed in the header and the selector identifies the signing key. After constructing the DKIM signature header, the signer calculates a hash of the message body using hash algorithm. The next choice is to go into the signature header is the canonicalization algorithm for the headers and for the body. Canonicalization stage is used for minimal transformation of the message that will give its best chance of producing the same canonical value at the receiving end.

Header fields are the parts of the message that are most vulnerable to change in transit. In the next step, the signer can choose the header field to sign using DKIM by leaving insignificant header fields unsigned. This may increase the chances that the signature verified successfully.

SHA256, hash algorithm is used to generate a cryptographic hash for the canonical message. A hashing algorithm takes a variable length data message and creates a fixed size message digest. Then the signer signs the hash using the RSA encryption algorithm in the signature header, and adds it to the beginning of the message header fields. Finally the encrypted content will be added in the DKIM header. This completes the task of creation of evidence by the sender.

3.2 Evidence validation by the receiver

In checking signature phase, valid signature header must be checked by the receiver. The desired key identity is determined and retrieved from the specified key store. It is then validated and the public key is extracted from it.

Policy retrieved from the receiver should be through the DNS query. From address has the domain for the query.
In evidence validation phase, the receiver can use the TPK to immediately check the validity of the attached evidence when receiving a packet.

Algorithm at Sender Side
Input: DKIMBevidgen (E-Mail);
Output: Evidgen E-Mail

//Evidence attached E-Mail (Evidgen E-Mail)
1) Check the E-Mail.

2) Encrypt the message body using Digital Signature Algorithm (DSA).

3) Initialize Temporal Public Key (TPK) & Temporal Secret Key (TSK) for storing the evidence.

4) Create the evidence of a packet using TSK chain.

5) Calculating generation of evidence, hash & concatenates P with Q //P represents packet which is going to transfer, Q represents sequence number of the packet.

6) Append the evidence in the usage table //Usage table contain (TPK) and (TSK).

7) Generate n random numbers.

8) Hashing the random value using SHA256.

9) Building the DKIM header and canonicalize the message.

10) Selection of the header.

11) Digital signature of the hashed value using RSA.

12) Concatenate the message to DKIM header.

Algorithm at Receiver Side
Input: DKIMBevidcheck (Received E-Mail)
Output: Original E-Mail

13) Check the signature

14) Validate the evidence using TPK

15) Renewal of the used TSK elements.

16) Decryption of the encrypted message using Receiver's private key.

The attacker must generate proper evidence for a packet to forge, which is difficult without the knowledge of the TSK chain. The receiver must also maintain a usage table for each column index of the TSK elements array based on received packets like sender. Periodic renewal of used TSK elements ensures secure communications between the sender and its receiver. The final phase is to decrypt the encrypted message using the receiver's private key.

4 EXPERIMENTAL SETUP

The implementation of the system is done in windows platform using JAVA on the publicly available spam corpus-Ling spam. The Ling spam corpus consists of 2171 legitimate E-Mail and 432 spam E-Mails in which 50% of the datasets is taken for implementation. The evidence generation value of random numbers contains E-Mail details, canonicalization value, the part which is generated by the header, selecting header in E-Mail details, SHA value, hash value and DKIM header value of the E-Mail in Table 1. Figure 2 represents the graph showing the misclassification percentage reduced by DKIM. As the E-Mail increases the misclassification percentage also increases. But the usage of DKIM decreases the misclassification percentage thereby combating the good word attack.

E-Mail Details	Header Generation	Canonicalization	Selecting header
To:dd@gmail.com From:ffff@rediff.com Subject:conference Date:Tue Apr 19 12:17:17 IST 2011 Message:International Conference on IS	i=@gmail.com	i=@gmail.comỹ c=simple/simple	i=@gmail.comỹ c=simple/simple t=Tue Apr 19 121717 IST 2011 h=ToFromSubjectDateMessage

Hash value	Sha value	DKIM header	Usage table
VG8= RnJvbQ == U3Via mVjdA == RGF0Z Q== TWVzc 2FnZQ ==	2047 06504745195334 42373552043352 37 13375151333937	i=@gmail.comỹ c=simple/simple t=Tue Apr 19 121717 IST 2011 a=rsa-sha256 h=VG8=RnJvbQ==U3ViamVj dA==RGF0ZQ==TWVzc2FnZ Q== 2047 06504745195334423735520 43 35237 133751513339370,×ä< "~&§Äü¯£åÂſ¼),ÜUEï¯§ö™L ÷äljäöEl Z™§	86:Q0ybWuUUZGu9kbUAMsp Xav7I8vwtKrBLmWZfEGOD dY=

Table 1: DKIM based architecture for combating good word attack.

Figure 2: Misclassification Percentage for Ling Spam Corpus

5 CONCLUSION

This paper presented novel solution for good word attack by employing DKIM mechanism. It differs from the other counterattack strategy as it incorporates sender evidence in the E-Mail messages thereby making Spam detection possible. The result shows that the misclassification percentage decreases as the mail increases with the help of DKIM to eradicate the insertion of good words in spam E-Mail which makes as legitimate.

REFERENCES

[1] Enrico Blanzieri, Anton Bryl, "A Survey of Learning-Based Tec niques of Email Spam Filtering", January 11, 2008.

[2] Fabrizio Sebastiani, "Machine Learning in Automated Text Categorization", ACM computing surveys, Vol 34, No 1, 2002.

[3] B. Sirisanyalak and O. Sornil, "An artificial immunity-based spam detection system", Proc. of IEEE Congress on Evolutionary Computation, 2007, pp.3392-3398.

[4] Gregory Lee Wittel , "Evaluating and Attacking Statistical Spam Filtering Systems", Thesis, B.S. (University of California, Davis) 2002.

[5] D. Lowd and C. Meek, "Good word attacks on statistical spam filters", In Proceedings of the 2nd Conference on Email and Anti-Spam, 2005.

[6] Y. Zhou, Z. Jorgensen and M. Inge, "Combating Good Word At-tackson Statistical Spam Filters with Multiple Instance Learning", Proc.of 19th IEEE International Conference on Tools with Artificial Intelligence, 2007, pp.298-305.

[7] Allman, E., Delany, M., and J. Fenton, "DKIM Sender Signing Practices", Internet Draft, http://www.ietf.org/internetdrafts/draft-allman-dkim-ssp-02.txt (work in progress), August, 2006.

[8] Barry Leiba, Jim Fenton, "DomainKeys Identified Mail (DKIM) Using Digital Signatures for Domain Verification", Journal of Foundations and Trends in Information Retreival, pp. 538-549, January 2008.

[9] Erkut Sinan Ayla Havesan, Ankara, Attila Ozgit, "An Architecture for End-to-End and Inter-Domain Trusted Mail Delivery Service", Communications of the ACM, pp. 24-33, February 2007.

[10] Ya-Jeng Lin, Shiuhpyng Shieh, Warren W. Lin, "Lightweight, Pollution-Attack Resistant Multicast Authentication Scheme", ASIAACCS'06, March 21-24, November 2006.

[11] http://www.di-mgt.com.au/rsa_alg.html

[12] http://www.ocean-logic.com/pub/OL_SHA256.pdf

[13] http://dsmc.eap.gr/members/pkitsos/papers/Kitsos_c09.pdf

Low-Power 1-bit CMOS Full Adder Using Subthreshold Conduction Region

Vishal Sharma, Sanjay Kumar

Abstract— In balancing the trade-off between power, area and performance, numerous efforts have been done. However, not much study has been done at the two extreme ends of the design spectrum, namely the ultra low-power with acceptable performance at one end (the main concern of this paper), and high performance with power within limit at the other. This paper is based on the exclusive use of subthreshold conduction currents to perform circuit operations, yielding a dramatic improvement in power consumption compared to traditional circuit design approaches. This improvement makes it feasible to design extreme low-power circuits with such an approach. The CMOS digital circuits for this work have been designed using standard TSMC 0.18 μm Technology.

Index Terms— Low-Power, Subthreshold Conduction Region, Full Adder.
.

———————————————— ◆ ————————————————

1 INTRODUCTION

IN most VLSI applications, arithmetic operations play an important role. Commonly used operations are addition, subtraction, multiplication and accumulation, and 1-bit Full Adder is the building block for most implementations of these operations. Obviously, enhancing the building block performance is critical for enhancing overall system performance [1] and in present, the power consumption has become a critical concern in today's VLSI system design. The need for low-power VLSI systems arises from two main forces. First, with the steady growth of processing capacity per chip, large current has to be delivered and the heat due to large power consumption must be removed by proper cooling techniques. Second, battery life in portable electronic devices is limited. Low-power design directly leads to prolonged operation time in these portable devices [2].

Subthreshold circuit design provides an efficient solution to significantly reduce the power. So this work is based on subthreshold conduction principle to design a low-power 1-bit Full Adder. In subthreshold circuits, the supply voltage is reduced well below the threshold voltage of a transistor. Due to the quadratic reduction in power with respect to the supply voltage, subthreshold circuits are classified as *ultra low-power circuits*.

Specifically in application areas where performance can be sacrificed for low-power, subthreshold circuits are an ideal fit. Some of the applications include devices such

as digital wrist watches, radio frequency identification (RFID), sensor nodes, pacemakers and battery operated devices such as, cellular phones [3].

2 MOTIVATION OF POWER REDUCTION

Up until now, the power consumption has not been of great concern because of the availability of large packages and other cooling techniques having the capability of dissipating the generated heat. However, continuously increasing density as well as the size of the chips and systems might cause to difficulty in providing adequate cooling and hence, might either add significant cost to the system or provide a limit on the amount of the functionality that can be provided [4].

Another factor that fuels the need for low-power chips is the increased market demand for portable consumer electronics powered by batteries. For these high performance portable digital systems, running on batteries such as-laptops, cellular phones and personal digital assistants (PDAs), low-power consumption is a prime concern because it directly affects the performance by having effects on battery longevity.

Hence, low-power 1-bit Full Adder design has assumed great importance as an active and rapidly developing application in VLSI. Due to their extreme low-power consumption, subthrehsold design approaches are appealing for a widening class of applications which demand low-power consumption and can tolerate larger circuit delays.

3 SUBTHRESHOLD CONDUCTION FOR LOW-POWER VLSI DESIGN

————————————————

- *Vishal Sharma is currently working as Assistant Professor in Electronics and Communication Engineering Department in Sir Padampat Singhania University, Udaipur, Rajasthan, India. E-mail: vishalfzd@gmail.com*
- *Sanjay Kumar is currently working as Assistant Professor in Electronics and Communication Engineering Department in Thapar University, Patiala, Punjab, India. E-mail: sanjay.kumar@thapar.edu*

Digital subthreshold circuits are currently used for some low-power applications such as hearing aids, wrist watches, pacemaker and wireless communication systems [5]. In subthreshold circuit design, supply voltage must be scaled down below the threshold voltage. Therefore, the load capacitance is charged or discharged by the subthreshold leakage current. When the Adder circuits operate in the subthreshold region, they should have different performance characteristics than those of the Adders working in the strong inversion region [6]. In this paper, we will see the characteristics of a 1-bit Full Adder cell operating in subthreshold region of operation comparing it with the 1-bit Full Adder cell operating in strong inversion.

In traditional digital VLSI design, subthreshold region of operation is avoided, since it contributes toward leakage power consumption when the device is in stand-by. But the power can be reduced significantly by exclusively utilizing this subthreshold leakage current to implement circuits. This is achieved by actually setting the circuit power supply V_{dd} to a value less than or equal to V_{th}. The subthreshold current is exponentially related to gate voltage giving the exponential reduction in power consumption, but also an increase in circuit delay [7]. So, we use the circuits operating in subthreshold conduction region where the power is main concern and large delay can be tolerated.
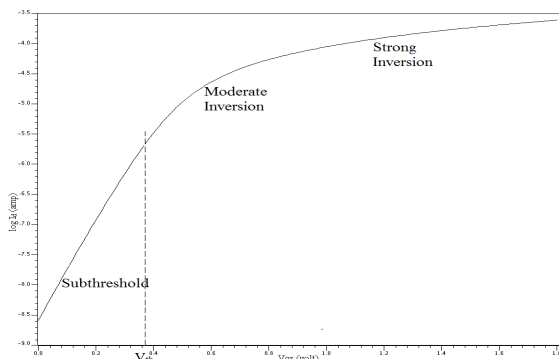
Fig. 1 CMOS Conduction region for an NMOS with V_{ds}= 1.8 V and V_{gs} varying from 0 V to 1.8 V.

Unlike moderate and strong inversion, in which the drift component of current dominates, subthreshold conduction is dominated by diffusion current [8].

3.1 Modelling of Subthreshold Current

In subthreshold conduction, the channel of the transistors is not inverted and current flows by diffusion. Subthreshold current can be expressed by the following basic equation:

$$I_{on-sub} = I_0 \, exp\left(\frac{V_{gs} - V_{th}}{nV_T}\right) \tag{1}$$

Equation (2) shows the same basic equation with low V_{ds} roll-off:

$$I_{on-sub} = I_0 \, exp\left(\frac{V_{gs} - V_{th}}{nV_T}\right)\left(1 - exp\left(\frac{-V_{ds}}{V_T}\right)\right) \tag{2}$$

where I_0 is the drain current when $V_{gs} = V_{th}$ given below:

$$I_0 = \mu_{eff} C_{ox} \frac{W}{L_{eff}} (n-1) V_T^2 \tag{3}$$

As expected for the diffusion current, (1) shows that I_{on-sub} depends exponentially on V_{gs}. Here, W is the width of the transistor, L_{eff} is the effective length, μ_{eff} is the effective mobility, C_{ox} is the oxide capacitance, n is the subthreshold slope factor $\left(n = 1 + \frac{C_d}{C_{ox}}\right)$, V_{th} is the transistor threshold voltage and V_T is the thermal voltage, $V_T = (kT/q)$.

4 MINIMUM ENERGY POINT MODEL

In this section, we will discuss a closed form solution for the optimum V_{dd} and V_{th} for a given frequency and technology operating in the subthreshold region means $(V_{dd} \, V_{th})$.

The total energy per operation of a digital CMOS circuit consists of two components: switching and leakage energy [9]. Here, we discuss in terms of an inverter.

Hence, total energy per operation can be expressed as:
$$E_{Total} = E_{SW} + E_L$$

$$= V_{dd}^2 \left(C_{eff} + W_{eff} K C_g L_{DP} exp\left(\frac{-V_{dd}}{nV_T}\right)\right) \tag{4}$$

where, C_{eff} is the average effective switched capacitance per operation, K is a delay fitting parameter, C_g is the output capacitance of the inverter and L_{DP} is the depth of the critical path.

By using (1), (2), (3) and (4), the equations for optimum supply voltage $V_{dd,opt}$ and optimum threshold voltage $V_{th,opt}$ can be derived as shown in (5) and (6) respectively [3].

$$V_{dd,opt} = nV_T \left(2 - \text{lambert W}\left(\frac{-2C_{eff}}{W_{eff} K C_g L_{DP}} exp(2)\right)\right) \tag{5}$$

Also, we can find the optimum value of $V_{th} = V_{th,opt}$ for a given frequency f as:

$$V_{th,opt} = V_{dd,opt} - nV_T \, log_e\left(\frac{f K C_g L_{DP} V_{dd,opt}}{I_0}\right) \tag{6}$$

If the argument to the natural log in (6) exceeds 1, then the assumption of subthreshold operation no longer holds because $V_{th,opt} \, V_{dd,opt}$. This constraint shows that there is a maximum achievable frequency for a given circuit in the subthreshold region [9].

Swanson and Meindl analysed the VTC of an inverter

and showed that the inverter operation could be simulated down to 100 mV [10]. The VTC curves for different supply voltages for an inverter are shown in Fig. 2. To find the minimum voltage, Swanson equated the *off* current of NMOS and PMOS and calculated the inverter gain in subthreshold. Since an inverter must have sufficient gain at $V_{dd}/2$, the minimum voltage to be used was estimated to be 8(kT/q) or 0.2 V [10].
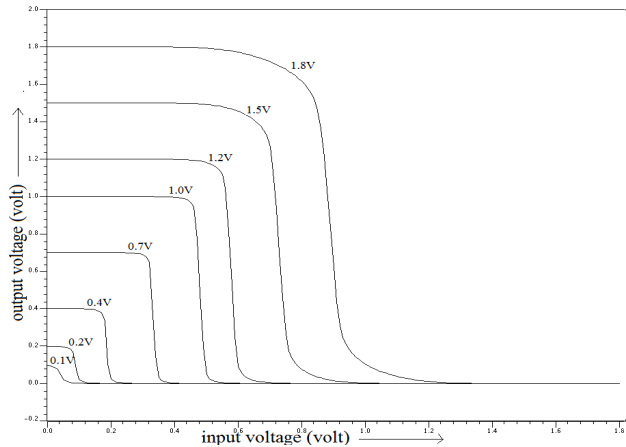
Fig. 2 Inverter VTC showing operation down to 100 mV in a 0.18 μm process.

5 DESIGNING OF LOW-POWER 1-BIT CMOS FULL ADDER

This section describes the design of 1-bit CMOS Full Adder operating in subthreshold conduction region. Before this design, the basic CMOS Inverter, shown in Fig. 3, is analysed in detail and then, based on this analysis, 1-bit Full Adder can be designed by calculating the values of *W/L* with the help of these *W/L* values of Inverter which has been designed for symmetric output and equal charging and discharging current.
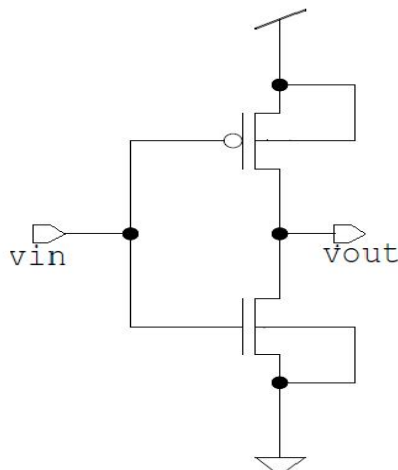
Fig. 3 Basic CMOS Inverter

By simulating this CMOS Inverter using TSMC 0.18 μm technology, the Inverter's values of *W/L* for PMOS = (1.2 μm/0.18 μm) and *W/L* for NMOS = (0.27 μm/0.18 μm) were obtained for strong inversion operation. While the Inverter's values of *W/L* for PMOS = (3.0 μm/0.18 μm) and *W/L* for NMOS = (0.27 μm/0.18 μm) were obtained for subthreshold operation.
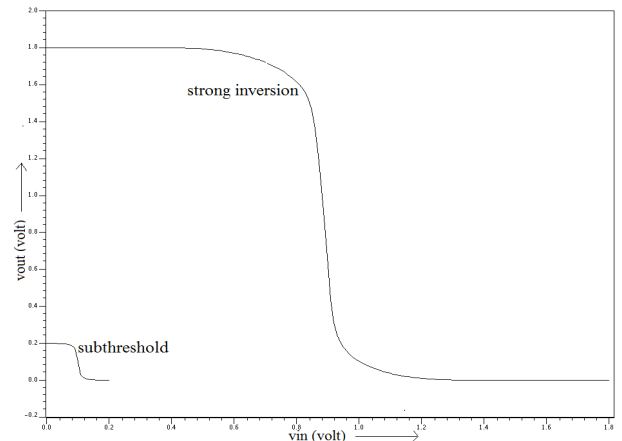
Fig. 4 VTC curves for Inverters operating in strong inversion and subthreshold.

Using these values of *W/L* of basic Inverter, Full Adder circuit can be designed having the equivalent *W/L* values equal to that of this basic Inverter [3].

6 DESIGN AND SIMULATION RESULTS

6.1 CMOS Inverter

The load capacitance, for the inverter described in previous section, for strong inversion region is 5 *fF* while the load capacitance for subthreshold conduction is 11 *fF*.
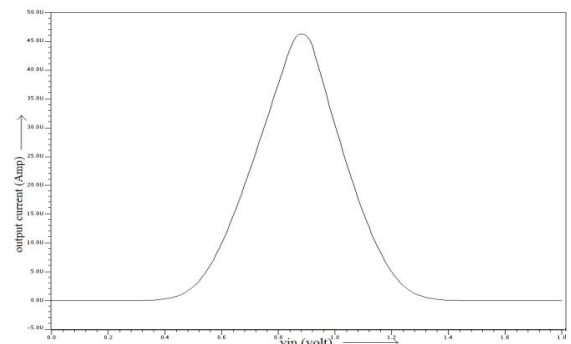
Fig. 5 Output current variation with input voltage in strong inversion region with V_{dd} = 1.8 V.

From Fig. 5 and 6, it is clear that the current depends on input (gate) voltage linearly in strong inversion region and exponentially in subthreshold region.
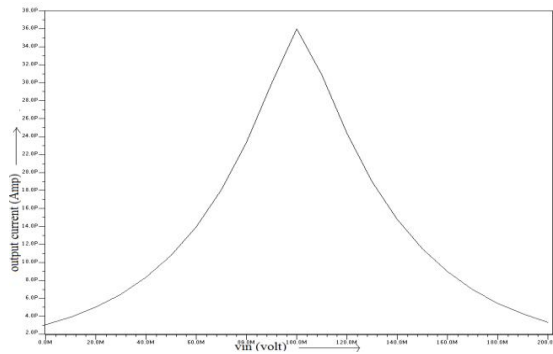
Fig. 6 Output current variation with input voltage in subthreshold region with V_{dd} = 0.2 V.
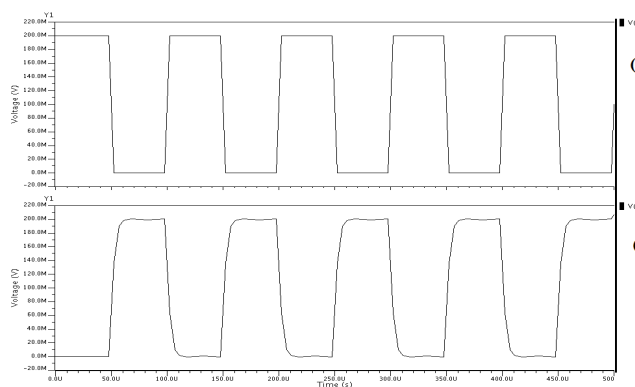
Fig. 7 Simulation result of transient analysis for CMOS Inverter in subthreshold region: (a) Input Signal, (b) Voltage waveform of Output Signal.

Fig. 8 shows the dynamic power variation with different clock frequencies for a 1-bit Full Adder operating in subthreshold conduction region.
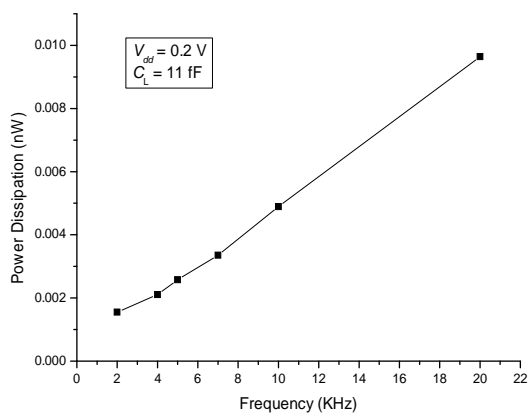
Fig. 8 Power dissipation results for a CMOS Inverter operating in subthreshold conduction region.

6.2 1-Bit CMOS Full Adder

A Full Adder is basic structure for any arithmetic circuit, so the design of a Full Adder is very necessary.

The basic structure and simulation result for a 1-Bit CMOS Full Adder are shown in figures given below.
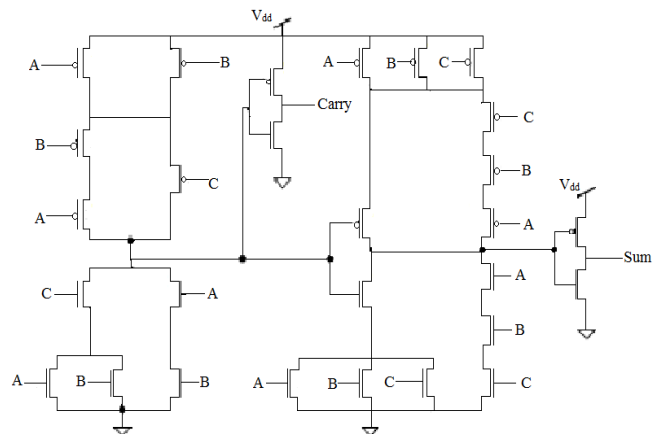
Fig. 9 Basic Structure of a 1-Bit CMOS Full Adder.
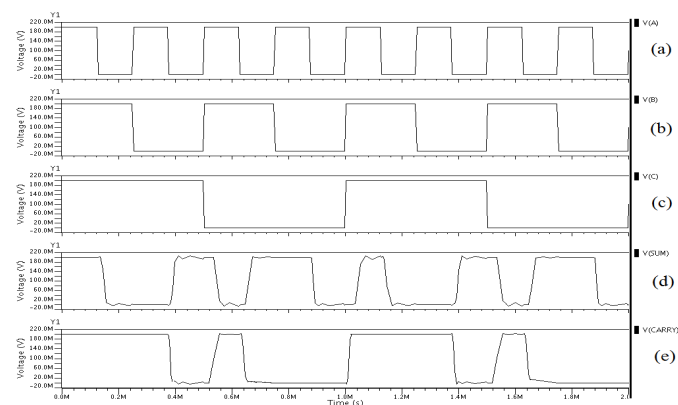
Fig. 10 Simulation result of transient analysis for a 1-Bit CMOS Full Adder in subthreshold conduction region: (a) Input Signal (A), (b) Input Signal (B), (c) Input Signal (C), (d) Voltage waveform of Output (SUM) Signal, (e) Voltage waveform of Output (CARRY) Signal.

TABLE I
POWER DISSIPATION RESULTS

	Operating Frequency (kHz)	Power (nW)	
		Superthreshold region	Subthreshold region
CMOS Inverter	2 k	275.87 n	0.00156 n
	4 k	389.54 n	0.00211 n
	5 k	486.93 n	0.00258 n
1-bit Full Adder	2 k	535.54 n	0.01652 n
	4 k	971.15 n	0.02441 n
	5 k	1100.94 n	0.02793 n

Fig. 11 shows the dynamic power variation with different clock frequencies for a 1-bit Full Adder operating in subthreshold conduction region.
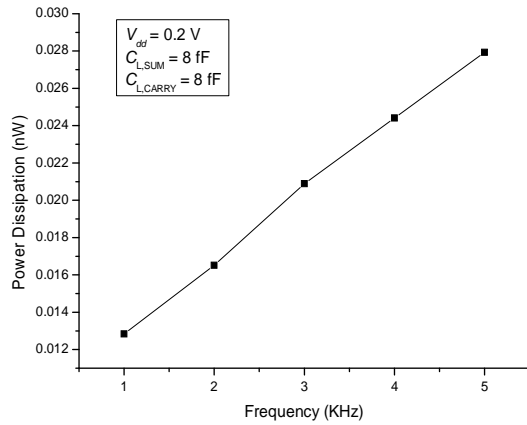
Fig. 11 Power dissipation results for a 1-bit CMOS Full Adder operating in subthreshold conduction region.

7 LAYOUT DESIGNS AND POST-LAYOUT SIMULATION RESULTS

After completing the physical layout design of all the structures, they are matched with schematics using layout versus schematic (LVS) simulation. LVS is done to determine whether a particular layout design corresponds to the original schematic of circuit diagram of the design.

7.1 CMOS Inverter

For a CMOS Inverter operating in subthreshold conduction region, layout and post-layout simulation results are shown in Fig. 12, 13 and 14 respectively.
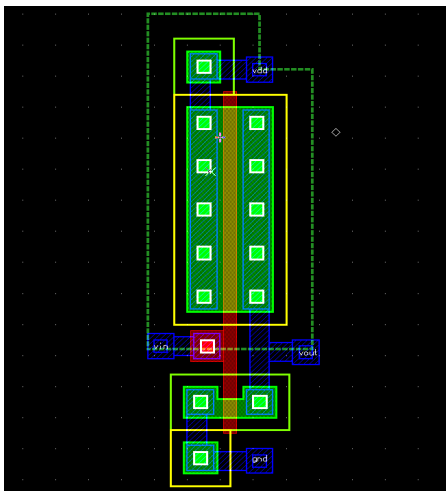
Fig. 12 Layout of CMOS Inverter operating in Subthreshold conduction region.
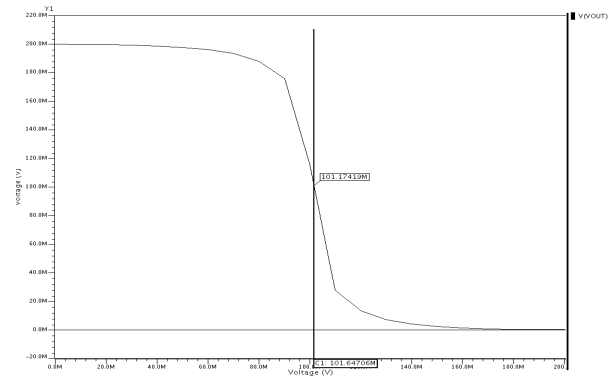
Fig. 13 Post-Layout Simulation – VTC curve for CMOS Inverter operating in subthreshold conduction region.
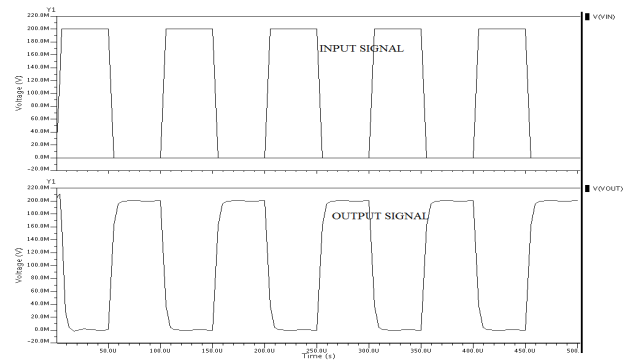
Fig. 14 Post-Layout Simulation – Transient analysis for CMOS Inverter operating in subthreshold conduction region.

7.2 1-Bit CMOS Full Adder

For a 1-bit CMOS Full Adder operating in subthreshold conduction region, layout and post-layout simulation results are shown in Fig. 15 and 16 respectively.
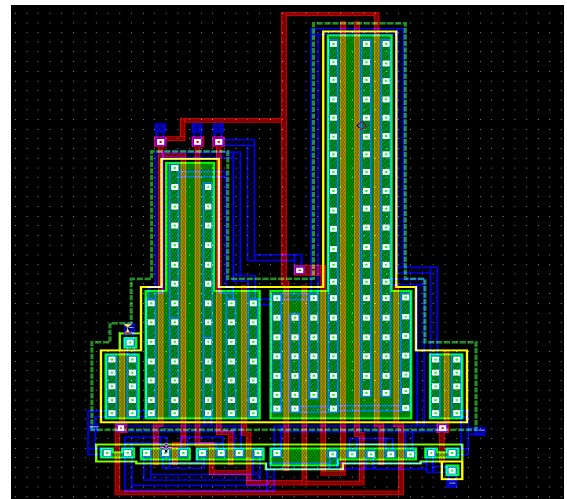
Fig. 15 Layout of 1-bit CMOS Full Adder operating in subthreshold conduction region.
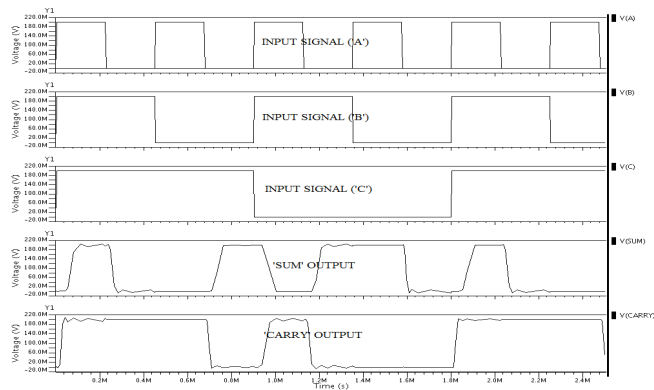
Fig. 16 Post-Layout Simulation – Transient analysis for 1-bit CMOS Full Adder operating in subthreshold conduction region.

8 CONCLUSIONS

Based on the subthreshold conduction region, the designing of a 1-bit CMOS Full Adder has been done. The supply voltage used for the circuits operating in subthreshold conduction region is 0.2 V. The power analysis has also been carried out for the circuits operating in subthreshold conduction region and in strong inversion region. It is found that the circuits operating in subthreshold conduction region provide the significant power reduction than the strong inversion region.

Also, it can be found that subthreshold conduction region is advantageous in applications where power is the main concern and performance can be sacrificed to achieve the low-power because the speed of a circuit operating in subthreshold conduction region becomes significantly slow.

REFERENCES

[1]   K. Granhaug and S. Aunet, "Six Subthreshold Full Adder Cells characterized in 90 nm CMOS technology" *IEEE Design and Diagnostics of Electronic Circuits and Systems*, pp.25-30, 2006.

[2]   S. Hanson, B. Zhai, K. Bernstein, D. Blaauw, A. Bryant, L. Chang, K. K. Das, W. Haensch, E. J. Nowak and D. Sylvester, "Ultralow-voltage, minimum-energy CMOS," *IBM Journal of Research and Development*, vol. 50, no. 4-5, pp. 469–490, 2006.

[3]   V. Sharma and S. Kumar, "Design of Low-Power CMOS Cell Structures Using Subthresold Conduction Region" *International Journal of Scientific and Engineering. Research*, vol. 2, Feb 2011.

[4]   A. P. Chandrakasan, S. Sheng and R. W. Brodersen, "Low Power CMOS Digital Design," *IEEE Journal of Solid-state Circuits*, vol. 27, no. 4, pp. 473-484, April 1999.

[5]   H. Soeleman, K. Roy and B. C. Paul, "Robust subthreshold logic for ultra low-power operation," *IEEE Trans.VLSI Syst.*, vol. 9, pp. 90-99, Feb. 2001.

[6]   V. Moalemi and A. Afzali-Kusha, "Subthreshold 1-bit Full Adder Cells in 100 nm Technologies," *IEEE Computer Society Annual Symposium on VLSI*, pp. 514-515, May 2007.

[7]   K. Ragini and B. K. Madhavi, "Ultra Low Power Digital Logic Circuits in Subthreshold for Biomedical Applications," *Journal of Theoretical and Applied Information Technology*, 2005.

[8]   K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer cmos circuits," in *Proceedings of IEEE*, vol. 91, no. 2, pp. 305-327, 2003.

[9]   A. Wang, B. H. Calhoun and A. P. Chandrakasan, *Subthreshold Design For Ultra Low-Power Systems,* Springer, USA, 2006.

[10] R. M. Swanson and J. D. Meindl, "Ion-Implanted Complementary MOS Transistors in Low-Voltage Circuits," *IEEE Journal of Solid-State Circuits,* vol. 7, no. 2, pp. 146-153, April. 1972.

[11] S. M. Kang and Y Leblebici, *CMOS Digital Integrated Circuits: Analysis and Design*, 3rd ed, McGraw Hill, 2004.

Thermal Performance of Thermosyphon Charged by Nanofluid for Cooling Electronic Component

M. G. Mousa

Abstract— In present work a experimental study based on thermosyphon is investigated using a solid nano-particles added to water as a working fluid. Tests are made on a thermosyphon. The experiment was performed in order to measure the temperature distribution and compare the heat transfer rate of the thermosyphon with the nanofluid and with pure water. The tested concentration level of the nano-particles is 0%, to 1.2%. Results show that the addition of 1.2% (by volume) of Al2O3 nano-particles in water presented improved thermal performance compared with the operation with pure water. Results showed that the total thermal resistance, using Al2O3-water based the nanofluid is enhanced by 57% at f equal 1.2 % compared to that of pure water, depending on filling ratio and volume fraction of the nano-particles in the base fluid. The experimental data are compared with the available literature and discrepancies between results are discussed.

Keywords - Two phase loop Thermosyphon, Thermal Performance, Nanofluids, Cooling of Electronic Devices.

Nomenclature

Cp   Specific heat, J/kg K
FR   Filling ratio
I   Current, Amp
k   thermal conductivity
L   Length, m
N   Number of thermocouples
Q   Input heat rate, W
TR   Total thermal resistance, K/W
T   Temperature, K
V   Applied voltage, Volt
Z   Vertical coordinates parallel to the test section ,m

Greek Symbols

ϕ Volume fraction of nano-particles, %
ρ Density, kg/m^3
μ Dynamic viscosity, $N.S/m^2$

Subscript

c   Condenser
e   Evaporator
nf   Nanofluid
m   Base fluid
RR   Reduction factor in thermal resistance, --

———————————— ◆ ————————————

1. INTRODUCTION

Heat dissipation from high performance microprocessors is approaching 100 W/cm². At this level, conventional cooling solution cannot match the demand from this high performance of components. In the present a new way to resolve the thermal challenges by developing and implementing an innovative low cost compact loop thermosyphon. The progress in the PC industry of high thermal loading coupled with high flux levels demands exploration of new heat transfer augmentation mechanisms. A nanofluid is emerging as alternative heat transfer fluids. The term 'Nanofluids' is used to indicate a special class   of heat transfer fluids   that contain nano-particles (<~100 nm) of metallic or non metallic substances uniformly and stably suspended in a conventional fluid. To solve the problem of growing heat generation by such electronic equipment, two-phase change devices such as heat pipe and thermosyohon cooling systems are now used in electronic industry and attracted the attention of several researchers **[1].**

The traditional working fluids have poor heat transfer properties compared to most solids. Therefore, Argonne National Laboratory has developed a new class of heat transfer fluids called "Nanofluids", which are engineered by suspending ultra-fine of metallic or nonmetallic nanometer dimension particles in traditional fluids such as water, engine oil, ethylene glycol by **Choi, and Eastman [1].** Some experimental investigations have revealed that nano-fluids have remarkably higher thermal conductivity and greater heat transfer characteristics than conventional pure fluids

Two phase loop Thermosyphon, TPLT, is a very simple and very efficient heat transfer device. Basically, it can be considered as a super-thermal conductor that transmits heat at small temperature difference by evaporation and condensation of the working fluid. The thermosyphon utilizes the local gravitational acceleration for the return of condensate from the condenser to the evaporator. Thermosyphon is considered as a passive device because heat is transferred inside a tube by buoyancy force only. Thermosyphon has a simple construction with no moving parts. Thus, it can be easily manufactured at low fabrication cost. Because of its small thermal resistance, thermosyphon transfers large amount of heat rates at small temperature differences. Heat transfer in thermosyphon is studied by many researchers. But, there is a need to investigate the effect of some parameters such as nanofluid concentration and other parameter especially from experimental point of view.

The heat transfer in a two-phase thermosyphon, TPLT, operating with a fluid in the near critical state was studied experimentally by **Gross and Hahne,** [2]. The fluid pressure, heat flow rate and tube inclination angle were experimentally investigated. Refrigerant-115 was used as a working fluid in the near critical state.

Steady state performance of a thermosyphon with tube separator was presented by **Lin and Faghri,** [3]. A one-dimensional steady-state mathematical model describing natural circulation two-phase flow thermosyphon with tube separator was presented. The theoretical results showed that a single liquid region in the lower part of the evaporator was

available if the heat rate and operating temperature were fairly low.

Single-phase and two-phase heat exchangers for power electronic components were presented by **Gillot et al**., [4]. The experiment performed to assess the feasibility of single-phase and two-phase micro heat sinks applied to the cooling of power components. After a brief recall of the principal characteristics of a power component insulated gate bipolar transistor, experimental measurements were described for multi-chip modules cooled by single-phase or two-phase heat sinks machined in a piece of copper.

Two-phase loop compact thermosyphon was presented by **Beitelmal et al.,** [5]. Heat dissipation from high performance microprocessors had a limiting value of 100 W/cm². They adopted a new way to resolve the thermal challenges by developing and implementing an innovative low cost compact loop thermosyphon.

Integrated thermal management techniques for high power electronic devices were presented by **Mcglen et al.,** [6]. In electronics industry, miniaturization of silicon components and enhanced performance has led to high power electronic devices with high packing densities. That has identified the future development of very high heat flux components. They reviewed the development of micro-processing technology and proposed a project to develop combined two-phase heat transfer and heat pipe technology with forced air convection and liquid condenser systems

An experimental investigation of thermosyphon thermal performance carried out using water and dielectric heat transfer liquids as the working fluid was performed by **Jouhara et al**., [7]. The copper thermosyphon was 200 mm long with an inner diameter of 6 mm. Each thermosyphon was charged with 1.8 ml of working fluid and tested usin an evaporator length of 40 mm and a condenser length of 60 mm. The thermal performance of the water charged thermosyphon outperformed the other three working fluids (FC-84, FC-77 and FC-3283) in both effective thermal resistance as well as maximum heat transport. However, these fluids, offered the advantage of being dielectric which may be better suited for cooling applications of sensitive electronics. Furthermore, they provided adequate thermal performance up to approximately 50 W, after which liquid entrainment compromised the thermosyphon performance.

Visual observation of flow patterns in the condenser and heat transfer measurements were obtained by **Maezawa and Takuma,** [8] using a vertical annular thermosyphon. Refrigerant-115 was used as a working fluid. As a result of visual observations, ripples (interfacial waves) were generated on the condensate film surface when the corresponding film Reynolds number exceeded approximately 20, and condensation heat transfer was promoted.

Ueda et al. [9] performed an experiment with vertically arranged closed two-phase thermosyphon. The characteristics of heat transfer of boiling in the heating section and condensation in the cooling section were investigated. The condensation heat transfer coefficient of the cooling section showed a decreasing trend with increasing wall temperature difference. However, its value was considerably lower under high vapor velocities. The boiling heat transfer characteristic

of the heating section could be approximated by the empirical correlation given for pool boiling.

Tanaka and Koshino, [10] performed series of experiments with a copper-water wickless heat pipe. The influence of working fluid charge and inclination angle on the heat transfer coefficients of the evaporator and the condenser were studied. The experimental results indicated that, as the heat transfer coefficient in the case of liquid film is larger than that in case of liquid pool.

Lin and Shyu, [11] experimentally studied the geyser boiling in a vertical annular closed two-phase thermosyphon. The experimental parameters considered were the heat load, condenser temperature, degree of liquid fill and length of the evaporator. It was clear that, geyser boiling i occurred more frequently and irregularly at high heat load. Wall temperature measurements indicated that the period of geyser boiling is shorter at higher heat load, shorter evaporator length and smaller liquid fill.

Noie et al, [12], studied experimentally the two phase closed thermosyphon in heat transmission. Fluids, with nano-particles suspended in them, were used as working fluids. The experimental results, at different input power, showed that the used of nanofluids increases the efficiency of two phase closed thermosyphon by 14.7%.

Mehta and Khandekar [13] and Khandekar et al [14and 15] investigated experimentally the overall thermal resistance of a closed two-phase thermosyphon using pure water and various water based nanofluids (of Al_2O_3, CuO and Laponite clay) as working fluids. All these nanofluids were showing the thermal performance greater than pure water. The behavior of nanofluids is explained in the light of pool boiling dynamics and the interplay of nucleating cavities with wettability of the nanofluids.

Zhen, et al [15] the experimental study was carried out to understand the heat transfer performance of a miniature thermosyphon using water-based carbon nanotube (CNT) suspensions as the working fluid. The suspensions consisted of deionized water and multi-wall carbon nanotubes with an average diameter of 15 nm and a length range of 5-15 Im. Experiments were performed under three steady operation pressures of 7.4 kPa, 13.2 kPa and 20 kPa, respectively.

The aim of the present work is to study experimentally, the thermal performance of thermosyphon cooling system in electronic equipment. The effect of different operating conditions on the thermal performance of the constructed thermosyphon is study. These conditions involve the type of working fluid (pure water and Al_2O_3 - water based nanofluid); filling ratio of the working fluid; volume fraction of nano- particles in the base fluid, and heat input rate. The additional aim of the loop thermosyphon is to reduce acoustic noise in the system by eliminating the existing heat sink fan.

2. Description of Experimental Apparatus

The nano-particles used in these experiments are AL_2O_3. To reduce the contact thermal resistance between the heater and the thermosyphon surface, the heated and adiabatic sections are insulated by Bakelite (0.233 W/m-K) and glass fiber wool (0.034 W/m-K). The ratio between the thermal resistances of heat flow from heater to the surrounding air compared to that from heater to thermosyphon is about 0.1%.

In this work adopts pure water and Al_2O_3-water based the nanofluid as working fluids. The size of used nano particles is 40 nm. A schematic layout of the experimental test rig is shown in Fig.1. The thermosyphon is divided into three sections; the evaporator and condenser are made of copper cube of 100.0 mm length and 1.0mm wall thickness and the adiabatic tubes which are 100.0mm length, 6.37 mm diameter and 0.85 mm thick. In the thermosyphon, the heat flux is generated by an electric heater (10). Pure water or nanofluid (16) is used as working fluid, and the vapor generated at the evaporator (1) is moved to the condenser (3) through the insulated vapor tube (2). The condensate returns back to evaporator by gravity through the insulated downcomer (4). The temperature along the thermosyphon walls are measured by Copper-Constantan thermocouples (K-type). Thermocouples are connected to a digital temperature recorder via multi-point switch connected to a data acquisition system. The thermosyphon is evacuated through a vacuum line (14) by vacuum pump and charged with working fluid through charging line (13). The electric power supplied to the evaporator heater which is similar to the heat of CPU (10) is measured by an ammeter (9) and voltmeter (8). The value of voltage can be adjusted by using an autotransformer (7). The voltage drop across the heater is varied from 5 to 35 Volt and the resistance of the heater is 20 Ω. A.C. voltage stabilizer (6) is used to ensure that no voltage fluctuation occur during the experiments. The operating pressures are 10, 12, and 15 kPa. The distribution of thermocouple along the thermosyphon is shown in figure 1.

2.1. Experimental Procedure

In the experimental work, the effect of filling ratio, FR, (volume of liquid over total evaporator volume) and heat flux are examined. The experimental runs are carried out according to the following steps:

1. The thermosyphon is evacuated and charged with a certain predetermined amount of working fluid and then the current and voltage are adjusted.
2. The supplied electrical power is manually adjusted to the desired value using the autotransformer.
3. As noted, the steady state condition is achieved after a time period of about 120 min, during which necessary adjustments in the input power is made. Thereafter, the readings of thermocouples, voltage and current are recorded as follows:
   ➢ The reading of thermocouples at the thermosyphon surface and ambient temperature are recorded sequentially using the selector switch.
   ➢ The voltage and current of the heater are measured to determine the value of the applied heat flux.
4. After finishing the above experimental steps, power is changed to another value and step 3 is repeated.
5. The thermosyphon is charged with another adjusted value of working fluid such that the filing ratio, FR, is changed using the following values: 32.0%, 40.0%, 48.0%, 56.0%, 64.0%, 72.0%, 80.0%, 88.0%, 96.0%, and 100.0% respectively. Then steps 1 through 4 are repeated.

Preparation of nano-particle suspensions is applied the nanofluids in heat transfer enhancement. In the present study, Al_2O_3 nano-particles are dispersed in pure water by ultra sonic without using any dispersant or stabilizer to prevent any possible changes of chemical properties of the nanofluid due to presence of additions. Nanofluids of 0.1%, to 1.2% volume fraction of particles are prepared. After measuring the equivalent volume to the required mass of nano-particle powder, they were mixed with distilled water in a flask subject to ultra sonic for about 4 h. No precipitation or settlement of the nano-particles is observed after 24 h of settling the suspension. The micrograph of Al_2O_3 nano-particles in pure water is shown in figure 2 at concentration, ϕ, equal to 1.0%.

2.2. Calculations procedure:

The obtained measured data of temperature and power are used to calculate the heat rate and the total thermal resistance.

At steady state condition, the produced heat generated by the electrical heating is given by the following equation to calculate the heat transfer rate and check the error analysis in experimental data may have resulted from:

$$Q = I \times V \qquad (1)$$

Where, Q, I, and V are the input heat rate, the electrical current in Amp, and the applied voltage drop a cross the heat element terminals.

Thus of the

$$Q = f(I, V) \qquad (2)$$

Error analysis of heat transfer rate is obtained by the following relation;

$$W_Q = \left[\left(\frac{\partial Q}{\partial V} W_V \right)^2 + \left(\frac{\partial Q}{\partial I} W_I \right)^2 \right]^{1/2} \qquad (3)$$

The experimental determination of the thermal performance of the thermosyphon requires accurate measurements of evaporator and condenser temperatures as well as the power transferred. Characterizing evaporator and condenser temperatures is, relatively, a straightforward task. One can obtain the average temperature along the evaporator and condenser surfaces are expressed as:

$$\mathbf{T}_e = \frac{\sum T_{ei}}{N_e} , \quad \mathbf{T}_c = \frac{\sum T_{ci}}{N_c} , \qquad (4)$$

Where N_e, N_c are the number of thermocouples on the evaporator, seven thermocouple, and condenser, seven thermocouple, respectively. The obtained data for temperatures and heat input rate are then used to calculate the total thermal resistance using the following relation:

$$\mathbf{TR} = \frac{T_e - T_c}{Q} \qquad (5)$$

The physical properties used for nanofluid are calculated from water and nano-particle properties at average bulk temperature using following correlations for density,

specific heat, viscosity, and thermal conductivity after [21-23]. One can assume that the nano-particles are well dispersed within the base-fluid. Thus, the effective physical properties are described by classical formulas as;

$$\rho_{nf} = (1 - \phi)\rho_m + \phi\rho_p \qquad (6)$$

$$\rho_{nf} Cp_{nf} = (1 - \phi)\rho_m Cp_m + \phi \, \rho_p Cp_p \qquad (7)$$

The effective dynamic viscosity of the nanofluid can be calculated using different existing formulas that have been obtained for two-phase mixtures. The following relation is the equation for a viscous fluid containing a dilute suspension of small, rigid, spherical particles.

$$\mu_{nf} = \mu_m \frac{1}{(1-\phi)^{2.5}} \qquad (8)$$

$$k_{nf} = k_m \left(\frac{k_p + 2k_m - 2\phi(k_m - k_m)}{k_p + 2k_m + (k_m - k_m)} \right)$$

Where, ϕ is the ratio of the volume of nano-particle to volume of the base fluid.

The relevant thermo-physical properties of the solid nano-particles (Al_2O_3) used in the present study are: Cp_p =773 J/kg.^{o}C, ρ =3880 kg/m^3 ,and k_p =36 W/m.^{o}C .from [21,23]

One can calculate the reduction factor in total thermal resistance of thermosyphon charged with nanofluid by referring its thermal resistance to that charged with pure water, expressed as

$$RR = (TR_m - TR_{ef})/TR_m \qquad (9)$$

The average temperature of the surface of the evaporator section increased from 60 to 100 ^{o}C. Data are recorded when the system reached steady-state condition. The maximum error in thermal resistance is $\pm$8.4 %

Experiments were performed under three steady operation pressures of 10.0 kPa, 12 kPa and 15 kPa, respectively

3. Results and Discussion

In the present section, experimental results involve the study of convective heat transfer in thermosyphon. Total thermal resistance is introduced. Using wall temperature readings, the total resistance of the tested thermosyphon is calculated for different values of operating parameters. The experiments for thermosyphon are carried out varying the values of heat source from 10 to 60 W. Pure water and the nanofluid used as working fluids are charged with filling ratios, FR, of 32.0%, 40.0%, 48.0%, 56.0%, 64.0%, 72.0%, 80.0%, 88.0%, 96.0%, and 100.0% respectively. Nano-particle concentration, ϕ are varied from 0.0% to 1.2%. Finally, the obtained experimental results are compared with the values obtained from available literature.

3.1. Thermal performance of thermosyphon using pure water as a working fluid

Figure 3 illustrates the surface temperature along the thermosyphon for different heat input rates (10, 20, 30 and 60 W) at fixed filling ratio, FR of 48% and operating

pressure of 15 kPa. As expected, the surface temperature decreases with increasing the distance from the evaporator, for any heat input rate. The surface temperature increases with the increase of heat input rate at a certain position along the thermosyphon. The adiabatic operating temperatures ranged from 40°C to 65°C. Externally controlled conditions are kept identical for baseline experiments with pure water.

Figure 4 illustrates the variation of the total thermal resistance, TR, with the filling ratio, FR, at different heat input rates, Q. It can be noticed that, TR decreases with increasing FR up to a value of 48%, after which the total thermal resistance, TR, starts to increase with increasing FR due to the increase of liquid inside the evaporator. The total thermal resistance, TR, is decreases as the heat input rate increases.

Figure 5 illustrates the surface temperature along the thermosyphon for different heat input rates (10, 20, 30 and 60 W) at fixed filling ratio, FR of 48% and operating pressure of 12 kPa.. Moreover, the surface temperature decreases with increasing the distance from the evaporator, for any heat input rate. At any certain position along the thermosyphon, the surface temperature increases with increasing heat input rate. The temperatures measured in case of operating pressure 12 kPa is lower than those measured in operating pressure 15 kPa see in (figures 3 and5)

3.2. Thermal performance of thermosyphon using nanofluid as a working fluid

Figure 6 illustrates the surface temperature along the thermosyphon for different heat input rates (10, 20, 30 and 60 W) at fixed filling ratio, FR of 48% and operating pressure of 15 kPa. The surface temperature decreases with increasing the distance from the evaporator, for all heat input rate.

Figure 7 illustrates the variation of the total thermal resistance of the heat pipe, R with the filling ratio for Al_2O_3-water based nanofluid of $\phi = 0.5\%$ at different heat input rates (10, 20, 30, and 60 W). As shown in the figure, TR decreases with the increase filling ratio up to a value of FR equals to 48%, where its value starts to increase with increasing FR. It can be also noticed in the figure that the thermal resistance, TR, gets decreases as the heat input rate increases.

Figure 8 illustrates the variation of the total thermal resistance, TR, of the thermosyphon, with the filling ratio for Al_2O_3-water based nanofluid of $\phi = 0.5\%$ at different heat input rates (10, 20, 30, and 60 W). As shown in the figure, TR decreases with the increase filling ratio up to a value of FR equals to 48%, where its value starts to increase with increasing FR. It can be also noticed in the figure that the thermal resistance, TR, decreases as the heat input rate increases.

Figure 9 shows the variation of the total thermal resistance of the thermosyphon, TR, with volume fraction of the nano-particle in the base fluid, ϕ at different heat input rate (10, 20, 30, and 60 W). Over the tested range of ϕ, while keeping the filling ratio of 48%, the percentage enhancement in TR ranged from 67% at $\phi = 1.2\%$ compared to that calculated using pure water. The addition of nano-particles has illustrated that during nucleate boiling some nano-particles

deposit on the heated surface to form a porous layer. This layer improves the wet ability of the surface considerably. The thermal conductivity of the working fluid is preferably high in order to minimize the temperature gradient. Thus, the thermal resistance of the fluid will be minimized. Thermal resistance of thermosyphon is caused by the formation of vapor bubble at the liquid–solid interface. A larger bubble nucleation size creates a higher thermal resistance that prevents the transfer of heat from the solid surface to the liquid [23]. The suspended nano-particles tend to bombard the vapor bubble during the bubble formation. Therefore, it is expected that the nucleation size of the vapor bubble is much smaller for the fluid with suspended nano-particles than that without them. One can estimate the relation between the percentage reduction factor in total thermal resistance, TR, RR, as a function of concentration as follow;

$$RR = 0.84\,\phi - 0.3\phi^2 \qquad (10)$$

The error in calculated thermal resistance is predicted by the above suggested around ±12 %.

3.3. Comparison between present and previous results

Figure 10 shows a comparison between the present experimental results with those reported by **Noie et al [12]** in case of using pure water as a working fluid with FR equals to 0.5. It can be observed that the present experimental results for the used two working fluids have the same trend as those reported by **Noie et al [12]**.The difference between both results when using pure water may be attributed to the difference in heat uncertainty in measurements.

4. Conclusions

In this paper the thermal enhancement of the thermosyphon performance, using Al_2O_3 nanofluid as the working fluid. In the present case, the DI-water with Al_2O_3 nano-particles in thermosyphon is experimentally tested. Conclusions may be drawn from the results of the performance tests. The more Al_2O_3 nano-particles are dispersed in the working fluid enhancement of the thermosyphon performance expressed by of the performance curves (thermal resistance as a function of the volume fraction, ϕ). The thermal resistance of the thermosyphon with nano-particle solution is lower than that with DI-water. It is shown that the thermal resistance decreases as the concentration increases. Results indicate that the Al_2O_3 nanofluid has remarkable potential as working fluid for thermosyphon of higher thermal performances. In order to obtain high heat transfer performance, it is necessary to fill between 40% and 60% of the evaporator inner volume. This conclusion is true for thermosyphons. For most examined conditions, the thermal resistance is improved by about to 60 % by increasing the nanofluid concentration to ϕ

References

(1) **Choi, S.U.S, "**Enhancing thermal conductivity of fluids with nano-particles", in: D.A. Siginer, H.P. Wang (Eds.), Developments and Applications of Non-Newtonian Flows, in: FED, vol. 66, ASME, NY, pp. 99-103,1995.

(2) **Gross, U., and Hahne, E**.,"Heat transfer in a two-phase thermosyphon operating with a fluid in the near critical state," Int. J. Heat Mass Transfer, Vol. 28, pp.589-601, 1985

(3)**Lin, L., and Faghri, A**.," Steady – state performance of a thermosyphon with tube separator ", Applied Thermal Engineering Vol. 17. No. 7, pp. 661-679, 1996.

(4) **Gillot, C., Bricard, A., and Schaeffer, C**.," Single- and two-phase heat exchangers for power electronic components", Int. J. Thermal Sciences Vol. 39,PP 826–832, 2000.

(5) Beitelmal, M. H., and Patel, C. D.," Two-Phase Loop: Compact Thermosyphon", Internet Systems and Storage Laboratory, 2002.

(6) Mcglen, R. J., Jachuck, R., and Lin, S.," Integrated thermal management techniques for high power electronic devices", Applied Thermal Engineering ,Vol.24 ,PP 1143–1156, 2004.

(7) Jouhara, H., Martinet, O., and Robinson, A.J., "Experimental Study of Small Diameter Thermosyphons Charged with Water", FC-84, FC-77 & FC-3283- 5th European Thermal-Sciences Conference, The Netherlands, 2008.

(8) Maezawa, S. and Takuma, M., "Heat transfer characteristics of the R113 annular two-phase closed thermosyphon," JSME series II, Vol. 31, pp. 469-476, 1988.

(9) Ueda, T., Miyashita, T., and Chu, P., "Heat transport characteristics of a closed two-phase thermosyphon," JSME series II, Vol. 32, pp. 239-246, 1989.

(10) Tanaka, O., and Koshino, H., "Heat transfer characteristics of a copper-water wickless heat pipe," JSME series B, Vol. 60, pp. 2826-2832, 1994.

(11) Lin, T. F., and Shyu, R. J., "Experimental investigation of geyser boiling in an annular two-phase closed thermosyphon," Int. J. Heat Mass Transfer, Vol. 38, pp. 295-307, 1995.

(12) Noie, S.H., Zeinali, S., Kahani, M., and Nowee, S. M., " heat transfer enhancement using Al_2O_3/ water nanofluid in a two phase closed thermosyphon International Journal of Heat and Fluid Flow 30 , pp.700–705, 2009.

(13) Mehta, B., and Khandekar, **S.,** 'Two-phase closed thermosyphon with nanofluids' 14th International Heat Pipe Conference (14th IHPC), Florianópolis, Brazil, April 22-27, 2007.

(14) Khandekar, S., Joshi, Y., and Mehta, B., "Thermal performance of closed two-phase thermosyphon using nanofluids", International Journal of Thermal Sciences, 47 pp.659-667, 2008.

(15) Khandekar, S., Joshi, Y., and Mehta, **B.,"** Thermal performance of closed two-phase thermosyphon using nanofluids "International Journal of Thermal Sciences 47, pp. 659–667, 2008.

(16) Liu,Z. , Yang, X., Wang,G., and Guo, G. "Influence of carbon nanotube suspension on the thermal performance of a miniature thermosyphon" International Journal of Heat and Mass Transfer vol.,53 , pp. 1914–1920, 2010.

(17) Liu, Z., Yang, X., and Guo, G., " Effect of nanoparticles in nanofluid on thermal performance in a miniature thermosyphon, Journal of Applied Physics, 102, 013526, 2007.

(18) Huminic, G., Huminic, A., Morjan, I., and Dumitrache, F., "Experimental study of the thermal performance of thermosyphon heat pipe using iron oxide nano-particles" International Journal of Heat and Mass Transfer 54, pp. 656–661, 2011.

(19) Xue, H., Fan, J., Hu, Y., Hong, R., and Cen, K., "The interface effect of carbon nanotube suspension on the thermal performance of a two-phase closed thermosyphon", Journal of Applied Physics, 100, 104909, 2006.

(20) Parametthanuwat, T., Rittidech, S., and Pattiya A., "A correlation to predict heat-transfer rates of a two-phase closed thermosyphon (TPCT) using silver nanofluid at normal operating conditions" International Journal of Heat and Mass Transfer 53, pp. 4960–4965, 2010.

(**21)Collier, G.**, and **Thome, J.,**" Convective Boiling and Condensation, Clarendon Press, Oxford, 1996.

(22) Faghri A., Heat Pipe Science and Technology, Taylor and Francis, 1995.

(23)Das, S. K., Choi, U.S., Yu,W., and Pradeep, T., "Nanofluid Science and Technology", Wily-Interscience, 2007.

Fig.2 micrograph of Al_2O_3 the nano-particles
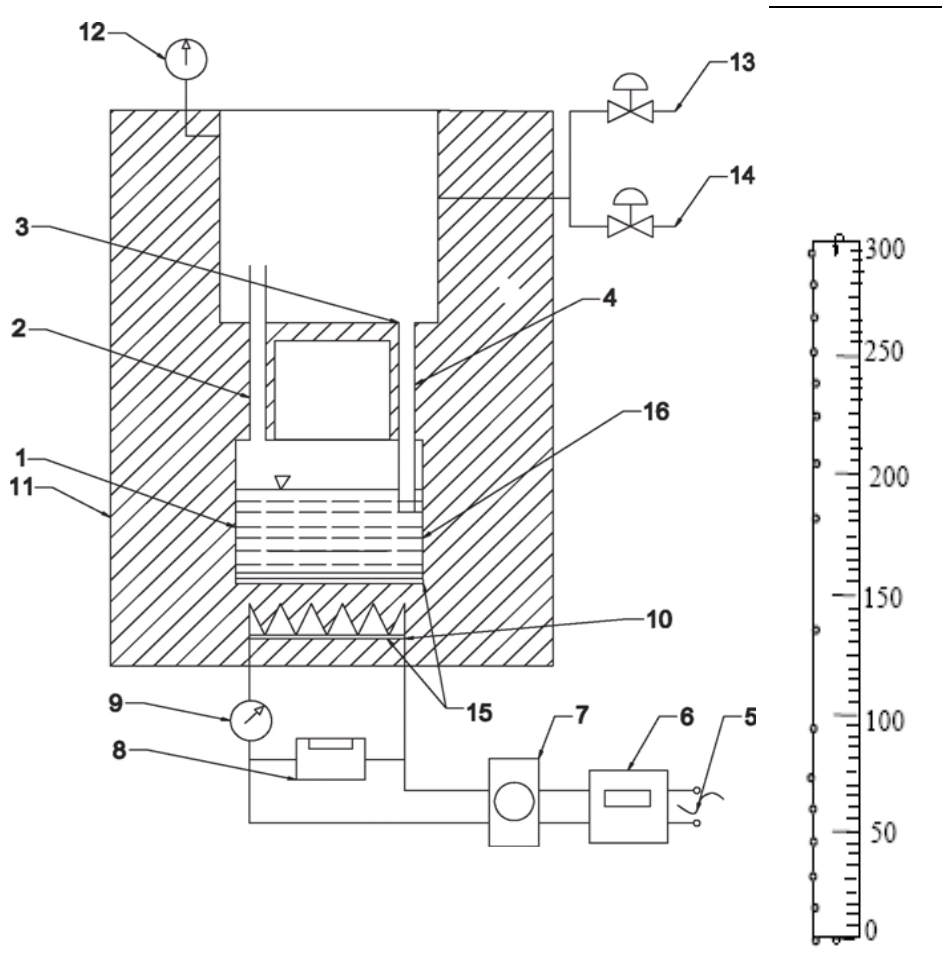
1-evaporator	5-supply	10- electric heater	14- vacuum line
2- vapor tube	6- stabilizer	11 - insulation	15 -mica sheet
3-condenser	7 - autotransformer	12- pressure gage	16-liquid in evaporator
4-liquid tube	8,9- multi-meter	13- charge line	

Figure (1) Schematic layout of the test rig used in experimental work and thermocouple distribution.
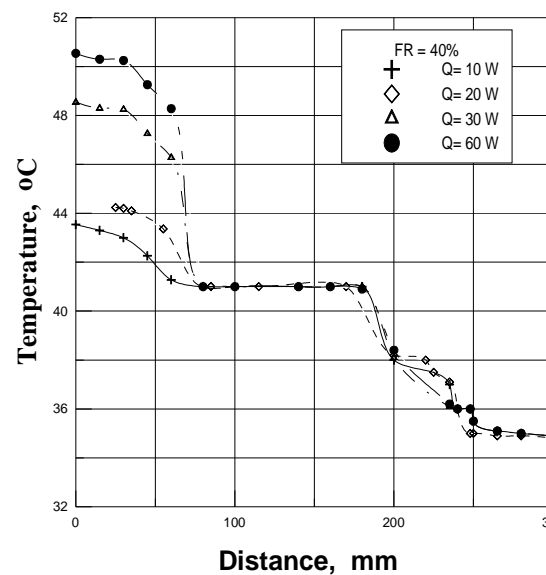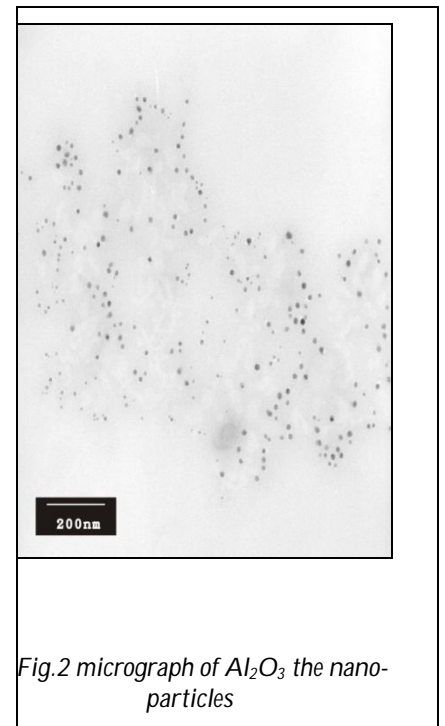
Figure 3 Temperature distribution along the thermosyphon surface for different heat input rate at operating pressure equal kPa.
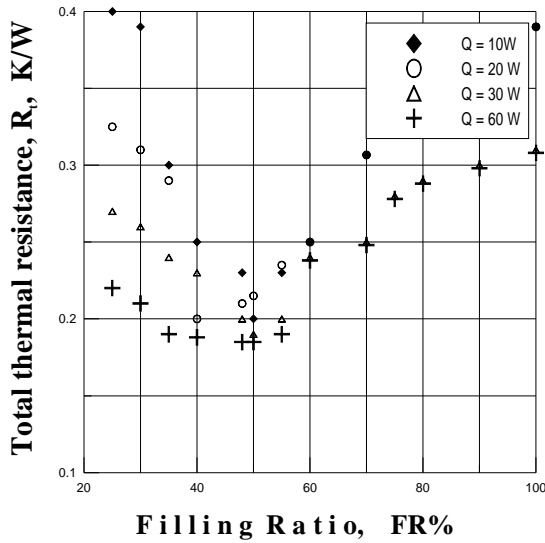
Figure 4 Variation of the total thermal resistance with filling ratio at different heat input rate
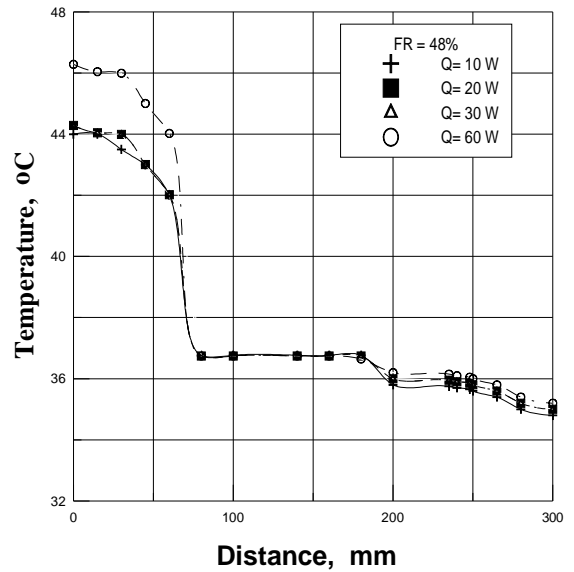
Figure 6 Temperature distribution along the thermosyphon surface for different heat input rate at operating pressure equal kPa.
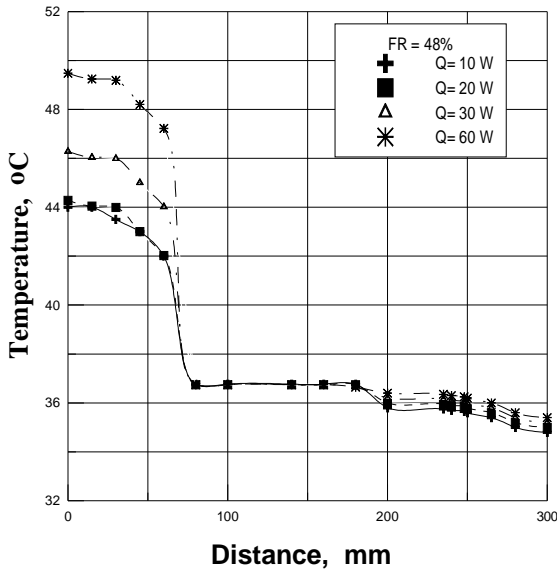
Figure 5 Temperature distribution along the thermosyphon surface for different heat input rate at operating pressure equal kPa.
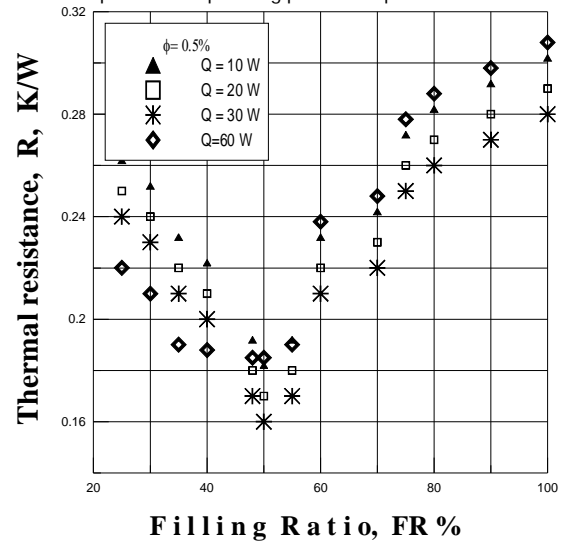
Figure 7 Variation of t hermal resistance with filling ratio at different heat input rate
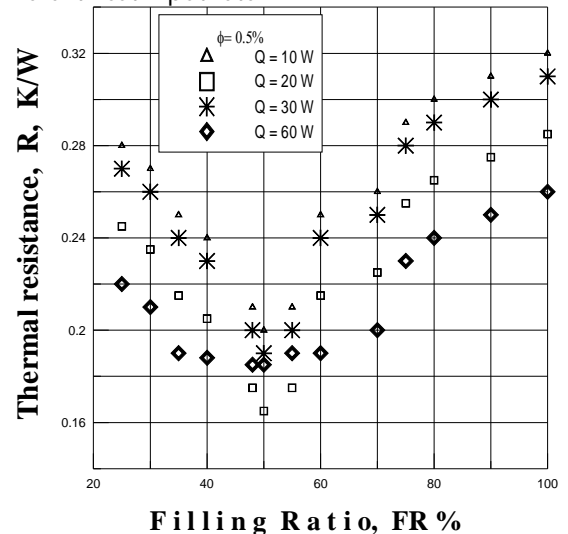
Figure 8 Variation of thermal resistance with filling ratio at different heat input rate
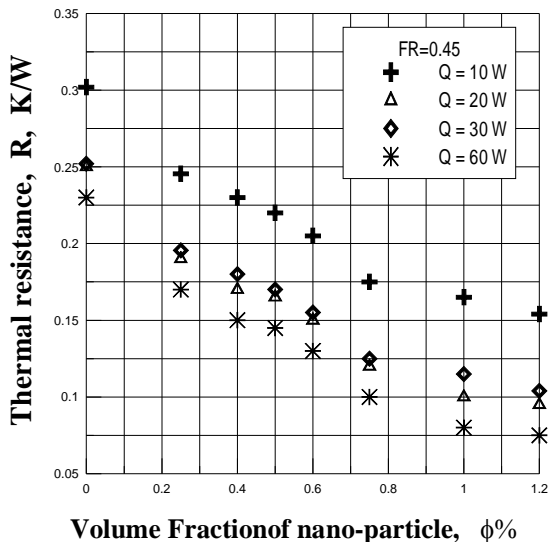
Figure 9 Variation of thermal resistance with volume fraction of nano-particles
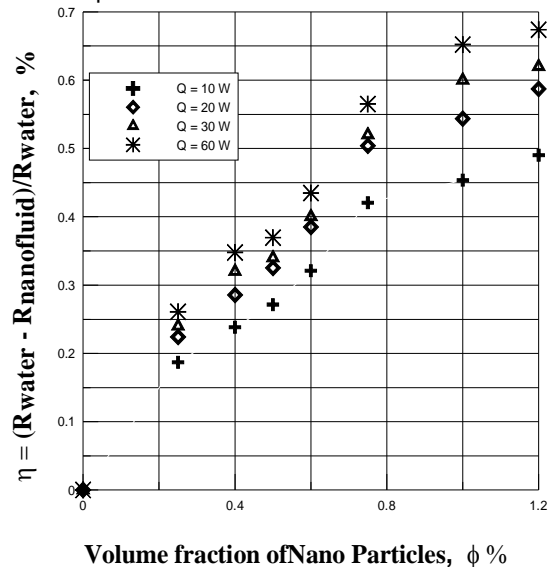
Figure 10 Reduction factor of total thermal resistance at different volume fraction of the nano-particles, ϕ
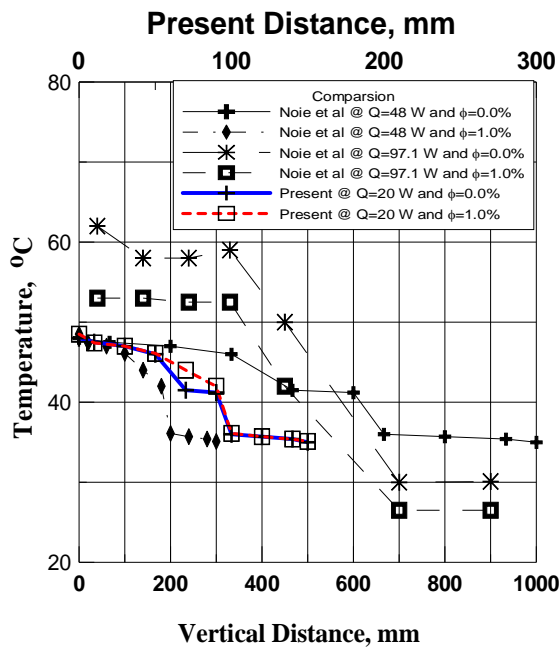
Figure 11 Comparison of the present results with available literature at different value of heat input rate

The Effect Of High Temperature On Viscosity Of Palm Oil During The Ripening Process Of Fresh Fruits

Afshin.Keshvadi*[1] Johari.Bin.Endan [1] Haniff.Harun [2] Desa.Ahmad [3] Farah Saleena[1]

[1]Department of Process and Food Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[2]Malaysian Palm Oil Board Bandar Baru Bangi 43000 Kajang, Selangor, Malaysia
[3]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia

Universiti Putra Malaysia, keshvadiafshin@gmail.com

Universiti Putra Malaysia, johariendan@gmail.com

Universiti Putra Malaysia, desa@eng.upm.edu.my

Universiti Putra Malaysia, saleena@eng.upm.edu.my

Malaysian Palm Oil Board, mhaniff@mpob.gov.my

ACKNOWLEDGEMENT

Thanks to the almighty God who had given me the strength to further my studies. I would like to convey my heartiest thanks and appreciation to the chairman of my supervisory committee, Associate Professor Ir. Dr. Johari Endan, Department of Process and Food Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM) for his encouragement and support throughout the study.

Special thanks to my supervisory committee, Professor. Dr. Desa Ahmad, Dr. Haniff Harun and Dr. Farah Saleena for their constructive comments and help and also to the Professor Dr. Miskandar, Mrs. Rosnah and Mr. Santiago for their help related to using the Malaysian Palm Oil Board (MPOB) field, Oil Analyzing Laboratory and related instruments.

Thanks to Professor. Dr. Badlishah and Mr. Soaib for help in using instruments of the Food Technology Laboratory. I am thankful to my dear mother and father. Also I'm grateful to my beloved wife, brothers, sister and brother in law for their everlasting support and courage.

Cover Letter

Dear Dr,

I wish to submit my regular article for intended publication in *International Journal of Scientific & Engineering Research* for your kind consideration.

I declare that the work submitted for publication is original, previously unpublished, and not under consideration for publication elsewhere. I also certify that all the authors have approved the paper for release and are in agreement with its content.

Address: Department of Process and Food Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM) 43400, Serdang, Selangor, Malaysia.
Mobile: +60147156207
E-mail: keshvadiafshin@gmail.com

I recently completed my graduation (PhD) from UPM. I'm so glad to work with you. My final defense is near and this experience given me confidence needed to succeed in my study. I would be grateful if my article is considered favorably.

Thank you.

Sincerely yours, Afshin Keshvadi

Abstract

This research was done on *Tenera* oil palm variety (A cross between Dura and Pisifera) on eight-year-old palms planted in 2003 at the Malaysian Palm Oil Board (MPOB) Research Station. Fresh fruit bunches were harvested and were removed the fruits from outer and inner layers of them randomly, during the ripening process between 8, 12,16 and 20 weeks after anthesis for oil extraction and determine to the affect of high temperature on palm oil viscosity as a Non-Newtonian fluid during the ripening process. The soxhlet extraction tubes with hexane were used to palm oil extraction.

Oil viscosity as a one of the important rheological properties of palm oil during the oil development in the bunch and constant function of shear rate, exhibiting Non-Newtonian fluid behaviors was measured by Rheostress, (Thermo Hakke, RS 600, Germany) with rotary parallel plates (35 mm diameter and 0.5 mm gap distance) in order to the effect of high temperature on palm oil viscosity.

The oil was sheared at a constant shear rate of 100 s^{-1} over temperatures ranging from 20 to 70 °C and viscosity was then measured.

All experiments were carried out in Microsoft Excel and MSTAT-C to statistical analysis and Regression coefficients.

KEYWORDS: Oil palm FFB, high temperature and viscosity

Introduction

In the food industry, viscosity is one of the most important parameters required in the design of technological process. On the other side, viscosity is also an important factor that determines the overall quality and stability of a food system. From the physicochemical point of view, viscosity means the resistance of one part of the fluid to move relative to another one. Therefore, viscosity must be closely correlated with the structural parameters of the fluid particles.

The raw palm oil as pressed from the fruit is a product even cruder than the commercial crude palm oil. It contains a lot of fiber, dirt, water, soluble impurities and considerable amount of debris. The average composition of the raw palm oil as received from the screw press is 40-75% oil, 10-40% water and 6-25% non-organic solids ([4]K.H Lim and Whiting, 1977). The crude palm oil is left in a clarifier to allow for the separation of oil from water. Water and its content of debris from immiscible liquids with the oil.

The heavier water and debris settle to the bottom leaving the relatively clean and light oil on the top to be skimmed off later. The effect of the operating conditions, such as temperature and shear rate on the density is relatively small compared to their effect on viscosity ([1]B.Marcia, et al. 2002). Hence, the density may be treated as a constant and the viscosity is considered the main operating condition influencing the separation process. Regardless of the numerous researches conducted previously, a reliable model to predict the effect of the independent operating conditions on the viscosity of crude palm oil is still lacking. This is due to the difficulty in studying crude palm oil, which behaves as a non-Newtonian fluid ([4]K.H Lim and Whiting, 1977), and also due to the composition complexity of the mixture itself (oil, water and debris). The fact is that, in all previous works to determine the design equations of crude palm oil settlers, the viscosity of crude palm oil mixture was always assumed, incorrectly to be an additional variable, independent of temperature and shear rate ([4,5]K.H Lim and Whiting, 1977.1982). Previous experimental work proved that crude palm oil is a non-Newtonian fluid. The viscosity was shown to change with shear rate applied. The power law, Eq.1, is widely used in literature to express the effect of shear rate on the viscosity of non-Newtonian liquids at constant temperature ([6]Aksel, 2002. and [2]Wallis, 1969).

$$\mu_\alpha = \kappa \gamma^{n-1} \qquad\qquad \text{(Eq. 1)}$$

The viscosity of crude palm oil was found to decrease with shear rate ([4]K.H Lim and Whiting, 1977), this behavior is called shear thinning, which is characterized by the value of the power index n being less than unity.

On the other hand, temperature changes also affect the apparent viscosity of crude palm oil. Like any other liquid, the viscosity of crude palm oil is expected to reduce as temperature increases. This phenomenon is explained by the reduction of the inter-molecular attractions with increasing temperature.

In other research, the flow behaviors of vegetable oils were measured over temperature ranging 20 to 70 °C to finding effect of high temperature on vegetable oil viscosity ([3]Juyoung K. *et al.* 2009).

Materials and Methods

Fruit Collection

Two eight-year oil palms (*Elaeis guineensis*), *Tenera* variety were obtained for each stage of ripening time 8,12,16 and 20 weeks after anthesis (total of 8 palms with almost same stage of bunch's anthesis) from MPOB Research Station.

Two bunches were harvested from two palms (one bunch from each palm) for each stage of ripening, from January till may 2010 and then fruits (total of 60 samples) were collected during the ripening process from outer and inner layers randomly, to oil extraction from mesocarp and kernel (after 12 weeks with kernel formation).

Oil Extraction

The samples were weighed and chopped, then were dried in the oven under 70°C for a day to remove the water in the fruits. The dry mesocarp and kernels were weighed and blended to get particle separately .The Oil was extracted in soxhlet extractor available in MPOB oil analyzing laboratory, using chemical solvent namely hexane.

Viscosity Measurement

The flow behaviors of palm oil during the ripening stages (8,12,16 and 20 weeks after anthesis) were measured by using a controlled-stress rheometer (Rheostress RS600, Thermo Hakke, Germany) with rotary parallel plates (35 mm diameter and 0.5 mm gap distance) at constant shear rate 100 s^{-1}, Over temperatures ranging from 20 to 70 °C. Also, after 12 weeks, Simultaneous with kernel formation, the kernel oil samples from 12,16 and 20 weeks after anthesis were impressed by high temperature to investigate of oil viscosity changes.

Results and Discussion

The Flow Behaviors of Mesocarp Oil Viscosity at Different Sampling Times

To analyze and compare of the changes in mesocarp oil viscosity during the ripening phase from 8 to 20 weeks after anthesis were analyzed and compared. The results of variance analysis showed significant differences in mesocarp oil viscosity between sampling times (8, 12, 16 and 20 weeks after anthesis) ($p < 0.01$), (Table 1).

Table 1. Analysis of variance for mesocarp oil viscosity at different times of sampling

Source	Degrees of Freedom	Sum of Squares	Mean of Squares
Replication	2	0.3311	0.55 ns
Time of sampling	3	9.2278	3.0759 **
Error	6	0.403	0.0992
C.V		13.63	

*, **Significant levels at 5 and 1%, ns – not significant

The mesocarp oil viscosity over sampling time showed that with growth over time mesocarp oil viscosity increased significantly. Minimum viscosity at the beginning stages of fruit development (8 weeks) was 0.019 Pascal, while at 12 weeks of growth it was

7

increased to 0.040 Pascal. At 16 weeks after anthesis the mesocarp oil viscosity was further increased by 18.6% to 0.047 Pascal. The maximum mesocarp oil viscosity recorded at the end of ripening at 20 weeks after anthesis was 0.91 Pascal. The difference in mesocarp oil viscosity between 12 and 16 weeks was not significant (Table 2).

Table 2. Mean comparison of viscosity during the ripening stages
(Means with the same letters are not significantly different at $p < 0.01$)

Time of sampling	Viscosity (pas)
8 weeks	0.019 c
12 weeks	0.040 b
16 weeks	0.047 b
20 weeks	0.91 a

Figure 1, Shows the changes of mesocarp oil viscosity by passing of growth periods until end of ripening stage.

$$y = 0.2105x^2 - 0.7845x + 0.6365$$
$$R^2 = 0.93$$
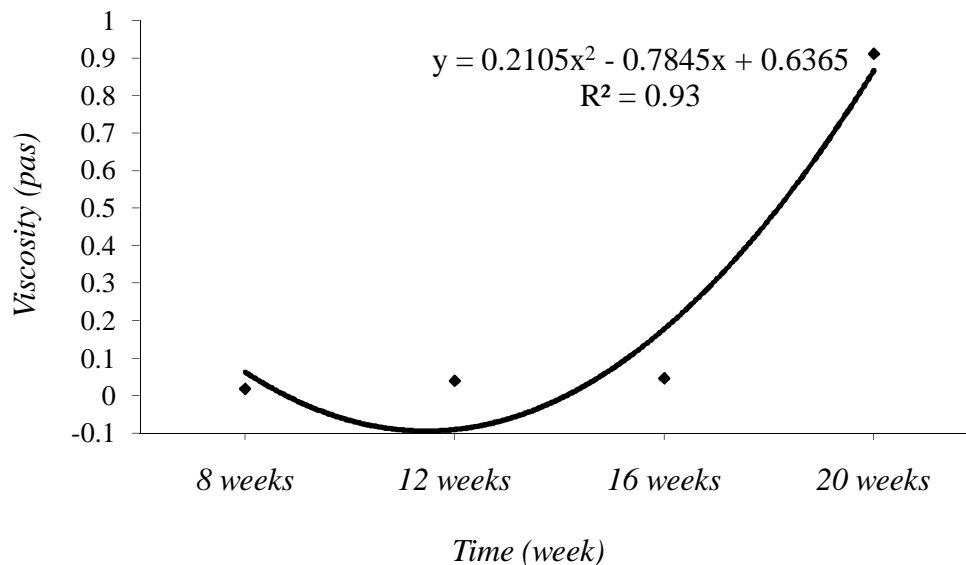
Figure 1. Effect of sampling time on flow behavior of mesocarp oil viscosity
($p < 0.01$)

8

The results showed that the changes in mesocarp viscosity followed an polynomial function; while the trend in mesocarp oil viscosity showed an upward movement, although, somewhat slow between 8 - 16 weeks, but subsequently between 16-20 weeks (ripening stage) the oil viscosity increased significantly at an polynomial rate. There was a high positive correlation (R^2= 0.93) between mesocarp oil viscosity and increase in ripening. The process showed that the main part of fruit content changes occurred between 16-20 weeks. This was recognized as the sensitive stage when any changes in environmental conditions will have a major effect on oil composition and oil yields.

The Flow Behaviors of Mesocarp Oil Viscosity at Different Temperatures

The flow behavior of mesocarp oil viscosity at different sampling times was investigated as a function of temperature ranging from 20 to 70°C. All of the four mesocarp oil samples (at 8, 12, 16 and 20 weeks after anthesis) exhibited the same viscosity pattern over temperature, which was a non-linear decrease in viscosity with increasing temperature. At each sampling time, the oil viscosity value decreased with increase in temperature from 20 – 70°C. The reductions recorded were as follows: at 8 weeks from 0.019 to 0.008 Pas, at 12 weeks from 0.040 to 0.013 Pas, at 16 weeks from 0.047 to 0.014 Pas, and at 20 weeks from 0.91 to 0.016 Pas.

Regression analysis on the changes in oil viscosity with increasing temperature showed that these changes over sampling time followed a non-linear exponential function. The high correlation between increasing temperature and decreasing viscosity in equations showed the high accuracy of the prediction equation to explain the changes in viscosity under increasing temperature (Figures 2 to 5).

9
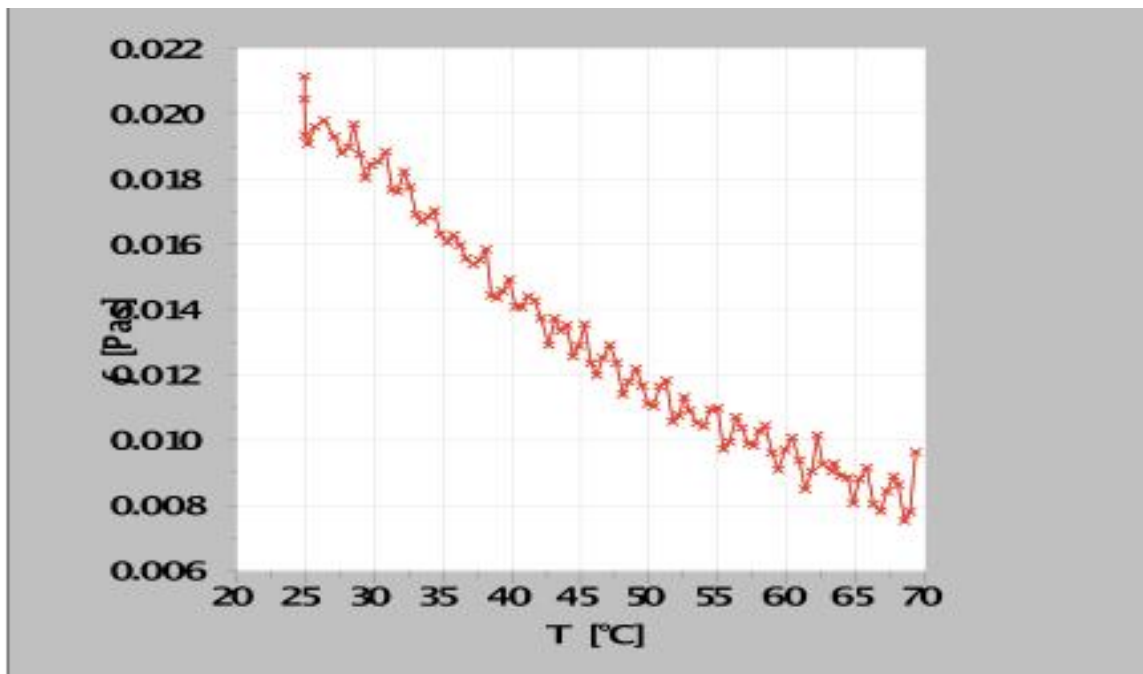
Figure 2. Effect of temperature on the flow behavior of mesocarp oil viscosity in 8 weeks after anthesis
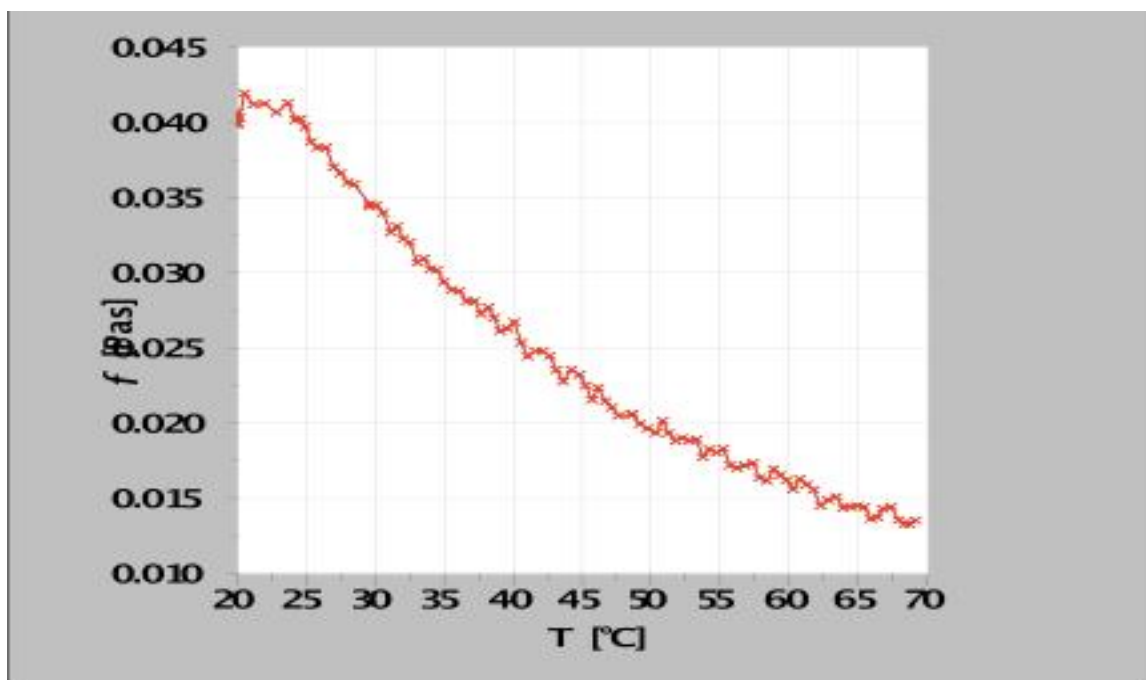
Figure 3. Effect of temperature on the flow behavior of mesocarp oil viscosity in 12 weeks after anthesis
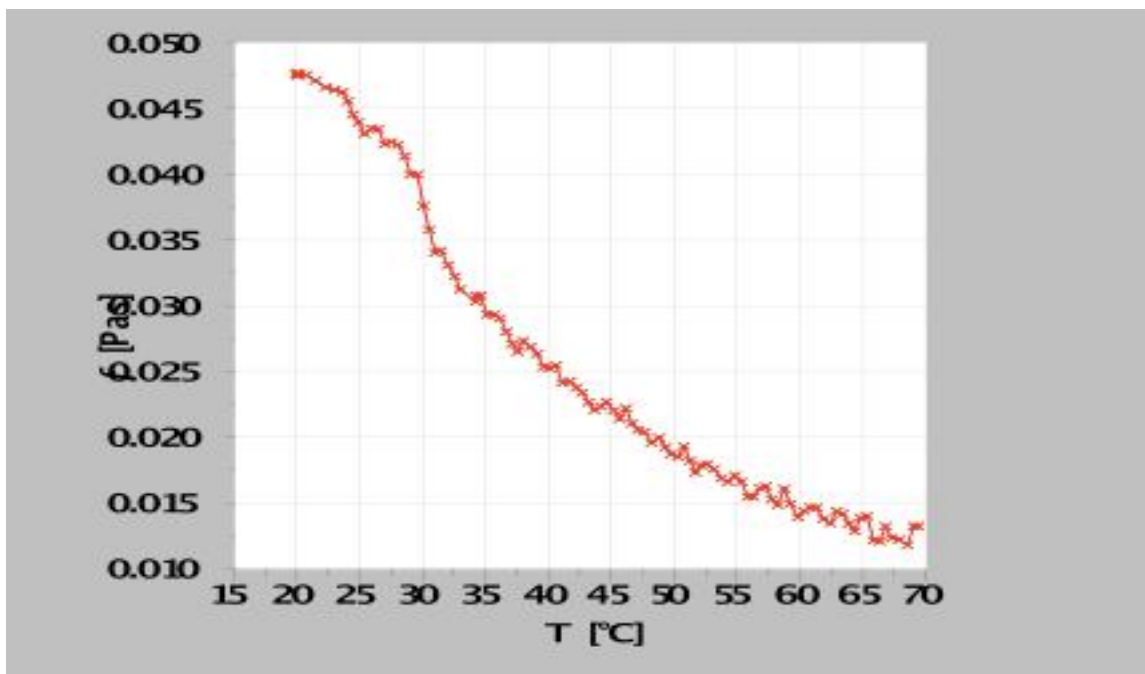
Figure 4. Effect of temperature on the flow behavior of mesocarp oil viscosity in 16 weeks after anthesis
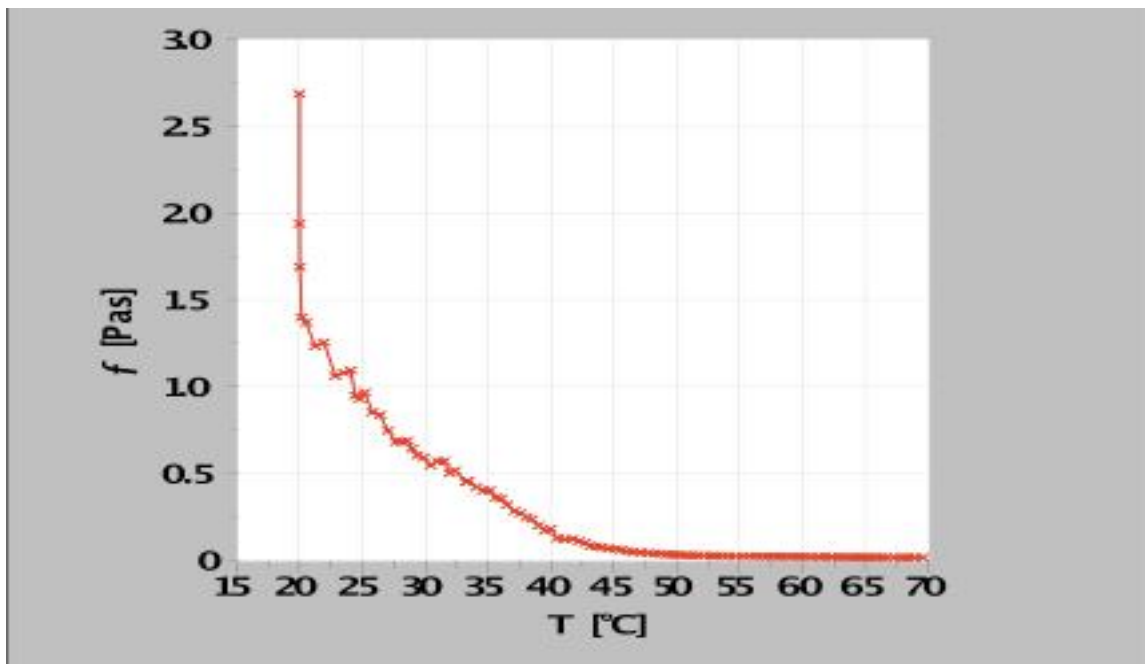
Figure 5. Effect of temperature on the flow behavior of mesocarp oil viscosity in 20 weeks after anthesis

The Flow Behaviors of Kernel Oil Viscosity at Different Temperatures

The flow behaviors of kernel oil viscosity at different sampling times were investigated as a function of temperature ranging from 20 to 70° C. All of the three kernel oil samples extracted at 12, 16 and 20 weeks after anthesis exhibited the same viscosity pattern over temperature, which was a non-linear decrease in viscosity with increasing temperature. At each sampling time, kernel oil viscosity value decreased when the temperature was increased between $20 - 70^{o}$C. The reductions recorded were as follow: at 12 weeks from 0.018 to 0.007 Pas; at 16 weeks from 0.035 to 0.011 Pas; and at 20 weeks from 0.046 to 0.012 Pas.

Regression analysis of changes in kernel oil viscosity with increasing temperature showed that these changes over sampling time followed a non-linear exponential function. The high correlation between increasing temperature and decreasing viscosity in the equations showed the high accuracy of above equation to explain the changes in viscosity under temperature changes (Figures 6 to 8).
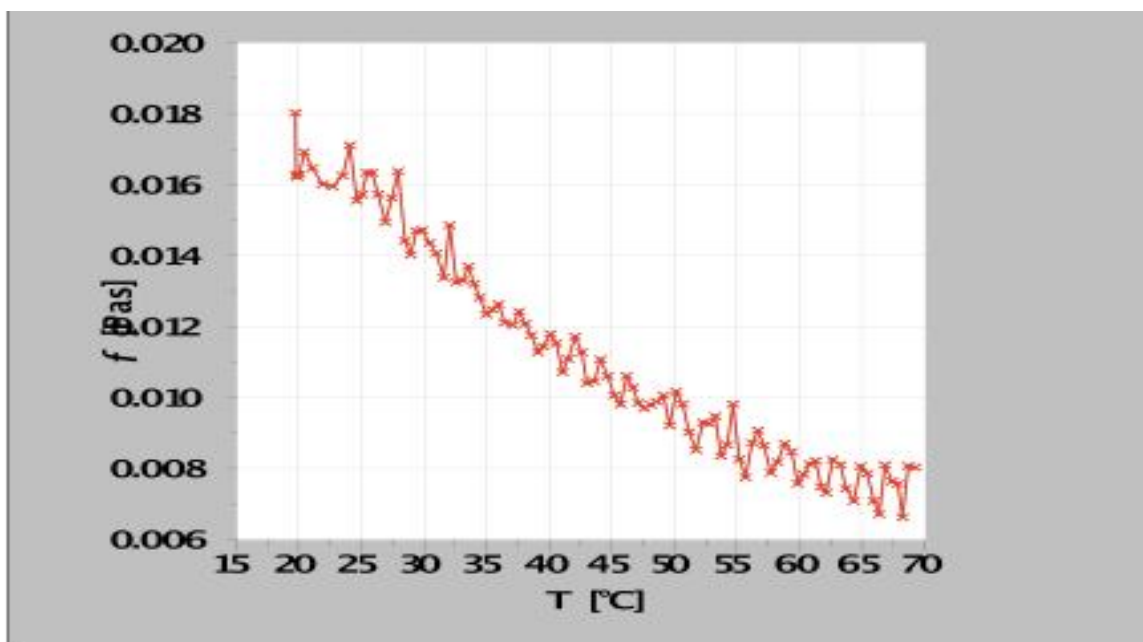
Temperature ^{o} C

Figure 6. Effect of temperature on the flow behavior of kernel oil viscosity in 12 weeks after anthesis

12
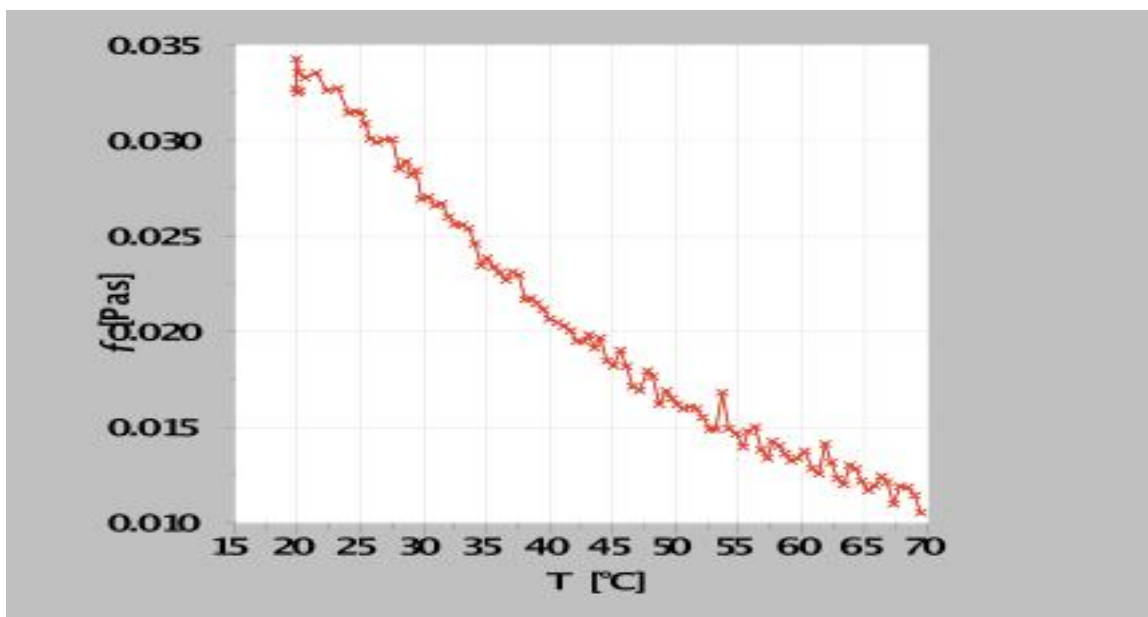
Temperature ^{o} C

Figure 7. Effect of temperature on the flow behavior of kernel oil viscosity in 16 weeks after anthesis
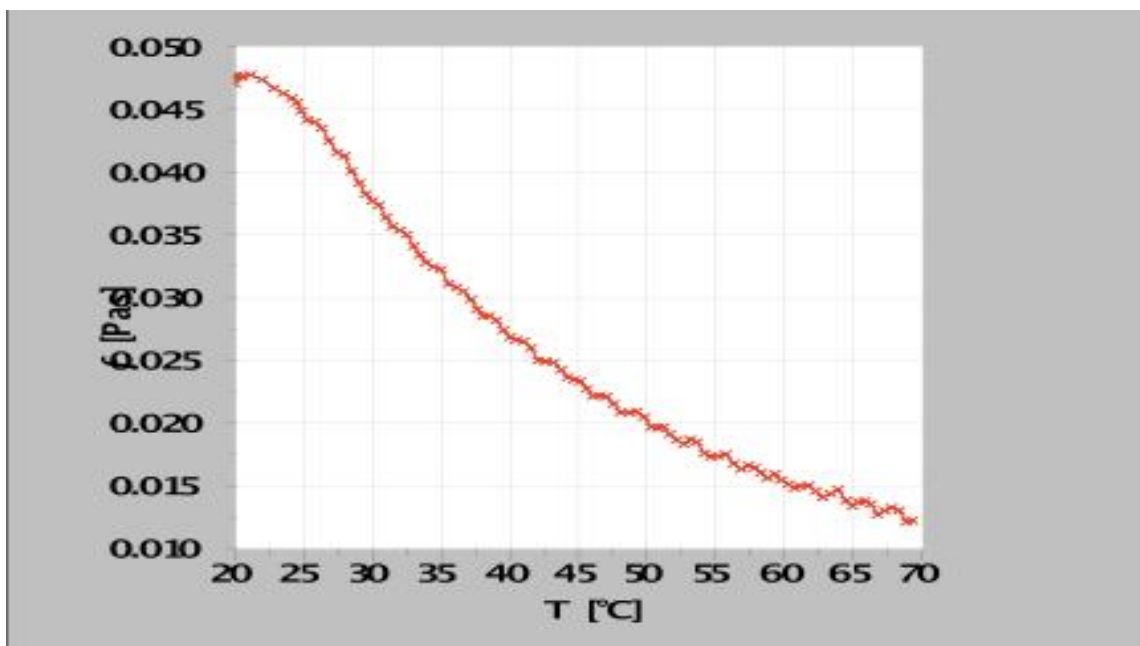
Temperature ^{o} C

Figure 8. Effect of temperature on the flow behavior of kernel oil viscosity in 20 weeks after anthesis

Conclusion

All mesocarp oil samples (at 8, 12, 16 and 20 weeks after anthesis) had a non-linear decrease in viscosity with increasing temperature. This reduction was recorded as follow; for oil in 8 weeks from 0.019 to 0.008 Pas, oil in 12 weeks from 0.040 to 0.013 Pas, for 16 weeks from 0.047 to 0.014 Pas and 20 weeks from 091 to 0.016 Pas in temperature ranged $20 - 70^o$ C. All of three oil samples (kernel oil at 12, 16 and 20 weeks after anthesis) exhibited the same viscosity pattern over temperature, which was a non-linear decrease in viscosity with increasing temperature. This reduction was recorded as follow; for oil in 12 weeks from 0.018 to 0.007 Pas, for 16 weeks from 0.035 to 0.011 Pas and for 20 weeks from 0.046 to 0.012 Pas in temperature ranged $20 - 70^o$ C.

Analysis of regressive model of changes of oil viscosity, under temperature increasing showed that these changes on different four sampling times followed a non-linear and exponential function. High correlation between increasing temperature and decreasing viscosity in equations showed the high accuracy of above equation to explain the changes of viscosity under temperature changes.

([3]Juyoung K and Deok N 2009) showed that the highest viscosity was observed in hazelnut oil, followed by olive, canola, corn, soybean, sunflower and grape seed oils. In addition, a high correlation ($R^2 = 0.94$) demonstrated that the flow behaviors of vegetable oils were positively governed by their major components also shown that a more rapid change in viscosity with temperature was observed in the oils containing more double bonds ($R^2 = 0.71$).

As we can see, since the stability of oils over time could vary depending on the oil type and further research would be necessary to investigate how pertinent parameters of the oils and their consequent oil uptake patterns are changed by frying conditions (time and temperature) and repeated use.

REFERENCES

[1]B.Marcia,S.Gnter, J.S.Milan, O.A.Elseoul, Vegetable oils-based micro emulsions: Formation, Properties and Application for Soil Decontamination, colloid polymer sci, vol.280, pp.973-983, 2002.

[2]G.B.Wallis, One-Dimensional Two-phase Flow, McGraw-Hill, New York.1969.

[3]Juyoung Kim a, Deok Nyun Kim a, Sung Ho Lee b, Sang-Ho Yoo a, Suyong Lee a, Department of Food Science and Technology and Carbohydrate Bio product Research Center, Sejong University, South Korea. 2009.

[4]K.H Lim and .A.M Whiting, The influence of Non-Newtonian behaviour of crude palm oil on the design of the clarification station Equipment. The Proceeding of the Malaysian international symposium on palm oil proceeding and marketing, international developments in palm oil, Malaysia, 1977.

[5]K.H.Lim,Theory and Application of Clarification in Palm Oil Mills. Harrison and Crosfield (M)Sdn.Bhd.Malaysia,1982.

[6]N.Aksel, A Model for Bulk Viscosity of a Non-Newtonian Fluid, Continuum Mechanics and Thermodynamics, Vol.7, No 3,pp.333-339, 2002.

A New Old Concept for Bussines Progress: Physical Capital Maintenance

Ionel Jianu, Iulia Jianu, Liviu Geambasu

Abstract—This study aims at promoting the physical capital maintenance concept, as a solution for the development of economic entities, by demonstrating the reliability of this concept in obtaining performance at the level of economic entity due to the technological upgrade. Physical capital maintenance is a relatively new concept in the accounting theory and practice. Very fast, convenient and handy indeed, we associate production capacity with technology, devices, networks, with the staff and its knowledge, and it seems that we are not far away from the truth. Physical capital maintenance requires the recognition of the profit only if the entity was able to maintain its current activity level which is measured by production capacity. This article aims to highlight the results that the practical application of the concept of physical capital maintenance strictly related to the technological aspect has in the organizational behavior changing. In this sense, the empirical research will demonstrate the changes in the current forms of business management through the use of the physical capital maintenance concept.

Index Terms— physical capital maintenance, technological upgrade, change management, change in organizational behavior

———————————————— ◆ ————————————————

1 INTRODUCTION

Technological advance has marked and will mark the historical evolution of the world's societies and peoples. It was, it is and it will be a factor of development for the human activities, a factor of individual and institutional emancipation, a factor of change on the fly of the ideas to do business and to think economically. It is the domain which has produced and produces the greatest paradigm shifts in business and economic thinking, the domain for whose relaxation energies and egos are compressed in research and development departments.

The world we live in is conditioned and animated by the peoples cultures and histories that are united on civilization aspect. Talking about more cultures and just one civilization, we can say, without the fear of being wrong, that we are witnessing a major cultural dispute, attempts to promote its own culture through everything that one nation or another has best and most valuable. On this background, we are lately seeing a fierce competition for dominance of technological development, the only capable to assure access, strong presence and to increase all markets for economic entity.

The present work brings into discussion the concept of physical capital maintenance which could be a viable alternative to provide the funding necessary to ensure continuity of the economic entity activity (implicit for investment policy – technological update).

———————————————

- *Ionel Jianu, Lecturer PhD Candidate, Bucharest Titu Maiorescu University, Romania, Email: ionel_j@yahoo.com*
- *Iulia Jianu, Lecturer PhD, Bucharest Academy of Economic Studies, Romania, Email: jianu.iulia@cig.ase.ro*
- *Liviu Geambaşu, PhD Candidate, Bucharest Academy of Economic Studies, Romania, Email: liviugeambasu@gmail.com*

2 PRIOR WORK

The permanent technological update, at the level of economic entities, is possible by approaching and implementing at their level the concept of physical capital maintenance. Physical capital maintenance is a relatively new concept in accounting theory and practice, not through its origins (the 30s of last century), but through the lack of problematic in the literature, situation that has leaded to the absence of an algorithm to enable the application of the concept in the economical entities practice. Extensive discussions on the physical capital maintenance were in the mid '70s when the U.S. standardization body wanted to improve the conceptual framework [1]. Due to the combined efforts of the international body for issuing International Financial Reporting Standards Board (IASB) on the harmonization of international accounting, physical capital maintenance concept was introduced in 1989 in the IASB's concept as an accounting alternative to financial capital maintenance, but from 1989 until the present there are no scientific papers to explain the consequence of maintaining the physical capital on the performance of the economic entity [2].

Physical capital maintenance requires recognition of the profit only if the production capacity of the entity at the end of the period is higher than the production capacity at the beginning of the period, after excluding any capital contributions or distributions from or to shareholders. The productive capacity is primary provided, at least at present, by the technology of the economic entity. And to recognize the profit (the rationale of any economic entity) it is necessary to increase the productive capacity. In our opinion this thing is possible through innovation, Beach [3] published the results of a study to reflect the

results of the measurement technique of physical capital in the U.S. rail industry. The study was based on official data from journals, statistical institutes and different associations in the period 1920-1929. For the valuation of physical capital it was used as measurement unit the number of years of using tangible assets of American railways. The productive capacity involves technology, and the technology, translated into specific terms of accounting (as management science) means tangible assets.

Break [4] believes that physical capital maintenance should be performed only by maintaining the physical characteristics of tangible assets held by the entity to carry out productive work, without trying the monetary evaluation of the physical capital that should be maintained. According to this concept, any change in the existing tangible assets leads to a real positive or negative result. When an entity acquires or produces a new tangible asset, it is added to the real profit. But changes on monetary assets, that do not affect tangible assets used in the productive activity of the entity, do not affect the physical capital.

Gynther [5] associates physical capital with the production capacity of an entity which assumes that all assets used in production must be evaluated individually at their current cost, and where no current cost can be determined, then specific price indices must be used.

3 RESEARCH METHODOLOGY

The paper is based on a positivist research to demonstrate the changes in the current forms of business management through the use of the physical capital maintenance concept (its substantiation on production capacity, discussed here under technological aspect). In our approach to validate the formulated hypothesis, the development of the economic entity will be reflected by the turnover. We will try therefore to establish a relationship between changes in the value of fixed assets and changes in turnover. We believe that there are more objective and representative the considerations related to the involution of the two concepts in question, the reduction of fixed assets, leading to the reduction of turnover. It is a reality of the current economic environment, in which the selling of assets deprives the entity of the technological factor that is so necessary to the development of economic entities. The study will be based on empirical research in order to validate the following hypothesis:

Hypothesis: There is a relationship between the physical capital maintenance and the development of economic entities.

We want to highlight the fact that the decrease of the fixed assets, at the level of an entity, leads to the decrease of the turnover, and consequently to the dependency between the decrease of physical capital and the entity de-capitalization. This confirms the relationship between physical capital maintenance and development of economic entities.

The testing of hypothesis was done using one of the most popular econometrics tools simple regression. The regression analysis shows whether there are relationships or not between the categories in question and how strong they are.

4 RESULTS AND DISCUSSIONS

4.1. The Quantitative Analysis

Wanting to capture the realities of the Romanian economic environment, we conducted the study on economic entities in Romania below the limit given by the weak diversification of the financial information provided by their annual reports (not having information about the type, nature, number, technical characteristics, etc.). To obtain the data with which we worked on the study, we accessed the Bucharest Stock Exchange (BSE) website. The valuation of physical capital was based on the value of fixed assets and on turnover. The data were taken from the financial information of economic entities published on the website. A number of 26 economic entities listed on the BSE which have a decrease in fixed assets and turnover have been under study for testing the correlation between decrease of fixed assets and decrease of turnover (Appendix 1).

4.2. Hypothesis Testing

The regression function presented in Figure 1 shows that in the case of a change with a 1% of the negative variation of the fixed assets, the negative variation of the turnover changes by 6.58%.

Regression function:

$$Y = 59721457 + 6,58X \qquad \text{(Formula 1)}$$

Regression Statistics			
Multiple R	0,5067		
R Square	0,2568		
Adjusted R Square	0,2258		
Standard Error	2E+08		
Observations	26		
ANOVA	Regression	Residual	Total
df	1	24	25
SS	3,22E+17	9,32E+17	1,25E+18
MS	3,22E+17	3,88E+16	
F	8,2909		
Significance F	0,00825		
	Intercept	X Variable 1	
Coefficients	59721457	6,58853	
Standard Error	4,4E+07	2,28816	
t Stat	-1,34454	2,879406	
P-value	0,191346	0,00825	

Fig.1 Summary output for simple regression between negative variation of the fixed assets and negative varia-

tion of the turnover

Since Multiple R has a positive value close to 0.5, this shows that between negative variation of fixed assets and negative variation of the turnover is a direct correlation of medium intensity. R Square shows us that 25,68% of the negative variation of the turnover is explained by the negative variation of the non-current assets. Since Significance F has a low value, below the threshold limit of 0.05, and F is a high value, we can accept the simple regression model presented in Formula 1.

4.3. The Qualitative Analysis

Physical capital maintenance concept requires an entity to recognize profit only if it maintained the production capacity and just for the share that surpasses the physical capital originally invested. This study showed that there is a connection between the physical capital maintenance and the development of an economic entity, which entitles us to say that, by physical capital maintenance at the level of economic entities, the performance of these entities will be much higher and the risk of bankruptcy because of fictitious dividends distribution will be as nonexistent. It is obvious that it is easily avoided the confusion that can occur by raising the role of physical capital maintenance concept to preserve the substance of technology, equipments in production, being noted the aspect of maintaining their power to participate in achieving performance. This comes from today's economic reality that offers the chance only to those who operate with equipment, technology and new concepts.

We managed to demonstrate that there is a strong link between the level of fixed assets and the performance of the economic entity, which gives us a chance to fight for change in the way of doing business, to be effective and efficiency on a growing and selective international market. The confirmation of the working hypothesis supports our attempt to demonstrate that there is a connection between physical capital maintenance and the development of economic entities. By this, we consider necessary the adjustment of decisions, of organizational and decisional climate to the new concept of physical capital maintenance and through this to continuous technological change and not to disinvestment. We consider it a strong enough argument for reconsidering the physical capital maintenance concept, having as actionable premises the socio-economic realities in which the entity operates.

Lately we are witnessing the divestment (sale of fixed assets) with the purpose to fill the gaps caused by a bad management. But we forget that they are short-term results and that failing to replace the sold ones with other technologies, more efficient, can bankrupt any business. Managers need a reliable tool for a new way of doing business, a tool that does not lead to disinvestment, and even more, that allow a continuous technological change and thus the development of the economic entity.

5. CONCLUSIONS AND SUGGESTIONS

The physical capital maintenance is a business management tool very useful to the senior management of the economic entity. It represents a paradigm change (a possible renouncement of financial capital maintenance), a solution to the current economic environment in which the inflation periods alternate more often than ever. It is the concept which does not allow the payment of unjustified dividends because it takes into account the real profit. We have a new managerial tool for business management. The paper highlights the quality of this new form of business management, its reliability in achieving performance of the economic entity, as well as the advantage offered by ensuring the capabilities required by continuous technological change.

Our study provides a useful basis for discussion, of further in-depths and developments, necessary for a complete perspective of the determinants of this concept. A first assessment can be made in connection with the provision of production capacity, at least in terms of technology, using the concept in question. It is widely accepted that the huge advances in living standards, caused by global economies, is largely due to the technological progress. Now, there are produced technologies that haven't even been imagined by the generations of the 1800s. The physical capital maintenance concept creates the idea, produces change, and ensures the framework for continuous technological change.

The cornerstone of economic growth is the technological progress. The standard of living is much higher now in most nations of the world than a hundred years ago. The reason is related to the fact that productivity, at the level of the economic entities, increased significantly and the cause of this progress is seen by the domain specialists as being technological changes that took place.

ACKNOWLEDGMENT

REFERENCES

[1] J.J. Young, "Making ul users", *Accounting, Organization and Society*, vol. 31, no. 6, pp. 579-600, 2006

[2] IASB, *Standarde Internaționale de Raportare Financiară*, Bucharest: Ceccar, 2009.

[3] E.F. Beach, „A measure of physical capital", *The Review of Economics and Statistics*, vol. 20, no. 1, pp. 11-20, 1938

[4] G.F. Break, „Capital maintenance and the concept of income", *Journal of Political Economy*, vol. 62, no. 1, pp:. 48-62, *1954*.

[5] R.S.. Gynther, „Capital Maintenance, Price Changes, and Profit Determination", *The Accounting Review*, vol. 45 , no. 4, pp. 712-730, 1970.

APPENDIX 1: The negative variation of the fixed assets and the negative variation of the turnover

No.	Entities' name	Non current assets for the year 2009	Non current assets for the year 2008	Negative variation of the non current assets (X - independent variable)	Turnover for the year 2009	Turnover for the year 2008	Negative variation of the turnover (Y - dependent variable)
1	ALRO S.A.	1.529.251	1.613.272	-84.021	1.410.482	1.968.016	-557.534
2	ARMATURA S.A.	18.780	19.978	-1.198	20.277	29.960	-9.683
3	AZOMURES S.A.	223.274	234.021	-10.747	743.310	1.144.100	-400.790
4	BOROMIR PROD SA	93.390	94.103	-713	133.669	162.445	-28.776
5	COMCM SA CONSTANTA	274.255	279.386	-5.131	38.568	64.422	-25.854
6	COMELF S.A.	30.455	32.054	-1.599	71.057	125.959	-54.902
7	COMPA S. A.	354.929	356.485	-1.556	253.632	310.818	-57.186
8	CONTOR GROUP S.A. Arad	58.190	59.173	-983	43.873	44.533	-660
9	DAFORA SA	325.506	336.952	-11.446	185.652	269.018	-83.366
10	GRUPUL INDUSTRIAL	7.790	10.267	-2.477	9.524	12.570	-3.046
11	IMPACT DEVELOPER	87.624	125.657	-38.033	54.233	151.688	-97.455
12	MECANICA CEAHLAU	25.832	31.898	-6.066	17.394	31.703	-14.309
13	MECHEL TARGOVISTE	234.344	255.019	-20.675	529.628	1.107.164	-577.536
14	MEFIN S.A.	15.464	21.990	-6.526	22.945	27.405	-4.460
15	MJ MAILLIS ROMANIA	24.025	31.192	-7.167	29.501	66.228	-36.727
16	OIL TERMINAL S.A.	170.699	173.650	-2.951	127.484	134.573	-7.089
17	OLTCHIM S.A. RM.	1.098.810	1.106.990	-8.180	1.077.519	1.946.943	-869.424
18	PRODPLAST S.A.	10.057	11.149	-1.092	46.001	66.133	-20.132
19	ROMCARBON SA BUZAU	181.598	183.243	-1.645	76.880	93.812	-16.932
20	SANTIERUL NAVAL	56.533	68.863	-12.330	99.895	125.967	-26.072
21	SC TRANSILVANIA	133.994	137.556	-3.562	28.866	58.940	-30.074
22	SINTEZA S.A.	160.722	162.198	-1.476	12.227	12.287	-60
23	SOCEP S.A.	62.601	64.470	-1.869	45.842	58.818	-12.976
24	T.M.K. - ARTROM S.A.	423.412	437.148	-13.736	441.697	678.876	-237.179
25	UAMT S.A.	35.418	38.043	-2.625	36.026	46.116	-10.090
26	VES SA	19.935	20.937	-1.002	41.312	51.002	-9.690

Improved Audio Watermarking Using DWT-SVD

N.V.Lalitha, G.Suresh, Dr.V.Sailaja

Abstract— Digital audio watermarking involves the concealment of data within a discrete audio file. Applications for this technology are numerous. Intellectual property protection is currently the main driving force behind research in this area. In this paper we present an efficient audio watermarking algorithm in the frequency domain by embedding an inaudible audio water mark. Comparison of two different algorithms i.e. Discrete Cosine Transform (DCT)-Singular Value Decomposition (SVD) and Discrete Wavelet Transform (DWT)-SVD is presented here. The effectiveness of these algorithms is verified by conducting experimentation. Experimental results show that the watermarked audio has good imperceptibility and is robust against different kinds of attacks, such as noise adding, re-sampling, cropping.

Index Terms— Audio Watermarking, Attacks, DCT, DCT-SVD, DWT, DWT-SVD, MOS evalauation, Singular Value Decomposition.

———————————————— ◆ ————————————————

1 INTRODUCTION

THE rapid development of the Internet and the digital information revolution cased significant changes in the global society, ranging from the influence on the world economy to the way people nowadays communicate [1]. Digitizing of multimedia data has enabled reliable, faster and efficient storage, transfer and processing of digital data. It also leads to the consequence of illegal production and redistribution of digital media. Digital watermarking is identified as a partial solution to related problems which allow content creator to embed hidden data such as author or copyright information into the multimedia data [2]. In cryptographic techniques significant information is encrypted so that only the key holder has access to that information. Once the information is decrypted the security is lost. Information hiding is unlike cryptography, message is embedded into digital media, which can be distributed and used normally. Information hiding doesn't limit the use of digital data. Within past few years several algorithms for embedding and extraction of watermark in audio sequence have been published [3-7]. Almost all audio watermarking algorithms work by exploiting the perceptual property of Human Auditory System (HAS). The simplest visualization of the requirements of information hiding in digital audio is possible via a magic triangle [8]. Inaudibility, robustness to attacks and the watermark data rate are in the corners of the

magic triangle. In order to satisfy the requirements of magic triangle, watermarks are seen embedded in Fourier domain [3], time domain [4], sub-band domain [6], wavelet domain [7] and by echo hiding [5].

2 DCT, DWT & SVD TECHNIQUES

The DCT and DWT transforms have been extensively used in many digital signal processing applications. SVD is a useful tool of linear algebra with several applications in image compression, watermarking and other areas of signal processing. A few years ago, SVD is explored for image watermarking applications [9, 10]. The brief introduction of these three techniques are presented in this section .

The DCT transform: The discrete Cosine Transform is a technique for converting a signal into elementary frequency components [11]. The most common DCT definition of a 1-D sequence of length N is

$$C(u) = \alpha(u)\sum_{x=0}^{N-1} f(x)\cos\left[\frac{\pi(2x+1)u}{2N}\right], \quad (1)$$

For u=0,1,2,…,N-1. Similarly, the inverse transformation is defined as

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)C(u)\cos\left[\frac{\pi(2x+1)u}{2N}\right], \quad (2)$$

for x=0,1,2,….N-1. In both equations (1) and (2) $\alpha(u)$ is defined as

$$\alpha(u) = \begin{cases} \sqrt{\dfrac{1}{N}} & for\ \ u=0 \\ \sqrt{\dfrac{2}{N}} & for\ \ u\neq 0 \end{cases} \quad (3)$$

————————————————

- *N.V.Lalitha , Asst.Prof, ECE Dept., GIET, Rajahmundry,A.P. INDIA, PH-+91 99850 01587. E-mail:lalithanarla.ece@gmail.com*
- G.Suresh, Assoc.Prof, ECE Dept., Krishna's Pragati Institute of Technology, Rajahmundry, A.P.INDIA. PH-09885837385. E-mail: sureshg_ece@yahoo.co.in
- *Dr. V.Sailaja ,Prof, ECE Dept., GIET, Rajahmundry,A.P. INDIA, E-mail:lalithanarla.ece@gmail.com*

It is clear from (1) that for u=0, $C(u=0) = \sqrt{\dfrac{1}{N} \sum_{x=0}^{N-1} f(x)}$.

Thus the first transform coefficient is the average value of the sample sequence. In literature, this value is referred to as the DC Coefficient. All other transform coefficients are called the AC Coefficients [12].

In particular, a DCT is a Fourier-related transform similar to the Discrete Fourier Transform (DFT), but using only real numbers. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real and even function is real and even), where in some variants the input and/or output data are shifted by half a sample.

The DWT Transform: The discrete wavelet transform has received a tremendous amount of interest in many important signal processing applications including audio and image watermarking [13, 14 and 15]. With the DWT, the audio signal can be transformed into frequency domain ranging from low frequency to high frequency. Besides, the high frequency spectrum is less sensitive to human ear. That is the reason why the high frequency component is usually discarded in the compression process. Therefore, information to be hidden can be embedded into the low frequency component to against the compression attack.

The DWT is defined by the following equation

$$W(j,k) = \sum_{j} \sum_{k} x(k) 2^{-j/2} \Psi(2^{-j} n - k)$$

Where $\Psi(t)$ is a time function with finite energy and fast decay called the mother wavelet. The DWT analysis can be performed using a fast, pyramidal algorithm related to multirate filterbanks [16].

As a multirate filterbank the DWT can be viewed as a constant Q filterbank with octave spacing between the centers of the neighboring higher frequency subband. In the Pyramidal algorithm the signal is analyzed at different frequency bands with different resolution by decomposing the signal into a coarse approximation and detail information. The coarse approximation is then further decomposed using the same wavelet decomposition step. This is achieved by successive highpass and lowpass filtering of the time domain signal and is defined by the following equations:

$$y_{high}[k] = \sum_{n} x[n] g[2k - n]$$

$$y_{low}[k] = \sum_{n} x[n] h[2k - n]$$

Where $y_{high}[k]$, $y_{low}[k]$ are the outputs of the highpass (g) and lowpass (h) filters, respectively after sub-sampling by 2.

The SVD Transform: A few years ago, a third transform called the Singular Value Decomposition (SVD) was explored for watermarking [17]. The SVD for square matices was discovered independently by Beltrami in 1873 and Jordan in 1874, and extended to rectangular matrices by Eckart and Young in the 1930s. It was not used as a computational tool until the 1960s because of the need for sophisticated numerical techniques. In later years, Gene Golub demonstrated its usefulness and feasibility as a tool in a variety of applications [18]. SVD is one of the most useful tools of linear algebra with several applications in image compression and other signal processing fields.

3 DCT-SVD ALGORITHM

Watermark embedding procedure

Algorithm for embedding an audio in original audio using DCT-SVD:

Step 1: Sample the original audio signal at a sampling rate of particular number of samples per second. Then, partition the sampled file into frames each having certain samples.

Step 2: Performing DCT transformation on original audio signal. This operation produces a Two sub-bands: A, D. The D represents the Details sub-band, and A represents the approximation sub-band.

Step 3: Apply SVD to the DCT performed approximation sub-band A. SVD decomposes the DCT coefficients into three matrices namely, U, S, V^T. Where U is Unary matrix, S is Singular matrix.

Step 4: Perform the steps 2 and 3 to the watermark signal also.

Step 5: Embed the watermark audio bits into the DCT-SVD-transformed original audio Signal according to the formula $S_{em} = S + k * S_w$ . . . . . .(3)

Where S = singular matrix of original audio signal
S_w = singular matrix of watermark audio signal
S_{em} = singular matrix of watermarked audio signal

Step 6: Produce the final watermarked audio signal as follows:

→ Apply the inverse SVD operation using the **U** and V^T matrices, which were unchanged, and the **S** matrix, which has been modified according to Equation (3).

→ Apply the inverse DCT operation to obtain each watermarked audio frame. The overall watermarked au-

dio signal is obtained by summing all Watermarked frames.

V^T matrices, which were unchanged, and the S matrix, which has been modified according to Equation (4).

→ Apply the inverse DCT operation to obtain each watermarked audio frame.

Algorithm for extracting an audio in original audio using DCT-SVD:

Step 1: Perform steps 2 and 3 of the embedding procedure until the S matrix is obtained for all frames of the watermarked audio signal.

Step 2: Compose the singular matrix of watermark audio signal in the DCT-SVD transformed watermarked audio signal according to the formula

$$S_{ex}= (S_{em}-S)/0.01 \quad \dots \dots \dots (4)$$

Where S_{ex} = singular matrix of extracted watermark audio signal.

Step 3: Produce the final watermark audio signal as follows:

→ Apply the inverse SVD operation using the U and

4 DWT-SVD ALGORITHM

Algorithm for embedding an audio in other audio:

Step 1: Sample the original audio signal at a sampling rate of particular number of samples per second. Then, partition the sampled file into frames each having certain samples.

Step 2: Performing DWT transformation on original audio signal..This operation produces Two sub-bands: A, D.The D represents Details sub-band, and A represents the Approximation sub-band.

Step 3: Apply SVD to the DWT performed approximation sub-band A. SVD Decomposes the DWT coefficients into three matrices namely, U, S, V^T. Where U is Unary matrix, S is Singular matrix.

Step 4: Perform the steps 2 and 3 to the watermark signal also.

Step 5: Embed the watermark audio bits into the DWT-SVD-transformed original audio signal according to the formula $S_{em} = S + k * S_w$ . . . . . .(3)

Where S = singular matrix of original audio signal
        S_w = singular matrix of watermark audio signal
        S_{em} = singular matrix of watermarked audio signal

Step 6: Produce the final watermarked audio signal as follows:

→ Apply the inverse SVD operation using the **U** and **V^T** matrices, which were unchanged, and the **S** matrix, which has been modified according to Equation (3).

→ Apply the inverse DWT operation to obtain each watermarked audio frame. The overall watermarked audio signal is obtained by summing all Watermarked frames.
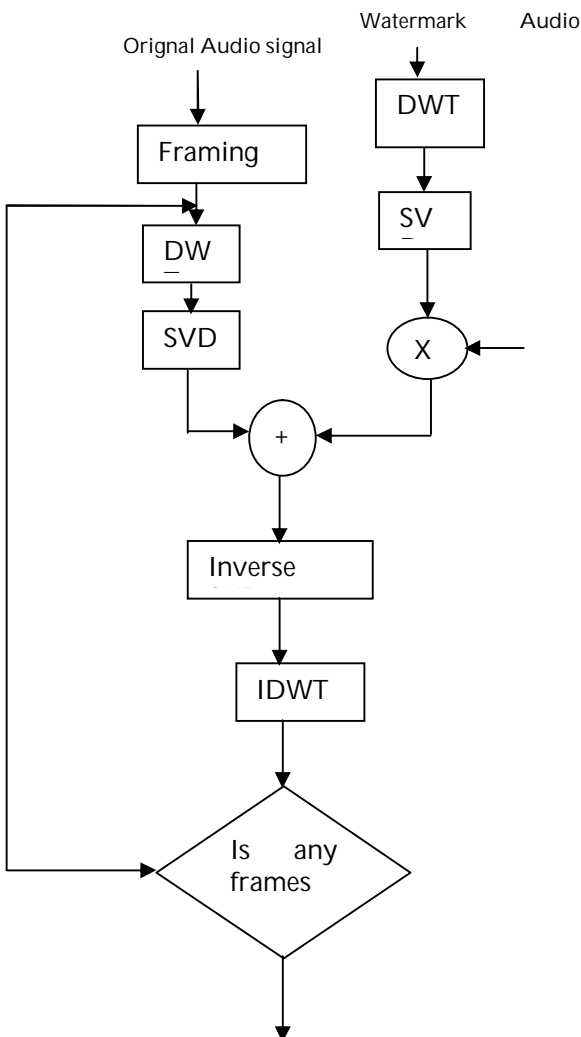
Algorithm for extracting an audio in other audio using DWT-SVD:

Step 1: Perform steps 2 and 3 of the embedding procedure until the **S** matrix is obtained for all frames of the watermarked audio signal.

Step 2: Compose the singular matrix of watermark audio signal in the DWT-SVD transformed watermarked audio signal according to the formula
$S_{ex} = (S_{em}-S)/0.01$ . . . .(4)
Where S_{ex} = singular matrix of extracted watermark audio signal.

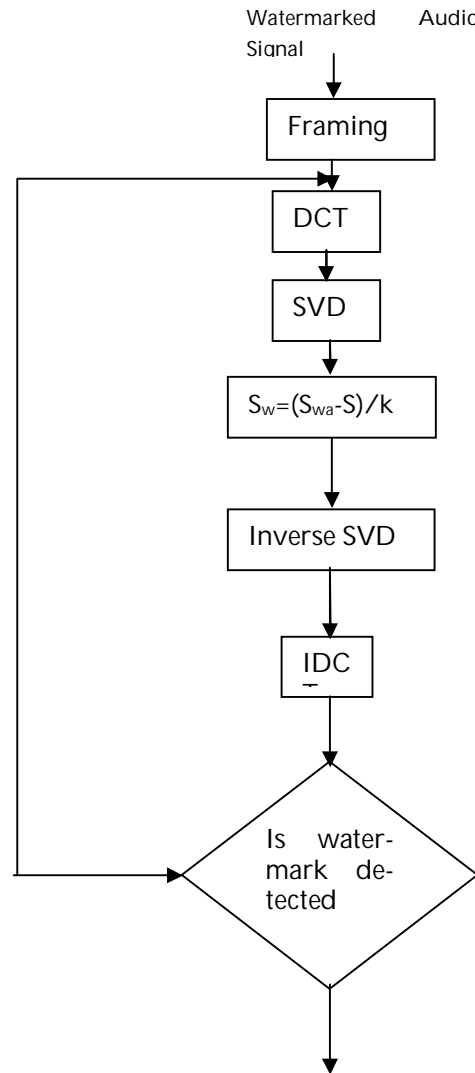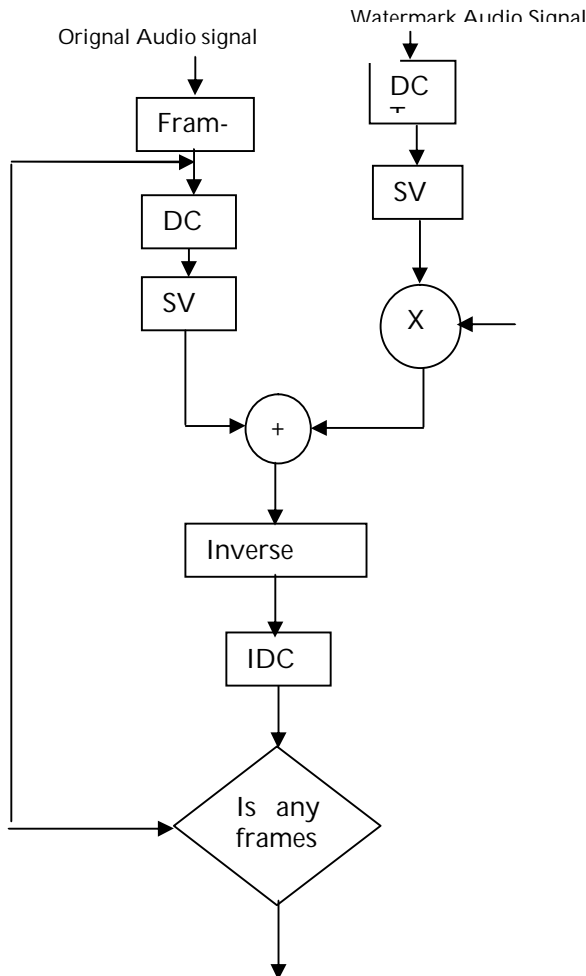**Step 3:** Produce the final watermark audio signal as follows:

→ Apply the inverse SVD operation using the **U** and **V^T** matrices, which were unchanged, and the **S** matrix, which has been modified according to Equation (4).

→ Apply the inverse DWT operation to obtain each watermarked audio frame.
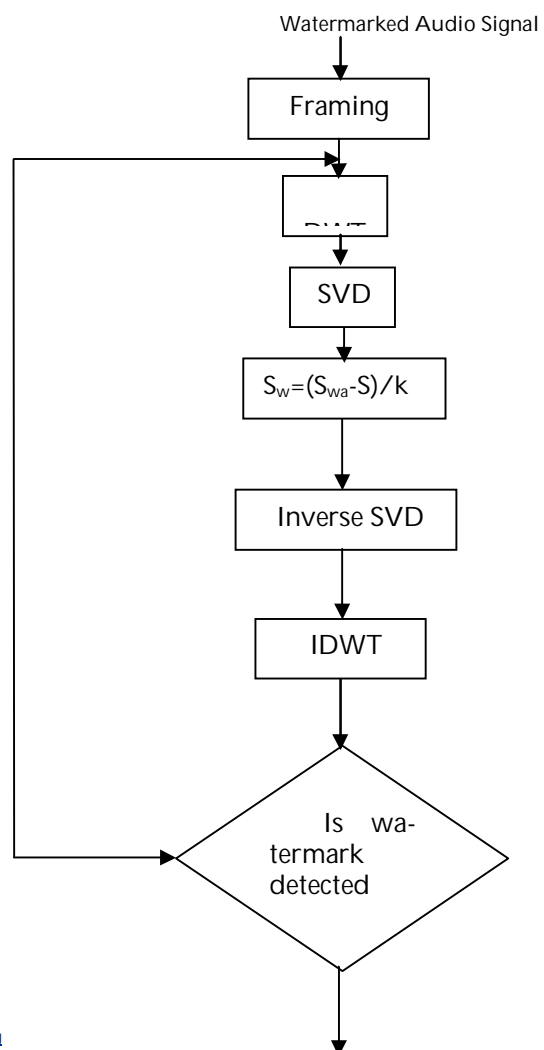
5 EXPERIMENTAL RESULTS

Four classes of audio signals like speech, pop music, rock music and instrumental were used to study the performance of the DCT-SVD and DWT-SVD algorithms. These classes were chosen because each class has different spectral properties.

Imperceptibility Test: Imperceptibility is related to the perceptual quality of the embedded watermark audio within the original audio signal. To measure imperceptibility, we use Signal-to-Noise Ration (SNR) as an objective measure, and a listening test as a subjective measure.

For subjective quality evaluation, a listening test was performed with five listeners to estimate the Mean Opinion Score (MOS) grade of the watermarked signals for four different signals. Each listener was presented with the pairs of original signal and the watermarked signal and asked to report whether any difference could be detected between the two signals. The five people listed to each pair for 15 times, and they gave a grade for the pair. The average grade for each pair from all listeners corresponds to the final grade for the pair. MOS evaluation criterion and MOS for the two techniques are listed in the Tables1 and 2 respectively.

Table 1. MOS evaluation criterion

Score	Watermark imperceptibility
5	Imperceptibility
4	Perceptibility but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

Table 2. MOS for the two techniques

Audio Signal	DCT-SVD	DWT-SVD
SPEECH	4.95	5
POP MUSIC	4.81	4.93
ROCK MUSIC	4.76	4.95
INSTRUMENT	4.54	4.9

SNR in db is calculated using the following equation [2].

$$SNR = 10\log_{10}\sum_{n=0}^{N-1}\frac{x^2(n)}{[x(n)-x^l(n)]^2}$$

Where, N is the length of audio signal
x(n) is the original signal
x^l(n) is the watermarked signal

n=0,1,2,3,……N

Signal to Noise Ratio (SNR) is a statistical difference metric which is used to measure the similitude between the undistorted original audio signal and the distorted watermarked audio signal.

Table 3. PSNR Evaluation

Class of Audio Signal	Attack Type	DCT-SVD	DWT-SVD
SPEECH	Without Attack	56.3394	**61.9459**
	Resample	**54.1057**	47.7226
	Noise	19.5943	**19.8746**
	Gaussian Noise	29.2365	**31.5413**
	Cropping	41.6865	**59.6636**
POP MUSIC	Without Attack	42.9722	**52.9715**
	Resample	18.806	**26.5770**
	Noise	**6.191**	4.664
	Gaussian Noise	14.8786	**16.2860**
	Cropping	38.6206	**39.2680**
ROCK MUSIC	Without Attack	48-1049	**50.6182**
	Resample	20.9047	**52.1388**
	Noise	**4.0980**	2.1915
	Gaussian Noise	14.2720	**18.1219**
	Cropping	15.6720	**21.6730**
INSTRU-STRU-MENT	Without Attack	54.4549	**69.5287**
	Resample	29.3643	**31.3434**
	Noise	**12.3940**	11.7206
	Gaussian Noise	20.2885	**23.7522**
	Cropping	37.5636	**42.2612**

6 CONCLUSION

An efficient audio watermarking algorithm in the frequency domain by embedding the inaudible audio water mark is presented here. It is verified that the DWT-SVD technique is robust for most of the attacks rather than the DCT-SVD. By means of combining the two transforms DWT-DCT along with SVD, inaudibility and different levels of robustness can also be achieved.

REFERENCES

[1] " Algorithms for Audio Watermarking and Stegano Graphy", NEDELJKO, CVEJIC Department of Elec Trical and Information Engineering, Information Processing Laboratory, University of OULU 2004.

[2] L.T.Bruton, J.D.Gordy, Performance Evaluation of Dig ital Audio Watermarking Algorithms, Proceedings of the 43rd IEEE Midwest Symposium, Michigan, Volume 1, pp.456-459, 8-11 August 2000.

[3] M.D.Swanson, B.Zju,A.H.Tewfik, Robust audio wa termarking using perceptual masking, Signal processing, vol.66,pp.337-355,1998.

[4] P.Bassia, I.Pitas, Robust audio watermarking in time domain, IEEE Trans. On Mutimedia, vol.3,No.2, pp.232-241, June 2001.

[5] W.Bender, D.Grul,N.Morimoto, A.Lu, Techniques for data hiding,IBM System Journal,vol.35, pp.313-336, 1996.

[6] Xin Li, Hong Heather Yu, Transparent and Robust Audio Data Hiding in Sub-band domain, Proc. Inter national Conference on Information Technology: cod ing and computing, Las Vegas, pp.74-79, March 27-29,2000.

[7] M.F.Mansour,A.W.Tewfik, Audio watermarking by time scale modification, Proc International Conference on Acoustic, Speech, and Signal processing, USA, vol.3 pp.1353-1356, May 7-11, 2001.

[8] Nedeljko Cvejic, Tapio Seppnen, Watermark Bit Rate in Diverse Signal Domains, Interna tional Journal of Signal Processing, vol.1, 2004.

[9] Bao P, Ma X (2005) Image adaptive water marking using wavelet domain singular value decomposition. IEEE Trans Circuits Syst Vid eo Technol 15(1):96-102

[10] Yavuz E, Telatar Z (2007) Improved SVD- DWT based digital image watermarking against watermark am biguity. In:Proceedings of the ACM symposium on applied computing, pp 1051-1055

[11] Rao K and P.Yip, Discrete Cosine Transform algo ithms, advantages, applications. Academic Press, USA, 1990.

[12] The Discrete Cosine Transform(DCT): Theory and Application, Syed Ali Khayam, De partment of Elec trical & Computer Engineering, Michigan State Uni versity, March 10th 2003.

[13] Michael A., Geraon and Peter G. Graven, "A high-rate butied-data channel for audio CD", Journal of the Audio Engineering Society, 43(1/2):3-22,January-February 1995.

[14] Shaoquan Wu, J.Huang, Y.Q. Shi, "Self-Synchronized Audio Watermark in DWT Domain",IEEE (ISCAS), v-712-v-715,2004.

[15] N. Sriyingyong, and K. Attakitmongcol, "Wavelet-Based Audio Watermarking Using Adaptive Tabu Search" Wireless Pervasive Computing, 2006 1st In ternational Symposium on 16-18 an. 2006 Page(s):1-5,2006.

[16] S.G Mallat "A Theory for Multiresolution Signal De

composition: The Wavelet Representation" IEEE

Transactions on Pattern Analysis and Machine Intelli

gence, Vol.11,1989,674-693.

[17] R. Liu and T.Tan, "A SVD-Based watermarking

Scheme for Protecting Rightful Ownership", IEEE

Transactions on Multimedia, 4(1),March 2002,pp.121-

128.

[18] D.Kahaner, C. Moler and S.Nash, Numerical Methods

and Software (New Jersey: Prentice-Hall, Inc, 1989).

[19] ITU-R, "Methods for objective measurements of per

ceived audio quality", BS.1387.

Handoff Analysis for UMTS Environment

Pankaj Rakheja, Amanpreet Kaur, Dilpreet Kaur

Abstract— UMTS is in one of the third generation mobile telecommunication technologies. And it supports various multimedia applications and services at an enhanced data rate with better security. It also supports mobile users and for that there is a process called handover where new channels are assigned to the user when it moves from a region covered by one node to a region covered by other. In this paper we analyse the effect of handover over the performance of the system.

Index Terms— DPCH, Handover, UTRA

——————————————— ◆ ———————————————

1 INTRODUCTION

Universal Mobile Telecommunications Service is a third-generation broadband which supports packet-based transmission of text, digitized voice and video, and the multimedia here can reach data rates up to 2 megabits per second (Mbps). It also offers a consistent set of services to mobile computer and phone users, no matter where they are located in the world. It is based on the Global System for Mobile (GSM) communication standard that is it is overlaid on GSM. It is also endorsed by major standards bodies and manufacturers as the planned standard for mobile users around the world. It can ensure a better grade of service and Quality of service on roaming to both mobile and computer users. Users will have access through a combination of terrestrial wireless and satellite transmissions.

Cellular telephone systems used previously were mainly circuit-switched, meaning connections were always dependent on availability of circuits. A packet-switched connection uses the Internet Protocol (IP), which uses concept of virtual circuits that is a virtual connection is always available to connect an endpoint to the other end point in the network. UMTS has made it possible to provide new services like alternative billing methods or calling plans. For instance, users can now choose to pay-per-bit, pay-per-session, flat rate, or asymmetric bandwidth options. The higher bandwidth of UMTS also enabled other new services like video conferencing. It may allow the Virtual Home Environment to fully develop, where a roaming user can have the same services to either at home, in the office or in the field through a combination of transparent terrestrial and satellite connections.

2 OVERVIEW

The term handover is also known as handoff. Whenever a user terminal moves into area covered by a different RNC while the conversation is still going on, then new channels are allocated to the user terminal which is now under

different control node or MSC. This is carried out to ensure continuity of communication and to avoid call dropping. For this to take place the system need to identify the user terminal and monitor its signal strength and setting of a threshold value below which a call or applications drops and enabling new channel allocation before this level.

There is Handoff margin which needs to be optimized for proper synchronization .It is the difference between signal strength at which handover should occur and the minimum required signal strength .If it is too low then there will be insufficient time to complete the process and if it is too large then unnecessary handovers will occur. The most important thing is the handovers are not visible to the users.

2.1 Handover types

Handovers can be broadly classified into two types namely: Intracellular and Intercellular Handover. In the Intracellular handover, mobile or user terminal moves from one cellular system to another. And in the Intercellular handover, user terminal moves from one cell to other. This is further classified into soft and Hard Handover.
Soft handover

Here we follow make before break concept where the user terminal is allocated new channels first then previous channels are withdrawn. Here chances of losing continuity are very less. But it needs user terminal or mobile to be capable of toning to two different frequencies. Here complexity at user end increases a lot. It is quite reliable technique but here channel capacity reduces.

Hard Handover

Here we follow break before make concept where from the user terminal previously allocated channels are first withdrawn then new are allocated. Here chances of call termination are more than in soft handover. Here at user

terminal complexity is less as it need not be capable of toning to two different frequencies. It provides advantage over Soft handover in Terms of channel capacity but it is not as reliable as soft handover

2.2 Prioritizing handoffs

Handoff requests are more important than new call requests or application requests as call dropping in between will be more annoying for the user then not being able to make a new call. So a guard channel is especially reserved for the handoffs. We also queue the requests made for proper flow and order control.
The most obvious cause for performing a handover is that due to its movement a user can be served in the another cell more efficiently (like less power emission, less interference etc). It may however also be performed for other reasons which may be system load control

2.3 Classification of cells

Active Set: Is defined as the set of Node-Bs the UE is simultaneously connected to (i.e., the UTRA cells currently assigning a downlink DPCH to the UE constitute the active set).

Monitored set: is defined as the set of nodes not in the active set but are included in CELL_INFO_LIST

Detected set is defined as the set of nodes neither in the active set nor in CELL_INFO_LIST but are detected by UT

2.4 Special considerations in UMTS environment

In UMTS environment the different types of air interface measurements are:

Intra-frequency measurements: those measurements which are carried out on downlink physical channels at the same frequency as that of the active set. And the measurement object here corresponds to one cell.

Inter-frequency measurements: those measurements which are carried out on downlink physical channels at frequencies that differ from the frequency of the active set. And the measurement object here corresponds to one cell.

Inter-RAT measurements: those measurements which carried out on downlink physical channels belonging to another radio access technology than UTRAN, e.g. GSM. And the measurement object here corresponds to one cell.

Traffic volume measurements: those measurements which are carried out on uplink channels to analyse the volume of traffic on them. And the measurement object here corresponds to one cell.

Quality measurements: these Measurements are carried out on downlink channels to obtain the various quality parameters, e.g. downlink transport block error rate. And the measurement object here corresponds to one transport channel in case of BLER. A measurement object corresponds to one timeslot in case of SIR (TDD only).

UE-internal measurements: Measurements of UE transmission power and UE received signal level.

UE positioning measurements: Measurements of UE position.

The UE supports a number of measurements running in parallel. The UE also supports that each measurement is controlled and reported independently of every other measurement.

3 WORK DONE

We have designed three scenarios where handovers occur when the user terminal moves from area of one node to other in order to enable communication between source and destination .we made analysis for effect of speed and number of handovers over throughput, average jitter, average end to end delay etc. Here in scenario 1 , shown in figure 1, terrain is 1500 sqms and there are two nodes and UEs moves from area of one node to that covered by the other while application is still in active state so one handover has occurred here, in scenario 2, shown in figure 2, terrain is 2500 sqms and there are three nodes and UEs moves from area covered by one to second than to third one while application is still in active state so two handover occurred and in the scenario 3, shown in figure 3, terrain is 3500 sqms and there are four nodes and UEs moves from area covered by one to second than to third to fourth one while application is still in active state so three handover have occurred. Here we have taken two users one travelling at 16 m/s and other at 20 m/s respectively. The later one is called fast UT (User terminal) while the former one is referred as Slow UT (User terminal)

The screenshots of three scenarios designed to analyse impact of handover on the overall performance of the system are:
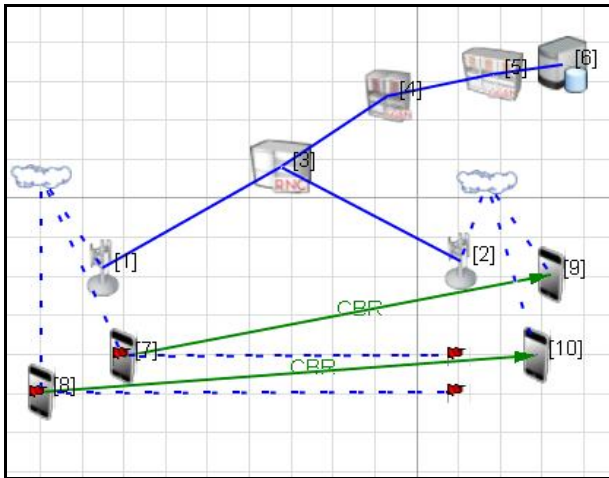
Figure 1: Screenshot of scenario for one Handover
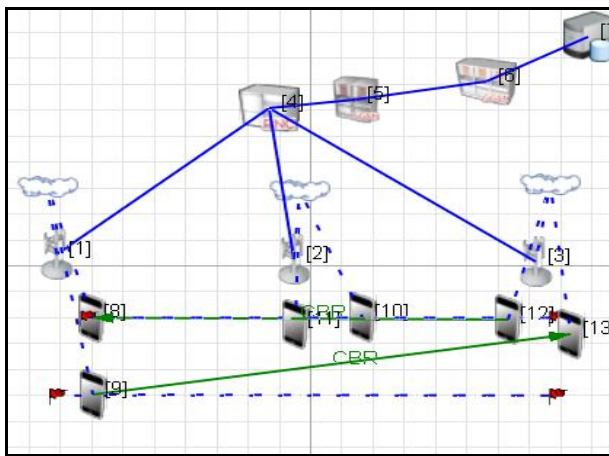
Figure 2: Screenshot of scenario for two Handovers
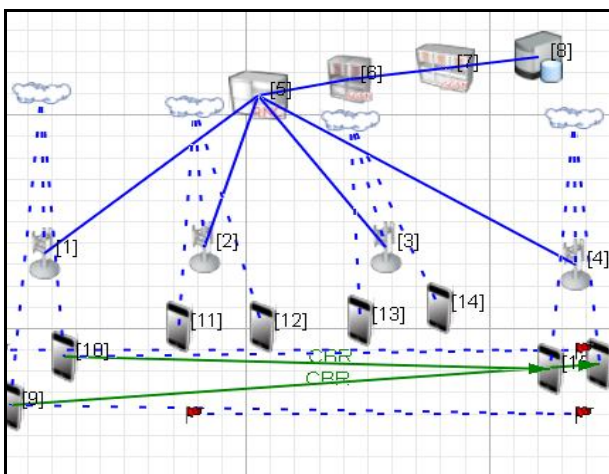
Figure 3: Screenshot of scenario for three Handovers

The values obtained by running the simulations for throughput, Average jitter and End to end delay for various scenarios shown in figures (1-3) are:

Table 1: Throughput values for handovers

Throughput	Fast UT	Slow UT
One hand-over	4171	4188
Two hand-overs	4169	4189
Three hand-overs	4170	4188

Table 2: Average jitter values for Handover

Throughput	Fast UT	Slow UT
One hand-over	0.233	0.213
Two hand-overs	0.0556	0.009
Three hand-overs	0.17	0.038

Table   3: `End to end delay values for Handovers

Throughput	Fast UT	Slow UT
One hand-over	0.67517	0.6755
Two hand-overs	0.45	0.39
Three hand-overs	0.52	0.44

The plots drawn using these values obtained are:
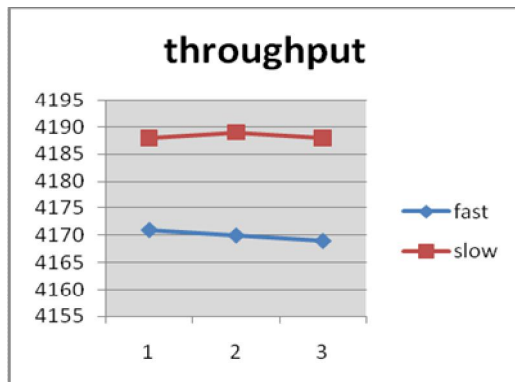
Figure 4: Graph of Throughput for Handover
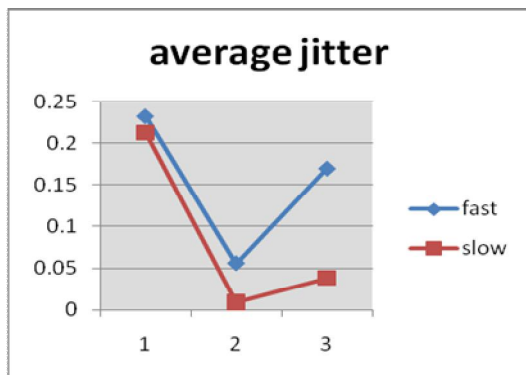
Figure 5: Graph of Average jitter for Handovers
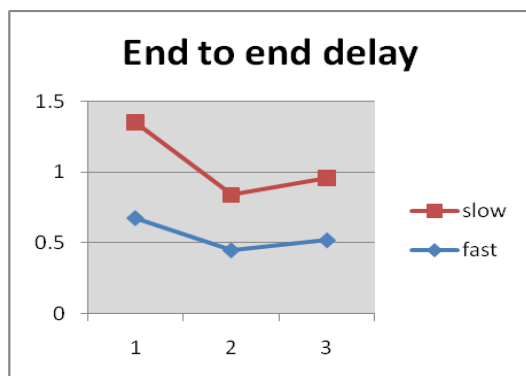
Figure 6: Graph of Average jitter for Handovers

4 CONCLUSION

On analysing the results obtained after running different simulations. We can say that the system has better performance for slow speed users as values of performance determining parameters like Throughput, Average jitter and End to end delay better than that for high speed us-

ers. With number of handovers throughput does not vary much, And Average jitter and End to end delay first fall down a bit latter increases. So, overall performance of system is good in case of handover as jitter and End to end delay are not much and Throughput is also good does not vary much.

5 REFERENCES

[1] Falciasecca, G. Frullone, M. Grazioso, P. Riva, G. Serra, A.M.," Performance evaluation of a UMTS in a urban environment", 1991, Sixth International Conference on Mobile Radio and Personal Communications

[2] van Nielen, M.J.J. ,R. PTT Nederland NV, Leidschendam ," UMTS: a third generation mobile system", Third IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 1992. Proceedings, PIMRC '92

[3] Oudelaar, j, " Evolution towards UMTS", 5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 1994

[4] Fan, L. Sheriff, R.E. Gardiner, J.G," Satellite-UMTS service provision using IP-based technology", 2000 IEEE 51st Vehicular Technology Conference Proceedings.

[5] Jin Yang Kriaras, "Migration to all-IP based UMTS networks", First International Conference on 3G Mobile Communication Technologies, 2000

[6] Hyeyeon Kwon Kyung-yul Cheon Aesoon Park," Analysis of WLAN to UMTS Handover", 2007 IEEE 66th Vehicular Technology Conference, 2007

A High Performance and Low Power Hardware Architecture for H.264 Transform Coding

Jubli Kashyap,Virendra Kumar Yadav

Abstract— In the search for ever better and faster video   compression standards H.264 was created. H.264 promises to be an excellent video format for use with a large range of applications and need for hardware acceleration of its very computationally intensive parts. To address this need, this paper proposes architecture for the discrete transform (DCT) and quantization blocks from H.264. The first set of architectures for the DCT and quantization were optimized for power, which resulted in transform and quantizer blocks that use 10.5623 mW Power. All of the designs were synthesized for Cadence BuildGate Synthesis CMOS technology, as well as the combined DCT and Quantization blocks went through comprehensive place and route flow.

Index Terms—CMOS Technology, DCT, H.264, JVT, ITU-T, SoC, Quantization, YUV System, Zero Shift.

—————————————— ◆ ——————————————

1 INTRODUCTION

Due to the remarkable progress in the development of products and services offering full-motion digital video, digital video coding currently has a significant economic impact on the computer, telecommunications, and imaging industry . This raises the need for an industry standard for compressed video representation with extremely increased coding efficiency and enhanced robustness to network environments. Since the early phases of the technology, international video coding standards have been the engines behind the commercial success of digital video compression. ITU-T H.264/MPEG-4 (Part 10) Advanced Video Coding (commonly referred as H.264/AVC) is the newest entry in the series of international video coding standards. It was

developed by the Joint Video Team (JVT), which was formed to represent the cooperation between the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) [3]-[5]. Compared to the currently existing standards, H.264 has many new features that makes it the most powerful and state-of-the-art standard . Network friendliness and good video quality at high and low bit rates are two important features that distinguish H.264 from other standards. The usual 8x8 DCT is the basic transformation in H.264. This eliminates any mismatch issues between the encoder and the decoder.

The demand of multimedia communications on mobile and portable applications is growing nowadays. To realize multimedia communications, implementing a video compression standard is essential in any multimedia processing system-on-a-chip (SoC). There have been reports on the very large-scale integration (VLSI) implementation of MPEG-4 video recently. The emerging efficient H.264 or MPEG-4 Part 10 standard can greatly reduce the bandwidth and storage requirements for multimedia data. The VLSI implementation of H.264 is a challenge since an

H.264 baseline decoder is approximately three times more complex than an H.263 baseline decode. Implementational flexibility is an important factor of concern for SoC designs. Since the traditional hardwired design is less flexible, the processor-based implementation is a preferred choice. VLSI implementation can be categorized into three types, hardwired, digital-signal-processor-based, and hybrid. To achieve higher performance with flexibility, the hybrid architecture has been proposed.

2 DESIGN REQUIREMENTS OF THE H.264 TRANSFORM
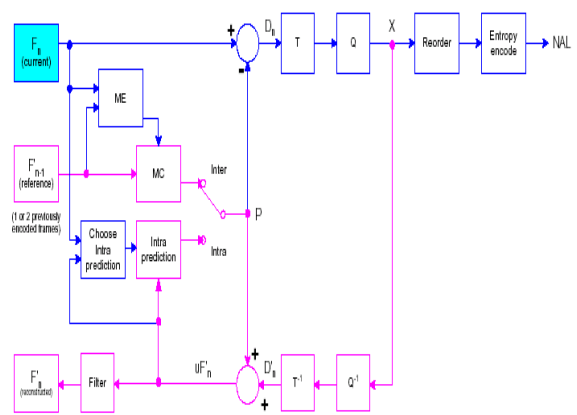
Block Diagram of H.264 Encoder

2.1 Discrete Cosine Transform

The DCT is conceptually similar to the DFT, except:

$$[f]_{NxN} = [T^T]_{NxN}[F]_{NxN}[T]_{NxN}$$

$$[F]_{NxN} = [T]_{NxN}[f]_{NxN}[T^T]_{NxN}$$

The DCT does a better job of concentrating energy into lower order coefficients than does the DFT for image data. The DCT is purely real, the DFT is complex (magnitude and phase).The DCT is purely real, the DFT is complex (magnitude and phase).

A DCT operation on a block of pixels produces coefficients that are similar to the frequency domain coefficients produced by a DFT operation. An N-point DCT has the same frequency resolution as and is closely related to a 2N-point DFT. The N frequencies of a 2N point DFT correspond to N points on the upper half of the unit circle in the complex frequency plane.

Assuming a periodic input, the magnitude of the DFT coefficients is spatially invariant (phase of the input does not matter). This is not true for the DCT. For most images, after transformation the majority of signal energy is carried by just a few of the low order DCT coefficients. These coefficients can be more finely quantized than the higher order coefficients. Many higher order coefficients may be quantized to 0 (this allows for very efficient run-level coding).

$$F(u,v) = \frac{2}{N}C(u)C(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

$$f(i,j) = \frac{2}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} C(u)C(v)F(u,v)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

Which can be written in matrix form, where the rows of [T] are the DCT basis vectors, as:

$$[f]_{NxN} = [T^T]_{NxN}[F]_{NxN}[T]_{NxN}$$

$$[F]_{NxN} = [T]_{NxN}[f]_{NxN}[T^T]_{NxN}$$

In the formulas, F(u,v) is the two-dimensional NxN DCT. u, v, x, y = 0,1,2,...N-1 where x, y are spatial coordinates in the sample domain and u, v are frequency coordinates in the transform domain.

C(u), C(v) = 1/(square root (2)) for u, v = 0.
C(u), C(v) = 1 otherwise.

MPEG specifies the spatial samples to be represented in 9 bits and the coefficients to be represented in 12 bits. The dynamic range of the coefficients is specified as [-2048:+2047].

————————————————

- *Jubli Kashyap is currently working as Assistant Professor(EC Deptt.) in NIEC, Guru Gobind Singh Inderprasth University, Delhi,India, Mobile No.- 09213356485. E-mail: jubli_k@rediffmail.com*
- *Virendra Kumar Yadav is currently working as TechnicalServices professional, IBM, India Mobile No.- 09212428516, E-mail: virendra7n@rediffmail.com.*

2.3 Color Transformation

The human visual systems are most sensitive to changes in luminance and less to changes in chrominance.

RGB must be converted to the other color systems.

YUV System

Y = 0.299R + 0.587G + 0.114B ;
U = B - Y ;
V = R - Y ;

Zero Shift

After transformation Y, U and V are in the range of [0,255] Zero shift changes this range to [-128,127]

DCT The strength of transform coding in achieving data compression is that the image energy of the most natural scenes is mainly concentrated in the low frequency region and hence into a few transform coefficients.

2.4 Quantization

Quantization allows us to reduce the accuracy with witch the DCT coefficients are represented when converting the DCT to an integer representation.

It tends to make many coefficients zero, specially those for high spatial frequencies.

Two standard tables are available for quantization.

2.5 Zigzag Scan

The zigzag pattern used in the JPEG algorithms orders the basis functions from low to high spatial frequencies .
It facilitate entropy coding by encountering the most likely non-zero coefficient first.

2.6 Picture Format

The picture is divided into number of Macro Blocks (MB)
In 4:4:4 sampling each MB contains four Y-blocks, four R-Y blocks and four B-Y blocks.
In 4:2:2 sampling each MB contains four Y-blocks, two R-Y blocks and two B-Y blocks.
* Color difference signal is sub-sampled in Horizontal direction
* The bit rate=720X576X25X8+360X576X25X (8+8) = 166Mb/s

In 4:2:0 sampling each MB contains four Y-blocks, one R-Y block and one B-Y block

 * Color difference signal is sub-samples in both horizontal and vertical direction

 * Bit rate=720X576X25X8+360X288X25X(8+8)= 124Mb/s.

Motion vectors are associated with MB's(Microblocks)

[2] Joint Video Team (JVT) of ITU-T VCEG and ISO/IEC MPEG, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification,   ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC, May 2003.

[3] ITU-T Rec. H.264 / ISO/IEC 11496-10, "Advanced

Video Coding", Final Committee Draft, Document JVTE022, September 2002

[4] INTRODUCTION TO DATA COMPRESSION 3RD EDITION BY KHALID SAYOOD.

[5] H.264 AND MPEG-4 VIDEO COMPRESSION BY IAIN E G

Richardson.Published by John Wiley & Sons, September 2003.

[6] Multimedia Communications by Fred Halsal, Pearson Educaton,2004.

3 IMPLEMENTATION RESULTS

The proposed architecture is implemented in Verilog HDL. The implementation is verified with RTL simulations using Cadence ncLaunch. The Verilog RTL is then synthesized to a Cadence BuildGate Synthesis tool. The resulting netlist is placed and routed to the Cadence SoC Encounter. The first set of architectures for the DCT and quantization were optimized for power, which resulted in transform and quantizer blocks that use 10.5623 mW Power.

4 REFERENCES

[1] I. Richardson, H.264 and MPEG-4 Video Compression,Wiley, 2003

Computer Simulation of Forging Using the Slab Method Analysis

S. B. Mehta, D. B. Gohil

Abstract— Forging is a very complex process and the measurement of actual forces for real material is difficult and cumbersome. The main objective of this document is to use the analytical methods for measuring parameters such as load, and stress distribution of forging process and use them to make simple, reliable, fast and non-expensive simulation tools, contrary to the commercial software's which require much means, time and a perfect knowledge of the process. Of the various methods used for analysing forging operations, the most often used SLAB method techniques are described here.

Index terms—analysis, simulation, modelling, forging, closed-die forging, open die forging, process.

———————————————— ◆ ————————————————

1 INTRODUCTION

THE developments in forging technology have increased the range of shapes, sizes and properties of the forged products enabling them to have various design and performance requirements. Also Closed-Die forging is an extremely complex forming process from the point of view of deformation mechanics. The non-steady state and non-uniform material flow, the considerable interface friction, and the heat transfer between the deforming material and the tooling are very difficult to analyze. To ensure the quality of the final product, even a highly experienced engineer spends a lot of time on optimizing the design of the process through a time-consuming trial and error method. Sometimes, the design engineers take the help of FE analysis to fine-tune the process and avoid the costly physical trials. In this context, it is worthwhile to mention that the FE analysis of metal forming is a time-consuming process, even on a powerful PC. Interpretation of the results of FE analysis requires in-depth knowledge and experience of both forging process and FE method. Other than the forging process parameters, the results of FE analysis depend on proper selection of a large number of FE parameters, e.g., element type and size, mesh topology, node numbering, and others. A proper selection of these FE parameters requires in-depth knowledge [1]. A feasible solution is the use of analytical methods, of which the most widely used, 'Sachs' or 'Slab' method, is used. The slab method here is used to divide the work piece into various slabs. The parameters such as load and stress are then easily calculated for these shapes and then added to get the value of final forging load.

- *S.B.Mehta is currently pursuing masters degree program in , Mechanical Engineering Department, S. V. National Institute of Technology*
  Surat, India. Sumeet.b.m@gmail.com
- *D.B.Gohil is currently working as anAssociate Professor, Mechanical Engineering Department, S. V. National Institute of Technology Surat, India, , dbgsvnit@gmail.com*

2 THE SLAB METHOD

The work piece being deformed is decomposed into several slabs. For each slab, simplifying assumptions are made mainly with respect to stress distributions. The resulting approximate equilibrium equations are solved with imposition of stress compatibility between slabs and boundary tractions. The final result is a reasonable load prediction with an approximate stress distribution [Kobayashi et al .1989]. The following assumptions are made in using the slab method of analysis [2]:

- the deforming material is isotropic and incompressible,
- the elastic deformations of the deforming material and tool are neglected,
- the inertial forces are small and are neglected,
- the frictional shear stress, τ, is constant at the, die/material interface,
- the material flows according to the von Mises rule,
- the flow stress and the temperature are constant within the analyzed portion of the deforming material.

2.1 Open die forging

1) Plane strain

In applying slab analysis to plane strain upsetting, a slab of infinitesimal thickness is selected perpendicular to the direction of metal flow (Fig. 1). Assuming a depth of "1" or unit length, a force balance is made on this slab. Thus, a simple equation of static equilibrium is obtained [Thomsen et al., 1965] [Hoffman et al., 1953].

Summation of forces in the X direction is zero or

$$\sum F_x = \sigma_x h - (\sigma_x + d\sigma_x)h - 2\tau dx$$

Or

$$d\sigma_x = -\frac{2\tau dx}{h}$$

Thus, by integration one gets:
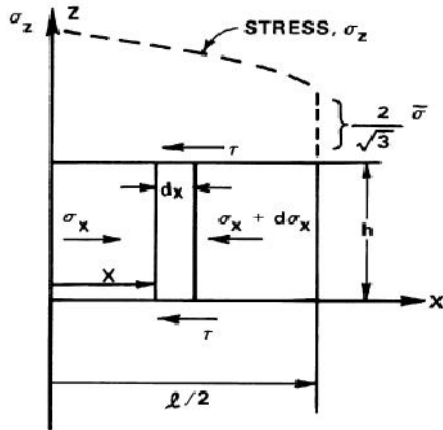
Fig 1. Equilibrium of forces in plane strain homogenous upsetting

$$\sigma_x = -\frac{2\tau}{h} x + C$$

From the flow rule of plane strain, it follows that:

$$\sigma_x = -\frac{2\tau}{h} x + C + \left| \frac{2}{\sqrt{3}} \sigma \right|$$

The constant C is determined from the boundary condition at x= l/2, where $\sigma_x = 0$, and from equation

$$\sigma_x = \left| \frac{2}{\sqrt{3}} \sigma \right|$$

Thus,

$$\sigma_x = -\frac{2\tau}{h}\left(\frac{l}{2} - x\right) - \frac{2}{\sqrt{3}}\sigma$$

Equation illustrates that the vertical stress linearly increases from the edge(x=l/2) of the figure towards the centre (x=0) In equation, the frictional shear stress, τ is equal to $m\bar{\sigma}/\sqrt{3}$. Thus, integration of the equation over the entire width l of the strip of unit depth gives the upsetting load per unit depth [2]:

$$L = \frac{2\sigma}{\sqrt{3}}\left(1 + \frac{ml}{4h}\right)l$$

2) *Axisymetric*

The flow rule for Axisymetric deformation is obtained by using a derivation similar to that used in plane strain deformation. The equilibrium of forces in the r direction gives [Thomsen et al., 1986][Hoffman et all., 1953] :

$$\sum F_r = \sigma_r(d\theta)rh - (\sigma_r + d\sigma_r)(r + dr)hd\theta + 2\sigma_\theta \sin\frac{d\theta}{2}hdr - 2\tau r d\theta dr$$

Further simplification using appropriate boundary conditions gives the result as

$$\sigma_z = \frac{2\tau}{h}(r - R) - \sigma$$

The equation says that stress increases linearly from the edge towards the centre. The upsetting load can now be found out as,

$$L = \sigma\pi R^2\left(1 + \frac{2mR}{2h\sqrt{3}}\right)$$
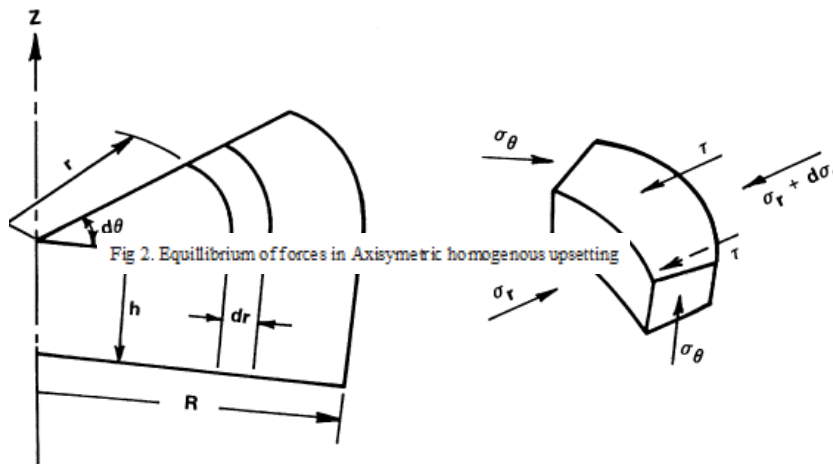
Fig 2. Equilibrium of forces in Axisymetric homogenous upsetting

2.2 Closed die forging

Fig. 1 shows axi-symmetric upsetting under inclined planes, where the flow is diverging. Writing the equilibrium equations in the *r* and *z* directions for the element shown in Fig. 3, using von mises rule, doing some simplifications, integrating the obtained differential equation and applying the boundary conditions, one can derive the following relationship for the distribution of σ_z

$$\sigma_z = \frac{K_2}{K_1} \ln\left(\frac{h_e}{K_3 - rK_1}\right) + \sigma_{ze}$$

Where,

$$K_1 = tan\alpha + tan\beta, \qquad K_2 = \sigma_0 K_1 + \tau\left(\frac{1}{cos^2\alpha} + \frac{1}{cos^2\beta}\right),$$

$$K_3 = h_b - r_b K_1$$

In the above equations τ is the frictional shear stress at the interface $(f\sigma_0)$, σ_0 the flow stress of the material , f the shear friction factor whose value is between 0 and 0.577, σ_{ze} the axial stress (σ_z) at r = r_e. All the geometrical parameters are shown in Fig. 3. It should be noted that the angles α and β in fig. 1 are positive. When the lower and upper surfaces are parallel, then, $K_1 = 0$.

Integrating the above equation for deformation zone, the required forming load (*L*) can be obtained as follows [3]

$$\frac{L}{2\pi} = \frac{K_2 h_e}{K_1{}^3} \{h_e[0.25 + 0.5S_b{}^2(\ln S_b - 0.5)]$$

$$- K_3[1 + S_b(lnS_b - 1)]\} + \frac{\sigma_{ze}}{2}(r_e{}^2 - r_b{}^2)$$

Where,

$$S_b = h_b/h_e$$
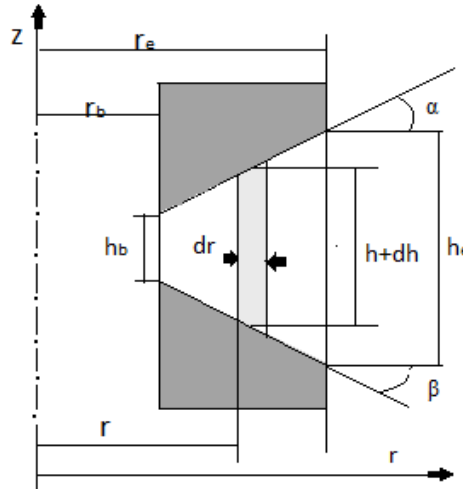
Fig. 3 Axi-symmetric upsetting between two inclined surfaces

1) Material flow in closed die forging:

When a deep die cavity is filled in closed-die forging, further movement of dies will reduce the extension of the plastic region. A part of the work material in the cavity becomes rigid, this part of the material forming the so-called 'dead-metal region' or 'dead zone'. In this stage, the plastic deformation is localized in a small part of the workpiece. Very often the deformation zone is localized at the centre of the workpiece. so that the material in the segregation zone might flow into the flash of the forgings. Fig. 2 here shows the deformation zone for a simple forging having two different values of flash [4]. The deformation zone could be of any shape based on the geometry of the die, but for calculation purposes, it is approximated into a closest fitting polygon, which is then used for finding the forging load.
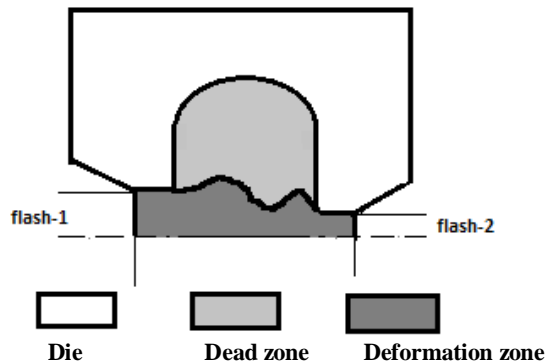
Fig. 4 Schematic diagram of the dead metal region and the deformation zone

As shown in Fig 3, the actual lens shaped deformation region is approximated into a quadrilateral shape (as shown by dotted region).

2.3 Determination of flow model

To estimate the final forming load in closed die forging operation, the flow of metal at the end stage of the process should be known. Previous researchers showed that the real and approximate shear surfaces are like those illustrated in Fig. 3. h and α could be calculated based on the following equations

$$\tan\alpha = +\sqrt{1 - \frac{(h/t) - 1}{(h/t)\ln(h/t)}}$$

$$\frac{h}{t} = 0.8\left(\frac{L}{t}\right)^{0.92}$$

When the forging section is more complex and some parts of the dies are closer to the parting line than $h/2$, the geometry of the flow model follows the die surface. Each opposite pair of lines of flow model, make up a deformation zone shown in Fig 4.Once the profile of the dead zone .i.e the real shear surface is found out, another approximate shear surface can be found out as shown in Fig. 4. In this forming the actual lens shaped deformation zone is approximated into a quadrilateral shape(as shown by dotted region). The approximate shear surface is then divided into further simpler shapes as shown in fig 5. The slab method is then individually applied to these regions for getting the load and stress value.
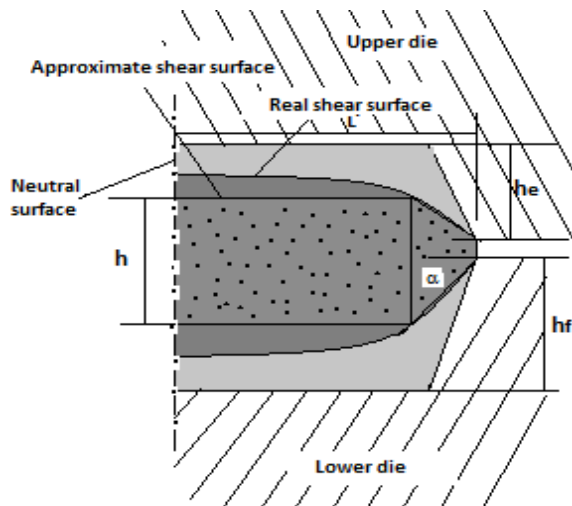
Fig 6 A plane strain model in steep sided die impression [5]
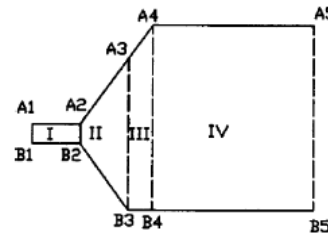
Fig. 7 The division of flow metal into standard deformation zones

3 THE DEVELOPED CODE

So far what we have seen is that closed die forging process is a much more complex process then open die. There are a number of parameters involved and finding the exact flow model could be tricky and time consuming for complex shapes. Hence a simple algorithm can be formed with its input being taken from an interactive environment of a computer simulation package. The user should firstly draw the profiles of the product in the GUI as shown in fig 5. With this regard, in closed die forging process, in the flash gutter where there is no resistance to metal flow should not be drawn. Once that is done, the geometry of the product is exported to the workspace. The algorithm does the simple calculation if it is open die forging. For closed die forging stepwise method for finding the forging parameters has been shown here

After calculations of the flash width, the flash thickness and the maximum radius of the round forging, the values of h and α are determined by the program to achieve the initial geometry of the flow model. Since each deformation zone is located between two inclined lines, each arc of the forging section is then changed into appropriate number of line segments, based on the radius and length of arc. Therefore the upper and lower parts of the final flow model are completely determined and again in each part, the drawing entities are sorted form the flash gap to the centre of the component.

In the next stage based on the technique shown in figure 7, the program divides the final flow model into standard deformation zones. The program for each zone specifies four points making up the corners of the deformation zone. These points together with the flow stress and friction factor, which are specified by the user, are employed by the program for the determination of the axial stress distribution and the forming load.
Next the stress distribution and the forging load can be calculated for whole component. At the final stage of the program, the distribution of the axial stress together with the flow model is drawn in ACAD. The total forming load is also reported to the user [3]
Similarly a code can be setup for open die forging process giving outputs.
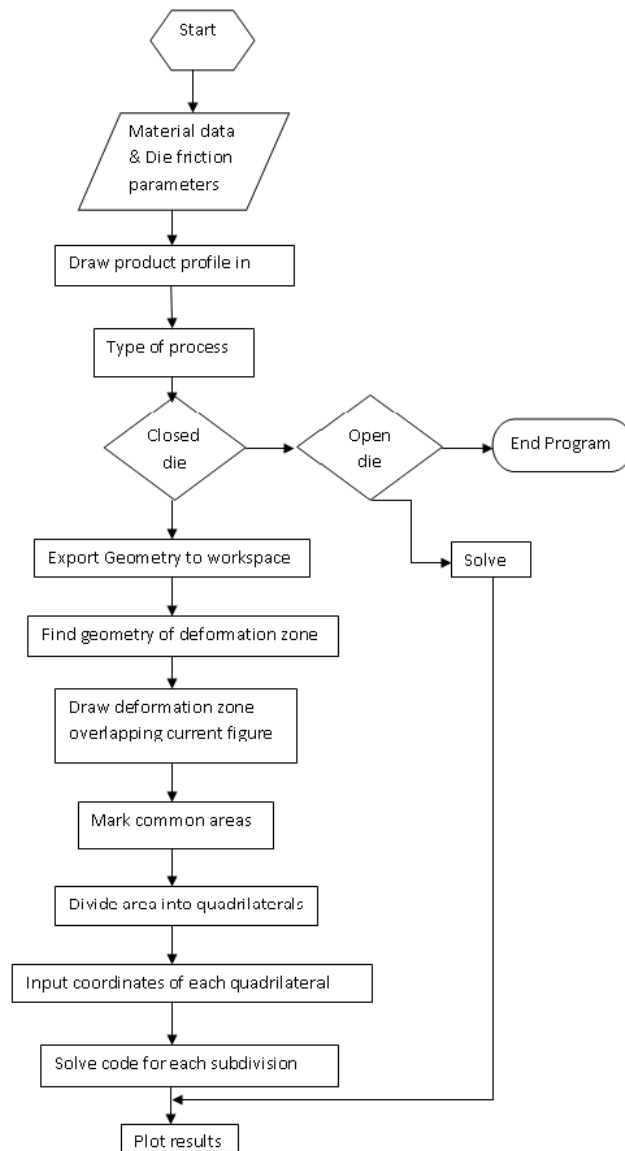
Fig 8. Flow chart of steps for forging analysis

4 CONCLUSIONS

Based on the research work presented in the previous sections, the conclusions are as follows:

1. The developed code is useful in the determination of the forming load and stress distribution for simple parts and only requires inexpensive hardware.
2. The use of the developed slab method code is very beneficial for determining the capacity of the required forging machine

5 APPENDIXES

τ = frictional shear stress

h = final height of strip

σ_0 = flow stress

σ_{ze} = axial stress

$\dot{\varepsilon}_x$ = strain rate

f = shear friction factor

v_t = velocity of the dies

x, y Cartesian coordinates

x_{np} = x position of the neutral plane

$\acute{h}_x$ = first derivative of h_x

$\dot{W}_{tot}$ = total work done

$\dot{W}_P$ =, power dissipated due to plastic deformation

$\dot{W}_s$ = power dissipated due to shear along velocity discontinuity

ACKNOWLEDGMENT

REFERENCES

[1] T. Gangopadhyay, D. Pratihar, I, Basak, Expert system to predict forging load and axial stress, Applies soft computing, Jan 6 2010

[2] T. Altan, G. Ngaile, G. Shen, Cold and hot forging: fundamentals and applications, ASME International, 2005

[3] F. Fereshteh-Saniee, M. Jaafari, Analytical, numerical and experimental analysis of the closed die forging, Journal of material processing technology, 125-126(2002) 334-340

[4])J. Hou, A plane strain UBET analysis of material flow in a filled deep die cavity in closed die forging, Journals of material processing technology,70(1997) 103-110

[5] T. Altan, J. Fiorentino , Prediction of loads and stress in closed die forging, Transactions of the ASME , may 1971

[6] J. Hou, H. Keib, U. Stahlberg, Determination of boundary between dead metal region and deformation zone by the upper bound method, Internal report, Division of material forming, Royal institute of technology 1993

[7] T. Altan, G. Shen, G. Ngaile, Cold and hot forging: fundamentals and applications, ASME International 2005

[8] Kobayashi, S., Oh, S.I., Altan, T., Metal Forming and the Finite Element Method, Oxford University Press, 1989.

[9] Nefissi N, Bouaziz Z, Zghal A, Prediction and simulation of Axisymetric forging load of aluminium , APEM journal, 2008

[10] B Tomov, Hot closed die forging-State-of-art and future development, AMME Journal, September 2007

[11] T Suzuki, Recent developments of forging in japan, International journal of machine tools manufacturing, vol 29, No. 1,pp 5-27 1989

Speech Recognition By Using Recurrent Neural Networks

Dr.R.L.K.Venkateswarlu, Dr. R. Vasantha Kumari, G.Vani JayaSri

Abstract -Automatic speech recognition by computers is a process where speech signals are automatically converted into the corresponding sequence of characters in text. In real life applications, however, speech recognizers are used in adverse environments. The recognition performance is typically degraded if the training and the testing environments are not the same. The study on speech recognition and understanding has been done for many years. The aim of the study was to observe the difference of English alphabet from E-set to AH-set. The aim of the study was to observe the difference of phonemes. Neural network is well-known as a technique that has the ability to classify nonlinear problem. Today, lots of researches have been done in applying Neural Network towards the solution of speech recognition. Even though positive results have been obtained from the continuous study, research on minimizing the error rate is still gaining lots of attention. This research utilizes Recurrent Neural Network, one of the Neural Network techniques to observe the difference of alphabet from E- set to AH - set. The purpose of this research is to upgrade the peoples knowledge and understanding on phonemes or word by using Recurrent Neural Network (RNN) and backpropagation through Multilayer Perceptron. 6 speakers (a mixture of male and female) are trained in quiet environment. The English language offers a number of challenges for speech recognition [4]. This paper specifies that the performance of Recurrent Neural Network is better than Multi Layer Perceptron Neural Network.

Keywords: Frames, Mel-frequency cepstral coefficient, Multi Layer Perceptron (MLP), Neural Networks, Performance, Recurrent Neural Network (RNN), Utterances.

———————————— ◆ ————————————

1. INTRODUCTION

Speech is human's most efficient communication modality. Beyond efficiency, humans are comfort and familiar with speech. Other modalities require more concentration, restrict movement and cause body strain due to unnatural positions. Research work on English speech recognition, although lagging that other language, is becoming more intensive than before and several researches have been published in the last few years [11]. Automatic speech recognition is a process by which a machine identifies speech. The conventional method of speech recognition insist in representing each word by its feature vector & pattern matching with the statistically available vectors using neural network [3]. The promising technique for speech recognition is the neural network based approach. Artificial Neural Networks, (ANN) are biologically inspired tools for information processing [15]. Speech recognition modeling by artificial neural networks (ANN) doesn't require a priori knowledge of speech process and this technique quickly became an attractive alternative to HMM [19]. RNN can learn the temporal relation ship of Speech – data & is capable of modeling time dependent phonemes [5].

The conventional neural networks of Multi- Layer Perceptron (MLP) type have been increasingly in use for speech recognition and also for other speech processing applications. Those networks work very well as an effective classifier for vowel sounds with stationary spectra, while their phoneme discriminating power deteriorates considerably for consonants which are characterized by variations of their short-term spectra. This may be attributable to a fact that feedforward multi-layer neural network are inherently unable to deal with time varying information like time-varying spectra of speech sounds. One way to cope with this problem is to incorporate feedback structure in the networks to provide them with an ability to memorize incoming time-varying information. Incorporating feedback structure in feedforward networks results in so-called Recurrent Neural Networks (RNNs) which have feedback connections between units of different layers or connections of self-loop type [6]. Speech recognition is the process of converting an acoustic signal, captured by microphone or a telephone, to a set of characters. The recognized characters can be the final results, as for applications such as commands and control, data entry and document preparation. They can also serve as the input to further linguistic processing in order to achieve speech understanding, a subject covered in section [19]. As we know, speech recognition performs their task similar with human brain. Start from phoneme, syllable, word and then sentence which is an input for speech recognition system [14].

2. PRESENT WORK

The objective of speech recognition is to determine the sequence of sound units from the speech signal so that the linguistic message in the form of text can be decoded from the speech signal. The steps used in the present speech recognition system are discussed below.

2.1 Input Acquisation

After capturing the speech by using microphone the speech data is saved in .wav files. The speech data is converted to

analog signal by using Praat object software tool. The signal is then converted into mono speech signal with 11kHz.

2.2 Front – End Analysis

The acoustic speech signal exists as pressure variations in the air. The micro phone converts these pressure variations into an electric current that is related to the pressure. The ear converts these pressure variations into a series of nerve impulses that are transmitted to the brain. Selection of features is very important for speech recognition task. Good features are required to achieve good result for recognition. Basic problem with speech recognition is identification of proper features for speech recognition task, and a strategy to extract these features from speech signal.

2.3 The Speech Utterance (Data Collection)

The source of data is a database consisting of 18 characters taken from 4 major sets and spoken 10 times by 6 speakers; those are 3 males and 3 females of various ages. Four major sets are

E set: B C D E P T G V Z
A set: J K
EH set: M N F S
AH set: I Y R

The data, which is speaker dependent, will be used for training and testing phases. In speaker dependent form, the first four utterances of each of the 18 characters spoken by every speaker are used to train the network and the remaining utterances are used to test the network. Therefore, the speech database contains 1080 utterances, which can be used for training the network, and 1080 utterances, which are available for testing. These characters are recorded by:- 1- Using Praat Object Software with sampling rate 11 kHz, 8-bit and mono is used to record the utterance. 2- In a closed room, the same microphone is used to record the spoken characters. 3- The files are saved in a .wav format.

2.4 Preprocessing

The speech signals are recorded in a low noise environment with good quality recording equipment. The signals are samples at 11kHz. Reasonable results can be achieved in isolated word recognition when the input data is surrounded by silence.

2.5 Sampling Rate

150 samples are chosen with sampling rate 11kHz, which is adequate to represent all speech sounds.

2.6 Windowing

To minimize the discontinuity of a signal at the beginning and end of each frame, we window each frame to increase the correlation of the MFCC. MFCC Spectral estimates between consecutive frames [8,16]. Inorder to avoid discontinuties at the end of speech segments the signal should be tapered to zero or near zero and hence reduce the mismatch. To the given 12 Mel-Frequency coefficients, and for time 0.005 seconds, a window length of 0.015 is selected by the Praat Object software tool.

2.7 Feature Extraction

Feature extraction consists of computing representations of the speech signal that are robust to acoustic variation but sensitive to linguistic content. The Mel-filter is used to find band filtering in the frequency domain with a bank of filters. The filter functions used are triangular in shape on a curvear frequency scale. The filter function depends on three parameters: the lower frequency, the central frequency and higher frequency. On a Mel scale the distances between the lower and the central frequencies and that of the higher and the central frequencies are equal. The filter functions are

$$H(f) = 0 \ \ for \ f \leq f_l \ \ and \ \ f \geq f_h$$

$$H(f) = (f - f_l)/(f_c - f_l) \ \ for \ \ f_l \leq f \leq f_c$$

$$H(f) = (f_h - f)/(f_h - f_c) \ \ for \ \ f_c \leq f \leq f_h$$

Mel frequency cepstral coefficients are found from the Discrete Cosine Transform of the
Filter bank spectrum by using the formula given by Davis and Mermelstein[1980].

$$c_i = \sum_{j=1}^{N} P_j \cos(i\pi / N(j - 0.5))),$$

Pj denotes the power in dB in the jth filter and N denotes number of samples.

12 Mel frequency coefficients are considered for windowing. Mel-Frequency analysis of speech is based on human perception experiments. Sample the signal with 11 kHz, apply the sample speech data to the mel-filter and the filtered signal is trained. Number of frames are obtained for each utterance from frequency coefficients by using Praat object software tool.

2.8 Neural Networks

Neural networks model some aspects of the human brains, where thinking process is achieved in synaptic connections between neurons. The structure of the network is layered and capable of high parallelism. Neural networks are useful in classification, function approximation and generally in complex problems, which do not require accurate solution. Neural networks must be taught before they can be used, which correspond to how humans learn. A Neural network consists of units that are interconnected with several other such units; they function independently on the input they are given and their local data. Usually all of the units are homogenous, but also heterogeneous networks exists.

Neural networks use a set of processing elements loosely analogous to neurons in the brain. These nodes are interconnected in a network that can then identify patterns in data as it is exposed to the data. In a sense, the network learns from experience just as people do. This distinguishes neural networks from traditional computing programs that simply follow instructions in a fixed sequential order. The structure of a neural network is given below:
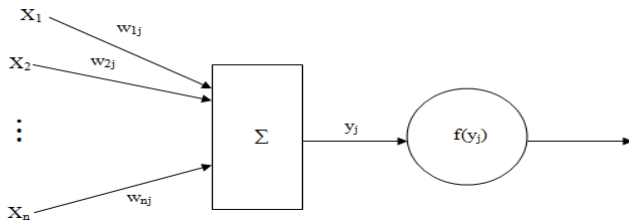
Fig 1: General Structure of Neural Network

A set of inputs is applied to each node representing the inputs from outside world or, alternatively, they may be outputs from other nodes. Each input is multiplied by a weight associated with the node input to which it is connected and the weighted inputs are then summed together. A threshold value local for each node is added to the weighted summation and the resulting sum is then passed through a hard limiting activation function. The sigmoid function is used as a transfer function.

2.9 Learning Method

Learning is necessary when the information about inputs/outputs is unknown or incomplete. Learning is the method of setting the appropriate weight values. There are two types of training namely supervised and unsupervised. The supervised learning method is used to train the neural network in this paper. Supervised learning requires the network to have an external teacher. The algorithm adjusts weights using input-output data to match the input-output characteristics of a network to the desired characteristics. In the learning without supervision, the desired response is not known and in supervised learning at each instant of time when the input is applied, the desired response of the system provided by the teacher is assumed. The distance between the actual and desired response serves as an error measure and is used to correct network parameters externally.

2.10 Training

The networks are usually trained to perform tasks such as pattern recognition, decision-making, and motory control. The original idea was to teach them to process speech or vision, similar to the tasks of the human brain. Nowadays tasks such as optimization and function approximation are common. Training of the units is accomplished by adjusting the weight and threshold to achieve a classification. The adjustment is handled with a learning rule from which a training algorithm for a specific task can be derived.

3. RECURRENT NEURAL NETWORK ARCHITECTURE

RNN have feedback connections and address the temporal relationship of inputs by maintaining internal states that have memory. RNN are networks with one or more feedback connection. A feedback connection is used to pass output of a neuron in a certain layer to the previous layer(s) [12]. The difference between MLP and RNN is RNN have feedforward connection for all neurons (fully connected). Therefore, the connections allow the network show the dynamic behavior. RNN seems to be more natural for speech recognition than MLP because it allows variability in input length [17].

The motivation for applying recurrent neural network to this domain is to take advantage of their ability to process short-term spectral features but yet respond to long-term temporal events. Previous research has confirmed that speaker recognition performance improves as the duration of utterance is increased [9]. In addition, it has been shown that in identification problems. RNNs may confer a better performance and learn in a shorter time than conventional feedforward networks [10].

Recently a simple recurrent neural network, which has feedback connections of self-loop type around hidden layer units, has been proposed as an attractive tool for recognizing speech sounds including voiced plosive sounds [12].This network has three layers such as input layer, hidden layer and output layer. Each of the output layer units has feedback connection with itself, i.e., a self-loop as shown in Fig. 2.

Fully recurrent networks feed back the hidden layer to itself. Partially recurrent networks start with a fully recurrent net and add a feedforward connection that bypasses the recurrency, effectively treating the recurrent part as a state memory. These recurrent networks can have an infinite memory depth and thus find relationships through time as well as through the instantaneous input space. Most real-world data contains information in its time structure. Recurrent networks are the state of the art in nonlinear time series prediction, system identification, and temporal pattern classification.
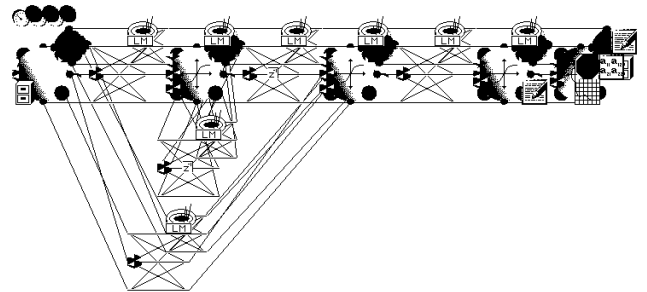
Fig 2: RNN Architecture

Actually this architecture has been used in visual pattern recognition but we use this architecture for speech recognition especially for English speech recognition by using Backpropagation Through Time (BPTT)as learning algorithm. This architecture also have been proved that this architecture is better than MLP in phoneme recognition accuracies [13] by using Backpropagation algorithm.

The Backpropagation Through Time (BPTT) algorithm is based on converting the network from a feedback system to purely feedforward system by folding the network over time. Thus, if the network is to process a signal that is time steps long, then copies of the network are created and the feedback connections are modified so that they are feedforward connections from one network to the subsequent network. The network can then be trained if it is one large feedforward network with the modified weights being treated as shared weight [18].

3.1. Multi-Layer Perceptron

This is perhaps the most popular network architecture in use today, due originally to Rumelhart and McClelland (1986). The units each performed a biased weighted sum of their inputs and pass this activation level through a transfer function to produce their output, and the units are arranged in a layered feedforward topology. The network thus has a simple interpretation as a form of input-output model, with the weights and thresholds (biases) the free parameters of the model. Such networks can model functions of almost arbitrary complexity, with the number of layers, and the number of units in each layer, determining the function complexity. Important issues in MLP design include specification of the number of hidden layers and the number of units in these layers.

The number of hidden units to use is far from clear. As good a starting point as any is to use one hidden layer, with the number of units equal to half the sum of the number of input and output units. Again, we will discuss how to choose a sensible number later.
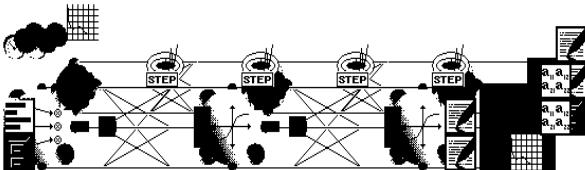
Fig 3: MLP Network architecture with step learning rule.

This network has an input layer (on the left) with three neurons, one hidden layer (in the middle) with three neurons and an output layer (on the right) with three neurons.

There is one neuron in the input layer for each predictor variable. In the case of categorical variables, N-1 neurons are used to represent the N categories of the variable.

Input Layer — A vector of predictor variable values $(x_1...x_p)$ is presented to the input layer. The input layer (or processing before the input layer) standardizes these values so that the range of each variable is -1 to 1. The input layer distributes the values to each of the neurons in the hidden layer. In addition to the predictor variables, there is a constant input of 1.0, called the bias that is fed to each of the hidden layers; the bias is multiplied by a weight and added to the sum going into the neuron.

Hidden Layer — Arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight (wji), and the resulting weighted values are added together producing a combined value uj. The weighted sum (uj) is fed into a transfer function, σ, which outputs a value hj. The outputs from the hidden layer are distributed to the output layer.

Output Layer — Arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight (wkj), and the resulting weighted values are added together producing a combined value vj. The weighted sum (vj) is fed into a transfer function, σ, which outputs a value yk. The y values are the outputs of the network.

If a regression analysis is being performed with a continuous target variable, then there is a single neuron in the output layer, and it generates a single y value. For classification problems with categorical target variables, there are N neurons in the output layer producing N values, one for each of the N categories of the target variable.

4. TRAINING PHASE

The multilayer back propagation algorithm is used to train the neural network for spoken characters for each speaker. Six speakers are trained using the multilayer perceptron with 108 input nodes, 2 hidden layers and 4 output nodes each for one word, with the noncurvear activation function sigmoid. The learning rate is taken as 0.1, momentum rate is taken as 0.5. Weights are initialized to random values between +0.1 and -0.1 and accepted error is chosen as 0.009. Frames are obtained for each mel-frequency coefficients of the utterance for each speaker.

4.1 Performance Evaluation

The performance are obtained for each speaker and presented in table 1.

Table 1: RNN Performance of E-Set

E-Set	Speaker1 %	Speaker2 %	Speaker3 %	Speaker4 %	Speaker5 %	Speaker6 %
B	92	92	96	94	94	90
C	92	87	97	94	93	93
D	91	93	90	95	97	95
E	90	91	89	92	95	88
P	89	100	98	97	96	91
T	99	100	98	97	96	91
G	93	92	90	93	95	91
V	98	93	89	91	95	91
Z	93	90	92	91	93	93
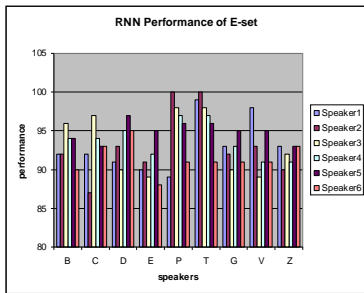
Fig 4: RNN Performance of E-Set

For E-Set, the maximum performances of speakers 1-6 are estimated as 99%, 100%, 98%, 97%, 97% & 95% respectively. For E-Set, the minimum performances of speakers 1-6 are estimated as 89%, 87%, 89%, 91%, 93% & 88% respectively.

Table 2: RNN Performance of A-Set

A-Set	Speaker1 %	Speaker2 %	Speaker3 %	Speaker4 %	Speaker5 %	Speaker6 %
J	93	99	93	90	92	90
K	92	99	96	93	92	94
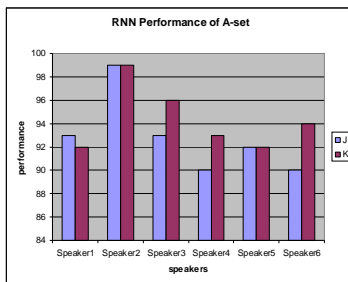
Fig 5: RNN Performance of A-Set

For A-Set, the maximum performances of speakers 1-6 are estimated as 93%, 99%, 96%, 93%, 92% & 94% respectively. For A-Set, the minimum performances of speakers 1-6 are estimated as 92%, 99%, 93%, 90%, 92% & 90% respectively.

Table 3: RNN Performance of EH-Set

EH-Set	Speaker1 %	Speaker2 %	Speaker3 %	Speaker4 %	Speaker5 %	Speaker6 %
M	91	95	91	91	95	98
N	100	89	92	95	91	92
F	92	90	93	90	95	90
S	89	90	98	90	98	96
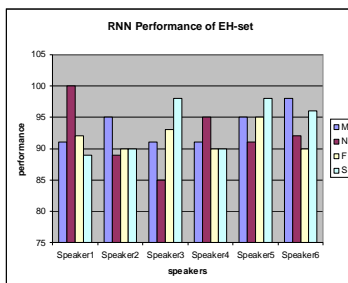
Fig 6: RNN Performance of EH-Set

For EH-Set, the maximum performances of speakers 1-6 are estimated as 100%, 95%, 98%, 95%, 98% & 98% respectively.

For EH-Set, the minimum performances of speakers 1-6 are estimated as 89%, 89%, 91%, 90%, 91% & 90% respectively.

Table 4: RNN Performance of AH-Set

AH-Set	Speaker 1 %	Speaker 2 %	Speaker 3 %	Speaker 4 %	Speaker 5 %	Speaker 6 %
I	93	95	96	93	95	91
Y	95	93	92	96	90	93
R	93	98	92	91	90	95
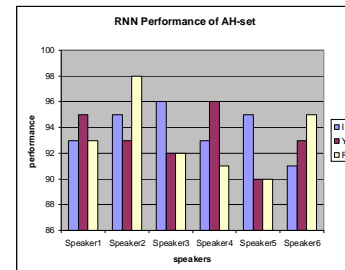
Fig 7: RNN Performance of AH-Set

For AH-Set, the maximum performances of speakers 1-6 are estimated as 95%, 98%, 96%, 96%, 95% & 95% respectively. For AH-Set, the minimum performances of speakers 1-6 are estimated as 93%, 93%, 92%, 91%, 90% & 91% respectively.

Table 5: MLP Performance of E-Set

E-Set	Speaker 1 %	Speaker 2 %	Speaker 3 %	Speaker 4 %	Speaker 5 %	Speaker 6 %
B	78	83	89	92	91	77
C	78	72	75	83	87	89
D	91	90	79	86	94	93
E	89	93	72	88	79	78
P	85	89	77	86	93	89
T	87	88	92	93	94	88
G	91	89	86	93	92	91
V	93	83	87	91	93	89
Z	92	89	90	78	85	83
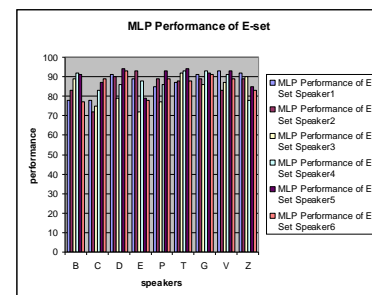
Fig 8: MLP Performance of E-Set

For E-Set, the maximum performances of speakers 1-6 are estimated as 93%, 93%, 92%, 93%, 94% & 91% respectively. For E-Set, the minimum performances of speakers 1-6 are estimated as 78%, 72%, 72%, 78%, 79% & 77% respectively.

Table 6: MLP Performance of A-Set

A-Set	Speaker 1 %	Speaker 2 %	Speaker 3 %	Speaker 4 %	Speaker 5 %	Speaker 6 %

| J | 87 | 93 | 89 | 82 | 90 | 86 |
| K | 87 | 98 | 90 | 83 | 86 | 89 |
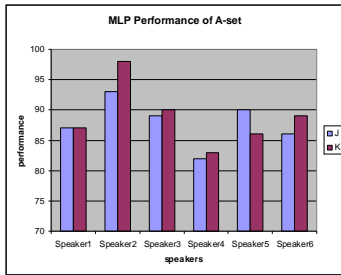
Fig 9: MLP Performance of A-Set

For A-Set, the maximum performances of speakers 1-6 are estimated as 87%, 98%, 90%, 83%, 90% & 89% respectively. For A-Set, the minimum performances of speakers 1-6 are estimated as 87%, 93%, 89%, 82%, 86% & 86% respectively.

Table 7: MLP Performance of EH-Set

EH-Set	Speaker 1 %	Speaker 2 %	Speaker 3 %	Speaker 4 %	Speaker 5 %	Speaker 6 %
M	89	91	89	79	93	95
N	88	85	90	93	79	89
F	90	88	93	86	91	84
S	86	89	95	89	95	93
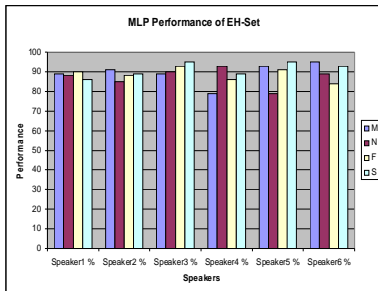
Fig 10: MLP Performance of EH-Set

For EH-Set, the maximum performances of speakers 1-6 are estimated as 90%, 91%, 95%, 93%, 95% & 95% respectively. For EH-Set, the minimum performances of speakers 1-6 are estimated as 86%, 85%, 89%, 79%, 79% & 84% respectively.

Table 8: MLP Performance of AH-Set

| AH-Set	Speaker 1 %	Speaker 2 %	Speaker 3 %	Speaker 4 %	Speaker 5 %	Speaker 6 %
I	79	93	92	93	90	90
Y	89	86	90	92	89	90
R	92	95	89	91	88	91
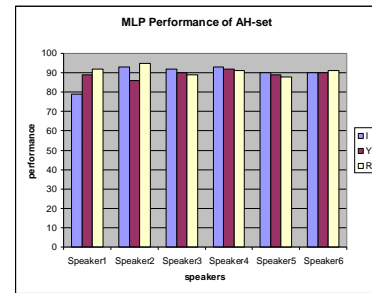
Fig 11: MLP Performance of AH-Set

For AH-Set, the maximum performances of speakers 1-6 are estimated as 92%, 95%, 92%, 93%, 90% & 91% respectively. For AH-Set, the minimum performances of speakers 1-6 are estimated as 79%, 86%, 90%, 91%, 88% & 90 % respectively.

5. TESTING PHASE

The same multilayer backpropagation algorithm is used to test the network of the spoken characters for the six speakers. Each speaker has to test the network by 18 characters repeated ten times. Each speaker, tests the word ten times and the node with the higher number in the output will be the winner node. The correct answer will be indicated by comparing this node with the input word to the network.

So by testing the characters said by each speaker the performance can be found by the equation.

Performance = Total succeeded number of testing characters / Total number of characters * 100%

Table9: The performance for the test phase for each speaker.

Speaker	Performance for E- Set	Performance for A-Set	Performance for EH-Set	Performance for AH-Set
1 (Male)	100%	100%	77%	95%
2 (Male)	100%	80%	93%	89%
3 (Male)	75%	82%	97%	94%
4 (Female)	85%	91%	93%	81%
5 (Female)	92%	93%	82%	81%
6 (Female)	87%	85%	85%	73%

The Maximum Performances of sets E, A, EH, AH are 100%, 100%, 97%, and 95% respectively. The Minimum Performances of sets E, A, EH, AH are 75%, 80%, 77%, and 73% respectively.

CONCLUSION

This paper shows that Recurrent Neural Networks are very powerful in classifying speech signals. Even with simplified models, a small set of characters could be recognized. The performance of the nets is heavily dependent on the quality of pre-processing. Mel Frequency Cepstrum Coefficients is very reliable. Both the Multilayer Feedforward Network with backpropagation algorithm and the Recurrent Neural Network are achieving satisfying results. The results

obtained in this study demonstrate that speech recognition is feasible, and that Recurrent Neural Networks used are well suited for the task.

REFERENCES

[1] Ben Gold and Nelson Morgan Speech and Audio Signal Processing, Wiley India Edition, New Delhi, 2007.

[2] B. Yegnanarayana, Artificial neural networks Prentice-Hall of India, New Delhi, 2006.

[3] John Coleman, "Introducing Speech and language processing", Cambridge university press, 2005.

[4] Mayfield T. L., Black A. and Lenzo K., (2003). "Arabic in my Hand: Small-footprint Synthesis of Egyptian Arabic." Euro Speech 2003, Geneva, Switzerland.

[5] D.A.Reynolds, "An overview of Automatic speaker recognition technology", proc. ICASSP 2002, orlands, Florinda, pp.300-304.

[6] Medser L. R. and Jain L. C., (2001). "Recurrent Neural Network: Design and Applications." London, New York: CRC Press LLC.

[7] R.O. Duda, P.E. Hart, and D.G. Strok Pattern Classification, 2nd edn, John Wiley, New York, 2001.

[8] Picton, P.Neural Networks, Palgrave, NY (2000).

[9] He J. and Liu L., (1999). "Speaker Verification Performance and The Length of Test Sentence." Proceedings ICASSP 1999 vol.1, pp.305-308.

[10] Gingras F. and Bengio Y., (1998). "Handling Asynchronous or Missing Data with Recurrent Networks." International Journal of Computational Intelligence and Organizations, Vol. 1, no. 3, pp. 154-163

[11] Jihene El Malik, (1998). "Kohonen Clustering Networks For Use In Arabic Word Recognition System." Sciences Faculty of Monastir, Route de Kairouan, 14-16 December.

[12] RuxinChenand Jamieson L. H., (1996). "Experiments on the Implementation of Recurrent Neural Networks for Speech Phone Recognition." Proceedings of the Thirtieth Annual Asilomar Conference on Signals, Systems and Computers, Pacific Grove, California, November, pp. 7790782.

[13] Koizumi T., Mori M., Taniguchi S. and Maruya M., (1996). "Recurrent Neural Networks for Phoneme Recognition." Department of Information Science, Fukui University, Fukui, Japan, Spoken Language, ICSLP 96, Proceedings, Fourth International Conference, Vol. 1, 3-6 October, Page(s): 326 -329.

[14] Joe Tebelskis, (1995). "Speech Recognition using Neural Network."Carnegie Mellon University: Thesis Ph.D.

[15] C.M.Bishop, Neural Networks for pattern recognition, oxford university press, 1995.

[16] Rabiner, L and Juang, B, -H; fundamentals of speech recognition, PTR prentice Hall, scan Francisco, N.J (1993).

[17] Lee S. J., Kim K. C., Yoon H. and Cho J. W., (1991). "Application of Fully Neural Networks for Speech Recognition." Korea Advanced Institute of Science and Technology, Korea, Page(s): 77-80.

[18] Werbos P., (1990). "Backpropagation Through Time: What It Does and How To Do It." Proceedings of the IEEE, 78, 1550.

[19] Lippman R.P., (1989). "Review of Neural Network for Speech Recognition." Neural Computation 1.1-38.

Balanced Ant Colony Algorithm for Scheduling DAG to Grid Heterogeneous System

Mrs. Smitha Jha

Abstract- Ant Colony Optimization can be used for scheduling tasks on resources in Grid. In earlier work this technique has been applied for independent task scheduling. This paper applies the above technique for dependent task scheduling. Here a hybrid algorithm by Sakellariou can be applied, where tasks in DAG (Directed acyclic graph) are upward ranked and sorted decreasingly. Then the sorted tasks are grouped along the sorted sequences and in every group, tasks are independent. Then independent task groups can be scheduled to resources using algorithm specified by author Chang.

Index Terms – Grid System, Directed Acyclic graph(DAG),Independent, Dependent task scheduling, Ant Colony Optimization, Grouping, Ranking.

———————————— ◆ ————————————

1 INTRODUCTION

Current scientific problems are very complex and need huge computing power and storage space. The past technologies such as distributed or parallel computing are unsuitable for current scientific problems with large amounts of data. Processing and storing massive volumes of data may take a very long time. Grid computing [3] is a new paradigm for solving those complex problems. In grids, we need to consider the conditions such as network status and resources status. If the network or resources are unstable, jobs would be failed or the total computation time would be very large. So we need an efficient job scheduling algorithm for these problems in the grid environment. Because the environment status may change frequently, traditional job scheduling algorithm such as ''First Come First Serve'' (FCFS), ''Shortest Job First'' (SJF), etc., may not be suitable for the dynamic environment in grids.

- *Mrs Smitha Jha is currently pursuing Ph.D degree program in Computer Science from BIRLA INSTITUTE OF TECHNOLOGY,EXT. CENTER NOIDA,UP,INDIA, PH-09953100878. E-mail: s,jha@bitmesra.ac.in*

In grids, users may face hundreds of thousands of computers to utilize. It is impossible for anyone to manually assign jobs to computing resources in grids. Therefore, grid job scheduling is a very important issue in grid computing.

Please refer to a survey [3], which also poses some open issues. A good schedule would adjust its scheduling strategy according to the changing status of the entire environment and the types of jobs. Therefore, a dynamic algorithm in job scheduling such as Ant Colony Optimization (ACO) [4,5] is appropriate for grids. ACO is a heuristic algorithm with efficient local search for combinatorial problems. ACO imitates the behavior of real ant colonies in nature to search for food and to connect to each other by pheromone laid on paths traveled. Many researches use ACO to solve NP-hard problems such as traveling salesman problem [6],graph coloring problem [7], vehicle routing problem [8], andso on. In [1] this technique has been applied for independent task scheduling. This paper apply the above technique for dependent task scheduling. Here a hybrid algorithm by Sakellariou[2] can be applied,where tasks in DAG(Directed acyclic graph) are upward ranked and sorted decreasingly. Then the sorted tasks are grouped along the sorted sequences and in every group,tasks are independent. We assume each job is an ant and the algorithm sends the ants to search for resources. We also modify the global and local pheromone update functions in ACO algorithm in order to balance the load for each grid resource.

The rest of the paper is organized as follows. Section 2 introduces the related work about many kinds of ACO algorithm and job scheduling in grids. Section 3 details the proposed ACO optimization algorithm, its pseudocode and

algorithm analysis in job scheduling for dependent task scheduling. Section 4 gives the comparison of proposed algorithm with other existing methods in terms of time complexity. Section 5 concludes paper and Section 6 lists references.

2.RELATED WORK:

2.1. Ant algorithms

There are many different kinds of ACO algorithm, i.e., AntColony System (ACS) [6], Max-Min Ant System (MMAS) [9],Rank-based Ant System (RAS) [10], Fast Ant System (FANT) [11] and Elitist Ant System (EAS) [12]. ACS uses the pseudo-randomproportional rule to replace state transition rule for decreasing computation time of selecting paths and update the pheromone on the optimal path only. It is proved that it helps ants search the optimal path.MMAS is based on the basic ACO algorithm but limiting the pheromone range to be greater than or equal to the low bound value (Min) and smaller than or equal to the upper bound value(Max). The low bound and upper bound are defined by the user.According to the low bound and upper bound values, MMAS could avoid ants to converge too soon in some ranges.

In the design of RAS, it sorts the ants by ant's tour length in ascending order after all ants completed their tours. It means that the first ant finds the shortest path to complete the tour and the last ant takes the longest tour. They give each ant a different density of pheromone to update their path by the ascending order: the higher the position of the ant, the more pheromone it could update; the lower the position of the ant, the less pheromone it has. By the idea of RAS, the shortest length gets more pheromone to attract more ants to follow and the system could get the optimal solution very soon, and it updates pheromone after each iteration. In order to avoid the sub-optimal solution, it applies a reset pheromone function.EAS update more pheromone on the best-so-far tour found inorder to attract more ants to follow the best-so-far tour.

2.2 Job scheduling in grids

Job scheduling is well studied within the computer operating systems [13]. Most of them can be applied to the grid environment with suitable modifications. In the following we introduce several methods for grids. The FPLTF (Fastest Processor to Largest Task First) [14] algorithm schedules tasks to resources according to the workload of tasks in the grid system. The algorithm needs two main parameters such as the CPU speed of resources and workload of tasks. The scheduler sorts the tasks and resources by their workload and CPU speed then assigns the largest task to the fastest available resource. If there are many tasks with heavy workload, its performance may be very bad. Dynamic FPLTF (DPLTF) [15] is based on the static FPLTF, it gives the highest priority to the largest task.DPLTF needs prediction information on processor speeds and taskworkload. The WQR (Work Queue with Replication) is based on the workqueue (WQ) algorithm [15]. The WQR sets a faster processor with more tasks than a slower processor and it applies FCFS and random transfer to assign resources. WQR replicates tasks in order to transfer to available resources. The amount of replications is defined by the user. When one of the replication tasks is finished,the scheduler will cancel the remaining replication tasks. The WQR's shortcoming is that it takes too much time to execute and transfer replication tasks to resource for execution.

Min-min [16] set the tasks which can be completed earliest with the highest priority. The main idea of Min-min is that it assigns tasks to resources which can execute tasks the fastest. Maxmin[16] set the tasks which has the maximum earliest completion time with the highest priority. The main idea of Max-min is that it overlaps the tasks with long running time with the tasks with short running time.

For instance, if there is only one long task, Min-min will execute short tasks in parallel and then execute long task. Max-min will execute short tasks and long task in parallel.The RR (Round Robin) algorithm focuses on the fairness problem. RR uses the ring as its queue to store jobs. Each job in queue has the same execution time and it will be executed in turn.If a job can't be completed during its turn, it will store back to the queue waiting for the next turn. The advantage of RR algorithm is that each job will be executed in turn and they don't have to wait for the previous one to complete. But if the load is heavy, RR will take long time to complete all jobs.

Priority scheduling algorithm gives each job a priority value and uses it to dispatch jobs. The priority value of each job depends on the job status such as the requirement of memory sizes, CPU time and so on. The main problem of this

algorithm is that it may cause indefinite blocking or starvation if the requirement of a job is never being satisfied.The FCFS (First Come First Serve) algorithm is a simple job scheduling algorithm. A job which makes the first requirement will be executed first. The main problem of FCFS is its convoy effect [13].If all jobs are waiting for a big job to finish, the convoy effect occurs.The convoy effect may lead to longer average waiting time and lower resource utilization.

3 BALANCED ANT COLONY ALGORITHM FOR DEPENDENT TASK SCHEDULING

3.1 ARCHITECTURE OF THE SCHEDULING SYSTEM

Model Description:

Suppose Directed acyclic graph(DAG) given below represents a set of dependent tasks.
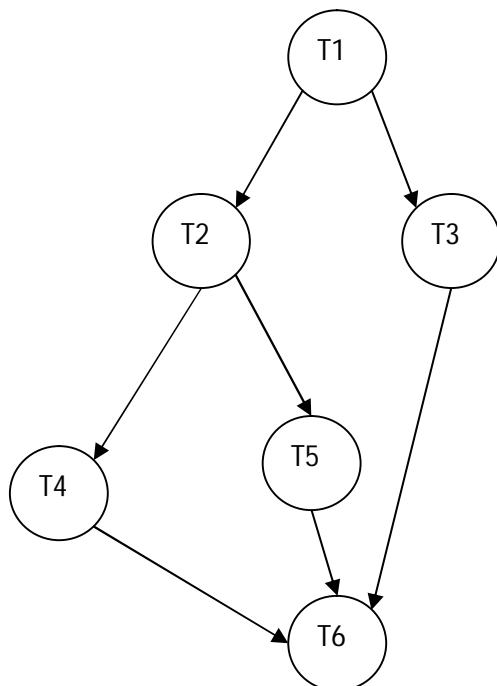
Fig-1(Given Directed Acyclic graph representing set of dependent tasks)

The DAG is divided into n no. of groups containing independent sets of tasks.

$$G = G_0 \cup G_1 \cup G_2 ... \cup G_{n-1}$$

With the condition that all the tasks in G_{i+1} will be executed after all the tasks in G_i has been executed for all i=0 to n-1.
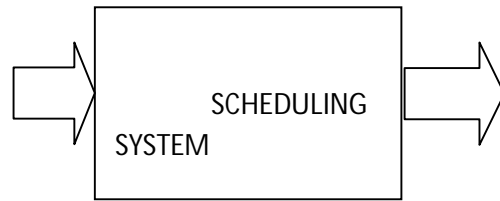
Input- [$G = G_0 \cup G_1 \cup G_2 ... \cup G_{n-1}$]

Output- MACHINES(A cluster of Workstations)

Fig-2 Scheduling Architecture

This can be solved by taking a cluster of workstations connected by internet as shown in Fig-3

$G_0, G_1, G_2 ... G_{n-1}$ can be applied to these workstations with a condition that there should be synchronization among them,so that G_{i+1} group uses the result of tasks completed in G_i group

$\forall i = 0 \, to \, n - 1$.Each group G_i tasks can be executed using Balanced Ant Colony Optimization.
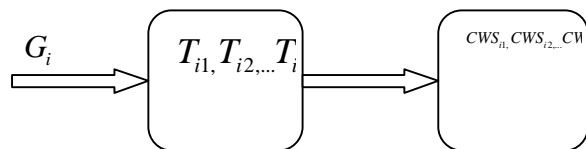
Fig-4 Assigning tasks to cluster of workstations.

Here each of the task is modelled as an ant and each of the machine in the cluster of workstation is

the food . Each ant select food through shortest path i.e. where the pheromone density is more.

The balanced antcolony algorithm for scheduling tasks to resources is described in the next section.

3.1 Proposed algorithm:

In this section we are presenting the algorithm for the dependent task scheduling in Grid heterogeneous system using Ant Colony Optimization technique.

Step 1: Input the Directed acyclic graph(DAG) G depicting how different tasks are dependent to each other.[In the graph each node representing task node and the edges between them represents the dependency among them]

Step 2: Rank the nodes in the graph using Ranking() method.

Step 3:Using GroupCreation() method ,a set of groups of independent tasks are made.

Step 4:for Groupno=1;Groupno<n;Groupno++ [for n no. of groups of independent tasks]

                Apply AntColony(Groupno.) method on each group.

Ranking() method algorithm

{

1. Assign a weight to each node as the average computation cost across all machines.
2. Assign a weight to each edge as the average communication cost across all machines.

3. Use upward ranking to compute the rank value for each node, The rank value R_i of a node I is recursively defined as follows

    $R_i = W_i + \max(C_{ij} + R_j)$ for all j present in S_i

    Where

W_i=weight of the node.

S_i=the set of immediate successor of node i.

C_{ij}=Weight of the edge connecting nodes i and j.

}

GroupCreation() method algorithm

{

1.Sort nodes in descending order of their rank values.

    $G_0=\{\},i=0;$

2.Scan nodes in descending order of the rank values
If current node has a dependence with a node G_i
    Then i++; and $G_i=\{\}$; Add node to G_i
3.Repeat step 2 until there are no more nodes.

}

SchedulingAntColony(Group no.) algorithm

{

[An Ant in the Ant system is a job in the Grid system. Pheromone value on a path in the Ant system is a weight for a resource in the Grid system. A resource with a large weight value means that the resource has a better computing power. The scheduler collects data from information server and uses the data to calculate a weight value of the resource. The pheromone(weight) of each resource is stored in the scheduler and the scheduler uses it as the parameter for the Balanced Ant Colony Optimization algorithm. At last the selects the resource according to the algorithm .]

The pheromone indicator(PI)(the initial pheromone value of each job to the resource)

$$PI_{ij} = \left[\frac{M_j}{BW_i} + \frac{T_j}{CS*(1-LD_i)} \right]$$

(1)

Where

M_j =Size of a given job.

BW_i =Network Bandwidth between scheduler and the resource i.

T_j =CPU time needed for the job j.

LD_i =Current load at resource i.

CS=CPU speed of resource i.

The load,Bandwidth and CPU speed can be obtained by NWS(Network Weather Service).

1. Do steps 2 to 4 until all the tasks are assigned
2. The resource status(How many jobs are there in resource, i. e. how much is it busy,the program execution time and the size of jobs) is told by PI. The larger the value of PI_{ij} , the more efficient it is to assign job j to the resource i.
3. The PI matrix is as follows for m resources and n jobs.

$$\begin{bmatrix} P_{11} & P_{12} & ... & PI_{1n} \\ P_{21} & P_{22} & ... & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{m1} & P_{m2} & \cdots & P_{mn} \end{bmatrix}$$

(2)

The largest entry from the PI matrix is chosen say P_{ij} ,then the jth job assigned to the resource ith. After this equation (2) is updated using equation (1) for m resources and (n-1) jobs. This is called local pheromone update.

4. The global pheromone update is to recalculate the entire PI matrix,after the resource completes the job.

3.2 Pseudocode of the proposed algorithm

In this section we are presenting the pseudocode for the dependent task scheduling using Ant Colony Optimization technique.

Procedure SchedulingDependenttask_Antcolony(Graph G)

{

Procedure MakingIndependent taskgroup();

For(i=1;i<groupno.;i++){

Procedure SchedulingAntcolony(groupno.);

}

}

Procedure Makingindependent taskgroup()

(

Procedure Ranking();

Procedure GroupCreation();

}

Procedure Ranking()

{

1.Assign a weight to each node as the average computation cost across all machines.
2.Assign a weight to each edge as the average communication cost across all machines.
3.Use upward ranking to compute the rank value for each node.}

Procedure GroupCreation()

{

1.Sort nodes in descending order of their rank values.

G0={},i=0;

2.Scan nodes in descending order of the rank values
If current node has a dependence with a node Gi
        Then i++; and Gi={}; Add node to Gi

    3.Repeat step 2 until there are no more nodes.

}

The rank value Ri of a node I is recursively defined as follows

Ri= Wi + max(Cij +Rj) for all j present in Si

Where

Wi=weight of the node.

Si=the set of immediate successor of node i.

Cij=Weight of the edge connecting nodes i and j.

Procedure SchedulingAntColony(Group no.)

{

[An Ant in the Ant system is a job in the Grid system. Pheromone value on a path in the Ant system is a weight for a resource in the Grid system. A resource with a large weight value means that the resource has a better computing power. The scheduler collects data from information server and uses the data to calculate a weight value of the resource. The pheromone(weight) of each resource is stored in the scheduler and the scheduler uses it as the parameter for the Balanced Ant Colony Optimization algorithm. At last the selects the resource according to the algorithm .]

The pheromone indicator(PI)(the initial pheromone value of each job to the resource)

$$PI_{ij} = \left[\frac{M_j}{BW_i} + \frac{T_j}{CS*(1-LD_i)} \right]$$

(1)

Where

M_j =Size of a given job.

BW_i =Network Bandwidth between scheduler and the resource i.

T_j =CPU time needed for the job j.

LD_i =Current load at resource i.

CS=CPU speed of resource i.

The load,Bandwidth and CPU speed can be obtained by NWS(Network Weather Service).

No_of_tasks_assigned=0;

Do
{
1.The resource status(How many jobs are there in resource, i. e. how much is it busy,the program execution time and the size of jobs) is told by PI.
The larger the value of PI_{ij} , the more efficient it is to assign job j to the resource i.
2.The PI matrix is as follows for m resources and n jobs.

$$\begin{bmatrix} P_{11} & P_{12} & ... & PI_{1n} \\ P_{21} & P_{22} & ... & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{m1} & P_{m2} & \cdots & P_{mn} \end{bmatrix}$$

(2)

The largest entry from the PI matrix is chosen say P_{ij} ,then the jth job assigned to the resource ith.
After this equation (2) is updated using equation (1) for m resources and (n-1) jobs. This is called local pheromone update.
3.The global pheromone update is to recalculate the entire PI matrix,after the resource completes the job.
No_of_tasks_assigned++;

}while(no_of_ tasks _assigned<totalno_tasks)

3.3 Algorithm analysis

Let T represent the Time complexity of the proposed algorithm.

T =T(Ranking)+T(Grouping)+T(ACO)

Where ACO is Ant Colony Optimization algorithm.

T(Ranking)=O(logn)
T(Grouping)=O(nlogn)[Tasks are sorted using heap sort)
T(ACO)= $o(n^2 P)$

4 COMPARISON WITH OTHER METHODS:

There exist other clustering heuristics for dependent task scheduling[3].Dominant Sequence Clustering(DSC) algorithm proposed by Yang et al[4] has time complexity for a DAG as $O((e+n)\log n)$ which is equal to $O(n^2 \log n)$ for a dense graph.. Similarly CLASS-II,a cluster algorithm proposed by Liou et al [5] has time complexity as $O(e+n\log n)$ which is equal to $O(n^2 + n\log n)$ for a dense graph. The fig-5 depicts a comparison of these methods with respect to no. of tasks to be assigned.

5 CONCLUSION

From above algorithm analysis it is concluded that ACO gives better performance with respect to DSC and CLASS II clustering heuristic methods for a fixed no. of processors. In this paper pseudocode of the proposed algorithm is written which has to be coded in C in next paper.

6 REFERENCES

[1] An ant algorithm for balanced job scheduling in grids Ruay-Shiung Chang_, Jih-Sheng Chang, Po-Sheng Lin Department of Computer Science and Information Engineering, National Dong Hwa University, Shoufeng Hualien, 974 Taiwan, ROC

[2].A Hybrid heuristic for DAG Scheduling on Heteorogeneous Systems. Rizos Sakellario and Henan Zhao ,Department of Computer Science,University of Manchester,Oxford Road,Manchester M13 9PL,U.K.

[3]. Fangpeng Dong and Selim G. Akl "Scheduling Algorithms for Grid Computing: State of the Art and Open Problems" Technical Report No. 2006-504

[4]. M. Dorigo, C. Blum, Ant colony optimization theory: A survey, Theoretical
Computer Science 344 (2–3) (2005) 243–278.

[5]. M. Dorigo, Ant colony optimization, http://www.aco-metaheuristic.org.

[6]. M. Dorigo, L.M. Gambardella, Ant colony system: A cooperative learning approach
to the traveling salesman problem, IEEE Transactions on Evolutionary
Computation 1 (1) (1997) 53–66.

[7]. E. Salari, K. Eshghi, An ACO algorithm for graph coloring problem, in: Congress
on Computational Intelligence Methods and Applications, 15–17 Dec. 2005,
p. 5.

[8]. Xiaoxia Zhang, Lixin Tang, CT-ACO—hybridizing ant colony optimization
with cycle transfer search for the vehicle routing problem, in: Congress on
Computational Intelligence Methods and Applications, 15–17 Dec. 2005, p. 6.

[9]. T. Stutzle, MAX-MIN Ant System for Quadratic Assignment Problems Technical
Report AIDA-97-04, Intellectics Group, Department of Compute Science,
Darmstadt University of Technology, Germany, July 1997.

[10]. B. Bullnheimer, R.F. Hartl, C. Strauss, A new rank-based version of the ant
system: A computational study, Central European Journal for Operations
Research and Economics 7 (1) (1999) 25–38.

[11]. E.D. Taillard, L.M. Gambardella, Adaptive memories for the quadratic assignment
problem, Technical Report IDSIA-87-97, IDSIA, Lugano, Switzerland,
1997.
[12]. M. Dorigo, V. Maniezzo, A. Colorni, The ant system: Optimization by a colony of
cooperating agents, IEEE Transactions on Systems, Man, and Cybernetics, Part
B 26 (1) (1996) 29–41.

[13]. Abraham Silberschatz, Peter Baer Galvin, Greg Gagne, Operating System
Concepts, 7th edition, John Wiley & Sons, 2005.

[14]. D. Saha, D. Menasce, S. Porto, Static and dynamic processor scheduling
disciplines in heterogeneous parallel architectures, Journal of Parallel and

Distributed Computing 28 (1) (1995) 1–18.

[15]. D. Paranhos, W. Cirne, F. Brasileiro, Trading cycles for information: Using replication to schedule bag-of-tasks applications on computational grids, in: International Conference on Parallel and Distributed Computing (Euro-Par), in: Lecture Notes in Computer Science, vol. 2790, 2003, pp. 169–180.

[16]. M. Maheswaran, S. Ali, H.J. Siegel, D. Hensgen, R. Freund, Dynamic matching and scheduling of a class of independent tasks onto heterogeneous computing
system, Journal of Parallel and Distributed Computing 59 (1999) 107–131
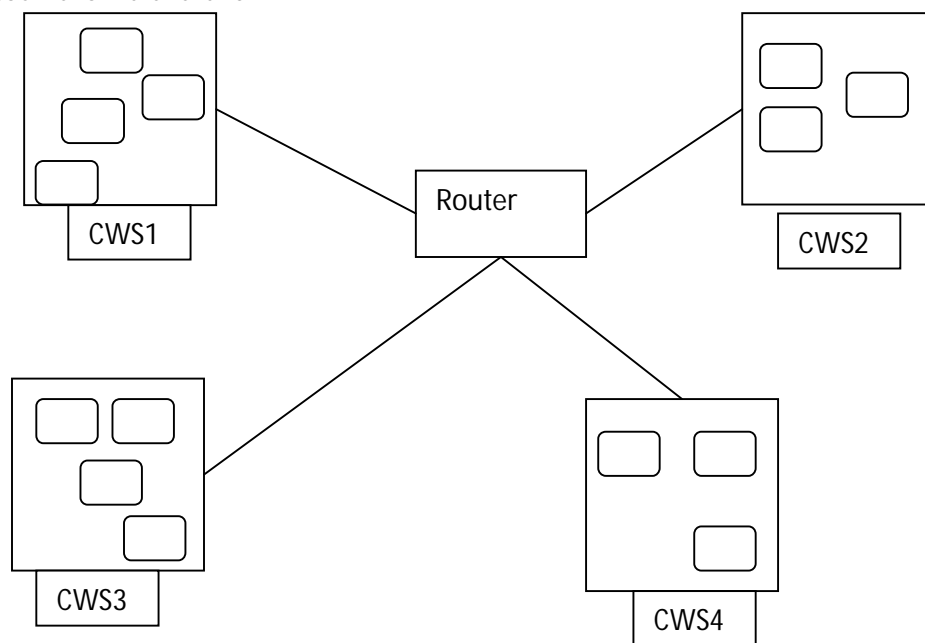
Figure-3
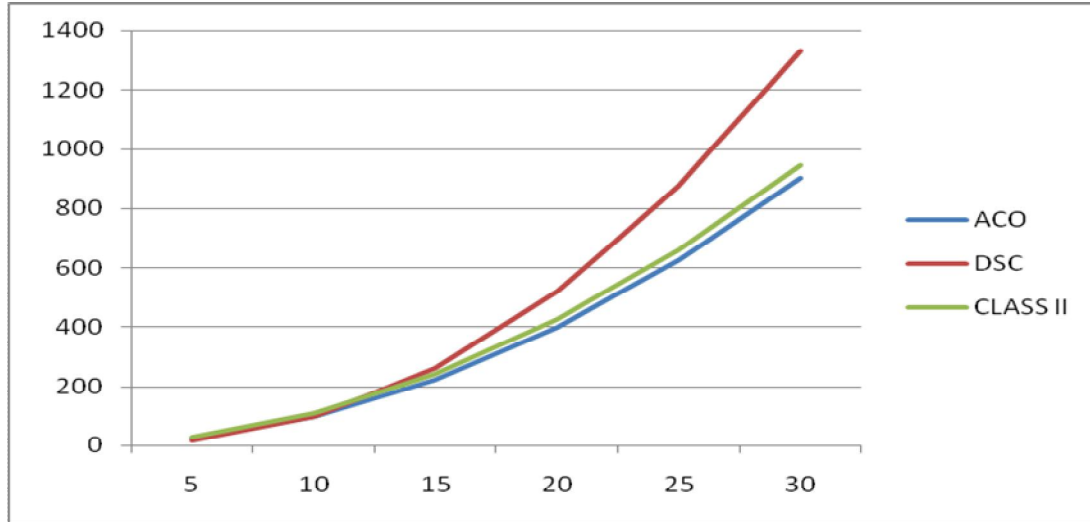
X AXIS-NO. OF TASK

Y AXIS-TIMECOMPLEXITY

Fig-5

Analysis of Optical Wireless Communication for Underwater Wireless Communication

P.Vijaya Kumar, S.S.K.Praneeth, Romarsha.B.Narender

Abstract- Recently, the importance of underwater wireless optical communication has been grown for applications of underwater observation and sea monitoring systems. This communication technology is expected to play a prominent role in investigating climate changes, prediction of natural disasters, discovery of natural resources, marine biology in lake, sea, and ocean environment. Though acoustic modems have long been the default wireless communication method for under water applications due to their long range, the need for high speed communication has prompted the exploration of non-acoustic methods that `have previously been overlooked due to their distance limitations. Optical wireless communication has been proposed as the best alternative in order to overcome the limitations in acoustic underwater communication. In this paper we have presented three scenarios regarding underwater optical wireless communication. They are a) Line of sight (LOS) b) Modulating Retro Reflector link c) Reflective link. In this analysis, each scenario has been tested with water having different extinction coefficient and various noise backgrounds. We analysed these links in OPTI SYSTEM. Opti System is an innovative, rapidly evolving, and powerful software design tool that enables users to plan, test, and simulate almost every type of optical link in the transmission layer of a broad spectrum of optical networks from LAN, SAN, MAN to ultra-long-haul. It offers transmission layer optical communication system design and planning from component to system level, and visually presents analysis and scenarios.

Index Terms- Line of sight scenario, Modulating Retro Reflecting link, Reflective link, Bit Error Rate (BER), extinction coefficient, clean ocean water, coastal ocean water, turbid ocean water.

———————————— ◆ ————————————

1. INTRODUCTION

THE underwater world is an important part of our planet, which covers 70 % of it.Nowadays acoustic technology is mostly used for establishing wireless communication link Among divers and ships, or sending long range remote signals. This is because sound waves travel through water faster than in air, receiving very little attenuation. However, due to frequency attenuation characteristic of acoustic waves in water, it is difficult to expand its bandwidth. Therefore acoustic approach cannot achieve high data rate, and also portable communication devices are difficult to be designed at lower cost. Even though the electromagnetic waves are so widely used there are some portions of our planet that are not suitable for these types of transmissions. In this realm the use of electromagnetic waves is very difficult, if not impossible, due to large attenuations of the electric and magnetic filled, this limits it range and effectiveness to transmit data. The other types of wave known are mechanical waves, for the underwater channel acoustics waves are the only solution for wireless communication. Acoustics waves can travel so much easily especially in sea water, where salinity shows strong conductivity.

Since humans are limited in their ability to work underwater, remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) have been in service since the 1950s to perform underwater tasks, such as collecting data and retrieving items. Operation of these vehicles is challenging , but oil resources are found further off shore, ROVs and AUVs are required to go deeper and stay deployed longer in perform critical tasks. One such task is monitoring a deep sea oil well-sending tethered ROVs thousands of meters below the surface in order to conduct survey is expensive and time consuming. To overcome this challenge we need underwater optical wireless communication system.

Cochenour, Mullen and Laux [8] measure both the spatial and temporal effects of scattering on a laser link in turbid underwater environments. Using Monte Carlo simulations and measurement results, they predict longer- range underwater optical free-space optical performance with bandwidths greater than 5 GHz for a range of 64 m in clear ocean water, dropping to a range of 1 GHz for a range of 8 m in turbid harbour water. J.H.Smart [9] examines the fundamental physics and natural variability of underwater optical attenuations and discusses the design issues underwater optical wireless communication associated with ocean physics and parameter variability. D.Kedar and S. Arnon [11] examine the potential of subsea free-space optics for sensor networks applications. Leveraging the emerging technologies of high sensitive photon counting detectors and semiconductor LED and laser light source in the solar blind UV. The same authors propose a novel non-line-of-sight network concept in which the optical link is implemented by means of back reflection of the propagation optical

signal at the ocean-air interface, which would help to overcome obstructions.

The paper by jaruwatanadilok [11] presents the modelling of an underwater wireless optical communication channel using the vector radiation transfer theory. The vector radiative transfer equation link. All the three links we discuss provide required data rate. We also present the performance analyses on the three links. From the analyses it is clear that water absorption increases due to changes in turbidity, the communication performance decrease dramatically for all three link types, but the modulated reflector link is the most affected. However, the absorption coefficient increases more moderately than does the water turbidity. We conclude from the analyses that a high data rate underwater optical wireless network is a feasible solution for emerging applications such as UUV to UUV links and networks of sensors. The analyses also include opti system. This software design tool have some benefits such as Providing global insight into system performance, Assessing parameter sensitivities aiding design tolerance specifications, Visually presenting design options and scenarios. This enables users to plan, test, and simulate the WDM/TDM or CATV network designs, SONET/SDH ring designs, Transmitter, channel, amplifier, and receiver designs, Estimation of BER and system penalties with different receiver models.
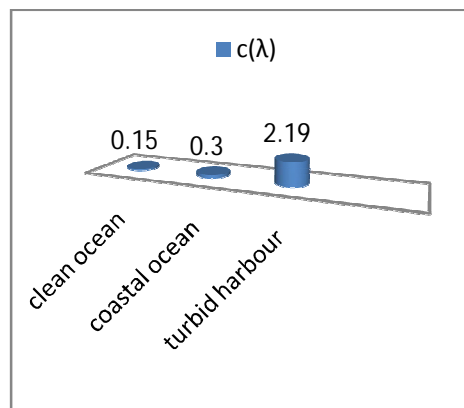
Fig.1 These are the extinction coefficient values for different kinds of water.

2.PROPERTIES OF THE UNDERWATER OPTICAL WIRELESS COMMUNICATION CHANNEL

capture the multiple scattering in natural water, and also includes the polarization of light.

Here in this paper we present about three models of optical wireless communication 1.line of sight link (LOS) 2.Modulating retro reflector link 3.Reflective

Light pulses propagating in aquatic medium suffer from attenuation and broadening in the spatial, angular, temporal and polarization domains. The attenuation and broadening are wavelength dependent and result from absorption and multi-scattering of light by water molecules and by marine hydrosols (mineral and organic matter).

The extinction co-efficient $c(\lambda)$ of the aquatic medium is governed by the absorption and scattering coefficient $\alpha(\lambda)$ and $\beta(\lambda)$, respectively, and we have

$$C(\lambda) = \alpha(\lambda) + \beta(\lambda) \qquad (1)$$

Figure (1) depicts the absorption, scattering, and extinction coefficient for three types of water – clean ocean water, coastal ocean water, and turbid harbor water. It is clear that increase in the turbidity dramatically increases the extinction coefficient from less than $0.1 m^{-1}$ for pure water up to more than $2 m^{-1}$ for turbid harbor water. However the absorption coefficient increases more moderately than does the turbidity.

The propagation loss factor as a function of wavelength and distance z is given by

$$L_{pr}(\lambda, z) = \exp(-c(\lambda) z) \qquad (2)$$

3. COMMUNICATION LINK MODELS

We now consider three types of communication links: the line of sight, the modulating retro reflector, and the reflective. In addition, we perform a bit error rate (BER) calculation.

3.1. Line-of-Sight Communication Link

Line of sight is a straight and unobstructed path of communication between transmitter and receiver. This is the most common link between two points in optical wireless communication system. The figure represents line of sight (LOS). In this scenario, the transmitter directs the light beam in the direction of the receiver. The optical signal reaching the receiver is obtained by multiplying the transmitter power, telescope gain, and losses and is given by

$$P_{R\_los} = P_T \eta_T \eta_R L_{pr}\left(\lambda, \frac{d}{\cos\theta}\right) \frac{A_{Rec}\cos\theta}{2\pi d^2(1-\cos\theta_0)}, \tag{3}$$

where PT is the average transmitter optical power, η_T is the optical efficiency of the transmitter, η_R is the optical efficiency of the receiver, d is the perpendicular distance between the transmitter and the receiver plane, θ is the angle between the perpendicular to the receiver plane and the transmitter-receiver trajectory, A_{Rec} is the receiver aperture area, and θ_0 is the laser beam divergence angle. When the Transmitter beam divergence angle is very narrow ($\theta_0 \ll \Pi/20$), Eq. (3) can be approximated as

$$P_{R\_los} = P_T \eta_T \eta_R L_{pr}\left(\lambda, \frac{d}{\cos\theta}\right) \frac{A_{Rec}\cos\theta}{\pi(d\tan\theta_0)^2}.$$

$$\tag{4}$$
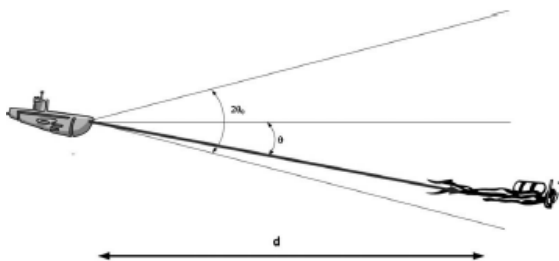
Fig.2. Shows the line of sight communication link.

3.2. Modulating Retro Reflector Communication Link

A Modulating Retro Reflector (MRR) system combines an optical retro reflector and an optical modulator to allow optical communications and sometimes other functions such as programmable signage. The modulating retro-reflector link is used when one party (for example, a submarine) has more resources another one (for example, a diver), as in Fig (2). In this case, the submarine has more energy, payload, and lifting capacity than the diver. Therefore it would be wise to put most of the complexity and power requirement of the communication system into the submarine. In a modulating retro-reflector link, the interrogator sits at one end (in our case, in the submarine), and a small modulating optical retro-reflector sits at the remote end. In operation, the interrogator illuminates the retro-reflecting end of the link with a continuous wave beam. The retro reflector in actively reflects this beam

back to the interrogator while modulating the information on it. The received power in this scenario is given by

$$P_{R\_Retro} = P_T \eta_T \eta_{Rec} \eta_{Retro} L_{pr}\left(\lambda, \frac{2d}{\cos\theta}\right)\left[\frac{A_{Retro}\cos\theta}{2\pi d^2(1-\cos\theta_0)}\right]$$
$$\times \left[\frac{A_{Rec}\cos\theta}{\pi(d\tan\theta_{0retro})^2}\right], \tag{5}$$

Where η_{Retro} is the optical efficiency of the retro reflector, θ is the angle between the perpendicular to the receiver plane and the transmitter-receiver trajectory, A_{Retro} is the retro reflector's aperture area, and θ_{0Retro} is the retro reflector's beam divergence angle.
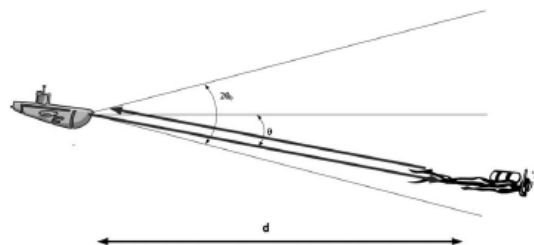
Fig.3. Shows the modulating retro reflective communication link

3.3. Reflective Communication Link

In some communication scenarios the line of sight is not available due to obstructions, misalignment, or random orientation of the transceivers. To address this problem a reflective communication link could be used. In this case, the laser transmitter emits a cone of light, defined by inner and outer angles θ_{min} and θ_{max}, in the upward direction. Here θ_i and θ_t are the angles of incidence and of transmission, respectively. (The latter is derived from the former using Snell's law).

The light reaching the ocean-air surface illuminates an annular area and is partially bounced back in accordance with the reflectivity. Since the refractive index of air is lower than that of water, total internal reflection (TIR) can be achieved above a critical incidence angle. When the transmitter is at depth h, the illuminated annular surface with equal power density at depth x is given by

$$A_{ann} = 2\pi(h+x)^2(1-\cos\theta_{max} - 1 + \cos\theta_{min})$$
$$= 2\pi(h+x)^2(\cos\theta_{min} - \cos\theta_{max}). \tag{6}$$

Equation (6) describes an annular area taken from a sphere of radius $h+x$, which would have uniform power density in free space .If we model the ocean-air

surface as smooth, then $\theta=\theta_i$, and we can derive the link budget by using the variables defined in Eq. (3). Then we can define the auxiliary function and calculate the received power as

$f_{R\_ref}(\theta) = P_T\cos\theta/A_{ann}=\{\Box_T\Box_R L_{pr}(\lambda,h+x/\cos\theta)$
$1/2\{[\tan(\theta_t-\theta)/\tan(\theta_t+\theta)]^2+[\sin(\theta-\theta_t)/\sin(\theta+\theta_t)]^2\}$
$\qquad\qquad \theta min \le \theta \le \theta_c$
$\{\Box_T\Box_R L_{pr}(\lambda,h+x/\cos\theta),\ \theta_c \le \theta \le \theta_{max}$ (7)

At the plane of the receiving sensor, node coverage is provided with in an annular area bounded by radii $(h+x)$ Tan θ_{min} and $(h+x)$Tan θ_{max}. Equation (7)can be simplified on the assumption that the receiver aperture is small relative to $h+x$, yielding the approximate received power as

$$P_{R\_ref}(\theta) \approx A_{Rec}f_{R\_ref}(\theta).$$

(8)
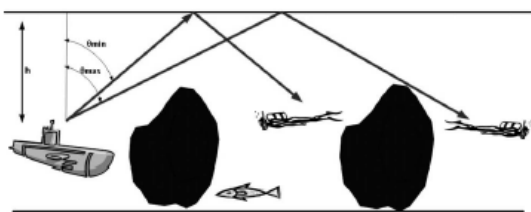
Fig.4 .Shows the reflective communication link

4. STIMULATION AND RESULTS

By using the values in table.1 we have simulated the loss at receiver end in the three scenarios using MATLAB. The results have been shown in the tables below from table II to table XVI. This gives the consistency of receiver power in the three scenarios. We have implemented the MATLAB results in OPTI SYSTEM. In opti system we have implemented spatial optical transmitter, spatial optical receiver, and OWC channel with required losses which were obtained in MATLAB. We have analysed the output of the receiver using different analysers such as RF spectrum analyser, oscilloscope visualizer, eye diagram analyser, BER calculator.

From fig (4) we can infer that for a time (bit period) of 0.3 to 0.75 the quality factor (Q) raises more than 2000.

Fig (5) represents the eye diagram for LOS link for c (λ)=0.3.since the eye is widely opened, we came to know that the ISI value is more. This leads to the reduction of expected data rate. In the fig (9) and fig (12) the eye patterns of the retro reflector link and reflective link has been given respectively. In both of them the noise levels are very high which results in ineffective communication between transmitter and receiver and data rate is reduced.

Table II, III, IV gives the receiver power of line of sight link at $\theta=30^0$ at different distances and for three extinction coefficients. From this if we consider for distance d=20, the loss value is more in turbid harbour water than the other two. Table V, VI, VII gives the receiver power for line of sight link at different θ values at a constant distance d=50m. From this we can know that the effectiveness of each water type changes with the change in value of θ. The same we can observe in modulating retro reflector link and reflective link also. Hence the turbid harbour water creates a bad impact on the data rate than the other two water types. Now if we consider table II and VIII the receiver power of LOS link and retro reflector link for coastal ocean water at constant $\theta=30^0$ the loss is more in retro reflector link. Hence line of sight link is more effective than the other two links.

TABLE 1
These are the parameters used in the numerical calculation

Parameter	value
Extinction coefficient for	
• Clean ocean	0.15
• Coastal ocean	0.30
• Turbid harbour	2.19
Critical angle (deg)	48.44
Optical efficiency for	
• Retro reflector	0.9
• Transmitter	0.9
• Receiver	0.9
Transmitter power (W)	1
Receiver aperture area (m^2)	0.01

Retro reflector aperture area (m^2)	0.01
Retro reflector beam divergence (θ_{retro}) (deg)	10
Beam divergence angle θ_0 (deg)	68
Transmitter inclination angles $\theta_{min},\theta_{max}$ (deg)	0,68
Transmitter depth h (m)	20
Receiver depth x (m)	20
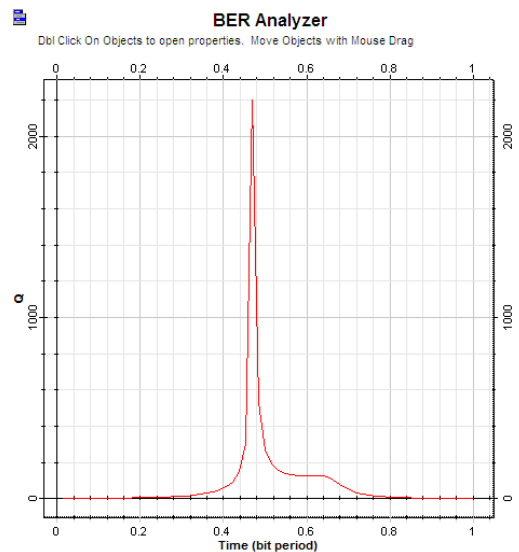
Fig.5.This graph shows BER analysis of Line of sight link at receiver side.
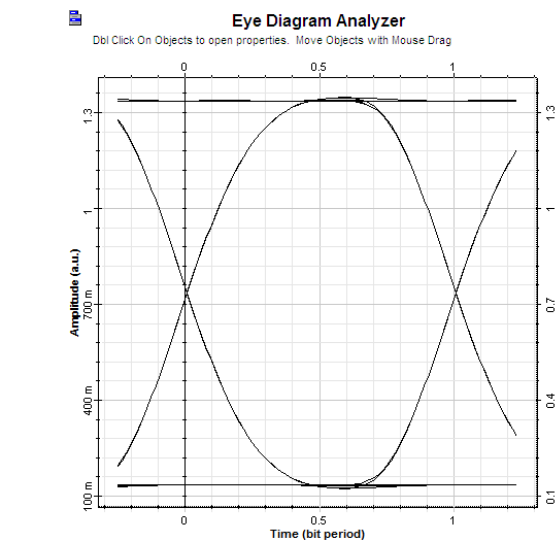
Fig.6. This graph shows eye diagram for Line of sight link.
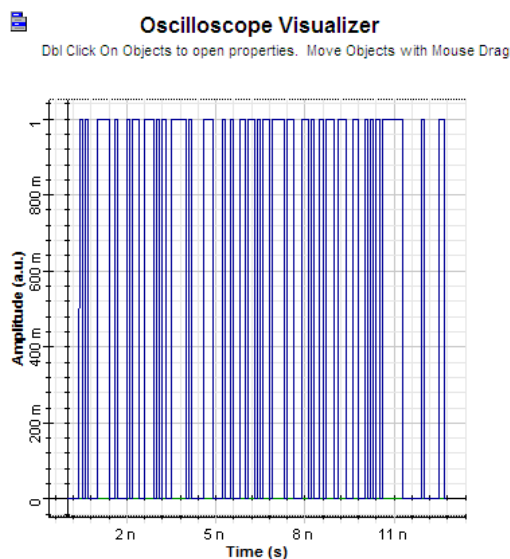
Fig.7. This graph shows oscilloscope visualizer output for Line of sight link
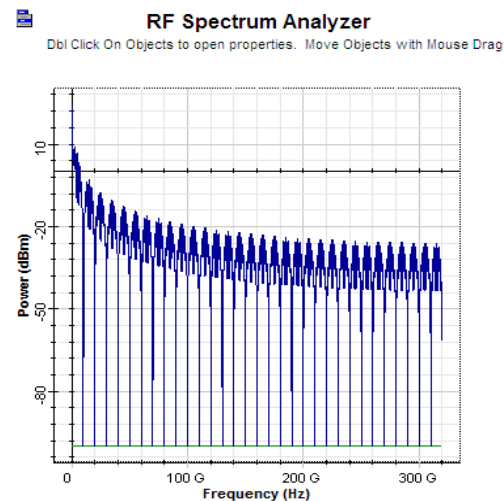
Fig.8. This graph shows RF Spectrum analysis for line of sight link
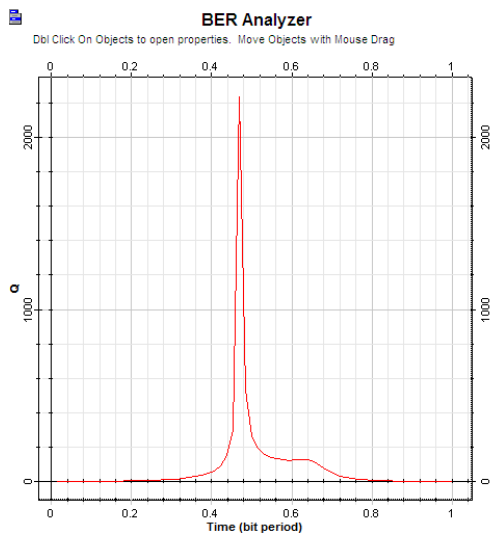
Fig .9. This graph shows BER analysis for Modulated Retroreflector Link.
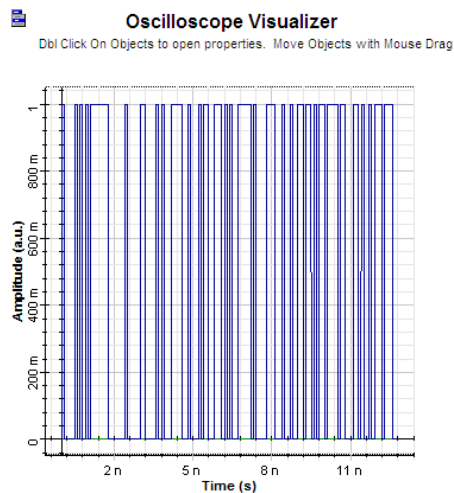
Fig.11. This graph shows oscilloscope visualizer for Modulated Retro reflector Link
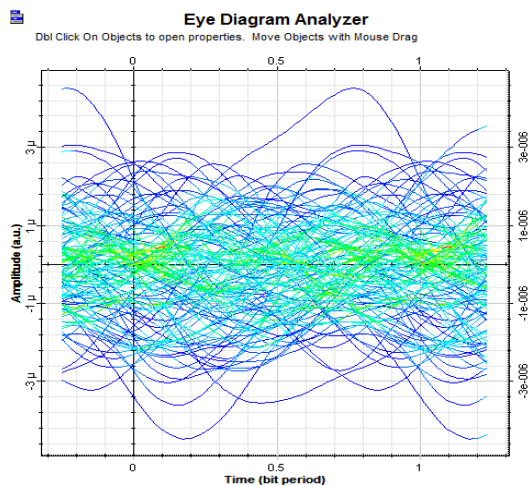
Fig .10.This graph shows eye diagram analysis for Modulated Retro reflector Link
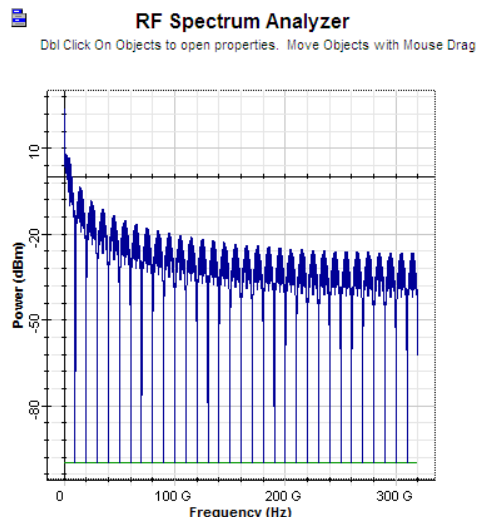
Fig .12. This graph shows RF Spectrum Analyser for Modulated Retro reflector Link
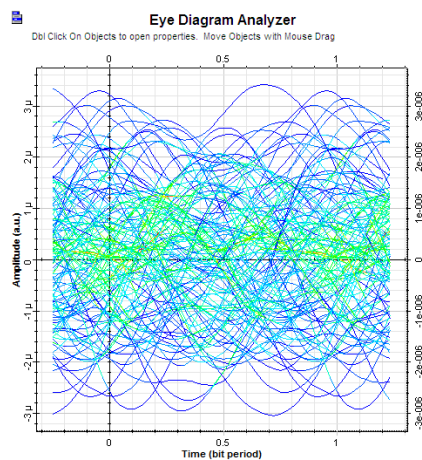
Fig.13. This graph shows eye diagram analysis for Reflective link
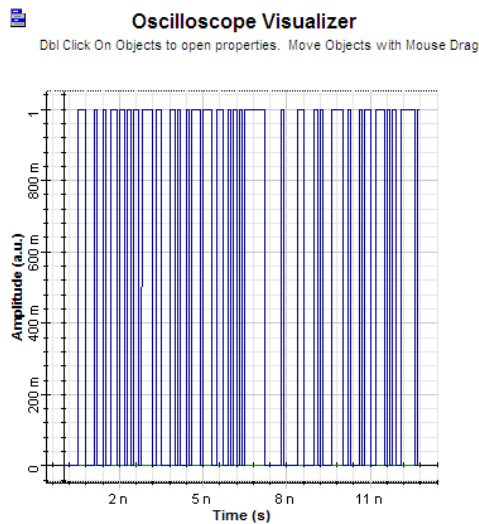
Fig.14. This graph shows Oscilloscope Visualizer for reflective link.
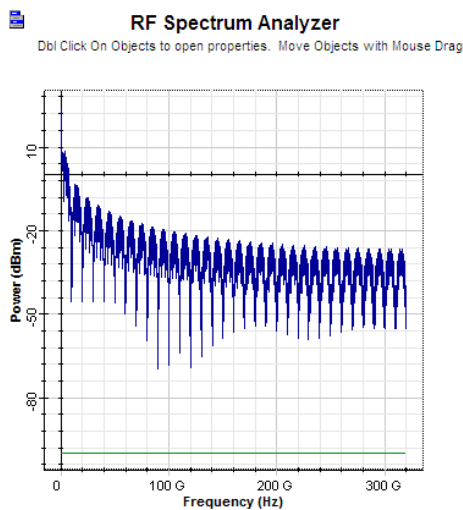
Fig .15. This graph shows RF Spectrum Analysis for Reflective Link

TABLE 2

This gives the receiver power(pr) in db at different distance(d) for line of sight link at constant θ=30 in coastal ocean water.

d(m)	0	10	20	30	40
Pr(db)	-27.4	-42.5	-57.5	-72.6	-87.6

TABLE 3

This gives the receiver power(pr) in db at different distance(d) for line of sight link at constant θ=30 in clean ocean water.

d(m)	0	10	20	30	40
Pr(db)	-27.4	-35	-42.5	-50	-57.5

TABLE 4

This gives the receiver power (pr) in db at different distance(d) for line of sight link at constant θ=30 in turbid harbour water.

d(m)	0	10	20	30	40
Pr(db)	-27.4	-137	-247	-356	-466

TABLE 5

This gives the receiver power (pr) in db at various t values for line of sight link at constant distance d=50m in coastal ocean water.

θ(deg)	0	15	30	45	60
Pr(db)	-126	23	-491	-187	7.3

TABLE 6

This gives the receiver power (pr) in db at various t values for line of sight link at constant distance d=50m in clean ocean water.

θ(deg)	0	15	30	45	60
Pr(db)	-93.4	-19.1	-280	-125	-26.8

TABLE 7

This gives the receiver power (pr) in db at various t values for line of sight link at constant distance d=50m in turbid ocean water.

θ(deg)	0	15	30	45	60
Pr(db)	-536	563	-3152	-968	432

TABLE 8

This gives the receiver power (pr) in db at different distance(d) for modulated retro reflector link at constant θ=30 in coastal ocean water.

d(m)	10	20	30	40	50
Pr(db)	-108	-150	-187	-222	-256

TABLE 9

This gives the receiver power (pr) in db at different distance(d) for modulated retro reflector link at constant θ=30 in clean ocean water.

d(m)	10	20	30	40	50
Pr(db)	-93.4	-120	-142	-162	-181

TABLE 10

This gives the receiver power (pr) in db at different distance(d) for modulated retro reflector link at constant θ=30 in turbid harbour water.

d(m)	10	20	30	40	50
Pr(db)	-297	-528	-753	-977	-1200

TABLE 11

This gives the receiver power (pr) in db at various t values for modulating retro reflector link at constant distance d=50m in coastal ocean water.

θ(deg)	0	15	30	45	60
Pr(db)	-235	63.93	-966	-358	31.19

TABLE 12

This gives the receiver power (pr) in db at various t values for modulating retro reflector link at constant distance d=50m in clean ocean water.

θ(deg)	0	15	30	45	60
Pr(db)	-170	-21.8	-543	-234	-37.2

TABLE 13

This gives the receiver power (pr) in db at various t values for modulating retro reflector link at constant distance d=50m in turbid harbour water.

θ(deg)	0	15	30	45	60
Pr(db)	-1056	1144		-1921	893

TABLE 14

This gives the receiver power (pr) in db at various t values for reflective link at constant h+x=40m in coastal ocean water.

θ(deg)	0	15	30	45	60
Pr(db)	-113	5.95	-407	-163	-6.94

TABLE 15

This gives the receiver power (pr) in db at various t values for reflective link at constant h+x=40m in clean ocean water.

θ(deg)	0	15	30	45	60
Pr(db)	-87.3	-28.1	-238	-113	-34.3

TABLE 16

This gives the receiver power (pr) in db at various t values for reflective link at constant h+x=40m in turbid harbour water.

θ(deg)	0	15	30	45	60
Pr(db)	-441	438	-2536	-788	337

V.CONCLUSION

The results above indicate the losses at the receiver end for medium distances. From this we can observe that as the turbidity of the water increases (i.e. extinction coefficient) the absorption of the water increases. So the losses at the receiver end also increases. Thus the communication in turbid harbour water reduces the data rate and results in an ineffective communication.Among the scenarios we discussed, the line of sight scenario is most effective than other two. Additional improvements to the availability of the network could be achieved by a hybrid communication system that would include an optical transceiver and an acoustical transceiver. A hybrid communication system can provide high data rate transmission by using optical transceiver. When the water turbidity is high or the distance between the terminals is large, the system can be switch to low data rate using the acoustic transceiver, thereby increase in the average data rate and availability.

REFERENCE

[1] Shalomi Arnon "Underwater optical communication" (2010).

[2] Milica Stojanovic "Wireless underwater communication system and networks: current achievements and future research challenges" (1994).

[3] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," Ad Hoc Networks **3**(3), 255-256 (2005).

- *P.vijaya kumar is working as assistant professor (AP(SRG)) in electronics and communication engineering. In SRM.university. PH NO:7358927171,email: vijayakumarp@ktr.srmuniv.ac.in*
- *S.S.K.prneeth is pursuing his B.Tech in electronics and communication engineering in SRM university.PH NO: 9176082100,email:pranusatya@gmail.com*
- *Romarsha.B.narender is pursuing his B.Tech in electronics and communication engineering in SRM university. PH NO:9791023649,email:romarshab@gmail.com*

[4] J. Heidemann, W. Ye, J. Wills, A. Syed, and Y. Li, "Research challenges and applications for underwater sensor networking," in *Proc. IEEE Wireless Communications and Networking Conf.*, pp. 228–235(2006).

[5] T. Dickey, M. Lewis, and G. Chang, "Optical oceanography; recent advances and future directions using global remote sensing and in situ observations," *Rev. Geophys.* **44**(1), RG1001 (2006).

[6] C.Detweiller, I. Vasilescu, and D. Rus, "AquaNodes: an underwater sensor network," in *Proc. Second Int. Workshop on Underwater Networks*, pp. 85–88, IEEE (2007).

[7] F. Hanson and S. Radic, "High bandwidth underwater optical communication," Appl. Opt. **47**(2), 277—283 (2008).

[8] B. Cochenour, L. Mullen, and A. Laux, "Spatial and temporal dispersion in high bandwidth underwater laser communication links," in *Proc. IEEE Military Communications Conf.*, pp. 1–7 (2008).

[9] J. H. Smart, "Underwater optical communication systems part 1: variability of water optical parameters," in *Proc. IEEE Military Communications Conf.*, pp. 1140–1146(2005).

[10] J. W. Giles and I. N. Bankman, "Underwater optical communications systems part 2: basic design considerations," in *Proc. IEEE Military Communications Conf.*, pp. 1140–1146(2005).

[11] D. Kedar and S. Arnon, "Subsea ultraviolet solar-blind broadband free-space optics communication," *Opt. Eng.* **48**(4), 046001(2009).

[12] S. Jaruwatanadilok, "Underwater wireless optical communication channel modelling and performance evaluation using vector radiative Transfer theory," *IEEE J. Sel. Areas Commun.* **26**(9), 1620–1627 (2008).

Analysis of Surface Roughness of Machined Surface of Powder Metallurgy Components

P.K.Bardhan, R.Behera, S.Patra, G.Sutradhar

Abstract - The surface roughness value (Ra) of sintered iron P/M components at different cutting speed has been investigated. Surface finish may also be critical for component assembly or system performance. Dimensional fit and mating surface interaction may require certain surface finish requirements to meet performance specifications. Experimental results of surface roughness (Ra) of P/M components at different cutting speed have been analyzed through the various process parameters during manufacturing using response surface model. It has been observed that the compaction pressure, sintering temperature and sintering time strongly influence the response variable, surface roughness. A second order response surface model (RSM) has been used to develop a predicting equation of surface roughness based on the data collected by a statistical design of experiments known as central composite design (CCD). The analysis of variance (ANOVA) shows that the observed data fits well into the assumed second order RSM model.

Keywords - Surface roughness, Sintered components, Hardness, Response surface, Central composite design, ANOVA.

——————————— ◆ ———————————

1. INTRODUCTION

Powder Metallurgy Steel components often have to be machined after heat treatment in order to obtain the correct shape as well as the required surface finish. Surface quality influences characteristics such as fatigue strength, wear rate, corrosion resistance, etc**.**

The surface finish of a component may be critical for certain applications, affecting properties such as wear resistance, fatigue strength, and coefficient of friction.

The surface analysis is one of the most important factors of the metal machining process due to the tolerance and geometry

————————————————

- Department of Mechanical Engineering, JIS College of Engineering, Kalyani, West Bengal, India.
- Department of Mechanical Engineering, Seemanta Engineering College,Mayurbhanj, Orissa,India
- CWISS IIT Kharagpur, West Bengal, India
- Department of Mechanical Engineering, Jadavpur University, West Bengal, India.

requirements.[1-2] The characteristic of powder metallurgy (P/M) surface geometry is the main issue, because the porous structure affects the surface quality. The overall smoothness and surface reflectivity depend on density, tool finish, and secondary machining operations. A discontinuous cutting path and some vibration occur when cutting tool passes from the edge of one pore to that of another [3]. Experimental results of surface roughness (Ra) of P/M components at different cutting speed has been analyzed through the various process parameters using response surface model. It has been observed that the compaction pressure, sintering temperature and sintering time strongly influence the response variable, surface roughness. A second order response surface model (RSM) has been used to develop a predicting

equation of surface roughness based on the data collected by a statistical design of experiments known as central composite design (CCD). The analysis of variance (ANOVA) shows that the observed data fits well into the assumed second order RSM model.

2. Experimental procedures

Kawasaki Steel Corporation Chiba Works, Chiba, Japan, supplied the iron Powder The relevant certification of chemical analysis and powder particle size distribution was performed by the same company and is presented in Table 1.

Table 1

Chemical Analysis of iron powder

C	Si	Mn	P	S	O	Total Fe
0.001	0.02	0.17-	0.013	0.010	0.129	Balance

Powder Properties: Apparent Density (gm/cc): 2.94
Flow (s/50gm)   : 24.7

Sieve Distribution:

Sieve Number	Size	Cumulative wt%
+ 100#	>150 um	8.5
+ 150#	>106 um	20.1
+ 200#	> 75 um	22.9
+ 250#	> 63 um	9.5
+ 325#	> 45 um	16.8
- 325#	<45 um	22.2

The iron powder was compacted in a closed cylindrical die using 120-Ton hydraulic press (make-Lawrence & Mayo) for green stage product (fig.1). During compaction, the die was lubricated with Zn-stearate. The sintering process was carried out in a tubular vacuum furnace of capacity 1450°C using argon as an inert atmosphere (fig.2). Since one of the major objectives of present investigations is to shade light on the hardness of the compacted sintered samples, 60 different P/M iron components (dia-25 mm) were produced according to design of experiment (DOE). Related surface roughness value of these samples were studied by Surftest SJ-301 (Mitutoyo) machine (fig.3)

against the variation of controllable process parameters like compaction, sintering time and sintering temperature. The results obtained through the experiments are given in Table 2 &. 3 and the available data have been analyzed by response surface method using Minitab software (version 14).

Fig.1.120 Ton. Hydraulic press.

Fig.2: Tubular Vacuum Furnace

Fig.3. Surface Roughness   Surf test – SJ-301

3. Results and Discussions

Table 2 and Table 3 depict a variation of surface analysis against the process parameters.

Table 2
Actual and coded values of process parameters and symbols used.

Process parameters (Independent variables)	Symbols		Levels					
	Actual	Coded	Actual			Coded		
Compaction load (Ton)	z_1	x_1	17.66	20.075	26.4 9	−1	0	+1
Sintering temperature (° C)	z_2	x_2	975	1050	1125	−1	0	+1
Sintering time (hrs)	z_3	x_3	1	1.5	2	−1	0	+1

Table 3.

Observed Surface roughness (Ra) – values for different settings of process parameters.

Sl. No.	Coded Value of Parameters			Actual Value of Parameters			Response variables Surface roughness Ra μm		
	x_1	x_2	x_3	Compact-ion Ton	Sintering Temp. °c	Sintering Time hour	R2 (@ cutting speed 18.37 m/min.	R2 (@ cutting speed 27.95m/min	R2 (@ cutting speed 4.24 m/min
1	-1	-1	-1	17.66	975	1	7.92	7.45	8.92
2	1	-1	-1	26.49	975	1	3.34	4.73	4.25
3	-1	1	-1	17.66	1125	1	5.71	7.16	6.86
4	1	1	-1	26.49	1125	1	2.31	2.75	2.29
5	-1	-1	1	17.66	975	2	6.31	6.88	8.57
6	1	-1	1	26.49	975	2	3.34	4.15	4.65
7	-1	1	1	17.66	1125	2	5.08	5.02	6.35
8	1	1	1	26.49	1125	2	3.25	4.40	5.65
9	-1.6818	0	0	14.6499	1050	1.5	7.08	7.22	8.59
10	1.68179	0	0	29.5001	1050	1.5	2.62	2.32	3.07
11	0	-1.6818	0	22.075	923.87	1.5	6.06	6.44	6.77
12	0	1.68179	0	22.075	1176.13	1.5	4.08	4.21	4.77
13	0	0	-1.6818	22.075	1050	0.6591	7.26	5.00	8.30
14	0	0	1.68179	22.075	1050	2.3409	5.40	4.98	6.52
15	0	0	0	22.075	1050	1.5	6.92	6.14	9.11
16	0	0	0	22.075	1050	1.5	6.75	4.78	7.73
17	0	0	0	22.075	1050	1.5	7.16	7.16	9.34
18	0	0	0	22.075	1050	1.5	7.56	7.92	6..32

| 19 | 0 | 0 | 0 | 22.075 | 1050 | 1.5 | 7.25 | 6.88 | 8.72 |
| 20 | 0 | 0 | 0 | 22.075 | 1050 | 1.5 | 6.88 | 7.02 | 7.24 |
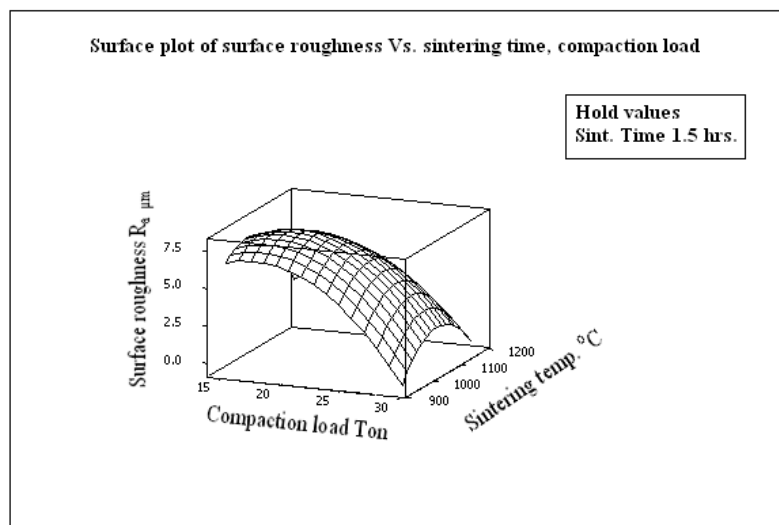
Fig. 4. Surface Plot of Surface roughness R_a $_{\mu m}$ (R1) vs Compaction load(X1), Sintering time
(X3) Sintering temperature. X2

From the fig. 4. it is quite evident that with gradual increase of compaction load and sintering temperature, surface roughness initially increases up to a maximum value(R_a 7.92)and then decreases, following a non-linear function up to the experimental limit, sintering time being kept constant at the value of 1.5 hrs. Similar behavior has also observed on the surface roughness value with the variation of sintering time and sinter temperature, which is depicted in fig. 5.

keeping the compaction load fixed at 22.08 Ton. Surface roughness attains a minimum value at high sintering temperature (1176^0C) over entire range of sintering time (0.659hr.to 2.34 hrs). The observed variation of surface roughness as manifested in figure 5 is nonlinear in nature. A change in surface roughness of the P/M components against sintering time and compaction load for a fixed sintering temperature of 1050^0C has shown in fig. 6. It is evident from the figure 6, that the

surface roughness of the P/M components under study shows an initial increase in surface roughness and then gradual decrease with increase in compaction load. This trend is observed for almost the entire range of sintering time. On the contrary at low compaction load (14.65 Ton) surface roughness starts with a high value($R_{a\ \mu m}$ 7.08) at low sintering time(0.659 hr.) and changes very little with increase in sintering time.
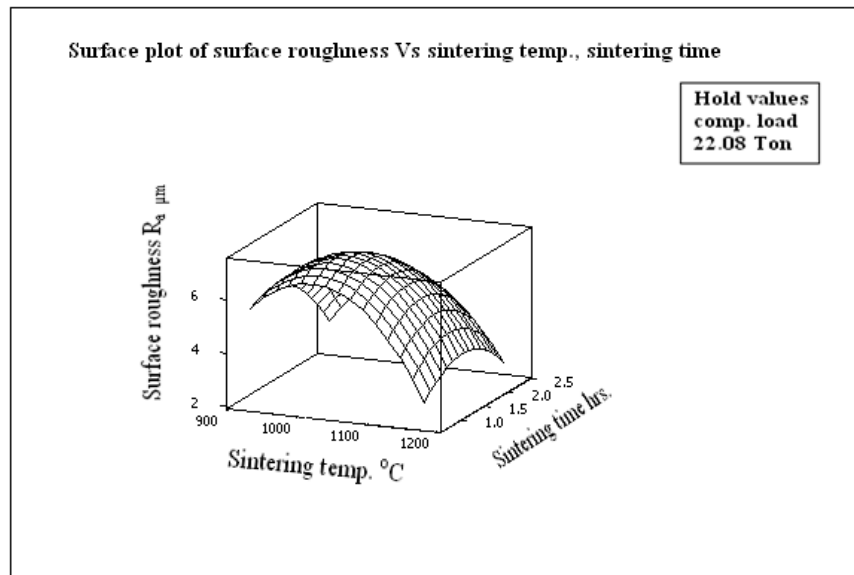
Fig.5 . Surface Plot of surface roughness $R_{a\ \mu m}$ (R1) vs Sintering time (X3) Sintering temperature(X2).
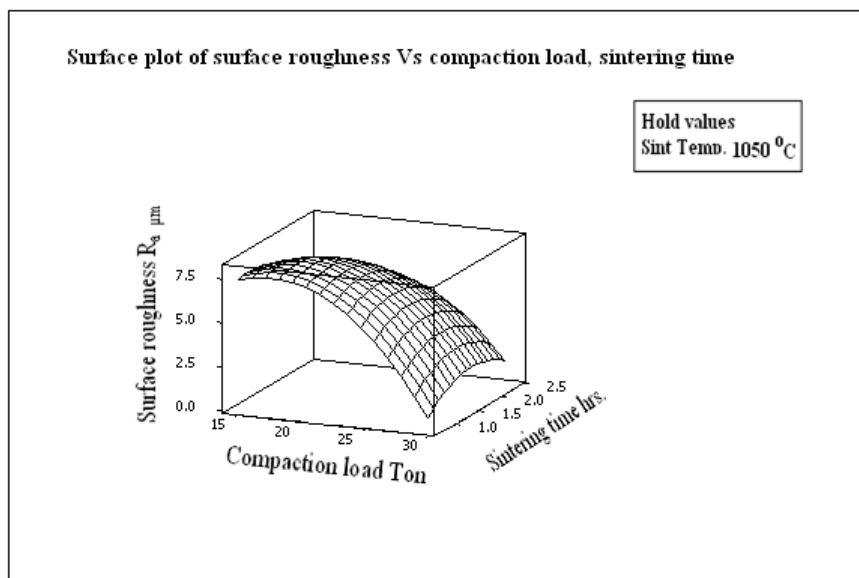
Fig. 6. Surface Plot of surface roughness $R_{a \ \mu m}$ (R1) vs Compaction load (X1) Sintering time (X3).

From the fig. 5, it is evident that sintering time has very little effect on surface roughness. Variation of surface roughness against sintering temperature and compaction load is presented in fig.4. The figure exhibits an increasing tendency is due to change in sintering temperature from 975°C to 1176°C and compaction load from 14.65 Ton to 29.50 Ton at a fixed sintering time of 1.5 hrs. Identical nature of variation has noted in simultaneous increase of compaction load and sintering time has illustrated in fig. 6. In this fig., Compaction load has altered between 14.65 Ton to 29.50 Ton and sintering time has been changed between 0.6 to 2.3 hrs at invariant sintering temperature of 1050°C. The response variable i.e. surface roughness under consideration shows a non linear nature when it is plotted against sintering temperature and sintering time at a fixed compaction load of 22.08 Ton. (fig. 5). In this case, the range of variation of the parameters is similar to that of previous two cases. It is worth mentioning, in all the cases the hold values are mean value of the range of variation corresponding to each variable.

3.1. Statistical analysis

Table 4

Analysis of Variance for Surface roughness (Ra)

Source	DF	Seg SS	Adj SS	Adj MS	F	P
Regression	9	57.3536	57.3536	6.37262	30.01	0.000
Linear	3	34.3304	7.9507	2.65023	12.48	0.001
Square	3	20.8885	20.8885	6.96284	32.79	0.000
Interaction	3	2.1346	2.1346	0.71155	3.35	0.064
Residual Error	10	2.1235	2.1235	0.21235		
Lack-of-Fit	5	1.6719	1.6719	0.33437	3.7	0.089
Pure Error	5	0.4516	0.4516	0.09032		
Total	19	59.4771				

A significance test was conducted to examine the effect of different process parameters and their inter-actions terms on the said response.

Table 5 shows the results of the significance test. The different terms used in Table 5 are as follows.

Table 5
Coefficients, standard errors, T statistics and p value for the response, Surface roughness (Ra).

Term	Coef	SE Coef	T	P
Constant	-129.949	27.8842	-4.660	0.001
X1	0.386	0.5963	0.648	0.532
X2	0.274	0.0471	5.821	0.000
X3	-4.222	5.0644	-0.834	0.424
X1*X1	-0.047	0.0062	-7.524	0.000
X2*X2	-0.000	0.0000	-6.884	0.000
X3*X3	-1.560	0.4855	-3.214	0.009
X1*X2	0.001	0.0005	2.187	0.054
X1*X3	0.150	0.0738	2.033	0.069
X2*X3	0.005	0.0043	1.066	0.311

R-Sq = 96.4%

From the results of ANOVA a mathematical model has been proposed for the evaluation of surface roughness of the powder metallurgy components. The proposed model is expressed as

$SRccd = -129.949 + 0.386 X_1 + 0.274 X_2 - 4.222 X_3 - 0.047 X_1^2 - 1.560 X_3^2 + 0.001 X_1 X_2 + 0.150 X_1 X_3 + 0.005 X_2 X_3$

Table 4 and Table 5 presents the ANOVA (Analysis of variances) and the significance test for the second order response surface equations, which quite clearly shows that second order response surface model fit well into the observed data. This is evident from the findings that co-efficient of determination (R^2) value is 96.4 %. Hence, it may be concluded that the prediction made by this developed model corroborates well with the experimentalobservations.
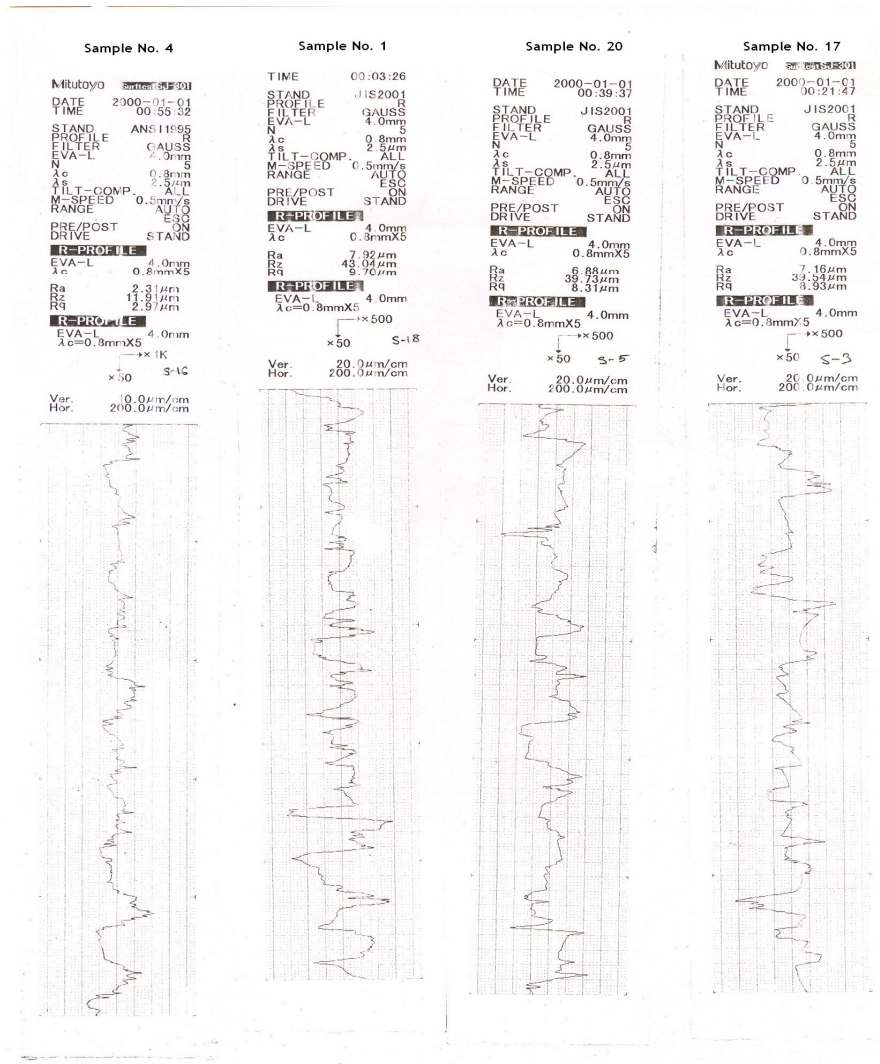
Fig.7. Comparison of surface roughness R_a of P/M samples

4. Conclusions

It is evident that the surface roughness of the P/M components under study shows an initial increase in surface roughness and then gradual decrease with increase in compaction load. This trend is observed for almost the entire range of sintering time. On the contrary at low compaction load (14.65 Ton) surface roughness starts with a high value ($R_{a\ \mu m}$ 7.08) at low sintering time (0.659 hr.) and changes

very little with increase in sintering time. It is also evident that sintering time has very insignificant effect on surface roughness. The ANOVA (Analysis of variances) for the second order response surface equations, which quite clearly shows that second order response surface model fit well into the observed data. This is evident from the findings that co-efficient of determination (R^2) value is 96.4 %. Hence, it may be

concluded that the prediction made by this developed model corroborates well with the experimental observations.

References:

[1]..ASTM B946 - 06Standard Test Method for Surface Finish of Powder Metallurgy (P/M) Products, 2010.

[2] The effect of machining on the surface integrity and fatigue life**.** A. Javidi, U. Rieger and W. Eichlseder, International Journal of Fatigue, Vol. 30, Issues 10-11, Pages 2050-2055, Oct. 2008.

[3] Some aspects in the surface integrity associated with turning of powder metallurgy compacts **,** G.T. Smith and M.J. Allsop**,** Wear, Vol.150, Issues 1-2, Pages -302, Oct 1991.

[4] K. S. Narasimhan, "Recent Advances in Ferrous Powder Metallurgy," *Advanced Performance Materials*, Vol. 3, No. 1, pp. 7-27, 1996.

[5] K. S. Naransimhan, "Sintering of Powder Mixtures and the Growth of Ferrous Powder Metallurgy," *Materials Chemistry and Physics*, Vol. 67, No. 1-3, pp. 56-65, 2001.

[6] H. Rutz, J. Khanuja and S. Kassam, "Single Compaction to Active High Density in Ferrous P/M Materials in Au-tomatic Applications," *PM2TEC*'96 *World Congress*, Washington, D.C.,1996.

[7] D. Chatterjee, B. Oraon, G. Sutradhar and P. K. Bose, "Prediction of Hardness for Sintered HSS Components Using Response Surface Method," *Journal of Materials Processing Technology*,Vol. 190, No. 1-3, pp. 123- 129, 2007.

[8] G. E. P. Boxes and N. R. Draper, "Emperical model building and response surfaces," Wiley, New York, 1987.

[9] D. C. Montgomery, "Design and Analysis of Experiments," John Wiley & Sons, New York, 1991.

X-ray Satellites Spectra in the Lα1 Region of 4d Transition Elements

Dr. Sameer Sinha, Vinay Kumar Pandey, Ajay Vikram Singh

Abstract- We have used Plasmon theory to explain   Energy Satellites and relative intensity of high energy X-ray satellites α3 , α4 & α5 with respect to Lα1 parent line in 4d Transition Metal (Zr , Nb , Mo , Ru , Rh) and estimated values are in agreement with the calculated values of Surendra Poornia and S.N.Soni.

Keywords - L α1 X-ray satellites, M vacancy Lαtransitions, X-Ray emission spectra.

——————————————   ◆   ——————————————

INTRODUCTION

IN the characteristic X-ray Spectra, Diagram as well as non Diagram lines are present. Those lines which fit in the conventional energy level diagram are called Diagram lines. & those lines which do not fit in the conventional energy level diagram are called non diagram lines. It is also known as "Satellites or Second order lines". Satellites are generally of weak intensity lines & are found close to more intense parent line. The satellites which are observed on higher energy side are called high energy satellites (HES) whereas those are observed on lower energy side are called lower energy satellites (LES). First Siegbahn & Stenstroem observed these satellites in the K-Spectra of element from Cr (24) to Ge (32) while coster theraeus & Richtmyer in the L-Spectra of element from Cu (29) to Sb (51) & Hajlmar, Hindberg & Hirsch in the M-Spectra of elements from Yb (70) to U (92). Several theories were proposed from time to time to explain the origin of these satellites. Out of these theories the plasmon theory is found to be the most suitable theory especially for those satellites.

Plasmon theory was first proposed by Bohm & pines which are extended by Housten, Ferrel, Noziers & Pines. According to this theory the low energy plasmon satellites are emitted when valence electron excites a plasmon during the annihilation of core hole conversely if Plasmon pre exists, its energy add up to the energy of diagram line.

 * Reader , Ganpat Sahai Post Graduate College , Sultanpur ,U.P. India
 ** Lecturer **,** Ganpat Sahai Post Graduate College , Sultanpur ,U.P. India
*** Associate Professor, Rajarshi Rananjay Sinh Institute of Management &Technology, Amethi , CSJ Nagar, U.P. , India
 E-Mail ID –
*** ajay_gspgcs@rediffmail.com
 ** vinay_gspg@rediffmail.com

The radiation less reorganization of electronic shell of an atom is known as Auger effect. Auger satellites have also been observed by Korbar and Mehlhorn [1] Haynes et at. [2] Edward and Rudd [3]. Theoretical explanation for K series Auger spectrum was given by Burhop and Asaad [4] using intermediate coupling. Later on more refined theory, using relativistic and configuration interaction has been used by Listengarter [5] and Asaad [6]

In Auger primary spectra, one can also observe secondary electron peaks close to the primary peaks are produced by incident electrons which have undergone well energy losses. The most common source of such energy loss in the excitation of collective plasma oscillations of the electrons in the solid. This gives rise to a series of plasma peaks of decreasing magnitude spaced by energy $\hbar\omega_p$ where ω_p is the frequency of plasma oscillation.

 Auger peaks are also broadened by small energy losses suffered by the escaping electrons. This gives rise to a satellite on the low energy of the Auger peak. Energy loss peaks have well defined energy with to primary energy.

The involvement of Plasmon oscillation in the X-ray emission or absorption spectra of solids has been widely studied during the last few decades and has been recognized that the electron –electron interaction has played an important role.

This Paper is devoted to Plasmon theory to explain the Energy Satellites and relative intensity of **high energy X-ray satellites with respect to Lα$_1$ parent line in 4d Transition Metal (Zr , Nb , Mo , Ru , Rh , Rd)** and estimated values are in agreement with the calculated values of Surendra Poornia and S.N.Soni.

According to Plasmon theory , if the valence electron , before filling the core vacancy , also excites a Plasmon ,then the energy $\hbar\omega_p$ needed for the excitation of Plasmon oscillation is taken from the transiting valence electron so that the emitted radiation will be derived off an energy $\hbar\omega_p$ and a low energy satellites will emitted whose sepration from the main X-ray line will correspond to $\hbar\omega_p$. On the other hand if the Plasmon pre exists , during the X-ray

emission process , then , on its decay it can give its energy to the transiting valence electron before it annihilates the core vacancy . Thus the energy of emitted X-ray photon will be higher than the main emission line and by an amount $\hbar\omega_p$ giving rise to high energy satellite .

Surendra Poornia and S.N.Soni have observed low and high energy satellite peaks in 4d Transition Metal **(Zr , Nb , Mo , Ru , Rh , Rd)** A close approximation of their tables shows that some satellites are at a distance of $\hbar\omega_p$ (Plasmon energy) from the main emission line . This observation forced us to think that these might be due to Plasmons emission and absorption .

MATHEMATICAL CALCULATION –

In order to confirm the involvement of Plasmon in the emission of X-ray satellites the relative intensity of single Plasmon satellites must be calculated . In this process first we deal with mathematical details of canonical transformation carried out over the model Hamiltonian of the system . Thus the energy separation ΔE of the low and high energy Plasmon satellite from the corresponding main line should be equal to the quantum of Plasmon energy $\hbar\omega_p$ which is given by [10]

$$\Delta E = \hbar\omega p = 28.8 \ \sqrt{\left(\frac{Z.\sigma}{w}\right)} ev \qquad 1$$

Where Z = No.of unpaired electrons taking part
 in plasma oscillation

σ = Specific gravity

ω = Molecular Weight

 This equation can be derived as given below .

From the classical consideration , we get the frequency of Plasmon oscillation as

$$\omega p = \left(\frac{4\pi n e^2}{m}\right)^{1/2} \qquad 2$$

 Hence the amount of energy given to Plasmon becomes

$$E_p = \hbar\omega_p = \hbar \ \left(\frac{4\pi n e^2}{m}\right)^{1/2}$$

 In this equation we can write

$$n = \frac{L\sigma Z}{w}$$

, Z and W are defined above and L is the Avogadro number .By putting the numerical value of constant , we get the Plasmon energy as

$$\Delta E = \hbar\omega p = 28.8 \ \sqrt{\left(\frac{Z.\sigma}{w}\right)} ev \qquad 3$$

Our calculated values of ΔE have been compared with the Scrocco's experimental value. And We have also calculated the relative intensity of plasmon satellites, which is different in different processes. If the excitation of plasmon occurs during the transport of the electron through the solid , it is known as extrinsic process of plasmon excitation. The plasmon can also be excited by another method known as intrinsic process. In this process, excitation of plasmon takes place simultaneously with creation of a hole. Bradshaw et al have further divided core hole excitation into two classes,

 1 - Where the number of slow electrons are conserved.

 2 - Where the number of slow electrons are not conserved

The Author has calculated relative intensity in both the cases with new modification in the light of Bradshaw [11] and Lengreth [12] work, which explains that not only intrinsic process but extrinsic process and their relative contribution may also contribute in relative intensities. The combined effect of intrinsic and extrinsic plasmon excitation intensity variation was also suggested by various Authors [12 , 15 , 16 , 17 18] as**:**

$$i = \frac{is}{Im} = a^n \sum_{m=0}^{n} \frac{\left(\frac{\beta}{\alpha}\right)^m}{m!} \qquad 4$$

The value of β is taken as $\beta = 0.12 r_s$ which is purely intrinsic, $r_s = (47.11/\hbar w_s)^{2/3}$ is dimensionless parameter and $\alpha = 0.47 r_s^{1/2}$ in the place of $\alpha = (1+l/L)^{-1}$ used by Pardee et. al.(13) . The equation (3) contains a series of terms. The first term of the equation is purely extrinsic, while second term is purely intrinsic. The other terms are containing the relative contributions of both extrinsic and intrinsic. The specialty of this formula is that each term alone or simultaneously with other terms is able to give the relative intensity. This formula also includes both the categories mentioned by Bradshaw and gives better results as compared than traditional methods for calculation of the relative intensity. Using the values of α, β and rs in equation (4)

Using the equation (4), the author has for the first time, calculated the relative intensity **of high energy X-ray satellites with respect to Lα_1 parent line in 4d Transition Metal (Zr , Nb , Mo , Ru , Rh , Rd)** metals. Our calculated and estimated values are in agreement with the calculated values of Surendra Poonia and S.N.Soni [14].

Reference

1. Korbar H. & Mehlhorn W.A. ; Phys. 191, (1966) 217.
2. Haynes S.K. & Velinsky, M & Velinsky L.J. ; Nucl. Phys. A99 (1967), 537.
3. Rudd M.E. & Edward & Volz, D.J. ; Phys Rev. 151, (1966), 28.
4. Asaad, W.N. & Burhop E.H.S. ; Proc. Phys. Soc., London 71, (1958), 369.
5. Listengarten, M.A. ; Bull Acad. Sci. U.S.S.R., Phys. Ser. 26 (1962), 182.
6. Asaad, W.N. ; Nucl. Phy. 66, (1965b), 494.

7. M.Scrocco in photoemission spectra of Pb.(II) halide; Phys. Rev. B25 (1982) 1535-1540 .

8. M.Scrocco , Satellites in X-ray Photo electron spectroscopy of insulator I 32 (1985) 1301-1306

9. M.Scrocco , Satellite in X-ray Photo electron spectroscopy of insulators II 32 (1985) 1307-1310

10. L.Marton , L.B.Lader and H. Mendlowitz; Adv. Electronic and Electro Physics; edited by L.M   arton Academic , New York 7 (1955) , 225

11. A. M. Bradshaw, Cederbaurn S.L,Domeke W. & Krause Jour. Phys C: Solid State   **Phys. 7, 4503, 1974**

12. D. C. Lengreth, Phys. Rev. Letter, 26, 1229, 1971

13. W. J. Pardee, G.D. Mahan, D. E. Eastman, R.A. Pollak, L. Ley, F.R. McFeely, S.P. Kowalczky and D.A. Shirely, Phys. Rev. B, 11, 3614, 1975.

14. Surendra poonia and S.N.Soni , Indian journal of pure and applied physics , vol.45, feb.2007 pp-119-126

15. S. K. Srivastava, Amar Bahadur, *Fizika A*, 17, 117, 2008

16. S. K. Srivastava Leonardo Journal of Sciences , 13 , 39 , 2008.

17. Sameer Sinha & Ajay Vikram Singh , International Journal of Scientific & Engineering Research, Volume 2, Issue 2, April -2011

18. Sameer Sinha & Ajay Vikram Singh , International Journal of Scientific & Engineering Research, Volume 2, Issue 2, February-2011

S.No.	NAME	SYMBOL	Z	W	Sp.Gravity	Exper. Value [14]	Author Value
1	ZIRCONIUM	Zr(40)	1	91.224	6.51	6.54	5.44
2	NIBOLIUM	Nb(41)	1	92.906	8.58	6.51	6.19
3	MOLYBDENUM	Mo(42)	1	95.94	10.28	6.54	6.67
4	RUTHENIUM	Ru(44)	1	101.07	12.45	6.85	7.15
5	RHODIUM	Rh(45)	1	102.91	12.41	7.36	7.07
6	PALLADIUM	Pd (46)	1	106.42	12.02	7.59	6.84

Energy sepration ΔE at $L\alpha_1$ satellite of 4-d Transistion element

Relative Intensity of $L\alpha_1$ satellite of 4-d Transition element

S.No.	Name of Metal	Surface Energy	Rs	Alpha	Beta	Author intensity	Exp. Intensity [14]	Symbol
1	ZIRCONIUM	5.44	4.22	0.97	0.51	0.78065	0.780012	$\beta+0.1+\beta^2/2\alpha$
2	NIBOLIUM	6.19	3.87	0.92	0.46	0.7780	0.770182	$\beta+0.1+\beta^2/2\alpha$
3	MOLYBDENUM	6.67	3.68	0.90	0.44	1.3530	1.368272	$2*(\beta+0.1+\beta^2/2\alpha)$
4	RUTHENIUM	7.15	3.51	0.88	0.42	0.7834	0.7434	$2*\beta - 0.1$
5	RHODIUM	7.07	3.54	0.88	0.42	0.8655	0.8498	$2*\beta$
6	PALLADIUM	6.84	3.77	0.91	0.45	0.8623	0.9044	$2*\beta$

A Pipelined Implementation for Full-Search Motion Estimation for Video Compression Using Variable Block-Size

Vasavi Sandya. G
M.TECH(VLSI Design)
SRM UNIVERSITY
CHENNAI-INDIA
sandhya.gangineni@gmail.com

K.SUGANTHI
ASST.PROF
SRM UNIVERSITY
CHENNAI-INDIA
mailtosuganthik@gmail.com

Abstract — **Variable block-size motion estimation (VBSME) has become an important technique in H.264/AVC to improve video quality. The hexagonal search method recently developed, an enhanced hexagonal search algorithm is proposed to further improve the performance in terms of reducing number of search points and distortion where a novel fast inner search is employed by exploiting the distortion information of the evaluated points. Our experimental results substantially justify the merits of the proposed algorithm**.

Index Terms — **Advanced video coding (AVC), Variable block size motion estimation (VBSME), H.264/AVC standard, VLSI architecture, fast search algorithm, Hexagonal search.**

I. INTRODUCTION

Block-matching motion estimation is vital to many motion compensated video coding techniques/standards, which is aimed to exploit the strong temporal redundancy between successive frames. However, the motion estimation is quite computational intensive and can consume up to 80% of the computational power of the encoder if the full search (FS) is used by exhaustively evaluating all possible candidate blocks within a search window. Therefore, fast search algorithms are highly desired to significantly speed up the process without sacrificing the distortion seriously. Many computationally efficient variants were developed, typically among which are three-step search (TSS), new three-step search (NTSS) [1], four-step search (4SS) [2], block-based gradient descent search (BBGDS) [3] and diamond search (DS) [4] [5] algorithms. In TSS, NTSS, 4SS and BBGDS algorithms, square-shaped search patterns of different sizes are employed. On the other hand, the DS algorithm adopts a diamond-shaped search pattern, which has demonstrated faster processing with similar distortion in comparison with NTSS and 4SS. This inspires us to investigate why diamond search pattern can yield speed improvement over some square-shaped search patterns and what is the mechanism behind. As a result, one may wonder whether there is any other pattern shape better than diamond for faster block motion estimation. In the following, we propose a hexagon-based search algorithm that can achieve substantial speed improvement over the diamond search algorithm with similar distortion performance.

2. HEXAGON BASED SEARCH ALGORITHM
2.1 Hexagonal Search Pattern

Ideally, a circle-shaped search pattern with a uniform distribution of a minimum number of search points is desirable to achieve the fastest search speed.
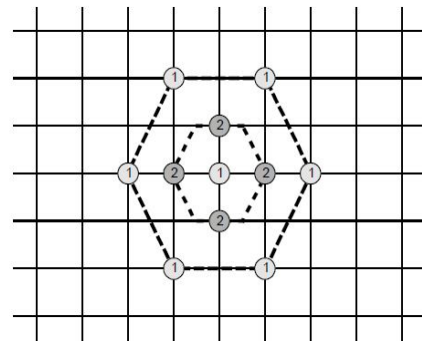
Fig. 1. Hexagon-based search (HEXBS): large (1) and small (2) hexagonal patterns.

Each search point can be equally utilized with maximum efficiency. Referring to the diamond search pattern [4] [5], we can see that the diamond shape is not approximate enough to a circle, which is just 90 degree rotation of a square. Consequently, a more circle approximated search pattern is expected in which a minimum number of search points are distributed uniformly. A hexagon based search pattern is devised which is depicted in Fig. 1. There are two different sizes of hexagonal search patterns, of which the larger one consists of 7 check points marked as 1 with the center surrounded by 6 endpoints of the hexagon with the two edge points (up and down) being excluded. Like the shrunk diamond pattern, a smaller shrunk hexagonal pattern covering 4 check points marked as 2 is inside the large one, which is finally applied in the focused inn we can see the 6 endpoints are approximately uniformly distributed around the center with similar

distances to the center. Note that the hexagonal search pattern also contains 2 fewer check points than the 9-point diamond search pattern. In the search process, the large hexagon-based search pattern keeps advancing with the center moving to any of the six endpoints. Whichever endpoint the center of the large search pattern moves to, there are always three new endpoints emerging and the other 3 endpoints being overlapped. Recall that in diamond search, 3 or 5 new points appear each time the diamond moves in different directions, in average 4 new points being evaluated for each move.

2.2 Algorithm Development

With the designed hexagonal pattern, we develop the search procedure as follows. In the first step, the large hexagonal pattern with 7 check points is used for search. If the optimum is found at the center, we switch to use the shrunk hexagonal pattern including 4 check points for the focused inner search. Otherwise, the search continues around the point with minimum block distortion (MBD) using the same large hexagonal pattern. Note that while the large hexagonal pattern moves along the direction of decreasing distortion, only 3 new non-overlapped check points will be evaluated as candidates each time. Fig. 2 shows an example of the search path strategy leading to the motion vector (+4, -4), where 20 (=7+3+3+3+4) search points are evaluated in 5 steps sequentially. The proposed hexagon-based search (HEXBS) algorithm can be summarized in the following detailed steps.

Step i) **Starting**: The large hexagon with 7 check points is centered at (0,0), the center of a predefined search window in the motion field. If the MBD point is found to be at the center of the hexagon, proceed to Step iii) (Ending); otherwise, proceed to Step ii) (Searching).

Step ii) **Searching**: With the MBD point in the previous search step as the center, a new large hexagon is formed. Three new candidate points are checked, and the MBD point is again identified. If the MBD point is still the center point of the newly formed hexagon, then go to Step iii) (Ending); otherwise, repeat this step continuously.

Step iii) **Ending**: Switch the search pattern from the large size of hexagon to the small size of hexagon. The four points covered by the small hexagon are evaluated to compare with the current MBD point. The new MBD point is the final solution of motion vector.

From the above procedure, it can be easily derived that the total number of search points per block is

$$N_{HEXBS}(m_x, m_y) = 7 + 3 \times n + 4 \qquad (1)$$

Where (m_x, m_y) is the final motion vector found, and n is the number of execution of Step ii).

2.3 Analysis of the HEXBS Algorithm

In this subsection, we will examine the proposed HEXBS algorithm as compared with the diamond search (DS) algorithm in term of number of points evaluated to find same motion vectors. For block motion estimation, computational complexity can be measured by number of search points required for each motion vector estimation. In contrast with the HEXBS algorithm by which the number of search points used is indicated in Equation (1), the diamond search method requires the following number of search points per block

$$N_{DS}(m_x, m_y) = 9 + M \times \tilde{n} + 4 \qquad (2)$$

where M is either 5 or 3 depending on the search direction and $\tilde{n}$ depends on the search distance. The $\tilde{n}$ in Equation (2) is
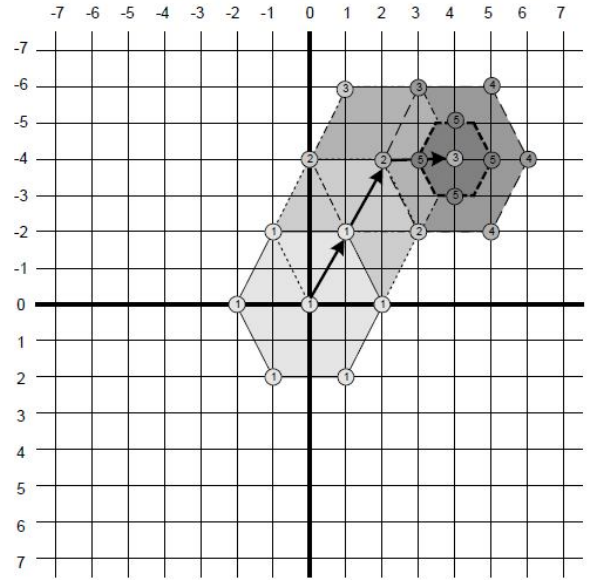
Fig. 2. Search path example locating the motion vector (+4, -4) by HEXBS. Note that a small HEXBS pattern is applied in the final step after the best candidate search point at step 3 remains the best at step 4. Totally 20 (=7+3+3+3+4) search points are evaluated in 5 steps.

always greater than or equal to the n in Equation (1)

By analyzing the minimum possible number of search points of HEXBS and DS, we show in Fig. 3 the number of search points saved by HEXBS as compared with DS for each motion vector location within ±4 region. From the figure, we can see that for the stationary and quasi-stationary motion vectors the HEXBS

algorithm has 2 fewer search points, while more search points can be saved for locating medium to large motion vectors beyond the ±1 region. For example, 7 search points (block matches) can be saved by HEXBS for locating motion

vectors (±1, ±3), (±4, ±2), (±2, ±4) and (±2, ±5). In short, the new HEXBS scheme can find any motion vector in motion field with fewer search points than the DS algorithm. Generally speaking, the larger the motion vector, the more search points the HEXBS algorithm saves, which is justified by the following experimental results.
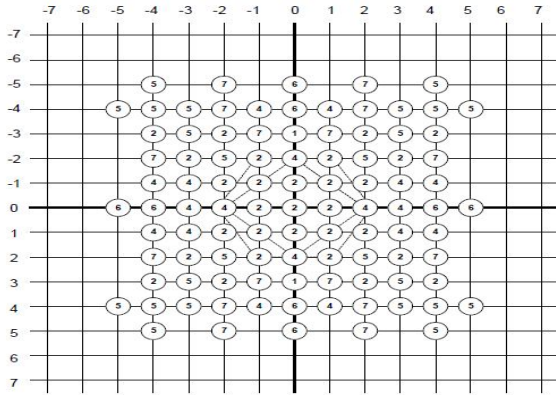
Fig. 3. Number of search points saved by HEXBS as compared to DS for each motion vector location.

3. Hardware description

Fig. 4 shows the architecture of a configuration with 6 execution units: four integer pel execution units (64x4 bits data path), one fractional-pel execution unit (64-bit data path) and one interpolation unit.
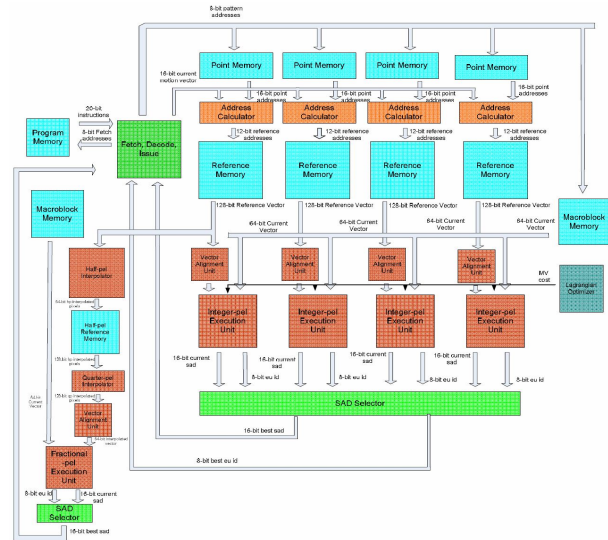
Fig. 4. Hardware Architecture

Motion estimation algorithms have not been standardized and over the years many fast motion estimation algorithms have been proposed by both industry and academia. Well-known fast motion estimation algorithms include logarithmic search, three-step search, diamond search, hexagon search, etc. along with more complex methods such as PMVFAST and UMH. These algorithms work by searching only a subset of all candidate positions for the best match, using a dynamic search pattern iteratively with each local winner set as the starting point for the following iteration. The same principles of calculating motion vectors by block-matching the current frame with other reference frames in the video sequence are shared by MPEG-2, MPEG-4, VC-1 and H.264, although the more recent standards have more sophisticated ways of achieving this.

The application specific instructions generated by the compiler specify the parallelism available at the search point level, and the resulting level of performance depends on the number of execution units present. Similar to VLIW processors, parallelism is extracted at compile time using a simple fetch, decode and issue unit which deals with one instruction at a time, and just like superscalar processors, different implementations of the microarchitecture remain binary compatible so that hardware upgrades will not need a software recompile. Processor is fully synchronous, and its top-level interface is shown in Fig 5.
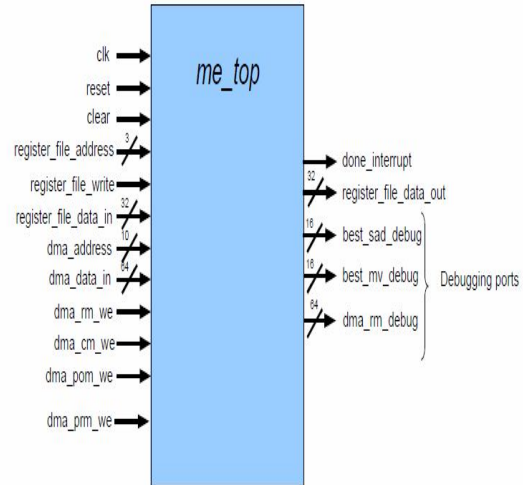
Fig. 5.Top Level Interface

4. CONCLUSION

In this paper a novel fast algorithm using hexagon based search pattern in block motion estimation, which demonstrates significant speedup gain over the diamond-based search and other fast search methods while maintaining similar distortion performance. The proposed HEXBS consistently has a faster search performance than DS regardless of no-, small-, medium- or large-motion. In other words, the new hexagon based search scheme may find any motion vector in motion field with fewer search points than the DS algorithm. Generally speaking, the larger the motion vector, the more significant the speedup gain for the new method will be. The experimental results have verified the statement and convincingly demonstrated the superiority of the proposed HEXBS to the other fast methods in terms of using the smallest number of search points with marginal increase in distortion.

REFERENCES

[1] R. Li, B. Zeng, and M. L. Liou, "A new three-step search algorithm for block motion estimation," IEEE Transactions on Circuits & Systems for Video Technology, vol. 4, pp. 438-442, Aug. 1994.

[2] L. M. Po and W. C. Ma, "A novel four-step search algorithm for fast block motion estimation," IEEE Transactions on Circuits & Systems for Video Technology, vol. 6, pp. 313-317, June 1996.

[3] L. K. Liu and E. Feig, "A block-based gradient descent search algorithm for block motion estimation in video coding," IEEE Transactions on Circuits & Systems for Video Technology, vol. 6, pp. 419-423, Aug. 1996.

[4] Shan Zhu, Kai-Kuang Ma, "A new diamond search algorithm for fast block-matching motion estimation," IEEE Transactions on Image Processing, vol. 9, no. 2, pp. 287-290, Feb. 2000.

[5] Jo Yew Tham, Surendra Ranganath, Maitreya Ranganath, and Ashraf Ali Kassim, "A novel unrestricted center-biased diamond search algorithm for block motion estimation," IEEE Transactions on Circuits & Systems for Video Technology, vol. 8, no. 4, pp. 369-377, Aug. 1998.

[6] T. Wiegand, G. J. Sullivan, "Overview of the H.264/AVC video coding standard", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, n. 7, July 2003.

[7] S. Zhu and KK Ma, "A New Diamond Search Algorithm for Fast Block-Matching Motion Estimation", *IEEE Transactions on Image Processing,* vol.9, no.2, February 2000.

[8] A. M. Tourapis, O. C. Au and M. L. Liou, "Highly Efficient Predictive Zonal Algorithms for Fast Block.- Matching Motion Estimation", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, n. 10, pp 934-947, October 2002.

[9] H-Y C. Tourapis, A. M. Tourapis, "Fast Motion Estimation within the JVT codec", *Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG 5th meeting*, Geneva Switzerland, 09-17 October 2002.

[10] C. Zhu, X. Lin and L.P Chau, "Hexagon-Based Search Pattern for Fast Block Motion Estimation", *IEEE Transactions on Circuits and Systems for Video Technology,* vol.12, no.5, pp 349-355, May 2002.

[11] P. De Pascalis, "Stima del Moto per Codificatori H.264", *Laurea Degree Thesis, Dept. of Information Engineering, University of Padova*, Italy, October 2003.

Analysis of Software Cost Estimation using COCOMO II

T.N.Sharma

Abstract - COCOMO II is an objective cost model for planning and executing software projects. It is an important ingredient for managing software projects or software lines of business. A cost model provides a framework for communicating business decisions among the stakeholders of a software effort. COCOMO II supports contract negotiations, process improvement analysis, tool purchases, architecture changes, component make/buy tradeoffs and several other return-on-investment decisions with a credible basis of estimate. COCOMO II incorporates several field-tested improvements to both broaden its applicability and improve its estimating accuracy for modern software development approaches. COCOMO II includes two underlying information models. The first is a framework for describing a software project, including models for process, culture, stakeholders, methods, tools and the size/complexity of the software product. The second is an experience base that can be used to estimate the likely includes significant updates to COCOMO to improve its applicability to modern processes, methods, tools and technologies. It also includes a much larger, more pertinent database of modern precedents and improves the adaptability of the model so it can be optimized across a broad spectrum of domains and project circumstances. This paper presents cost estimation of various projects using COCOMO II. This article also presents statistical analysis for relevance of base COCOMO II model for effort estimation in present scenario.

Index Terms— Software Cost Estimation, COCOMO II, Scale Factors, Cost Drivers, Case Studies.

———————————————— ◆ ————————————————

1 INTRODUCTION

Software cost estimation is a prediction of the cost of the resources that will be required to complete all of the work of the software project.

Software has a bad reputation about cost estimation. Large software projects have tended to have a very high frequency of schedule overruns, cost overruns, quality problems, and outright cancellations. Instead of it bad reputation, it is important to note that some large software projects are finished on time, stay within their budgets, and operate successfully when deployed.

Currently a new generation of software processes and products is changing the way organizations develop software. The new approaches – evolutionary, risk driven and collaborative software processes; fourth generation languages and application generators; commercial off the

shelf (COTS) and reuse driven software approaches; fast track software development approaches; software process maturity initiatives – lead to significant benefit in terms of improved software quality and reduced software cost, risk and cycle time.

COCOMO II model tailored to these new forms of software development, including rationales for the model decisions. The major new modeling capabilities of CO-COMO II are a tailorable family of software sizing models, involving Object Points, Function Points, and Source Lines of Code; nonlinear models for software reuse and reengineering; an exponent-driver approach for modeling relative software diseconomies of scale; and several additions, deletions, and updates to previous COCOMO effort-multiplier cost drivers. This model is serving as a framework for an extensive current data collection and analysis effort to further refine and calibrate the model's estimation capabilities.

2. ESTIMATION EQUATIONS

_____

T.N.Sharma is undergoing his research work from Department of Statistics, University of Rajashthan, Jaipur (India).Mob.:9414248794. Email: tnsharma@rediffmail.com

In COCOMO II, the amount of effort in person-months, PM, is estimated by the formula:

The amount of calendar time, TDEV, it will take to develop the product is estimated by

TDEVNS = C x (PMNS)F

$$\text{where } F = D + 0.2 \times 0.01 \times \sum_{j=1}^{5} SFj$$

$$= D + 0.2 \times (E - B)$$

In COCOMO-II effort is expressed as person month(PM). COCOMO II treats the number of person-hours per month, PH/PM, as an adjustable factor with a nominal value of 152 hours/PM.

- The value of n is 16 for the Post-Architecture model effort multipliers, Emi, and 6 for the Early Design model, the number of SFi stands for exponential scale factors.
- The values of A, B, C, D, SF1 …, and SF5 for the Early Design model are the same as those for the Post-Architecture model.

Baseline Effort Constants:
    A = 2.94;   B = 0.91

$$PM = A \times Size^E \times \prod_{i=1}^{n} EM$$

$$\text{where } E = B + 0.01 \times \sum_{j=1}^{5} SF_j$$

Baseline Schedule Constants:
    C = 3.67;   D = 0.28

2.1 SCALE FACTORS

The application size is exponent is aggregated of five scale factors that describe relative economies or diseconomies of scale that are encountered for software projects of dissimilar magnitude.

- Precedentedness(PREC)
- Development Flexibility (FLEX)
- Architecture / Risk Resolution (RESL)
- Team Cohesion (TEAM)
- Process Maturity (PMAT)

2.2 COST DRIVERS/ EFFORT MULTIPLIERS

Cost drivers are characteristics of software development that influence effort in carrying out a certain project. Unlike the scale factors cost drivers are selected based on the rationale that they have a linear affect on effort. There are 17 effort multipliers that are utilized in the COCOMO II model to regulate the development effort.

- Required Software Reliability (RELY)
- Data Base Size (DATA)
- Developed for Reusability (RUSE)

- Documentation Match to Life-Cycle Needs (DOCU)
- Execution Time Constraint (TIME)
- Main Storage Constraint (STOR)
- Platform Volatility (PVOL)
- Analyst Capability (ACAP)
- Programmer Capability (PCAP)
- Personnel Continuity (PCON)
- Applications Experience (APEX)
- Platform Experience (PLEX)
- Language and Tool Experience (LTEX)
- Use of Software Tools (TOOL)
- Multisite Development (SITE)
- Required Development Schedule (SCED)

3. CASE STUDIES
PROJECT 1 ADVOCATE'S DESKTOP
As first case study we have taken a project in consideration which was developed for an very famous advocate

firm. The project takes care billing and case information of the firm.

After completion of project we calculated the efforts (Person-Month) using COCOMO II and got the actual time taken to develop the project. Total line of code of C# is 7187 i.e. 7.1 KLOC.

	VL	L	N	H	VH	XH	Our Value
PMAT	7.8	6.24	4.68	3.12	1.56	0	4.68

Here

$$\sum_{j=1}^{5} SFj = 18.97$$

$$E = B + 0.01 \times \sum_{j=1}^{5} SFj = 1.097$$

Scale Factors (SFj)	Very Low	Low	Nominal	High	Very High	Extra High	Our Value
PREC	thoroughly unprecedented	Largely unprecedented	somewhat unprecedented	generally familiar	largely familiar	thoroughly familiar	3.72
	6.2	4.96	3.72	2.48	1.24	0	
FLEX	rigorous	occasional relaxation	some relaxation	general conformity	some conformity	general goals	3.04
		4.05	3.04	2.03	1.01	0	
	5.07						
RESL	little (20%)	some (40%)	often (60%)	generally (75%)	mostly (90%)	full (100%)	4.24
	7.07	5.62	4.24	2.83	1.41	0	
TEAM	very difficult interactions	some difficult interactions	basically cooperative interactions	largely cooperative	highly cooperative	seamless interactions	3.29
			3.29				
	5.48	4.38		2.19	1.1	0	

Baseline Effort Constants: **A** = 2.94; **B** = 0.91
Baseline Schedule Constants: **C** = 3.67; **D** = 0.28

Driver	Symbol	VL	L	N	H	VH	XH	Our Value
RELY	EM_1	0.82	0.92	1.00	1.10	1.26		.82
DATA	EM_2		0.90	1.00	1.14	1.28		0.9
CPLX	EM_3	0.73	0.87	1.00	1.17	1.34	1.74	.87
RUSE	EM_4		0.95	1.00	1.07	1.15	1.24	1.0
DOCU	EM_5	0.81	0.91	1.00	1.11	1.23		1.0
TIME	EM_6			1.00	1.11	1.29	1.63	1.0
STOR	EM_7			1.00	1.05	1.17	1.46	1.0
PVOL	EM_8		0.87	1.00	1.15	1.30		0.87
ACAP	EM_9	1.42	1.19	1.00	0.85	0.71		0.85
PCAP	EM_{10}	1.34	1.15	1.00	0.88	0.76		0.88
PCON	EM11	1.29	1.12	1.00	0.90	0.81		0.90
APEX	EM12	1.22	1.10	1.00	0.88	0.81		0.88
PLEX	EM13	1.19	1.09	1.00	0.91	0.85		0.85
LTEX	EM14	1.20	1.09	1.00	0.91	0.84		0.91
TOOL	EM15	1.17	1.09	1.00	0.90	0.78		0.90
SITE	EM16	1.22	1.09	1.00	0.93	0.86	0.80	0.80
SCED	EM17	1.43	1.14	1.00	1.00	1.00		1.0

$$\prod_{i=1}^{n} EM = 0.184295$$

$$PM = A \times SizeE \times \prod_{i=1}^{n} EM$$

$$\text{where } E = B + 0.01 \times \sum_{j=1}^{5} SFj$$

Applying the values on formula :
Here we have

A = 2.94
Size = 7.1

$$E = 1.097$$
$$B = .91$$
$$\prod EM = 0.184295$$

$$PM = 4.67724$$

Actual time taken for this project is 4 Months.
We have applied above COCOMO II formula on 4 software projects which produces us the results as per following table. Also we have taken the actual person month in the table.

S.N.	Name of Project	Technology	KLOC	$\prod_{i=1}^{n} EM$	$\sum_{j=1}^{5} SF_j$	Estimated PM using COCOMO II	Actual PM
1	Advocate's Desktop	C#.Net	7.1	0.184	18.97	4.6	4
2	Online TrueLogic	DotNet	15.01	0.204	18.97	11.83	9.5
3	Online Project Management	Java	7.8	.204	18.97	5.76	4
4	Unit Converter	Android	1.6	.1546	18.97	.76 (22 Days)	.53 (16 Days)

4. CONCLUSION

This brief article shows how to make cost estimates using COCOMO II for a sample project, and outlines basic steps, terms, and tools used. Obviously, ad hoc estimates are prone to error. COCOMO II make it easy for you to clarify not only an expected project cost and duration, but also prompt you to verify all basic sides of a software project by providing clear, compact, and concise terms, methodology, which are tested on a wide range of real-life projects and thus reduce essentially project risks and provide reasonable grounds for communication with a project stockholder. Paper presents difference in between estimation by COCOMO II and actual time taken by the project.

5. REFERENCE:

- Arlene F. Minkiewicz, The Evolution of Software Size: A Search for Value, CrossTalk, The Journal of Defense Software Engineering, April 2009 p.p. 23-26

- Basavaraj, M.J. and Shet, K.C. (2008). Empirical Validation of Software Development effort multiplier of Intermediate COCOMO Model. Journal of Software, vol.3, No. 5 pp. 65-71.

- Boehm Barry (2000). Software Cost Estimation using COCOMO. Prentice Hall, US.

- Barry Boehm, Ricardo Valerdi, COCOMO Suite Methodology and Evolution CROSSTALK The Journal of Defense Software April 2005, p.p. 20-25

- Coe, David j. Improving Consistency of Use Case Points Estimates. CrossTalk, The Journal of Defense Software Engineering, March 2008 pp 8-12

- Demirors, O.; Gencel, C, Conceptual Association of Functional Size Measurement Methods , IEEE Society, Volume 26, Issue 3, May-June 2009 p.p. 71 – 78

- Evans Michael, Alex Abela, and Tom Beltz(2002). Seven Characteristics of Dysfunctional Software Projects. The Journal of Defense Software Engineering.

- Ye Yang, Zhihao Chen, Ricardo Valerdi, Barry Boehm, "Effect of Schedule Compression on Project Effort", *CSSE Tech Report, 2005*

Power Infrastructure Monitoring System on Embedded Web

Mr Ajay K. Kakde, Mr Ajay P. Thakare

Abstract— PIMS NOC system is located centrally at the tower infrastructure service provider premises to monitor the various RMSs. The PIMS NOC is designed for modularity and scalability and comprises of a Communication Server, a Data Server and various User Terminals. Each Communication Server is connected to the GPRS/CDMA network service provider via a leased line. The bandwidth requirements of the leased line would be based on the total number of PIMS agents being monitored. PMS are received and processed by the Communication Server. The processed data is then passed onto the Data Server which primarily manages the tasks of storage and management of the PMS data. Various filtering options are provided to 'slice the data' as per required views. The filtered data serves as input for report generation. Various types of reports can be generated based on user requirements.

Index Terms— PIMS, PIMS Agent, NOC, Data Server, Embedded Web, GPRS, socket server.

————————————————— ◆ —————————————————

1 INTRODUCTION

Each cell-site tower is equipped with a Power Management System (PMS) which is a critical component of the cell-site tower infrastructure. It's primary function is to ensure uninterrupted power supply to the various equipments present at the cell-site tower. Due to the criticality associated with the proper operation of the PMS, it is required that any malfunction in its operation is immediately reported to a centralized NOC as well as service personnel to initiate remedial measures.

The PIMS NOC system is located centrally at the tower infrastructure service provider premises to monitor the various PMSs. The PIMS NOC is designed for modularity and scalability and comprises of a Communication Server, a Data Server and various User Terminals. Each Communication Server is connected to the GPRS/CDMA network service provider via a leased line. The bandwidth requirements of the leased line would be based on the total number of PIMS agents being monitored / controlled.

Asynchronous events / alarms that occur at the PMS are received and processed by The Communication Server. The processed data is then passed onto the Data Server for storing in the database. In the case of alarms, appropriate visual alarms (based on criticality) are dis-

played on User terminals and pre-configured actions (like sending SMS to pre-configured numbers) are initiated. Periodic poll of the PIMS agents are carried out to obtain current health and status information.

The following diagram illustrates the system:

————————————————
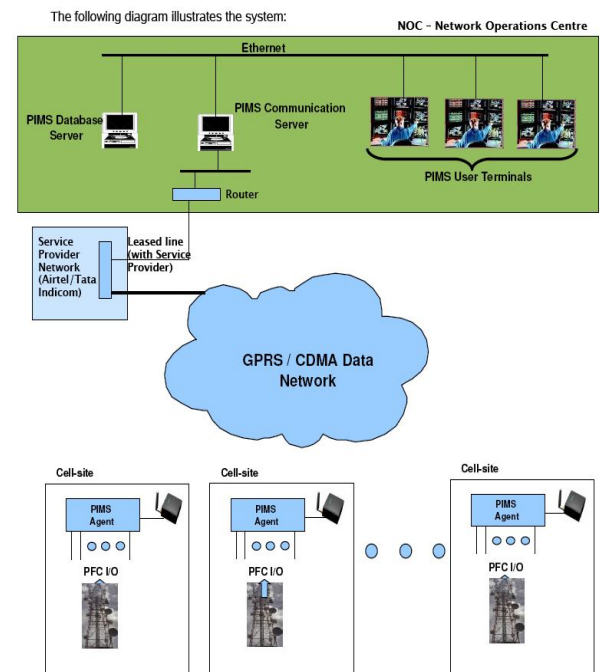
- *Mr Ajay K. Kakde is currently pursuing masters degree program in digital electronics in Amravati University, India, (ISTE Member LM-70512) E-mail Id - ajay.kakde@gmail.com*
- *Mr Ajay P. Thakare is currently pursuing Phd program in Nagpur University, India, (IEEE Member (USA) -41492816) IETE Fellow (F-122123) E-mail Id - apthakare40@gmail.com*
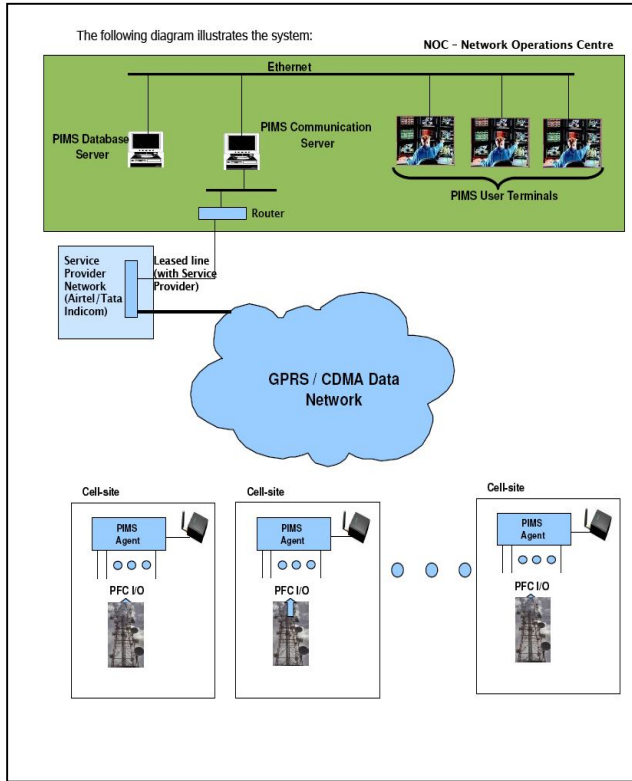
3 SYSTEM ARCHITECTURE

The embedded system hardware and software depends on the product type, analysis required and the communication port available with the product or not. The user has to specify what is normal and what is abnormal. The timely report even if the product is functioning normal. Service call alerts etc. At the centre a server   with suitable software to handle internet and a huge database is used and a provision for sending EMAILS on a given address as per the product is sent for immediate action so that no operator is required at the server centre.

A robust communication framework for communication between the PIMS and its agents is put in place. The framework is based on GPRS or CDMA data connectivity between the PIMS communication server and the PIMS agents with the entire communication being TCP/IP based. An application based protocol framework is present for message exchange between the PIMS communication server and the PIMS agents. Messages are defined for event reporting, health & status information, and control operations.

The PIMS agents are devices that co-habit with the PMS system at each cell tower location. They are connected to the PMS system for status / health monitoring. The following diagram depicts the structure of the PIMS agent:
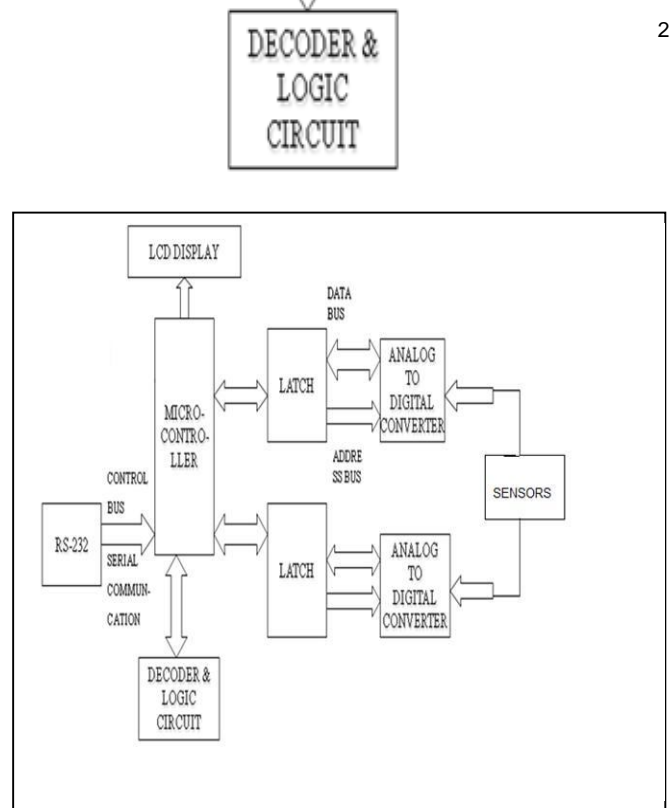
FIGURE: - 2. BLOCK DIAGRAM OF PIMS AGENT

The PIMS agents are devices that co-habit with the PMS system at each cell tower location. They are connected to the PMS system for status / health monitoring. The following diagram depicts the structure of the PIMS agent. We monitor a given product by using the best (reliable) and cheap mode of communication from the site to the central monitoring point in case occurrence of any abnormalities in the product or after a period specified by the user of the system. Every site is added with a client unit which collects data from the product and identifies the abnormalities and communicates with the centre via GPRS

2.link on GSM/GPRS MODEM ,the centre is having a internet connection with static IP and number of such clients can communicate with the centre as large as 2000

4. IMPLEMENTATION

Sequence of operation

- Analog inputs from the sensors are given to the ADC.
- The digital inputs and the ADC output is given to the microcontroller.
- In the microcontroller these inputs are processed using logic and digital circuits.
- The processed output is given to the computer using RS232 and HyperTerminal functionality is used to observe the output.

SIGNAL CONDITIONING
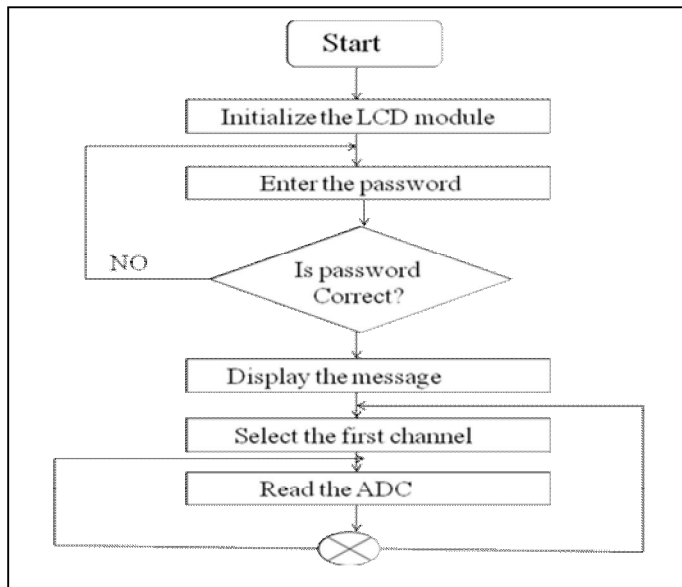
Flowchart 1 -Initialization of system
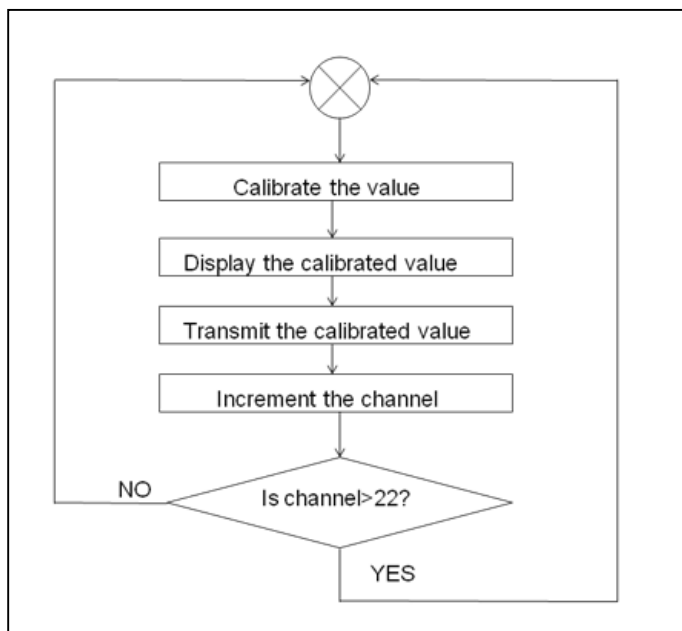
Flowchart 2 – Reading data from all channels

PIMS has 22 analog inputs and 16 digital inputs to which various sensors are used.Once the system is initialized PIMS reads data from all inputs. All the 22 analog and 16 digital channels are scanned and the data is sent to the processor. The Processor sends this data to monitoring system. It reads the data from input and sends it to monitoring system continuously

DEVELOPMENT OF THE WEB BASED SYSTEM

The embedded remote monitoring system completes the data Collection in the embedded platform and provides the data to remote host through the TCP/IP protocol from Web server. It creates condition to realize unattended management through providing Web-based graphical management interface for the Internet or LAN users

It eliminated the special client management software and realized unified management of various equipments in the network. Through existing public communications networks, without geographical restrictions, using a standard Web browser, Users can directly access to the Web server in embedded devices and performance remote monitoring, diagnosis and maintenance of all nodes on the network Embedded system is a kind of special computer system which has limited resources and functions. To implement Web server in embedded system is characteristic of itself. Considering the need of large dynamic data exchange during equipment monitoring and controlling is implemented through VB and ASP environment

THE KEY FEATURES

1. System based on distributed architecture (for scalability)
2. Instantaneous notification of fault / alarm conditions to the PIMS central server.
3. Communication framework based on GPRS/CDMA data network – with the underlying communication being TCP/IP based connectivity.
4. Modular scalability – The PIMS is designed to ensure that that the system is scalable in a modular fashion.
5. Logging the occurrence and clearing of all fault / alarm conditions.
6. Visual overview of the status and health of all PMS at the PIMS central system.
7. Remote Control functions – the system can be extended to carry out control functions from the NOC on a selected PMS

RESULTS

The microcontroller sent processed data to computer using RS 232 and data is monitored on HyperTerminal window. Here date, time and analog and digital output is displayed.

ADVANTAGES

- ➢ Monitoring system is developed to supply practical needs, is suitable for supporting maintenance and other services and for prevention of failures. It provides early warning of the problems.
- ➢ It reduces the manual work done in the cell sites. Hence making monitoring easy.
- ➢ PIMS does not require complicated solutions and high investment. Hence its maintenance is cheap.
- ➢ Ability to Monitor and Control remotely
- ➢ Efficient use of available Internet Network
- ➢ Real time monitoring and controlling
- ➢ Cost effective

APPLICATIONS

1. Home protection systems
2. Manufacturing
3. Health Care
4. Military
5. Bio-Technology

FUTURE SCOPE

1. Can be expanded into GSM, Satellite and Wi-Fi networks
2. Wireless sensors can be used
3. Implementation using Wi-Fi Ethernet within Organisation
4. Enabling text messages with desired parameters

TOOLS USED

Hardware specification

- ➢ PC, Microchip programmer, debugger, CRO, Multimeter.
- ➢ Microchip – PIC18F4520
- ➢ High Performance, Enhanced FLASH Microcontrollers with 10-Bit A/D
- ➢ On-Chip Program Memory: FLASH (bytes):- 32K
- ➢ # Single Word Instructions:- 16384
- ➢ On-Chip RAM (bytes):- 1536
- ➢ Data EEPROM:- 256
- ➢ Multiplexer – CD4051 BM 8:1 MUX for analog inputs.
- ➢ Serial communication: - RS 232.

Software specifications

- ➢ Compiler for 'C' programming.
- ➢ MS Hyperterminal.
- ➢ Internet data logger
- ➢ Socket Server etc.

REFERENCES

1. Yong-tao ZHOU, Xiao-hu CHEN, Xu-ping WANG, Chun-jiang YAO, 2008″: Design of Equipment Remote Monitoring System Based on Embedded Web″, IEEE, pp 73-78
2. Ali Ziya Alkar, 2009″ An Internet Based Interactive Embedded Data- Acquisition System for Real-Time Applications″, IEEE, Vol .58.No.3, pp.522- 529.
3. D. Comer, *Internetworking with TCP/IP Volume I*(Prentice-Hall, 1995)
4. Seifedine Kadry, Khaled Smaili, 2008″ Control System for Internet Bandwidth Based on Java Technology,″ Journal of Theoretical and Applied information Technology.
5. Mo Guan,Wei,Ying Bao, 2008 ″ A Monitoring System Based on Embedded Internet Technology for Embedded Devices″, IEEE ,Vol.pp.5-8

Hydrodynamic Free Convection Flow of A Rotating Visco-elastic Fluid Past An Isothermal Vertical Porous Plate With Mass Transfer

S. Biswal, G.S. Ray, A. Mishra

Abstract - The effect of free convection and mass transfer in the unsteady flow of an incompressible electrically conducting visco-elastic past an isothermal vertical porous plate with constant suction normal to the plate has been studied. The effects of permeability parameter (Kp) of the porous medium, rotation parameter (R), Grashof number for heat transfer (Gr), Grashof number for mass transfer (Gm), frequency parameter (ω) and the heat source parameter (α0) on the transient primary and secondary velocity field, temperature field and the rate of heat transfer have been investigated with the help of graphs and tables.

Key words - Hydrodynamic flow, mass transfer, rotating fluid, porous medium, isothermal plate.

———————————————— ◆ ————————————————

1. INTRODUCTION

THE problem of hydrodynamic free convection flow of a rotating viscoelastic fluid has received a considerable attention of many researchers because of its applications in cosmical and geophysical science. These problems are of general interest in the field of atmospheric and oceanic circulations, nuclear reactors, power transformers and in the field of scientific and industrial research. Permeable porous plates are used in the filtration process and also for a heated body to keep its temperature constant and to make the heat insulation of the surface more effective.

Several authors have discussed the flow of a viscous fluid in a rotating system in the presence and absence of magnetic field. Reptis and Singh [1] have reported the effect of rotation on the free convection MHD flow past an accelerated vertical plate. Singh[2] has studied the unsteady free convection flow of a viscous liquid through a rotating porous medium. Dash and Biswal [3] have investigated the effect of

at infinite vertical porous plate with time dependent temperature and concentration.

Rath and Bastia[7] have analysed the steady flow and heat transfer in a visco-elastic fluid between two coaxial rotating disks. Mukherjee and Mukherijee[8] have studied the unsteady axisymmetric rotational flow of elastico-viscous liquid. Datta and Jana[9] have investigated the problem of flow and heat transfer in an elastico-viscous liquid over an oscillating plate in a rotating frame.

The present study considers the simultaneous effect of heat and mass transfer on the hydrodynamic free convection flow of a rotating viscoelastic fluid past an infinite vertical isothermal porous plate.

2. Formulation of the problem

Let us consider the flow of a rotating viscoelastic incompressible fluid σ in a medium past an infinite vertical isothermal porous plate. Let x and y-axis be in the plane of the plate and

z-axis normal to the plate and let u, v, w be the velocity components of the fluid in *x, y* and z direction respectively. Both the liquid and the plate are considered in a state of rigid body rotation about z-axis with uniform angular velocity Ω. Further, Let us assume that the constant heat source Q (absorption type $Q = -Q_0$ $(T-T_\infty)$ is at $z = 0$ and the suction velocity at the plate $w = -w_0$ where w_0 is a positive real number. Here we have neglected buoyancy effect. Since the plate is infinite in extent, all physical variables depend on z and t only. Considering *u* + *iv = q* and using non-dimensional quantities the equations governing the flow (dropping the asterisks) are

$$\frac{\partial q}{\partial t} - \frac{\partial q}{\partial z} + \left(2iR + \frac{1}{K_p}\right)(q - U) = \frac{\partial U}{\partial t} + \frac{\partial^2 q}{\partial z^2} + R_c \frac{\partial^3 q}{\partial z^2 \partial t} + G_r T + G_m C$$

--------------------(1)

$$\frac{\partial^2 T}{\partial z^2} + \Pr \frac{\partial T}{\partial z} - \Pr \frac{\partial T}{\partial t} - \alpha_0 T = 0$$

--------------------(2)

$$\frac{\partial^2 C}{\partial z^2} + Sc \frac{\partial C}{\partial z} - Sc \frac{\partial C}{\partial t} = 0$$

--------------------(3)

The non-dimensional boundary conditions are

$q = 1 + \varepsilon e^{i\omega t}$, $(T = 1 + \varepsilon e^{i\omega t}, \omega<0)$, $C = 1 + \varepsilon e^{i\omega t}$ at $z = 0$

$q = (1 + \varepsilon e^{i\omega t}, \omega>0)$, $T \to 0$, $C = 0$ at $z \to \infty$ --------------------- (4)

The non-dimensional quantities introduced in equations (1) – (3) are defined as

$$z^* = w_0 \frac{z}{v}, \quad t^* = w_0 \frac{t}{v}, \quad u^* = \frac{u}{U_0},$$

$$K_p^* = K_p' \frac{w_0^2}{v^2}$$

Where K_p' is the dimensional permeability parameter and K_p^* is the non-dimensional permeability parameter.

$$q = \frac{u}{U_0} + i\frac{v}{U_0}, \quad T = \frac{T - T_\infty}{T_w - T_\infty}, \quad P_r = \frac{v\rho C_p}{K},$$

$$R = \frac{\Omega v}{W_0^2}, \quad R_c = \frac{K_0 \omega_0^2}{v^2}$$

$$G_m = \frac{\rho g \beta^* (C_w - C_\infty)}{U_0 w_0^2}, \quad G_r = \frac{v g \beta_0}{U_0 w_0^2}(T_w - T_\infty),$$

$$V = \frac{v'}{U_0}$$

Where P_r : Prandtl number

G_r : Grashof number,

M : Magnetic parameter,

R : Rotation parameter

K_p : Permeability parameter,

α_0 : heat source parameter,

U : free steam velocity

R_c : non-Newtonian parameter

3. Solution of the equations:

In order to solve the equation [1] – [3] we assume velocity, temperature and concentration of the liquid in the neighborhood of the plate as

$$Q = (1 - q_0) + \varepsilon e^{i\omega t}(1 - q_1)$$

--------------------(5)

$$T = T_0 + \varepsilon T_1' e^{i\omega t}$$

--------------------- (6)

$$C = C_0 + \varepsilon e^{i\omega t} C_1$$

---------------------- (7)

Where

$(1 - q_1) e^{i\omega t} = (M_r + I M_i) (\text{Cos } \omega t + i \sin \omega t)$

$T_1 e^{i\omega t} = (T_r + iT_i) (\text{Cos } \omega t + i \sin \omega t)$

$C_1 e^{i\omega t} = (C_r + iC_i) (\text{Cos } \omega t + i \sin \omega t)$

For free stream (when $\varepsilon << 1$)

$U(t) = 1 + \varepsilon e^{i\omega t}$ ------------------------(8)

The transient primary velocity and temperature profiles can be deduced from equations (5) and (6) for $\omega t = \pi/2$

Hence $u(z, t) = u_0(z) - \varepsilon M_i$

---------------------- (9)

$T = T_0 - \varepsilon T_1$

---------------------- (10)

Where $u_0(z) + iv_0(z) = 1 - q_0$

Using equations (5) – (8) in equations (1), (2) and (3) and equating harmonic and non-harmonic terms, we get

$(1 + i\omega R_c)$

$$q_1'' - Vq_1' - \left[\frac{1}{K_p} + (2iR + i\omega)\right] q_1 = -G_r T_1 - G_m C_1$$

---------------------- (11)

$$q_1'' - Vq_0' - \left(\frac{1}{K_p} + 2iR\right) q_0 = -G_r T_0 - G_m C_0$$

---------------------- (12)

$$T_0'' - P_r V T_0' - \alpha_0 T_0 = 0$$

---------------------- (13)

$$T_1'' - P_r V T_1' - (i\omega P_r + \alpha_0) T_1 = 0$$

---------------------- (14)

$$C_0'' + S_c C_1' = 0$$

---------------------- (15)

$$C_1'' + S_c C_1' - i\omega S_c C_1 = 0$$

---------------------- (16)

The boundary conditions are

$q_0 = 0$, $T_0 = 1$, $C_0 = 1$, $q_1 = 0$, $T_1 = 1$, $C_1 = 0$ at $z = 0$

$q_0 \to 0$, $T_0 \to 0$, $C_0 \to 0$, $q_1 \to 0$, $T_1 \to 0$, $C_1 \to 0$ at $z \to \infty$ ---------------------- (17)

Solving equations (11) – (16) with boundary conditions (17) we obtain for $R_c << 1$ and ω small,

$$C_0 = e^{-S_c z}$$

$$C_1 = e^{D_1 z}$$

$$T_0 = e^{D_2 z}$$

$$T_1 = e^{D_3 z}$$

$q_0 =$

$$\frac{Gr\left(e^{D_2 z} - e^{D_2 z}\right)}{(D_2 - D_4)(D_2 - D_5)} + \frac{Gm\left(e^{-scz} - e^{D_3 z}\right)}{(Sc + D_4)(Sc + D_5)}$$

$q_1 =$

$$\frac{Gr\left(e^{D_3 z} - e^{D_7 z}\right)}{(D_3 - D_6)(D_3 - D_7)} + \frac{Gm\left(e^{D_3 z} - e^{D_7 z}\right)}{(D_1 - D_6)(D_1 - D_7)}$$

Putting values of T_0, T_1 in equation (6)

C_0, C_1 in equation (7) and q_0, q_1 in equation (5) we get

$$T = e^{D_2 z} + \varepsilon e^{i\omega t}\left(e^{D_3 z}\right)$$

$$C = e^{-S_c z} + \varepsilon e^{i\omega t}\left(e^{D_1 z}\right)$$

$$q = 1 - \frac{Gr\left(e^{D_2 z} - e^{D_2 z}\right)}{(D_2 - D_4)(D_2 - D_5)} + \frac{Gm\left(e^{-scz} - e^{D_3 z}\right)}{(S_c + D_4)(S_c + D_5)}$$

$+ \varepsilon e^{i\omega t}$

$$\left[1 - \frac{Gr\left(e^{D_3 z} - e^{D_7 z}\right)}{(D_3 - D_6)(D_3 - D_7)} - \frac{Gm\left(e^{D_3 z} - e^{D_7 z}\right)}{(D_1 - D_6)(D_1 - D_7)}\right]$$

Separating real and imaginary parts

$u = 1 - G_r (B_4 B_{12} - B_5 B_{13}) - G_m (B_6 B_{14}$

$- B_7 B_{13}) + \varepsilon H_1 \, Cos \, \omega t - \varepsilon H_2 \, Sin \, \omega t$

$v = \varepsilon H_2 \, Cos \, \omega t + \varepsilon \, H_1 \, Sin \, \omega t - G_r \, (B_5 \, B_{12}$
$+ B_4 \, B_{13}) - G_m \, (B_7 \, B_{14} + B_6 \, B_{13})$

Where

$H_1 = 1 - (B_8 \, B_{16} - B_9 \, B_{17}) \, G_r - (B_{10} B_{18} - B_3 \, B_{11}) \, G_m,$

$H_2 = - G_r \, (B_0 \, B_{16} + B_8 \, B_{17}) - G_m \, (B_{11} \, B_{18} + B_3 \, B_{10}),$

$$D_1 = \sqrt{\left(1 + \frac{4}{K_p}\right)^2 + 64 R^2}$$

$$D_2 = \sqrt{\left(P_r^2 + 4\alpha_0\right)^2 + 16 \omega^2 P_r^2},$$

$$D_3 = \sqrt{\left(1 + 4M^2 \frac{4}{K_p}\right)^2 + \left(8R + 4\omega\right)^2},$$

Taking the value of V=1.

The values of the other constants involved are omitted here to save space.

4.   Results and Discussions

The problem of hydrodynamic free convection and mass transfer flow of a rotating viscous fluid past an isothermal vertical porous plate has been considered. The effects of porosity parameter (K_p), rotation parameter (R), Grashof numbers (G_r, G_m), frequency parameter (ω) and the heat source parameter (α_0) on the transient primary velocity, secondary velocity, temperature distribution profiles and the rate of heat transfer have been discussed to observe the physical significance and the mystery of the problem. To be realistic Prandtl number is chosen as 0.71 (the water vapour) and 7.0 (for water) approximately at 1 atmosphere and 25^0 C). From physical point

of view $G_r < 0$ corresponds to cooling of the plate and $G_r < 0$ corresponds to heating of the plate by free convection currents.

The transient primary velocity profiles are shown in Fig. 1. It is observed that the transient velocity u increases regressively and after attaining the maximum value decreases asymptotically and ultimately attains steady state. Comparing the curves (4 and 5) of figures (1), it is observed that the porosity parameter (K_p) increases the primary velocity. The rotation parameter has negligible effect on the primary velocity while it decreases the secondary velocity appreciably (curves 4 and 6). Comparing the curves (4, 8) and (2, 3), it is evident that the Grashof number for heat transfer (G_r) and the Grashof number for mass transfer (G_m) have accelerating effect on both the components of the velocity field. On careful observation of curves (4, 9) and (2, 4), it is observed that Prandtl number (P_r) and heat source parameter (α_0) have decelerating effect on the transient velocity. The frequency parameter has negligible effect on the velocity field.
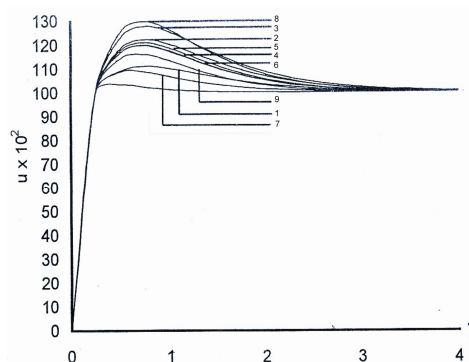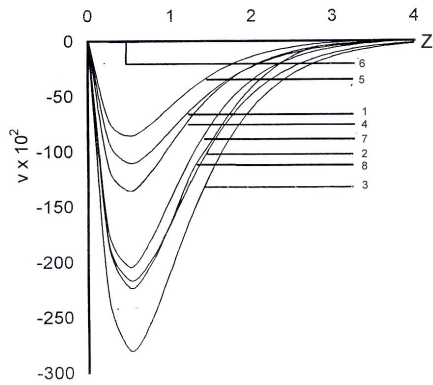

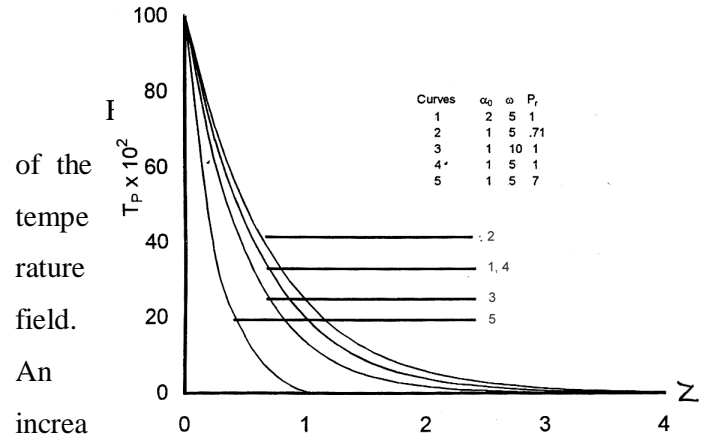**Fig 1 :** Effect of different parameters on transient

primary velocity profile when ε= 0.02

Curves	K_p	R	G_r	α_0	ω	P_r	G_m
1	1	.2	2	1	5	.71	0
2	1	.2	2	1	5	.71	2
3	1	.2	2	1	5	.71	3
4	1	.2	2	1	5	1	2
5	2	.2	2	1	5	1	2
6	1	.4	2	1	5	1	2
7	1	.2	2	1	5	1	2
8	1	.2	4	1	5	1	2
9	1	.2	4	1	10	1	2

Fig 2: Effect of different parameters on transient secondary velocity profile when ε=0.02

The effects of different parameter on the transient secondary velocity have been shown in Fig.2. It is marked that the velocity V decreases with modified Grashof number G_m (curves 1, 2 and 3). This is valid for water vapour or air. But, the velocity 'V' rises for liquid sodium (P_r=1). As the value of rotation parameter R increase, the transient secondary velocity also rises (curve 5). The variation of frequency parameter ω does not produce any appreciable variation in the velocity 'V'. An increase in Grashof number G_r reduces the velocity 'V'. For water (P_r = 7.0), the velocity 'V' falls in comparison to that of liquid sodium (P_r = 1.0)

of the temperature field. An increase in frequency parameter leads to an oscillatory effect on the transient secondary temperature d for the transient ves 3 and 4). The ates the transient oes not affect the
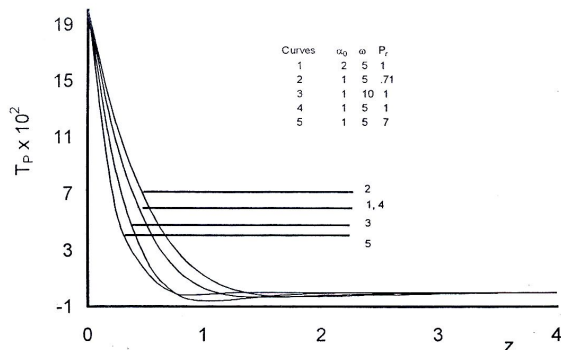
Fig 3 : Effect of different parameters on transient primary temperature profiles when ε= 0.02, ωt = $\pi/2$

Table (1 and 2) show the effects of α_0, P_r and ω on the transient primary rate of heat transfer (Nu_p) and transient secondary rate of heat

Curves	K_p	R	G_r	α_0	ω	P_r	G_m
1	1	.2	2	1	5	.71	0
2	1	.2	2	1	5	.71	2
3	1	.2	2	1	5	.71	3
4	1	.2	2	1	5	1	2
5	1	.4	2	1	5	1	2
6	1	.2	2	1	5	1	2
7	1	.2	4	1	5	1	2
8	1	.2	2	1	5	7	2

1 and ωt = s reveal that eter (α_0>0),

α_0 increase Nu_p and Nu_s and reverse effect is

noted for generation type that source parameter ($\alpha_0 > 0$). It is interesting to note that an increase in frequency parameter leads to an oscillatory effect on both components of transient heat transfer.

Fig. 4 : Effect of different parameters on transient secondary temperature profiles when $\varepsilon = 0.02$, $\omega t = \frac{\pi}{2}$

Table: 1

Variation of transient primary (Nu_p) and secondary (Nu_s) heat transfer at $\varepsilon = 0.01$ and $\omega t = \frac{\pi}{2}$ (absorption type)

α_0	P_r	ω	Nu_p	Nu_s
1.0	0.71	5.0	– 0.011391	0.019139
2.0	0.71	5.0	– 0.010032	0.021261
1.0	7.00	5.0	– 0.034770	0.085340
2.0	7.00	5.0	– 0.034309	0.086018
1.0	0.71	10.0	– 0.017410	0.023950
2.0	0.71	10.0	– 0.016239	0.025409

Table: 2

Variation of transient primary (Nu_p) and secondary (Nu_s) heat transfer at $\varepsilon = 0.01$ and $\omega t = \frac{\pi}{2}$ (generation type)

α_0	P_r	ω	Nu_p	Nu_s
– 1.0	0.71	5.0	– 0.015049	0.015359
– 2.0	0.71	5.0	– 0.017161	0.013911
– 1.0	7.00	5.0	– 0.035720	0.084011
– 2.0	7.00	5.0	– 0.036209	0.083340
– 1.0	0.71	10.0	– 0.020042	0.021279
– 2.0	0.71	10.0	– 0.021469	0.020091

5. Conclusion

The following conclusions are drawn for the transient primary and secondary velocity field, temperature field and the rate of heat transfer.

1. Heat source parameter (α_0) and Prandtl number (P_r) have decelerating effect on both components of the transient velocity field. The effect of the heat source parameter (α_0) is more pronounced than the Prandtl number (P_r). The Prandtl number also increases the temperature field.

2. The Grashof number for heat transfer

(G_r) and Grash of number for mass transfer (G_m) enhance both components of the velocity of the flow field.

3. The porosity parameter (K_p) accelerates the transient primary velocity field while it does not affect the transient secondary velocity field.

4. Both components of the velocity field decrease with the increase in rotation parameter but the effect is more pronounced in the case of the secondary velocity. The effect of frequency parameter(ω) is significant for both the components of temperature field and the rate of heat transfer and it has a negligible effect on the velocity field.

5. The heat source parameter enhances both components of rate of heat transfer for heat sink and the reverse effect is observed for heat source.

References

1. Raptis, A. and Singh, A.K., *Mechanics Res. Comm.,* Vol. **12**, pp 31-41 (1985)

2. Singh, A.K., *Proc. Ind. Nat. Sci. Acad*., Vol. **52** A, pp 221-226 (1986)

3. Dash, G.C. and Biswal, S., *AMSE, Modelling, Simulation and Control*, B, Vol. **21**, No.4 pp 13-24 (1989)

4. Dash, G.C. and Ojha, B.K., *AMSE, Modelling, Simulation and Control*, B, Vol **21** No. 1, pp 1-12 (1989)

5. Sacheti, N.C. and Singh, A.K., *Int. Comm. Heat and Mass transfer*, Vol. **19,** 423-433 (1992)

6. Sattar, M.A., *Indian J. Pure Appl. Math*., Vol **23**, No. 1, 759-766 (1994)

7. Rath, R.S. and Bastica, S.N., *Proc. Indian Acad. Sci.*, **87A**, No.9, 227-236 (1978)

8. Mukherje, S. and Mukherjee, Srikumar, *Indian J. pure appl. Math*., **14 (12)**, 1534-1541 (1983).

9. Datta, N. and Jana, R.N., *Istanbul Univ. Fen Fak. Mec. Seri A*, **43**, 121-130 (1978)

Thermodynamic Adsorption of Herbicides on Eight Agricultural Soils

Rounak M. Shariff, Kafia M. Shareef

Abstract— Thermodynamics of adsorption of three pesticides using a kinetic approach was investigated. The objective of this study is to study the effect of temperature on the sorption behavior of three commonly used pesticides in Kurdistan onto eight natural soil samples collected from different agricultural locations. To elucidate the effect of temperature on the sorption process, the experiments were done at 15, 25, 35 ±1ºC. Values of the standard thermodynamic functions: equilibrium constant K_o, free energy change ΔG^o, enthalpy change ΔH^o and entropy change ΔS^o revealed that adsorption of atrazine, picloram and propanil was spontaneous exothermic and physical in nature to some extent. Data obtained revealed that adsorption coefficient decreases with the increasing temperature. Values of $\ln K_o$ were in the range 1.238 to 12.75, 1.674 to 14.032 and 29.83 to 173.1 for atrazine, picloram and propanil respectively. The ΔG^o values were in the range -30.549 to -3.172 $KJmol^{-1}$, -33.614 to -4.2897 $KJmol^{-1}$, and - 414.72 to - 76.412 $KJmol^{-1}$ for atrazine, picloram, and propanil respectively. Values of ΔH^o followed the range -5.010 to -0.738 to $KJmol^{-1}$, -4.769 to -0.848 $KJmol^{-1}$ and -75.779 to - 1.628 $KJmol^{-1}$ for atrazine, picloram, and propanil respectively. ΔS^o followed the range-16.138 to -1.757 $Jmol^{-1} k^{-1}$, -14.711 to -2.502 $Jmol^{-1} k^{-1}$ and -199.01 to -62.457 $Jmol^{-1}k^{-1}$ for atrazine, picloram, and propanil respectively.

Index Terms — Adsorption isotherms, Adsorption thermodynamic, atrazine, picloram, propanil, HPLC.

———————————————— ◆ ————————————————

1 INTRODUCTION

THERE is considerable public concern about potential adverse impact of pesticide used on ecosystem and human health. To minimize any such detriments, sound understanding of environmental fate and behavior of pesticides is necessary under local soil and environmental conditions [1], [2]. Temperature is an important factor governing the rate of adsorption in soil pore. The existence of a number of solid structures of picloram were suggested and discussed with the application of the values of ΔH^o_{sol} as a correction of solubility- temperature effect on the standard enthalpy of the pesticide adsorption processes [3], [4] . The major factors that determine the extent to which herbicides are adsorbed by soil include: i) physical or chemical characteristic of the adsorbent, ii) physical or chemical properties of the pesticides, and iii) properties of the soil system, such as cleay mineral composition, pH, kinds and amounts of exchangeable cations, and temperature [5], [6]. Since information on the sorption behavior of pesticides in soil is essential in predicting their leaching potential and contamination of ground water and no data are available in literature for sorption kinetics equilibrium parameters of the three commonly used pesticides (atrazine, picloram and propanil) in Kurdistan. Studies were conducted on their sorption and determining the thermodynamic parameters

———————————————

- *Rounak M.Shariff, Departmen of Chemistry, College of Science, University of Salahaddin- Erbil-Kurdistan region -Iraq. E-mail: rounakm2000@yahoo.com*
- *Kafia M.Shareef, Departmen of Chemistry, College of Science, University of Salahaddin- Erbil-Kurdistan region -Iraq. E-mail: skavemy@yahoo.com*

associated their sorption onto natural soil samples.

2 METHODLOGY

2.1 Soils

Fresh soil samples were taken from eight main agricultural locations, representing awide range of physicochemical properties. Subsamples of homogenized soils were analyzed for moisture content, organic matter content, particle size distribution, texture, pH, loss on ignition and exchangeable basic cations the detail were charactrerized in previous article [7].

2.2 Pesticides

Analytical grad substituted atrazine, picloram and propanil were purchased from Riedal-de Haen, Sigma-Aldrich Company. All chemicals used were of analytical grade reagents and used without pre-treatments. Standard stock solutions of the pesticides were prepared in deionised water.

2.3 Adsorption Experiments

The effect of temperature on adsorption of pesticides from aqueous solution was determined at 15, 25, 25±1 C° employing a standard batch equilibrium method [8]. Duplicate air-dried soil samples were equilibrated with different pesticide concentrations (2, 5, 10, and 15 µgml-1) at the soil solution ratios: 4:10, 4:8, and 1:10 for atrazine, picloram and propanil respectively. The samples plus blanks (no pesticide) and control (no soil) were thermostated and placed in shaker for 24 h for atrazine and picloram and for 10 h for propanil. The tubes were centrifuged

for 20 min. at 3500 rpm. Supernatants were analyzed by PerkinElmer series 200 USA family high performance liquid chromatography (HPLC) for each pesticide concentration. The detailed information about the soil characteristics and their sorption process has been reported in our previous work [8], [9].

3.1 Equilibrium constant

The equilibrium constant K_o was enabled us to calculate the thermodynamic parameters for physico-chemical equilibrium between pesticides and soils [10].

$$K_o = \frac{a_s}{a_e} = \frac{\gamma_s}{\gamma_e}\frac{C_s}{C_e} \quad (1)$$

Where a_s: activity of the adsorbed solute, a_e: activity of the solute in the equilibrium solution, C_s: µg of solute adsorbed per milliliter of solvent in contact with the adsorbent surface, C_e: µg of solute per milliliter of solvent in the equilibrium solution, γ_s: activity coefficient of the adsorbed solute. γ_e: activity coefficient of the adsorbed solute in the equilibrium solution. As the concentration of the solute in the solution approaches zero, the activity coefficient, γ, approaches unity. "Equation (1), "may then be written as:

$$C_s \xrightarrow{\lim} 0 \ \frac{C_s}{C_e} = \frac{a_s}{a_e} = K_o \quad (2)$$

Values of $\ln K_o$ were obtained from the plot of $\ln(C_s/C_e)$ vs. C_s, $\ln K_o$ was obtained at $C_s = 0$, as described by Biggar and Chenung [11]. The results were summarized in Table 1, 2, and 3 for atrazine, picloram and propanil respectively. Values of $\ln K_o$ were in the range 1.238 to 12.75, 1.674 to 14.032 and 29.83 to 173.1 for atrazine, picloram and propanil respectively. It is well known that K_o is constant at constant temperature and the position of equilibrium depends only on thermodynamic quantities and is independent of any consideration of kinetics or mechanism. Values of K_o obtained can vary among soils due to the quantities and composition of soil components. The K_o values were decreased with rise in temperature, confirming that the pesticides had a high preference for adsorption at low temperature.

3.2 Standard Free Ernergy Change

Adsorptions equilibrium constant K_o can be expressed in terms of the standard Gibbs free energy ΔG^o for adsorption [12], [13].

$$\Delta G^o = -RT\ln K_o$$

(3)

The values of ΔG^o at 288.15, 298.15 and 309.15K were summarized in Tables 1, 2, and 3 for adsorption of atrazine, picloram and propanil respectively. The values of ΔG^o for adsorption of studied pesticides were negative and decreased with temperature rise indicating that adsorption of pesticides on the soils were spontaneous with a high preference of the soil surface. The data revealed that adsorption of pesticides were in the following order: propanil > atrazine > picloram. The magnitude of ΔG^o also showed that the interactions of pesticides with the soil were thermodynamically spontaneous process and adsorption occurred through a bonding mechanism. The ΔG^o values were in the range -30.549 to -3.172 KJmol^{-1}, -33.614 to -4.2897 KJmol^{-1}, and -14.72 to -76.412 KJmol^{-1} for atrazine, picloram and propanil respectively. The results obtained in the present study are similar to those of Gupta et al [14] who reported an increase in the values of ΔG^o with temperature. Variation of ΔG^o with temperature may be due to the increase in the degree of freedom of adsorbed molecules, which enhances desorption rather than adsorption at higher temperatures [15], [16].

3.3 Standard Enthalpy Change

The standard enthalpy change of adsorption ΔH^o represents the difference in binding energies between the solvent and the soil with the pesticides. Values of ΔH^o were determined graphically from the slope of the plot of $\ln K_o$ vs. $1/T$ Fig. 1 a, b, and c using the following "equation" [11].

(4)

$$\left[\frac{dLnK_o}{d(\frac{1}{T})}\right] = \frac{-\Delta H^o}{R}$$

Values of ΔH^o were summarized in Table 4, followed the range -5.010 to -0.738, -4.769 to -0.848 and -75.779 to -1.628 KJmol^{-1} for atrazine, picloram and propanil respectively. The negative values of ΔH^o indicated the exothermic behaviors of the reaction. The linear nature of the plot indicates that the mechanism of adsorption is not changed as temperature is changed. The values of R^2 were in the range 0.9999 to 0.902, 0.989 to 0.765, and 0.902 to 0.999 for atrazine, picloram and propanil respectively. The negative enthalpy of adsorption indicates an exothermic binding [17], [18]. Showing that the interaction of pesticides with the soil is an energetically stable exothermic process and the adsorption occurred through a bonding mechanism. The ΔH^o values explain the binding strength of pesticides to the soil; the lower negative value of ΔH^o indicates stronger binding. Thus low values of ΔH^o pointed towered chemisorptions; hence the herbicides adsorption may be due to coordination and /or protona-

tion, hydrogen bonding and dipole association and van der Waal's forces. This indicates that the interactions between the pesticides and the studied soil samples were stronger at lower temperature.

Fig 1 **a Variation of ln K_o with 1/T for adsorption of atrazine on the eight soil samples (♦ S_1, ■ S_2, ▲ S_3, x S_4, * S_5, •S_6, +S_7, -S_8).**

Fig.1. b **Variation of ln K_o with 1/T for adsorption of picloram on the eight soil samples (♦ S_1, ■ S_2, ▲ S_3, x S_4, * S_5, •S_6, +S_7, -S_8).**

Fig. 1 c **Variation of ln K_o with 1/T for adsorption of propanil, on the eight soil samples (♦ S_1, ■ S_2, ▲ S_3, x S_4, * S_5, •S_6, +S_7, -S_8).**

3.4 Isosteric Enthalpy Of Adsorption

The isosteric enthalpy of adsorption ΔH is the standard enthalpy of adsorption at a fixed surface coverage. Values of ΔH were calculated by the expression [17].

$$\Delta H = R[\frac{dLnC_e}{d(\frac{1}{T})}]x \tag{5}$$

Where x is amount of pesticide adsorbed, and the average was calculated for each concentration. The values of ΔH Table 5 were in the following range -0.05 to -0.0084 kJmol^{-1}, -0.0057 to - 0.014 kJmol^{-1}, and -0.0141 to -0.0034 kJmol^{-1} for atrazine, picloram and propanil respectively. The values of ΔH of adsorption as a function of amount of pesticide adsorbed was almost the same for all the three pesticides on eight soils that support our inference regarding the mechanism of adsorption. These values were relatively small and were of the order which was consistent with a physical type of adsorption [18].

3.5 Standard Entropy Change

The values of standard entropy change ΔS^o of adsorption were determined by using the "equation bellow" [18].

$$LnK_o = \frac{\Delta H^o}{RT} + \frac{\Delta S^o}{R} \tag{6}$$

The values of ΔS^o were determined from the plot of -lnK_o against 1/ T and the results were summarized in Table 4. The values of ΔS^o followed the range -16.138 to -1.757 Jmol^{-1} k^{-1}, -14.711 to -2.502 Jmol^{-1} k^{-1} and -199.01 to - 62.457 Jmol^{-1}k^{-1} for atrazine, picloram, and propanil respectively. The negative values of ΔS^o indicate that the adsorption of the herbicide formed on all eight soil samples were stabilization, association, fixation or immobilization of the pesticides molecules as a result of adsorption decreased the degree of freedom, causing negative entropy effect[19], [20]. The values of ΔS^o pointing to the formation of the complexity by coordination or association of the herbicides and an exchangeable cation with the resultant of the loss in the degree of freedom of the pesticide.

3.6 Organic Matter Normalized Free Energy Change Of Adsorption

Organic matter was the most important factor that governed the adsorption of pesticides on soils. The organic
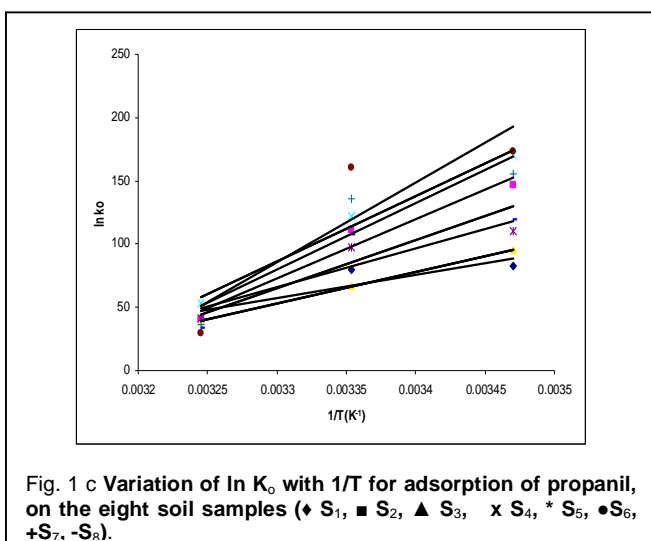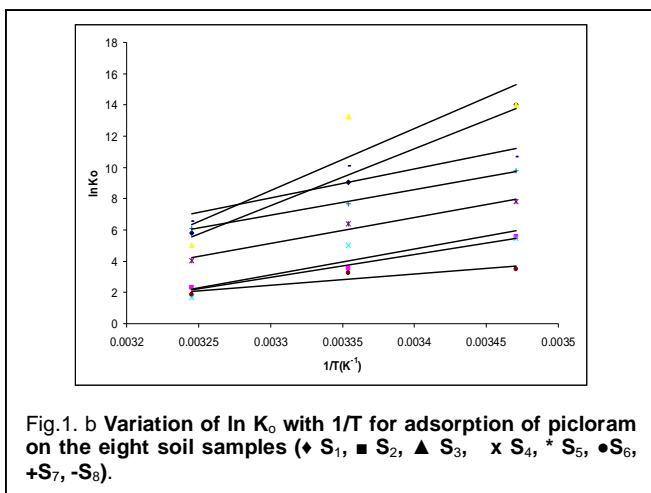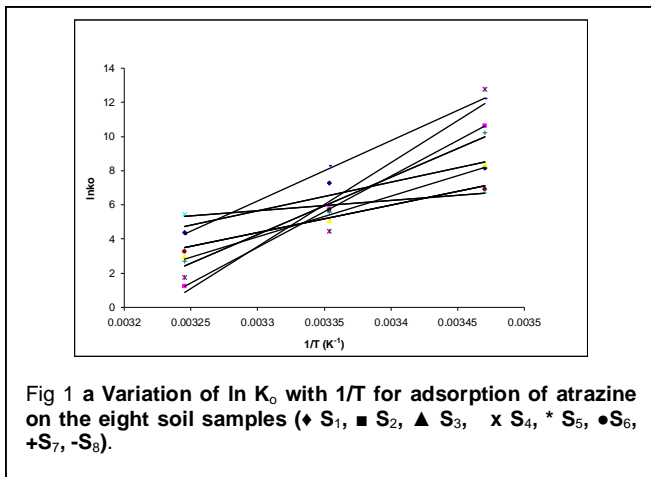
matter normalized free energy changes ΔG^o_{OM} of adsorption of atrazine, picloram and propanil were calculated by using the following equation [21].

$$\Delta G^o{}_{OM} = -RTLnK_{OM} \quad (7)$$

The values of ΔG^o_{OM} in Table 6 were in the range -8.0106 to -10.188, -10.629 to -7.7422, and - 7.7984 to - 2.0458 KJmol^{-1} for atrazine, picloram and propanil respectively. The negative values of ΔG^o_{OM} revealed that the adsorption of the three pesticides on the eight soil samples are spontaneous process and the adsorption has physical nature [22]. This signifies that there is a constant partitioning of the three pesticides. Between soil and solution, and water molecules do not pose strong competition for the adsorption sites, and also more affinity of the studied pesticides towards soil particulate matter than soil solution. Because most of the available sites in these soils are probably present at the surface of SOM and are therefore readily available for adsorption. The values of ΔG^o_{OM} for adsorption under the effect of temperatures were in the order $T_1 > T_2 > T_3$ for the studied pesticides (with some exceptions). This indicates that as the temperature increases the adsorption coefficients decrease for atrazine, picloram and propanil.
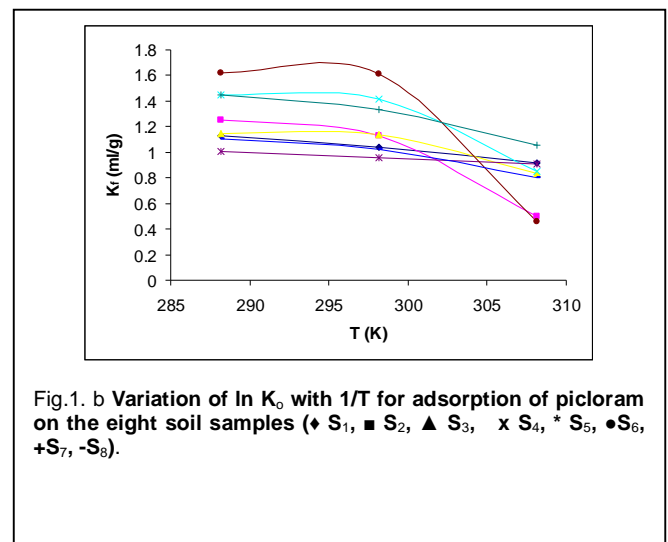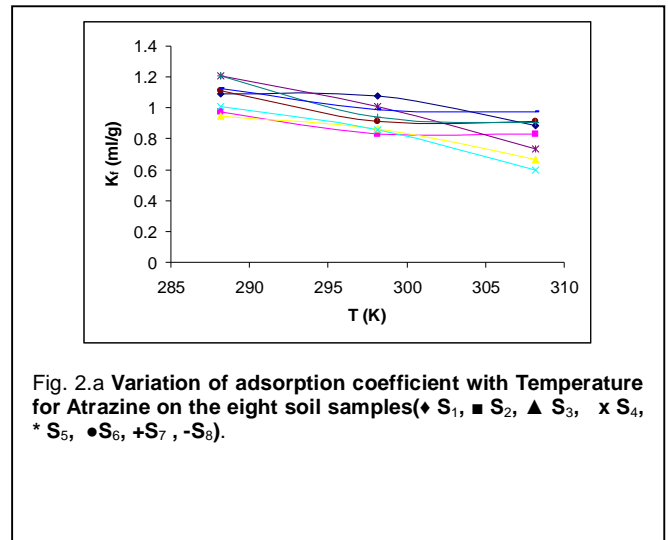
4 TEMPERATURE DEPENDENCY OF ADSORPTION COEFFICIENT

Variation of adsorption coefficient with temperature Fig. 2 a, b and c indicted that as the temperature increased the values of adsorption coefficient for these herbicides decreased. This can be explained by the types of the interactions between the pesticides and soil through adsorption that may occur as a result of two types of forces: enthalpy-related and entropy-related forces [23]. Hydrophobic bonding is anexample of an entropy-driven process; it is due to a combination of London dispersion forces (instantaneous dipole-induced dipole) associated with large entropy changes resulting from the removal of the sorbate from the solution. For polar chemicals, the enthalpy-related forces are greater, due to the additional contribution of electrostatic interactions. A small temperature effect was detected with the adsorption of studied pesticides by the soil samples; this behavior has been interpreted as due to physical adsorption. Results obtained in the present study are similar to those reported by Biggar et al [11].

4 CONCLUSION

Adsorption experiments were conducted at 15, 25, and 35°C to study the thermodynamic (equilibrium) parameter, associated the adsorption of the studied pesticides on the selected soil samples. The net effect of pesticide ad-

sorption associated to the temperature. Adsorption efficiency for the studied pesticides was found to depend on the nature of adsorbent and adsorbate and the nature of the interactions between them.

Fig. 2.a **Variation of adsorption coefficient with Temperature for Atrazine on the eight soil samples(♦ S$_1$, ■ S$_2$, ▲ S$_3$, x S$_4$, * S$_5$, ●S$_6$, +S$_7$, -S$_8$).**

Fig.1. b **Variation of ln K$_o$ with 1/T for adsorption of picloram on the eight soil samples (♦ S$_1$, ■ S$_2$, ▲ S$_3$, x S$_4$, * S$_5$, ●S$_6$, +S$_7$, -S$_8$).**

Fig 2 c **Variation of ln K_o with 1/T for adsorption of propanil, on the eight soil samples ($\diamond$ S_1, $\blacksquare$ S_2, $\blacktriangle$ S_3, x S_4, * S_5, $\bullet S_6$, +S_7, -S_8).**
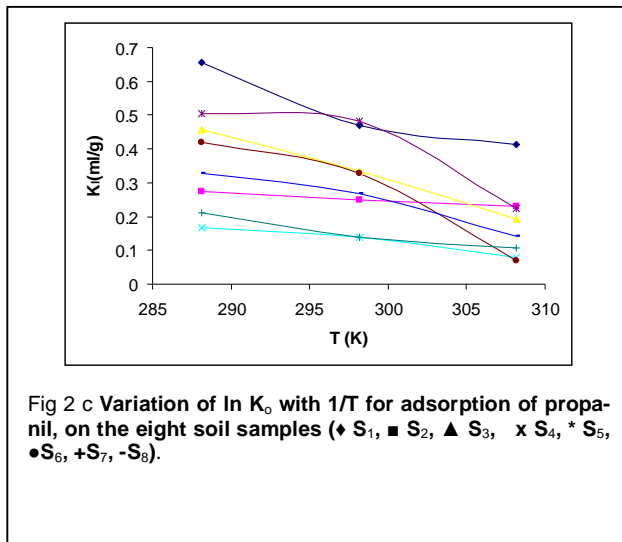
REFERENCES

[1] Smith, M.C., D.R., Shaw, J.H. Wassey, M. Boyette, and W. Kingery. "using Nonequilibrium Thin-disc and Batch Equilibrium Techniques to Evaluate Herbicide Sorption". J. Environ. Qual. 32 (2003):1393-1404

[2] B. S. Ismail, A. O. S. Enoma, U. B.Cheah, K. Y. Lum, and Zulkifli Malik." Adsorption, Desorption, and Mobility of Two Insecticides in Malaysian Agricultural Soil". J. Environ. Sci. Health. 4 B37(2002):355-364

[3] A. Bala Subramaniam, C. Mejean, M. Abkarian, H. A. Stone." Microstusture, Morphology and Lifetime, of Armored Bubbles Exposed to Surfactants". J. Colloids, Bubbles and Surfactant. Environ. 4(2005): 553-556

[4] Braida, W. J., J.J. Pignatello, Y. Lu, P.I. Ravikovitch, A. V. Neimark, and B. Xing. "Sorption Hysteresis of Benzene in Charcoal Particles". Environ. Sci. Technol. 37(2003): 409-417

[5] J.B. Alam, A. K. Dikshit and M. Bandy Opadhayay. "Evaluation of Thermodynamic Properties of Sorption of 2,4-D and Atrazine by Trirubber Gramules". Sepration and Purification Technology. 42(2005):80-40.

[6] R. Ahmad, R.S. Kookana, and A.M. Alston, bulletin of Environ. Contaminat and toxicol. 66 (2001):313-318.

[7] M. Kafia and M. Rounak, "Adsorption of Herbicides on eight Agricultural Soils". J. of university of anbar for pure since. 3 (2009):67-75.

[8] G. Suman, and V. T. Gajbhiye, J. of Environ. Sci. and health, 6 B37 (2002):573-586.

[9] Alexandre G. S. Prado, Andreia H. Tosta and Glaudio Airoldi. "Asorption, Separation, and Thermodynamic Data on Herbicide Picloram Anshored on Silica Gel and its Cation Interaction Behavior."Journal of Colloidal and Interface Science. 2 269(2004):259-264

[10] OP. Bansal. "Kinetics of Interaction of three Carbamate Pesticides with Indian soils: Aligarh district. Pest Manag Sci. 60(2004):1149-1155

[11] Biggar, J. W. and M. W. Chenung "Adsorption of picloram (4-Amino 3, 5, 6-Trichloropineolinic Acidon Panoche. Ephrata and Palouse Soils a Thermodynamic Approach to Adsorption Mechanism". Soil Sci. Soc. Am. Proc. 37(1973): 863-868

[12] K. Chaudhary and B. Prasad. "Thermodynamics of Potassiume Exchange Reaction in Entisol and vertisol using a Kinetic Approach by Miscible Displacement Technique ". Journal of the Indian Society of Soil Science 2 47(1999) : 221-229

[13] T. Sismanoglu, A. Ercage, S. Pura and E. Ercage. "Kinetics and Isotherms of Dazomet Adsorption on Natural Adsorbents" J. Braz. Chem. Soc. 5 15 (2004):669-675

[14] H. K. Gupta, Devprakash, P. K. M. Ishra and vishal pathak. "Studies on the Kinetics and Thermodynamics of Adsorption of Cypermethrin on Montmorillonite Clays". Asian Journal of Chemistry . (3-4)16 (2004):1391-1397

[15] Michael Sander, Yuefeng Lu and Josseph J. Pingnatello. "A Thermodynamically Based method to Quantify Trine Sorption Hystersis. J. Environ. Qual. 34(2005):1063-1072

[16] Young –HakKim, Thomas M. Heinze, Seong-Jae Kim, and Carl E. Cerniglia. "Adsorption and Clay – Catalized Degradation of Erythromycin A on Homoionic Clays". J. Environ. Qual. 33(2004):257-264

[17] W. J. Farmer and Y. Aochi. Soil Sci. Soc. Am. Proc. 38(1974): 418

[18] Elsayed A. Elkhatib, A. M. Mahdy and N. H. Bbrakat "Thermodynamics of Copper Desorption from Soils as Affected by Citrate and Succinate" Soil & Water Res. 4 (2007):135-140

[19] J. Phys Chem .Ref Data. 30(1)(2001):231-235

[20] John O. Agbenin and Bernardo van Raij" Kinetics and Energetic of Phosphate Release from Tropical Soils Determined by Mixed Ion- Exchange Resins ".Soil Sci. Soc. Am. J. 65(2001):1108-1114

[21] Sarah Taylor –lovel, Gerald k. Sims, and Loyd M. Wax."Effect of Moisture, Temperature, and Biological Activity on Degradation of Isoxaflutole in Soils: Aligath District". J. Agric. Food Chem.50(2002):5626-5633

[22] Adlophe Monkiedje and Micheal Spiteller. "Fungicides, Mefenoxam and Meta Laxly, and their Acid Metabolite in Typical Cameroonian and German Soils". Chemosphere. 49(6)(2002):659-668

[23] C.T. Chion, J. peters and V. H. Freed."A physical concept of soil water Equilibrium of Nonionic Organic Compounds. 206 (1979): 205

TABLE 1
EQUILIBRIUM CONSTANTS AND STANDARD FREE ENERGY CHANG AT THREE TEMPERATURES FOR ADSORPTION OF ATRAZINE ON THE SELECTED SOIL SAMPLES

Soil	T_1 K		T_2 K		T_3 K	
	lnK_0	ΔG^o (KJ/mol)	lnK_0	ΔG^o (KJ/mol)	lnK_0	ΔG^o (KJ/mol)
S_1	8.154	-19.534	7.271	-18.023	4.365	-11.185
S_2	10.64	-25.482	5.783	-14.334	1.238	-3.172
S_3	8.365	-20.039	5.106	-12.657	2.973	-7.618
S_4	6.820	-16.339	5.713	-14.162	5.446	-13.953
S_5	12.75	-30.549	4.467	-11.073	1.754	-4.493
S_6	6.918	-16.573	5.703	-14.136	3.273	-8.386
S_7	10.24	-24.526	5.557	-13.775	2.711	-6.944
S_8	12.20	-29.218	8.26	-20.465	4.24	-10.856

TABLE 2
EQUILIBRIUM CONSTANTS AND STANDARD FREE ENERGY CHANG AT THREE TEMPERATURES FOR ADSORPTION OF PICLORAM ON THE SELECTED SOIL SAMPLES

Soil	T_1 K		T_2 K		T_3 K	
	lnK_0	ΔG^o (KJ/mol)	lnK_0	ΔG^o (KJ/mol)	lnK_0	ΔG^o (KJ/mol)
S_1	14.03	-33.616	9.029	-22.382	5.817	-14.902
S_2	5.57	-13.349	3.561	-8.8273	2.287	-5.8602
S_3	14.03	-33.614	13.28	-32.919	5.013	-12.842
S_4	5.46	-13.076	5.009	-12.416	1.674	-4.2897
S_5	7.79	-18.678	6.407	-15.881	4.021	-10.300
S_6	3.48	-8.333	3.256	-8.0710	1.876	-4.8065
S_7	9.79	-23.472	7.653	-18.969	6.107	-15.648
S_8	10.7	-25.612	10.09	-25.024	6.520	-16.704

TABLE 3
EQUILIBRIUM CONSTANTS AND STANDARD FREE ENERGY CHANG AT THREE TEMPERATURES FOR AD-
SORPTION OF PROANIL ON THE SELECTED SOIL SAMPLES

Soil	T_1 K		T_2 K		T_3 K	
	$\ln K_0$	ΔG^o(KJ/mol)	$\ln K_0$	ΔG^o(KJ/mol)	$\ln K_0$	ΔG^o(KJ/mol)
S_1	83.00	-198.85	79.87	-197.99	41.35	-105.93
S_2	146.2	-350.18	111.3	-275.97	40.54	-103.87
S_3	95.06	-227.74	66.88	-165.79	38.87	-99.593
S_4	170.6	-408.79	122.3	-303.13	53.42	-136.87
S_5	110.4	-264.51	97.73	-242.28	41.35	-105.93
S_6	173.1	-414.72	160.6	-397.98	29.83	-76.412
S_7	155.7	-372.98	135.6	-336.15	36.40	-93.257
S_8	119.3	-285.73	107.3	-265.95	33.09	-84.764

TABLE4
STANDARD ENTHALPY CHANGE AND STANDARD ENTROPY CHANGE (DETERMINED GRAPHICALLY) FOR ADSORPTION
OF ATRAZINE, PICLORAM, AND PROPANIL ON THE SELECTED SOIL SAMPLES

Soil	atrazine			picloram			propanil		
	ΔH^o (KJ/mol)	R^2	ΔS^o J/mol.k	ΔH^o (KJ/mol)	R^2	ΔS^o (J/mol.k)	ΔH^o (KJ/mol)	R^2	ΔS^o (J/mol.k)
S_1	-2.010	0.902	-5.953	-4.396	0.989	-13.597	-22.025	0.791	-65.57
S_2	-5.019	0.999	-16.14	-1.757	0.988	-5.442	-56.168	0.955	-176.64
S_3	-2.885	0.989	-9.023	-4.769	0.797	-14.711	-29.995	0.999	-92.628
S_4	-0.738	0.901	-1.757	-2.003	0.823	-6.235	-1.628	0.986	-62.457
S_5	-5.904	0.931	-19.057	-2.009	0.972	-6.015	-36.605	0.869	-112.86
S_6	-1.938	0.957	-5.869	-0.848	0.838	-2.502	-75.779	0.799	-239.79
S_7	-4.029	0.985	-12.781	-1.973	0.995	-5.679	-63.204	0.859	-199.01
S_8	-4.247	0.999	-13.266	-2.208	0.841	-6.318	-45.634	0.838	-142.76

TABLE 5
ISOSTERIC HEAT CHEAHG OF ADSORPTION OF ATRAZINE, PICLORAM, AND PROPANIL ON THE SELECTED
SOIL SAMPLES.

Soil	atrazine X(J/mol)		picloram X (KJ/mol)		propanil X(J/mol)	
	$T_1:T_2$	$T_2:T_3$	$T_1:T_2$	$T_2:T_3$	$T_1:T_2$	$T_2:T_3$
S_1	-0.0183	-0.0135	-0.0105	-0.0087	-0.0141	-0.0093
S_2	-0.0243	-0.0042	-0.0115	-0.0082	-0.0113	-0.0069
S_3	-0.0289	-0.0120	-0.0119	-0.0059	-0.0096	-0.0089
S_4	-0.0291	-0.0202	-0.0140	-0.0057	-0.0112	-0.0030
S_5	-0.0501	-0.0084	-0.0116	-0.0091	-0.0117	-0.0739
S_6	-0.0215	-0.0119	-0.0119	-0.0108	-0.0131	-0.0034
S_7	-0.0327	-0.0100	-0.0129	-0.0085	-0.0126	-0.0050
S_8	-0.0183	-0.0145	-0.01231	-0.0102	-0.0130	-0.0056

TABLE 6
ORGANIC MATTER NORMALIZED FREE ENERGY CHANGE OF ADSORPTIO OF ATRAZIN, PICLORAM, AND PROPANIL ON THE
SELECTED SOIL SAMPLES

Soil	atrazine ΔG^o_{OM} (KJ/mol)			picloram ΔG^o_{OM} (KJ/mol)			propanil ΔG^o_{OM} (KJ/mol)		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
S_1	-9.001	-8.935	-8.746	-8.757	-8.854	-8.8203	-7.449	-6.890	-6.792
S_2	-9.967	-9.546	-9.861	-10.22	-10.311	-8.5742	-6.566	-6.564	-6.564
S_3	-9.905	-9.662	-9.297	-10.01	-10.325	-9.8902	-7.798	-7.297	-6.116
S_4	-10.19	-9.707	-9.100	-10.63	-10.942	-10.011	-5.459	-5.181	-3.908
S_5	-9.029	-8.545	-8.011	-8.247	-8.427	-8.5593	-6.597	-6.719	-4.978
S_6	-9.720	-9.204	-9.517	-10.26	-10.599	-7.7422	-7.031	-6.669	-8.424
S_7	-9.795	-9.137	-9.354	-9.863	-10.009	-9.7379	-5.257	-4.376	-2.046
S_8	-9.319	-8.969	-9.236	-8.928	-9.051	-8.7286	-5.829	-5.732	-5.759

Economic Water Rate and Optimum Performance of Two Stage Azeotropic Refrigerating System

Prof. D.V. Mahindru, Priyanka Mahendru

Abstract –With a view to conserve energy, the use of azeotropes in a multistage refrigerating system is quite timely. Depending upon the requirement, such a system incorporates conventionally either a water cooled or air cooled condenser. The total operating cost of a refrigerating system with a water cooled condenser comprises the cost of water and the cost of electricity needed to drive the compressor(s). There is enough potential for research in finding out the ways to achieve maximum coefficient of performance and the least operating cost simultaneously for multi-stage azeotropic system. However, to avoid overloading of sewage facilities and to comply with municipal codes for the use of water, the water flow rate required in refrigerating system should be minimized. In the present investigation, economic water rates for two stage refrigerating systems, operating on most commonly used azeotropes R-500 and R-502, have been searched out over a wide range of operating limits. Such economic rates, if followed, would produce maximum COP and consume minimum power. The effects of controlling variables, e.g. approach, cost ratio etc have also been studied on the heat transfer to condenser, optimum condensing temperature and economic water rate. The results have been presented in tabular form only.

Index Terms— Azeotropes, multistage refrigerating, compressors, water cooled condenser.

--------------------- ◆ -------------------

1. INTRODUCTION

THE total cost of a multistage refrigerating system with water cooled condenser comprises the cost of power required to run the compressors and the cost of water needed in the condensing unit to make the system work in a closed and continuous cycle. To minimize the total operating cost, the power consumed by compressors may be reduced by properly selecting their compression ratio, while cost on water can be diminished by consuming cooling water as low as possible. The compression ratio of a compressor gets affected, if either the condensing pressure or inter-stage pressure is changed. The quality of cooling water gets controlled by the quantity of heat to be transferred to the condenser. It is evident that larger is the water flow rate through a condenser, the lower will be the condensing temperature resulting in lower compression ratio for high pressure (HP) compressor and hence lower cost for compressor power, but more cost on water. To the contrary, if the lower water is used , condensing temperature would be higher and thereby expenditure on water decreases while that on compressor power increases. Hence it calls for a compromise between condensing temperature and cooling water rate to achieve minimum total operating cost. The cooling water rate that minimizes the total operating cost is usually termed as Economical Cooling Water Rate.

Prof. D.V. Mahindru is currently working as Professor (Mech. Engg.) in Mech. Engg Deptt of SRMGPC, Tewari Ganj, Lucknow-227105 U..P. India.

Priyanka Mahendru is currently working as Sr. Lecturer (E&C) in E.I.deptt of SRMGPC, Tewari Ganj, Lucknow-227105-227105 U.P.India

2. REVIEW OF PREVIOUS WORK

To cope with the existing energy shortages and the need to conserve the expended energy to the maximum possible extent, attempts have been made by Macharnen and chapman (4) and Downing (5) on various refrigerants and their mixtures. Among the mixtures of refrigerants, R-500 and R-502 have become very common. These are known as azeotropes. An azeotrope, by definition, is the mixture of refrigerants that does not separate in to their original components with pressure/temperature changes. It has fixed thermodynamic properties unlike those of their components.

Azeotrope R-500 consists of 73.8% R-12 and 26.2% R-152. Its normal boiling point is about 3.5℃ lower than that of R-12. It produces refrigerating effect per unit of swept volume about 18% more than that of R-12 .A Freon-12 system designed for 60 cycle current can be shifted to 50 cycle current by using azeotrope R-500. It would result in approximately the same refrigerating capacity and evaporator and condenser conditions.

Azeotrope R-502 is a mixture of 48.8% refrigerant R22 and 51.2% refrigerant R-115. It boils at a temperature of about 4.8℃ lower than that of R-22. Significantly lower discharge temperatures and lower winding temperatures are realized because of the higher capacities and lower values of compression ratio associated with R-502. Further R-502 decreases the swelling or softening effect on the common electrical insulating materials caused by

the presence of R-115.The inter-stage pressure for the two stage refrigerating system is conventionally selected as the geometric mean of operating pressure limits to minimize the total compression work. But it has been established in (6) that if power input to the system is to be minimized, the inter-stage pressure should be optimized with coefficient of performance (COP) as the objective function.

2. THERMO DYNAMIC CONCEPT

In general, one may write the heat rejected to condenser for a refrigerating system as :

$$Q_h = P(1+ COP) \qquad \text{------------1.1}$$

But Q_h per unit of cooling is expressed by :

$$Q_h/ Q_c = P(1+COP)/ Q_c = (1+1/COP) \qquad \text{-------------1.2}$$

Further, for a two stage refrigerating system, COP becomes maximum if inter-stage pressure is optimized for minimum power input. Equation 1.2 may be written as :
$$Q_{ho}/ Q_c = (1+1/COP_o)$$

As $COP_o > COP$, We get

$$Q_h/ Q_c = (1+1/COP) \quad Q_{ho} < Q_h \quad \text{from equations 1.2 and 1.3} \qquad \text{-------1.3}$$

It means that heat rejection to condenser would be minimum and hence minimum quantity of cooling water would be required for given condenser with a two stage system operates with optimum inter-stage pressure/temperature as decided on the basis of minimum power input.

Thus the problem of finding out economical cooling water rate for a two stage refrigerating system is coupled optimization problem, that is, first the system needs to be optimized for its minimum power consumption, and then optimum condensing temperature is to be searched out to minimize the total operating cost. On the other hand, if the system employs air cooled condenser, the total operating cost of the system would be that of power cost only. It would turn out to be minimum, if the COP of the system is maximum.

3. PRESENT WORK

In the present investigation, azeotropes R-500 and R-502 have been selected as the working fluids for two stage refrigerating system. Economic water rates that minimize the total operating cost and maximum COP are searched out over a wide range of operating temperature limits. Systems are also optimized for maximum COP, if they incorporate air cooled condensers in place of water cooled condensers. Optimum design quantities of interest are presented in the form of tables. Effects of operating variables on the design quantities are also displayed through tables.

4. SYSTEM ANALYSIS

5.1 System Employing Water Cooled Condenser (case-1)
Figure -1(a) shows the schematic of idealized two stage refrigeration system. The various heat and work quantities and pressure levels are indicated in the figure. The following simplifying assumptions are made for this system analysis:

a) The thermodynamic cycle of the system is a standard one comprising isentropic compression, isentropic expansion and absence of superheating of the suction vapour and sub cooling of the high pressure (HP) condensate.
b) The pressure drop in evaporator, compressor valves, condenser piping etc are neglected.
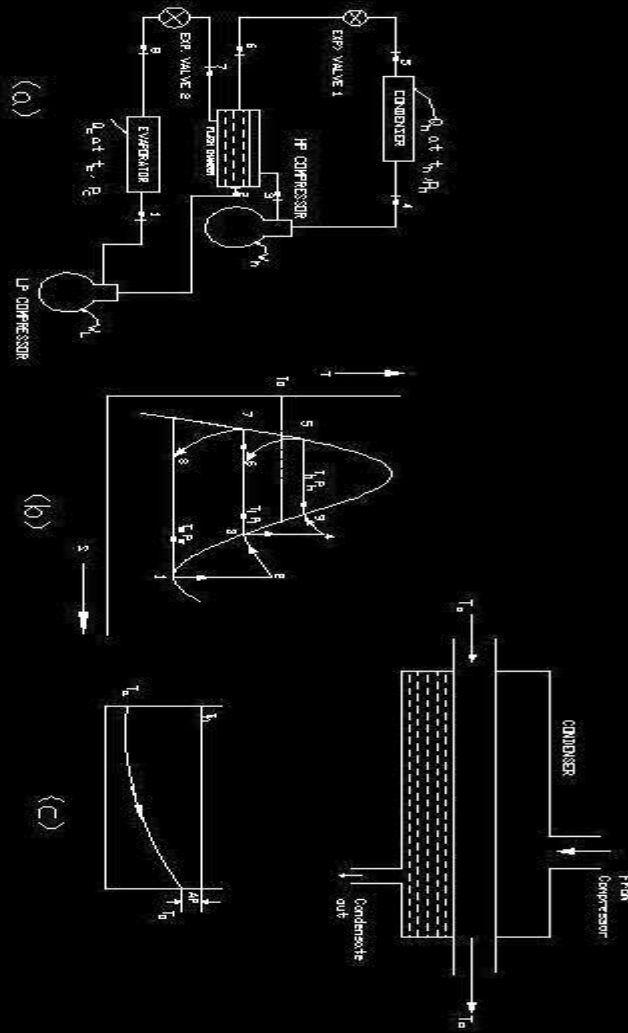c) Entire condensation of HP gas inside the condenser takes place at a fixed temperature (T_h).

Fig1. (a) Two stage Azeotropic Refrigerator System
(b) Temperature Entropy (T-S) diagram
(c) Condensation in the Condenser

Referring to fig-1(b), one may write refrigerant mass flow through LP compressor on per ton-hour basis as :

$M_1 = 12,600/(h_1 - h_5)$ ----------5.1

By energy balance on the flash chamber, refrigerant mass flow through HP compressor turns out to be :

$M_3 = M_1 (h_2 - h_1)/((h_3 - h_6)$

$= 12,600(h_2 - h_1)/(h_1 - h_5)(h - h_5)$ -----------5.2a

If we consider that refrigerant mass flow through LP compressor is unit kg, then mass flow through HP compressor based on similar lines, would be :

$M_3 = (h_2 - h_1)/((h_3 - h_6)$ -----------5.2b

and, total compression work shall be :

$W_T = (h_2 - h_1) + M_3((h_4 - h_3)$ -----------5.3a

However, on the basis of per ton hour, the total compression work may be written as :

$W = W_1 + W_2$

$= M_1 (h_2 - h_1) + M_1 (h_4 - h_3)$ -------------5.3b

Power consumption of the system :

$P = W/(3600)$ -------------5.4

For preliminary design purposes, the enthalpies per unit mass of superheated vapour at points 2 and 4 can be approximately related to the enthalpies per unit mass of the saturated vapours at points 3 and 4, respectively as:

$h_2 = h_3 + T_i (s_1 - s_3)$ -------------5.5

$h_4 = h_9 + T_h (s_3 - s_9)$ -------------5.6

saturated properties of both the azeotropes are estimated from the correlations available in reference (10)

The coefficient of performance of system shall be :

$COP = Q_c / W = 12,600/ W$ -------------5.7

Economic water rate expression as developed in ref (7,8) per unit ton of refrigeration, when total operating costs are minimized, is given by :

$M_{we} = 15.45 \ (Q_h._{\Delta P}.C/C_w)^{0.5}$ -------------5.8

When Q_h is heat rejection to condenser per ton per hour and is given as :

$Q_h = _3 (h_4 - h_5)$ -------------5.9

ΔP is the increase in power 'P' per degree rise in condensing temperature.

Optimum condensing temperature is expressed as :

$T_{ho} = T_{wi} + DT_o$ -------------5.10

Where $DT_o = 0.1545 \ (Q_h/ P.C_w/C)^{0.5} + AP$ -------------5.11

With AP as the approach, Fig 1(c) representing the end temperature difference.

It is evident that the expressions given in equation 5.4 and 5.7 to 5.9 can be expressed in temperature alone. An explicit expression has not been attempted at as it becomes extremely involved. Moreover, it serves no useful purpose because We can directly feed the above expressions in computer program to evaluate the objective function. The governing performance quantities in terms of operating parameters/variables can be expressed as:

$P = P \ (T_i, T_h, T_e, T_{wi}, AP, Q_c)$ -------------5.12

$\Delta P = P \ (T_h + 1, T_i, T_e, T_{wi}, AP, Q_c) - P \ (T_h, T_i, T_e, T_{wi}, AP, Q_c)$ -------------5.13

$Q_h = Q_h(T_h, T_i, T_e, T_{wi}, AP, Q_c)$ -------------5.14

$_{we} = _{we} \ (T_h, T_i, T_e, T_{wi}, AP, Q_c, C, C_w)$ -------------5.15

$COP = (T_h, T_i, T_e, T_{wi}, AP, Q_c)$ -------------5.16

5.2 System Employing Air Cooled Condenser (Case-2)

All the assumptions as mentioned above are also valid for this case. Since System is not incorporating water cooled condenser, only power consumed by compressors needs to be minimized to minimize the total operating cost.It indirectly suggests that COP should be maximized. In equation 5.7, compression work (W) has been substituted from equation 5.3a. It has been established in ref.6 that COP is strongly dependent upon the inter-stage pressure/temperature besides other factors. Thus in order to achieve maximum COP, inter-stage pressure/temperature should be optimized. Equations 5.2 to 5.7 have been used for this.

5.3 Objective Function and Optimization

For case 1, the objective function is the total operating cost together with the COP as given by equations (5.15) and (5.16) above. The total operating costs are to be minimized producing maximum COP as well.Since total operating costs
have been minimized while deriving expressions for economic water rate   ($\circledcirc$we).from equations 5.15 and 5.16, it is clear that $\circledcirc$we and COP depend upon inter-stage and condensing temperatures (Ti and Th) if other parameters are   kept fixed. It leads to a two dimensional maximization problem with the two decision variables (Ti and Th) subject to the constraints:

Te < Ti < Th   ------------5.17

And

Th > Twi + AP   ------------5.18

For case -2 the objective function is the coefficient of performance as given by   (5.7) but in this equation (W) has been calculated from equation (5.3a). There is only one decision variable (Inter-stage pressure/temperature) and hence would be a one dimentional maximization problem subjected to the constraint as given by equation (5.17) only.

5.4 Solution Technique

To find Tho, where total operating cost is minimum together with the optimum system performance, initially some convenient Th >Ta was assumed. With the help of this Th and given values of evaporator temperature(Te) subroutine maximises the COP and transfers required optimum quantities ($\circledcirc$3 and P) to the main program. Now Th is increased by unit degree and the above process is repeated. P is determined. Q☐h is found from equation 5.9 in the main program. Thereafter, DTo is estimated from equation 5.11 to determine Tho from equation 5.10. With this new value of Tho, the above computations are repeated till two successive values of Tho differ by + 0.1% . Condensing temperature , thus predicted , is the required optimum condensing temperature (Tho) because it produces minimum operating cost for maximum COP. Finally, at Tho , $\circledcirc$we is determined from equation 5.8, Fig 1.2(a) and 1.2(b).

For case-2, only subroutine OPTIMUM is used. Different values of condensing and evaporating temperatures are fed to the sub-routine and inter-stage temperatures are optimized for maximum COP's. Later at these optimal points, main design quantities are computed and are termed as the optimum design quantities as they yield maximum COP of the system or minimum operating cost. Method of successive iterations has been used to optimize the systems for both the cases. During computation, the necessary properties of the liquid and gaseous azeotropes have been found from the co-relations given in ref (9).

Different ranges of various operating parameters considered for the azeotropic systems are based upon practical considerations and their commonly adopted values. They are as follows:

Evaporator temperature   : 5   t0 -50$\circledcirc$
Ambient/Cooling Water Temperature   : 15   to 75$\circledcirc$
Approach Temperature   : 2   to 5$\circledcirc$
Cost ratio   : 0.5 to 10

6. RESULTS AND DISCUSSIONS

6.1 SYSTEM WITH WATER COOLED CONDENSER:

Besides the direct use of Tables 6-1 to 6-3 for preliminary optimum design of the systems, they also exhibit the quantitative effects of operating variables on the design quantities for a specified set of operating parameters. Not only this, the feasible operating conditions can also be achieved with the help of the figures achieved. The graphical presentation of thermodynamic performance, display of main optimum design quantities for R-500 and R-502 respectively when the systems have been optimized for COP's only over a quite wide range of operating temperature limits is under the scope of future work..

The approach (AP) has been kept at 3⊚. For a fixed set of Rc, AP, ta and te values, tho for R-500 is found to be slightly higher than that of R-502. On the other hand, economic water rate and heat rejection to condenser Q□ho are seen to be higher in case of R-502 for given Rc (except equal to 10) ,AP, ta and te refer Tables 6.1 to 6.3. the detailed graphical presentation is under the scope of future research.

6.2 SYSTEM WITH AIR COOLED CONDENSER

Tables 6-4 to 6-5 can be directly used for preliminary design for preliminary optimum design of the systems operating on R-500 and R-502 using air Cooled Condenser. They also exhibit the quantitative effects of operating variables on the design quantities for a specified set of operating parameters. Not only this, the feasible operating conditions can also be achieved with the help of the figures achieved. The graphical presentation of thermodynamic performance, display of main optimum design quantities for R-500 and R-502 respectively when the systems have been optimized for COP's only over a quite wide range of operating temperature limits, is under the scope of future research.
COP's of R-500 system is observed to be higher than that of R-502 systems (Ref. Tables 6.4 and 6.5)

TABLE 6-1: THE OPTIMUM DESIGN QUANTITIES FOR TWO-STAGE AZEOTROPIC REFRIGERATING SYSTEM INCORPORATING WATER COOLED CONDENSER

Design Parameters : R_c = 3.0   AP = 3℃

Azeotrope	R-502					R-500				
t_e (℃)	t_{io} (℃)	t_{ho} (℃)	$\square_{we}$ (Kg/ton-h)	$Q\square_{ho}$ (KJ/ton-h)	COP_o	t_{io} (℃)	t_{ho} (℃)	$\square_{we}$ (Kg/ton-h)	$Q\square_{ho}$ (Kg/ton-h)	COP_o
-50	-7.66	30.34	595.01	18,289.2	2.21	-12.25	31.13	527.79	18096.5	2.29
-40	-3.01	30.53	547.61	17,280.1	2.69	-3.38	30.95	512.50	17067.8	2.82
-30	1.73	30.73	506.17	16,392.8	3.32	1.46	31.11	477.37	16215.0	3.49
-20	6.55	30.93	469.5	15,602.9	4.20	6.27	31.27	446.45	15462.3	4.40
-50	4.12	49.72	725.90	20,435.5	1.61	2.91	49.94	680.87	19808.5	1.75
-40	8.55	49.92	662.23	19,193.4	1.91	7.43	50.8	628.80	18663.4	2.08
-30	13.06	50.12	606.87	18,106.1	2.29	12.28	50.22	583.64	17663.9	2.49
-20	17.56	50.33	558.34	17,143.3	2.77	16.95	50.36	544.00	16784.4	3.01
-50	16.17	69.13	911.86	23,414.2	1.17	15.10	69.32	831.87	22028.1	1.34
-40	20.27	69.32	824.51	21,817.0	1.37	19.69	69.47	761.94	20659.7	1.56
-30	24.45	69.51	749.28	20,429.9	1.614	24.27	69.63	700.83	19458.9	1.84
-20	28.71	69.71	683.87	19,210.3	1.91	28.81	69.79	647.39	18404.9	2.17

TABLE 6-2: THE OPTIMUM DESIGN QUANTITIES FOR TWO-STAGE AZEOTROPIC REFRIGERATING SYSTEM INCORPORATING WATER COOLED CONDENSER

Design Parameters : $R_c = 3.0$ $t_a = 30℃$

Azeotrope	R-502					R-500				
t_e (℃)	t_{io} (℃)	t_{ho} (℃)	$\square_{we}$ (Kg/ton-h)	$Q\square_{ho}$ (KJ/ton-h)	COP_o	t_{io} (℃)	t_{ho} (℃)	$\square_{we}$ (Kg/ton-h)	$Q\square_{ho}$ Kg/ton-h)	COP_o
-50	-2.44	39.10	645.29	19,184.6	1.91	2.57	40.29	527.79	18096.5	2.01
-40	2.10	39.30	591.51	18,081.3	2.30	2.02	40.57	512.50	17067.8	2.42
-30	6.72	39.50	544.63	17,113.0	2.79	6.83	40.74	477.37	16215.0	2.94
-20	11.43	39.71	503.26	16,252.9	3.45	11.63	40.91	446.45	15462.3	3.63
-50	-1.32	41.03	658.27	19,397.3	1.85	-1.82	41.90	680.87	19808.5	1.97
-40	3.34	41.23	602.83	18,271.2	2.22	2.86	42.09	628.80	18663.4	2.37
-30	7.87	41.44	554.58	17,283.0	2.69	2.69	42.23	583.64	17663.9	2.88
-20	12.50	41.65	512.05	16,405.7	3.31	3.31	42.39	544.00	16784.4	3.54
-50	-0.71	42.0	665.00	19,505.8	1.82	-1.41	42.72	831.87	22028.1	1.95
-40	3.86	42.20	608.76	18,367.9	2.18	3.30	42.89	761.94	20659.7	2.34
-30	8.50	42.41	559.74	17,359.5	2.64	8.14	43.03	700.83	19458.9	2.84
-20	13.03	42.62	516.64	16,483.5	3.24	12.85	42.7	647.39	18404.9	3.49

TABLE 6-3: THE OPTIMUM DESIGN QUANTITIES FOR TWO-STAGE AZEOTROPIC REFRIGERATING SYSTEM INCORPORATING WATER COOLED CONDENSER

Design Parameters : AP= 3.0(℃) t_a = 30℃

Azeotrope	R-502					R-500				
t_e (℃)	t_{io} (℃)	t_{ho} (℃)	$\square_{we}$ (Kg/ton-h)	$Q\square_{ho}$ (KJ/ton-h)	COP_o	t_{io} (℃)	t_{ho} (℃)	$\square_{we}$ (Kg/ton-h)	$Q\square_{ho}$ (Kg/ton-h)	COP_o
-50	3.92	49.48	295.59	20,407.2	1.61	2.90	50.013	277.90	19,806.0	1.75
-40	8.57	49.94	270.49	19,198.4	1.91	7.54	50.34	257.40	18,677.5	2.07
-30	13.2	50.41	248.71	18,138.0	2.28	12.43	50.67	239.05	17,698.2	2.47
-20	17.82	50.89	229.41	17,196.7	2.74	17.29	51.01	223.10	16,832.8	2.98
-50	-2.78	38.51	828.24	19,121.7	1.93	-2.96	39.62	679.22	18,832.1	2.02
-40	1.73	38.67	759.07	18,020.7	2.32	1.62	39.87	617.21	17,765.7	2.44
-30	6.42	38.83	698.75	17,054.8	2.83	6.42	40.02	572.77	16,843.3	2.97
-20	11.00	38.99	645.65	16,196.9	3.50	11.2	40.17	534.11	16032.8	3.67
-50	-3.75	36.92	1153.26	18,953.4	1.98	-4.83	36.49	1271.04	18,564.3	2.11
-40	0.82	37.03	1057.20	17,865.8	2.39	0.11	36.82	1098.81	17,575.4	2.53
-30	5.38	37.15	973.64	16,912.1	2.92	4.54	37.39	904.37	16,616.3	3.14
-20	10.02	37.26	899.93	16,065.0	3.64	9.41	36.79	998.34	15,847.3	3.88

TABLE 6-4 OPTIMUM DESIGN QUANTITIES FOR TWO-STAGE R-500 AZEOTROPIC REFRIGERATING SYSTEM INCORPORATING AIR COOLED CONDENSER

t_h (℃)	t_e (℃)	t_{io} (℃)	r_1	r_2	M_3 (Kg/Kg)	V_1 (m³/kg)	V_3 (m³/kg)	W_T (KJ/Kg)	COP_o
15	-20	-2.22	1.902	1.718	1.14	0.1123	0.0695	26.14	6.71
	-30	-7.00	2.426	2.024	1.19	0.1645	0.846	35.56	4.93
	-40	-12.20	3.146	2.436	1.24	0.2484	0.1055	46.25	3.79
	-50	-16.82	4.320	2.893	1.30	0.3905	0.1297	58.23	3.01
30	-20	5.62	2.455	2.023	1.23	0.1123	0.0584	36.87	4.52
	-30	0.90	3.169	2.356	1.28	0.1645	0.0705	46.90	3.56
	-40	-3.84	4.22	2.760	1.34	0.2484	0.0858	58.23	2.86
	-50	-12.21	5.129	3.704	1.43	0.3905	0.1214	72.60	2.34
45	-20	14.00	3.172	2.292	1.33	0.1123	0.495	46.76	3.36
	-30	9.33	4.141	2.638	1.39	0.1645	0.0594	57.41	2.74
	-40	4.65	5.584	3.064	1.46	0.2484	0.0717	69.47	2.26
	-50	-0.05	7.818	3.557	1.54	0.3905	0.0873	83.07	1.89
60	-20	22.71	4.076	2.503	1.48	0.1123	0.0429	58.99	2.49
	-30	18.11	5.376	2.851	1.55	0.1645	0.0512	70.50	2.09
	-40	13.60	7.353	3.253	1.63	0.2484	0.0613	83.50	1.76
	-50	9.06	10.451	3.733	1.72	0.3905	0.0739	93.17	1.50

TABLE 6-5 OPTIMUM DESIGN QUANTITIES FOR TWO-STAGE R-502 AZEOTROPIC REFRIGERATING SYSTEM INCORPORATING AIR COOLED CONDENSER

t_h (℃)	t_e (℃)	t_{io} (℃)	r_1	r_2	M_3 (Kg/Kg)	V_1 (m³/kg)	V_3 (m³/kg)	W_T (KJ/Kg)	COP_o
15	-20	-2.02	1.84	1.65	1.169	0.0611	0.0396	21.97	6.46
	-30	-6.80	2.31	1.92	1.221	0.0877	0.0481	29.86	4.76
	-40	-11.60	3.00	2.25	1.279	0.1302	0.0588	38.69	3.67
	-50	-16.42	4.05	2.66	1.345	0.2004	0.0726	48.73	2.92
30	-20	6.02	2.34	1.84	1.278	0. 0611	0.0340	30.91	4.29
	-30	1.40	2.99	2.12	1.339	0.0877	0.0408	39.31	3.37
	-40	-3.42	3.95	2.44	1.404	0.1302	0.0494	48.74	2.72
		-7.91	5.41	2.84	1.481	0.2004	0.0604	59.49	2.23
45	-20	14.40	2.98	2.04	1.431	0. 0611	0.0299	39.68	3.09
	-30	10.03	3.86	2.30	1.499	0.0877	0.0355	48.64	2.52
	-40	5.55	5.17	2.62	1.577	0.1302	0.0425	58.74	2.08
	-50	1.15	7.21	2.99	1.667	0.2004	0.0513	70.24	1.74
60	-20	23.21	3.77	2.19	1.652	0. 0611	0.0267	48.76	2.29
	-30	18.91	4.94	2.45	1.736	0.0877	0.0318	58.47	2.91
	-40	14.70	6.71	2.75	1.830	0.1302	0.0379	69.34	1.61
	-50	10.56	9.52	3.09	1.938	0.2004	0.0452	81.74	1.36

7. CONCLUSSIONS:

1.For a preliminary design of two stage azeotropic refrigerating system, the Tables 6-1 to 6-5 presented can directly be used.

2. Economic water rate and heat transfer to condenser turns out to be relatively lower in case of R-500 for a given set of condenser, evaporator, ambient and approach temperatures and cost ratio.

3. R-500 system produces comparatively higher COP than R-502 system for specified operating conditions.

4.The effect of approach temperature is more pronounced on the economic water rate than the other quantities. It should be selected quite carefully.

5.Though, the initial investment in case of R-500 system turns out to be more than R-502 system, it would get compensated over a small span of time because of lower operating cost of the R-500 system.

.

8. REFERENCES:

1. Buchler,Leon, Economical use of condense water for Amonia Compression refrireration systems, Ice and Refrig, January,1947.

2. Bochmer,Andrew P., Condenser pressure for Air Conditioning,Heating,Piping   and air-Cond., p77 and 94,1946.

3. Buchler,Leon, Andrew P., economical of condenser water in compression refrigeration system, journal ASRE, 57,pp. 251-254 March,1949.

4.Mcharness,R.C. and chapman, D.D., Refrigerating capacity and performance data,ASRE J., pp 49-58, January,1962.

5.Downing,R.C. Mixed refrigerants, Service Manual RT-38, El. Du Pont Nemours company., Wilmington, Delaware U.S.A.

6. Gupta, V.K and Prasad, Manohar, Numerical Estimation of main parameters for realistic Two stage ammonia refrigerating systems, J. Indian Innstitute of Science Bangalore,64,10,pp.219-227,1983.

7. Jordan ,R.C. and Priester, G.B.refrigeration and Air conditioning, Prentice-Hall of India, New delhi,1969.

8. Prasad, Manohar, Refrigeration and air Conditioning, Wiley Eastern, New Delhi,1983.

9.Gupta, V.K., Thermodynamic design data for optimum performance of Two Stage Azeotropic Refrigerating systems, J.Heat Recovery Systems (U.K)

10. Kern,D.Q., Process Heat Transfer, Mv Graw-Hill, Kogakusha, Tokyo,1960.

11.Mahindru, D.V., Economic Water rate and Optimum Performance of Two
   Stage Azeotropic Refrigerating Systems, a Thesis submitted to Kanpur
   University in partial fulfillment of the requirements   for the degree of Master of
   Technology, Mechanical Engineering (Design) ,1985.

9. NOMENCLATURE:

AP   : Approach℃
C   : Cost of Electricity (Rs/kw-h)
c_w : Cost of Water (Rs/k- litre)
COP_o: Maximum co-efficient of performance.
h_x : Enthalpy per unit mass at state condition ' x' (kJ/kg)
$\square_1$: Refrigerant mass flow through LP compressor 9kg/ton-h).
$\square_3$: Refrigerant mass flow through HP compressor (kg/ton-h)
$\square_{we}$:Economic water rate (kg/ton-h)
M_3 : Refrigerant mass flow through HP compressor per unit mass
      flow through Lp compressor(kg/kg).
P   : Power 9kw/ton)
P_o : Minimum power (kw/ton)
ΔP : Power increase per ton per degree rise in condensing

      Temperature (kw/ton-℃)

P_h : Condensing pressure (bar)

P_e : Evaporating pressure (bar).

P_{io} : Optimum inter-stage pressure (bar).

Q_h : Heat transfer to condenser (kJ/ton-h).

Q_{ho} : Heat transfer to condenser pertaining to minimum
         operating cost and maximum COP (kJ/ton-h)

Q_c : cooling effect per unit mass flow through LP compressor
         (kJ/kg).

r_1 : Compression ratio (P_{io}/P_e) of LP compressor.

r_2 : Compression ratio (P_h/P_{io}) of HP compressor.

R_c : Cost ratio (c/c_w).

s_x :Entropy per unit mass at state condition (kJ/kg-k)

t_a, T_a : Ambient temperature (℃,K)

t_{w1}, T_{w1}: Inlet water temperature (℃,K)

t_h, T_h : condensing temperature (℃,K)

t_e, T_e :evaporating temperature (℃,K).

t_{ho}, T_{ho} : Optimum condensing temperature(℃,K).

t_i, T_i : Interstage temperatures (℃,K)

t_o, T_o : Outlet water temperature (℃,K)

t_{io}, T_{io} : Optimum interstage temperature(℃,K).

V_1 : Minimum volume flow through LP compressor per unit
         mass flow through LP compressor (m^3/kg).

V_3 : Minimum volume through HP compressor per unit mass
         flow through LP compressor (m^3/kg).

W_1 : compression work of LP compressor (kJ/ton-h).

W_2 : compression work of HP compressor(kJ/ton-h).

W : Total compression work (kJ/ton-h).

W_T : Total compression work (kJ/ton-h).

Speed Scheduling of Autonomous Railway Vehicle Control System Using ANN

Nida Anwar, Muhammad Saleem Khan, Khalil Ahmed, Aiesha Ahmad and Atifa Athar

Abstract— This paper presents the artificial neural network base design approach for speed scheduling of autonomous railway vehicle control system the proposed system enhances the efficiency of vehicle under constraints and contributes in controlling issues, considering relative safety and operational requirements to avoid accidents. This design approach provides the vehicle's speed management and scheduling under certain conditions: junction track information JTI, environment monitoring EM, vehicle tilting VT, track condition TC, track clearance TCL. The proposed system will be capable to full fill the needs of modern trends, to enhance the vehicle control automation system.

Index Terms— ANN, Autonomous Railway Vehicle System, Spped Scheduling.

———————————— ◆ ————————————

1 INTRODUCTION

Nowadays human factor plays a significant role in control of railway system in a whole and a rolling stock in particular. Over the last three decades railway vehicle control technology is improving and in some countries it is being partially implemented. The modern technologies has reviewed the subject of Mechatronics and Electronics for the development of intelligent devices and controllers that have interfaces to work in global network, wireless networks and are programmed to use methods of artificial intelligence. Intelligent devices have possibilities to negotiate with each other and to coordinate their work to get better decisions. Some crashes with harmful outcomes happened in Latvia in the last three years. Main reason is the driver's inattention passing the read signals. The task is to prevent such accidents by reducing the human factor [1]. The general benefits and features of control and monitoring for the railway and transportation vehicle are well understood.

The agent base technology is needed in Mechatronics to achieve flexible, network enabled collaboration between-centralized and distributed system for high level auto-

nomous vehicle control [2]. Traditional ways of predicting railway safety are very expensive in term of time consuming which make them inefficient under certain circumstances. The advance concepts for monitoring and control of railway vehicle dynamics can provide better performance to facilitate speed scheduling and management. The proposed autonomous system gives solution for track condition with the acceptable ride quality over lower quality tracks, and tilting of trains around the track with speed adjustment [3]. The behavior of the train on the track at any time is not only affected by the state of the track at that time but also affected by the status of the track [4]. This leads us to include track parameters from some prior points as input to the neural network along with the parameters of the current point. Different neural network structures and techniques have already been created and tested to find the best structure for predicting railway safety [5]. The design model of multi-agents based autonomous railway vehicles control system has already been proposed that provides the vision of highly accurate system for autonomous control of railway vehicles, as that system is based on heavy information communication, therefore proper coding and filtering techniques need to be exercised [6]. The proposed system implements real time telemetry services for railway system using the information of equipment for rail, monitoring the signals of junction track information JTI, crossing gates, track condition TC, Track clearance TCL, environment monitoring EM, vehicle tilting VT, on track vehicle to vehicle and between vehicle and main control information exchange this information is helpful for vehicle speed control and scheduling [7]. The main innovation behind is to have the track irregularities as input and we can then predict to solve which part of the track is safe to use and which part is not. The true power of neural networks lies in their ability to represents both linear and non-linear relationships and their ability to learn these relationships

• **Nida Anwar** is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan. She is currently working as faculty member in VU, Lahore, Pakistan (e-mail: nidaanwar85@gmail.com).
• **Dr. M. Saleem Khan** is with the Computer Science Department as Director in GC University Lahore, Pakistan (e-mail: mskgcu@yahoo.com).
• **Dr. Khalil Ahmed** is with the School of Computer Science at NCBA&E Lahore, Pakistan; He is an expert academician and passionately engaged in research. (e-mail: drk@ncbae.edu.pk).
• **Aiesha Ahmad** is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan (e-mail: aiesha@ncbae.edu.pk).
• **Atifa Athar** is working as a faculty member at the School of Computer Science, NCBA&E Lahore, Pakistan. She is a research fellow at SCS, NCBA&E (e-mail: atifaathar@yahoo.com).

directly from data being modeled. Another important feature of Neural Networks is in their intrinsic fault tolerance, even if some neurons fail, the network can perform well because knowledge is distributed across all the elements of the network.

The frame work of this paper is: Section 2 describes the brief introduction of adaptive networks which includes fuzzy inference system (FIS) and Artificial Neural Network (ANN). Structure of proposed speed scheduling system is discussed in section 3, Section 4 describes the neural network based intelligent speed scheduling system and section 5 explains the design algorithm of the proposed system. Section 6 gives the results and discussion. Acknowledge of this work is mentioned in Section 7.

2 ADAPTIVE NETWORKS

2.1 Fuzzy Systems

The fuzzy inference system is a popular computing frame work based on the concept of fuzzy set theory, fuzzy if-then rules, and fuzzy reasoning. It has been successfully applied in fields such as automatic control, data classification, decision analysis, expert system and vision, because of its multi-disciplinary nature, the fuzzy inference system is known by a number of names such as "fuzzy-rule-base system", "Fuzzy Expert system"[8]. The basic structure of a fuzzy inference system consists of three conceptual components: a rule base which contains selection of fuzzy rules a database or dictionary, which defines the membership function used in fuzzy rules, and a reasoning mechanism, which perform the inference procedure upon the rules and a given condition to derive a reasonable output or conclusion. The actual purpose of fuzzy inference is mapping of input to a given output by using fuzzy logic that provides the base for better, quicker and more accurate decisions/output (s) as compared to traditional approach. Fuzzy inference system is not only capable to deal with knowledge in linguistic rules from human experts but also improve performance by adapting itself using numeric data [9].

2.2 Neural Network

This section describes the architecture and learning procedures of adaptive networks, which are a super set of all kind of neural networks with supervised learning capabilities. As the name implies adaptive network is a network structure whose overall input-output behavior is determined by the values of a collection of modifiable parameters. More specifically, the configuration of an adaptive network is composed of a set of nodes connected through directed links, where each node is a process unit that performs a static node function on its incoming signals to generate a single node output and each link specifies the direction of signal flow from one node to another.

Usually a node function is a parameterized function with modifiable parameters; by changing these parameters, we are actually changing the node function as well as the overall behavior of the adaptive network. The data set is usually called the training data set and the procedure we follow in adjusting the parameters to improve the performance of the network is known as learning rules [8]. Many NN models and learning algorithms have been proposed. Typical network structures include feedback and feed-forward NNs. Learning algorithms are categorized into supervised learning and unsupervised learning [10].

3 BLOCK DIAGRAM OF PROPOSED STRUCTURE OF THE RAILWAY VEHICLE SPEED CONTROL SYSTEM

The development to enhance the capabilities of the train control system is under progress at various platforms. The proposed system is based on: the starting station, initial time, complete root chart loading with all the utilities, observations; junction track change information JTI, crossing gate information CG, track clearance TCL, track condition TC, environment monitoring EM, vehicle tilting VT, request from non-stop stations, vehicle to vehicle, and from vehicle to control room communication facilities, autonomous decisions and operations. All this information is help full for scheduling and controlling vehicle speed. Junction track information JTI, crossing gate information CG, and track clearance TCL are the hard conditions, in the presence of anyone of them makes the vehicle to stop. Fig. 1 shows flexible conditions: VT, TC and EM for speed scheduling system. The hard conditions: JTI, CG and TCL for speed scheduling are shown in Fig. 2.
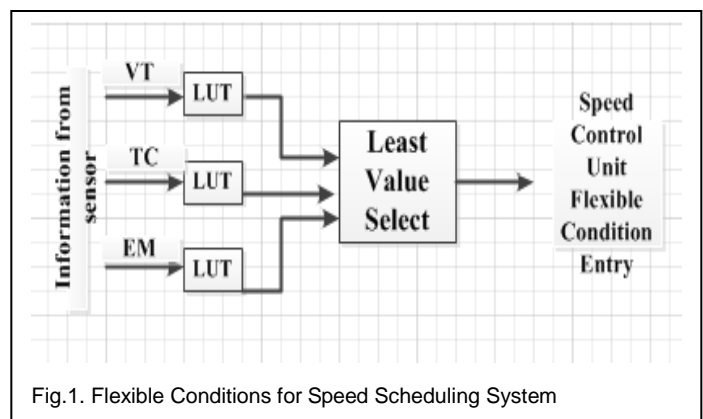

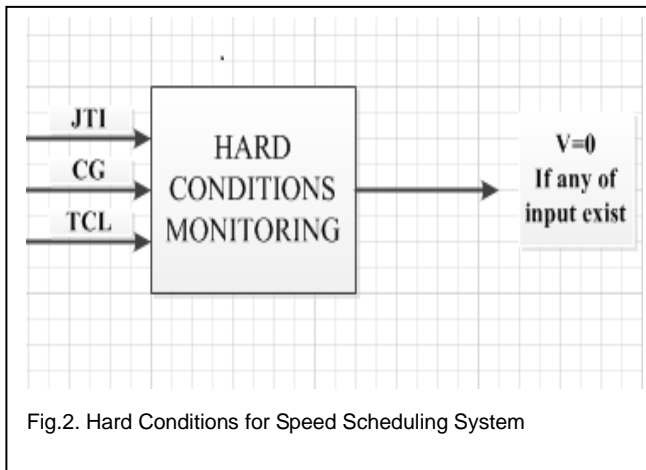Fig.1. Flexible Conditions for Speed Scheduling System

Fig.2. Hard Conditions for Speed Scheduling System

For vehicular traffic control on a track or on adjacent tracks joining at junction, conditions for normal traffic or rear-end collision must be examined. A mathematical modular approach is used for the safe vehicle speed adjustment to reduce any time delay due to any reason i.e. hard or flexible conditions in the journey [6]. The characteristics of mathematical formulation for normal vehicle traffic encountered in practice are proposed for proper speed control and adjustment. The possible ways to increase the speed and improve the synchronization of trains are required to be adopted for the stability of networks [11]. This model has the capacity through its speed scheduling unit to include the consideration about random delays of trains and the propagation of delay across the network.

Fig.3. Block Diagram of Speed Scheduling Control Unit

This discussion shows that the final speed adjustment in these types of special cases requires the information exchange between vehicles in automation, and another option proposed to be used is to post the sensors at the front end or the rear end of the train to collect the information for speed estimation. The inputs and outputs for speed scheduling control unit are shown in Fig. 3.

4 NEURAL NETWORK BASED PROPOSED STRUCTURE OF THE RAILWAY VEHICLE SPEED CONTROL SYSTEM

The proposed structure of speed control system is composed of 3 layers: input layer, hidden layer and the output layer. The input data is presented to the input layer that is connected with the hidden layer; and the last hidden layer is connected to the output layer. In each layer each is neuron is independent of each other and the weights between each layer represents the knowledge of ANN. ANN get these weights by training which means that ANN must be trained before it contains useful knowledge. In the proposed structure neural network get trained by flexible conditions that will control the speed of the vehicle and these flexible conditions door function with hard conditions i.e. if any one of the hard condition get occur will make the vehicle to stop. Fig.4 shows the NNW based proposed structure of railway vehicle speed control system. This training of neurons will be helpful in future when the same condition occurs. Control part of rolling stock is represented as neural network mathematical model.

Fig.4. Neural Network Base Proposed Structure of Railway Vehicle Control System.

Input data set for neural network:

$$X = \{x1, x2 \ldots xn\};$$

ANN had hidden layers:

$$L = \{l1, l2 \ldots lk\};$$

Neural network output:

$$Yij\ [Speed]$$

Weights for each input of ith neuron of jth layer:

$$W_{ij} = \{w_{i1}, w_{i2} \ldots w_{in}, w_{in}\};$$

Bias for each i-th neuron of j-th layer: **bi j;**

Input summation function for each i-th neuron of j-th layer:

$$Y_{ij} = \sum (W_{ij} * X) + b_{ij};$$

5 DESIGN ALGORITHM

This system is designed for two input variables: that are Hard Conditions and the Soft conditions. Hard Conditions consists of further three input variables i.e. JTI, CG, TCL, occurrence of any one of them makes the

TABLE 1

RANGES OF FLEXIBLE HARD CONDITIONS

MEMBERSHIP FUNCTIONS

	Flexible Conditions			Hard Conditions			
Ranges	TC	VT	EM	JTI	CG	TCL	STOP
0-50	Good	Low	Pleasant	N	N	N	STOP
0-100	Average	Medium	Average	N	N	Y	STOP
50-100	Worst	High	Windy	Y	Y	N	STOP

railway vehicle to stop, whereas Soft Conditions are Vehicle Tilting (VT), Track Condition (TC) and Environment Monitoring (EM) are used for controlling and adjustment of railway vehicle speed in km/h.

Table-1 and Table-2 show the specified ranges for the adjustment of final speed according to the input parameter ranges and their respective memebership

TABLE 2

RANGES OF OUTPUT VARIABLE

MEMBERSHIP FUNCTIONS

Speed Ranges (km/h)	Membership Function
0-20	Very Slow
0-40	Slow
20-60	Average
40-80	Normal
60-100	Fast
80-120	Very Fast

function.

6 RESULT AND DISCUSSION

For neural network based vehicle control system shown in Fig.4, controlling part of rolling stock contain 10 neurons in the hidden layer that gives the adjusted speed of the vehicle. Back propagation neural network technique is being used; the term back propagation refers to the way the learning procedure is being performed, i.e. by propagating gradient information from the network's outputs to its inputs. Each node in such network has the same node function, that is the combination of linear and sigmoidal function and the network is trained by using the Levenberg Marquardt algorithm. `trainlm` is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization. Trainlm is often the fastest back propagation algorithm and is highly recommended as a first-choice supervised algorithm. ANN controller takes input from the sensor and its outputs adjusted speed signals from actuators. The central part of the learning rule for adaptive network concerns how to recursively obtain a gradient vector in which each element is defined as the derivative and error measure with respect to parameter i.e. change in the input parameter will make changes in the related output node containing specified input parameter, until the error rate is not equal to zero, zero error rate means that network is able to adapt itself according to the input parameters and is able to produce exactly the desired output vector in the training data pair. As such method is generally referred to as "Back Propagation Learning Rule" because the gradient vector is calculated in the opposite direction to the flow of output signal.
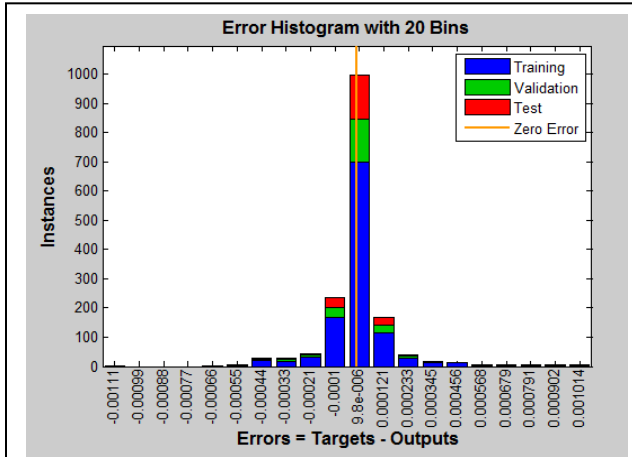
Fig.5 Error histogram shows the mimum error rate after training, validating and testing the data set
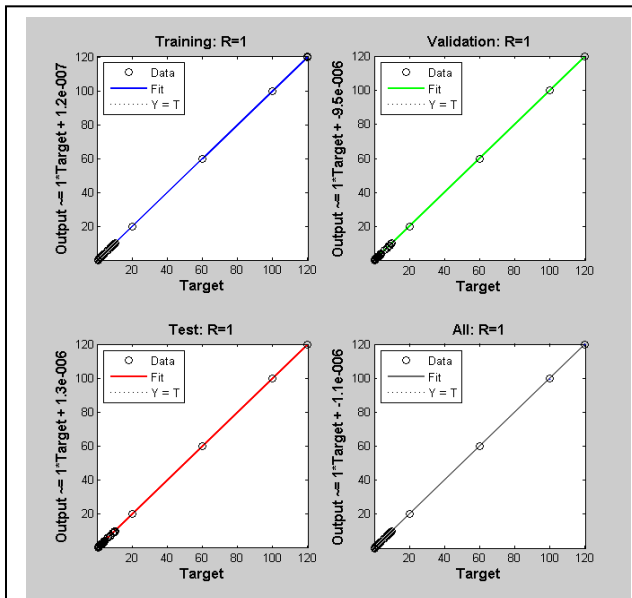
Fig.6 Regression graph of training, validating and testing of data set with respect to the target output

After doing the MATLAB simulations on data set the results have been generated. Fig. 5 shows error histogram graph of minum errors at the training, validation and testing stages which is 2.26038e-8, 2.97033e-8and 2.86677e-8 respectively with repect to the target/output and instances.
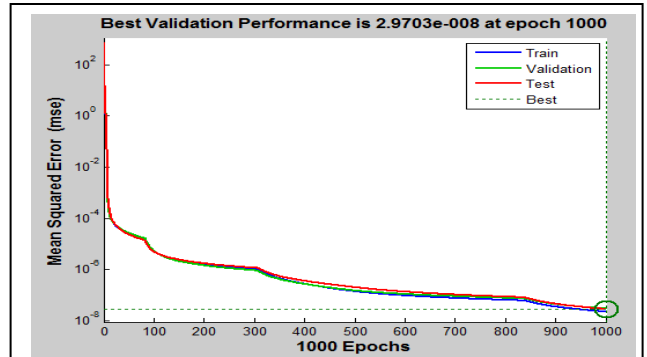

Fig.7 The best validation perormance graph of the network with repect to mean square error that is minimun after 1000 epoch
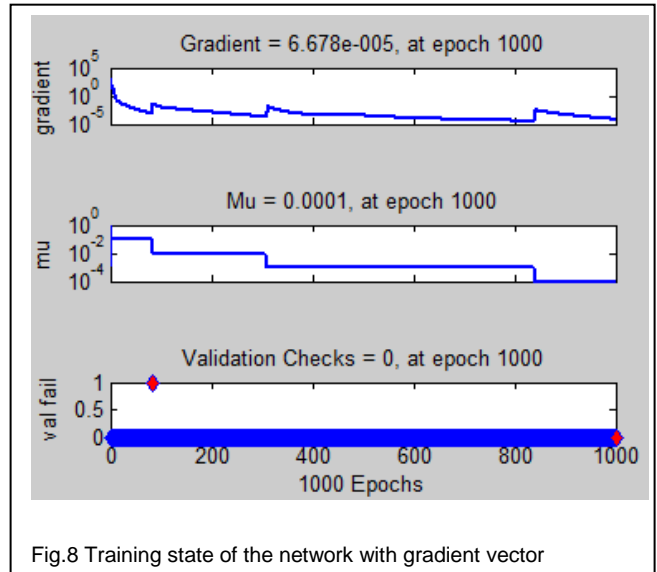
Fig.8 Training state of the network with gradient vector

Regression Graph shows the closeness between input and the output parametrs. Fig.6 show the regression graphs of data training ,its validation and testing, indivisualy and altogether results are shown in one separate graph. Data is best validated at 1000th epoch with repect to minimum means square error that is given in Fig. 7. Fig.8 explaines overall training state of the network with minimum gradient vector.

7   CONCLUSION AND FURTURE WORK

This work provides the vision of flexible architecture for the autonomous control of railway vehicle. The main objective of this paper is to develop ANN that is capable of predicting safety measurements of trains on railway

track. The experimental results shows that input data is get trained, validate and tested and able to predict the desired output with very low error. Minimum error rate shows that predictive cautions can be taken to minimize the accidents; ANN can be used as tool in taking decisions. Errors can be minimized but not eliminated. The hardware for speed scheduling and management system can be developed in future using FPGAs based Microelectronics chips to provide the state of the art solution for this autonomous real time control system.

ACKNOWLEDGMENT

REFERENCES

[1] Mikihail Gorobetz, l. R. (2009). Method of Accidents Prevention using Intellegent Control system for Railway Transports.*Proceedingd of the 6th International Scientific Conference TRANSBALTICA 2009*.

[2] R.Harrison, S. (2006). Reconfigurable modular automation system for automotive. *Springer Science 1 March 2007*, 175-190.

[3] Stefano Bruni, T. (2007). Control and Monitoring for Railway vehicle Dynamics. *45*, 743-779.

[4] L.Vulliet, F. a. (2002). Neural Networks for slope movement prediction. *International Journal of Geomechanics*, 153-173.

[5] S.Nefti, M. (2004). A Neural Network approach for Railway safety Rediction. *IEEE international Conference on systems, Man and cybernetics*.

[6] M. Saleem Khan and Khaled Benkrid" Design Model of Multi-Agents Based Autonomous Railway Vehicles Control System" Lecture Notes in Engineering and Computer Science Vol. 2175 2009, p.p. 1260-1265.

[7] Ben Schiller. (2008). Model approach to Running a Railway".

[8] JYH-Shing Roger Jang, C.-T. S. (1995). Neuro Fuzzy Modelling and Control. *Proceeding of IEEE*.

[9] Maaref, K. M. (2009). "INTELLIGENT CONTROL FOR A DRONE BY SELF-TUNABLE FUZZY INFERENCE SYSTEM". *6th int. Multi-conference on Systems, Signals and Devices,IEEE,*.

[10] TAKAGI, H. (1997). Introduction to Fuzzy Systems, Neural Networks, and Genetic Algorithms. *Kluwer Academic Publishers, Norwell,Massachusetts, USA*, 1-33.

[11] Funke, O. . (2004). Analysis stability and investment in railway networks using advance evolutionary algorithms. *Int. Trans.Oper.Res*, 381.

Online Graphical Display of Blood Oxygen Saturation and Pulse Rate

Dilpreet Kaur, Sukhwinder Kumar, Shashi Sharma

Abstract—This paper will design a non-invasive pulse oximeter using the ADuC842 microcontroller. A pulse oximeter is a medical device that indirectly monitors the oxygen saturation of a patient's blood and heart rate. The hardware has been developed for the pulse oximeter and programming/coding has been done for calculating blood oxygen saturation and pulse rate of a patient. The results are displayed on the OLED or transported to PC. The software used is ASPIRE (Advanced Systems Programming Integrating Real-time Emulation) verion 1.05. Assessing a patient's need for oxygen is the most essential element to life; no human life thrives in the absence of oxygen.

Index Terms— Heart rate, microcontroller, oxygen saturation, pulse oximeter.

————————————— ◆ —————————————

1 Introduction

In the earlier days, the common method used to measure blood oxygen saturation was arterial blood gas measurement. An Arterial Blood Gas (ABG) is a blood test that is performed using blood from an artery. It involves puncturing an artery with a thin needle and syringe and drawing a small volume of blood. This method was invasive, expensive, difficult, painful and potentially risky.

The pulse oximeter [1] [5] was introduced in the early 1980s. It is particularly a convenient, non-invasive measurement instrument. Pulse oximeter is less expensive compared to the old health monitoring devices, simple to use as it needs no user calibration, small enough to be wearable and accurate enough for clinical use. For these reasons, in almost every hospital, critical care units and surgical theatres; the pulse oximeter is acknowledged as a standard monitoring device.

Pulse oximeter is recommended for the monitoring of patients during anesthesia or those with conditions such as asthma. It is used to assess the viability of limbs after plastic and orthopaedic surgery. It is recommended as an important care tool for new born infants and patients during surgery. Pulse oximeters are also useful for pilots operating in a non-pressurized aircraft above 10,000 feet where supplemental oxygen is required. It is also useful for mountain climbers and athletes whose oxygen levels may decrease at high altitudes or with exercise.

2 Overview

The principle [6] of pulse oximetry is based on the red and infrared light absorption characteristics of oxygenated and deoxygenated haemoglobin. Oxygenated haemoglobin absorbs more infrared light and allows more red light to pass through whereas deoxygenated (or reduced) haemoglobin absorbs more red light and allows more infrared light to pass through. Red light is in the 600-750 nm wavelength light band whereas infrared light is in the 850-1000 nm wavelength light band. The absorption relationship of oxygen levels in the blood for the red and infrared wavelengths is shown in figure 1.
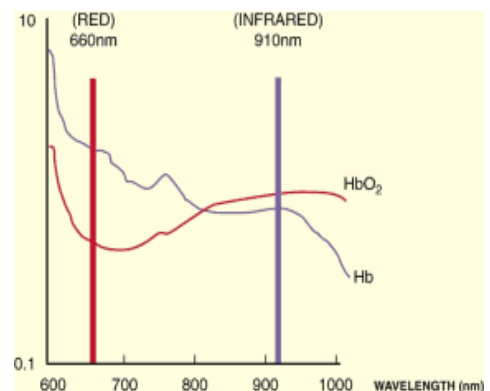
Figure 1: Absorption relationship of oxygen levels in the blood for the red and infrared wavelengths

—————————————————

- *Dilpreet Kaur is currently pursuing masters degree program in electronics and communication engineering at Institute of Technology and Management, Gurgaon, Haryana, India. E-mail: er.dilpreetkaur@gmail.com*

There are two methods [2] of sending light through the measuring site, transmission and the reflectance. In the transmission method, the emitter and the photo detector are opposite of each other with the measuring site in-between. The light can then pass through the site. In the reflectance method, the emitter and photo detector are next to each other on top of the measuring site. The light bounces from the emitter to the detector across the site.

With each heart beat, the heart contracts and there is a surge of arterial blood, which momentarily increases arterial blood volume across the measuring site. This results in more light absorption during the surge. If light signals received at the photodiode are looked at 'as a waveform', there should be peaks with each heartbeat and troughs between heartbeats.

A newborn's heart rate [4] [7] is typically around 120 beats per minute (bpm). A heart rate in the vicinity of 70 beats per minute (bpm) is considered normal for an adult. When a person enters his golden years, the heart rate slows to approximately 50 bpm. When exercising, the heart rate may double. Accounting for all of this data, to say, 50 to 200 bpm are considered good readings for the heart rate.

For the blood oxygen saturation [3] [9], the following table must be taken into account:

TABLE 1

$S_P O2$ INTERPRETATION

SpO_2 Reading (%)	Interpretation
95-100	Normal
91-94	Mild Hypoxemia
86-90	Moderate Hypoxemia
<85	Severe Hypoxemia

Where hypoxemia is defined as decreased partial pressure of oxygen in blood and low oxygen availability to the body or an individual tissue or organ.

3 Work done

The block diagram of pulse oximeter is shown in figure 2. It consists of seven stages namely; PPG probe, switching and LED driver circuit, I/V converter, amplifier stage, ADC stage, microcontroller and LCD display.
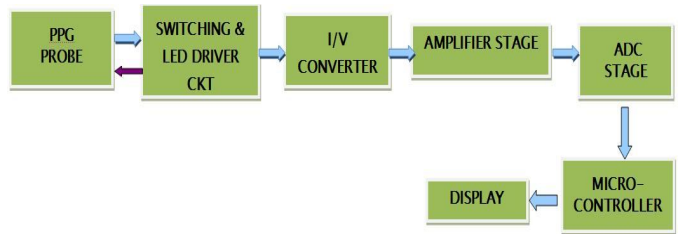
Figure 2: Block diagram of pulse oximeter

1. PPG Probe: A PPG (Photoplethysmograph) probe [8] senses the PPG signal from the flow rate of our blood and converts it into the electrical signal with the help of receiver and the source. The signal produced upto here is from 0 to 1.5 volts. The probe model number used is DS-100A manufactured by Nellcor. It is a reusable sensor for spot checks or short-term monitoring. PPG probe has adult finger sensor which is used for pulse oximeters. A DB9 connector is connected to the probe which has 9 pins. The application site of the probe includes ear, head, finger and toe. The cable length is 3ft/0.9m - 10ft/3.0m. There are various features of the PPG probe such as easy to use and design, built in shielding which protects the signal from electronic noise, high quality LEDs that maximize tracking capabilities and sensor head design which optimizes the signal and shields the detector from ambient light. Figure 3 shows the Nellcor SpO2 sensor.
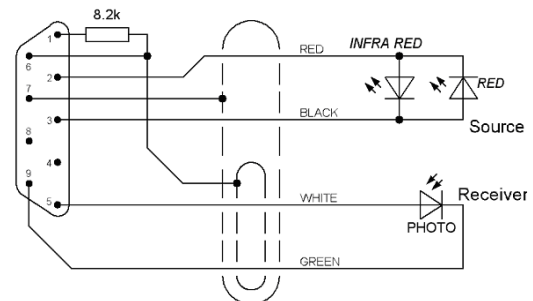
Figure 3: Nellcor SpO2 sensor

2. Switching and LED driver circuit: In switching and LED driver circuit, there are two transistors and two NOT gates which are used for switching the LEDs. IC 7404 is used as NOT gate. When one LED is on, the other will be off and vice-versa. Thus, when red LED is on, infrared LED will be off and when infrared LED is on, red LED will be off. Red LED corresponds to oxy-

genated haemoglobin whereas infrared LED corresponds to deoxygenated haemoglobin.

3. I/V converter: The signal obtained upto here is in the form of current. We have to change this signal to its voltage form. For this we need, current to voltage converter. A current to voltage converter is an electrical device that takes an electric current as an input signal and produces a corresponding voltage as an output signal. It is also known as trans-impedance amplifier. IC UA741 is used as a current to voltage converter.

4. Amplifier stage: The signal obtained upto here is of the order of 0 to 1.5 volts. In order to amplify the signal, we need an amplifier. An amplifier amplifies the signal. IC UA741 is used as an amplifier with two potentiometers to adjust the dc level of the signal. We have to set the gain for both the potentiometers such that the signal is above ground level and no clipping of the waveform is there in both the positive as well as the negative half of the cycle.

5. ADC stage: ADC stands for analog to digital converter. The signal obtained is in the analog form. In order to give the signal to microcontroller, we have to change the analog signal to its digital form. We are using ADuC842 microcontroller which has on-chip ADC. The key features of ADC include on-chip reference of 2.5V, single supply, fast, 8 channels, 12-bit resolution, multi-channel mux, track and hold, calibration features etc. The ADCCON1 and ADCCON2 are two SFRs that need to be configured while writing the program. The ADC converter consists of a conventional successive approximation converter. The converter accepts an analog input range of 0V to Vref. The external reference can be in the range of 1V to AVdd.

6. Microcontroller: The microcontroller used is ADuC842. The features of ADuC842 mainly include increased performance i.e. single-cycle 20 MIPS 8052 core and high speed 420 kSPS 12-bit ADC, increased memory i.e. up to 62 kBytes on-chip Flash/EE program memory and 4 kBytes on-chip Flash/EE data memory, in-circuit reprogrammable of 2304 bytes on-chip data RAM, smaller package i.e. 8 mm × 8 mm chip scale package and 52-lead PQFP, 8-channel, 420 kSPS high accuracy, 12-bit ADC, DMA controller, high speed ADC-to-RAM capture, two 12-bit voltage output DACs, dual output PWM Σ-Δ DACs, on-chip temperature monitor function, 8051 compatible instruction set (20 MHz max), 32 kHz external crystal, on-chip programmable PLL, 12 interrupt sources, 2 priority levels, dual data pointers, extended 11-bit stack pointer, on-

chip peripherals like time interval counter (TIC) and watchdog timer (WDT), core clock of 2.098 MHz. Figure 4 shows the ADuC842 microcontroller card which is used in the circuit.

Figure 4: ADuC842 microcontroller card

7. Display: The display used for showing the values for blood oxygen saturation and pulse rate along with the waveform is OLED display. An Organic Light Emitting Diode (OLED) is a LED in which the emissive electroluminescent layer is a film of organic compounds which emit light in response to an electric current. OLED is a new technology and has features like low thermal dissipation, high pixel density, excellent clarity, thinner and lighter than LCDs etc. In dark rooms, an OLED screen can achieve a higher contrast ratio than an LCD. We have used μOLED-32028-P1 in the circuit which has features like screen outline of 49.1 x 67.3 mm, active area of 43.2 x 57.6 mm, near 180 degree viewing angle, diagonally its size is of 2.83″, PCB size is of 49.1 x 67.3 x 11.0mm, voltage supply from 4.5V to 5.5V, current @ 90mA nominal when using a 5.0V supply source, powered by the fully integrated PICASO-GFX Graphics Processor etc.
Figure 5 shows the OLED display which is used in the circuit.
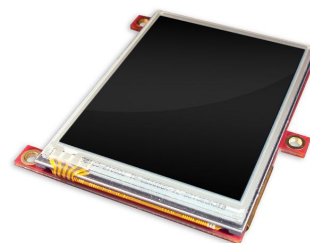

Figure 5: OLED display

We have interfaced the ADuC842 microcontroller with the PICASO-GFX microcontroller which is integrated on the OLED display.

Figure 6 shows the bread board testing of the circuit before the final implementation on the PCB.
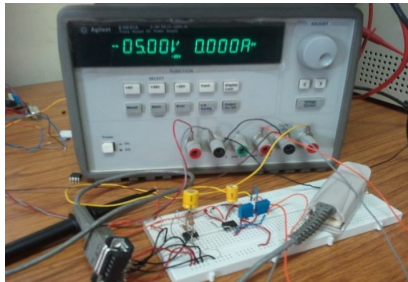

Figure 6: Bread board testing of the circuit

After making the hardware, next step was to write the code for the circuit. The code was written in assembly language and the PICASO libraries were added in the code. Figure 7 shows the flowchart for the calculation of blood oxygen saturation and heart rate of a patient.
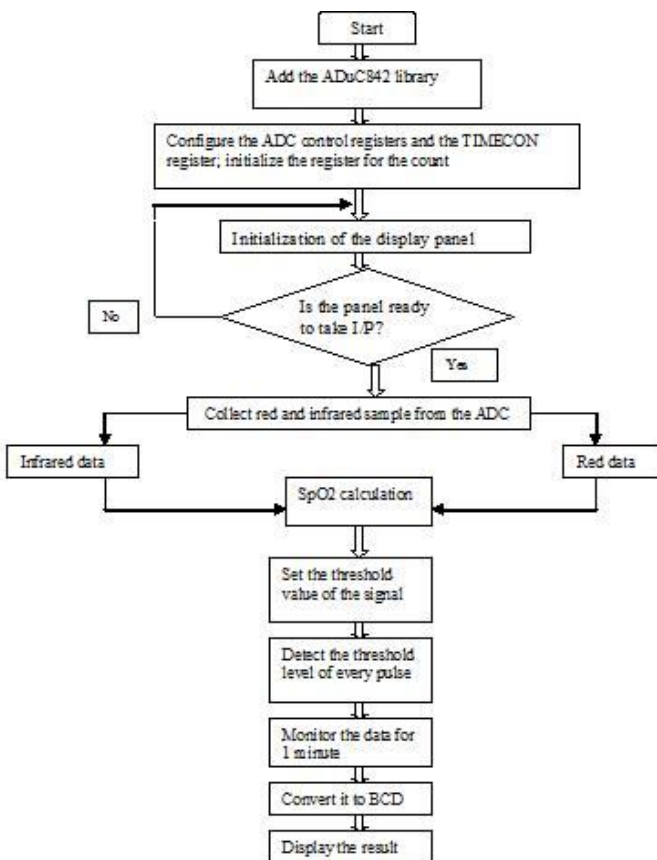
Figure 7: Flowchart for the calculation of SpO2 and heart rate

Blood oxygen saturation (SpO2) can be calculated according to the following formula:

$$SpO2 = HbO2 \times 100\% \,/\, Hb + HbO2$$

where SpO2 is the percentage of blood oxygen saturation

HbO2 is the amount of oxygenated hemoglobin
Hb is the amount of deoxygenated hemoglobin
Hb+HbO2 is the total hemoglobin

4 RESULTS

The result is the blood oxygen saturation and heart rate in digital values as well as the waveform on the OLED display. Figure 8 shows the waveform for the heart rate.
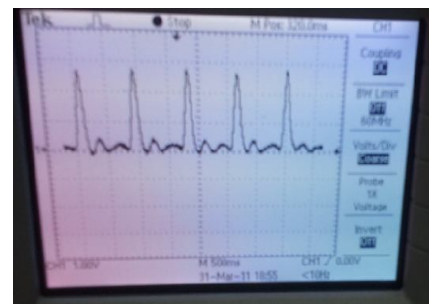
Figure 8: Waveform for the heart rate

5 CONCLUSION

The pulse oximeter is a very useful device for monitoring patients during anesthesia, intensive care, emergency departments, general wards or those with conditions such as asthma. The system uses the ADuC842 as a microcontroller and has very low power consumption. The system has attractive features to measure the SpO2 and the pulse rate. The programming is done in assembly language. The display panel used is OLED which is a new technology and has high pixel density.

REFERENCES:

[1] Cho Zin Myint, Nader Barsoum, Wong Kiing Ing, "Design a medicine device for blood oxygen concentration and heart beat rate",

Global Journal on Technology and Optimization, June 2010.

[2] Maziar Tavakoli, Student Member, IEEE; Lorenzo Turicchia, and Rahul Sarpeshkar, Senior Member, IEEE, "An Ultra-Low-Power Pulse Oximeter Implemented with an Energy-Efficient Transimpedance Amplifier", IEEE transactions on Biomedical circuits and systems, VOL. 4, NO. 1, February 2010.

[3] Pawan K. Baheti, Harinath Garudadri, "An Ultra low power pulse oximeter sensor based on compressed sensing", IEEE Computer Society, 2009.

[4] Susannah Fleming, Lionel Tarassenko, Matthew Thompson, and David Mant, "Non-invasive Measurement of Respiratory Rate in Children Using the Photoplethysmogram", 30th Annual International IEEE EMBS Conference, August 2008.

[5] Stephan Reichelt_, Jens Fiala, Armin Werber, Katharina Förster, Claudia Heilmann, Rolf Klemm, and Hans Zappe, "Development of an Implantable Pulse Oximeter", IEEE transactions on Biomedi-cal Engineering, VOL. 55, NO. 2, February 2008.

[6] Guowei Di, Xiaoying Tang, Weifeng Liu, "A Reflectance Pulse Oximeter Design Using the MSP430OF149", IEEE/ICME International Conference on Complex Medical Engineering, 2007.

[7] Hassan Deni, Diane M. Muratore, Robert A. Malkin, "Development of a Pulse Oximeter Analyzer for the Developing World", IEEE, 2005.

[8] Panayiotis A. Kyriacou, Sarah Powell, Richard M. Langford, and Deric P. Jones, "Esophageal Pulse Oximetry Utilizing Reflectance Photoplethysmography", IEEE transactions on Biomedical Engineering, VOL. 49, NO. 11, November 2002.

[9] Juan M. Lopera, Juan Diaz, Miguel J. Pneto and Fernando Nuno, "Pulse oximeter for homecare", IEEE Proceedings of the Second Joint EMBSBMES Conference, October 2002.

Speed Scheduling of Autonomous Railway Vehicle Control System using Neuro-Fuzzy System

Aiesha Ahmad, M.Saleem Khan, Khalil Ahmed, Nida Anwar and Atifa Athar

Abstract— This paper presents the speed scheduling system of autonomous railway vehicles using neuro-fuzzy system. Speed maintaining and scheduling system is considered integral part for successful development of autonomous railway vehicle control system. This work focuses on development of intelligent speed scheduling system to successfully cope with constraint of different conditions by improving performance and stability as compared to existing control systems of railway vehicles. This intelligent speed scheduling system has ability to learn, take decision and act according to hard conditions; junction track information (JTI), crossing gate information (CG) and track clearance (TCL), and flexible conditions; vehicle tilting (VT), track condition (TC) and environment monitoring (EM), uses neuro-fuzzy system (NFS) comprising main features of fuzzy inference system (FIS) and artificial neural network (ANN). Artificial neural network (ANN) and fuzzy inference system (FIS) are used to solve complex real time speed scheduling problems intelligently by learning, adaptation and human knowledge incorporation. The proposed speed scheduling system learns and adapts automatically under uncertain situations of railway track system. This helps to maintain successfully the speed of railway vehicles with environment monitoring, time scheduling and minimizes the risk of overturning.

Index Terms— Autonomous Railway Vehicle, Control System, Environment Monitoring, Learning and Adaptation, Neuro-Fuzzy System, Railway Track System, Speed Scheduling.

———————————————— ◆ ————————————————

1 INTRODUCTION

ADVANCEMENT in railway vehicles technology has been increased from last few years to facilitate passengers by improving performance with speed, time scheduling, traffic control and passenger management [1]. The importance of control, management and monitoring for railway vehicles progress gradually under modern solution of embedded systems, software based computer aided control systems, sensors and data communication technologies. Different methods are used to control speed at different stages under different conditions for high speed railway trains like the development of eddy current brake system (ECB) by generating constant torque with control algorithm and intelligent dispatching system using neuro-fuzzy technique [2], [3]. The design and development of agent base autonomous railway vehicle control system is considered important for flexible and well established network to enable collaboration between centralized and distributed systems of railway tracks [4].

In these autonomous railway control systems, agents are defined as condition monitoring units with capability to collect information independently and control system autonomously according to their design control [5].

The monitoring and control of railway vehicles are mainly focused on speed management and scheduling, traffic control and time scheduling. Speed scheduling plays vital role for successful development of autonomous railway control system by focusing on hard conditions; junction track information (JTI), crossing gate information (CG) and track clearance (TCL), and flexible conditions; vehicle (VT) tilting, track condition (TC) and environment monitoring (EM). To maintain information about track condition with support of acceptable ride quality and tilting of trains around curved tracks with speed adjustment have been done through some sensors, mathematical modular technique and kalman filtering for data estimation and tracking in existing systems. The complexity and dynamic nature of autonomous railway control system is needed some sophisticated method with domain knowledge representation, automatic learning and adaptation for successful handling of uncertain situations during running on track.

Intelligent systems incorporate the main features of artificial intelligence like learning and decision making under uncertain situations by using previous knowledge [3], [6], [7]. To make autonomous railway control system more flexible and accurate, it's necessary to develop intelligent speed scheduling system using neuro-fuzzy techniques. Neuro-fuzzy system consists of key features of

———————————————

- *Ms. Aiesha Ahmad is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan (e-mail: aiesha@ncbae.edu.pk).*
- *Dr. M. Saleem Khan is with the Computer Science Department as Director in GC University Lahore, Pakistan (e-mail: mskgcu@yahoo.com).*
- *Dr. Khalil Ahmed is with the School of Computer Science at NCBA&E Lahore, Pakistan, He is an expert academician and passionately engaged in research. (e-mail: drk@ncbae.edu.pk).*
- *Ms. Nida Anwar is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan. She is currently working as faculty member in VU, Lahore, Pakistan (e-mail: nidaanwar85@gmail.com).*
- *Ms. Atifa Athar is working as a faculty member at the School of Computer Science, NCBA&E Lahore, Pakistan. She is a research fellow at SCS, NCBA&E (e-mail: atifaathar@yahoo.com).*

fuzzy inference systems like capability to perform uncertain reasoning under incorporation of human knowledge and artificial neural network to learn and adapt in uncertain situations that form the hybrid intelligent system. The proposed intelligent speed scheduling system for high speed railway vehicles will be able to learn and adapt under uncertain situations of hard and flexible conditions by adjusting speed according to the requirement with high safety, performance and time management.

The arrangement of this research paper is as follows: section 2 describes the brief introduction of fuzzy inference system (FIS), Artificial Neural Network (ANN) and Neuro-Fuzzy systems (NF). Structure of proposed speed scheduling system is discussed in section 3 with design model of proposed system and MAT Lab simulation while section 4 consists of results and discussion of Design model of intelligent speed scheduling system. Conclusion and future work are discussed in section 5.

2 OVERVIEW OF FUZZY INFERENCE SYSTEM, ARTIFICIAL NEURAL NETWORK AND NEURO-FUZZY SYSTEM

Fuzzy inference system, artificial neural network and Neuro-fuzzy system provide the control strategies for the Various systems to run autonomously.

2.1 Fuzzy Inference System

There has been rapid growth in fuzzy logic and fuzzy set theory from last few decades to deal successfully with complex uncertainties inputs in many real time processes to achieve definite outputs (or solution/ decision) [6], [8], [9]. Basically fuzzy inference system (FIS) or fuzzy model is a famous computing framework that consists the idea of fuzzy IF-THEN Rules, and the selection of fuzzy rules. Fuzzy set theory deals with membership functions and fuzzy reasoning to perform inference mechanism upon the rules and given facts to obtain reasonable solution. The actual purpose of fuzzy inference is mapping of inputs to a given output using fuzzy logic that provides the base for better, quicker and more accurate decision/output (s) as compared to traditional approach. Fuzzy inference system is not only capable to deal with knowledge in linguistic rules from human experts but also improve performance by adapting itself using numeric data [10]. Artificial Neural Network (ANN) cannot deal with linguistic values which gives advantage to FIS over ANN. Fuzzy inference system further divided into three categories such as Bank Inference System, Medical Inference System and Boiler Controller System, on the basis of wide range of successful applications in different domains [8].

2.2 Artificial Neural Network

Artificial Neural Network (ANN) is considered as one of main areas in Artificial Intelligence (AI) to make systems intelligent with learning capability. It is inspired by biological nervous system as human brain has amazing parallel processing ability with effectively handling of incomplete and imprecise information [9]. The ANN is composed of number of neurons (or processes) that communicate with each other through weighted connections. Artificial Neural Network follows some rules which help to adjust the weights of connections on the basis of data. There are two main techniques used in ANN learning paradigm: Supervised Learning; which consists of input-output pairs for training to minimize the weights of connection between ANN outputs and Training data output, Unsupervised Learning; focused on input data sets with solution state space and adjust connections on the basis of output neurons with highest activation output [7]. ANN and FIS are widely used in the field of predication, identification, diagnostics and the control of linear and nonlinear systems.

2.3 Neuro-Fuzzy System

Neuro-Fuzzy System combines the advantages of FIS and ANN, therefore it is considered as powerful approach in area of Hybrid Intelligent systems to handle complex real time problems successfully [9], [11]. It is like neural network with equivalent functionality of fuzzy inference system because it has ability to combine the parallel learning and computation abilities of neural network with knowledge representation and explanation like humans of fuzzy model. This approach increases the transparency of neural network with infusion of learning capability in fuzzy models.

3 STRUCUTRE OF PROPOSED SPEED SCHEDULING SYSTEM

The proposed speed scheduling system consists of pre-loaded information of track in form of root chart and intelligent design to cope with uncertain conditions successfully. Railway track system receives two main inputs (hard and flexible conditions) from environment and uses sensors to differentiate between these inputs. In case of flexible conditions (FC): VT, EM and TC observed through sensors and given to FIS. These sensors are capable to monitor environment monitoring (EM), track condition (TC) and vehicle tilting (VT) individually by subdividing into 0 to 5 volt in which 0 volt represents the absence of these flexible conditions with no speed adjustment while from 1-5 volt shows gradual increase in these conditions with respective decrease in speed, like 5 volt represents worst condition with slow speed. Hard conditions (HC) like JTI, TCL and CG are sensed through sensors at particular distance and ultimately stop the train whatever adjusted speed may be after existence of flexible conditions. Brief description of input versus output is shown in Table 1 of railway vehicle speed scheduling.

The speed scheduling is achieved by using back-propagation algorithm in neuro-fuzzy for learning and

successful management of uncertain situations.

TABLE 1
OVERVIEW OF I/O RELATION

INPUT						OUTPUT
Flexible Conditions			Hard Conditions			Speed
TC	EM	VT	JTI	CG	TCL	
L	M	L	N	N	Y	Stop
L	L	H	N	Y	N	Stop
H	L	L	Y	N	N	Stop
VH	H	L	N	Y	Y	Stop
L	L	L	Y	Y	N	Stop
L	H	H	Y	N	Y	Stop
H	L	H	Y	Y	Y	Stop
L	L	L	N	N	N	V.Fast
L	L	M	N	N	N	Fast
M	M	L	N	N	N	Fast
L	L	H	N	N	N	Average
M	VH	L	N	N	N	Slow

3.1 Design Model of Intelligent Speed Scheduling System

This neuro-fuzzy based proposed speed scheduling system will be capable to reduce speed by increasing time after comparing with root chart in uncertain environment. Then it will compare the increase time with root chart estimated time to reach next junction and increase speed to its maximum possible limit to overcome the delay time after handling uncertain situation. The learning capability of proposed system will mature by handling uncertain situations with passage of time that ultimately increase the accuracy and response time. The design model of the proposed speed scheduling is shown in Fig. 1.
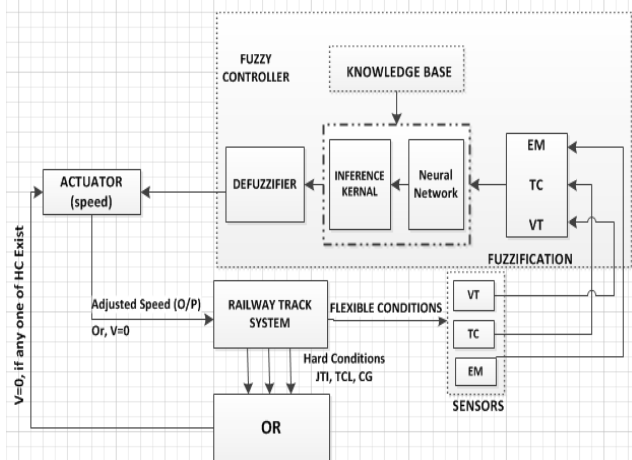
Fig. 1. Neuro-Fuzzy Based Intelligent Speed Scheduling System.

3.2 Speed Scheduling With Fuzzy Inference System

Basic structure of speed scheduling system with hard conditions (HC) and flexible conditions (FC) using fuzzy logic is shown in Fig. 2 which is first step towards neuro-fuzzy technique.
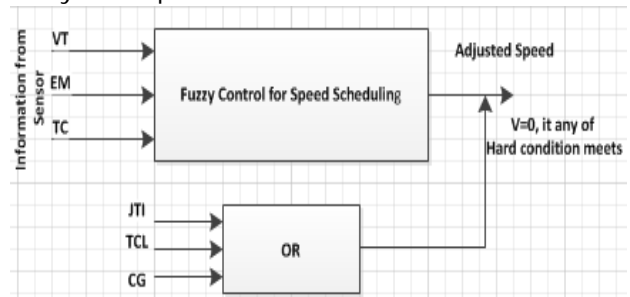
Fig. 2. Block Diagram of Speed Control System.

The two main conditions, HC and FC are further categorized into six input variables; Junction Track Information (JTI), Crossing Gate Information (CG) and Track Clearance (TCL) come under HC, and Vehicle Tilting (VT), Track Conditions (TC), Environment Monitoring (EM) are related to FC and categorized by sensors. Fuzzy control system is used to adjust speed quickly and precisely in the presence of any flexible condition or delay due to any hard condition after comparing with root chart. The fuzzy control system for the proposed speed scheduling system consists of fuzzifier, inference kernel with knowledge base including database, rule base and output membership functions, and defuzzifier block shown in Fig.3.
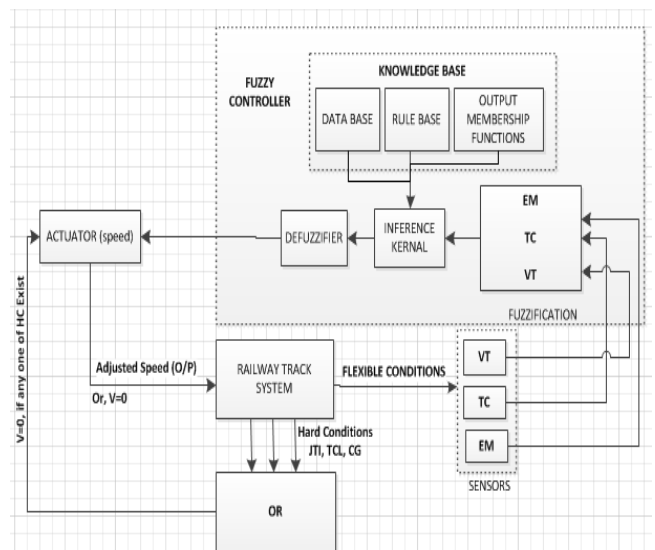
Fig. 3. Fuzzy Control Speed Scheduling.

The crisp values of input variables (VT, TC and EM) are reached to fuzzifier after passing through sensors to identify the types of these input variables [12]- [13]. In fuzzifier, comparison of input crisp values up to certain levels is done by generating linguistic values (Low, Medium, High and

Very High) against each input variable. These linguistic values are passed to inference kernel connected with knowledge base.

In knowledge base, key feature of database is to manipulate fuzzy data and provide essential definitions to describe the linguistic control rules which help the rule base to define the control goals and control policy of particular system such as speed scheduling of railway vehicle in this scenario while output membership functions define the strength of output variables with formulation to adjust speed. After receiving feedback from knowledge base, the next step of inference kernel is to simulate the human decision with fuzzy logic rules to make the control decision in term of adjusted speed, the final outcome. In the next step, defuzzifier maps fuzzy output variable (Slow, Average, Fast and Very Fast) to a crisp value which finally comes to railway track system after passing through actuator.

3.3 Speed Scheduling With Adaptive Neuro-Fuzzy Inference System

Learning capability of speed scheduling system with FIS shows its better performance with quick response in uncertain situations as compared to previous methods. Adaptive neuro-fuzzy inference system (ANFIS) is used to adjust speed according to the requirement with learning capability to manage successfully uncertain situations. ANFIS only supports Takagi/Sugeno type fuzzy inference system which generates single output. Training data set with 400 values for training of flexible conditions with respect to speed is used in ANFIS. The ANFIS model structure with three inputs (VT, TC and EM) shown in Fig. 4, with output speed. Each input further categorizes into four membership functions with 64 fuzzy rules and back propagation learning algorithm.
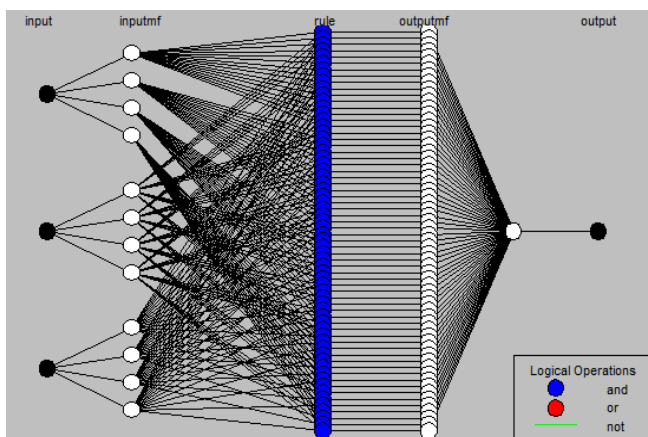
Fig. 4. ANFIS Model Strucutre for Speed Scheduling of Railway Vehicle.

After training of 400 data sets, testing and checking data sets have applied for validation of training data to support our proposed speed scheduling system and shown in Fig. 5.

There are 54 testing data pairs against training of 400 data sets with average data error less than 4, which is very low. The testing data results are shown in Fig.6.
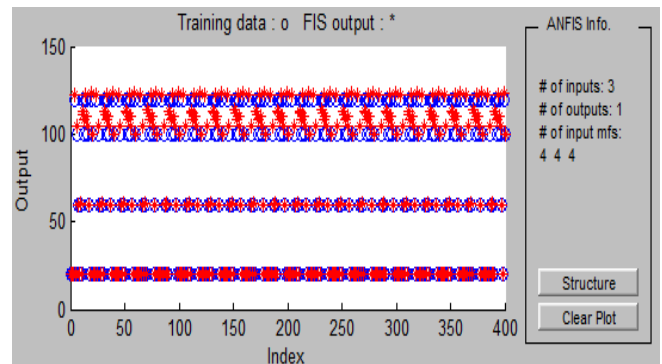
Fig. 5. ANFIS Model for Training Data with Soft Conditions.
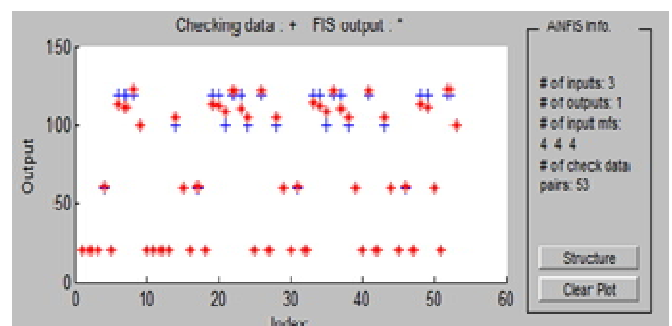
Fig. 6. ANFIS Model for Testing Data with Soft Conditions.

Checking data set with 53 check data pairs have done to show the authentication of training and testing data pairs. The checking data result has shown in Fig.7 that again shows minimum noise against the training data. After that, the importance of proposed intelligent speed scheduling system has been proved.
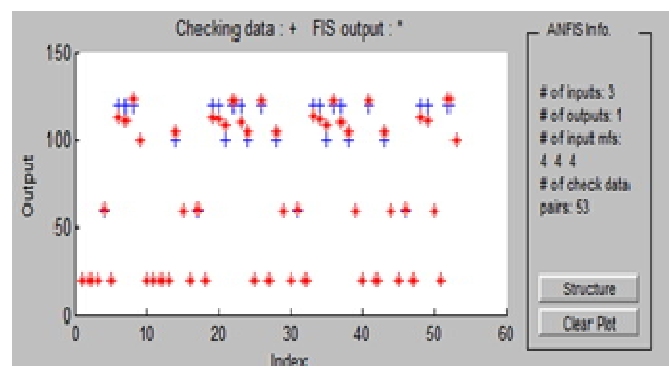
Fig. 7. ANFIS Model for Checking Data with Soft Conditions.

4 RESULTS AND CONCLUSION

The design scheme of the proposed speed scheduling system of railway vehicle has shown significant improvement using neuro-fuzzy system with MAT Lab. simulation as compared to earlier methods. Soft conditions like vehicle tilting, environment monitoring and

track conditions are inversely proportional with respect to the speed and have shown substantial reduction of speed in both ways individually as well as combined effect of these inputs. The effect of these input conditions on speed is shown in Fig. 8 which is according to the design ratio of input and output conditions.

Fig.8 (a) shows with gradual increase in EM and VT, the speed is going to reduce.

Fig. 8 (b) has shown same results with EM on x-axis and TC on y-axis.

Fig. 8 (c) shows, at 0 Volt of VT and EM, the speed is maximum120 km/h, while with gradual increase in VT and EM values, speed is reducing and ultimately reaches at 20 km/h which is very low against highest values of VT and EM.

Fig. 8 (d) shows the same results with TC and EM (as discussed with VT and EM).
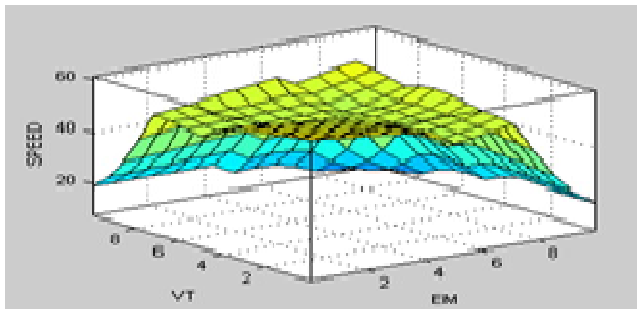
Fig.8. (a) Plot between Environment Monitoring and Vehicle Tilting.
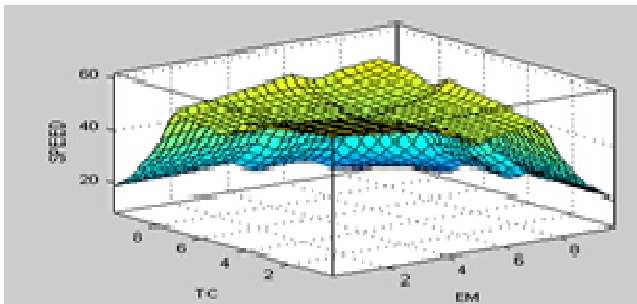
Fig.8. (b) Plot between Environment Monitoring and Track Conditions.
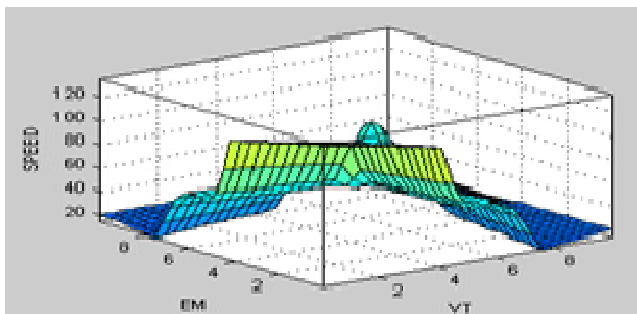
Fig.8. (c) Plot between Vehicle Tilting and Environment Monitoring.
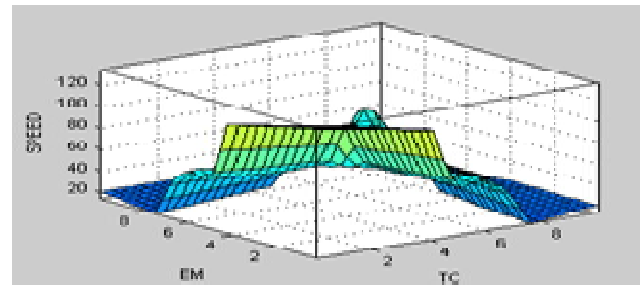
Fig.8. (d) Plot between Environment Monitoring and Track Conditions.

Figure 8 (a to d) MAT Lab Simulation Results.

5   CONCLUSION AND FUTURE WORK

Speed scheduling of railway vehicle has shown substantial improvement by focusing on hard and flexible conditions with learning capability due to neuro-fuzzy technique. In order to overcome on issues and make railway control system more secure and accurate, there is a need of some intelligent system to adjust speed according to situation which is the main part of railway control system. The MAT Lab. results have shown that this intelligent design model will work successfully in real time environment for speed scheduling of railway vehicles. In future, it will help to design whole railway control system intelligent with learning capability of neuro-fuzzy technique to develop secure and intelligent railway control system. State of the art Microelectronics technology can be used to develop FPGAs based control chips for this autonomous railway control system.

ACKNOWLEDGMENT

REFERENCES

[1]   M.Saleem Khan and, Khalid Benkrid, "Design Model of   Multi Agents Based Autonomous Railway Vehicles Control   System", in proc. Of int. Multiconference of Engineering and computer Scientists, Hong Kong, March 18-20, 2009, vol.2.

[2]   Ryoo .Hong-Je, Jong-Soo Kim, Do-Hyun Kang, Geun-Hie Rim, Yong-Ju Kim and Chung-Yuen Won, "Design and Analysis of an Eddy Current Brake for a High-speed Railway Train with. Constant Torque Control", IEEE, pp. 277-281, 2000.

[3]   SUN Jingdong and WANG Shunli, "Intelligent Train Dispatching for High-Speed Railway Based on Fuzzy Neural Network", IEEE, pp. 222-225, 2010.

[4]   R. Harrison, M. Lee, "Reconfigurable Modular Automation Systems for Automotive Power-train Manufacturer", int. Flex Manuf Syst, Springer Science 1 March 2007.

[5]   M. Saleem Khan and, Khalid Benkrid, "Development Tech-

niques of   Multi-Agents Based Autonomous Railway Vehicle Control System", World Academy of Science Engineering and Technology, pp. 558-568, 2010.

[6] S.H. Lee, R.J. Howlett, and S.D. Walters, "Fuzzy and Neuro-fuzzy Techniques for Modelling and Control", Springer-Verlag Berlin Heidelberg, pp.1206-1215 2006.

[7] Gary R. George P. E. and Frank Cardullo, "Application of Neuro-Fuzzy Systems to Behavioral Representation in Computer Generated Forces".

[8] Sandhya Tarar and Prof. Sanjay Jasola, "Modeling and Simulation of Fuzzy Inference System", 2nd int. conference on software Technology and Engineering (ICSTE), IEEE, pp.17-19, 2010.

[9] Zhan-Li Sun, Kin-Fan Au and Tsan-Ming Choi, "A Neuro-Fuzzy Inference System Through Integration of Fuzzy Logic and Extreme Learning Machines", IEEE transactions on systems, man, and cybernetics-part B:cybernetics,vol.37.No.5, pp.1321-1331,October 2007.

[10] K. M. Zemalache and H. Maaref, "Intelligent Control for a Drone by Self-Tunable Fuzzy Inference System", 6th int. Multi-conference on Systems, Signals and Devices, IEEE, 2009.

[11] Prashant Sharma and Dr.Preeti Bajaj,"Performance Analysis of Vehicle Classification System using Type-1 Fuzzy, Adaptive Neuro-Fuzzy and Type-2 Fuzzy Inference System", 2nd int. conference on Engineering and Technology, ICETET-09,pp. 581-584, IEEE, 2009.

[12] Morelli, Jordan E., Akira Hirose and Hugh C. Wood "Fuzzy-Logi-Based Plasma Position Controller for STOR-M" IEEE Transaction on Control Systems Technology, Vol. 13, No 2, March 2005.

[13] M. Saleem Khan, "Fuzzy Time Control Modeling of Discrete Event Systems", Proc. Of World Congress on Engineering and Computer Science 2008.

Autonomous Environment Control System using Fuzzy Logic

Abdul Salam Mubashar, M. Saleem Khan, Khalil Ahmad, Yousaf Saeed

Abstract—This research work presents an autonomous system for premises environment control using fuzzy logic. This proposed design of control system has four inputs: luminance intensity, luminance mode, temperature and humidity. There are six controlling outputs for luminance controller, air conditioner, ceiling fan, air-cooler fan, water-pump and heating unit. This design model can be applied for indoor and outdoor environments like office, work place, home, commercial areas and streets. This application of fuzzy logic would contribute in minimizing the energy wastages. Fuzzy rules are formulated, applied and tested using MATLAB simulation.

Index Terms— autonomous environment control system, fuzzy logic environment control, luminance intensity, temperature and humidity level control system.

———————————— ◆ ————————————

1 INTRODUCTION

THERE are focused considerations for controlling environment suitable to its living beings for not only easiness but also providing comfort to increase working capacities with fresh mind. It would save and enhance energy resources also. There have been criticalities for limitations of energy resources including ever increasing requirements for energy utilizations. There are wastages of energy during its utilizations even in needlessness specifically in deregulated environments [1]. It can be overcome through proper scheduling but it involves hazards. An appropriate solution is to make autonomous and controlled systems according to the requirements. The proposed system is based on fuzzy logic. Fuzzy Logic is suitable for uncertainties issues. The non-probabilistic problems are dealt with fuzzy logic [3]. There are different fuzzy inference system and defuzzification techniques were reported [2], [6]. However this research provides comprehensive application of fuzzy Logic for the particular appliances together involving the automation of light controller, air conditioner, heating unit ceiling fan and air-water-cooler controlling luminance, temperature and humidity. It would also save the energy cost. The feelings of temperature vary with the different levels of humidity. Humidity is the amount of water vapor in the air. In high humidity living beings feel the atmosphere very hot in summer because it reduces the effectiveness of sweating

————————————————

- *Mr. Abdul Salam Mubashar is doing research in the field of Computer Science at National College of Business Administration and Economics (NCBA&E), Lahore, Pakistan (e-mail: salammunim@yahoo.com)*
- *Dr. M. Saleem Khan, Director Computer Science Department is with GC University, Lahore, Pakistan (e-mail: mskgcu@yahoo.com)*
- *Dr. Khalil Ahmad, Director Computer Science Department National College of Business Administration and Economics (NCBA& E), Lahore, Pakistan (e-mail: ahmedkhalil08@gmail.com).*
- *Mr.Yousaf Saeed is the faculty member & Ph.D. Scholar in the field of Computer Sciences at National College of Business Administration and Economics NCBA& E, Lahore, Pakistan (e-mail: usafonline.email@gmail.com)*

to cool the body by reducing the evaporation of perspiration from the skin. At normal temperature with wet atmosphere, it would be felt cool by living beings because water vapors absorb energy rapidly [5]. Air-cooler appliance works on this phenomenon. It includes air-cooler fan and water pump. It is something like window fan [7]. The phenomenon is also implemented for other appliances to save energies. By decreasing or increasing humidity there is less energy required for controlling temperature and making appropriate luminance according to the specifications to minimize energy wastages and discomfort. The frame work of this paper comprises, design and structure of the proposed autonomous environment control system in section 2, section 3 gives the design algorithm of autonomous environment control system, section 4 describes simulation results and discussion whereas Section 5 gives conclusion and future work.

2 DESIGN AND STRUCTURE OF THE PROPOSED AUTONOMOUS ENVIRONMENT CONTROL SYSTEM

There are included Fuzzyfier, Inference Kernel connected with Knowledgebase and Defuzzifier in the proposed fuzzy logic system. The Knowledgebase contains Database, Rule base and Membership Functions.

The fuzzifier converts the Crisp values into Linguistics values. The linguistics values are manipulated for inference engine [4]. The kernel of the system provides the output according to the Rule Selector. The Rule Selector selects rules according to the Knowledgebase. Knowledgebase comprises of Database, Rule base and Membership Functions [3]. Rule base is built up carefully considering all possible effective situations. There have been incorporated four input variables: Luminance Intensity, Luminance Mode, Temperature and Humidity. There are six outputs controlling appliances: Luminance Controller, Air Conditioner, Ceiling Fan, Air-Cooler Fan, Water

Pump and Heating Unit. There have been considered sixty rules for the effective situations focusing energy saving. There are used sixty operators for the rule selector. Luminance mode can be set according to the requirement. However there are proposed Time Oriented Luminance Intensity Mode Specifications for three types of premises including indoor and outdoor environments.

The feelings of hotness depend upon temperature and humidity. During low temperatures, the heating systems work for long enough to convert the climate into dry resulting discomfort. At normal temperature heating system remains off while there wet climate effects cool.
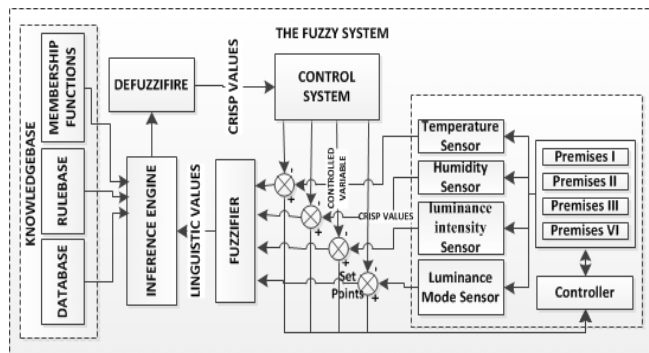
Fig.1. Block diagram of the proposed system to control the environment in different premises using fuzzy logic.

Similarly at warmer temperatures dry humidity demands comfort. Such situations indicate deregulations that result in not only discomfort but energy wastages also. The proposed system deals with such scenarios. It can be applied for different types of premises including office, workplace, study, display, sleep, candlelight and street modes whether the premises contain or don't contain all the appliances together. It would control luminance, temperature and humidity.

3 DESIGN ALGORITHM OF AUTONOMOUS ENVIRONMENT CONTROL SYSTEM

The system is designed for five appliances. It takes four inputs namely luminance intensity, luminance mode, temperature and humidity. The membership functions with their respective ranges for input variables are shown in the Table I.

TABLE I.

MEMBERSHIP FUNCTIONS OF INPUT VARIABLES: LUMINANCE INTENSITY, LUMINANCE MODE, TEMPERATURE AND HUMIDITY

Luminance Intensity (cd)	MF Range	Luminance Mode	MF Range	Temperature (°C)	MF Range	Humidity (RH)	MF Range
Dark	0-33	Sleep/Off	0-33	Cold	0-15	Wet	0-33
Low	0-66	Candle Light/Safety	0-66	Cool	0-30	Normal	0-66
				Nominal	15-45		
Normal	33-100	Normal Working	33-100	Warm	30-60	Dry	33-100
				Very Warm	45-80		
Bright	66-100	Work/Study /Display	66-100	Hot	60-100	Very Dry	66-100
				Very Hot	80-100		

There are six output variables to adjust and controle an environment for the five appliances. The membership functions for light control and air conditioner are represented in Table II.

TABLE II.

MEMBERSHIP FUNCTIONS OF OUTPUT VARIABLES: LUMINANCE CONTROLLER AND AIR CONDITIONER

Luminance Controller	MF Range	Air Conditioner	MF Range
Off	0-30	Off	0-20
Low	0-60	Fan Mod	0-40
Medium	30-90	Low	20-60
High	60-100	Medium	40-80
		High	60-100
		Very High	80-100

The membership functions for output variables: ceiling fan, air-water fan, water pump and heating unit are shown in the Table III.

TABLE III.

MEMBERSHIP FUNCTIONS OF OUTPUT VARIABLES: CEILING FAN, AIR-COOLER FAN, WATER PUMP AND HEATING UNIT

Ceiling Fan	MF Range	Air-Cooler Fan	MF Range	Water Pump	MF Range	Heating Unit	MF Range
Off	0-30	Off	0-50	Off	0-50	High	0-50
Low	0-60	Low	0-100	Low	0-100	Low	0-100
Medium	30-100	High	50-100	High	50-100	Off	50-100
High	60-100						

Table IV shows the rule base for input luminance intensity, luminance mode and corresponding output for luminance controller.

TABLE IV.

RULE BASE FOR LIGHT INTENSITY CONTROL

Rule No.	Input		Output
	Luminance Mode	Luminance Intensity	Luminance Controller
1	Sleep/Off	Dark	Off
2		Low	Off
3		Normal	Off
4		Bright	Off
5	Candle light/Safety	Dark	Low
6		Low	Off
7		Normal	Off
8		Bright	Off
9	Normal working	Dark	Medium
10		Low	Medium
11		Normal	Off
12		Bright	Off
13	Work/Study / Display	Dark	High
14		Low	High
15		Normal	Medium
16		Bright	Off

Luminance intensity would be used to adjust the lumin-

ance of the premises according to the requirement and the type of the premises. The requirments can be set with respect to the type of the premeses with time oriented scheduling in terms of luminance mode. A model scheduling has been drawn for offic or work environmrnt and given in the Table V.

TABLE V.
TIME ORIENTED LUMINANCE MODE SPECIFICATIONS FOR OFFICE OR WORK PREMISES

Time Oriented Mode Specifications for Work Place Premises		
Time	**Scale**	**Luminance Mode**
Arrival	07:00-09:00	Normal Daily Working
Working Hours	09:00-17:00	Work/Study /Display
Closing	17:00-18:00	Normal Daily Working
Evacuation	18:00-18:30	Candle light/Safety
Work Place Closed	18:30-07:00	Sleep/Off

For home premises a modulation for luminance control mode is represented in Table VI.

TABLE VI.
Time Oriented Luminance Intensity Mode Specifications for Home premises

Time	Scale	Luminance Mode
early morning	04:00-06:00	Candle light/Safety
morning	06:00-10:00	Sleep/Off
noon	10:00-14:00	Sleep/Off
after noon	14:00-17:00	Sleep/Off
evening	17:00-20:00	Normal Working
night	20:00-23:00	Study/Display
late night	23:00-04:00	Sleep/Off

Controling light in the outdoor environments is manupulated with the dierect daylight luminance. It can be implimented by both time oriented and the luminance sensor. However the lator is convinient. Table VII presents both ways.

TABLE VII.
TIME ORIENTED LUMINANCE INTENSITY MODE SPECIFICATIONS FOR OUTDOOR PREMISES

Time	Scale	Luminance Mode
Evening to Morning	18:00-06:00 (< 200 Cd)	Normal Working
Morning to Evening	06:00-18:00 (> 200 Cd)	Sleep/Off

Table VIII presents rule base for inputs: temperature and humidity to controle output variables: ceiling fan, air-cooler fan its water pump, air conditioner and heating unit. Rules 17 to 24 are formulated for low temperatures when heating unit is critical while cooling appliances remain off.

TABLE VIII
RULE BASE FOR TEMPERATURE AND HUMIDITY CONTROL IN THE PREMISES

Rule No.	Inputs		Outputs				
	Temperature	Humidity	Ceiling Fan	Air-Cooler-Fan	Water-Pump	Air-Conditioner	Heating Unit
Rule Base for Temperature and Humidity Control in Indoor Premises at Low Temeratures When Heating Unit Might Be On							
17	Cold	Wet	Off	Off	Off	Off	High
18	Cold	Normal	Off	Off	Off	Off	High
19	Cold	Dry	Off	Off	Off	Off	Low
20	Cool	Wet	Off	Off	Off	Off	High
21	Cool	Normal	Off	Off	Off	Off	Low
22	Cool	Dry	Off	Off	Off	Off	Low
23	Nominal	Wet	Off	Off	Off	Off	Low
24	Nominal	Normal	Off	Off	Off	Off	Off
Rule Base for Temperature and Humidity Control in Indoor Premises Containing Air-Cooler and Ceiling Fan							
25	Warm	Wet	Off	Off	Off	None	Off
26	Warm	Normal	Low	Off	Off	None	Off
27	Very Worm	Normal	Medium	Off	Off	None	Off
28	Hot	Normal	High	Off	Off	None	Off
29	Hot	Dry	Medium	Low	Low	None	Off
30	Hot	Very dry	Low	High	High	None	Off
31	Very Hot	Normal	Low	Low	Low	None	Off
32	Very Hot	Dry	Medium	High	High	None	Off
33	Very Hot	Very Dry	High	High	High	None	Off
Rule Base for Temperature and Humidity Control in Indoor Premises Containing Air-Conditioner							
34	Warm	Wet	None	None	None	Off	Off
35	Warm	Normal	None	None	None	Fan Mod	Off
36	Very Warm	Normal	None	None	None	Fan Mod	Off
37	Hot	Normal	None	None	None	Low	Off
38	Hot	Dry	None	None	None	Medium	Off
39	Hot	Very dry	None	None	None	High	Off
40	Very Hot	Normal	None	None	None	Medium	Off
41	Very Hot	Dry	None	None	None	High	Off
42	Very Hot	Very Dry	None	None	None	Very High	Off
Rule Base for Temperature and Humidity Control in Indoor Premises Containing CeilingFan							
43	Warm	Wet	Off	None	None	None	Off
44	Warm	Normal	Low	None	None	None	Off
45	Very Warm	Normal	Medium	None	None	None	Off
46	Hot	Normal	Medium	None	None	None	Off
47	Hot	Dry	Medium	None	None	None	Off
48	Hot	Very dry	Low	None	None	None	Off
49	Very Hot	Normal	High	None	None	None	Off
50	Very Hot	Dry	High	None	None	None	Off
51	Very Hot	Very Dry	Medium	None	None	None	Off
Rule Base for Temperature and Humidity Control in Indoor Premises Containing Air-Cooler							
52	Warm	Wet	None	Off	Off	None	Off
53	Warm	Normal	None	Low	Off	None	Off
54	Very Warm	Normal	None	High	Off	None	Off
55	Hot	Normal	None	Low	Low	None	Off
56	Hot	Dry	None	High	Low	None	Off
57	Hot	Very dry	None	Low	High	None	Off
58	Very Hot	Normal	None	High	Low	None	Off
59	Very Hot	Dry	None	High	High	None	Off
60	Very Hot	Very Dry	None	High	High	None	Off

Rules 25 to 33 cover the premises containing all the appliances except heating unit. Rules 34 to 42 are for the premises that contain air conditioner and rules 43 to 51 reserve for just ceiling fan. Rules 52 to 60 cover the crticalities for air cooler including its water pump.

Plots of membership functions for input fuzzy variables are given in Fig. 2, Fig. 3, Fig. 4 and Fig. 5. The four membership functions Dark(1), Low(2), Normal(3) and Bright(4) are used to represent various ranges of input variable luminance intensity in Fig. 2.
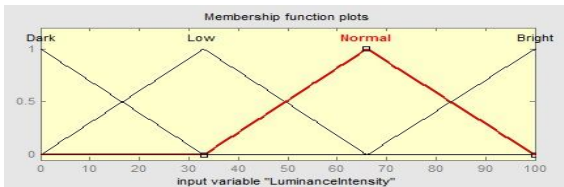
Fig. 2. Plot of membership functions for input fuzzy variable Luminance Intensity

The four membership functions Sleep / Off (1), Candle light/Safety (2), Normal daily Working (3) and Study/Display (4) are used to represent various ranges of input variable luminance intensity mode in Fig. 3.
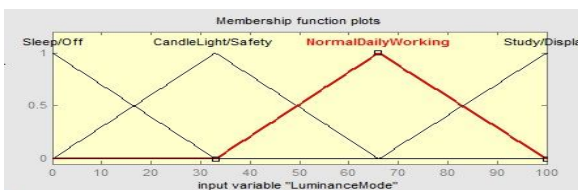
Fig. 3. Plot of membership functions for input fuzzy variable Luminance Mode

The seven membership functions Cold (1) Cool (2), Normal(3), Warm(4), Very Warm (5), Hot (6) and Very Hot(7) are used to represent input variable Temperature in Fig. 4 containing six regions.
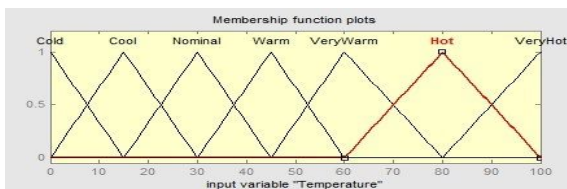
Fig. 4. Plot of membership functions for input fuzzy variable Temperature with their respective ranges

The four membership functions Wet (1) Nominal (2), Dry(3) and Very Dry(4) used to represent input variable Humidity in Fig. 5. It consists of three regions.
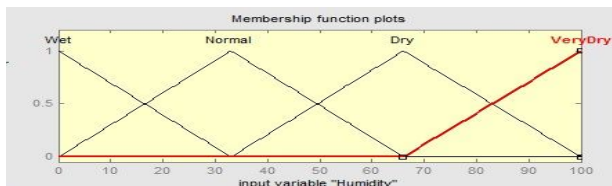
Fig. 5. Plot of membership functions for Humidity with the ranges

The four membership functions Off(1), Low(2), Medium(3) and High(4) luminance are used to represent fuzzy output variable Light Controller in Fig. 6. It contains three regions.
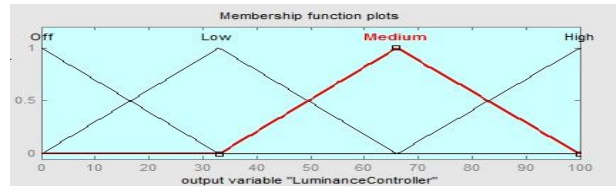
Fig. 6. Plot of membership functions for output Luminance Controller with their respective ranges

There are used six membership functions Very High (1) High(2), Medium(3), Low(4), Fan Mode (5) and Off (6) to represent output controlling variable for Air Conditioner in Fig. 7 containing five regions.
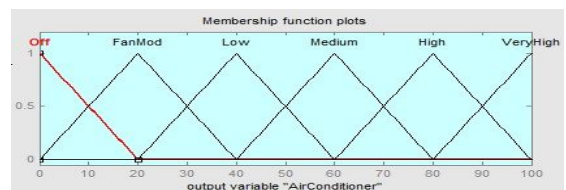
Fig. 7. Plot of membership functions for output variable Air conditioner with their respective ranges

There are used four membership functions High(1) Medium(2), Low(3) and Off(4) to represent output controlling variable for Ceiling Fan in Fig. 8 consisting three regions. The third region is critical at high humidity.
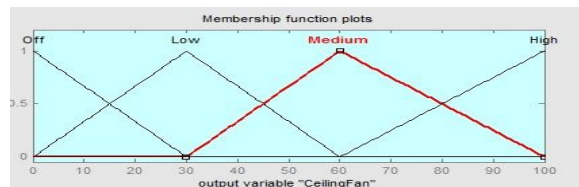
Fig. 8. Plot of membership functions for fuzzy output variable Ceiling Fan with their respective ranges

Although the rang values of Air-cooler Fan, Water Pump and Heating Unit are taken according to the requirment, however these are same. Therefore the shape of of the plot of membership functions for output variable Air-cooler Fan, its Water Pump and Heating Unit are occued same in this proposed design and are shown in Fig. 9.
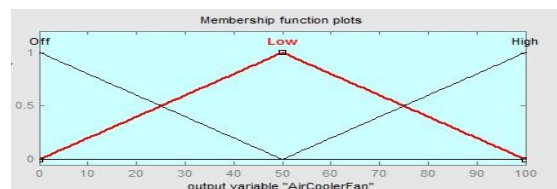
Fig. 9. Plot of membership functions for output variable Air-Cooler Fan, its Water Pump and Heating Unit with their respective ranges

There are used three membership functions High (1) Low (2) and Off (3) to represent output controlling variable for Air Water Fan, its Water Pump and Heating Unit in Fig. 9 consisting of two regions.

The Fuzzy Fier will convert the three input crisp values into six output linguistic values. There are sixty rules are fired for the values of one set of variables. A total of 60 rules have been generated. The inference engine consists of sixty operators. The rule selector receives four input values luminance intensity, luminance mode, temperature and humidity. It provides singleton values of output functions under algorithm rules applied on this design model. The Defuzzifier converts the output values into crisp values to control the environment.
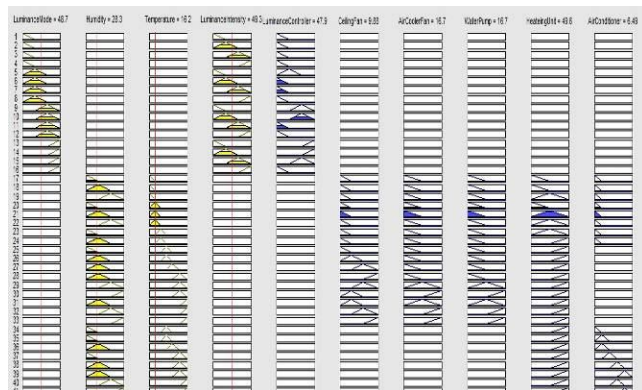
Fig. 10. MATLAB rule viewer and simulations result for autonomous environment control system

The rule viewer represents output variations for different values of input variables according to the rule base.

4 SIMULATION RESULTS AND DISCUSSION

The surface view of the plots of output variables luminance controller, air conditioner, ceiling fan, air-cooler fan, water pump and heating unit on input variables luminance intensity, luminance mode, temperature and humidity has been drawn according to the design scheme of the rule base. The plots in Fig. 11 indicate the inter dependencies of the effects of temperature and humidity on controlling units.

Fig. 11(a). shows that luminance intensity and luminance mode is directly proportional to the light controller. It means energy can be saved only through proper scheduling and does not by other than light resource.

Fig. 11(b). describe that ceiling fan directly proportional to temperature in most of the ranges of the membership functions while ceiling fan is not directly proportional to humidity in all the ranges of the membership functions. It shows that energy would be less consumed in the situations when ceiling fan is not depending on temperature and humidity.

Fig. 11(c). indicates the plot of air-cooler fan on temperature and humidity while energy would be saved in the situations when air-cooler fan is not depending on temperature and humidity.

Fig. 11(d). is the plot of water pump on temperature and humidity showing that it works smoothly in medium-low temperatures and high humidity while it is directly proportional to the humidity in high temperatures.

Fig. 11(e). indicates that energy is to be saved from heating unit having small values in high humidity with low temperature.

Fig. 11(f). depicts that energy is saving by reducing air conditioner in low temperatures.
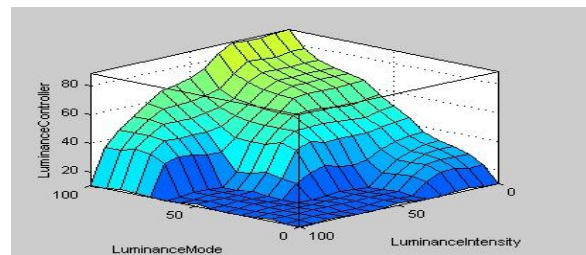
Fig. 11(a). Plot view of luminance mode, luminance intensity on luminance controller.
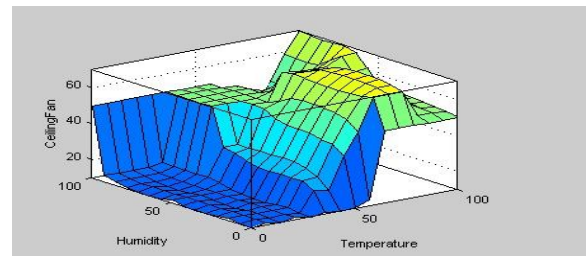
Fig. 11(b) Plot view of ceiling fan on temperature and humidity.
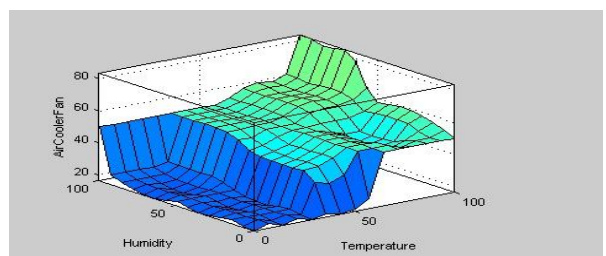
Fig. 11(c) Plot view of air-cooler fan on temperature and humidity.
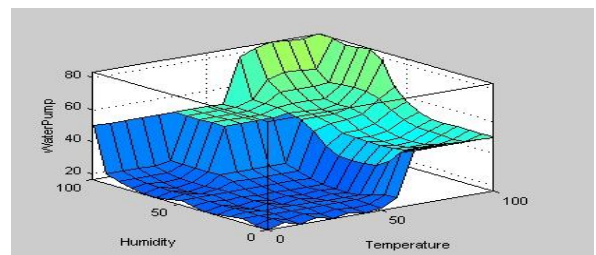
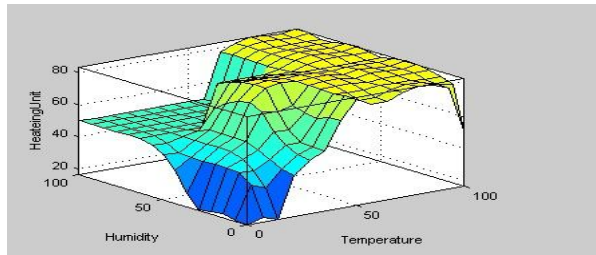Fig. 11(d) Plot view of water pump on temperature and humidity.

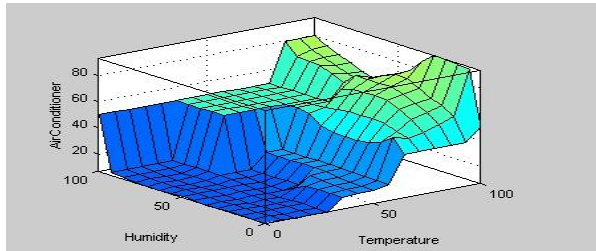Fig. 11(e) Plot view of heating unit fan on temperature and humidity.

Fig. 11(f) Plot view of air conditioner on temperature and humidity.

5 CONCLUSION AND FUTURE WORK

The simulated results show the effectiveness of the Autonomous Environment Control System. It automates the appliances and contributes in better utilization of energy. The proposed system is suitable for deregulation environments also. It works from low humidity to high humidity, low temperature to high temperature and dark luminance to bright maintaining environment pleasant to the user satisfaction and comfort. It would enhance working capacities also.

For future work neural network simulation of Autonomous Environment Control System would be manipulated including more appliances.

REFERENCES

[1]   H. Salehfar, B. J. LaMeres et. al., "Fuzzy Logic Based Direct Load Control of Residentional Electric Heaters and Air Conditioners Recognizing Customer Prefrences in a Deregulated Environment", Proc. 1999 IEEE

[2]   Shabiul Islam, Shakowat,"Development of a Fuzzy Logic Controller Algorithm for Air-conditioning System", ICSE 2006, Proc. 2006 IEEE.

[3]   M. Saleem Khan et. al., "A Proposed Grinding and Mixing System using Fuzzy Time Control Discrete Event Model for Industrial Applications", Proc. 2009 IMECS

[4]   B. P. Zeigler, P. Herbert, "Theory of Modeling and Simulation, Integrating Discrete Event and Continuous Complex Dynamic Systems" 1994, IEEE Press

[5]   C.Michael Hogan., "Abiotic factor", Encyclopedia of Earth. eds Emily Monosson and C. Cleveland. National Council for Science and the Environment. 2010, Washington DC

[6]   Mircea Grigoriu, et.al. "Intelligent Buildings Energy Supply Following Climate Parameters Variation Fuzzy Control" 2010, Proc. WSEAS.

[7]   James Wiese, Maple Grove, "Window fan control system and method of controlling a fan unit", 2009 US Patent application publication.

Mr. Abdul Salam Mubashar is doing Research in Artificial Intelligence at National College of Business Administration and Economics (NCBA&E) Lahore, Pakistan. His research area of interest includes Machine Cognition.

Mr. M. Saleem Khan is an Assistant Professor at the GC University, Lahore, Pakistan. Currently he is working as director Computer Science Department in GC University, Lahore, Pakistan. He availed research fellowship at The School of Electronics & Engineering, University of Edinburgh, UK and completed his Ph.D. thesis in the field of control systems design, simulation and analysis in local and distributed environment. He contributed his services on various projects in the field of Advanced Electronics and Communication. His research interests include control systems design and industrial applications. He promoted a large team of Electronics researchers and organized this field in his country. Mr. Khan had also been served as a senior scientific officer in a classified defense research organization in his country.

Mr. Khalil Ahmed is the Director School of Computer Science at National College of Business Administration and Economics (NCBA&E) Lahore, Pakistan. He is an expert academician and passionately engaged in research. His area of research is machine consciousness, A.I. and knowledge management.

Mr. Yousaf Saeed is an Assistant Professor at National College of Business Administration and Economics (NCBA&E) Lahore. He has done his M.Phil. in Broadband Communication from University of Westminster, London, United Kingdom and now he is doing Ph.D. from National College of Business Administration and Economics Lahore. His research area include Broadband and high Speed Communication Networks.

Bio-inspired Neuro-Fuzzy Based Dynamic Route Selection to Avoid Traffic Congestion

Sagheer Abbas, M. Saleem Khan, Khalil Ahmed, M.Abdullah and Umer Farooq

Abstract— this paper presents the bio-inspired neuro- fuzzy based route selection system to avoid traffic congestion. The proposed neuro-fuzzy system selects the best multi-parameters direction between two desired nodes: source and the endpoint. This research practices a mixture of neuro-fuzzy logic and ant colony system (ACS) algorithm for the principal routing to fulfill all the preferred requirements of the user using operational traffic data directly carried by the traffic control room and predicted by artificial neural networks for forthcoming minutes of the traffic.

Index Terms— Bio-inspired, Traffic congestion, Dynamic route selection, artificial neural networks, Bio-inspired Fuzzy based traffic system.

———————————— ◆ ————————————

1 INTRODUCTION

Traffic congestion is becoming more alarming with every day as the number of vehicles on the roads is increasing rapidly. Route selection grounded on specific desires is a foremost problem for city travelers. In route selection systems, the source and the target are given. The objective of our system is to find out a route with the most favorable conditions satisfying all the needs and desires of a traveller. During the last decades, many researchers tried different approaches to find an optimized route, considering the important parameters that were the most advantageous for the city travellers. In theses systems, the distance and travel time were tried to minimize. Now-a-days many drivers are becoming concerned with not only these desires but also with rising fuel costs, to avoid traffic jamming and to adopt the route that satisfies their all desires like if users are facing road-blocking on the way then it should be possible to adopt the next favorable route. It is a non-polynomial NP hard problem to find out the route that fulfills all prerequisites that force to enumerate all the possible routes. Previous researchers have not spoken a fast, self-motivated and a applied except Salehinejad et al who tried a fuzzy based dynamic route selection approach [1].
In 2007, Barth et al have mechanized some pleasant navigation based environmental systems. It consists of the fuzzy logic for pervasive and authoritative device for the

ideal routing [1] .
In 2011, Raghavendra et al used a new mechanism for pheromones disappearance through a function based on the convergance of the time. It shows a better peformace as compared to all others traditional methodologies [4].
In this paper, Diogo et al attempted ADR algorithm for the routing owing to enhance the efficiency of the system. A new mechanism was used for the cost calculation based on the incentives [6].
In this paper, Saliba et al proposed the quality of service for internet traffic routing based on ants' colony optimization algorithm. The system's performance was compared with the other traditional approaches and it was observed that the presentation of this model was much better as compared to all other methods [7]. In it three paramertes were used for the optimizations that are time, diatance and the no of vans. This work has much importantance regarding time-based palnning. Regarding digital image and tracks detection was proposed on ants based algorithms. [8]. Salehinejad et al. have proposed much for the route selection dased on the ACS [9].
In this paper, Attia et al suggested two algorithms for routing the network traffic. First got the inspiration by the multi-ant parametres, based on the reactive and proactive approaches to adopt the path and second was the mix-up with the QoS constraints. Its performance was excellent for the real time scenario [12].
In this paper, xiaamong et al attempted a quantum mechanism for the travelling salesman problem, initializing the ants with the supeposition of qubit.this sytem shows the performance much better as compared to all other classical techniques [13]. Salehine jad et al introduced a fuzzy based route selection and ant colony system (ACS) based on "Distance", "Incident Risk" and "Traffic" variables [1].

————————————————

Sagheer Abbas is *with Computer Science department as research fellow at NCBA&E Lahore, Pakistan (E- mail: sagheer@ncbae.edu.pk).*

Dr. M. Saleem Khan *is with the Computer Science Department as Director in GC University Lahore, Pakistan (E-mail: mskgcu@yahoo.com).*

Dr. Khalil Ahmed *is with the School of Computer Science at NCBA&E Lahore, Pakistan, He is an expert academician and passionately engaged in research. (E-mail: drk@ncbae.edu.ok).*

Muhammad Abdullah *is a PhD Scholar with Computer Science department at NCBA&E Lahore, Pakistan. (E-mail: abdullah_lahore@yahoo.com)*

Umer Farooq *is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan (E-mail: umer@ncbae.edu.pk).*

This research work proposes a bio-inspired neuro-fuzzy based route selection system with six adjustable constraints. This system uses a combination of ANNs, neuro-fuzzy logic and ant colony optimization algorithm to find out an optimum path based on the costs calculated with time, aimed at all possible directions for the user between two nodes origin and the destination, valuing the user□s point of interests. An optimum route refers to a route that satisfies all the preferred parameters. In this paper, these parameters are "Distance", "Traffic-flow", "Environment monitoring", "Width", "Road condition", "Traffic lights". Due to the computational complexity six important parameters are taken else some more variables like "Incident risk" "Road condition" and the "View" (Home, building, park, river or any other scenic view the user wants to pass through). In this proposed system the present data is provided by the traffic control center and the coming traffic proceedings data is anticipated by artificial neural networks.

The arrangement of this research paper follows as: section II describes the overview of our suggested system, Artificial Neural Network (ANN) and Ant colony optimization (ACO) algorithm. System and Parameters of the proposed system are discussed in section III, while in section IV the simulation results are discussed. Section V provides the conclusion and future work. Acknowledgement of this work is declared in Section VI.

2 OVERVIEW OF PROPOSED SYSTEM

The proposed system is a combination of ANNs and neuro-fuzzy inference engine. All six inputs to the proposed system are provided from the traffic control room TCR and the average speed of the vehicle is given from the user. The system comprises of the previous data, TCR indications that are being received continiously, the present time and the typical speed of the vehicle. Artificial neural networks are casted for the forseen minutes of the traffic. Time delay estimation is used because time is required to move from base to destination. This estimation is done due to different loads of traffic among varied hours of day and night. This estimation is done as the system should be able to estimate the arrival time of the vehicle to the destination. This estimation is done on the basis of the average speed of the vehicle.

The system has the capability to avoid up coming traffic congestion as it is attentive to the up-to-date vehicle position system using GPS. Therefor, if congestion ensues the system recommends the user next direction according to the customized parametres. It is done by keep notifying the quantity of pheromones on the paths using ACS. The system has the flexibility for the customization of the parametres that are being used in this system.

In the Fig. 1 main block diagram of our proposed scheme is shown, how the TCR, present time, Avg. speed, previous data and the ANN collectively work to select the most favorable route using neuro-fuzzy system.
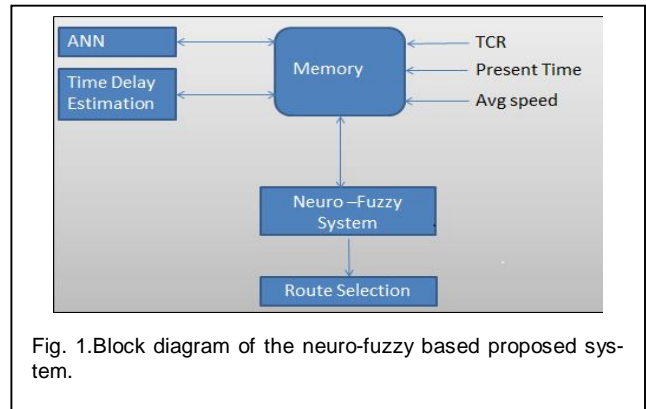
Fig. 1.Block diagram of the neuro-fuzzy based proposed system.

2.1 Neuro Fuzzy Inference System.

It is hybrid approach where neural and fuzzy advantages are combined for the paralle execution. The arrangement contains the structure of fuzzification and de-fuzzification. They both work with combined method.. The conforming grid has enhanced results as compared with the earlier work. The process of transforming crisp values into grades of membership for linguistic terms of fuzzy sets

Fuzzification: In it, the crisp values of the fuzzy system are convertred into grades of member functions.

Defuzzification: It is the reverse process of above to create a defuzzified value for the system.

2.2 Artificial Neural Networks (ANNs)

Scheming and executing intellectual systems is a complex factor for innovation and developments of enhanced products .Artificial neural networks are the approach that works higly parallel computations as compared to our current machines. For a dynamic system, these networks are highly attentionable and can resolve problems that are not possible in linear computation. A very atteactive attribute of these networks is, they are adaptive in nature. If a dynamic application is required, then these networks are considered most suitable due to adaptive nature and decidedly parallelism.
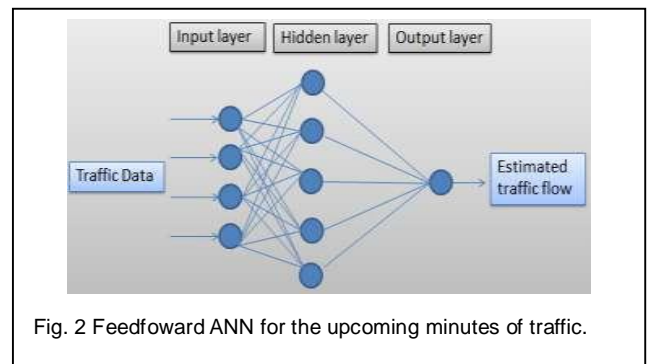

Fig. 2 Feedfoward ANN for the upcoming minutes of traffic.

Neural networks are being engaged in various applications including handwritten character recognition, human mind simulation and patterns appreciation [14]. ANNs

have diffirents types and these networks are trained, grounded on the training data to perform a specific task.

2.3 Ant Colony Optimization Algorithm

Ant colony optimization (ACO) is a probalistics technique that disentangles the cost for most favorable path selection. In ACO algorithm, artificial ants simulate intelligent behavior for discovering the optimized cost of the route selection. These anst are capable of solving complex problems like to find food using a chemical substance called pheromones. As a single ant moves between its nest and the source, it deposits some pheromones quantity on that route. This pheromones quantity contains the path intensity to be adopted for other ants. As the pheromones, quantity is being added frequently, becoming more noticeable for the ensuing ants.

An ant naturally adopts the way where the pheromones quantity is higly concentrated. With the passage of the time, if ants are not following the higly concentrated path, the pheromones quantity start decreasing, causing the path to be less predicted. Those trails that are navigated (to food and back) most quickly soon become the trails with the highest intensity of pheromones, thus boosting the next ants to make choice of these routes.

It is not necessary that the ants are bound to follow the route with the highest intensity of pheromones. Ants can disregard the previous route and search food on all other possible routes. If the ants are able to find a shortest path towards food, then the current solution is swapped with the previous one with highest pheromones quantity. As a new shortest path is defined and updated, the pheromones quantity will start evaporation with time, causing the path is less anticipated for the next ants and resulting the longer path to be skipped.

Route selection based on the above mechanism is the prime factor of ACO. The detailed description of ant behavior as it narrates to ACO is provided by Dorigoet-al [15].

3 DESIGN ALGORITHM OF PROPOSED SYSTEM

On the user defined parameters, for every single vehicel the proposed system is executed. In this system, the traffic indication is delivered by the TCR that containspresent traffic data and is updated dynamically. This system consists of average speed of vehicle, previous vehicles data, present time and the environment monitoring, no of lanes etc. ANNs are used for the future traffic minutes. The ANNs are trained as traffic predictor are used in this system causing helpful to the sytem with use of TCR and all other inputs to the system. The next section defines in details about the strategy algorithm of the system.

Initialize: it comprised of the initial values of the parameters such as number of ants and evaporation co-efficient etc.

LocateAnts: Ants are located on the origin.an ant which is on the way informs other ants, either it is congested or still on the way to destination. It is must that each ant can cross each junction only once. If an ant is blocked some where due to congestion or road blockage , no chance to continue further then will move back to source.

Paradigm 'ρ':
During this phase, cost based probability for each route is calculated for all active ants. This probability is calculated from the route souce to. For each ant, here it is elaborated with k.

$$p_{ij}^k = \begin{cases} \dfrac{\tau_{ij}^{\alpha} \prod_{l \in parameters} \xi_{ijl}^{-\alpha l}}{\sum_{h \in tabu_k} \tau_{ih} \prod_{l \in parameters} \xi_{ihl}^{-\alpha l}} & = \quad j \in tabu_k \\ \\ 0 & else \end{cases}$$

Where, τij denotes the pheromone intensity from source to destination 'i' and 'j' respectively. Constraint„α□ is showing the the route [15, 16], tabu k is showing the all routes that are not suitable for all ants to be adopted in future. Ξijl and 'I' are showing the cost of constraint l for route ij and the worth level of constraint l respectively. In this process list is consisted of all the jammed routes and the constainsts are the parametres for the user to adopt the route according to his customized way.. Currently the six most important parameters"Distance", "Trafficflow", "Environment monitoring", "Width", "Roadcondition", and "Trafficlights" are used in this system. However, it can be enhanced by incorporating other parametress that are desiredable such as "Road Quality" and "other desired facilities like hospital" etc. [8-10].All these parameters have different significance impact from i to j; distance ij, road-condition ij, Env-monitoring ij, *Traffic lights ij, Traffic flow ij, width ij.*

Distance: It is a distance between two junctions from the source to target at time t.

Traffic flow: It is load of traffic on the routes from the source to target at a time t.

Environment Monitoring: It is the environment condition on the different routes from the source to target at a time't'.

Width: It is the total no of lanes on the all-possible routes at a time't' from the source to target.

Traffic Lights: It is the total no of the lanes on all possible routes from the source to target at a time't'.

Road Condition: it is the scale of the road condition that is available at a specific time't' on all possible routes from the source to target

Select Route: a random parameter $0 \le p \le 1$ with uniform probability is matched with the parameter 'P' where $0 \le P \le 1$ and mostly fixed to 0.90. The comparison result

between 'P' and 'p' choices one from these.

a) If 'p' is larger than 'P' the ant adopts the route with highest pheromones intensity.

b) If 'p' is smaller than 'P' or equal, then ant opts the route with possibilities.

Update Tabu List: In this process, the route is updated and added to the list, causing other ants not to select that path. If an ant is on the way to source and is blocked due to traffic congestion or any other reseaon, that path is deactivated from the route for the following ants in order to provide the route reliability. More over its probability isnot considered any more.

Update Pheromone: Ant colony sytem consist of two main rules: first is updation of local pheromone and the other rule is updation of the global pheromone. This rule is applied after all the traffic is travelled to the destination. The pheromone amount is added to the route that is successful completed by an ant as,

$$\tau^{new}_{ij} = \tau^{old}_{ij} + (10 \times \Delta \tau)$$

Where, '$\Delta\tau$' is the amount of local pheromone updating. $\Delta\tau$ is the amount of pheromones of our neuro-fuzzy system. Six parameters of this system use three member functions that are taken by our systems and the pheromone intensity was used between '0' to '5'. If all parameters are in the most ideal situation then the pheromone intensity is maximum else varying between '0' to '5'.

It is the final step of the algorithm to keep posted the global pheromone.

$$\tau^{new}_{ij} = p\tau^{old}_{ij}$$

Where ρ is the evaporation constant and is usually set to 0.9.

In Table 1, some rules for the system are discussed. If all the parameters are found at their ideal values then the intensity of the pheromones is extraordinary. If three parameters comprised of maximum value and the other three contained minimum value then the pheromones intensity will be 3. If all the parameters have minimum degree then the pheromones intensity is minimum and the lowest probability to adopt that path. If the pheromone intensity is high then it is decidedly observed to adopt that route.

Select best route: after m loops, route with the minimum

coust is suggested by the system.

4 SIMULATION RESULTS AND DISCUSSION

According to the scheme of the neuro-fuzzy system for the route selection is designed and simulated on Matlab. All the parameters have significant effect on the system's output. A neuro-fuzzy system is designed for the optimum route selection between two nodes, called the origin and the destination.

In the fig.3, six parameters are used as inputs to the Neuro-fuzzy system. The system makes three member functions of each parameter itself. The rules for the system are denser due to six inputs and three-member function against each input as shown in the figure 3. The output member functions are also denser and the output is generated after all the computation is done.
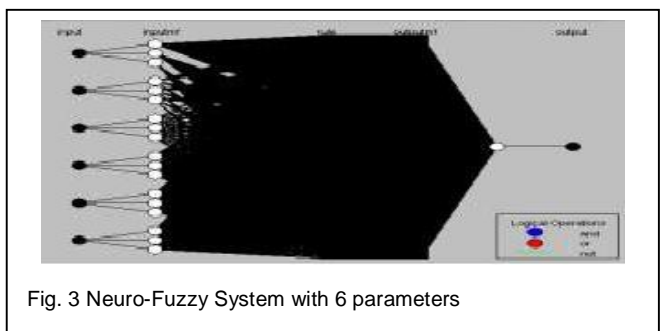
Fig. 3 Neuro-Fuzzy System with 6 parameters

In the following table, we showed the relational parametric relational among all constratints of the sytem.

In the fig.4, Some data is given to the system for the working out of the proposed neuro-fuzzy system. Total no of nodes 1503, participated during the training. In which there are 54 nonlinear parameters, 5103 linear parameters and 729 fuzzy rules are used.
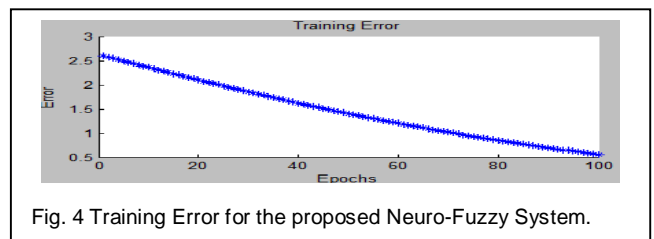

Fig. 4 Training Error for the proposed Neuro-Fuzzy System.

ANFIS training using back propagation technique completed at epoch 130 and the error is 0.56108, is too small.

Fig 5 shows blue circles for the data, offered for the testing purposes to the system and red stars are the outputs against the given inputs to the system. In the testing data fifteen pairs are randomly selected and presented to the system. Red stars are the neuro-fuzzy outputs of the testing data. The average testing error is 2.4827, remains too low and acceptable.

TABLE 1

PARAMETERS RELATIONAL TABLE OF THE SYSTEM

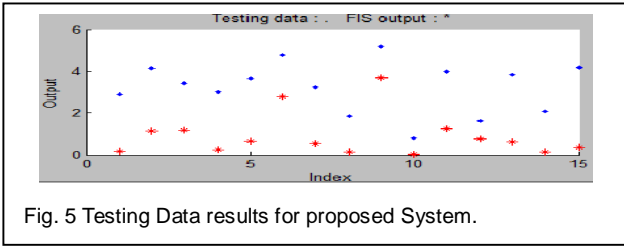| Rules	If						Then
Rule no	Distance	Traffic	R.C.	E .M	Lights	Width	Pheromns
1	low	high	high	Good	low	max	6
2	low	high	low	Good	medium	max	4
3	low	high	medium	Good	medium	max	5
4	medium	medium	medium	Avg	medium	medium	3
5	medium	medium	high	low	low	low	2
6	high	low	low	low	high	low	0
7	high	low	high	Low	low	low	1

Fig. 5 Testing Data results for proposed System.

In Fig 6 the checking data was presented to the system. Total no of pairs that are presented to the system, are 22 and the FIS output is shown against those pairs. Results against the checking data pairs are satisfied and the error is 0.54811.
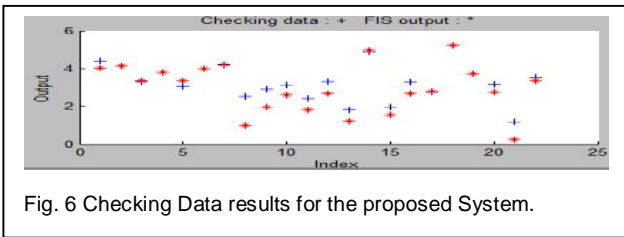

Fig. 6 Checking Data results for the proposed System.

In fig 7 surface graph among inputs road condition and traffic flow against output is shown. When the road condition is at the ideal value the pheromones amount will be max against these inputs that will enhance the probability to adopt that route against two parameters,
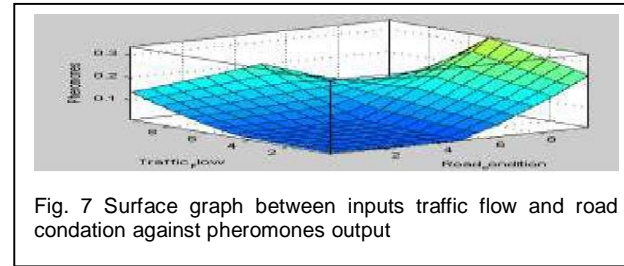
Fig. 7 Surface graph between inputs traffic flow and road condation against pheromones output

In figure 8,9,10 and 11 different parameters combination impact on the output is shown. These parameters impact on the pheromones are customizable.
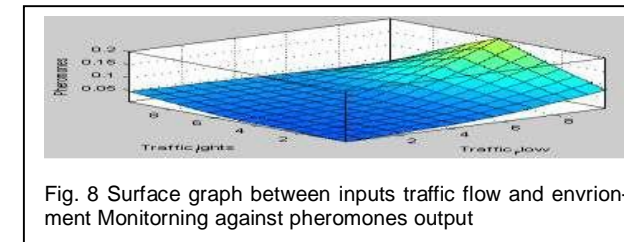
Fig. 8 Surface graph between inputs traffic flow and envrionment Monitorning against pheromones output
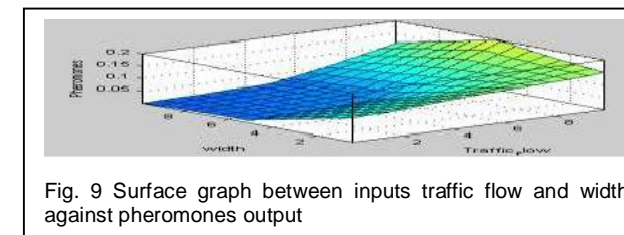
Fig. 9 Surface graph between inputs traffic flow and width against pheromones output
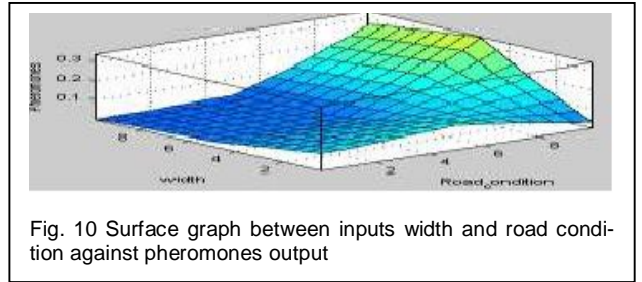
Fig. 10 Surface graph between inputs width and road condition against pheromones output
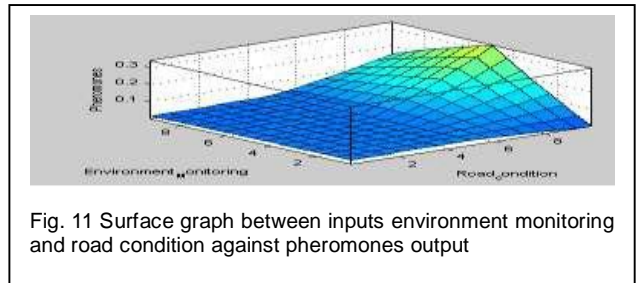
Fig. 11 Surface graph between inputs environment monitoring and road condition against pheromones output

Fig. 12 shows that Input variable traffic flow impact is directly relational to the pheromones intensity. Greater the intensity of the environment monitoring, the more impact will be on the pheromones intensity. Like this all others variables impact can be seen on the output variable.
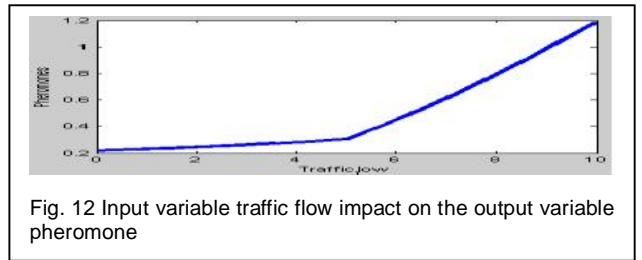
Fig. 12 Input variable traffic flow impact on the output variable pheromone

5   CONCLUSION AND FUTURE WORK

In this proposed system, a neuro-fuzzy sytem is carried out to represent the characteristics of the constraints owing to select the most encouraging route according to the customized preferences of the driver who is utilizing this system. The proposed system is aware of changing the route dynamically when a congestion or obstructive traffic is perceived on the selected route. A route ranking can be provided to the user forecasting the decision making more adaptive and naturist. This work will reduce the cost handsomely in order to select the route according to the user's customized preferences.

In this proposed sytem, costs of all the possible routes are studied to accept the favorable one. In conclusion, the route with the lowest cost is adopted using our system. This research can be continued for day-to-day life practices, by employing vehicle to vehicle (V2V) Communication. In future microelectronics, state of the art technology can be used to develop the hardware of this sytem.

ACKNOWLEDGEMENT

References

[1] Talebi, H. S. (2010). Dynamic Fuzzy Logic-Ant Colony System-Based. Applied Computational Intelligence and Soft Computing. Hindawi Publishing Corporation.

[2] Prasana Kumar and Raghavendra,(2011) "On the evaporation mechanism in the ant colony optimization," Annals computer science series, pp. 51-56

[3] Diogo Alves, Jelmer van Ast, Zhe Cong (2010). Ant Colony Optimization for Traffic Dispersion Routing, 13th International IEEE

[4] Annual Conference on Intelligent Transportation Systems, pp. 683-688

[5] Saliba, C. Farrugia, R.A (2010 quality of service aware Ant colony optimization routing algorithm, 15th IEEE Mediterranean Electrotechnical Conference , 343 - 347

[6] A. Broggi, M. C. (2003). An evolutionary approach to visual sensing for vehicle navigation. IEEE Transactions on Industrial Electronics, vol. 50, no. 1, pp. 18–29.

[7] H. Salehinejad, F. P. (2008). A new route selection system: multi-parameter ant algorithm based vehicle navigation approach. Vienna, Austria: in Proceedings of the International Conference on Computational Intelligence for Modeling, Control and Automatio, pp 1089-1094.

[8] Swann, S. A. (1979). Analysis of freeway traffic times-series data by using Box Jenkins techniques. Transportation Research, no. 72, pp. 1–9.

[9] I. Ohe, H. K. (1995). A method for automatic detection of traffic incidents using neural networks. in Proceedings of the Vehicle Navigation and Information Systems Conference in Conjunction with the Pacific Rim TransTech Conference, A Ride into the Future, pp 231-235.

[10] Attia, R. Rizk, R. Mariee, M. . (2009). Wireless and Optical Communications Networks, 2009. WOCN '09. IFIP International Conference on ,pp, 1-5

[11] Meldrum, C. T. (1995). Freeway data prediction using neural networks. in Proceedings of the Vehicle Navigation and Information Systems Conference in Conjunction with the Pacific Rim TransTech Conference, A Ride into the Future, pp. 225–230.

[12] Xiaoming You Xingwai Miao Sheng Liu. (2009) .quantum computing based ant colony optimization algorithm for travelling sales man problem, pp,359-362

Sagheer Abbas is M.Phil. Scholar in the field of Computer Sciences at National College of Business Administration and Economics NCBA& E, Lahore, Pakistan . His research area includes Artificial Intelligence, Quantum Computing and Sparse Distributed Memory.

M. Saleem Khan is an Assistant Professor at GC University, Lahore, Pakistan. He is working as director Computer Science Department in GC University, Lahore, Pakistan. He availed research fellowship at The School of Electronics & Engineering, University of Edinburgh, UK. and completed his Ph.D. thesis in the field of control systems: design, simulation and analysis in local and distributed environment. He contributed his services on various projects in the field of Advanced Electronics and Communication. His research interests include control systems design and industrial applications. He promoted a large team of Electronics researchers and organized this field in his country. Mr. Khan had also been served as a senior scientific officer in a classified defense research organization in his country.

Khalil Ahmed is the Director School of Computer Science at National College of Business Administration and Economics (NCBA&E) Lahore, Pakistan. He is an expert academician and passionately engaged in research. His area of research is machine consciousness, A.I. and knowledge management. He completed his PhD in the fields of Computer Sciences and Robotics.

Muhammad Abdullah is Ph.D. Scholar in the field of Computer Sciences at National College of Business Administration and Economics Lahore, Pakistan and his area of research is Artificial Intelligence, Machine Reflexes and Fuzzy Modeling.

.

Mr. UmerFarooq, is M.Phil scholar at School of Computer Science National College of Business administration & Economics(NCBA&E)Lahore, Pakistan and his area of research is Artificial Intelligence, Machine Reflexes and Fuzzy Modeling.

Smart Home Security System using Fuzzy Logic

Muhammad Anwaar Saeed, Muhammad Saleem Khan, Khalil Ahmed, Umer Farooq

Abstract—Secure home provide a sense of security to its owner. Home security has two aspects, inside and outside. Inside security covers the concept of securing home from threats like fire etc. whereas, outside security is meant to secure home against any burglar/intruder etc. This work is aimed to provide a solution for home security that takes decision dynamically using the pervasive devices. Also this solution has the feature to intimate security analysis results anywhere in the world using internet.

Index Terms —Home Security, Fuzzy logic, Smart Home, Pervasive devices, Dynamic Decision, Sensors, Threats.

————————————— ◆ —————————————

1. INTRODUCTION

IT is supposed that in future people will have an invisible and ubiquitous computing infrastructure to perform different activities both at work and home. Modern home requires easy to use and synergistic devices [1]. Spinellis [1] has proposed the idea of Information Furnace to integrate different available devices in home for different services. Currently, a variety of devices is available in modern home with different access modes and interfaces which results in complexity for end user. The Information Furnace model proposes the synergies among these devices [1].

A smart homecare system using smart phones, wireless sensors, web servers and IP webcams is proposed by Leijdekkers et al [2]. Proposed smart homecare system provides facility to elderly people to check their health and status and provides an easy way to contact to hospital in an emergency [2].

Ghorbel et al [3] have proposed the integration of networking and communication technologies in the smart homes concept dedicated to people with disabilities. Proposed concept is based on the UPnP protocol to discover and control devices indoor and uses wireless technologies to enhance mobility [3].

Popescu et al [4] have proposed a security architecture allowing digital rights management in home networks consisting of consumer electronic devices. In the proposed model, devices are allowed to establish dynamic groups in an environment where legally acquired copyrighted content are seamlessly transmitted between devices. Popescu et al [4] have claimed that connectivity between devices has a minimal reliance on public key cryptographic operations.

Gao et al [5] have suggested the concept of a self-programming thermostat that without any human intervention creates a best possible setback schedule by sensing the possession statistics of a home. The system monitors possession using simple sensors in the home and the user defines the desired balance between energy and comfort using a single knob. It is observed that this approach has an advantage over EnergyStar setback schedule approach by reducing the heating and cooling demand by up to 15% [5].

Use of wireless sensor networks is a low cost, easy way to monitor physical environments. By integrating the context-aware capability of wireless sensor networks into surveillance systems is an attractive trend. Tseng et al [6] have proposed iMouse system, which combines wireless sensor networks, and surveillance technology, to support intelligent mobile surveillance services. Proposed (iMouse) system consists of mostly inexpensive static sensors to monitor environment and few expensive mobile sensors to perform some advanced actions [6].

Kim et al [7] have proposed a Home Security system based on Sensor Network (HSSN) configured by sensor nodes including radio frequency (RF), ultrasonic, temperature, light and sound sensors. Proposed system has the capability to acknowledge security alarm events that are acquired by sensor nodes [7].

Initially fuzzy logic control was introduced to model free control design approach but was criticized due lack of systematic stability analysis and controller design. G. Feng [8] has shown the current improvement in the analysis and design of model based fuzzy control systems.

2. SYSTEM ARCHITECTURE:

From the literature, it is observed that previous home security models have considered some limited security concerns. Therefore one security model may be good in one situation but cannot provide the required results in other situations. To provide optimal home security solution, a new model is proposed. In this proposed model, sensors are used to detect abnormalities within the house or outside the house. There is a dedicated server for the sensors used to collect data inside the house. This server is responsible to collect information transmitted by the sensors and then analyze to detect any abnormality. Similarly, a separate server is used to process the information transmitted by sensors located

————————————————
- **Muhammad Anwaar Saeed** *is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan. He is currently working as faculty member in VU, Lahore, Pakistan (e-mail: m_a_saeed@yahoo.com)*
- **M. Saleem Khan** *is with the Computer Science Department as Director in GC University Lahore, Pakistan (e-mail: mskgcu@yahoo.com).*
- **Khalil Ahmed** *is with the School of Computer Science at NCBA&E Lahore, Pakistan; He is an expert academician and passionately engaged in research. (E-mail: drk@ncbae.edu.pk).*
- **Umer Farooq** *is with Computer Science department as a faculty member and a research fellow at NCBA&E Lahore, Pakistan (e-mail: umer@ncbae.edu.pk).*

outside the house. Both these servers are connected to a main server which process the information provided by these servers. Fuzzy logic is used to detect any abnormality. In case a threat is detected then main server report about the threat to concern people using internet besides setting the alarms on. Following is the graphical representation of basic logic of the proposed system:
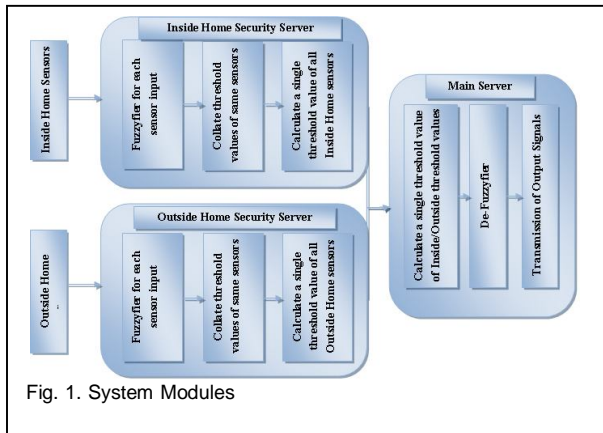

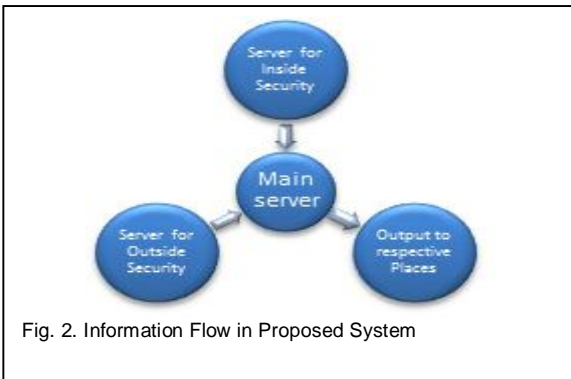Fig. 1. System Modules

Following diagram depicts the flow of information between different major modules

Fig. 2. Information Flow in Proposed System
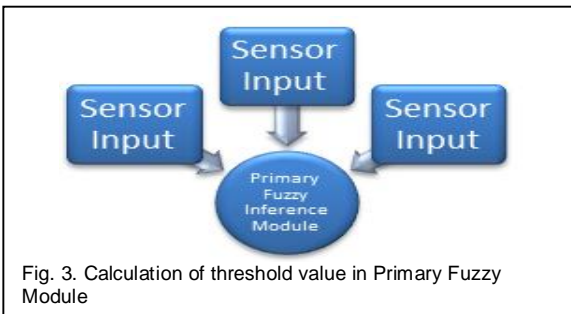
Fig. 3. Calculation of threshold value in Primary Fuzzy Module

Six type of input is provided to the system. Multiple sensors of each type are used to collect data. All inputs of same sensor type are provided to an initial fuzzy inference module, which is responsible to calculate the threshold value. These calculated threshold value of each input type is then provided to respective server responsible for inside or outside security. An overall threshold value of these six initial threshold values is

separately calculated using fuzzy logic module on inside/outside security servers respectively. Both inside/outside security threshold values are provide to main server for analysis. Final decision is made based on these values. If any of the value is above the critical value then alarm signal is generated to respective person/department. Using this method, it is possible to generate different output alarms considering the intensity and relevance of threshold value to that specific person/department. Different processes in decision making are described in the following figures. Figure 4 and figure 5 represent the concept of fuzzy inference
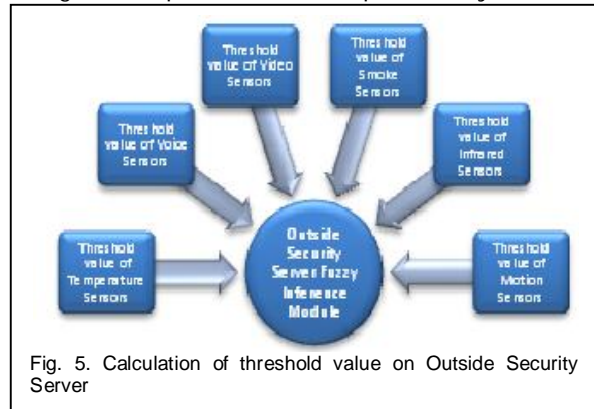
Fig. 5. Calculation of threshold value on Outside Security Server

module at Inside/Outside Home Security Server respectively.

Threshold values calculated at Inside/Outside Home Security Servers are collated at main server for decision making process. This concept is represented in figure 6. After collation process, threshold value is calculated and
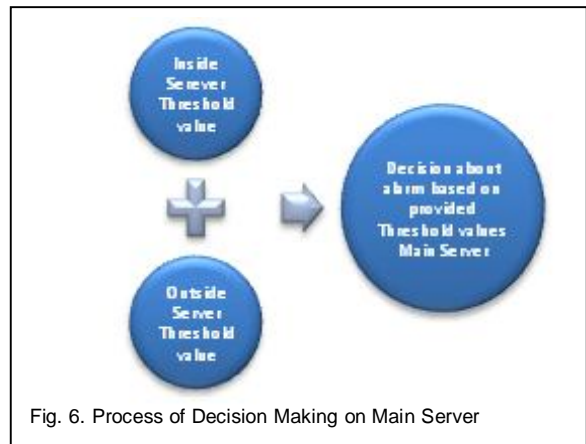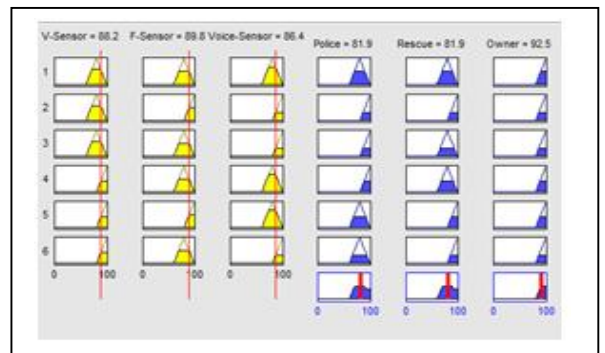

Fig. 6. Process of Decision Making on Main Server

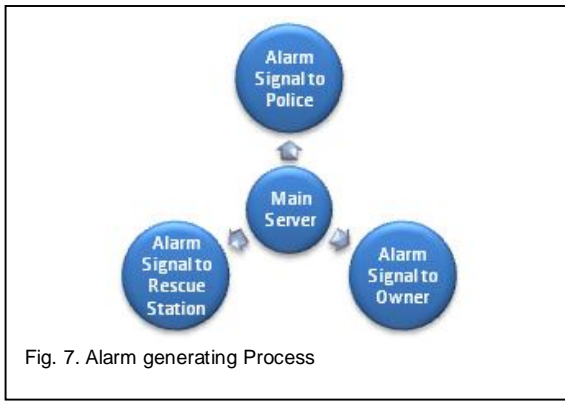alarm signal type for each desired destination is calculated as shown in figure 7.

Fig. 7. Alarm generating Process

3. FUZZY PARAMETERS FOR DECISION MAKING

Configuration of fuzzy parameters for the proposed system is discussed in this section. Home security system is configured by sensor nodes connected to server. These sensor nodes include radio frequency, ultrasonic, temperature, light, sound and video sensors. Threshold value for each input is above 90% and for a video sensor, used in outside security, distance threshold is taken as 1 feet. If value is increased from any threshold value then alarm is on, and notified to specified location through internet. Inputs to the system and respective outputs from fuzzy inference system are shown in figure 8. Respective member functions are shown in figure 9 for the sample scenario.
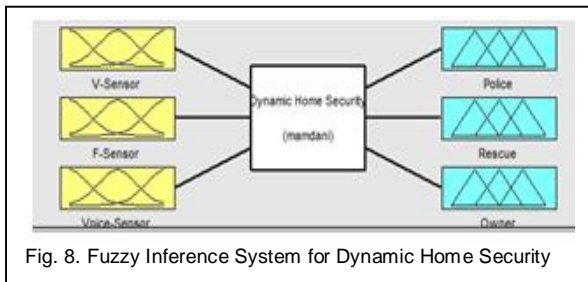
Fig. 8. Fuzzy Inference System for Dynamic Home Security
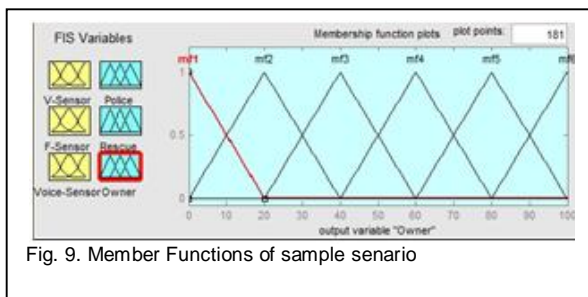
Fig. 9. Member Functions of sample senario

For a sample scenario, where only three types of sensors are used namely video, fire and voice. Effect of threshold values of these sensors and respective fuzzy logic based output using MATLAB is shown in figure 10.

In figure 11, 3D graph show the relationship between voice sensors, fire sensors and output threshold for rescue.

In figure 12, 3D graph show the relationship between voice sensors, fire sensors and output threshold for police.
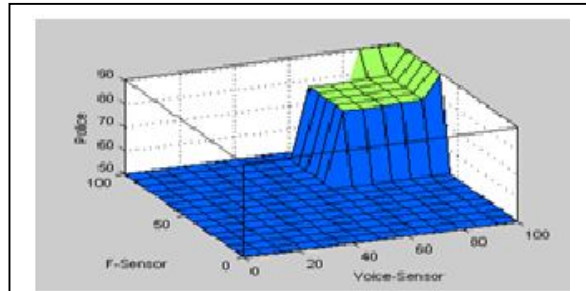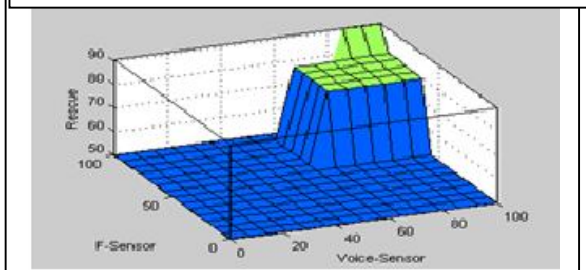

Fig. 12.

Fig. 11.

In figure 13, 3D graph show the relationship between voice sensors, fire sensors and output threshold for owner.
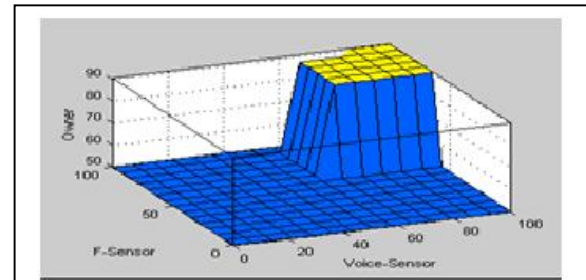

Fig. 13.

In figure 14, 3D graph show the relationship between voice sensors, video sensors and output threshold for police
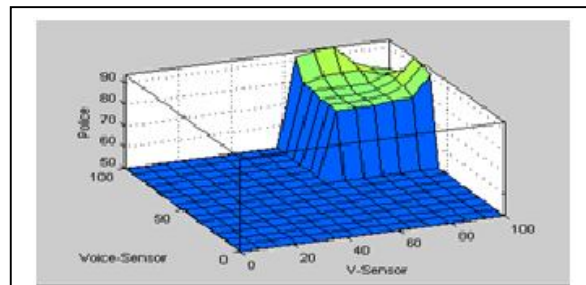

Fig. 14.

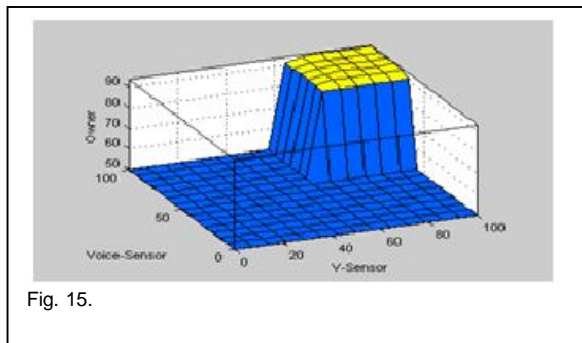In figure 15, 3D graph show the relationship between voice sensors, video sensors and output threshold for Owner.

Fig. 15.

In figure 16, 3D graph show the relationship between voice sensors, video sensors and output threshold for
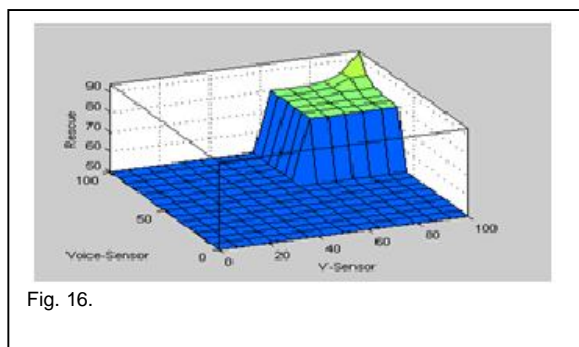

Fig. 16.

Rescue.

From the sample scenario, it is observed that fuzzy logic based home security system provides flexibility to detect different nature of threats and respective outputs. Like in the above scenario, it can be seen that for the same input different output thresholds are provide for owner, police and rescue.

4. CONCLUSION

A fuzzy logic based home security system is proposed. It is observed that using this proposed concept, a better and flexible home security is provided. Proposed system inherits the properties of fuzzy logic and thus provides intermediary values as compare to Boolean logic bi-value outputs.

REFERENCES

[1] D. D. Spinellis, "The information furnace: consolidated home control," PersUbiquitComput, vol. (7), pp. 53–69, 2003.

[2] P. Leijdekkers, V. Gay, and E. Lawrence, "Smart Homecare System for Health Tele-monitoring," Proceedings of the First International Conference on the Digital Society (ICDS'07), IEEE Computer Society 2007.

[3] M. Ghorbel, M. Segarra, J. Kerdreux, R. Keryell, A. Thepaut, and M. Mokhtari, "Networking and Communication in Smart Home for People with Disabilities," Computers Helping People with Special Needs, Springer Berlin / Heidelberg, pp.624, 2004.

http://dx.doi.org/10.1007/978-3-540-27817-7_138

[4] B. C. Popescu, B. Crispo, A. S. Tanenbaum, and F. L. A. J. Kamperman, "A DRM Security Architecture for Home Networks," DRM'04, Proceedings of the 4th ACM workshop on Digital rights management, October 25, 2004, Washington, DC, USA.

[5] G. Gao, and K. Whitehouse, "The Self-Programming Thermostat: Optimizing Setback Schedules based on Home Occupancy Patterns," Proceedings of BuildSys'09, November 3, 2009, Berkeley, CA, USA.

[6] Y. Tseng, Y. Wang, and K. Cheng, "An Integrated Mobile Surveillance and Wireless Sensor (iMouse) System and Its Detection Delay Analysis," MSWiM'05, October 10–13, 2005, Montreal, Quebec, Canada.

[7] Y. Kim, H. Kim, S. Lee, and K. Lee, "Ubiquitous Home Security Robot based on Sensor Network," Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'06), 2006

[8] G. Feng, "A Survey on Analysis and Design of Model-Based Fuzzy Control Systems," IEEE Transactions on Fuzzy Systems, vol. 14(5), pp. 676-697, October 2006.

Muhammad Anwaar Saeed is a Ph.D. scholar at school of computer science, NCBA&E, Lahore, Pakistan (phone: 092-333-4400473; e-mail: m_a_saeed@ yahoo.com).

M. Saleem Khan is with the GC University Lahore Pakistan, working as Director Department of Computer Science and currently availing a research fellowship at The School of Electronics &Engineering in Edinburgh University UK (e-mail mskgcu@yahoo.com).

Khalil Ahmed is heading the School of Computer Science at NCBA&E, Lahore, Pakistan. He is an expert academician and passionately engaged in research. His area of research is machine consciousness, A.I. and knowledge management. (e-mail drk@ncbae.edu.pk)

UmerFarooq, is M.Phil scholar at School of Computer Science National College of Business administration & Economics(NCBA&E)Lahore, Pakistan and his area of research is Artificial Intelligence, Machine Reflexes and Fuzzy Modeling. (e-mail: umer@ncbae.edu.pk).

Autonomous System Controller for Vehicles Using Neuro-Fuzzy

Umer Farooq, M. Saleem Khan, Khalil Ahmed, M. Anwaar Saeed and Sagheer Abbas

Abstract— this paper presents the approach of neuro fuzzy systems to design autonomous vehicle control system. The purposed intelligent controller deliberates obstacles avoidance, unstructured environment adaptation and speed scheduling of autonomous vehicle based on neuro-fuzzy with reinforcement learning mechanism. The purposed system provides the autonomous vehicle navigation and speed control in unstructured and unsafe environment..

Index Terms— Autonomous Vehicle Controller; Autonomous Vehicles Speed Control; Neuro-Fuzzy Vehicle Control System; Obstacles Avoidance System.

————————————— ◆ —————————————

1. INTRODUCTION

The autonomous vehicles system is an important research area in which a lot of work has been done in vehicles controlling systems. Autonomous vehicles have prospective application in many fields, such system are used for assisting human in dangerous environments, conducting military and performing routine tasks for industry[1]-[2].

Traffic congestion problem due to human driving involves reaction time weakness and error in judgment that causes accidents and affects the smooth traffic flow. The automated highway systems are purposed for the main objectives i.e. automated vehicle control and traffic control technologies. Intelligent cruise control is an advance vehicle control system that manages an appropriate actuator action for safe speed and distance control [3].

A lot of systems are purposed about the intelligent transport system, including smart vehicles, driving safety and traffic mobility. There are many applications for the vehicles control like driver assistants system [4]. GPS based navigation control, environment awareness system and others. GPS/INS sensor base path tracking system, steering controller incorporated with speed controller based on the finite preview optimal control method [5]

_____

- **Umer Farooq** is with Computer Science department as a faculty member and a research fellow at NCBA&E Lahore, Pakistan (e-mail: umer@ncbae.edu.pk).
- **M. Saleem Khan** is with the Computer Science Department as Director in GC University Lahore, Pakistan (e-mail: mskgcu@yahoo.com).
- **Khalil Ahmed** is with the School of Computer Science at NCBA&E Lahore, Pakistan; He is an expert academician and passionately engaged in research. (e-mail: drk@ncbae.edu.pk).
- **Muhammad Anwaar Saeed** is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan. He is currently working as faculty member in VU, Lahore, Pakistan (e-mail: m_a_saeed@yahoo.com)
- **Sagheer Abbas**, is with Computer Sciences department as faculty member and a research fellow at NCBA&E Lahore, Pakistan.(e-mail: sagheeer@hotmail.com).

This research paper is organized as follows: section II describes the brief introduction of fuzzy inference system (FIS), and Neuro-Fuzzy systems (NF). Structure of the proposed autonomous vehicle Speed Controller is discussed in section III, while section IV is the simulation in MAT Lab. Section V is having experimental results and discussion. This paper is concluded with remarks on the proposed model and future work in section VI. In section VII acknowledgement of this work is mentioned.

2. OVERVIEW OF FUZZY INFERENCE SYSTEM AND THE NEURO-FUZZY SYSTEM

Fuzzy logic is the formalization of human capabilities, reasoning and rational decision making in an environment of imprecision, uncertainty, incompleteness, inconsistent of information and performing physical and mental task without measurement and calculation [6].
The applications of fuzzy logic are multi-disciplinary in nature include automatic control, consumer electronics, signal processing, information retrieval, database management, computer vision, data classification, decision making and so on [7].

The remarkable perspectives of fuzzy logic systems are neuro-fuzzy systems, the first hybridization of fuzzy logic systems and artificial neural networks in soft computing. The neuro fuzzy systems have fuzzy logic knowledge representation ability with the learning capability of artificial neural networks. Fuzzy logic and artificial neural networks are used for the different intelligent systems. ANNs distributed processing capability of large no of computational simple element while fuzzy logic is related to reasoning on a higher level. Fuzzy systems do not have the capability of learning, adaptation and distributed computing. The neuro-fuzzy systems make the adaptation mechanisms of neural network with the symbolic components of fuzzy inference system [8].

Figure 1 shows a fuzzy based speed control with four sided sensors inputs. The two types of inputs are considered using hard and the flexible conditions, the variation of speed will accord on the flexible condition and the speed will be zero whenever the hard condition meets.

The inputs are fuzzifiered by the fuzzyfier module and then provide these inputs to the FIS module consisting of three conceptual components: a rule base, a data base and a reasoning mechanism. Fuzzification represents the member functions of the linguistic terms of the input variables and defuzzification to generate defuzzified values that represent the achieved level of a speed for the fuzzified variable
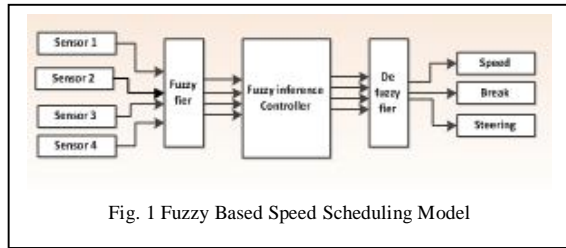

Fig. 1 Fuzzy Based Speed Scheduling Model

There are many reinforcement learning techniques to solve the problems. The selection of optimal policy computed by the reinforcement learning can be used to generate fuzzy rules. Neuro fuzzy systems such adaptive network based fuzzy inference system can be used to learn fuzzy if-then rules for the weight selection problem using the training data obtained from RL's optimal policy [9].

The autonomous vehicle should be able to manage the driver task therefore it has to manage complex traffic situations in real time The autonomous vehicles control systems research primary goal to make driving on road harmless and easy. The suggested neuro-fuzzy inference engine the adaptation of environment and controlling the speed of the vehicle.

3. DESIGN AND STRUCTURE OF PROPOSED SYSTEM

The figure 2 shows the data flow of the vehicle control. In the figure sensed data hard and flexible is provided to the fuzzy inference with incorporating the measurement. The fuzzy inference module groups measurement into the three major classes, far, medium and close. The action of the each class is different for each group. If the input leading distance close then the system will decrease the speed if it is in medium, then speed will remain constant and if the Leading distance is far then speed will be increased.
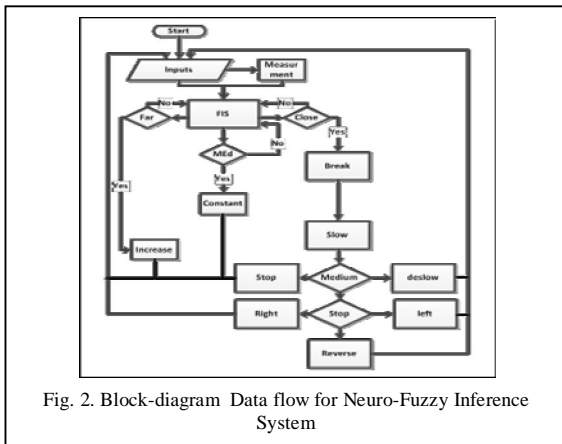

Fig. 2. Block-diagram Data flow for Neuro-Fuzzy Inference System

To control a vehicle speed autonomously, the control system must have the ability for the adaptation of environment. In the purposed controller of speed three types of sensors are used for the sensing the environment. The location of the current state can be determined by the GPS system or by calculating a travel path from its starting point with the use of electronic compass. These sensors tackle flexible conditions like road condition environment condition (pleasant or stormy), leading distance and hard condition like accident on road or blocked road. On the flexible condition the variation of speed will be calculated, it can be a high medium or low based range of the variable value and if the hard condition meets, the speed will be zero. On the sensed information steering, transmission, acceleration and braking modules will be activated. These modules are connected to the neuro fuzzy inference system to adapt the inputs and calculate the desired speed on the bases of the inputs and the desire input for the steering, transmission, acceleration and braking modules.
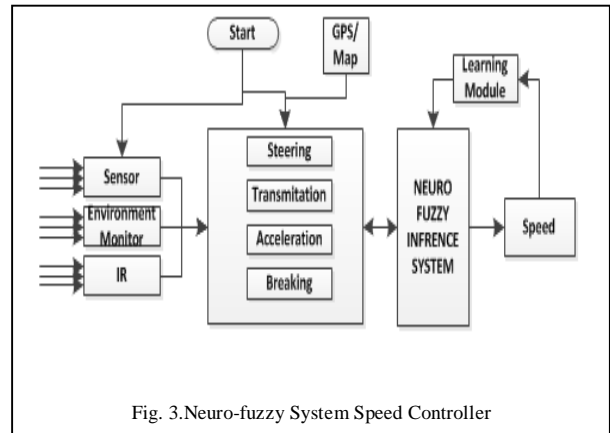

Fig. 3.Neuro-fuzzy System Speed Controller

4. NEURO-FUZZY SYSTEM STRUCTURE

The figure 4 shows the ANFIS structure. ANFIS is the hybrid of Artificial Neural Networks and Fuzzy inference system. ANNs map input to output through the collection of layered processing elements that are organized in parallel by synaptic junctions. Each pass data, signals propagate from the input to output layer to produce output which is compared to the desired output. The difference between output and desired output is then adjusting the synaptic connection (weight). This process gives ANNs, the ability of looking for patterns to provide the advantage of learning. FIS is based on Fuzzy Logic, IF-THEN rules and fuzzy reasoning which can link to human reasoning through linguistics variables. The combination of ANNs and FIS of ANFIS, integrates the benefits of the individual intelligent systems to develop the superior technique that can optimally model the dynamics of difficult systems such as autonomous vehicle speed controller.
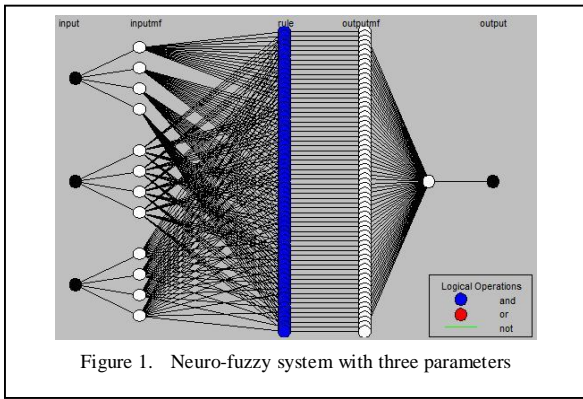
Figure 1.   Neuro-fuzzy system with three parameters

ANFIS is 6 layers Feedforward neural network and of the sugeno FIS type. In this paper we present 3 inputs and 1 output for the ANFIS with 3-dimensional input-output space. The layer 1 represents the input layer of raw data entry for ANFIS. Neural network clasifies data and looks for patterns within it. The input data in 3- dimentional space, it is classified in groups calld fuzzy space. The raw input data or crisp value are compared with the membership function in the antecedent of the antecedent of the rulesof ANFIS, to determined the degree to which inputs. EM, RC and LD belong to fuzzy sets S1 S2 and S3 respectively. The degree to which the inputs lie in the fuzzy space is given the value between 0 and 100. This process is known as fuzzification and takes place in layer 2 as given in Equation (1), (2) and (3).

$$L_i^2 = \mu_{Ai}(RC) \tag{1}$$

$$L_i^2 = \mu_{Bi}(LD) \tag{2}$$

$$L_i^2 = \mu_{Ci}(EM) \tag{3}$$

where $i = 1,2,3 ....$

Once the location in the fuzzy space identified, then the product of the degrees to which inputs satisfy the membber ship function is found. This product called the firing strength of a rule and represented in layer 3 shown in Equation (4), called rule layer.

$$l_i^3 = w_i = \mu_{Ai}(RC) \cdot \mu_{Bi}(LD) \cdot \mu_{Ci}(EM) \tag{4}$$

where $i = 1,2,3 ....$

In layer 4, the normalization layer, the ratio of each rule's firing strength is calculated according to Equation (5).

$$L_i^4 = \overline{w}_i = \frac{w_i}{w_1 + w_2 + w_3 + \cdots + w_{64}} \tag{5}$$

Layer 5 is the defuzzification layer; the output of each input node is the weighted consequent value as in Equation (6).

$$L_i^5 = \overline{w} f_i = \overline{w}(p_i + q_i + r_i) \tag{6}$$

Summation layer , layer 6 is an output layer which is the sum of all the outputs of the layer 5 as in Equation (7).

$$L_i^6 = \sum \overline{w} f_i = \frac{\sum w_i f_i}{\sum w_i} \tag{7}$$

The number of epochs was 10000 for training. The numbers of MFs for the input variables are 4. The number of rules is then 64(4×4×4) then the triangular MF is used for three input variables.

Therefore, the ANFIS used here contains a total 292 parameters, of which 36 (4×3+4×3+4×3=36) are the premise parameters and 256 (4×64) are the consequent parameters.

The training testing and checking errors obtained from the ANFIS are 4.194, 4.2913and 4.2134 respectively.

5.   RESULTS AND DISCUSSION

The result form the ANFIS, grid partitioning technique showed training error, testing error and checking error as in Figure 5 Figure 6 and Figure 7.

Table 1 show; if all variables are found at their best value then the speed will be very high to be adopted. Likewise if any varibale has the low value among sensing parameters then the speed will be low.

TABLE I.   INPUT OUTPU LINGUISTIC REPRESENTAION.

RC	EM	LD	Speed
V.High	V.High	V.High	V.High
High	High	High	High
Medium	Medium	Medium	Medium
Low	Low	Low	Low
V. Low	V. Low	V. Low	V.Low

The data taken for system from 0 to 100 for all inputs variable and 0 to 120 for the output variable. The Sample of 400 pairs out of the total possiable combination 105 used for the traning of the ANFIS with the backpropagation technique, 53 pairs of data is used for testing and 40 pairs of data is used for the checking of the network.

Figure 5 shows the traning results, the sample of 400 pairs is used for traning of the neuro-fuzzy system. The error rate is 4.1943 on the 25000 epochs.
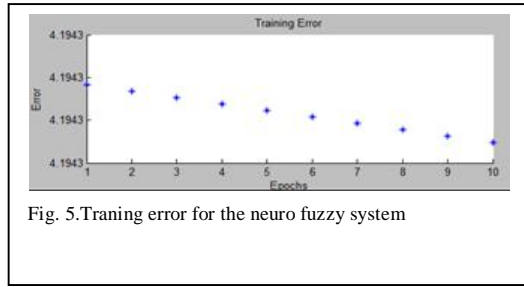

Fig. 5.Traning error for the neuro fuzzy system

Figure 6 shows the testing results, the sample of 53 pairs of data set is used for the testing of the neuro fuzzy system. Error rate of the testing of the system was 4.2913.
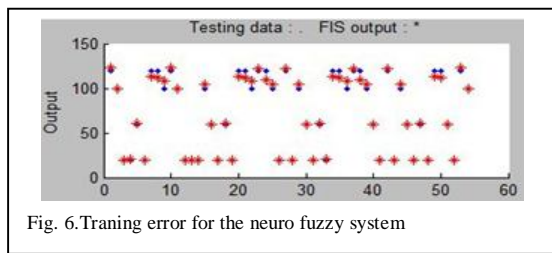

Fig. 6.Traning error for the neuro fuzzy system

Figure 7 shows the checking result of the neur-fuzzy system. Sample of 53 data pairs is used for the checking the resultant erro is 4.2134.
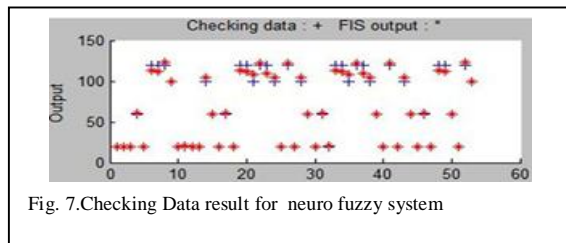

Fig. 7.Checking Data result for neuro fuzzy system

The Figure 8, Figure 9, and Figure 10 show the surface graps of the input variables. Figure 8 shows the relation of EM and RC on the speed. If EM and RC have the low sensing intesity the speed will be high. Figure 9 shows the relation of LD, RC and figure 10 shows EM and LD.
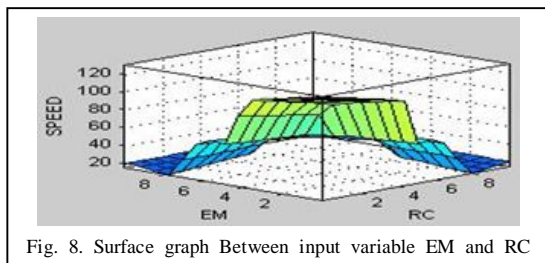

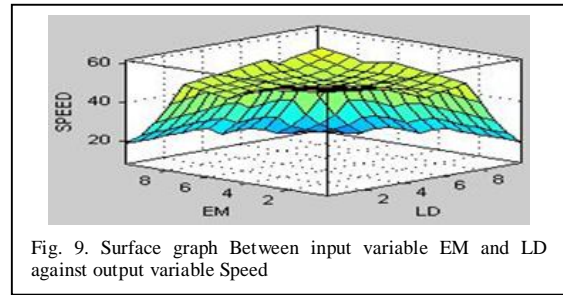Fig. 8. Surface graph Between input variable EM and RC

Fig. 9. Surface graph Between input variable EM and LD against output variable Speed
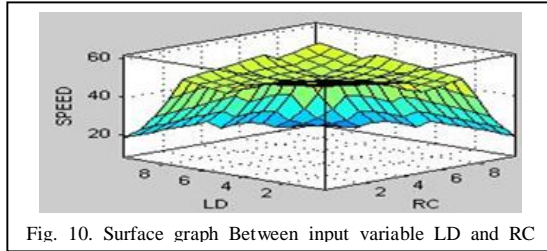

Fig. 10. Surface graph Between input variable LD and RC

Figure 11 concludes the resutls, if the speed lies in the low category then systen will check the LD. If LD is detected in very high intensity then the speed will increase, if it is perceived in medium then speed will be normal and LD is sensed very low the system will propose for the lane change.
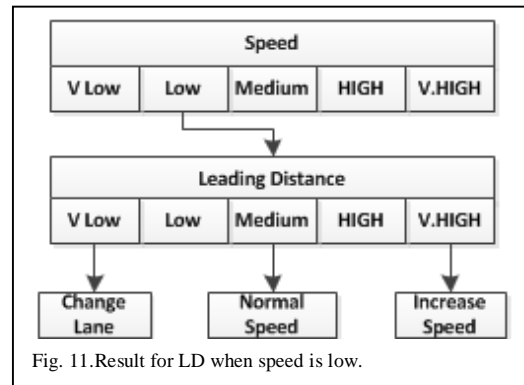

Fig. 11.Result for LD when speed is low.

6. CONCLUSION AND FUTURE WORK

The human reaction time, weakness and error in judgment cause accident and affect the smooth traffic flow. The autonomous speed controller deliberates obstacles avoidance and unstructured environment adaptation for the autonomous vehicle.The autonomous vehicle will manage the driver task with complex traffic situations in real time to make driving on road harmless and easy.A neuro fuzzy approached is used for the autonomous vehicle speed controller. The system knowledge representation and the learning mechanism make easy to adapt the environment and control the speed to avoid the traffic congestion and collision.

In future state of the art Microelectronics technology can help to develop the hardware of the system to achieve the required goal.
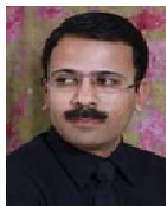
REFERENCES

[1]   D.C.Slaughter; Giles, D.K.; Downey, D. "Autonomous Robotic Weed Control Systems": Elsevier, pp. 63-78, 2008.

[2]   Real-time indoor autonomous vehicle test environment. How, J.P; BEthke, B; Frand, A.; Dele, D. April 2008, Cambridge : IEEE, 2008, Vol. 28 issue:2.

[3]   An Intelligent Longitudinal Controller for Application in Semiautonomous Vehicles. Lin Cai;AhmadB.Rad. s.l. : IEEE, 2010.

[4]   An Image Processing System for Driver Assistance. U.Handmann; T.Kalinke;C. Tzomakas; M.Werner;W.v.Seelen. s.l. : ELSVIES, 2000.

[5]   Design and Testing of a Controller for Autonomous Vehicle Path Tracking Using GPS/INS Sensors. JuyongKand; Rami Y. Hindiyeh; seugnwuk Moon; J. Christian Gerdes .Kyongsuyi. Seoul,Korea : international Dederation of Automatic Control, 2008. Proceedings of The 17th World Congress.

[6]   Is there a need for fuzzy logic? Zadeh, Lotfi A. s.l. :Exsevier, 2008, pp. 2751-2779.

[7]   Application of Neuro-Fuzzy Systems to Behavioral Representation in Computer Generated Forces, Gary R. George P. E. and Frank Cardullo,

[8]   An Overview on Soft Computing in Behavior. Hoffmann, Frank.2003Volume 2715/2003, 544-551, DOI: 10.1007/3-540-44967-1_65

[9]   Neuro-Fuzzy Learning of Strategies for Optimal Control Problems. Kamali, Kaivan, et al., et al. 2002.

Mr. UmerFarooq, is M.Phil scholar at School of Computer Science National College of Business administration & Economics(NCBA&E)Lahore, Pakistan and his area of research is Artificial Intelligence, Machine Reflexes and Fuzzy Modeling.

(e-mail: umer@ncbae.edu.pk).

Mr. M. Saleem Khan is an Assistant Professor at the GC University Lahore Pakistan. Currently he is working as director Computer Science Department in GC University, Lahore, Pakistan. He availed research fellowship at The School of Electronics & Engineering, University of Edinburgh, UK and completed his Ph.D. thesis in the field of control systems design, simulation and analysis in local and distributed environment. He contributed his services on various projects in the field of Advanced Electronics and Communication. His research interests include control systems design and industrial applications. He promoted a large team of Electronics researchers and organized this field in his country. Mr. Khan had also been served as a senior scientific officer in a classified defense research organization in his country. (e-mail: mskgcu@yahoo.com).

Khalil Ahmedis the Director School of Computer Science at National College of Business Administration and Economics (NCBA&E) Lahore, Pakistan. He is an expert academician and passionately engaged in research. His area of research is machine consciousness, A.I. and knowledge management. He earned his PhD degree in the fields of Computer Sciences and Robotics from USA and Japan.

 (e-mail: drk@ncbae.edu.pk).

Muhammad Anwaar Saeed is a Ph.D. scholar at school of computer science, National College of Business Administration and Economics (NCBA&E), Lahore, Pakistan. His research area is A.I. and knowledgeManagment, Data security and Fuzzy Modeling.

(E-mail: m_a_saeed@yahoo.com)

Mr. Sagheer Abbas, M.Phil. Scholar in the field of Computer Sciences at National College of Business Administration and Economics NCBA& E, Lahore, Pakistan . His research area includes Artificial Intelligence, Quantum Computing and Sparse Distributed Memory.

(E-mail: sagheeer@hotmail.com).

FIS Based Speed Scheduling System of Autonomous Railway Vehicle

Aiesha Ahmad, M.Saleem Khan, Khalil Ahmed, Nida Anwar and Umer Farooq

Abstract—This paper presents the design model of speed scheduling system of autonomous railway vehicle control using fuzzy inference system (FIS). Successful development of speed scheduling and maintaining system plays crucial role to make autonomous railway vehicle control system more effective under constraint of uncertain conditions. This research work emphasis to develop the speed scheduling system with capability to adjust in uncertain conditions magnificently by improving performance, stability, controllability and safety of railway vehicles and ultimately reduce the risk to meet the needs of modern trend of autonomous control system. The proposed design model is capable to successfully cope with hard conditions; junction track information (JTI), crossing gate information (CG) and track clearance (TCL), and flexible conditions; vehicle tilting (VT), track conditions (TC) and environment monitoring (EM) using fuzzy inference system with better and quicker response with human knowledge incorporation. This system will be helpful to successfully maintain the speed of railway vehicles with environment monitoring, time scheduling and minimizes the risk of overturning.

Index Terms— Autonomous Railway Vehicle, Control System, Environment Monitoring, Fuzzy Inference System, Speed Scheduling, Time Scheduling.

———————————————— ◆ ————————————————

1 INTRODUCTION

ADVANCEMENT in railway vehicles technology has been increased from last few years to facilitate passengers by improving performance with speed, time scheduling, traffic control and passenger management [1]. The importance of control, management and monitoring for railway vehicles progress gradually under modern solution of embedded systems, software based computer aided control systems, sensors and data communication technologies. The design and development of agent base autonomous railway vehicle control system is considered important for flexible and well established network to enable collaboration between centralized and distributed systems of railway tracks [2]. In these autonomous railway control systems, agents are defined as condition monitoring units with capability to collect information independently and control system autonomously according to their design control [3]. The monitoring and control of railway vehicles are mainly focused on speed management and scheduling, traffic control and time scheduling. Speed scheduling plays vital role for successful development of autonomous railway control system by fo-

cusing on hard conditions; Junction track information (JTI), crossing gate information (CG) and track clearance (TCL), and flexible conditions; vehicle (VT) tilting, track condition (TC) and environment monitoring (EM). To maintain information about track condition with support of acceptable ride quality and tilting of trains around curved tracks with speed adjustment have been done through some sensors, mathematical modular technique and kalman filtering for data estimation and tracking in existing systems. Mathematical modular techniques sometimes unable to meet the need of real time environment due to lack of flexibility while kalman filtering technique also can't perform well in presence of noise in initial stages with modeling of system [4]. The complexity and dynamic nature of autonomous railway control system is needed some sophisticated method with domain knowledge representation and save time with quick response to handle uncertain situations successfully during running on track. Fuzzy inference system has capability to perform uncertain reasoning under incorporation of human knowledge in real time environment with better and quicker response. The proposed design model will be capable to adjust speed under uncertain conditions with high safety, performance and time management.

The arrangement of this research paper is as follows: section 2 consists of brief overview of fuzzy logic and fuzzy inference system that helps to understand the importance of fuzzy base systems. Structure of proposed speed scheduling system is discussed in section 3. Design algorithm is presented in section 4 and results and discussion is described in section 5 on the basis of design algorithm. Section 6 presents conclusion and future work on the basis of design algorithm.

————————————————————

- **Ms. Aiesha Ahmad** is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan (e-mail: aiesha@ncbae.edu.pk).
- **Dr. M. Saleem Khan** is with the Computer Science Department as Director in GC University Lahore, Pakistan (e-mail: mskgcu@yahoo.com).
- **Dr. Khalil Ahmed** is with the School of Computer Science at NCBA&E Lahore, Pakistan; He is an expert academician and passionately engaged in research. (e-mail: drk@ncbae.edu.pk).
- **Ms. Nida Anwar** is with Computer Science department as research fellow at NCBA&E Lahore, Pakistan. She is currently working as faculty member in VU, Lahore, Pakistan (e-mail: nidaanwar85@gmail.com)
- **Mr. Umer Farooq** is with Computer Science department as research fellow and faculty member at NCBA&E Lahore, Pakistan. (e-mail: umer@ncbae.edu.pk).

2 OVERVIEW OF FUZZY LOGIC AND FUZZY INFERENCE SYSTEM

Imprecise modes of reasoning are successfully demonstrated by fuzzy logic with appropriate answer that considered fundamental aspect of human rational decision making under ambiguous and uncertain environment based on incomplete, inexact or bit reliable knowledge [5], [6], [7], [8]. Fuzzy logic simulates human reasoning through mathematical theory of fuzzy sets with power of high precision while mapping of input values to an output values is done using process of fuzzy inference which drives to the final decision. Therefore fuzzy inference system (FIS) is considered as one of main computing framework in artificial intelligent (AI) that integrate human knowledge with idea of fuzzy IF-THEN rules to determine imprecise and uncertain reasoning in real time environment. FIS provides better, quicker and more appropriate solution as compared to traditional approaches because these approaches worked with crisp set of distinct and precise boundaries while in fuzzy set transition from non-membership to membership is gradual.

3 STRUCUTRE OF PROPOSED SPEED SCHEDULING SYSTEM

The fuzzy inference based proposed speed scheduling system consists of preloaded information of track in form of root chart and design to cope with uncertain conditions successfully with minimum response time. Railway track system receives two main conditions (hard and flexible) from environment and uses sensors to differentiate between these conditions. In case of flexible conditions (FC): VT, EM and TC observed through sensors and given to FIS. These sensors are capable to monitor EM, TC and VT individually by subdividing into 0 to 5 volt in which 0 volt represents the absence of these flexible conditions with no need of speed adjustment while from 1-5 volt shows gradual increase in these conditions with respective decrease in speed, like 5 volt represents worst condition against one or more flexible conditions (VT, TC or EM) with slow speed. Hard conditions (HC) like JTI, TCL and CG are sensed through sensors at particular distance and ultimately stop the train whatever adjusted speed may be after existence of flexible conditions. Fig.1 is representing the basic structure of FIS based speed scheduling system in presence of both FC and HC using fuzzy logic.

The proposed system will be capable to reduce speed or stop by increasing time according to requirement in uncertain environment. Then it will compare the increase time with root chart estimated time to reach next junction and increase speed to its maximum possible limit to overcome the delay after handling uncertain situation.
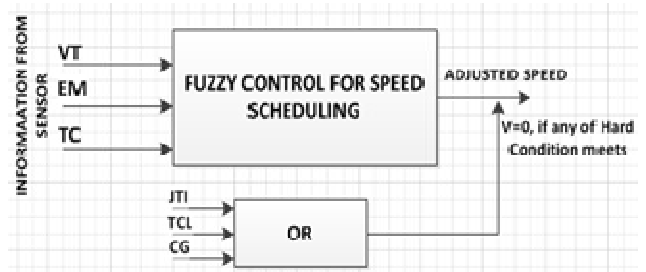

Fig.1. Block Diagram of Speed Control System using FIS.

The design procedure can be explained with the help of mathematical equations.

In the start, speed is

$$S = V * t$$

After uncertain conditions, the speed reduced as

$$S = (V + \Delta V)\, t$$

Change in velocity calculated as

$$S/t - V = \Delta V$$

Then this change in speed will compare with root chat estimated speed to overcome the delay to reach the next junction according to root chart calculated time. An overview of speed scheduling system with fuzzy control and speed adjustment is given in Fig. 2.
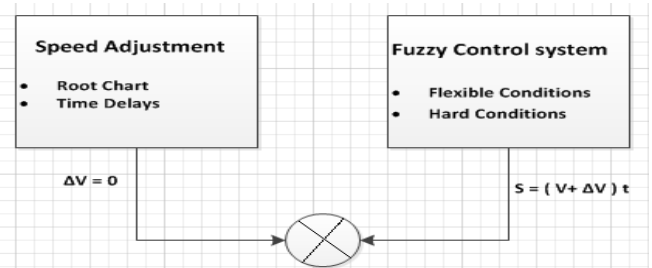

Fig.2. Block Diagram of Speed Scheduling System.

Fuzzy control system is used to adjust speed quickly and precisely in the presence of any flexible condition or delay due to any hard condition after comparing with root chart. The fuzzy control system for proposed speed scheduling system consists of fuzzifier, inference kernel with knowledge base including database, rule base and output membership functions, and defuzzifier block as shown in Fig.3. The crisp values of input variables VT, TC and EM are reached to fuzzifier after passing through sensors to identify the types of these input variables [7], [9], [10]. In fuzzifier, comparison of input crisp values up to certain levels is done by generating linguistic values (Low, Medium, Average, High and Very High) against each input variable. These linguistic values are passed to inference kernel connected with knowledge base which further categorize into database, rule base and output membership functions.

In knowledge base, key feature of database is to manipulate fuzzy data and provide essential definitions to describe the linguistic control rules which help the rule base to define the control goals and control policy of particular system such as speed scheduling of railway vehicle in this scenario while output membership functions define the strength of output variables with formulation to adjust speed. After receiving feedback from knowledge base, the next step of inference kernel is to simulate the human decision with fuzzy logic rules to make the control decision in term of adjusted speed, the final outcome. In the next step, defuzzifier maps fuzzy output variable (Slow, Average, Fast and Very Fast) to a crisp value which finally comes to railway track system after passing through actuator.
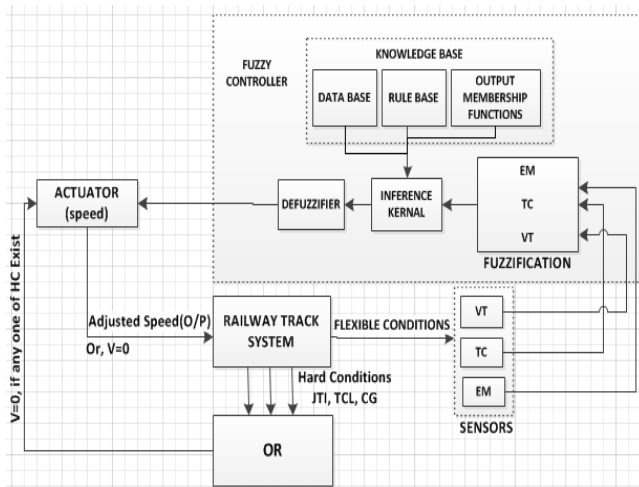

Fig.3. Design Model of FIS Based Speed Scheduling System.

4 DESIGN ALGORITHM OF PROPOSED MODEL

The proposed model is designed for three input variables (VT, EM and TC) which are derived from flexible conditions. The membership functions for these three input variables are shown in Table1.

TABLE 1.
INPUT MEMBERSHIP FUNCTIONS

Scale	VT. Membership Function-MF	EM. Membership Function-MF	TC. Membership Function-MF
0-30	Low	Pleasant	Good
0-60	Medium	Below Pleasant	Below Good
30-90	Above Medium	Average	Average
60-100	High	Below Storm	Bad
90-100	Very High	Strom	Worst

The five membership functions mf1 [1], mf1 [2], mf1 [3], mf1 [4] and mf1 [5] are used to represent the different ranges of input fuzzy variable "Vehicle Tilting" as shown

in Fig. 4. The plot of VT consists of four regions. The other two variables, EM and TC are design on the same pattern for simplification and better understanding.
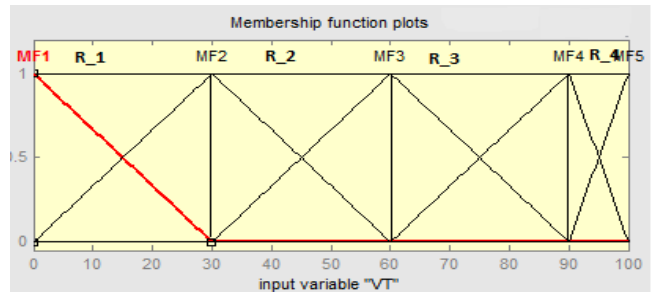

Fig. 4. Plot of Membership Functions for Input Fuzzy Variable- Vehicle Tilting.

The output variable speed consists of seven membership functions. The detail about each membership functions with scale and singleton values are shown in Table 2. The plot of seven membership functions of speed with maximum speed limit 120km/h is shown in Fig. 5.

TABLE 2
OUTPUT MEMBERSHIP FUNCTIONS

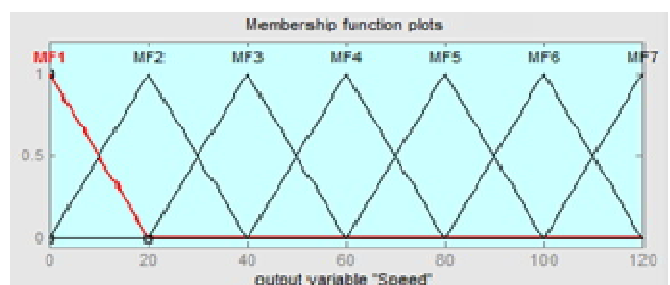| Membership Function	Range	Speed	Singleton Values
MF1	0-20	Very Slow	S1=0
MF2	0-40	Slow	S2=0.16
MF3	20-60	Average	S3=0.33
MF4	40-80	Above Average	S4=0.5
MF5	60-100	Normal	S5=0.66
MF6	80-120	Fast	S6=0.83
MF7	100-120	Very Fast	S7=1

Fig. 5. Plot of Membership Function for output Fuzzy Variable- Speed.

4.1 Fuzzification

There are three fuzzy input variables in proposed design model of speed scheduling and each variable is divided into four regions while f1 and f2 are linguistic values of fuzzy variable "Vehicle Tilting", f3 and f4 for "Environment Monitoring" and f5 and f7 for "Track Condition".

The linguistic values described as the mapping values of these three fuzzy input variables: VT, EM and TC with their membership functions categorize into four regions as shown in Table 3.

TABLE 3
LINGUISTIC VALUES OF FUZZIFIERS OUTPUTS IN ALL REGIONS

Input Variables	Linguistic Fuzzifier Outputs	R_1	R_2	R_3	R_4
VT	f1	mf1 [1]	mf1 [2]	mf1 [3]	mf1 [4]
	f2	mf1[2]	mf1 [3]	mf1 [4]	mf1 [5]
EM	f3	mf2 [1]	mf2 [2]	mf2 [3]	mf2 [4]
	f4	mf2[2]	mf2 [3]	mf2 [4]	mf2 [5]
TC	f5	mf3 [1]	mf3 [2]	mf3 [3]	mf3 [4]
	f6	mf3[2]	mf3 [3]	mf3 [4]	mf3 [5]

Input variables, VT, EM and TC are inversely proportional to speed while each variable has independent effect on speed, even with minor change in any one variable. The number of rules for complete simulation of design model is 64 and rule base will maintain the record of these rules. In this case, 8 rules are used for the values of specific region of input variables like VT=10, EM=15 and TC=25 are taken for region 1 with membership functions and corresponding mapping values mf1 [1], mf1 [2], mf2 [1], mf2 [2], mf3 [1], mf3 [2]. The Fig. 6 is representing the fuzzification process of these input variables crisp values to linguistic variables.
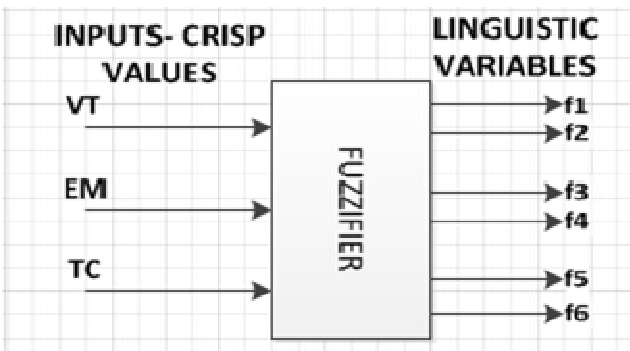

Fig. 6 Fuzzifier Representing, 3- Inputs- Crisp Values with 6- Outputs- Linguistic Variables.

4.2 Fuzzy Inference Engine

The inference engine consists of eight AND operators which select minimum value from linguistic values of three input variables. The working technique of these AND operators are different form logical ANDs. This inference engine accepts six inputs from fuzzifier and ap-

plies min-max composition to obtain the value to adjust speed. The min-AND operation is used from min-max technique to get minimum value from input variables TC, EM and TC as shown in Fig. 7 which gives the overview of fuzzy inference engine working strategy.
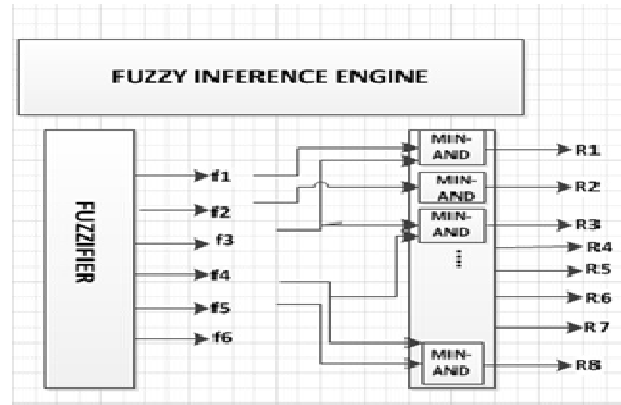

Fig. 7. Block Diagram of Fuzzy Inference Engine.

Rule1 = f1 ^ f3 ^ f5 = mf1 [1] ^ mf2 [1] ^ mf3 [1]
  = 0.33 ^ 0.5 ^ 0.83 = 0.3
Rule2 = f1 ^ f4 ^ f6 = mf1 [1] ^ mf2 [2] ^ mf3 [2]
  = 0.33 ^ 0.5 ^ 0.16 = 0.16
Rule3 = f1 ^ f3 ^ f6 = mf1 [1] ^ mf2 [1] ^ mf3 [2]
  = 0.33 ^ 0.5 ^ 0.83 = 0.33
Rule4 = f1 ^ f4 ^ f5 = mf1 [1] ^ mf2 [2] ^ mf3 [5]
  = 0.66 ^ 0.5 ^ 0.83 = 0.5
Rule5 = f2 ^ f3 ^ f5 = mf1 [2] ^ mf2 [1] ^ mf3 [1]
  = 0.66 ^ 0.5 ^ 0.83 = 0.5
Rule6 = f2 ^ f4 ^ f6 = mf1 [2] ^ mf2 [2] ^ mf3 [2]
  = 0.66 ^ 0.5 ^ 0.16 = 0.16
Rule7 = f2 ^ f3 ^ f6 = mf1 [2] ^ mf2 [1] ^ mf3 [2]
  = 0.66 ^ o.5 ^ 0.16 = 0.16
Rule8 = f2 ^ f4 ^ f5 = mf1 [2] ^ mf2 [2] ^ mf3 [1]
  = 0.66 ^ o.5 ^ 83 = 0.5

The ^ operator between membership function values is used for min-AND process to get minimum value of the function.

4.3 Rule Selector

The rule selector of proposed model receives three crisp values of VT, EM and TC, and provides singleton values of output function with specific rules of the design model requirement. In this case, 8 rules are required to find the required values S8, S9…, S15 according to division of regions for soft conditions while hard conditions are absent.

The rules are listed in Table 4 with existence of both hard and flexible conditions for quick overview of rules for speed control system.

4.3 Defuzzifier

In defuzzification process, crisp values for final estimated speed are obtained after estimating its inputs regarding

TC, EM and VT from the rule base.

TABLE 4
ILLUSTRATION OF APPLIED RULES

INPUT						OUTPUT	
FLEXIBLE CONDITIONS			HARD CONDITIONS			SPEED In km/h	Singleton Values
VT	EM	TC	JTI	TCL	CG		
1	1	1	N	N	Y	Stop	S1
1	1	2	N	Y	N	Stop	S2
1	2	1	Y	N	N	Stop	S3
1	2	2	N	Y	Y	Stop	S4
2	1	1	Y	N	Y	Stop	S5
2	1	2	Y	Y	N	Stop	S6
2	2	2	Y	Y	Y	Stop	S7
1	1	1	N	N	N	100-120	S8
1	1	2	N	N	N	80-120	S9
1	2	1	N	N	N	80-120	S10
1	2	2	N	N	N	80-120	S11
2	1	1	N	N	N	80-120	S12
2	1	2	N	N	N	80-120	S13
2	2	1	N	N	N	80-120	S14
2	2	2	N	N	N	60-100	S15

There are 16 inputs are given to the defuzzifier, eight values from eight rules (Rule1-Rule8) and eight values from rule selector (S8-S15). The center of average method (C.O.A) is used by each defuzzifier to estimates the crisp value with mathematical expression $\sum S_i * R_i / \sum R_i$, where $i = 1$ to 8 in the given scenario as shown in Fig. 8.
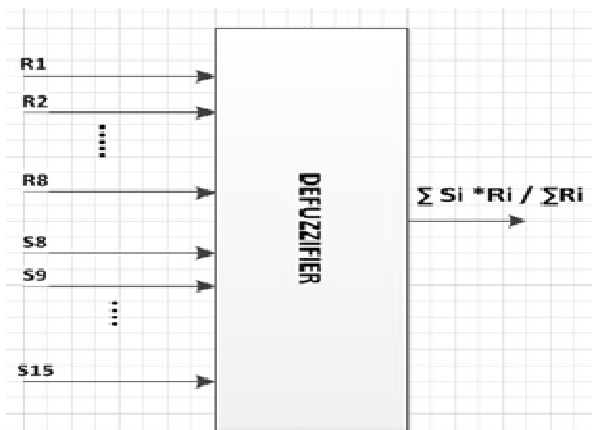

Fig. 8. Block Diagram of Defuzzifier.

The mathematical expression $\sum S_i * R_i / \sum R_i$ has used for crisp values of output variables which are according to MATLAB simulation results as shown in Fig. 9.

The values of input variables, VT, TC and EM are taken as the same for MATLAB simulation as for mathematical calculation that shows correct results with quick response.
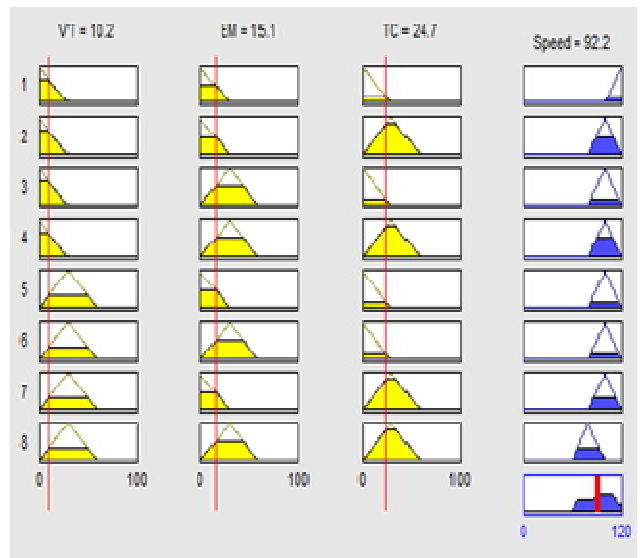

Fig.9. MATLAB- Rule Viewer and Simulation Results for Speed Scheduling System of Railway Vehicles.

5 RESULTS AND DISCUSSION

The design model of proposed speed scheduling system of autonomous railway vehicle has shown significant improvement in performance regarding safety and time to meet uncertain conditions with minimum delay using fuzzy inference system with MATLAB simulation as compared to traditional approaches. Soft conditions like VT, EM and TC are inversely proportional with respect to speed and have shown substantial effect on it in case of any change in input variables individually and combined as well. The effect of these inputs on speed has shown in Fig.10 which is according to rule base of design algorithm.

Fig. 10 (a) has shown that at 0 of both VT and EM, the speed is fast, more than 100 km/h which gradually decrease with increase in values of EM and VT.

Fig. 10 (b) has shown the same effect with input variables VT and TC which is according to rule base by assigning specific values to input variables to specific region and prove that these three variables have same effect on speed.
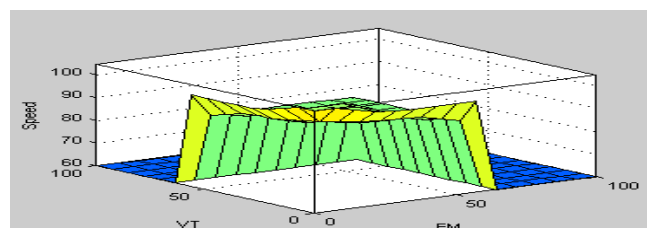

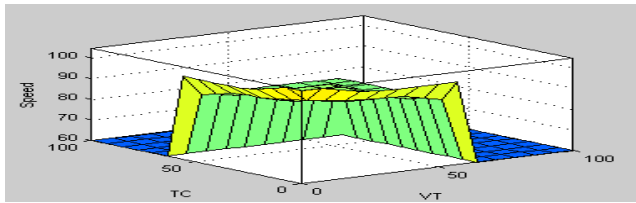Fig. 10 (a) Plot between Environment Monitoring and Vehicle Tilting.

Fig. 10 (b). Plot between Vehicle Tilting and Track Condition.

6 CONCLUSION AND FUTURE WORK

FIS based speed scheduling system of autonomous railway vehicles has shown remarkable improvement to compete the demands of new trends with safety and minimum time delay. Now day's railway control system requires some sophisticated method to handle real time problems without compromising on schedule time and security. MATLAB results have shown that proposed model of railway speed scheduling system will be capable to handle uncertain conditions successfully with better and quicker response as compared to existence methods with time management and safety by reducing risk. In future, fuzzy inference system base railway control system will be more secure with performance enhancement in real time environment. State of the art Microelectronics technology can be used to develop FPGAs based control chips for this autonomous railway control system.

ACKNOWLEDGMENT

REFERENCES

[1] M. Saleem Khan and, Khalid Benkrid, "Design Model of Multi Agents Based Autonomous Railway Vehicles Control System", in proc. Of int. Multiconference of Engineering and computer Scientists, Hong Kong, March 18-20, 2009, vol.2.

[2] R. Harrison, M. Lee, "Reconfigurable Modular Automation Systems for Automotive Power-train Manufacturer", int. Flex Manuf Syst, Springer Science 1 March 2007.

[3] M. Saleem Khan and, Khalid Benkrid, "Development Techniques of Multi-Agents Based Autonomous Railway Vehicle Control System", World Academy of Science Engineering and Technology, pp. 558-568, 2010.

[4] M. Malleswaran, J. Mary Anita, S. N. Sabreen and V. Vaidehi "Real-time INS/GPS Data Fusion Using Hybrid Adaptive Network Based Fuzzy Inference", 11th Int. Conf. Control, Automation, Robotics and Vision Singapore, pp. 1536-1540, December 2010.

[5] Lofti A. Zadeh, "Fuzzy Logic", pp. 83-92, IEEE, April 1988.

[6] S.H. Lee, R.J. Howlett and S.D. Walters, "Fuzzy and Neuro-fuzzy Techniques for Modeling and Control", pp.1206-1215, Springer-Verlag Berlin Heidelberg 2006.

[7] Tang Xianzhi and Wang Qingnian, "Driving Intention Intelligent Identification Method for Hybrid Vehicles Based on Fuzzy Logic Inference", Int. Symposium on Computational Intelligence and Design, pp. 16-19, IEEE, 2010.

[8] Gary R. George P. E. and Frank Cardullo, "Application of Neuro-Fuzzy Systems to Behavioral Representation in Computer Generated Forces".

[9] Morelli, Jordan E., Akira Hirose and Hugh C. Wood "Fuzzy-Logi-Based Plasma Position Controller for STOR-M" IEEE Transaction on Control Systems Technology, Vol. 13, No 2, March 2005.

[10] M. Saleem Khan, "Fuzzy Time Control Modeling of Discrete Event Systems", Proc. Of World Congress on Engineering and Computer Science 2008.

Optimizing Performance of Token Ring for Balanced and Unbalanced Load Using OPNET

Pankaj Rakheja, Dilpreet Kaur

Abstract— Token Ring is a local area network which resides in the Data Link Layer (DLL) of the OSI model. Stations on a token ring LAN are logically organized in a ring topology with data being transmitted sequentially from one ring station to the next with a control token circulating around the ring controlling access. Each station passes or repeats the special token frame around the ring to its nearest downstream neighbour. This token-passing process is used to arbitrate access to the shared ring media. Stations that have data frames to transmit must first acquire the token before they can transmit them. It can transmit the packet for fixed duration decided by token holding time. Here we have analyzed its performance for a given number of nodes for balanced and unbalanced traffic flow and varied the size of packet to find out the optimum packet size for both balanced and unbalanced traffic flow. We have used OPNET for simulation.

Index Terms—CSMA, Hub, Mac, Token Ring, topology, OPNET.

———————————————— ◆ ————————————————

1 INTRODUCTION

Token Ring [1-6] is a Local Area Network (LAN) protocol defined in the IEEE 802.5 section. Here all stations are connected in a ring and each station can just directly hear the transmissions only from its immediate neighbour. Permission to transmit is granted by a message or packet called token that circulates around the ring.

Token Ring as defined in IEEE 802.5 has originated from the IBM Token Ring LAN technologies. Both are based on the token passing technologies. They differ in minor ways; so are generally compatible with each other.

Token-passing networks move a small frame, called a token, around the network. Possession of the token grants the right to transmit. If a node receiving the token has no information to send, it seizes the token, alters 1 bit of the token (which turns the token into a start-of-frame sequence), appends the information that it wants to transmit, and sends this information to the next station on the ring. While the information frame is circling on the ring, no token is on the network, which means that the other stations wanting to transmit must wait. Therefore, collisions cannot occur on token ring networks and prioritization of nodes is done to avoid conflicts and impermissible delays for real time applications.

————————————————

- *Dilpreet Kaur, is currently pursuing masters degree program in electronics and communication engineering at Institute of Technology and Management, Gurgaon, Haryana, India E-mail: er.dilpreetkaur@gmail.com*
- *Pankaj Rakheja, is currently pursuing masters degree program in electronics and communication engineering at Institute of Technology and Management, Gurgaon, Haryana, India E-mail: rakheja.prince13@gmail.com*

The information frame circulates on the ring until it reaches the intended destination station, which copies the information for further processing. The information frame continues to circle the ring and is finally removed when it reaches the sending station. The sending station can check the returning frame to see whether the frame was seen and subsequently copied by the destination.

There is a monitor or head node which supervises the network. It captures the failed or erroneous packets circulating in the ring to prevent unavailability of the network to the stations in case a packet rotates there indefinitely when destination node fails.

Unlike Ethernet CSMA/CD networks, token-passing networks are deterministic, which means that it is possible to calculate the maximum time that will pass before any end station will be capable of transmitting. This feature and several reliability features make Token Ring networks ideal for applications in which delay must be predictable and robust network operation is important.

2 OVERVIEW

A Token Ring network is an implementation of IEEE standard 802.5. It's not their physical cable layout that distinguishes it from other networks but rather its token access method. The architecture of a typical Token Ring network begins with a physical ring. However, in its IBM implementation we have a star-wired ring arrangement of computers in the network connected across a central hub. Figure 1 be-

-low shows a logical ring and a physical star topology. The logical ring [2] [3] represents here the path of the token between computers or various nodes connected. The actual physical ring of cable is in the hub. Users are part of a ring, but are connected through a hub.
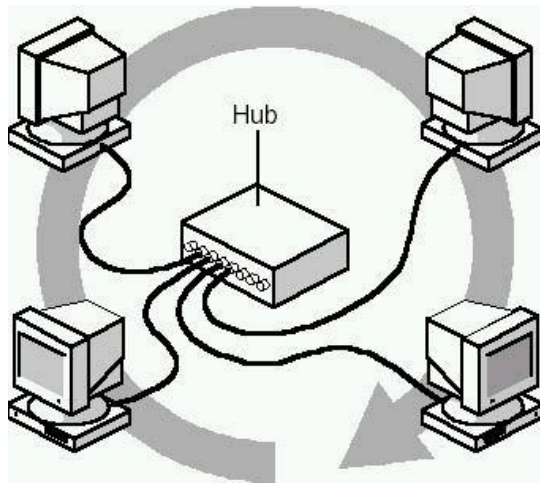

Figure 1: Logical and physical ring in a token ring

A Token Ring network includes the following features: star-wired ring topology, token-passing access method , shielded and unshielded twisted-pair (IBM Types 1, 2, and 3) cabling transfer rates of 4 and 16 Mbps, baseband transmission ,802.5 specifications.

When the first Token Ring computer comes online, the network generates a token. The token is a predetermined formation of bits that enables the computer to put data on the cables. The token travels around the whole ring and uses polling at each computer until one of the computer signals that it wants to transmit data and takes control of the token. A computer cannot transmit unless it possesses token and during that the other computers wait till the sending computer releases the token.

After the computer captures the token, it sends a data out on the network. The frame proceeds around the ring until it reaches the destination host or computer addressed in the frame. It copies the frame into receive buffer and marks the frame in the frame status field to indicate that the information was received.

The frame continues around the ring until it arrives at the sending computer, where the transmission is acknowledged as successful. The sending computer then removes the frame from the ring and transmits a new token back on the ring.

The first computer to come online is assigned by the Token Ring system as monitor. The monitoring computer makes it sure that frames are being delivered and received correctly. It does this by checking for frames that have circu-

lated the ring more than once and ensuring that only one token is on the network at a time.

This process of monitoring the network is known as beaconing. It also sends a beacon frame periodically and if it does not get this beacon that is a form of announcement marking presence of the monitor, then it sends a message around the ring to check if monitor failed and if it failed then reconfiguration of the network take place.

The traffic or data can be both balanced and unbalanced. In the former one, the inter arrival time for each node is the same that generally happens in theoretical conditions but in real life scenarios we have to support different services which have different needs. Some permit certain delay while for others, delay is not permissible. For example, in voice conversations or chat sessions delays are not permissible.

3 WORKDONE

We have designed a network with 32 nodes and Token holding time of 0.01 seconds with both balanced and unbalanced traffic. The scenario designed is as shown below in figure 2.
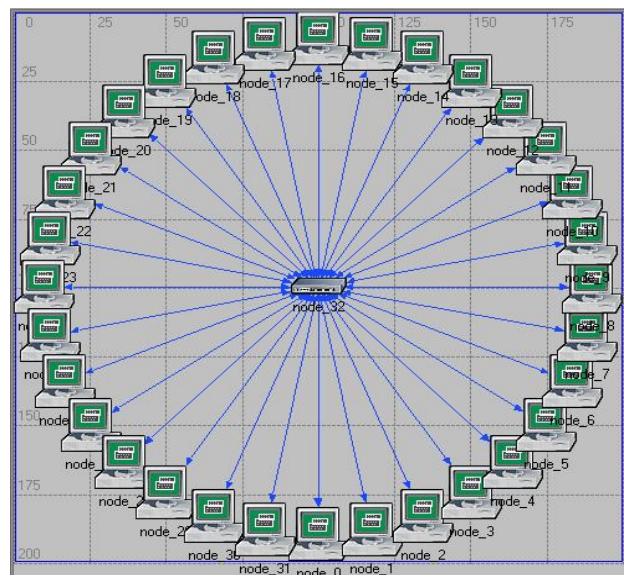
Figure 2: Scenario of Token ring

Here we have a Tr_32_hub_adv in the centre with 32 tr_stations connected through hub with duplex 10 links. This whole network lies in a campus or office. It supports all the requisite or default applications. We have fixed Token Holding Time (THT) to be 0.01 seconds. We have analysed the delay and MAC access time for various packet sizes for both balanced and unbalanced traffic to get the optimum or most effective packet size and to compare its results for both balanced and unbalanced traffic.

The parameters dealt at each work station are shown in Figure 3 below.

Figure 3: Workstation Parameters

We have used switched mode operational mode along with THT of 0.01 seconds. For balanced node we have the same inter arrival time for all nodes whereas it's different for different nodes in case of unbalanced load or traffic.

4 RESULTS

We have plotted graphs for delay encountered by packet from source to destination and MAC access delay for balanced and unbalanced traffic or load for different packet sizes. We have taken four values of packet sizes i.e.256, 512, 1024 and 2048 bytes.

Figure 4 and 6 are for balanced traffic or load. The former is for delay encountered by packets from node to destination and the later one is for media access delay for nodes. Figure 5 and 7 are for unbalanced traffic or load. The former is for delay encountered by packets from node to destination and the later one is for media access delay for nodes. Figure 8 and 9 compare the total delay and delay associated with media access for balanced and unbalanced load at different packet sizes.
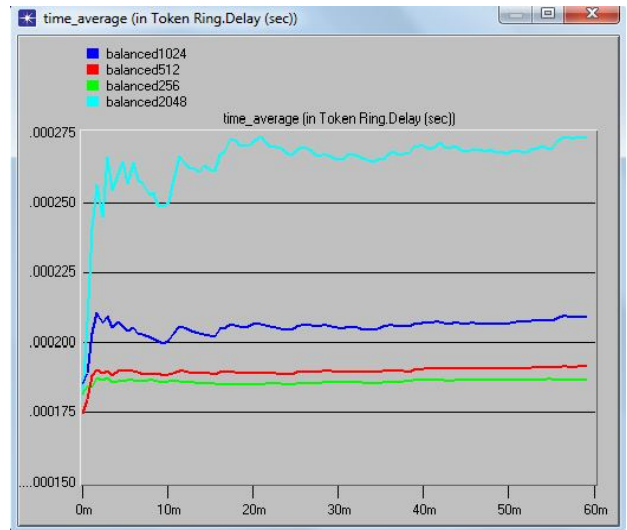

Figure 4: Delay for different packet sizes for balanced load
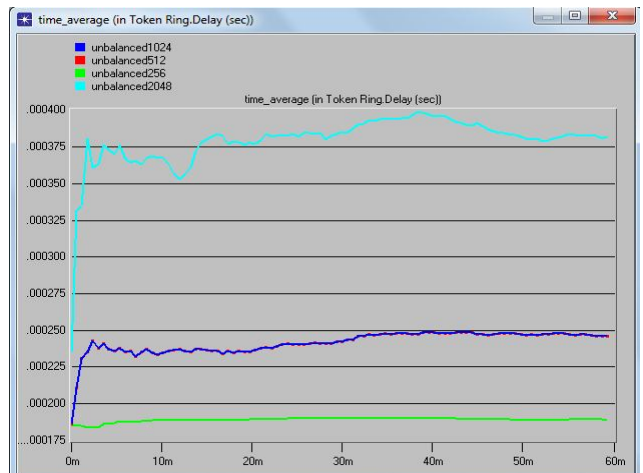
Figure 5: Delay for different packet sizes for unbalanced load
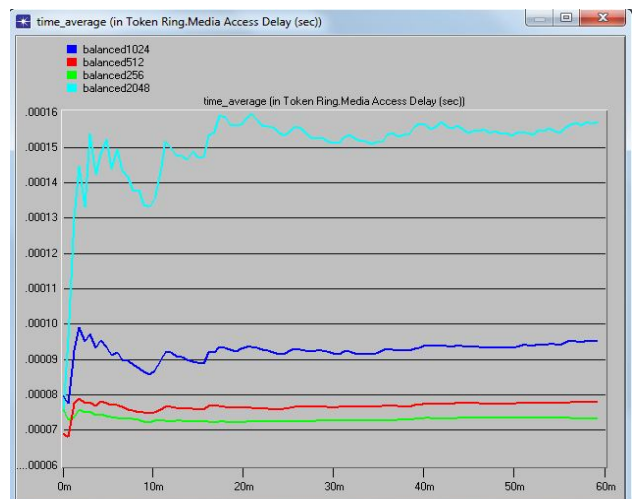
Figure 6: Media access delay for different packet sizes for balanced load
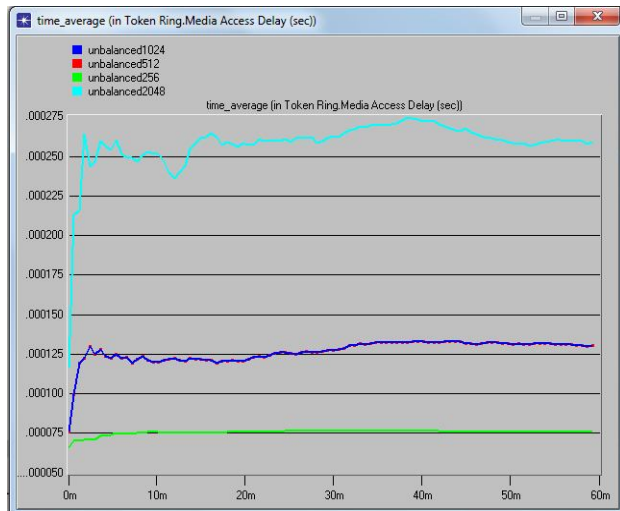
Figure 7: Media access delay for different packet sizes for unbalanced load
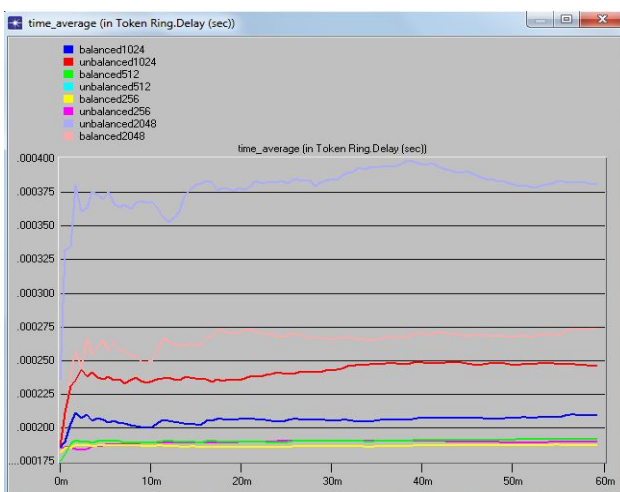
Figure 8: Delay for different packet sizes for balanced and unbalanced load
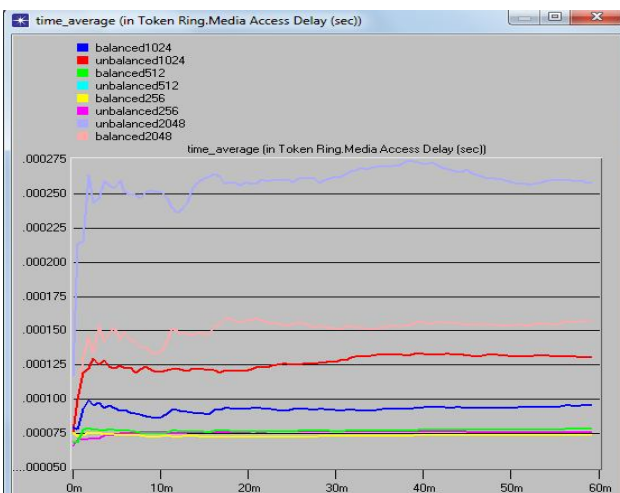
Figure 9: Delay for different packet sizes for balanced and unbalanced load

5 CONCLUSION

Token Ring is Local Area Network technology based on token access method where a node can send data only if it has token and transmission is just for time interval equal to Token Holding Time. Here we can have both balanced and unbalanced loads which enable it to support most of the real time applications but to ensure that the delays associated should be low. That in turn depends on packet size. For a given number of nodes packet size needs to be optimized as delay and throughput depends on it too. Here on analysing the result we can say that 256 bytes of packet size is optimum for balanced traffic or load and for unbalanced one we can choose 256 or 512 as delays are same for both. On comparing total delay encountered by packets and media access delay between balanced and unbalanced load, the former one has lower values of delays associated than the later one.

6 REFERENCES

1. Werner Bux, Felix H. Closs, C. Kuemmerle, H. J.Keller, and H. R.Mueller,"Architecture and Design of a Reliable Token-Ring Network," lEEE Journal on Selected Areas in communications, Vol. SAC-I, No. 5, pp. 756-765, November 1983.
2. ANSVIEEE Std.802.2-1985,"Logical Link Control,"1985.
3. ANSVIEEE Std.802.5-1985,"Token Ring Access Method and Physical Layer Specifications," 1985.
4. Bux, W.," Token-ring local-area networks and their performance" , Proceedings of the IEEE, Feb 1989.
5. Smythe, C. ," ISO 8802/5 token ring local-area networks", Electronics & Communication Engineering Journal 1999.
6. Nasro Min-Allah, Manzoor Elahi, Xing jiansheng, Wang Yongji," Enhancing Feasibility Analysis of IEEE 802.5 Token Ring", IEEE 2008.

Design of Viterbi Decoder for Noisy Channel on FPGA

Ms. M.B.Mulik, Prof. U.L.Bombale, Prof. P.C.Bhaskar

Abstract— The problem of survival memory management of a Viterbi Decoder (VD) was solved by introducing a pointer implementation for the register exchange (RE) method, where a pointer is assigned to each row of memory in the SMU. The content of the pointer which points to one row of memory is altered to point to another row of memory, instead of copying the contents of the first row to the second. In this paper, the one-pointer VD is proposed; if the initial state of the convolutional encoder is known, the entire survivor memory unit (SMU) is reduced to only one row. Because the decoded data are generated in the required order, even this row of memory is dispensable. Thus, the one-pointer architecture, referred to as memory less Viterbi Decoder (MLVD), reduces the power consumption of a traditional trace back (TB) VD by approximately 50 percent. A prototype of the MLVD with a one third convolutional code rate and a constraint length of nine is mapped into a Xilinx.

Index Terms—wireless, low power, Viterbi Decoder, register exchange, memory less.

————————————— ◆ —————————————

1. INTRODUCTION

THE Viterbi Algorithm is an efficient method for the realization of maximum likelihood decoding of convolutional codes [2],[3]. The digital VD is widely used in many digital wire line and wireless applications. Existing VDs are categorized by the way how the decoded data bits are stored in and retrieved from the SMU. Two methods are mainly used, the RE and the TB method [4]. In the literature, the RE method is acceptable for trellises with only a small number of states, whereas the TB approach is acceptable for trellises with a large number of states. Therefore, the TB method has been widely investigated and implemented. Although several attempts to reduce the power consumption of the TB VD have been proposed, only a few attempts to combine the advantages of the TB and RE methods have been reported.

The modified RE method [1], which is further improved in this work, utilizes the pointer concept, that is widely used in software engineering. Instead of moving the contents of one row of memory to a second row of memory, the pointer to the first row is altered to point to

• *Ms M.B.Mulik is currently pursuing master degree program in Electronics Technology in Shivaji University, Kolhapur, India,PH-09921660850, Email-mulik_madhuri2002@yahoo.co.in, Sponsored by Shard Institute of Technology, Polytechnic, Yadrav.*
• *Dr.U.L.Bombale is currently working as a Professor in Electronics Technology in Shivaji University, Kolhapur, India, PH-09049274380, Email-uttam_bombale@rediffmail.com.*
• *Prof.P.C.Bhaskar is currently working as a M.Tech Co-Coordinator in Electronics Technology in Shivaji University, Kolhapur, India, PH-09881248133, Email- pxbhaskar@yahoo.co.in*

the second row. The pointer to one row of memory simply carries the current state in the trellis of the VD. The pointer implementation avoids the power hungry register exchange operations of the traditional RE method, and is referred to as pointer Viterbi Decoder (PVD). In the next section the PVD is briefly reviewed, and then its memory less version is introduced.

2. THE MEMORYLESS VITERBI DECODER (MLVD)

The PVD keeps track of the current row position of the decoder in the memory. It makes use of the fact that the bit appended to each row of memory is exactly the bit that is shifted into the pointer to form the new pointer to that row of memory. To show the functionality a 4-state rate=1/3 convolutional encoder (G0=101, G1=111, and G2=111) is employed to encode the input sequence of (10110100100). The code stream, (111, 011, 000, 100, 100, 000, 011, 111, 111, 011, 111) is generated and transmitted over a noisy channel. The noisy code stream, (111, 011, 001, 100, 100, 000, 011, 111, 110, 011, 111), for example, is received at the decoder. The underlined bits are incorrect because of the noise encountered during transmission. Applying the modified RE method results in the diagram illustrated in Figure. 1. This figure displays the successive values for the pointers and rows of memory over time.

A closer look reveals that each row of memory is used to trace the decoded bits, if an initial state is assumed. The first row of memory decodes the data, if an initial state, S_0, is assumed. The last row records the decoded data, if an initial state, S_{255}, is assumed, and so on. If the initial state is known, are all these rows of memory necessary? Absolutely not. For example, if the initial state is zero, then only the first row of memory is

needed. In other words, the storage of the decoded bits is necessary in order to choose only one row of memory at the end to represent the actual decoded bits. If the required row of memory is predetermined, then there is

no need for the storage of the other rows. Furthermore, there is no need for the storage of the row that is assigned to the predetermined
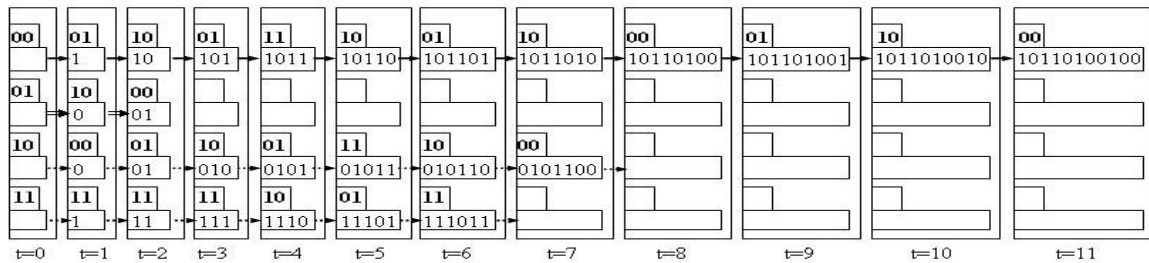
Fig. 1. New RE approach with pointer implementation (the upper register carries the pointer and the lower register carries the decoded bits)

then there is no need for the storage of the other rows.

Furthermore, there is no need for the storage of the row that is assigned to the predetermined initial state, because the RE approach generates the decoded bits in the correct order. The decoded bits are produced, and then read out from the decoder.

Thus, a memory free Viterbi decoder can be implemented by solely resetting the encoder contents for each L bits that are encoded. There is no need to interrupt the data sequence or to transmit a long sequence of zero data bits. The encoded data is continuous, but the contents of the encoder bits (eight bits for K = 9 convolutional encoder) are reset to zero for each L bits transmitted. The only overhead for such an implementation in a communication system is to synchronize between the transmitter and the receiver. The new VD implementation is called, the Memory Less Viterbi Decoder (MLVD). Since the MLVD needs to track only one row, the MLVD requires only one pointer to track the current position of the decoder in the trellis in Figure. 2.
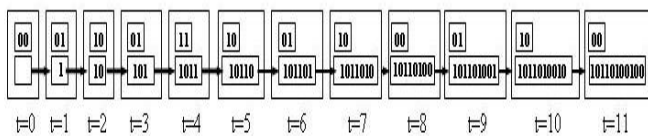
Fig. 2. MLVD approach with pointer implementation (the upper box carries the pointer and the lower box carries the decoded bits)

The MLVD is an extra low power design or a VD with the only restriction of resetting the encoder register at each L of the encoded data bits and providing the necessary synchronization. The block diagram of the MLVD, designed in VHDL, is shown in Figure .3.
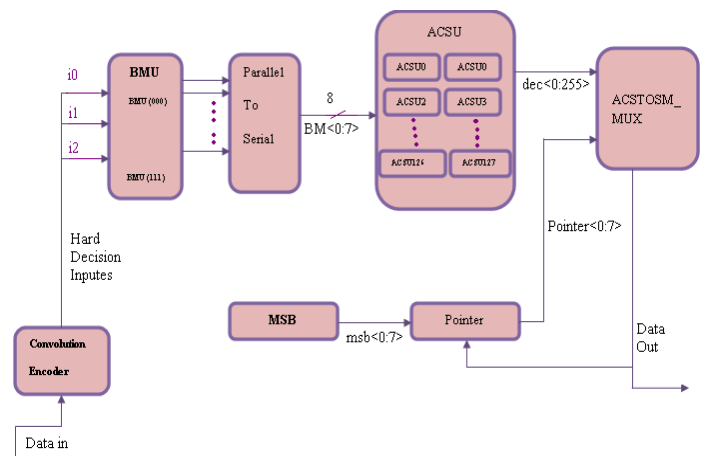
Fig. 3. MLVD block diagram

2.1 Convolutional Encoder

The convolutional encoder that is implemented is shown in Figure 4 It is a K=9 and r=1/3 convolutional encoder. Figure. 4. Shows the block diagram of the convolutional encoder. Three outputs are generated for each bit encoded. To implement a hard-decision VD, The output of the encoder is fed directly into the first block of the VD, the BMU.

2.2 Branch Metric Unit (BMU)

In the VA for decoding convolutional codes, however, for binary convolutional codes, it is proven that linear distances (Hamming distances) can be used as the optimum branch metrics. For three 3-bit hard decision input bits, (i0, i1, i2), each branch metrics are generated.

The BMU performs simple add and subtract
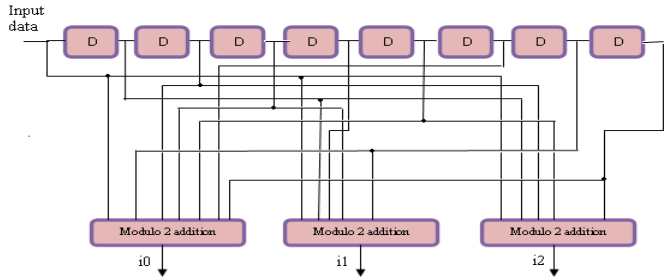
Fig. 4. Convolution encoder

operations on the decision bits to generate the output. The BMU performs the computations, as represented in Figure. 5. The bit serial format of the branch metrics is generated by the parallel to serial module at the output of the BMU the BMs is then fed into the ACSU. As shown in Figure. 5. The bit serial format of the BMs is then fed into the ACSU.
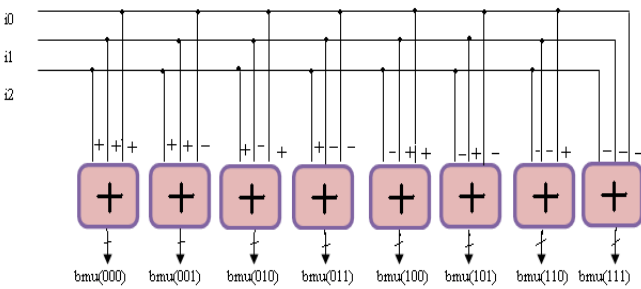
Fig. 5. Branch Metric Unit (BMU) operations

2.3 Add Compare Select Unit (ACSU)

The ACSU is composed of 128 units; each is composed of an ACS butterfly module, as shown in Figure 6.
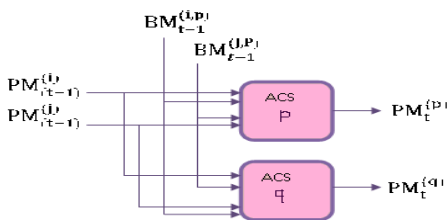
Fig. 6. One ACS butterfly module

2.3.1. Bit Serial ACS Architecture

The block diagram for one ACSU is displayed in Figure. 7. The bit serial approach significantly reduces the interconnection load, but does not allow for the reformulation of the ACS functions. Thus, the traditional addition and subtraction operations are performed in the bit serial ACSU. However, the bit serial approach

necessitates two 8-bit registers for each PM to temporarily store the two calculated PM values which are compared in the ACSU so that the smaller one can be chosen as the new PM. In addition, extra registers are required to store the carry bits that are generate during the addition and subtraction operations in the ACSU. A very simple shifter is introduced which requires only one control signal and is composed of eight master latches.
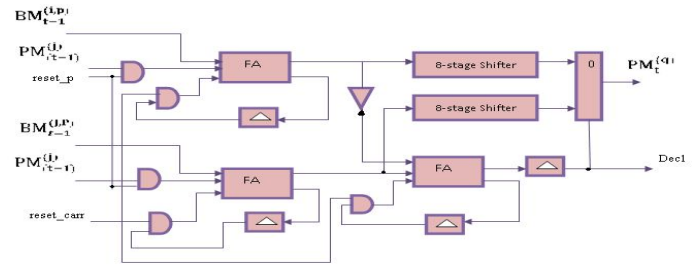
Fig. 7. Bit Serial ACS Architecture

2.3.2 Butterfly structure of the ACS
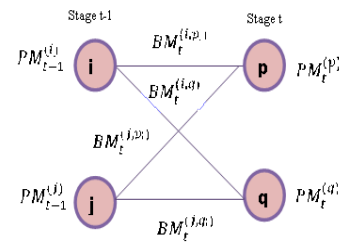
Fig. 8. Butterfly structure of the ACS

For example, if both paths from state j are considered to be survivor paths for the destination states p and q, the BMs of both paths are compared. Then the pointer, which carries the value j, changes to the destination pointer, whose path has the smaller BM. The pointer of i changes to the other destination pointer, while the path from state i receives a termination high signal. This indicates that the path from state i is terminated, and the decision bits are no longer appended to the path's register. Consequently, only the pointer value of the terminated path will continue to be modified in order to prevent the presence of duplicated pointer values.

The ACSU is composed of 126 units; each is composed of an ACS butterfly module which adds the BMs to the corresponding path metrics (PMs), compares the new PMs, and then feeds the selected PMs Back into the ACSU. The typical operation of one ACS butterfly module is as follows:

If

$$BM^{(i,p)} + PM^{(i,p)} < BM^{(j,p)} + PM^{(j,p)}$$

Then $Dec^{(p)} = 0$

$$PM^{(p)} = PM^{(i)} + BM^{(i,p)}$$

If

$$BM^{(i,p)} + PM^{(i,p)} > BM^{(j,p)} + PM^{(j,p)}$$

Then $Dec^{(p)} = 1$

$$PM^{(p)} = PM^{(j)} + BM^{(j,p)}$$

If $\quad BM^{(i,q)} + PM^{(i,q)} < BM^{(j,q)} + PM^{(j,q)}$
Then $Dec^{(q)} = 0$

$$PM^{(q)} = PM^{(i)} + BM^{(i,q)}$$

If $\quad BM^{(i)} + PM^{(i,q)} < BM^{(j)} + PM^{(j,q)}$
Then $Dec^{(q)} = 1$

$$PM^{(q)} = PM^{(j)} + BM^{(j,q)}$$

And according to the symmetric characteristics of the VD

$$BM^{(i,q)} = BM^{(j,p)}$$

and $\quad BM^{(i,p)} = BM^{(j,q)}$

Therefore, only two branch metrics *BMs* are connected to each butterfly unit as shown in Figure 4.8.The bit serial approach for the MLVD implementation to reduce the routing overhead among the different ACSU modules. The routing among the different ACSU modules represents an overhead, which is left for the automatic optimization of the FPGA router.

2.3.3 Branch Metric

For Hard-decision metric, $\mu_t\left(r_t, x^{(p,q)}\right)$ can be computed in advance. For example, for an n = 3 binary code, there are eight possible received values, 000, 001, 010, 011, 100, 101, 110, 111, & eight possible transmitted values. The metric is an 8 × 8 array, such as following Table 1 shows,

Table 1

Branch Metric

$\mu_t\left(r_t, x^{(p,q)}\right)$	000	001	010	011	100	101	110	111
$x^{(p,q)} = 000$	0	1	1	2	1	2	2	3
$x^{(p,q)} = 001$	1	0	2	1	2	1	3	2
$x^{(p,q)} = 010$	1	2	0	1	2	3	2	2
$x^{(p,q)} = 011$	2	1	1	0	3	2	2	1
$x^{(p,q)} = 100$	1	2	2	3	0	1	2	2
$x^{(p,q)} = 101$	2	1	3	2	1	0	2	1
$x^{(p,q)} = 110$	2	3	1	2	1	2	0	1
$x^{(p,q)} = 111$	3	2	2	1	2	1	1	0

2.4 Add Compare Select TO Survivor Memory (ACSTOSM)

The ACSTOSM is employed to route the decision of the appropriate ACS module to the output. The ACSTOSM is a 256 to one decoder. The select signal for this large decoder is the output of the pointer module. The output of the ACSTOSM module is already the decoded output sequence of the VD, but is fed back into the pointer module to update the current state in the decoding trellis.

2.5 Pointer

The pointer block contains the current state of the decoder (eight bits). For each bit decoded, the pointer content is updated. The exact position of the bit that will be updated is determined by the MSB block.

2.6 Most Significant Bit (MSB)

The MSB block has eight outputs that point to the eight bits of the pointer in parallel. At the reset, the eighth bit of the MSB block is enabled which, in turn, causes the eighth bit of the pointer to be the most significant bit. After accessing an incoming code word in the ACSU, a decision bit is written into the MSB of the pointer, and the output of the MSB block is shifted to the right so that only the seventh bit is enabled. For the pointer, this indicates that the seventh bit is now considered to be the most significant bit. In this case, pointer P should be read in the following order: P6P5P4P3P2P1P0P7, instead of the usual order: P7P6P5P4P3P2P1P0. This process continues with each incoming code word.

3. Comparison of the memory/register operations to decode codeword for various VD methods

Table 2

Comparison of various VD methods

Operation	TB (48)	Modified RE (48)	MLVD (168)
Writing decision bits into the	256 x 48 12288w	6144w	–

memory			
Reading from the memory	48 x 3 *r* 72*w*	48*r*/2 24*w*	–
Writing into the MSB of the pointers	–	256 x 48 *rg* 1536*w*	168*rg* 21*w*
Writing the termination bits into the termination registers	–	256 *rg* 32 *w*	–
8-bit register shifting (shift to the right and append a bit to the LSB)	48 x 3 x 8*rg* 144*w*	–	–
Total	12504*w*	7736*w*	21*w*

Where,

W = writing one bit into the memory block

r = reading one bit from the memory block

rg = writing one bit into a register

4. Simulation result of VD using modelSim

Simulation of MLVD is did using XILINX ISE simulator and corresponding results shown in figure 9. Simulation is done using command line of modelSim. Using commands provided by modelSim it is possible to force signal to desired value for required amount of time. Initially required signals are added into waveform window of modelSim and inputs ate forced using commands.

Figure. 9. Shows the VD output which is corrupted by eb1 error signal, whenever eb1 signal goes high it introduces error. Figure .10. Shows decoder output 'dom' i.e. sequence 011 110 001 000 010 000 that becomes 010 111 000 001 011 001 means each last bit of decoded output goes into error.

But using trellis diagram the branch metric calculated from that minimum path metric selected in viterbi decoder so that the error is corrected as shown in figure 10. In this waveform shows viterbi decoder corrected that corrupted data.

Fig. 9. Simulation result of MLVD using modelSim

Fig. 10. Simulation results of MLVD shows corrupted data is recovered back to original data

5. Conclusion

In this work, error detection and correction techniques have been used which are essential for reliable communication over a noisy channel. The effect of errors occurring during transmission is reduced by adding redundancy to the data prior to transmission.

A MLVD based on FPGA is proposed and implemented in this work .The MLVD is a memory less implementation of the VA, and successfully decodes the continuous data encoded by a convolutional encoder and synchronizing the VD with the resetting procedure. The new implementation is realized by applying the pointer concept to the RE implementation, and by reinforcing the

initial state of the convolutional encoder every L bits encoded. The hardware and computational overhead of the new implementation is only a 256 to 1 decoder, which is switching at the data rate frequency. The new MLVD is a memory less high speed, low latency, and low power variation to the VD

REFERENCE

[1] D. A. El-Dib and M. I. Elmasry, "Modified register-exchange viterbi decoder for low-power wireless communications," *IEEE Transactions on Circuits and Systems I*, vol. 51, no. 2, pp. 371-378, February 2004.

[2] A. Viterbi, "Error bounds for convolutional codes and asymptotically optimum decoding algorithm," *IEEE Transactions on Information theory*, vol. It-13, no. 2, pp. 260–269, April 1967.

[3] G. Forney, "The viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, March 1973.

[4] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall, 1995.

[5] J. H. et al, "Cdma mobile station modem asic," *IEEE Journal of Solid- State Circuits*, vol. 28, no. 3, pp. 253-260 March 1993.

[6] I. Kang and A. W. Jr, "Low-power viterbi decoder for cdma mobile terminals," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 3, pp. 473- 482, March 1998.

[7] Inyup Kang, Member, IEEE, and Alan N. Willson Jr. Fellow IEEE."Low-Power Viterbi Decode for CDMA mobile terminals" Ieee journal of solid-state circuits , vol.33, no. 3, march 1998,

[8] James Tang and Esam Abdel-Raheem "High Speed Viterbi Decoder Design and FPGA Implementation" .Department of Electrical and Computer Engineering , University of Windsor, Ontario, Canada N9B 3P4.

[9] C. B. Shung *et al.*, "Area-efficient architectures for the Viterbi algorithm Part I: Theory," *IEEE Trans. Commun.*, vol. 41, pp. 636–644, Apr. 1993.

[10] S.-S. Wang, "A state-reduction viterbi decoder for convolutional codes with large constraint lengths," Master's thesis, National Chiao Tung University, Hsinchu, Taiwan, June, 2002

[11] H.-L. Lou, "Linear distances as branch metrics for viterbi decoding of trellis codes," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3267-3270, June 2000

[12] R. Henning and C. Chakrabarti, "Low-power approach for decoding convolutional codes with adaptive viterbi algorithm approximations, "*Proc. IEEE International Symposium on Lower Power Electronics and Design*, pp. 68-71, August 2002.

A Multi-Agent Based Autonomous Traffic Lights Control System Using Fuzzy Control

Yousaf Saeed, M. Saleem Khan, Khalil Ahmed, Abdul Salam Mubashar

Abstract— This research work presents an application of fuzzy logic for multi-agent based autonomous traffic lights control system using wireless sensors to overcome problems like congestion, accidents, speed and traffic irregularity. The proposed agent based approach can provide a preferred solution by minimizing the vehicle waiting time especially the emergency vehicles using fuzzy logic control under situations that normally occur during emergency. The effectiveness of this approach is tested by taking two traffic junctions.

Index Terms— Traffic lights control system, application of fuzzy logic, autonomous sytems, congestion control, FGPA based microelectronic chips, multiple emergency vehicles, and multi-agents based autonomous traffic control.

———————————————— ◆ ————————————————

1 INTRODUCTION

URBAN areas nowadays have a great deal of traffic jams especially when a number of junctions are taken into consideration. Air and noise pollutions, accidents, time wastage and so many other factors are of serious concerns in traffic. There are many other methods of controlling traffic lights in junctions to minimize the amount of traffic. Traffic can be detected by traffic light controller intelligently [1]-[2]. The use of fuzzy logic control in traffic signals are essential and plays an important role, however, when network intersections are concerned, the control problem occurs which is an important issue in traffic engineering [3].

S. Coleri, S. Y. Cheung and P. Varaiya et al discussed the various sensor networks for monitoring traffic [4]. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, worked on a survey on sensor networks used for traffic control system [5]. A. N. Knaian developed a wireless sensor network system for smart roadbeds and intelligent transportation using wireless sensor package to instrument roadways for Intelligent Transportation Systems. The sensor package counted passing vehicles, measured the average roadway speed, and detected ice and water on the road. Clusters of sensors can transmit this information in near real-time to wired base stations for use, controlling and predicting traffic, and in clearing road hazards [6]. W. J. Chen, et al proposed a real time dynamic traffic control system based on wireless sensor network [7]. M. Tubaishat, et al developed an adaptive traffic light control with wireless sensor networks and discussed various scenarios [8]. Y. Lai, et al discussed the protocols for traffic safety using wireless sensor network and described the design and implementation of energy-efficient protocols that can be used to improve traffic for an intelligent traffic management system deploying low-cost wireless sensor nodes on the roadbed and worked collaboratively to detect potential collisions on the road [9].

In this proposed research work, two junctions are discussed and brought into focus along with the use of wireless sensors as many studies suggest the use of sensors. There are no traffic lights control system in these junctions and huge amount of traffic causing wait time and accidents. As a result, physical existence of traffic police is always required there which is inappropriate due to availability of technology today. Also, due to heavy traffic in these two junctions, emergency vehicles face hardships when they pass from there. The two junctions have been divided into different directions and separations with installation of traffic lights control system and wireless sensors. Regarding traffic flow conditions on these junctions, per minute minimum 50 and maximum 115 vehicles passes by. The proposed Multi-Agent Based Autonomous Traffic Lights Control System uses Fuzzy logic Control system to overcome problems like congestion, accidents and speed to minimize the vehicles waiting time especially under emergency conditions.

The framework of this paper consists of: overview of multi-agent based autonomous traffic lights control system in section II, in section III design model of fuzzy logic based traffic control system is discussed, in section IV design algorithm of fuzzy based traffic system is discussed, in section V, simulation results are discussed and in section

————————————————

- *Mr. Yousaf Saeed is a Ph.D. Scholar in the field of Computer Sciences at National College of Business Administration and Economics NCBA& E, Lahore, Pakistan (e-mail: usafonline.email@gmail.com)*
- *Dr. M. Saleem Khan, Director Computer Science Department is with GC University, Lahore, Pakistan (e-mail: mskgcu@yahoo.com)*
- *Dr. Khalil Ahmad, Director Computer Science Department National College of Business Administration and Economics (NCBA& E), Lahore, Pakistan (e-mail: ahmedkhalil08@gmail.com).*
- *Mr. Abdul Salam Mubashar is doing Research in the field of Computer Science at National College of Business Administration and Economics (NCBA&E), Lahore, Pakistan (e-mail: salammunim@yahoo.com)*

VI, conclusion results are discussed.

2 OVERVIEW OF MULTI-AGENT BASED AUTONOMOUS TRAFFIC LIGHTS CONTROL SYSTEM

For inputs, three agents have been taken that include police, ambulance and fire brigade vehicles. Upon arriving, they will set their route on Global Positioning System (GPS) first, which will be located in their vehicles. Once the route is set, they will start moving on their proposed route while switching on their emergency lights and specific emergency siren. Sensors will be located at a specific distance before the junctions which will detect the speed and sound waves of siren at a particular threshold. Based on the speed, sensors will communicate wirelessly with the traffic control system of the two junctions while realizing their routes. On that basis traffic control systems of the two junctions will be able to minimize the traffic flow by inter-communication thus assigning the right time for red and green lights so that emergency vehicles can pass quickly. Fig.1 shows the block diagram of multi-agent based traffic control system. In case of more than one emergency vehicle coming from different directions, the traffic control system will be able to avoid collisions.
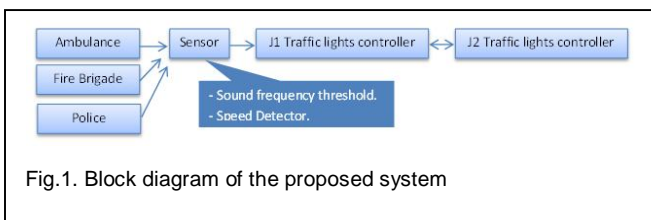

Fig.1. Block diagram of the proposed system

3 DESIGN MODEL OF FUZZY LOGIC BASED TRAFFIC CONTROL SYSTEM

The two junctions have been categorized as J1 and J2 as shown in Fig. 2. Wireless sensors S1, S2 and S3 have been used which are installed 500m away from the junctions. We are considering three emergency vehicles, ambulance, police and fire brigade, coming from three different directions at the same time with different speed ratio. Based on their speed detected by the sensors, two actions are performed. One, traffic flow is minimized on their routes so that they can pass by with their maximum possible speed and secondly, their collisions are avoided. In this scenario, first we consider an ambulance coming from the left side of Junction J1 having its route defined straight ahead. It will pass through roadways w1 and w7 with road codes 00001 and 00004 respectively. Secondly, we consid-

er a police vehicle coming towards junction J1 having its route defined and will pass through roadways w3 and w7 having road codes 00002 and 00004 respectively. Finally, we consider a fire brigade vehicle coming from the right side of junction J2 having its route defined and will pass through the road ways w11 and w5 having road codes 00006 and 00003 respectively.

Also, the speed of ambulance is detected by the sensor S1 which is 80 km/h, the speed of police vehicle is detected by sensor S2 which is 60 km/h and the speed of fire brigade is detected by sensor S3 which is 40 km/h. Each of these wireless sensors has been installed at 500m away from the junctions. However, the total distance between sensors S1 and S3 is 1.5 km. The expected time of these emergency vehicles are found with the help of the following formula:

$$Distance\ (s) = Velocity\ (v)/\ Time\ (t)$$

or

$$Velocity\ (v) = Distance\ (s)/\ Time\ (t)$$
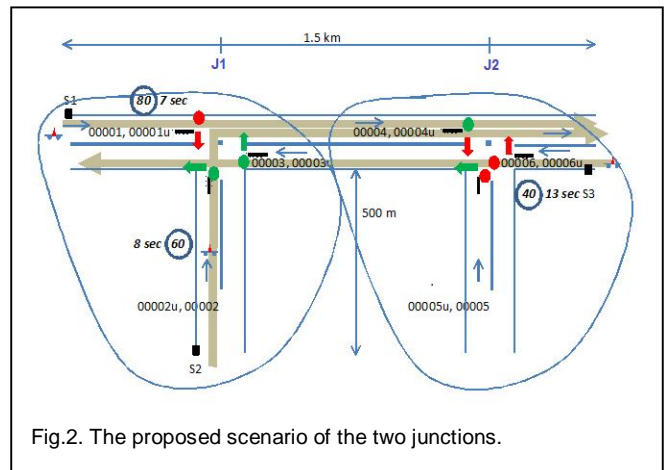
Fig.2. The proposed scenario of the two junctions.

4 DESIGN ALGORITHM OF FUZZY BASED TRAFFIC SYSTEM

The data at junction J1 containing sensors, traffic flow, road ways and traffic light status which is G for green and R for red are given in Table 1, however, yellow light has been ignored as it is marginal in this case and our main focus is on the green time to move the traffic as much as possible during the arrival of emergency vehicles.

TABLE 1
DATA OF JUNCTION J1

Junctions	J1					
Sensors	S1		S2			
Traffic Flow	100		60		95	
Road ways	w1	w2	w3	w4	w5	w6
Traffic Light Status	G	G	R	G	R	R
	G	G	R	G	R	R
G=Green Y=Yellow R= Red	R	R	G	G	G	G
	R	R	G	G	G	G

Table 2 shows the data at junction J2 containing sensors, traffic flow, road ways and traffic light status which is again G for green and R for red.

TABLE 2
DATA OF JUNCTION J2

Junctions	J2					
Sensors					S3	
Traffic Flow	115		50		80	
Road ways	w7	w8	w9	w10	w11	w12
Traffic Light Status	G	G	R	G	R	R
	G	G	R	G	R	R
G=Green Y=Yellow R= Red	G	G	R	G	R	R
	G	G	R	G	R	R

As each road has been assigned a special code, for u-turn there is 'u' written after the code, however, without the term 'u' means straight path. Also, categorization in the form of roadways and road codes has been defined in Table 3.

TABLE 3
ROADWAYS WITH THEIR RESPECTIVE ROADCODES

Road Ways	Road Codes
w1	00001
w2	00001u
w3	00002
w4	00002u
w5	00003
w6	00003u
w7	00004
w8	00004u
w9	00005
w10	00005u
w11	00006
w12	00006u

The code 00001 for instance contains five digits in which the first digit defines the existence of emergency vehicle. It is basically ON and OFF. If the code is in the form 00001, it means no emergency vehicle is detected whereas the code 10001 means that we have an emergency vehicle on this route.

The second digit defines the existence of ON and OFF of traffic light, which is 1 for green and 0 for red. For instance, the code 00001 means traffic light status is red on that route and 01001 means traffic light status is green on that particular route.

In case of emergency vehicle coming towards the road having code 00001 for instance will be of the form 11001 which again indicates the existence of emergency vehicle on the route and traffic light signal status which is green.

Upon getting all those relevant data, simulation software MATLAB is used for fuzzy control. Membership functions for both inputs and outputs have been defined in Table 4 and Table 5.

TABLE 4
INPUT MEMBER FUNCTIONS

0 – 20	Slow
0 – 40	Below Medium
20 – 60	Medium Fast
40 – 80	Fast
60 – 100	Very Fast

TABLE 5
OUTPUT MEMBER FUNCTIONS

0 – 20	Very Long
0 – 40	Long
20 – 60	Medium
40 – 80	Short
60 – 100	Very Short

Three inputs in the form of ambulance, police and fire brigade vehicles with their green time as their outputs are shown in Fig. 3.
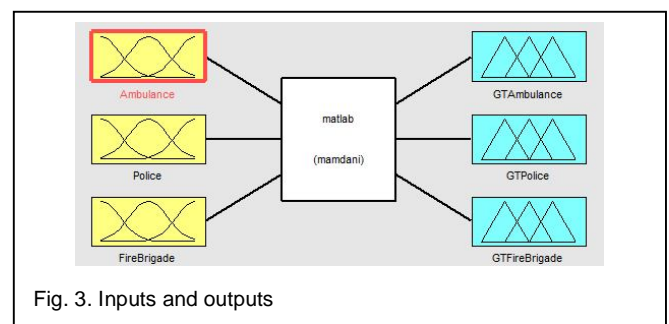

Fig. 3. Inputs and outputs

Five membership functions for each of input and output have been created. Fig. 4 shows the membership functions for input variable "Ambulance".
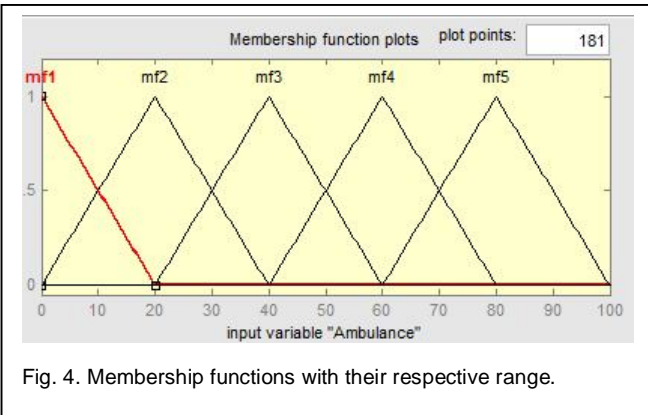

Fig. 4. Membership functions with their respective range.

A total of 60 rules have been generated. Rules were defined first by considering a single emergency vehicle with all possible speeds. Rules were then defined by considering two emergency vehicles with all possible speeds and finally, rules were defined for three emergency vehicles with all possible speeds. Moreover, by assigning the speed of our inputs, the rule viewer shows the respective green time for our outputs.

5 SIMULATION RESULTS AND DISCUSSION

Simulation software MATLAB generated rules in rule viewer shown in Fig. 5 on the basis of our data and finally surface view has been generated.
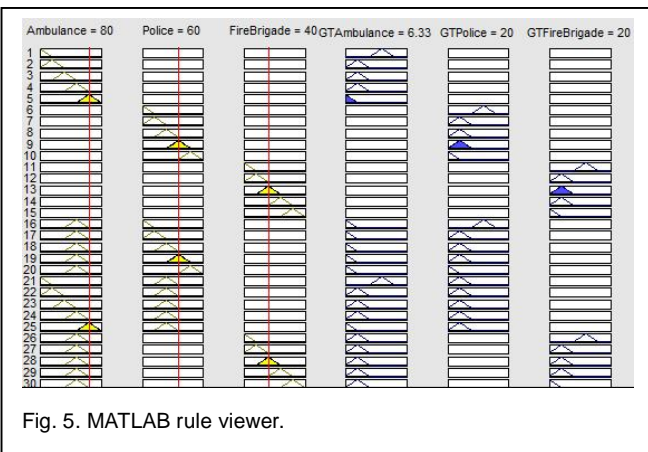

Fig. 5. MATLAB rule viewer.

Two surface views has been shown here that indicates the effect of output.
Fig. 6 shows a graph between the inputs: Ambulance and Police vehicles and output: Green time for ambulance.
Fig. 7 shows a graph between the inputs: Ambulance and

Fire brigade vehicles and output: Green time for ambulance.
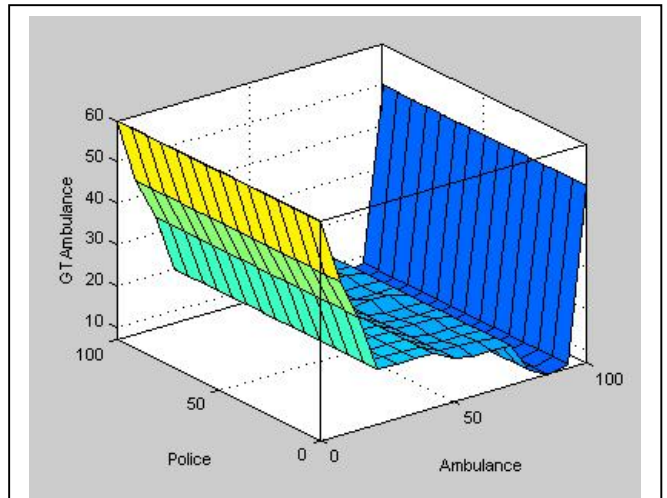
Fig. 6. Surface view for Ambulance and police vehicles.
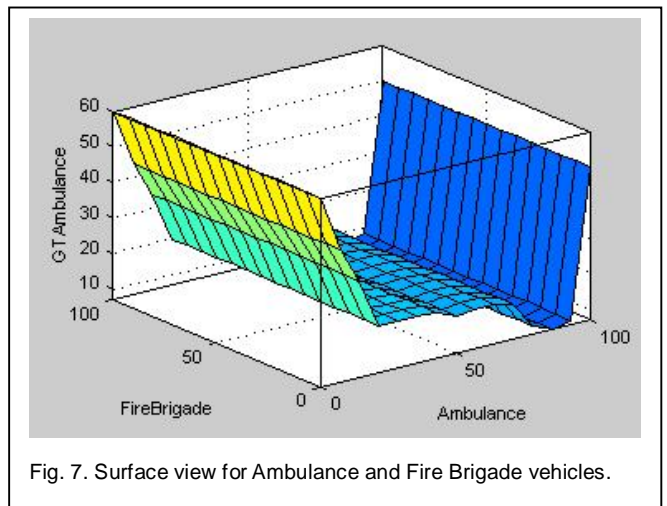
Fig. 7. Surface view for Ambulance and Fire Brigade vehicles.

6 CONCLUSION RESULTS AND DISCUSSION

Using proper data collection for the two junctions and proposing a traffic control system for multi-agents, the emergency vehicles passes from these two junctions quickly facing less traffic and at the same time collisions are avoided in case of multiple emergency vehicles coming from different directions.

In case of new emerging algorithms and hardware technology, the proposed system will be flexible enough to be enhanced in order to handle future traffic aspects using FPGAs based Microelectronics chips to control traffic signals.

ACKNOWLEDGEMENT

REFERENCES

[1] Cheonshik Kim, You-Sik Hong. "Traffic signal using smart agentsystem" *American Journal of Applied Sciences*, 2008: 1487-1493.

[2] Tzafestas et al. "Advances in Intelligent Autonomous Systems: Microprocessor based and Intelligent Systems Engineering" Kulwer Academic Publishers, 1999.

[3] Skowronski, W., L. Shaw. "Self-learning fuzzy traffic controller for a traffic junction. I." *European Congress on Intelligent Techniques and Soft Computing*-EUFIT 93, Aachen, Germany, 1993: 751-761.

[4] S. Coleri, S. Y. Cheung and P. Varaiya. "Sensor networks for monitoring traffic" *Proceedings of the 42nd Annual Allerton Conference on Communication, Control and Computing* 2004, pp. 32-40.

[5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. "A survey on sensor networks" *IEEE Communications Magazine*, Vol. 40, 2002, pp. 102-114.

[6] A. N. Knaian. "A wireless sensor network for smart roadbeds and intelligent transportation systems" *Technical Report, Electrical Science and Engineering*, Massachusetts Institute of Technology, June 2000.

[7] W. J. Chen, L. F. Chen, Z. L. Chen, and S. L. Tu, "A real time dynamic traffic control system based on wireless sensor network" *Proceedings of the 2005 International Conference on Parallel Processing Workshops*, Vol. 14, 2005, pp. 258-264.

[8] M. Tubaishat, Y. Shang, and H. Shi. "Adaptive traffic light control with wireless sensor networks" *Proceedings of IEEE Consumer Communications and Networking Conference*, 2007, pp. 187-191.

[9] Y. Lai, Y. Zheng, and J. Cao. "Protocols for traffic safety using wireless sensor network" Lecture Notes in Computer Science, Vol. 4494, 2007, pp. 37-48.

Experimental cost reduction in Modern Manufacturing Industries through Software based Mechatronics System: An overview

Ramamoorthy.C, Dr. V. Selladurai and Dr. Rajesh Ranganathan

Abstract — Nowadays, the changing demand patters by the customers are increasing the pressure on the manufacturers. In this situation, manufacturers to sustain in the market are forced to do the activities like (i) Increase the selling price of their product and (ii) Reduce their internal costs of their product. In this scenario, it is not possible to increase the selling price of the products. The only way is to reduce their internal costs and improve the quality of the products to survive in their field. This paper suggests a methodology to develop and implement software based mechatronics systems in their factory premises for taking necessary steps to reduce the experimental cost, internal cost and improve the efficiency of the organization. The software based mechatronics approach is used to predict and analyze the relevant data pertaining to the selected problems. The usefulness of this approach is not only applicable for pump manufacturing industries but also various industrial segments, which would enable to cater the current and future customer demanding needs.

Index Terms — Minimum Economic crisis, software based mechatronics system, experimental cost reduction, Internal cost reduction, efficiency improvement, economic growth.

———————————— ◆ ————————————

1 INTRODUCTION

AS global competition continues to intensify across industries, companies are actively pursuing strategies that will enable them to compete more effectively and improve profitability. In modern manufacturing industries, software based mechatronics system are widely used to operate, monitor and control of various processes. Modern electro mechanical system consists of sophisticated software and hardware components to achieve high accurate, precision and reliable manufacturing processes.

According to Danial (1998) [1], traditional computer programming involves setting down a list of tasks for the computer to execute in the given sequential order. Terutomo (2001) [2] stated that in the early 1980s, the mechatronics based computer aided flexible manufacturing systems were introduced in order to improve the efficiency of the organizations. Claimed by Anders (2004) [3], the complexity of the mechatronics system emphasizes the need for a virtual machine concept for dealing with this trade-off during product development. Godfrey (2005) [4] expressed that mechatronics responds to industry's increasing demand for engineers who are able to work

across the discipline boundaries of electronic, control and mechanical engineering to identify and use the proper combination of technologies for optimum solutions to today's increasingly challenging engineering problems.

As per the report given by Amuthakkannan (2007) [5], in high precision manufacturing industries, software based systems are widely used to obtain precision, accurate and reliable data. Based on the report given by Mark, et al., (2008) [6], mechatronics involved the fusion of mechanical technology with electrical and electronics, computer science, material technologies etc. As described by Sophie (2009) [7] mechatronics is the industry's principal route to meeting the ever increasing demand for more performance at lower cost and in a smaller footprint. Agoria (2010) [8] studies show that a mechatronics design approaches is a key to all research avenues in order to drastically improve their overall performance (adaptability, productivity, quality, reliability and lifecycle costs (i.e. reduced energy consumption, reduced waste).

2 PROBLEM STATEMENT

Modern manufacturers are under intense, unrelenting pressure to find new ways to cut costs, improve quality, and boost customer satisfaction. These parameters are usually achieved through developing creative and innovative ideas. The software based mechatronics system is a relevant approach which provides lots of newer, creative and innovative ideas to develop virtual machinery for the prediction and analyzes of relevant data in order to meet experimental cost reduction and performance targets. Realizing the potential for software based mechatronics

———————————————

- **RAMAMOORTHY.C**, Research Scholar, Department of Mechanical Engineering, Coimbatore Institute of Technology, Coimbatore – 641 014, Tamilnadu, India, E-mail: ramsun.turbo@gmail.com

- **Dr.V.SELLADURAI**, Principal, Coimbatore Institute of Technology, Coimbatore – 641 014, Tamilnadu, India, E-mail: selladurai.v@gmail.com

- **Dr. RAJESH RANGANATHAN**, Associate Professor, Department of Mechanical Engineering, Coimbatore Institute of Technology, Coimbatore – 641 014, Tamilnadu, India, E-mail: drrajeshranganadhan@gmail.com

system in the industrial environment, this paper is focused to cater future projected industrial needs. The developments of virtual machinery using software based mechatronics system for cater future demand and its parameters are taken as research problem and solution methodologies are proposed.

3 LITERATURE REVIEW

According to Ritchie and Lewis (2003) [9], the wider literature about the role of existing theory, practice and research will shape future research. Kothari (2004) [10] expressed that the researcher should undertake extensive literature survey connected with the problem. Russell (2006) [11] stated that the contemporary practice of posing research is a wider literature is consistent with more recent articulations of grounded theory methodology and of the closely related case study methodology. The literature review assesses the past and current status of research work in the area of mechatronics system and software based mechatronics approach. The work done by the earlier researchers in these areas are classified in the following aspects:

3.1 Mechatronics System

Dobrivoje (1999) [12] expressed that in the past several decades, the rapid development of electronics industry and ever increasing applications of computers and the automation of various industries, demand for mechatronics applications have increased exponentially. As stated by MBT (2000) [13], the use of mechatronics systems views and validates electro-mechanical processes for the most efficient production operation possible. Christopher (2004) [14] outlined that the stages of a product's life include engineering, production and overall testing helps validate appropriate designs, verify appropriate production, and thus reduce wastes. Godfrey (2005) [4] outlined that Mechatronics responds to industry's increasing demand for engineers who are able to work across the discipline boundaries of electronic, control and mechanical engineering to identify and use the proper combination of technologies for optimum solutions to today's increasingly challenging engineering problems.

As per the report given by Eiji and Tatsuo (2007) [15], the manual operation process is a waste of time, waste in transportation, waste of movement, overburden etc. for each task. Ernest and Stanislav (2008) [16] indicated that the integrated structures, mechatronics and robotics, combine computers and their programs into mechanisms are used to monitor them by sensors (transducers). MBT (2009) [17] suggested that making analysis more integral to design; blending of mechanical, electrical, and electronic design is called mechatronics. As per the report given by SPLMS (2010) [18], mechatronics system views and validates electro-mechanical processes for the most efficient production operation possible. Based on the literatures it is clear that software programming is essential to

achieve the desired innovative system.

3.2 Software based Mechatronics System

Programmers develop software applications every day in order to increase efficiency and productivity in various situations. Johnson (1998) [19] stated that the LabVIEW software has advanced the state of software development in graphical programming, user interface design, and development environment. NIC (2003) [20] reported that LabVIEW is a graphical programming language that uses icons instead of lines of text to create applications. As described by Christopher (2004) [14], the control software is to be designed to allow operators with little or no computer experience to monitor and control the system.

Eiji and Tatsuo (2007) [15] expressed that well designed software is an excellent computational tool to simulate dynamic mechatronics systems. Instead of measuring a process variable directly with one physical sensor, measurements form several sensors and other knowledge of the process are incorporated using software to obtain an even more accurate measurement. Bitter (2007) [21] point of view, the ability of the software to provide abstraction is also significant because it improves code readability. CNA (2008) [22] studies shows that the control electronics are designed and manufactured with tailored software adaptations to further improve operation in the machines or even to integrate the machine functions.

As indicated by NIC (2009) [23], LabVIEW is a powerful development environment for signal acquisition, measurement analysis, and data presentation, giving you the flexibility of a programming language without the complexity of traditional development tools. NIC (2010) [24] outlined that, with LabVIEW software, it can be quickly create user interfaces that give you interactive control of your software system. Claimed by Pedro and Fernando (2010) [25], LabVIEW is a graphical control, test, and measurement environment development package. It is one of the most important software platforms for developing engineering applications and could be connected with different hardware systems, as well as running standalone programs for simulating the controller's performance. Luis (2011) [26] pointed out that the intelligent and autonomous software programs are capable of interacting with other software components within a given application, and sharing a common goal.

4 BACKGROUND OF THE PAPER

At the starts of the paper, the analyses of various concepts which are essential are to be studied. Peter (1999) [27] expressed that the background concepts provides the research and researcher to make sense of the 'data' and the study to proceed in a systematic way. As described by Brian (2003) [28], the different backgrounds for the selective area are reducing the gap between the theoretical studies and practical analysis. In this paper the back-

grounds of mechatronics system and software based mechatronics approach were presented.

4.1 Background of mechatronics system

The new role of mechatronics in the 21st century is the one of the main technology to support people directly by intelligent machines and systems with sensor, information technology, network, and human sciences. IAG (2006) [29] reported that the Iskra Avtoelektrika Group increased their sales volume by 23 % in comparison with the year before mainly as a result of the mechatronics applications. The share of new products sales is by 5 percentage points higher than in the previous year and represents 30 % of the total sales. As suggested by John (2007) [30], about 15% of all investments in the process industry are currently allocated to mechatronics devices. There is an important tendency of some growing sectors, specifically the equipment industry, which investment has grown at a rate of 11.2% in 2005 and 31% in 2006. Gerard (2008) [31] expressed that an increase in automation, to detect early failure diagnostics of the machines, the mechatronics share of the production costs of machines is gradually increasing to about 10% to 20%.

4.2 Background of software based mechatronics system

Since the dawn of computing in the 1940s, the applications and uses of computers have grown at a staggering rate. According to ACM (2004) [32], software plays a central role in almost all aspects of daily life: in government, industries, banking and finance, education, transportation, entertainment, medicine, agriculture, law etc. As suggested by Godfrey (2005) [4], mechatronics is defined as the synergistic combination of precision mechanical, electronic, control, and systems engineering, in the design of products and manufacturing processes. Agoria (2006) [8] stated that most modern engineering products already make use of the close interaction between classical mechanics, electronics, control engineering and software that is known as mechatronics.

John (2007) [30] expressed that about 15% of all investments in e.g. the process industry are currently allocated to mechatronics devices. Application of mechatronics can provide a large competitive advantage. Production costs can be reduced by reducing waste costs but also labour costs. There is an important tendency of some growing sectors, specifically the equipment industry, which investment has grown at a rate of 11.2% in 2005 and 31% in 2006. Mark (2008) [6], et al., insisted that computer innovations increased from 26% to 74% between the years 1963 to 2008 in UK firms.

5 RESEARCH APPROACH

As stated by Kothari (2004) [10], research approach is a way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically. The various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them are to be taken. The solution given in this paper has been carried out by using the research approach as shown in Figure 1.

Fig.1. Research approach
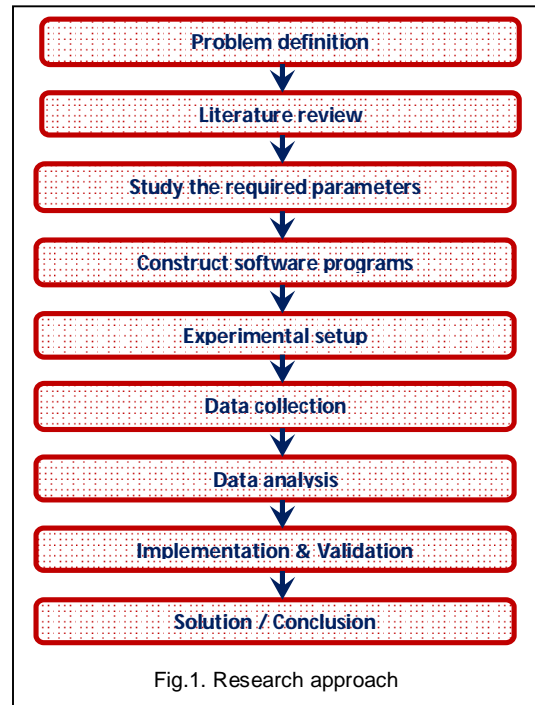
As per the field work approach, the relevant conceptual literatures concerning the concepts, theories, and empirical case studies were studied. After the conceptual stage, arrangements should be made for the selection of relevant manufacturing industry to carry out the case studies. The desired parameters of software based mechatronics system were analyzed before conducting the programme. The relevant data pertaining to the selected problems were predicted and analyzed though software based mechatronics system. The optimal solution was tested through proper validations before implementing in regular production.

6 CASE STUDY

Bill and John (2001) [32] expressed that case study leads to obtain very novel idea and no longer limited to the particular individual. Kothari (2004) [10] outlined that a case-study is an in-depth approach to reach the basic relationship between the theoretical and practical aspects.

6.1 Formation of team and time frame

As reported by WoS (2009) [34], the teams or work forces that can meet deadlines and solve the selected problems by planning and streamlining activities. A team comprising of skilled personnel from each department had been formed for planning, monitoring, evaluating and directing. This team had been discussed among the members and with the top management; the time frame for conducting integrative programme had been formulated.

6.2 Product selection

Understanding the importance of product selection, the team has to gone through the customer care records, customer complaints book, service engineers' log book and all. Marco (2009) [35] studies shows that during the product selection the literatures for product life cycle, providing knowledge based environment to support value creation to the customers in term of creation, innovation, lead time reduction, quality improvement, performance improvements were considered. A sample product called Shallowell jet pump is shown in figure 2.
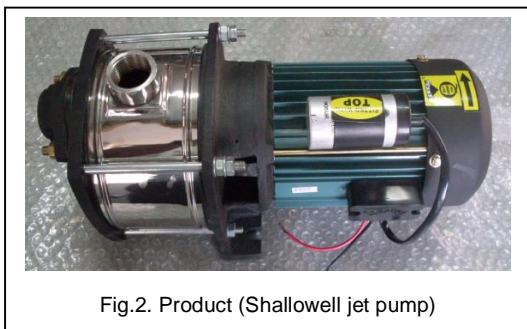

Fig.2. Product (Shallowell jet pump)

6.3 Apparatus selection

The selection of apparatus for manipulating the input signals and output signals are one of the important phases for experimentation. According to Robert (2008) [37], the accuracy of apparatus describes the amount of uncertainty when considering quantization error, unavoidable system noise, and hardware imperfections. Anders (2004) [3] stated that a measuring instrument was designed which had an accuracy of a millionth of an inch, increasing the accuracy of the machines dramatically. Claimed by BEC (2006) [38], high precise, accurate and reliable measuring apparatus will improve the confident level of the employers, employees and customers.

6.4 Data acquisition card

Christopher (2004) [14] expressed that choosing the right combination hardware and software can save hours of development time and strongly influence the performance of built application. As stated by Cory (2005) [39], data acquisition card is in some sense a hybrid device in that aspects of both the conventional digital receiver and

the sub-sampling receiver are used to acquire the desired signal. Arun (2006) [40] indicated that the Data Acquisition (DAQ) system has been interfaced to computer through parallel port for transferring the acquired data to the computer. As described by Bitter (2007) [21], taking data from a data acquisition card can now graph the outputs on a digital graph. Pedro and Fernando (2010) [25] explained that the data acquisition system plays a key role in the control system design. A sample DAQ (PCI-6034E) card is shown in Figure 3.
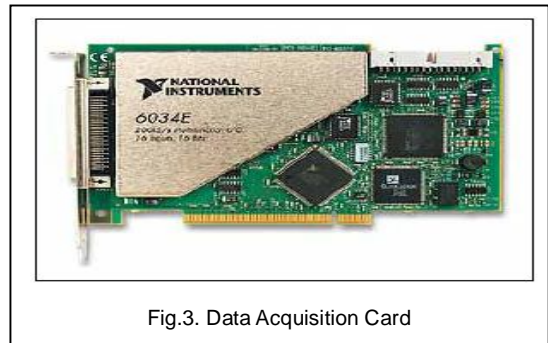

Fig.3. Data Acquisition Card

6.5 LabVIEW Software

The National Instruments' LabVIEW Software is a graphical development environment which provides advanced functionality and performance that engineer and scientists can use to develop sophisticated applications. Johnson (1998) [19] suggested that LabVIEW is a new kind of graphical programming language that has brought the power of the computer to the scientists and engineers in a powerful new metaphor that is natural to the user. As per the report given by Thomas (2003) [41], an experienced LabVIEW user can program the application easily with this information. Gary and Richard (2006) [42] outlined that LabVIEW brings the ease of graphical programming to the arcane world of real-time system programming. Jeffrey and Jim (2006) [43] point of view, LabVIEW create programs using a graphical notation (connecting functional nodes via wires through which data flows); in this regard, it differs from traditional programming languages like C, C++, or Java, in which program with text. Sumathi and Surekha (2007) [44] reported that LabVIEW introduced graphical user interfaces and visual programming into computerized instrumentation, joining simplicity of a user interface operation with increased capabilities of computers. NIC (2010) [24] reports show that LabVIEW is a highly productive graphical programming language for building data acquisition and instrumentation systems.

6.6 Software development

Johnson (1998) [19] explained that the increase in productivity comes from the interactive approach to software

development. As expressed by Christopher (2004) [14] software approach is used for hardware and other capabilities, which also minimizes the integration time and effort for future developments. VMWare (2006) [45] reports shows that software development teams to more efficiently utilize software development and test lab assets, accelerate software development cycles, and increase the quality of delivered software systems. Sumathi and Surekha (2007) [44] suggested that instrumentation software enabled the creation of a simulated physical instrument, having the capability to control physical sensing components. Pedro and Fernando (2010) [25] outlined that software platforms for developing engineering applications and could be connected with different hardware systems, as well as running standalone programs for simulating the controller's performance (validating the controller by simulation then implementing it). Sample software development using LabVIEW software is shown in figure 4.
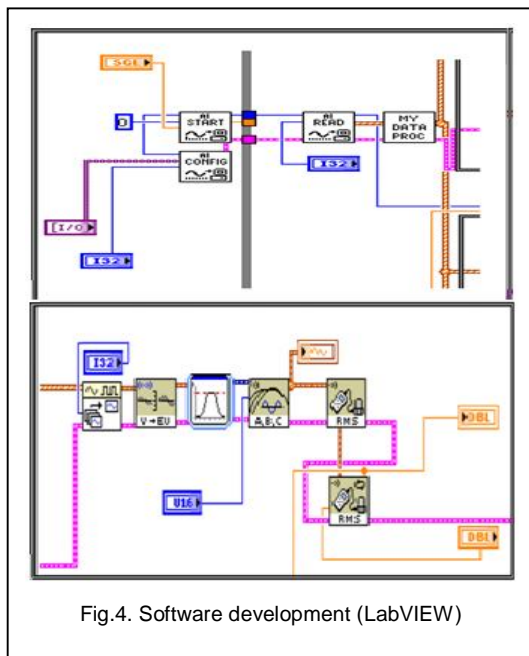

Fig.4. Software development (LabVIEW)

6.7 Experimental setup

The experimental method is a scientific method. It is oriented to the future in the sense that the researcher is seeking to evaluate something new. The experimental groups are given a description of the basic design rules derived from the literature like Yogesh (2006) [46], Lisa, et al. (2008) [47] and Lucienne and Amaresh (2009) [48] and applied these while designing. The designers in the control group are only given the design task. The task is based on one of the cases in the empirical study. A sample experimental setup is shown in figure 5.
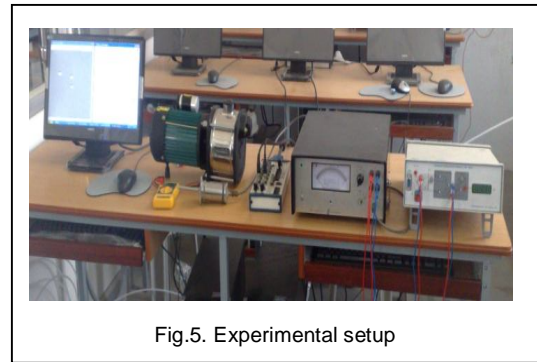

Fig.5. Experimental setup

7 PREDICTION AND ANALYSIS OF DATA USING SOFTWARE BASED MECHATRONICS SYSTEM

The prediction and analysis of present data in a process / product will play a vital role in an organization during the decision marking about the quality improvements, performance improvements, processes improvements, newer innovations etc. In order to examine the existing data in products / processes, this paper insists them to utilize the software based mechatronics system. Godfrey [4] stated that mechatronics is the synergistic combination of precision mechanical engineering, electronic control, and systems thinking in design of products and manufacturing processes. According to Mahalik (2003) [36], an automatic control system could perform a repetitive job; but could not take any decision in the event of variable circumstances. So, the introduction of software based system monitors the status of the process continuously and takes corrective action dynamically to stabilize the process.

7.1 Experimentation

As suggested by Auston (1997) [49], experimental methodology would also improve the quality of researches. Pande and Holpp (2002) [50] expressed that experimental is a method for developing and conducting controlled assessments of how a process or a product performs, usually testing two or more characteristics under different conditions. In addition to helping target causes of a problem, it can be essential to get maximum benefit out of solutions. USEPA (2007) [51] insisted that the experimentation may make sense to start testing techniques at the organizational levels help to identify synergies and opportunities and evaluate how well they worked. Adapt the validated results to fit into the organizational systems and culture.

Anders (2004) [3] pointed out that conducting experimentation is the purpose of achieving cost-efficient and accurate enough prediction of data to support optimization and well-informed design decisions during research and

development. Claimed by AISB (2005) [52], carrying experimental methods of the arts into research in order to create a new form of inquiry that has real agency on social, conceptual and economic levels. Agoria (2006) [8] studies show that the experimental models, being able to reproduce many typical experimental tests, permit, still in the design phase, to optimize the process to a much larger extent than traditional tools.

As stated by Nikolay (2007) [53] after the completion of successful experiment, the acquired knowledge reaches another point at which the individual is able to control the entity. Wolfgang and Stanislav (2008) [54] indicated that experiments include analytical or approximate calculation or simulation models derived from engineering sciences. As described by Sophie (2009) [7], experimentation is required to determine what parameters are necessary to be found by the research at long term to achieve the goal. Sample experimentation is shown in figure 6.
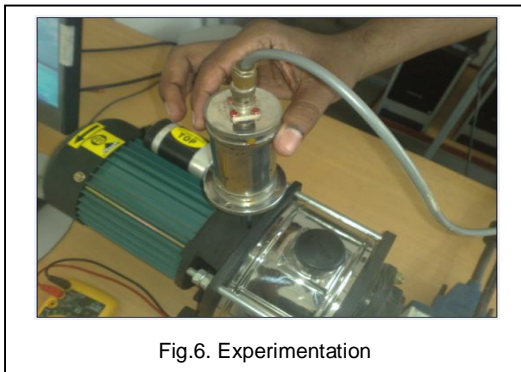

Fig.6. Experimentation

7.2 Experimental outcomes

The measurements of outcomes are usually needed accompanied by some investigation of process. The ability to turn information and data into effective action brings desirable outcomes for a research / organization. Anders (2004) [3] stated that the outcome of the experimentation is crucial for future research.
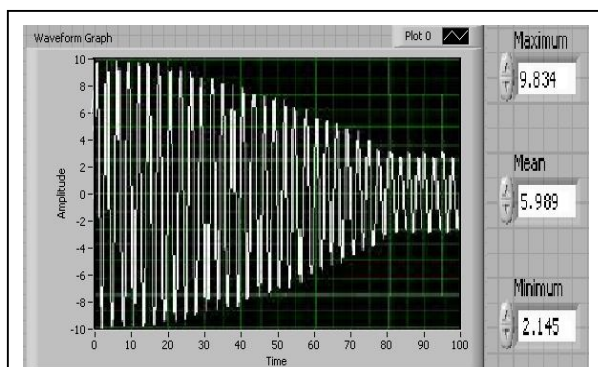

Fig.7. LabVIEW front panel

Claimed by AISB (2005) [52] the outcomes in order to evaluate a research's success and is crucial to shift the focus onto the dynamics of the evolving discourse that led to the concrete results. Wolfgang and Stanislav (2008) [54] expressed that designing more efficient, effective, rational, better directed, with a better outcome becomes a welcome addition to the information for researchers. The sample outcomes of vibration level through LabVIEW front panel are shown in figure 7.

TABLE 1
PREDICTION OF VIBRATION LEVEL

SI No	Pump No.	Time (Sec)	Amplitude (µm)		
			Max	Min	Avg
1	TRS10250011	60	5.83	2.14	3.98
2	TRS10250012	60	5.76	2.07	3.92
3	TRS10250013	60	5.90	2.21	4.05
4	TRS10250014	60	5.94	2.25	4.09
5	TRS10250015	60	5.66	1.96	3.81
6	TRS10250016	60	5.54	2.75	4.14
7	TRS10250017	60	5.90	2.21	4.05
8	TRS10250018	60	6.00	2.31	4.15
9	TRS10250019	60	5.80	2.01	3.91
10	TRS10250020	60	5.60	1.91	3.76

Max = maximum, Min = Minimum, Avg =Average

As stated by Mark and Peter (2001) [58] the innovations and changes that occur in all successful organizations are disruptive and dysfunctional unless their purpose is well articulated and communicated, they are coordinated, and those affected by them feel a sense of ownership and influence over their nature and outcomes. Based on the statement given by Luis (2011) [26], the outcomes of present research will involve not only researchers and academics in general, but also companies who wish to assess the potential for transfer of results to their applications. The sample outcomes of vibration analysis are shown in table 1.

7.3 Experimental validation

According to Janeri and Lewis (2003) [55], as a consequence, numerous suggestions are made about how to cross-check the validity of a finding or conclusion (validation) or to allow sufficient access to the research process for others to do so themselves (documentation). Checking accuracy of fit which involve deriving hypotheses from one part of the data and testing them on another by constant checking and comparison across different. Anders (2004) [3] pointed out that validation aims at checking if a model is suitable for its intended purpose by comparing the experimentation results to what is expected by the user or to results obtained from studying the real system.

Agoria (2006) [8] studies show that validations of the present outcomes are to be done prior to starting the regular operations. Eiji and Tatsuo (2007) [15] reported that the validation of the proposed model identified by experiments shows well corresponding to real attendant's propelling behaviour. As expressed by SPLMS (2010) [18] validation enables manufacturing companies to accurately and efficiently define the manufacturing process plan for their component parts. Luis (2011) [26] point of view, for ensuring high quality solutions; the proposed systems are to be undergone verification and validation.

8 RESULTS AND DISCUSSIONS

As stated by Murray (2008) [56], positive results helps create and maintain team spirit and commitment to the team objectives, most importantly, it may drive innovation within the team. Rouhani (2006) [57] insisted that the research results are significant, taking into the account that traditional industry has to be changed into global competitive, innovative and technology developed industry. As expressed by Russell (2006) [11], the experimental results of an investigation can be conducive to further innovative work. Janeri and Lewis (2003) [55] reported that the discussions suggest a revision of the research objectives, or a radical change in the way in which the data are collected, and then there may be more reason to consider or not. Mark and Peter (2001) [58] pointed out that the discussion aims to describe and explain some different funding options for qualitative research and to highlight the processes and issues involved in each. Lisa, et al., (2008) [47], discussion concerning a problem / solution often produces useful information. Various newer concepts can be developed through such an exercise.

8.1 Cost and time saving applications

AISB (2005) [52] outlined that the skills required to create reliable mechatronics design at low cost are acquired by experience. BEC (2006) [38] suggested that machine builders are expected to develop machines with higher throughput that, at the same time, reduce operational costs and increase safety. As expressed by Agoria (2006) [8] a mechatronics design approach is a key to all three avenues and there is likely to be hybrid integrations of two or more approaches in order to meet cost/performance targets for particular sectors and applications. Brian (2007) [59] explained that the simplicity of the mechatronics design provides relatively low experimental costs, less time to construct and more data reliability. The expected cost and time saving applications are shown in figure 8.
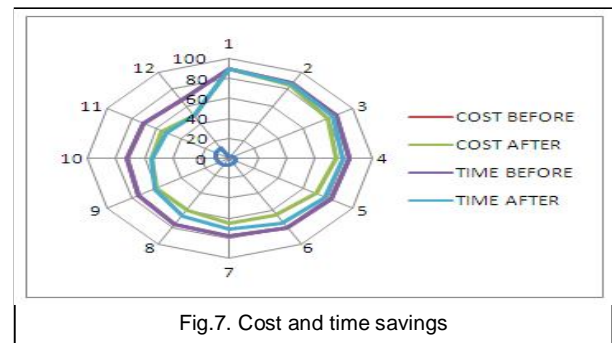

Fig.7. Cost and time savings

8.2 Reliability of software based mechatronics system

John (2007) [30] studies show that a more reliable process data will lead to higher quality products, processes and services. As described by Wolfgang and Stanislav (2008) [54] correct and reliable information, at the right time, and in a usable form is currently one of the problems. Luis (2011) [26] indicated that science and industry are constantly evolving that whether it is possible to create a system capable of facilitating communications, yet be portable, reliable, eventually self-powered and very simple to use. These parameters are fulfilled by mechatronics design.

9 CONCLUSIONS

This paper conclusion is based on the conceptual analysis of the software based mechatronics approach. Present manufacturing industry attention illustrates the need for the integration of manufacturing process and a software simulation for reaching a better solution within short time and low cost. This paper brings few ideas on software based mechatronics system to fulfill the manufacturing environments' expectations. This approach had bring the number of benefits like low cost for experimental, reliable data, lead time reduction, productivity improvements, quick change over, more profitability, more customer satisfaction and innovative technology developments.

This paper has an approach in that software based mechatronics approach to ensure an approach that will benefit end-users. This approach to manufacturing via process monitoring and control, supported strongly by product performances and characterization of the final product to ensure a component / process of high quality in all aspects has been applauded. This innovative approach has yielded extremely positive benefits to all manufacturing industries internationally.

nology, Coimbatore, India for his support and providing them with efficient lab facilities for perusing this research. They would like to thank the management and employees of M/s. Mayur Motor Industries, Coimbatore, India and M/s. National Instruments Corporation, Bangalore, India for their continuous and inestimable co-operations during the experimental setups, predictions and analysis of relevant data.

REFERENCES

[1] Danial J. Neebel, et al., *A Survey of Modern Computer-Based Experiments*, published by James Madison University, Texas, p.13, 1998.

[2] Terutomo Ozawa, *The hidden side of the flying Geese: Model of catch-up growth*, published by East West Centre, Hawaii, p.04, 2001.

[3] Anders Jönsson, *Lean Prototyping of Multi-body and Mechatronics Systems*, PhD thesis, Blekinge Institute of Technology, Sweden, p.12, 2004.

[4] Godfrey C. Onwubolu, *Mechatronics Principles and Applications*, published by Elsevier Publications Ltd, UK, p.13, 2005.

[5] Amuthakkannan.R, *Optimal design of Software based mechatronics system*, PhD thesis, Anna University of Technology, Chennai, p.64, 2007.

[6] Mark Dodgson, et al., *The Management of Technological Innovation*, published by Oxford University Press, New York, p.64, 2008.

[7] Sophie Vandebroek, *Xerox Innovation: Creating the Future Today*, published by Xerox Innovation Group, Webster, p.05, 2009.

[8] Agoria J.P., *Mechatronics: New Generation of Production Systems*, published by High Manu-future Group, Belgium, p.49, 2006.

[9] Ritchie Jane and Lewis Jane, *Qualitative Research Practice: A Guide for Researchers*, published by SAGE Publications Ltd, London, p.48, 2003.

[10] Kothari. C.R., *Research Methodology*, published by New Age International Publishers, New Delhi, p.219, 2004.

[11] Russell W. Belk, *Handbook of Qualitative Research Methods in Marketing*, Published by Edward Elgar Publishing Limited, United Kingdom, p.21, 2006.

[12] Dobrivoje Popovic and Ljubo Vlacic, *Mechatronics in Engineering Design and Product Development*, published by Marcel Dekker Inc., New York, p.101, 2000.

[13] MBT (Manufacturing Business Technology), *Impact on the midsize manufacturing enterprise*, published by Manufacturing Business Technology, p.05, 2000.

[14] Christopher G. Relf, *Image acquisition and processing with LabVIEW*, published by CRC Press, Boca Raton, p.60, 2004.

[15] Eiji Arai and Tatsuo Arai, *Mechatronics for Safety, Security and Dependability*, published by Elsevier Publishing Limited, UK, p.33, 2007.

[16] Ernest W Elder and Stanislav Hosnedl, *Design Engineering: A manual for enhanced Creativity*, published by CRC Press, New York, p.49, 2008.

[17] MBT (Manufacturing Business Technology), *Globalization's 21st century impact on the midsize manufacturing enterprise*, published by Manufacturing Business Technology, UK, p.05, 2009.

[18] SPLMS (Siemens Product Lifecycle Management Software Inc.), *Digital manufacturing solutions* published by Siemens Product Lifecycle Management Software Inc., USA, p.16, 2010.

[19] Johnson, Gary W, *LabVIEW Power Programming*, published by McGraw-Hill Professional, New York, p.33, 1998.

[20] NIC (National Instruments Corporation), *LabVIEW: Getting Started with LabVIEW*, published by National Instruments Corporation, USA, p.06, 2003.

[21] Bitter Rick, et al., *LabVIEW: advanced programming techniques*, published by CRC Press, Boca Raton, p.182, 2007.

[22] CNA (Crouzet North America), *Innovative D.C. motorization solutions adapted to your applications*, published by Crouzet North America, California, p.33, 2008.

[23] NIC (National Instruments Corporation), *LabVIEW Graphical Development Environment*, published by National Instruments Corporation, USA, p.43, 2009.

[24] NIC (National Instruments Corporation), *LabVIEW: Measurement and Automation Software*, published by National Instruments Corporation, Texas, p.05, 2010.

[25] Pedro Ponce-Cruz and Fernando D. Ramírez, *Intelligent Control Systems with LabVIEW*, published by Springer-Verlag London Limited, UK, p.04, 2010.

[26] Luis M. Camarinha-Matos, *Technological Innovation for Sustainability*, published by Springer Publication Inc., Heidelberg, p.384, 2011.

[27] Peter Woods, *Successful Writing for Qualitative Researchers*, Second edition published by Routledge Publishers Inc., Abingdon Oxon, p.49, 1999.

[28] Brian Atkin, et. al., *Construction Process Improvement*, published by Blackwell Publishing Company, London, p.33, 2003.

[29] IAG (Iskra Avtoelektrika Group), *Iskra Avtoelektrika in 2006: Annual Report*, published by Iskra Avtoelektrika Group, Slovenia, p.07, 2006.

[30] John Blankendaal, *Mechatronics Sector Study Brabant – Catalunya*, published by Mateo – Match Tech, London, p.05, 2007.

[31] Gerard C.M. Meijer, *Smart Sensor Systems*, published by John Wiley and Sons, Ltd, Publication, West Sussex, p.13, 2008.

[32] ACM (Association for Computing Machinery), *Software Engineering*, published by Association for Computing Machinery, UK, p.05, 2004.

[33] Bill Martinson and John A., *Molinar lean manufacturing training*, published by Minnesota Technology, Inc., West central Minnesota, p.41, 2001.

[34] WoS (Wings of Success), *Lean Manufacturing Principles Made Easy*, published by Wings of Success, London, p.41, 2009.

[35] Marco Taisch, et al., *Lean Product and Process Development*, PhD thesis, Cranfield University, UK, p.33, 2009.

[36] Mahalik. N.P., *Sensor networks and configuration: fundamentals, standards, platform and standards*, published by Springer publishers Inc, German, p.33, 2003.

[37] Robert H Bishop, *Mechatronics system control, logic and data acquisition*, published by CRC Press, Taylor & Francis Group, London, p.296, 2008.

[38] BEC (Baldor Electric Company), *Control design for machine builders*, published by Baldor Electric Company, USA, p.07, 2006.

[39] Cory L. Clark, *LabVIEW Digital Signal Processing and Digital Communications*, published by The McGraw-Hill Companies, Inc., New York, p.06, 2005.

[40] Arun Venkatesh, et al., *Design of FFT Spectrum Analyzer on PC Platform using Labview and Microcontroller based Dataacquisition System"*, Journal of Instrumentation Society of India, Volume 36, Issue 1, p.06, 2006.

[41] Thomas Klinger, *Image Processing with LabVIEW and IMAQ Vision*, published by Pearson Education, Inc., New Jersey, p.33, 2003.

[42] Gary W. Johnson and Richard Jennings, *LabVIEW Graphical Programming*, published by the McGraw-Hill Publishers Inc., New York, p.13, 2006.

[43] Jeffrey Travis and Jim Kring, *LabVIEW for Everyone - Graphical Programming Made Easy and Fun*, published by Pearson Education, Inc., New Jersey, p.33, 2006.

[44] Sumathi S. and Surekha P., *LabVIEW based Advanced Instrumentation*

Systems, published by Springer-Verlag Berlin, Heidelberg, p.29, 2007.

[45] VMWare, "Virtual Lab Automation: A Quantum Leap in IT Cost Reduction and Application Development Process Improvement", published by VMware, Inc., USA, p.11, 2006.

[46] Yogesh Kumar Singh, *Research Methodology*, published by New Age International (P) Limited., New Delhi, p.134, 2006.

[47] Lisa M. Given, et al., *Qualitative Research Methods*, Volume 1 & 2, published by SAGE Publications, Inc., California, p.30, 2008.

[48] Lucienne T.M. Blessing and Amaresh Chakrabarti, *DRM, a Design Research Methodology*, published by Springer-Verlag Limited, London, p.130, 2009.

[49] Auston Marmaduke Kilpatrick, *Lean Manufacturing Principles: A Comprehensive Framework for Improving Production Efficiency*, Master thesis, Massachusetts Institute of Technology, USA, p.09, 1997.

[50] Pande, P. and Holpp, L., *What is Six Sigma?*, published by McGraw-Hill Publishers Inc., New York, p.33, 2001.

[51] USEPA (United States Environmental Protection Agency), *The Lean and Environment Toolkit*, published by United States Environmental Protection Agency, USA, p.33, 2007.

[52] AISB (Artificial Intelligence and the Simulation of Behaviour Society), *Robotics, Mechatronics and Animatronics in the Creative and Entertainment Industries and Arts*, published by The Society for the Study of Artificial Intelligence and the Simulation of Behaviour, UK, p.38, 2005.

[53] Nikolay Avgoustinov, *Modelling in Mechanical Engineering and Mechatronics* published by Springer-Verlag London Limited, UK, p.18, 2007.

[54] Wolfgang Ernst and Stanislav Hosnedl, *Design Engineering: A manual for enhanced creativity*, published by CRC Press, Baca Raton, p.167, 2008.

[55] Janeri Jane and Lewis Jane, *Qualitative Research Practice*, published by SAGE Publications, London, p.274, 2003.

[56] Murray E. Jennex, *Knowledge Management: Concepts and Methodologies*, published by Information Science Reference, New York, p.60, 2008.

[57] Rouhani Sepi, *Research Innovation Technology*, published by Nelson Mandela Metropolitan University, South Africa, p.48, 2006.

[58] Mark Balnaves and Peter Caputi, *Introduction to quanitative research*, published by SAGE Publications, London, p.33, 2001.

[59] Brian S. Elliott, *Electromechanical Devices & Components Illustrated Sourcebook*, published by The McGraw-Hill Companies, New York, p.40, 2007.

[60] NIC (National Instruments Corporation), *LabVIEW Measurement and Automation Software*, published by National Instruments Corporation, USA, p.63, 2010.

AUTHORS' VITAE

Ramamoorthy.C., Research Scholar, Department of Mechanical Engineering, Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India. His research interest includes the integration of lean principles and mechatronics systems. He perused B.E. (Mechanical), M.E. (Advanced Manufacturing Technology) and M.B.A. (Production Management). He holds 10 years of professional experience including 7 years of industrial experiences, 3 years of teaching experiences and 3 years of research experience. He is the recipient of awards for his outstanding services, achievements and contributions, to include, best performance award from M/s. Trichy steel rolling mills limited, Tiruchirappalli, India, Cash award M/s. Autoprint machinery manufacturers private limited, Coimbatore, India and best performance award from M/s. Turboprint Machinery Manufacturers India (P) Ltd, Coimbatore, India. He published 2 research papers in various International Journals, presented 1 technical paper in a worldwide conference (Malaysia) and published 21 technical papers in various international / National conference proceedings in the area of manufacturing technology and mechatronics engineering.

Dr.V.Selladurai, Principal of Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India, his research interest includes the manufacturing systems, modeling and simulation, operation research, scheduling and optimization. He holds 28 years instituted by Coimbatore institute of technology, Coimbatore, best Journal Paper award from professional experience including 3 years of industrial experience and 30 years of teaching experience. He is the recipient of awards for his outstanding services, achievements and contributions to include, best teacher award in the academic year 1999-2000 Journal of Industrial Engineering, best session paper award in the National Conference on Challenges in Achieving Global Quality, best session paper Award from the International Conference on (EISCO 2005), best Paper Award from Twelth National Convention of Electrochemists (NCE-12). He received project funds at a sum of Rs.56.5 Lakhs from various funding agencies. He successfully guided 25 Ph.D research projects in various fields. He organized 5 International conferences, 3 National conferences and 10 International workshops in various fields. He published 95 International Journals, 55 National Journals, and more than 175 International / National conference proceedings. He is reviewer of 5 reputed International Journals.

Dr.Rajesh Ranganathan, Associate Professor, Department of Mechanical Engineering, Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India. His research interests include, supply chain management, rapid manufacturing, inventory reduction, lead-time reduction, product and process customization, strategic planning and multi criteria decision analysis. He holds work experience as a business consultant / knowledge transfer partnership associate working for oil and gas industry in Aberdeen, Scotland, U.K. He worked with National Health Service U.K. and with a leading health care industry in Europe in the area of strategic planning, technological forcasting, inventory reduciton, product customization, and manufacturing lead time reduction. He guides researchers in different disciplines. He published about 15 International and National journals.