

Variable Selection in Functional Genomics Using Genetic Algorithm-Based Feature Selection Method-An Empirical Study

¹V. Sujatha and ²Shaheda Akthar

¹Department of CSE, Acharya Nagarjuna University, Guntur, Vignan's Nirula Institute of Technology and science for Women, Pedapalikaluru, Guntur, AP, India

²Department of Computer Science, Government College for Women, Guntur, India

Abstract: Microarray information is an secondary dimensional dataset with a little example measure which holds not main pertinent as well as unimportant what's more excess genes. Gene interpretation profiling may be measuring those action about many genes simultaneously. Identikit right gene determination will be an significant errand to microarray information order. Genetic algorithms gets its ideas from biological world and the way genes interact to other genes to make new genes. In this research, Genetic algorithm is Applied on Lymphoblastic Leukaemia (ALL) data set. This fill in summarizes majority of the data signifying gene interpretation profile of five assemblies of patients (EMLLA, Hyp+50, MLL, T, also TEL), dispose of the The majority irrelevant genes, most extreme and least outflow esteem for each gene, furthermore, genes were positioned by least also maximam values, what's more assuming that they were inside the highest point 15% were chose to characteristic examination.

Key words: Microarray data, high dimensional data, Genetic algorithms, gene expression profiling, examination, highest, assuming

INTRODUCTION

A Genetic Algorithm (GA) is great for finding solutions to difficult problems and commonly used on data mining. Feature selection is a very important method which selects the useful features for used classification process as the fact that classifier performance is sensitive to the selection of the features used to build good-quality classifier from small or high dimension data that is naturally noisy (Dudoit and Fridlyand, 2003). The idea behind Genetic algorithms is given a pool of parents (genes) we select two and initiate a crossover between the parents generating two children which in turn can undergo mutation. The essential transform to an Genetic algorithm is.

Introduction: Frist make haphazardly an starting populace (Liu *et al.*, 2005) for fancied size, starting with best a couple people should thousands.

Assessment: Each part of the number assessed also wellness is ascertained to each unique. These prerequisites may be direct faster calculations (Alshamlan *et al.*, 2015; Ooi and Tan, 2003) would better, choice also selecting best people in the populace. Furthermore, discarding those awful outlines will move

forward our populaces general wellness. Populaces with a higher wellness will produce more posterity (Efron, 1983).

Hybrid: Joining together viewpoints about our chose people will generate certain qualities starting with two alternately that's only the tip of the iceberg people (Dudoit and Fridlyand, 2003) that will make best "fitter" posterity which will thus inherit best qualities structure their guardians.

Transformation: Mutations need aid presented under our populace haphazardly (Efron and Tibshirani, 1993). These mutations will prepare new mutated genes that need aid utilized within our number.

Repeatable: Those procedure may be rehashed starting with venture 2 until a exact number is acquired. The cycle determination also replication (Braga-Neto and Dougherty, 2004), crossover furthermore change will be known as era.

Literature review

Data set: This study utilization intense Lymphoblastic Leukemia (ALL) dataset which will be a large portion as a relatable point pediatric harm. The information situated

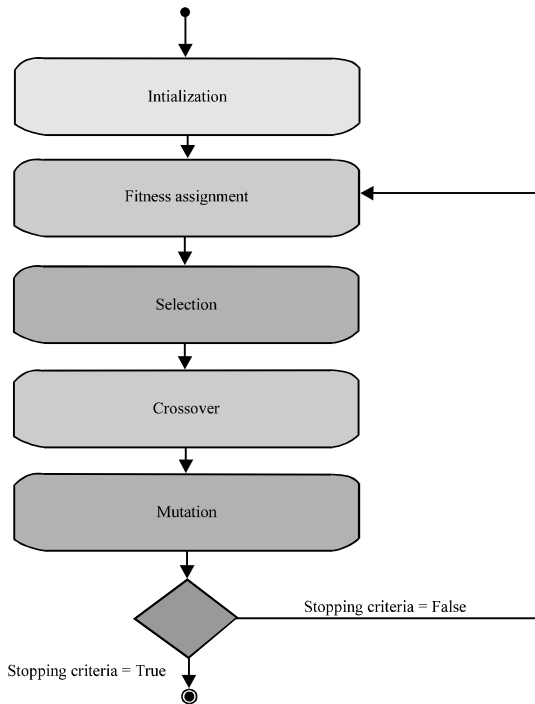


Fig. 1: Genetic algorithm schematic representation procedure

holds 327 tests of intense Lymphoblastic Leuctra (ALL) patients. There are a lot of people referred to hereditary predictive variables clinched alongside ALL which incorporate white blood cell ALL those hyperdiploid karyotype and the translocations: t(12;21)[TEL-AML1], t(4;11)[MLL-AF4], t(9;22)[BCR-ABL] and t(1;19)[E2A-PBX] speaking to 7 separate illness sub-classes which have been transformed on get outflow profiling information utilizing (Efron and Tibshirani, 1993; Ye, 2003) Affymetrix GeneChips. In this study, we summarize majority of the data signifying gene interpretation profile of five bunches of claiming patients (EMLLA, Hyp+50, MLL, T and TEL), kill those The majority irrelevant genes, most extreme what's more base statement quality for each gene furthermore, genes were positioned by base Furthermore maximam qualities (Fu, 1997), what's more assuming that they were inside the top banana 15% were chose for characteristic investigation. Those anlaysis may be completed for an arrangement from claiming phases. (Fig. 1).

MATERIALS AND METHODS

Phase 1; Test set up to the investigation: Make an object that will store 300 populaces (max solutions = 300) which will hold 5 genes (Population size = 5) that relate

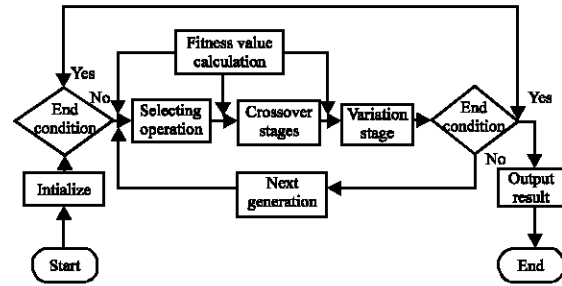


Fig. 2: Output for one Genetic algorithm cycle (200 generations)

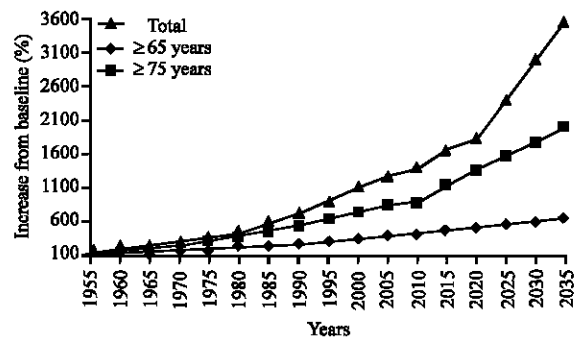


Fig. 3: Monitoring of accumulated population

should models created utilizing a close to separation centroid classifier (Lapointe *et al.*, 2004) (classification. Method = “neardistcentriod”) with a arrangement correctness about in any event 90% (goalFitness = 0. 9).

Phase 2; Evolving models/populaces: This will be a sample yield to particular case ga cycle (200 iterations): the yield indicates the emulating parameters (Braga-Neto *et al.*, 2004). No. From claiming evolutions: 300; No. About evolutions arrived at those objective fitness: 292; Value of the best chromosome: 0. 89103; Percentage relative (Li *et al.*, 2001) of the objective fitness: 99%; No. for generations required: 200; methodology period (Aakanksha *et al.*, 2012) went through previously, most recent evolution: 40 sec; Gathered procedure run through for all evolution: 4818 sec; Remaining period necessary to gather those formerly necessary to gather the formerly specified number about populations: 5748 sec (Fig. 2 and 3).

Figure 4 demonstrates real-time following of the Genetic algorithm the level hub of the, main what's more base plots visualizes unranked gene indexes. The verthandi hub of the highest point board is visualizes those populace list while those verthandi hub of the bottom board is visualizes those era number. In the center

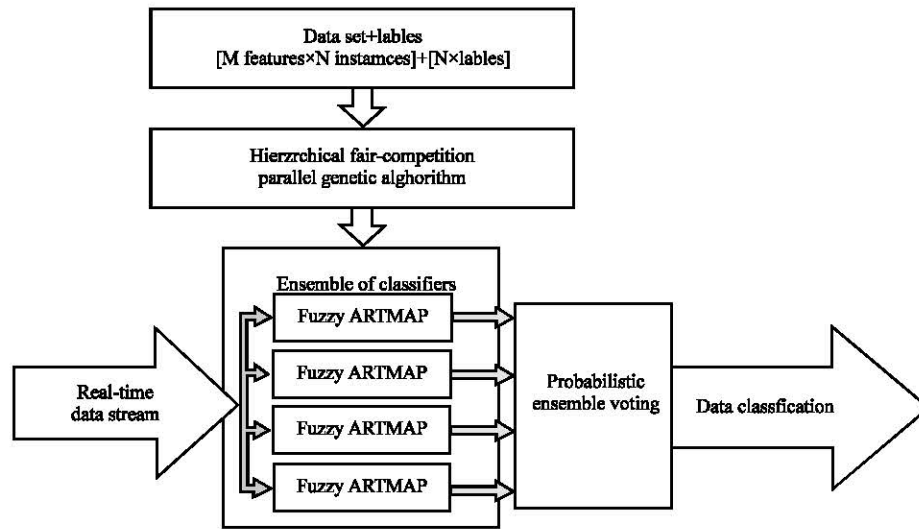


Fig. 4: Real-time monitoring of the Genetic algorithm

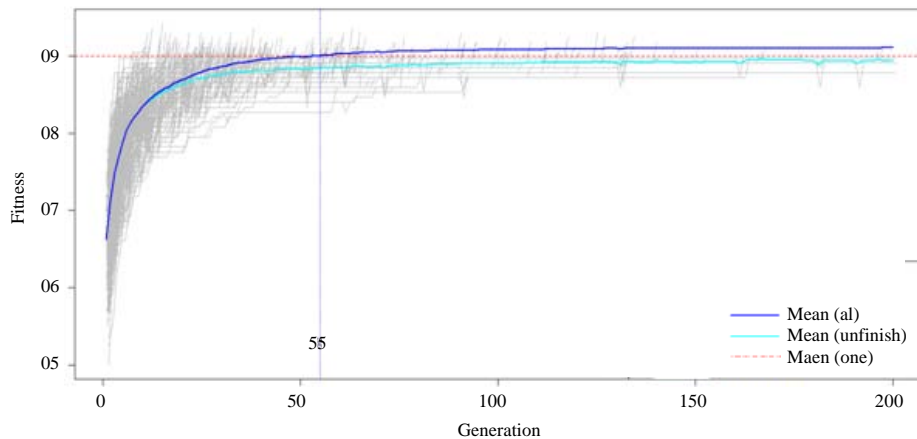


Fig. 5: Evolution of the maximum fitness across generations in 300 populations

plot the level hub (Peng *et al.*, 2003) will be visualizes those era in as much as those verthandi hub will be showing the wellness esteem.

Phase 3; Determiation and refinement for populaces (Chromosomes): Figure 5 provides for previously an average, we need aid getting will an answer previously, era 56 which may be thick as normal. Those lines demonstrates normal wellness to every last one of chromosomes which arrived at those objective furthermore likewise not arrived at objective. These accordance define a useful “confidence interval” for those wellness every last one of generations (Yeoh *et al.*, 2002). Trademark plateau stamped for dabbed line in Fig. 5 may be advantageous should concluded that look is not attempting should compass our objective. We could likewise plot separate those evolutions that bring arrived

Table 1: Overall classification accuracy

Measures	NBC (%)	SVM (%)	k-NN(%)
Acuurac	96.25	93.75	95
Specification	92.50	90	90
Sensitivity	100	97.50	100
Percision	93.2	90.70	90.91

at those objective which will be demonstrated in Fig. 6. Generally exactness of the number to chosen models (Braga-Neto and Dougherty, 2004) is plotted in Table 1 the level hub speaks to each unique example assembled (Chuchra, 2012) In light of their sickness classes while those verthandi hub speaks to the predicted classes (Lapointe *et al.*, 2004; Jain and Srivastava, 2013). Those barplots symbolize the % of the models that arrange every test clinched alongside a class. To example, specimens to second section (marked done red) have a place with the HYP+50 population. For average, effectively ordered 87%

of the times and likewise they need aid “wrongly” arranged 2% of the times Similarly as EMLLA, 5% of the times concerning illustration MLL, 1.5% as T and 4.7% as TEL.

Phase 4; With find climate rank of the genes stable or not: To discover that rank of the genes stable or not we investigate our calculation indicated clinched alongside (Fig. 7) stochastic searches which will be utilized within

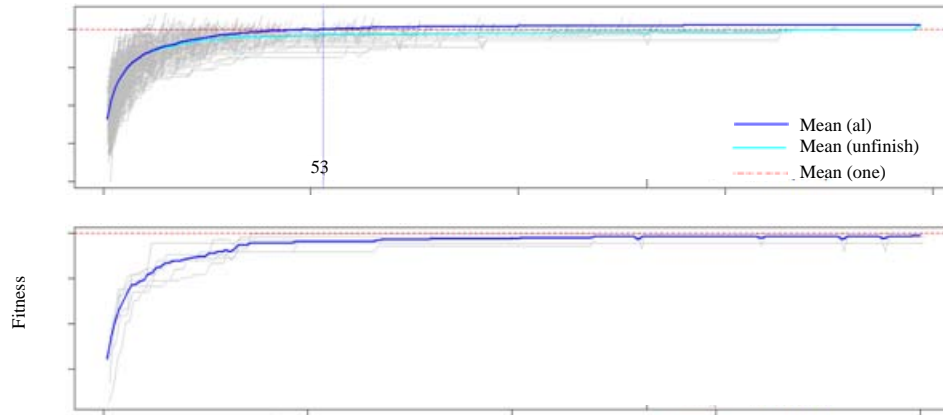


Fig. 6: Separate the evolutions that have reached the goal maximum fitness across generations in 300 Populations: a) Fitness 294 (solutions/chromosomes) [project]: nearcent-mean-0, 1-3 k folds and b) Fitness 6 (Non-solutions/chromosomes) [project]: nearcent-mean-0, 1-3 k folds

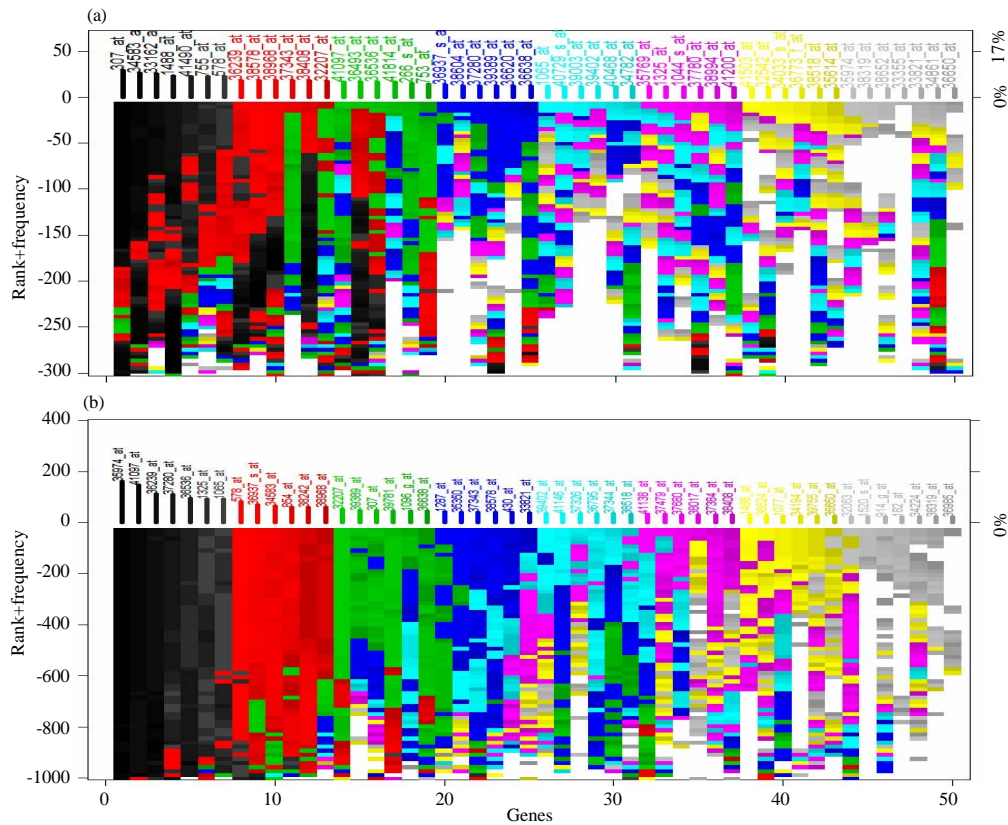


Fig. 7: a) Gene ranks across past evolutions and b) Rank stability in 1000 chromosomes

Genetic algorithm (Han and Kamber, 2001) would useful skilled techniques to recognizing answers for a streamlining issues. Those beginning stage for at whatever Genetic algorithm (Bharadwaj and Pal, 2011) may be should scan clinched alongside an irregular number. Our approach is should figure the recurrence that each gene shows up in the chromosome populace. Consequently observing the strength about gene ranks (based for their frequency) offers the likelihood will visualize model joining.

Over Fig. 7a the majority incessant 50 genes would demonstrated over 8 separate colors with regarding

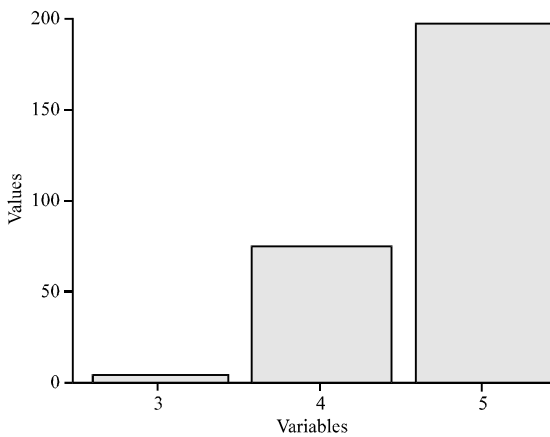


Fig. 8: Refinement of the chromosomes

6 alternately 7 genes for every color. Level hub in indicates the genes requested eventually Tom's perusing rank (Chuchra, 2012). Verthandi hub indicates the gene recurrence and the shade coded rank about each gene for past evolutions. Starting with Fig. 7 it will be watched that 7 dark genes need been stable no less than throughout the most recent 50 results while a few red genes need as of late swap from green. Consequently, we might reason that red furthermore green genes need aid not yet stable in light 300 chromosomes would not sufficient on settle these genes. For this 1000 chromosomes need aid recreated and outcomes are produced which indicates clinched alongside Fig. 7b which show all the more strength clinched alongside ranks. An additional property may be that Main genes need aid continuously settled to order; 1st bootleg genes, after that red, green etcetera.

Stage 6; Discovering genes incorporated in a chromosome that are helping of the model accurac:

Figure 8 demonstrates that an enormous measure of the chromosomes oblige all five genes arrange the tests faultlessly. By and large we manufacture models with additional genes over the classes infact larger part of the datasets really hold numerous just two classes (e.g., Treated-untreated, cancer-normal, wild-mutant, etc).

Phase 7; Creating delegate test models: Figure 9 indicates forward determination methodology by

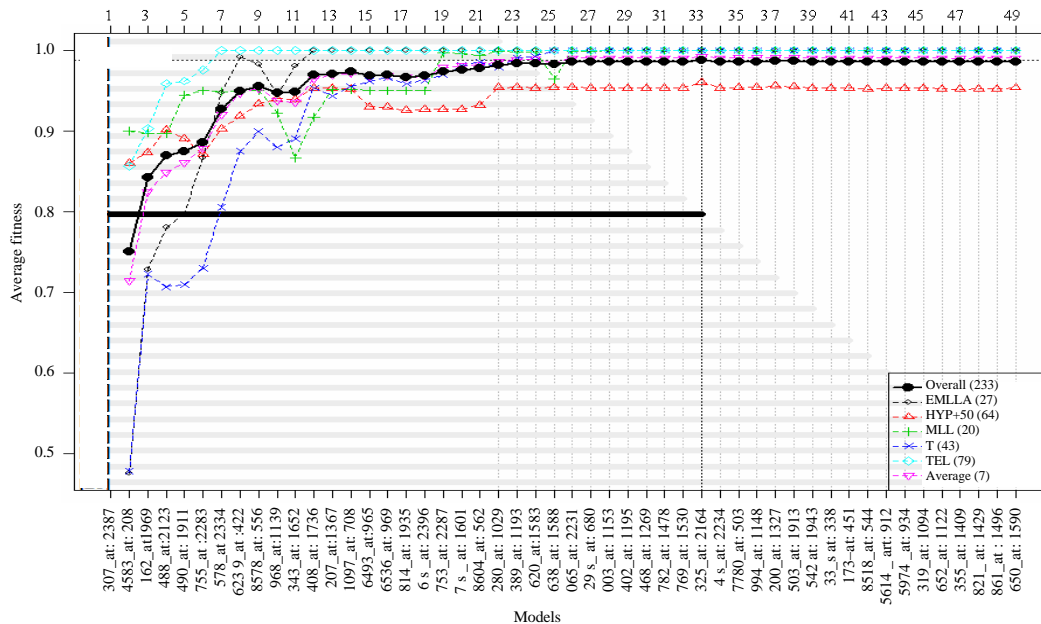
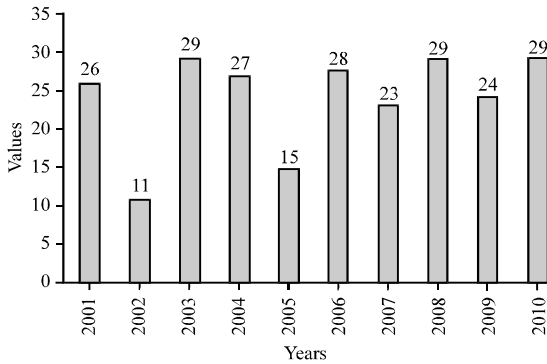


Fig. 9: Forward selection model: (Alshamlan *et al.*, 2015; Braga-Neto and Dougherty, 2004; Dudoit and Fridlyand, 2003; Efron, 1983; Efron and Tibshirani, 1993; Lapointe *et al.*, 2004; Li *et al.*, 2001; Liu *et al.*, 2005; Ooi and Tan, 2003; Peng *et al.*, 2003; Yeoh *et al.*, 2002; Han and Kamber, 2001; Bharadwaj and Pal, 2011; Jain and Srivastava, 2003; Dunham, 2006; Umamaheswari and Niraimathi, 2013)



assessing the test slip utilizing the wellness work altogether test sets. Level hub speaks to those genes requested by their rank. Verthandi hub indicates the order exactness. Strong transport speaks to those generally speaking exactness. Colored dashed lines representable those correctness for every population. Those effects indicates models which would higher over 99% of the most extreme out from claiming 29 best models. Model named as 12, holding the The majority 33 incessant genes. The opposite 28 models incorporated on fsm are 99% as near those best model (Fig. 10-13).

Fig. 10: From a model resulted from forward selection and an original evolved chromosome

Phase 8: Foreseeing population enrollment (Dunham, 2006) from claiming obscure specimens. Sham information

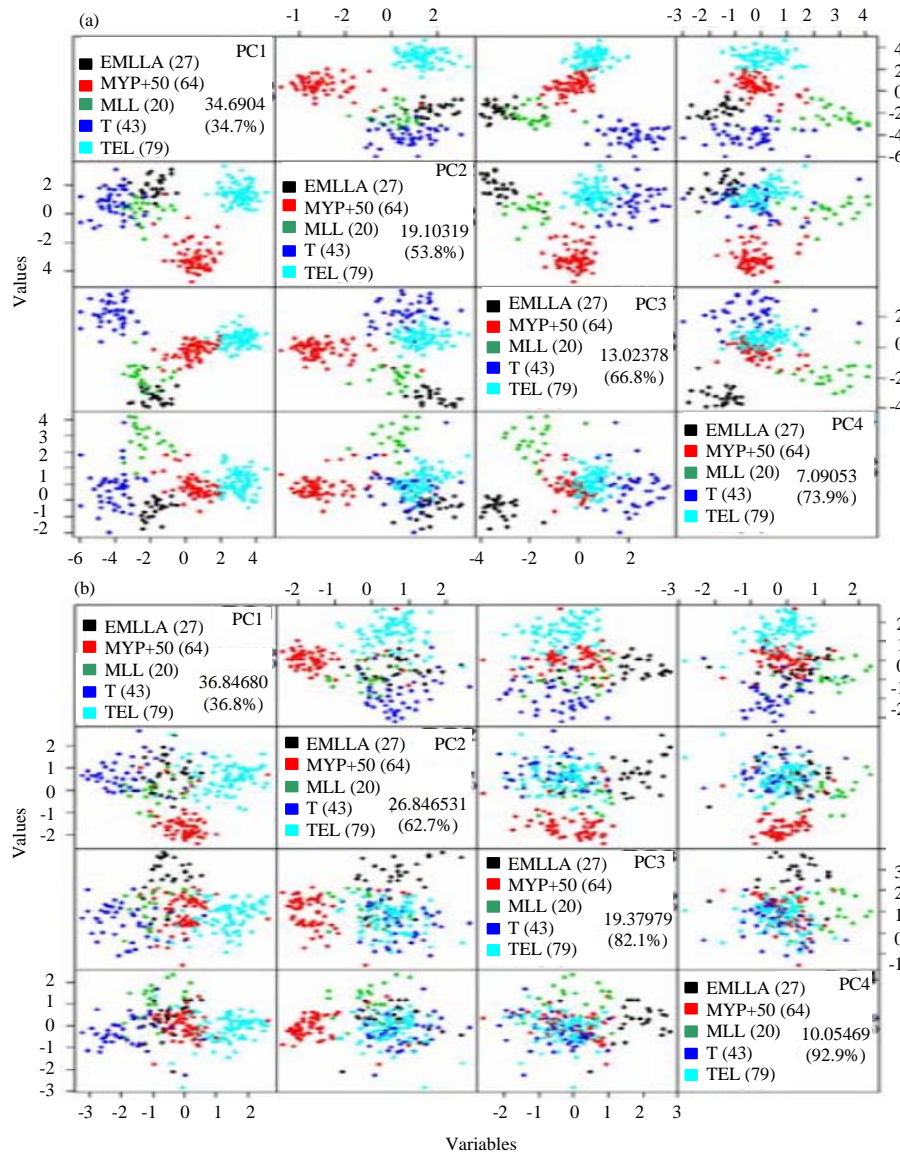


Fig. 11: Depiction of a model (left) and a chromosome (right) in PCA space

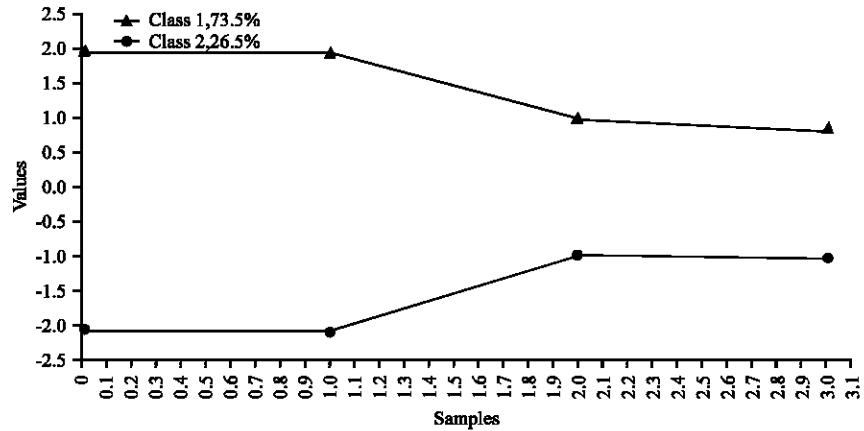


Fig. 12: Sample profiles per class

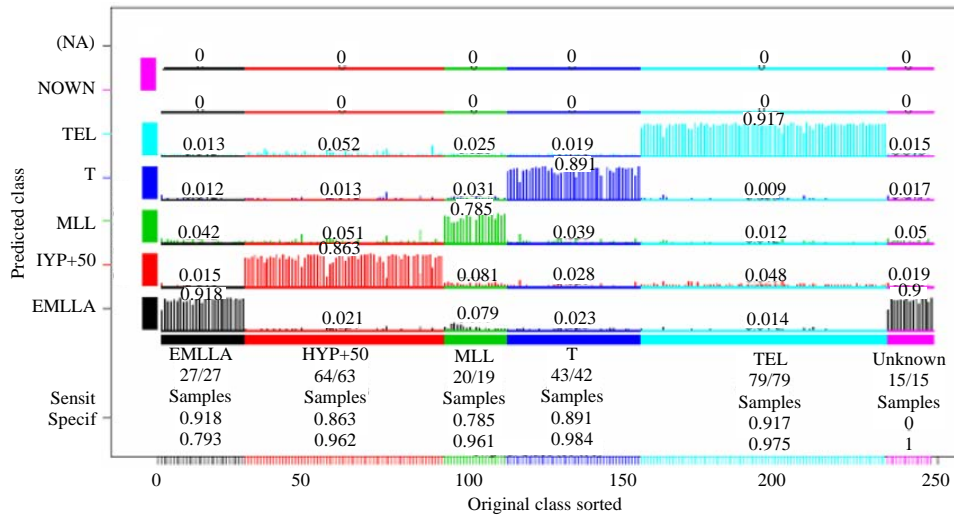


Fig. 13: Prediction for unknown samples (the last 15 samples in the right)

might have been temporally appended of the first information. That point class prediction matrix will run for every last one of number which results of the plot hint at on Fig. 13 parts is an parameter utilized within class predictionmatrix (Umamaheswari and Niraimathi, 2013). The bring about shortages of the plot may be demonstrated the place the new information might have been named likewise “UNKNOWN. Those bootleg bars for these specimens demonstrate that they were predicted likewise EMLLA (the final one 15 specimens in the right).

RESULTS AND DISCUSSION

Those point of this dissection may be with recognize gene sets that are predictive of malady sort for a board about leucra patients. In this worth, of effort we mimicked how on raise Genetic algorithms to an order issue. Genetic

algorithm with closest classifier may be connected on main positioning genes. Genetic algorithm selects practically educational genes and closest classifier arrangement exactness may be acknowledged concerning illustration the wellness capacity. This straightforward model in light of different factual measures what’s more Genetic algorithm figures A large portion applicable biomarkers to leucra patients classes.

CONCLUSION

Our technique is helpful when a rearing system can just phenotype a subset of the accessible genotyped people, yet goes for assessing the reproducing estimation of a (conceivably significantly bigger) gathering of genotypes. Genomic determination permits to assess the rearing estimation of plants or creatures utilizing

genotypic and phenotypic data from a preparation populace. By supplanting arbitrary examining with our advanced determination conspire while choosing the preparation set, the rearing esteems in the test set can be evaluated with higher exactnesses. In the event that the applicant and the test sets are both arbitrarily chosen from a similar populace, choosing an upgraded preparing test from the competitors with our technique enhances the correctnesses of GEBV for this populace. In any case, the utilization of our technique is additionally restricted, since, it requires that every one of the genotypes are known ahead of time and that the people that are chosen in the preparation set are accessible for phenotyping. In this study variable choice in quality articulation informational indexes by utilizing hereditary calculations has been researched. Not with standing it which just pick principle impacts we have additionally presented a technique for determination of interactions. It has been demonstrated that subsequent subset yields order methods that perform extremely well.

REFERENCES

- Aakanksha, B., P.J. Shweta and M.N. Madan, 2012. Data mining techniques and distinct applications: A literature review. *Intl. J. Eng. Res. Technol.*, Vol. 1, 1-10.
- Alshamlan, H.M., G.H. Badr and Y.A. Alohal, 2015. Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer classification. *Comput. Biol. Chem.*, 56: 49-60.
- Bharadwaj, B.K. and S. Pal, 2011. Mining educational data to analyze student's performance. *Int. J. Adv. Comput. Sci. Applic.*, 2: 63-69.
- Braga-Neto, U. and E. Dougherty, 2004. Is cross-validation valid for small sample microarray classification? *Bioinformatics*, 20: 374-380.
- Chuchra, R., 2012. Use of data mining techniques for the evaluation of student performance: A case study. *Intl. J. Comput. Sci. Manage. Res.*, 1: 425-433.
- Dudoit, S. and J. Fridlyand, 2003. Classification in Microarray Experiments. In: *Statistical Analysis of Gene Expression Microarray Data*, Speed T. (Ed.). Chapman and Hall/CRC, Boca Raton, Florida, pp: 93-158.
- Dunham, M.H., 2006. *Data Mining, Introductory and Advanced Topics*. Pearson Education, Delhi, India, ISBN:98-81-7758-785-2, Pages: 301.
- Efron, B. and R.J. Tibshirani, 1993. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, New York, USA., ISBN: 0412042312.
- Efron, B., 1983. Estimating the error rate of a prediction rule: Improvement on cross-validation. *J. Am. Stat. Assoc.*, 78: 316-331.
- Fu, Y., 1997. Data mining. *IEEE. Potentials*, 16: 18-20.
- Han, J. and M. Kamber, 2001. *Data Mining, Concepts and Techniques*. Morgan Kaufmann, Burlington, Massachusetts, ISBN:9787040100419, Pages: 550.
- Jain, N. and V. Srivastava, 2013. Data mining techniques: A survey paper. *Intl. J. Res. Eng. Technol.*, 2: 1163-2319.
- Lapointe, J., C. Li, J.P. Higgins, V.D.M. Rijn and E. Bair *et al.*, 2004. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc. National Acad. Sci. U.S. Am.*, 101: 811-816.
- Li, L., T.A. Darden, C.R. Weingberg, A.J. Levine and L.G. Pedersen, 2001. Gene assessment and sample classification for gene expression data using a genetic algorithm-k-nearest neighbor method. *Comb. Chem.. High Throughput Screening*, 4: 727-739.
- Liu, J.J., G. Cutler, W. Li, Z. Pan and S. Peng *et al.*, 2005. Multiclass cancer classification and biomarker discovery using GA-based algorithms. *Bioinf.*, 21: 2691-2697.
- Ooi, C.H. and P. Tan, 2003. Genetic algorithms applied to multi-class prediction for the analysis of gene expression data. *Bioinf.*, 19: 37-44.
- Peng, S., Q. Xu, X.B. Ling, X. Peng and W. Du *et al.*, 2003. Molecular classification of cancer types from microarray data using the combination of genetic algorithms and support vector machines. *FEBS. Lett.*, 555: 358-362.
- Umamaheswari, K. and S. Niraimathi, 2013. A study on student data analysis using data mining techniques. *Intl. J. Adv. Res. Comput. Sci. Software Eng.*, 3: 117-120.
- Ye, N., 2003. *The Handbook of Data Mining*. Taylor & Francis, Mahwah, New Jersey, ISBN:9780805855630, Pages: 720.
- Yeoh, E.J., M.E. Ross, S.A. Shurtleff, W.K. Williams and D. Patel *et al.*, 2002. Classification, subtype discovery and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*, 1: 133-143.

Precise Long-Term Prediction of Behavior in a 3-D Chaotic Map

M. Mammeri

Department of Mathematics, University of Kasdi Merbah, Ouargla, Algeria

Abstract: In this study, the prediction accuracy on long-term of future behavior in a 3-D chaotic map is possible. This study introduces and justifies by numerical simulation a new phenomenon shown by the 3-D map which is the behavior in the 3-D map is always regular and periodic and bounded on large time, i.e. that the behavior repeats itself regularly after cycles on large time intervals. The new 3-D map produces new several chaotic attractors obtained via quasi-periodic route to chaos.

Key words: Precise prediction, 3-D map, future behavior, symmetric intervals, regular behavior, time intervals

INTRODUCTION

Since, the discovery of the first famous three dimensional Lorenz chaotic attractor (Lorenz, 1963), chaos as a most fascinating phenomenon in non-linear dynamical systems, in recent years has been a considerable increase in their study (Rossler, 1976; Chua *et al.*, 1968; Lu and Chen, 2002; Hu *et al.*, 2012; Mandal *et al.*, 2013, 2011; Mammeri, 2016; Gonchenko *et al.*, 2005, 2007; Gonchenko and Meiss, 2006). The correct prediction on long-term of the behavior of solutions in non-linear dynamical systems is interesting and important in understanding of evolution of behavior of chaotic systems. It is well known the chaotic behavior to be strongly dependent on initial conditions, small changes in initial conditions can possibly lead to immense changes in subsequent on large time intervals and is what so ever difficult to precisely predict the future in next few decades. For example, in life sciences an interesting still open problem is to predict human future behavior from the actions they took in the past in other word it is possible to use someone past actions to predict his future behavior? Many other researchers have investigate the prediction of future behavior see for instance the research done by Lehnertz and Elger (1998), Romanelli *et al.* (1988) Jeong (2002), Arrow *et al.* (2008), Boettiger and Hastings (2013) and Dauwels *et al.* (2009). In this study, the accurate prediction of future behavior in 3-D chaotic discrete system on large time intervals is possible. This short study reports and investigate the effect of the sine map in 3-D discrete systems, the modified 3-D discrete map (Eq. 1) obtained via direct modification of the 3-D discrete system proposed by Dullin and Meiss (2000). The map (Eq. 1) is defined with two sine nonlinearities topologically different from any other know 3-D systems. In this study, we have studied some impotents basic

properties for a new 3-D map and introduces and justifies numerically a new physical phenomenon shown by the new 3-D map (Eq. 1) which is the behavior of the map (Eq. 1) is always regular and periodic, i.e. that the behavior repeats itself regularly after cycles (periods) on large time symmetry intervals. Furthermore the behaviors observed for the map 1 are bounded and symmetric about the origin. Considered essentially the following modified 3-D map Eq. 1:

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \\ z_{t+1} \end{pmatrix} = \begin{pmatrix} y_t \\ \sin z_t \\ a+bx_t+cy_t-\sin z_t \end{pmatrix} \quad (1)$$

where, $(a-c) \in \mathbb{R}^3$ are bifurcation parameters and $(x_t, y_t, z_t) \in \mathbb{R}^3$ are the state variables. The choice of sinusoidal map has an important role is to guarantee the boundedness of the orbits of the map (Eq. 1) for all values of a-c. Generally, the map (Eq. 1) is not symmetric and the associated map of the new 3-D map (Eq. 1) is continuous and differentiable on \mathbb{R}^3 . Furthermore the Jacobian matrix of the map (Eq. 1) is not constant and equal to $b \cos z$ ($b \neq 0$ and $\cos z \neq 0$). The system (Eq. 1) can be transform into a third-order difference Eq. 2 as follows:

$$z_{t+1} = a+bsin z_{t+2}+csin z_{t+1}-sin z_t \quad (2)$$

MATERIALS AND METHODS

Qualitative properties of the map: In this study, we will show that the all orbits of the map (Eq. 1) are bounded and are lies inside in a box and we investigate domains for the bifurcation parameters $(a, b, c) \in \mathbb{R}^3$ in which the fixed points of the map (Eq. 1) are asymptotically stable.

Theorem 1: The all orbits of the map (Eq. 1) are bounded for every parameters $(a-c) \in \mathbb{R}^3$ and $t > 2$ and for all finite initial conditions (x_0, y_0, z_0) .

Proof: We use the following standard results, the real sequence $(z_n)_n$ is bounded if there is one positive real such that $|z_n| \leq k$ for every $n \in \mathbb{N}$. In our case the sequence (z_t) , given in satisfies the following inequality, $|z_t| \leq 1 + |a| + |b| + |c|$ because $|\sin z| \leq 1$ for every $z \in \mathbb{R}$. Since, the real $1 + |a| + |b| + |c|$ is positive, thus, the sequence (z_t) , is bounded for every $(a, b, c) \in \mathbb{R}^3$ and $t > 2$. Thus, implies the all orbits of the map (Eq. 1) are bounded for every $(a, b, c) \in \mathbb{R}^3$, $t > 2$ and for all finite initial conditions $(x_0, y_0, z_0) \in \mathbb{R}^3$.

Conclusion 1: The all orbits of the map (Eq. 1) are lies inside in the box:

$$\{(x, y, z) \in \mathbb{R}^3 : |x| \leq 1, |y| \leq 1, |z| \leq 1 + |a| + |b| + |c|\} \quad (3)$$

Proof: It's very easy to prove this conclusion, since, the map (Eq. 1) is equivalent to:

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \\ z_{t+1} \end{pmatrix} = \begin{pmatrix} \sin z_t \\ \sin z_t \\ a + b \sin z_{t-2} + c \sin z_{t-1} - \sin z_t \end{pmatrix} \quad (4)$$

Theorem: The fixed point $A(x, y, z)$ of the map (Eq. 1) is asymptotically stable for all $a \in \mathbb{R}$ and if and only if $(b, c) \in \cup_{i=1}^3 \Omega_i$ where:

$$\Omega_1 : \begin{cases} -1 < b < 1 \\ \frac{b(b+1)\cos z^2 - 1}{\cos z} < c < \frac{1 - (1-b)\cos z}{\cos z} \end{cases} \quad (5)$$

$$\Omega_2 : \begin{cases} -1 < b < 1 \\ \frac{1 + (1-b)\cos z}{\cos z} < c < \frac{b(b+1)\cos z^2 - 1}{\cos z} \end{cases} \quad (6)$$

Proof: The characteristic polynomial of the Jacobian matrix of the map (Eq. 1) calculated at the fixed point $A(x, y, z)$ which takes the form, $P_A(\lambda) = \lambda^3$, according to the result available in Ogata (1995), we conclude that the fixed point A of the map (Eq. 1) is asymptotically stable if the following conditions hold $|b \cos z| < 1$, $1 + \cos z - c \cos z - b \cos z > 0$, $1 - \cos z - c \cos z + b \cos z - c \cos z > 0$ and $1 - b^2 \cos^2 z > b \cos^2 z - c \cos z - c \cos z$. From Eq. 1, we have Eq. 5 $|b| < 1$ and from Eq. 2, Eq. 3 and Eq. 5, we have Eq. 6 $c \cos z < 1 - (1-b) |\cos z|$, from Eq. 4 we have Eq. 7 $c \cos z > b$

$(b+1) \cos^2 z - 1$ and from Eq. 6 and 7, we get Eq. 8 $b(b+1) \cos^2 z - 1 < c \cos z < 1 - (1-b) |\cos z|$. Finally, the conditions Eq. 1 and 8 give the all conditions Eq. 5 and 4 of asymptotic stability for the fixed point A . For example, the fixed points of the map (Eq. 1) are the real solutions of the system:

$$x = y, y = \sin z, z = a + bx + cy - \sin z$$

Hence, one may easily obtain Eq. 7:

$$z - (b+c) \sin z + \sin z - a = 0$$

Can't be compute the fixed points of the map (Eq. 1) analytically, we remark if $a = 0$, the point $(0, 0, 0)$ it is fixed point of the map (Eq. 1) for all values of the bifurcation parameters $(b, c) \in \mathbb{R}^2$. Then we have the following theorem:

Theorem 3: If $a = 0$, the fixed point $(0, 0, 0)$ of the map (Eq. 1) is asymptotically stable if and only if the following conditions holds:

$$\begin{cases} -1 < b < 1 \\ b(b+1) - 1 < c < b \end{cases} \quad (7)$$

If we choose $a = 0$, $b = 0.8$ and $c = 0.1$. Then with this values the fixed point $(0, 0, 0)$ is asymptotically stable and we have the following three eigenvalues $\lambda_1 = -0.8566 - 0.6229i$, $\lambda_2 = -0.8566 + 0.6229i$ and $\lambda_3 = -0.3380$, thus $|\lambda_{1,2,3}| < 1$.

Bifurcation analysis of the map: In this study, we will illustrate some observed chaotic behaviors, the dynamical behaviors of the map (Eq. 1) are investigated numerically. Figure 1-6 show the bifurcation diagram and the diagram of the variation of the largest Lyapunov exponent of the map (Eq. 1) that are obtained at different values of parameter a , $a \in [-4, 4]$. However, we deduce from the bifurcation diagram Fig. 1a that the proposed map (Eq. 1) exhibit a quasi-periodic bifurcation scenario route to chaos for the selected values of the bifurcation parameter a .

First, we fix the initial condition $x_0 = y_0 = z_0 = 0.01$ and $b = 0.8$, $c = 0.9$ and let the parameter a vary in the interval $[-4, 4]$, the map (Eq. 1) exhibits the following dynamical behaviors as shown in Fig. 4a and b). For the quasi-periodic with periodic windows at the point $a = -2.72$ the dynamical behavior the map (Eq. 1) is in the

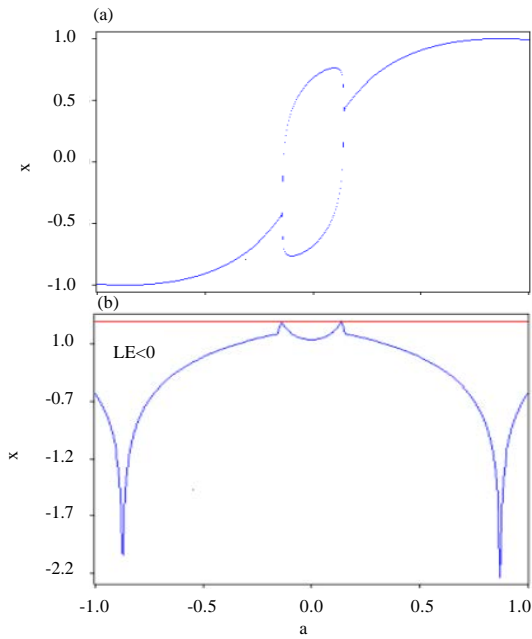


Fig. 1: a) The symmetric bifurcation diagram for the map (Eq. 1) obtained for $b = 0.8$, $c = 0.9$ and $-1 \leq a \leq 1$ and b) Variation symmetry of Lyapunov, exponent of the map (Eq. 1) for $b = 0.8$, $c = 0.9$ and $-1 \leq a \leq 1$

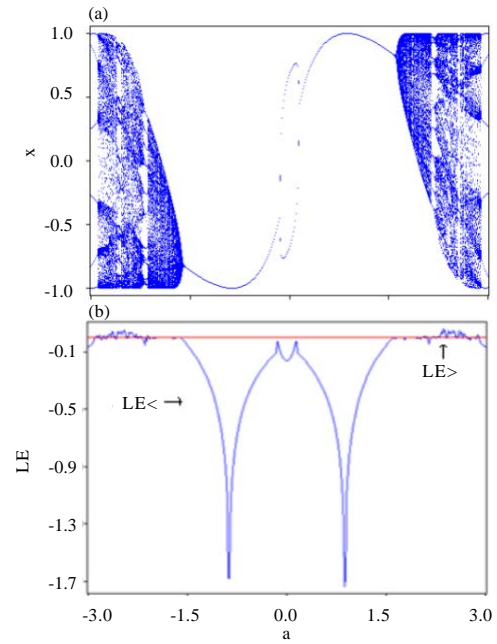


Fig. 3: a) The symmetric bifurcation diagram for the map (Eq. 1) obtained for $b = 0.8$, $c = 0.9$ and $-3 \leq a \leq 3$ and b) Variation symmetry of Lyapunov exponent of the map (Eq. 1) for $b = 0.8$, $c = 0.9$ and $-3 \leq a \leq 3$

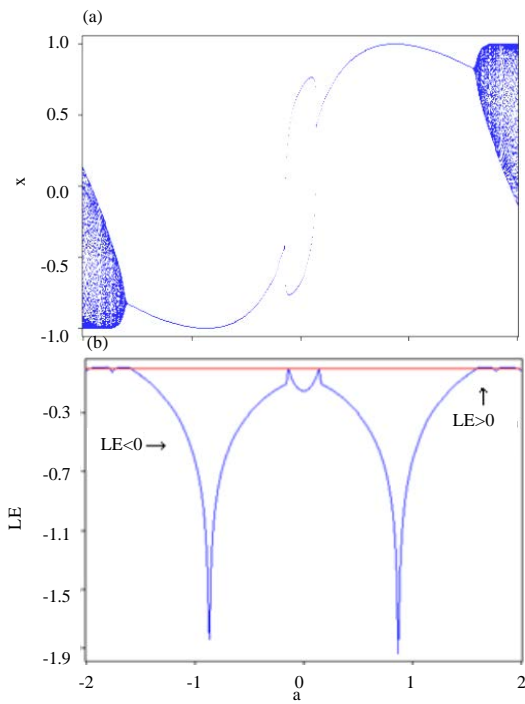


Fig. 2: a) The symmetric bifurcation diagram for the map (Eq. 1) obtained for $b = 0.8$, $c = 0.9$ and $-2 \leq a \leq 2$; and b) Variation symmetry of Lyapunov exponent of the map (Eq. 1) for $b = 0.8$, $c = 0.9$ and $-2 \leq a \leq 2$

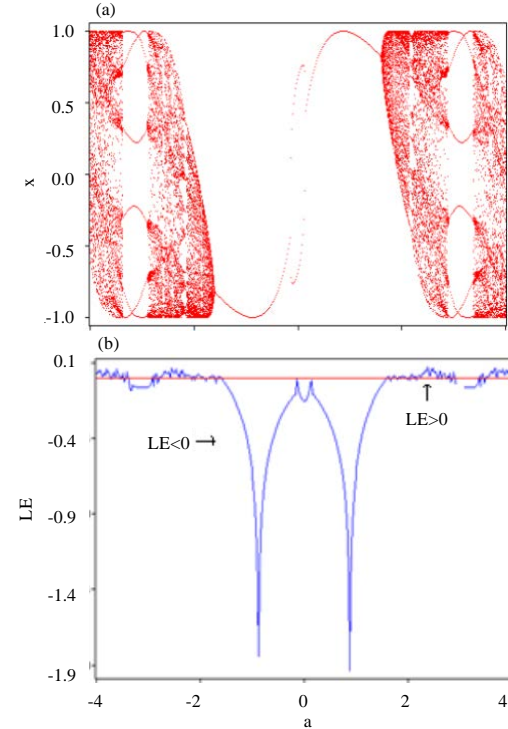


Fig. 4: a) The symmetric bifurcation diagram for the map (Eq. 1) obtained for $b = 0.8$, $c = 0.9$ and $-4 \leq a \leq 4$ and b) Variation symmetry of Lyapunov exponent of the map (Eq. 1) for $b = 0.8$, $c = 0.9$ and $-4 \leq a \leq 4$

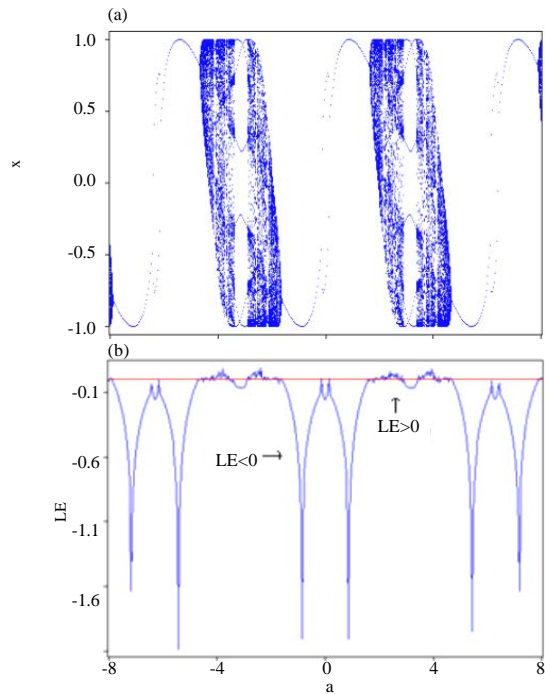


Fig. 5: a) The symmetric bifurcation diagram for the map (Eq. 1) obtained for $b = 0.8$ $c = 0.9$ and $-8 \leq a \leq 8$ and b) Variation symmetry of Lyapunov exponent of the map (Eq. 1) for $b = 0.8$, $c = 0.9$ and $-8 \leq a \leq 8$

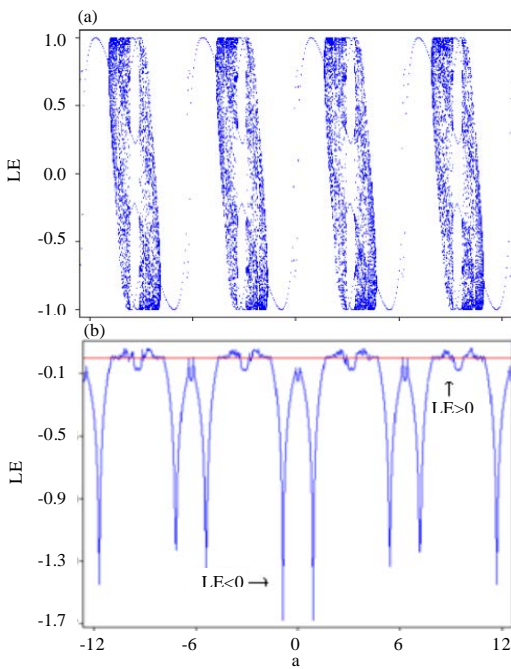


Fig. 6: a) The symmetric bifurcation diagram for the map (Eq. 1) obtained for $b = 0.8$ $c = 0.9$ and $-12 \leq a \leq 12$ and b) Variation symmetry of Lyapunov exponent of the map (Eq. 1) for $b = 0.8$, $c = 0.9$ and $-12 \leq a \leq 12$

quasi-periodic-40 as shown in Fig. 7d and at the point $a = -3.52$ the behavior is quasi-periodic as shown in Fig. 7c, for the range $-4 \leq a \leq -2.72$ the behavior of the map (Eq. 1) is chaotic with periodic windows in the chaotic band which is verified by the corresponding largest Lyapunov exponent is positive, Fig. 7a and b shows, respectively the chaotic behaviors of the map (Eq. 1) when $a = -3.76$ and $a = -3.6$ (Fig. 1-3).

Secondly, for the range $0 \leq a < 1.68$ the dynamical behavior of the map (Eq. 1) is periodic which is verified by the corresponding largest Lyapunov exponent is negative as shown in Fig. 4b for the range $1.68 \leq a < 2.32$ the dynamical behavior of the map (Eq. 1) is quasi-periodic orbits with periodic windows at the point $a = 2.24$ the dynamical behavior of the map (Eq. 1) is in the quasi-periodic-19 as shown in Fig. 8e and f shows the quasi-periodic attractor of the map (Eq. 1) when $a = 1.84$ for the range $2.32 \leq a < 2.72$ the dynamical behavior of the map (Eq. 1) is chaotic with periodic windows in the chaotic band which is verified by the corresponding largest Lyapunov exponent is positive. For the range $2.72 \leq a < 3.6$ the dynamical behavior of the map (Eq. 1) is quasi-periodic with periodic windows at the point $a = 3.52$ the dynamical behavior of the map (Eq. 1) is in range $-1.68 < a \leq 0$ the dynamical behavior of the map (Eq. 1) is periodic which is verified by the corresponding largest Lyapunov exponent is negative as shown in Fig. 4b for the range $-2.32 < a \leq -1.68$ the dynamical behavior the map (Eq. 1) is quasi-periodic with periodic windows at the point $a = -2.24$ the dynamical behavior the map (Eq. 1) is in the quasi-periodic-19 attractor as shown in Fig. 7e and f shows the quasi-periodic attractor of the map (Eq. 1) when $a = -1.84$ for the range $-2.72 < a \leq -2.32$ the dynamical behavior of the map (Eq. 1) is chaotic with periodic windows in the chaotic band which is verified by the corresponding largest Lyapunov exponent is positive. For the range $-3.6 < a \leq -2.72$ the dynamical behavior of the map (Eq. 1) is the quasi-periodic-40 as shown in Fig. 8d and at the point $a = 3.52$ the dynamical behavior is quasi-periodic as shown in Fig. 8c for the range $3.6 \leq a \leq 4$ the dynamical behavior of the system (Eq. 1) is chaotic with periodic windows in the chaotic band which is verified by the largest Lyapunov exponent is positive, Fig. 8a and b, respectively show the dynamical behavior of the map (Eq. 1) when $a = 3.76$ and $a = 3.6$.

Precise long-term prediction of behavior: Dynamical systems theory is an area of mathematics used to describe

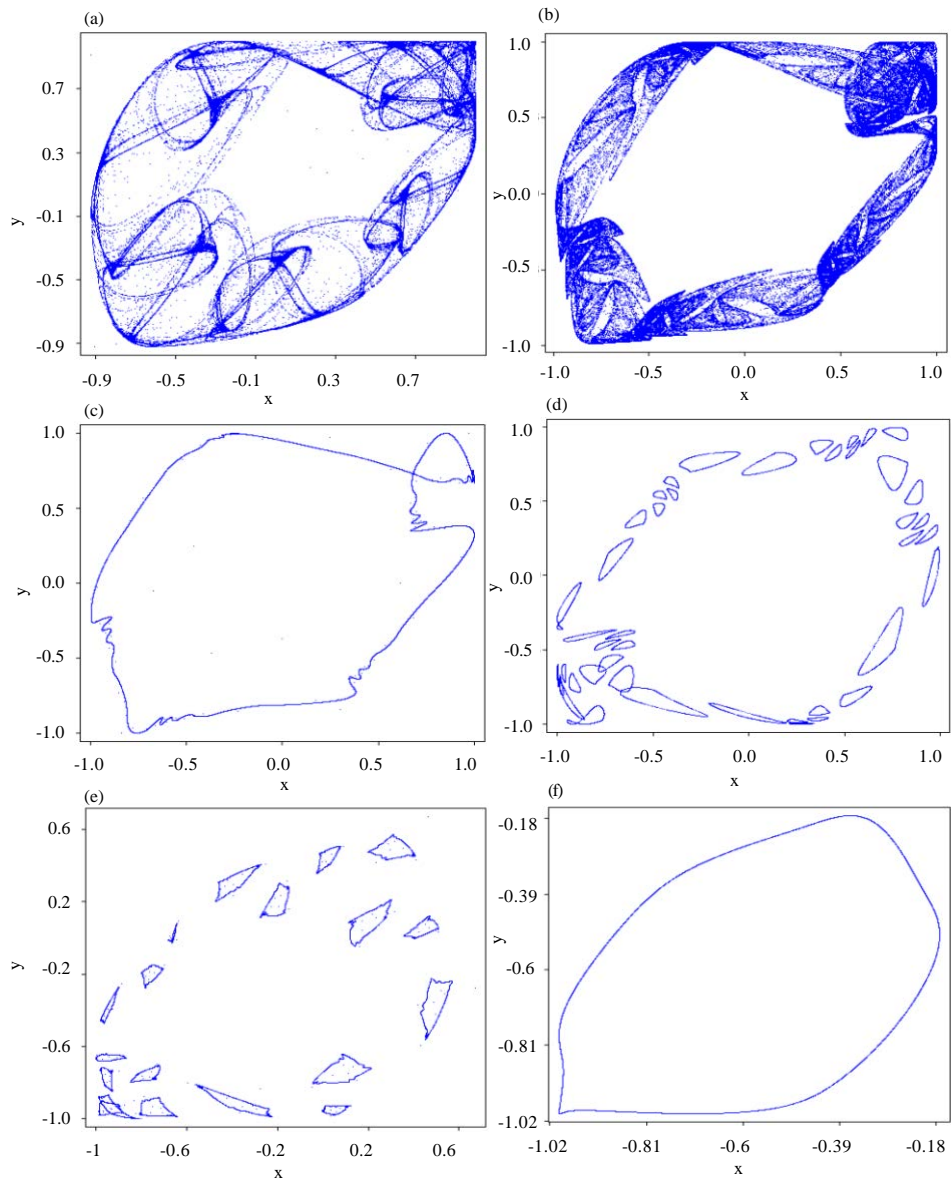


Fig. 7: a) $a = -3.76$; b) $a = -3.76$; c) $a = -3.52$; d) $a = -2.72$; e) $a = -2.24$ and f) $a = -1.84$

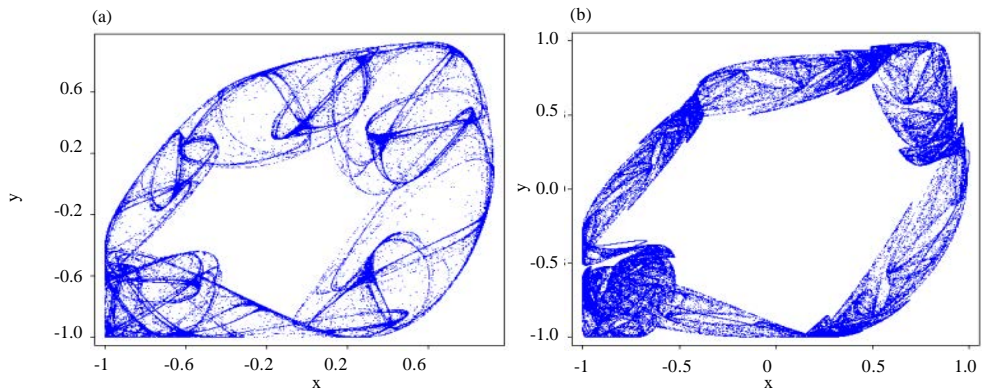


Fig. 8: Continue

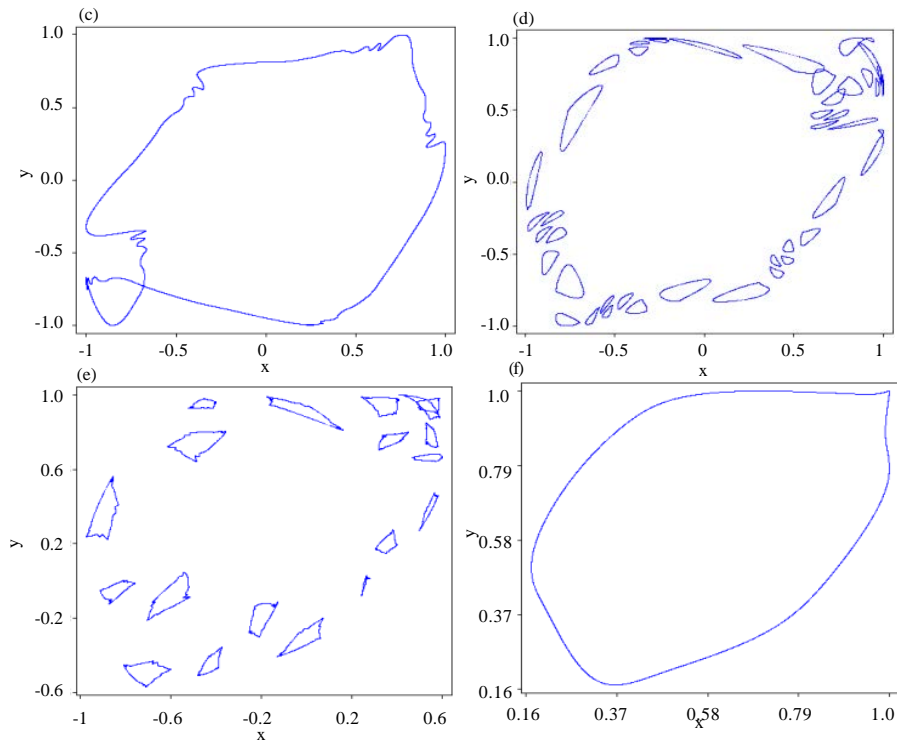


Fig. 8: a) $a = 3.76$; b) $a = 3.76$; c) $a = 3.52$; d) $a = 2.72$; e) $a = 2.24$ and f) $a = 1.84$

the behavior of dynamical systems. The correct prediction of the behavior of orbits of dynamical map (Eq. 1) is important in understanding of evolution process and reduces uncertainty. From the above analysis and the bifurcation diagrams in Fig. 1a, b and 6a, b we deduce that the dynamical behavior on long-term of the map (Eq. 1) is always regular periodic and bounded and has a horizontal stretch, i.e. that the dynamical behavior repeats itself on large time symmetry intervals $[-a, a]$, $a \in \mathbb{R}$ as it moves along the a -axis and the cycles of this regular repeating are called periods and the amplitude of the behavior bands will be 1 (confined between -1 and +1) as shown in Fig. 1a, b and 6a, b. Furthermore, the dynamical behavior of map (Eq. 1) given in Fig. 7 and 8 are respectively symmetric about the origin and inside the two finite symmetry intervals $[-a, 0]$ and $[0, a]$, $a \in \mathbb{R}$ (Fig. 7a-f and 8a-f).

CONCLUSION

In this study, the prediction accuracy on long-term of future behavior in 3-D systems is possible this phenomenon is justified by numerical investigation. In other hand an analytic properties for the dynamics of this system is also presented in term of a single bifurcation

parameter. Furthermore, the new 3-D system produces new several chaotic attractors observed via. quasi-periodic route to chaos.

REFERENCES

Arrow, K.J., R. Forsythe, M. Gorham, R. Hahn and R. Hanson *et al.*, 2008. The promise of prediction markets. *Sci. N. Y. Washington*, 320: 877-878.
 Boettiger, C. and A. Hastings, 2013. Tipping points: From patterns to predictions. *Nat.*, 493: 157-158.
 Chua, L.O., M. Komuro and T. Matsumoto, 1986. The double scroll family. *IEEE Trans. Circuits Syst.*, 33: 1072-1118.
 Dauwels, J., F. Vialatte, C. Latchoumane, J. Jeong and A. Cichocki, 2009. EEG synchrony analysis for early diagnosis of Alzheimer's disease: A study with several synchrony measures and EEG data sets. *Proceedings of the Annual IEEE International Conference on Engineering in Medicine and Biology Society EMBC*, September 3-6, 2009, IEEE, Minneapolis, Minnesota., ISBN:978-1-4244-3296-7, pp: 2224-2227.
 Dullin, H.R. and J.D. Meiss, 2000. Generalized Henon maps: The cubic diffeomorphisms of the plane. *Phys. D. Nonlinear Phenom.*, 143: 262-289.

- Gonchenko, S.V., I.I. Ovsyannikov, C. Simo and D. Turaev, 2005. Three-dimensional Henon-like maps and wild Lorenz-like attractors. *Intl. J. Bifurcation Chaos*, 15: 3493-3508.
- Gonchenko, S.V., J.D. Meiss and I.I. Ovsyannikov, 2006. Chaotic dynamics of three-dimensional Henon maps that originate from a homoclinic bifurcation. *Regul. Chaotic Dyn.*, 11: 191-212.
- Gonchenko, S.V., V.S. Gonchenko and J.C. Tatjar, 2007. Bifurcation of three-dimensional diffeomorphisms non-simple quadratic homoclinic tangencies and generalized henon maps. *Regul. Chaotic Dyn.*, 12: 233-266.
- Hu, W., G.H. Zhao, G. Zhang, J.Q. Zhang and X.L. Liu, 2012. Stabilities and bifurcations of sine dynamic equations on time scale. *Acta Phys. Sin.*, Vol. 61,
- Jeong, J., 2002. Nonlinear dynamics of EEG in Alzheimer's disease. *Drug Dev. Res.*, 56: 57-66.
- Lehnertz, K. and C.E. Elger, 1998. Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity. *Phys. Rev. Lett.*, 80: 5019-5022.
- Lorenz, E.N., 1963. Deterministic non-periodic flow. *J. Atmos. Sci.*, 20: 130-141.
- Lu, J. and G. Chen, 2002. A new chaotic attractor coined. *Int. J. Bifurcation Chaos*, 12: 659-661.
- Mammeri, M., 2016. A large chaotic region in a 3-D sinusoid discrete map. *Intl. J. Appl. Math. Stat.*, 55: 133-144.
- Mandal, K., S. Banerjee and C. Chakraborty, 2011. Symmetry-breaking bifurcation in load resonant DC-DC converters. *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, May 15-18, 2011, IEEE, Rio de Janeiro, Brazil, ISBN:978-1-4244-9473-6, pp: 1327-1330.
- Mandal, K., S. Banerjee and C. Chakraborty, 2013. Symmetry-breaking bifurcation in series-parallel load resonant DC-DC converters. *IEEE. Trans. Circuits Syst. Regul. Pap.*, 60: 778-787.
- Ogata, K., 1995. *Discrete-Time Control Systems*. 2nd Edn., Prentice Hall, Upper Saddle River, New Jersey, USA., ISBN:9780133286427, Pages: 745.
- Romanelli, L., M.A. Figliola and F.A. Hirsch, 1988. Deterministic chaos and natural phenomena. *J. Stat. Phys.*, 53: 991-994.
- Rossler, O.E., 1976. An equation for continuous chaos. *Phys. Lett. A*, 57: 397-398.

Energy Yield Analysis for Fixed Linear Fresnel Reflectors Type Concentrated Solar Power System in the UAE

Ahmet Ozturk and Zaki Iqbal
RAK Research and Innovation Center, Department of Engineering,
American University of Ras Al Khaimah, P.O. Box 31208, Ras Al Khaimah RAK,
United Arab Emirates

Abstract: Concentrated Solar Power (CSP) is a commonly researched branch of renewable energy technologies. Linear Fresnel Reflector (LFR) systems are also not a very niche topic in CSP. This study aims to approach the LFR systems in a quite different concept, by considering a new way of solar tracking. In this study, the Fresnel mirrors are fixed on a rotating platform which can track the sun's azimuth angle by rotation. An analytical model to populate the mirrors on a defined area is developed and using the model a mirror setup is generated. Moreover, the effects of the structural parameters on the solar yield are analyzed by using a ray tracing software. When analyzing the solar yield the study keeps focus on the optical parts of LFR systems.

Key words: CSP, LFR, rotation, optical, tracing, yield

INTRODUCTION

Global rising temperature as a result of huge carbon emissions, volatile oil prices and concerns of governments about security of energy supply caused rapid growth in the renewable energy market, especially in solar Thermal-Concentrated Solar Power (CSP) technologies (Crabtree and Lewis, 2007; Mills, 2004). In this study, the Fresnel mirrors are fixed on a rotating platform which can track the sun's azimuth angle by rotation. An analytical model to populate the mirrors on a defined area is developed and using the model a mirror setup is generated. Moreover, the effects of the structural parameters on the solar yield are analyzed by using a ray tracing software. When analyzing the solar yield the study keeps focus on the optical parts of LFR systems.

Motivation: Growth in the economy and increasing living standards are welfare indicators that every country is proud of. These developments are closely bound to their energy consumption. With the world having a population and economic growth trend as of now, we may assume that the energy consumption will increase even quicker. Increasing oil prices and global climate changes are closely affecting the supply of energy required by this inevitable growth. After the global oil crises in 1973 and 1979, the world focused on developing renewable energies. New technologies were experimented and

implemented in large scales which were researched in the past, some having a history of more than two centuries. The first big scale parabolic trough, dish, solar tower and solar chimney plants were installed as a consequence of these crises. Until that time due to market resistances and insufficient financial and political supports of governments, the developments in this field were delayed. Due to global rising temperature as a result of huge carbon emissions, volatile oil prices and concerns of governments about security of energy supply caused rapid growth in the renewable energy market, especially in solar thermal-concentrated solar power-technologies (Crabtree and Lewis, 2007; Mills, 2004).

The most common solar thermal power concept-parabolic trough- is more costly in comparison to linear Fresnel due to the curved mirrors and tracking mechanism. Using Fresnel reflectors seems as an attractive option to solve the cost problem Larbi *et al.* (2000). The classical Linear Fresnel Reflector (LFR) concept is cheaper in comparison to other solar thermal power technologies. But still in LFR technology every mirror is tracking the sun autonomously to reflect the incident sunrays to the receiver which also adds to the total cost of the systems. The particular specialty of this project is to contribute in developing a low cost CSP system which will generate electricity with lower cost per kWh of electricity. Omitting the tracking system for each mirror and populating extra flat mirrors on a platform can achieve this. The platform as a whole is capable of

rotating and tracking the sun's motion in azimuth angle. As a result to develop a robust, efficient, scalable and modular mirror design is aimed.

Objective: The study focuses on developing a mirror setup as Linear Fresnel Reflector (LFR) for concentrated solar power utilization purpose. Analyzing the effect of structural errors on the solar yield (sensitivity analysis) of the LFR module is also an important target of this study. The existing module is also based on the principle fixed mirrors on a rotating platform. Therefore, this study will show how to calculate a mirror setup for LFR. Then the calculated mirror setup will be compared with the existing mirror setup using ray tracing software. For the sensitivity analysis, the existing mirror setup will be used, so that, the results can be applied and experimented without waste of time and resource. The concept will be based on extra flat plate mirrors which will be fixed on a rotating platform to track the sun's motion in azimuth direction throughout the day. When positioning the mirror strips in the mirror setup, they should be positioned in a formation that blocking and shading of mirror strips will be minimally.

The RAK research and innovation center in Ras al Khaimah, United Arab Emirates can use the study as a base for further experiments on an existing LFR module available. Therefore, the solar data of this region will be used in the calculations and simulations. The existing module is also based on the principle fixed mirrors on a rotating platform, yet, the positioning methodology of the mirror setup and efficiency factors are unknown. Therefore, this study will show how to calculate a mirror setup for LFR. Then, the calculated mirror setup will be compared with the existing mirror setup using ray tracing software. For the sensitivity analysis the existing mirror setup will be used, so that, the results can be applied and experimented without waste of time and resource.

Literature review

Solar radiation: The solar radiation reaches outside of the atmosphere through vacuum of space. During its journey through the atmosphere to the earth surface solar radiation gets weakened because of scattering, reflection and absorption.

The incoming solar radiation is reduced approx. to half of its initial value when reflected and absorbed by environmental factors such as clouds, atmospheric gases and aerosols. Beside that the, earth radiates some amount of the energy from its surface to the sky. Again a significant amount of this radiation is returned to the earth by the greenhouse gases. From the extraterrestrial solar radiation the remaining solar rays

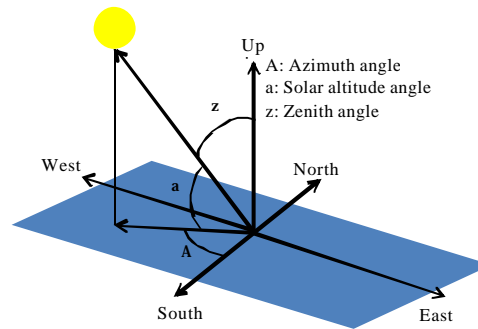


Fig. 1: Sun path and its corresponding angles

reaching the surface are called direct radiation or beam radiation. Some portion of the absorbed, reflected or scattered solar radiation in the atmosphere returns to earth surface in form of short wavelength radiation. Photovoltaic panels can utilize global solar radiation where for solar thermal systems direct beam radiation is useful (Garg, 2009).

Sun's path: While sun and earth are moving in their determined orbits in the space, solar systems require tracking the sun where their efficiency is directly related with incidence angle of the sun rays. For tracking the sun its position in the sky with respect to a reference point on the earth should be determined accurately. The sun's position can be defined by the following angles as also shown in Fig. 1.

A is the Azimuth angle and defines the angle between the projected position of the sun on the horizontal plane and the south direction. z is the zenith angle and is the deviation of solar position from the normal vector (up direction). α is the solar altitude angle or elevation angle. Zenith angle and solar altitude angle can be substituted for each other ($\alpha = 90-z$) (Mertins, 2009).

The sun angles are not only depending on the time of the day but also on the declination at the respective day of the year. To put in other words the pattern of sun's path in the sky changes throughout the year (Walter and George, 2005).

The sun's height (solar altitude angle) is ranging from $0-90^\circ$ where the azimuth is ranging from -180° to 180° . The East is defined per definition with $A = -90^\circ$ (Gharbi *et al.*, 2011).

Concentrated solar power: Solar power has relative low density which must be collected and concentrated for efficiently utilization. Solar power can be concentrated using refracting or reflecting optical surface. Reflecting optical surfaces are most commonly used method for commercial purposes so the study will focus on this side of CSP.

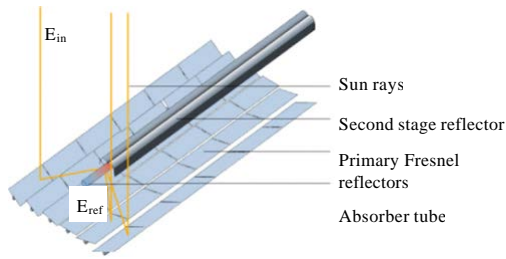


Fig. 2: Basic components of a linear Fresnel reflector module (Haberle *et al.*, 2002)

By line focusing plants, the receiver and the reflector mirrors are positioned parallel to each other. Parabolic trough and linear Fresnel systems count in this category of CSP.

Figure 2 presents a closer look to LFR. Primary Fresnel reflector mirrors reflecting (E_{ref}): the incident solar rays (E_{in}) to the receiver tube. In some designs, also, a second stage reflector above the receiver can be included to collect the escaping rays and redirect them back to the receiver.

In a classical linear Fresnel reflector design several mirror strips are populated in parallel along a receiver, which are also called primary mirrors. The primary mirrors are responsible to reflect the incident solar rays to the receiver which is placed at a determined height. The mirrors are populated having a certain distance between each other to reduce shadowing and blocking effects. The receiver consists of an absorber tube and optionally a secondary reflector and insulation. Secondary reflector facing with its reflective surface downwards to the primary mirrors assures to concentrate the escaping rays back on the absorber tube. Closing the secondary reflector with a glass plate from its downside prevents negative effects of dust and convection losses (Mertins, 2009). To achieve reasonable solar efficiency solar tracking is a must for this system. Most common way of tracking is installation of a separate one axis tracking mechanism for every single mirror strip. Another option is moving the receiver and keeping the mirrors fixed. Although, there is theoretical research done on fixed mirrors-moving receiver, commercially it is not applied (Pujol *et al.*, 2006). This study considers extra flat fixed mirrors populated on a rotating platform which is tracking the sun in one axis in other words tracking the azimuth angle of the sun.

MATERIALS AND METHODS

Mathematical modeling of the linear Fresnel reflector Parameters of mirror design to be used in mathematical modeling: Figure 3 shows a simplified schematic of the

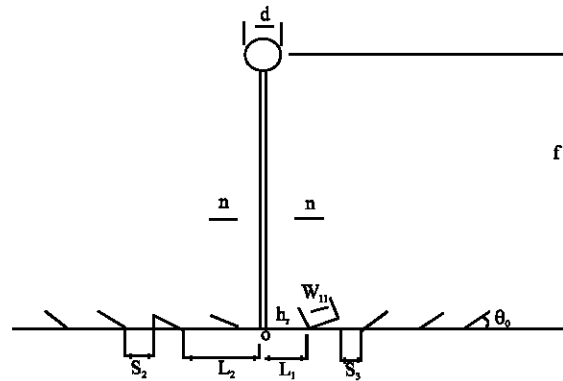


Fig. 3: Simplified schematic of the mirrors and the receiver tower

mirrors and the receiver tower from the front side. Significant parameters affecting the solar yield are d is the diameter of the receiver tube, n is the index and also the number of mirror strips on each side of receiver tower in one module, f is the focal height. The distance between center point of the receiver tube from the horizontal plane where bottom points of mirrors lie, W_n is the Width of the n th mirror, θ_n is the inclination angle of mirror measured counterclockwise from the horizontal plane, L_n is the distance of n th mirror's bottom end from the receiver tower, S_n is the Space between n th and $(n-1)$ th mirrors, measured horizontally from last point of the mirror to the first point of next mirror, h , is the horizontal position of receiver with respect to the origin, O is the origin of the module. Mirror's lower ends are positioned on the horizontal line which represents the x-axis for the calculations in following chapters. The receiver tower is aligned to the positive y-axis.

Building the mathematical model: For the modeling of the optical system only direct beam irradiance of the sun is considered. The concentrator-set of mirror strips populated on the platform is perfectly tracking the sun. Reflected Energy E_{ref} by the solar collector to the focal surface is a function of the incident solar Energy E_{in} , design, environment, and material parameters:

$$E_{ref} = f(E_{in}, \text{design, environment, material}) \quad (1)$$

Reflected energy is basically equal to the incident energy multiplied by an efficiency factor η :

$$E_{ref} = \eta \times E_{in} \quad (2)$$

The energy initiating from the sun and reaching the receiver tube after being reflected by the mirrors is a

function of incident solar energy on the mirrors, environmental effects and material properties of the mirrors. So, the reflected energy is proportional to the incident energy and multiplied by an efficiency factor. This efficiency factor η is also a function of the design, chosen materials and environmental effects:

$$\eta = f(\text{design,material,environment}) \quad (3)$$

The equation for the efficiency can be expressed as the multiplication of the efficiencies for the aforementioned factors design, material and environment:

$$\eta = \eta_{\text{design}} \times \eta_{\text{material}} \times \eta_{\text{environment}} \quad (4)$$

The efficiencies due to the material and environmental factors are not considered in this study. So, the design efficiency can be further divided in reflector and receiver efficiencies:

$$\eta_{\text{design}} = \eta_{\text{reflector}} \times \eta_{\text{receiver}} \quad (5)$$

The efficiency of reflector $\eta_{\text{reflector}}$ is a function of its physical dimensions and design where the parameters in following equation are stated before:

$$\eta_{\text{reflector}} = f(W, f, \theta, S, L, d) \quad (6)$$

Following equations are used in iteration to find a mirror design for predetermined receiver diameter, width of mirrors and height of receiver where receiver is placed on the focal line of the mirror setup.

Following five equations are based on the study by Sootha and Negi (1994). In that study, the solar irradiation has been considered, so that, there is an angular subense of the sun on the earth.. In this study, the sun rays reaching the linear reflector system has been assumed to be parallel to each other.

Equation 7 calculates the inclination angle of the nth mirror. A vertical beam is drawn downwards to the end point of a mirror and reflected off the mirror to the receiver where it hits the lower half of the receiver tangentially. x_n and y_n are the coordinates of this tangential point. L_n is the horizontal distance of the nth mirror's lower point from the symmetry plane:

$$\theta_n = \frac{1}{2} \times \tan^{-1} \left(\frac{L_n + x_{a,n}}{y_{a,n}} \right) \quad (7)$$

Using simple trigonometric equities one will obtain Eq. 8 and 9 which provide the required values of the tangential point on the receiver tube to calculate the inclination angle θ_n of mirror strip:

$$x_{a,n} = R \cdot \cos \theta_n \quad (8)$$

$$y_{a,n} = f - R \cdot \cos \theta_n \quad (9)$$

To prevent the blocking and shading effects of mirrors to each other a minimum space-shift-should be kept between each mirror strip. Following equation calculates the required space between nth and (n-1)th mirror strips:

$$S_n = \frac{(L_{n-1} + W_{n-1} \cos \theta_{n-1} + x_{n,a}) W_{n-1} \sin \theta_{n-1}}{y_{a,n} W_{n-1} \sin \theta_{n-1}} \quad (10)$$

Following Eq. 10 calculates where the lower end point of the mirror should be placed on horizontal plane which is required for installation. Namely this Eq. 10 defines location of the respective mirror. The width of the respective mirror strip, the shift and the location of previous mirror are required parameters:

$$L_n = L_{n-1} + W_{n-1} \cos \theta_{n-1} + S_n \quad (11)$$

Following Eq. 11 provide the factors to evaluate and compare the efficiency and functionality of the calculated mirror setup. The concentration factor defines the ratio of collected solar energy by mirrors to the concentrated energy on the receiver:

$$\text{Concentration factor} = \frac{A_{r_{\text{mirror effective}}}}{A_{r_{\text{receiver}}}} \quad (12)$$

The gross area of mirrors define the total area used for one side of the reflector including the sum of dead space -the necessary space between each mirror to prevent blocking and shading-between mirrors and sum of mirror's area on the horizontal plane:

$$A_{r_{\text{mirror, gross}}} = 2 \cdot \left[L_1 \sum_{n=1}^k W_n \cos \theta + S_n \right] \quad (13)$$

The effective mirror area is the sum of dead space between mirrors subtracted from the gross area of mirrors:

$$A_{r_{\text{mirror effective}}} = 2 \cdot \sum_{n=1}^k W_n \cos \theta \quad (14)$$

Area utilization factor is an important measure to determine how well the reflector area is used and how dense the mirrors are populated. A greater utilization factor allows greater reflection of incident solar energy to the receiver:

$$\text{Area utilization factor} = \frac{A_{r_{\text{mirror, effective}}}}{A_{r_{\text{mirror, gross}}}} \quad (15)$$

Computational model: For the calculation of mirror setup, MATLAB and MS Excel is chosen. The equations mentioned above require to be computed in a loop. Given these conditions using MATLAB was the optimum choice.

To determine the positions of mirror elements in a Linear Fresnel Reflector (LFR) design, one needs to determine four main parameters.

- Width of the mirror element (W)
- Angle of inclination of each mirror element (θ)
- Respective distance of each mirror element to a reference point (L)
- Space between each mirror element (S)

Calculating the mirror setup: In order to determine these parameters Eq. 7-11 has been used. Equation 7 is used to calculate the inclination angle of each mirror strip and takes $x_{a,n}$ and $y_{a,n}$ as input. But then Eq. 8 and 9 which calculate $x_{a,n}$ and $y_{a,n}$, take θ_n as input. To solve this problem an iteration procedure should be performed. Firstly, a θ_n value is assumed and using this $x_{a,n}$ and $y_{a,n}$ are calculated. Then putting $x_{a,n}$ and $y_{a,n}$ in Eq. 7 gives another θ_n value. The θ_n values of this and previous steps are compared. So, the calculation should be put in a loop and iterated until the input θ_n to Eq. 8 and 9 and result θ_n of Eq. 7 matches. Following this, location of mirror strip and shift between mirrors are calculated.

The code first takes input values for d, f, W width of the LFR module on one side and minimum shift between each mirror i from the user. Then initialization of variables is done. After taking inputs and initializing variables the above mentioned loop is executed. The loop calculates as many mirrors as they would fit in a somewhat wider area than the given width of the LFR module. Another small loop checks the positions of the mirrors and determines which one is the last mirror strip that fits within the defined module width.

After having the mirror setup calculated and number of mirror strips determined, next step is to round the significant figures of the results of the inclination angle, mirror locations and shift between mirrors to manufacturing limits. In the last step, the code imports the number of mirror strips θ , L, S and W to an MS Excel file.

Model validation: To implement the theoretical calculated mirror setup in the simulation environment some assumptions had to be made. Four cases of solar altitude angles were chosen as point of interest such as:

Table 1: Comparison of calculated power with the simulation results

Solar altitude angles	Direct horizontal Beam radiation (Wm^{-2})	Energy model calculation (W)	Simulation tracepro (W)	Difference (%)
90°	546	24056	24053	<0.01
60°	488	21501	21491	<0.01
45°	411	18108	18091	<0.01
30°	285	10156	9227	9

- 90°
- 60°
- 45°
- 30°

Simulation environment is built up using three types of building blocks:

- A surface source representing sun
- Mirror strips
- Receiver tube

A computational method was used to check the power yield results obtained by ray tracing on a reference setup. The calculation was crosschecked with simulation of TracePro software and since the results were correlating it was reliable to continue with the next step, sensitivity analysis to structural errors.

Table 1 shows the comparison results. For most of the solar altitude angles in the design interval of the receiver, the calculation and simulation results match. At 30° one can observe a difference of 9%. The reason is that the receiver length is not enough to absorb all the reflected power at this solar altitude angle and the computational model was not powerful enough to consider the loss due to this factor.

RESULTS AND DISCUSSION

Figure 4a shows the power change with respect to the vertical change in position of the receiver. The initial design position of the receiver is taken as origin. Then simulation is performed by linearly increasing-decreasing the height of the receiver with 1cm intervals until the received power decreases at least by 30% at one of the solar altitude angles. Same procedure is repeated with four solar altitude angles. As a result vertical change in position of receiver should be between -4 and +5 cm to prevent power loss more than 10%. Considering the 4 m height of the receiver tower, 4-5 cm position change means an error allowance of 1% which should not be ignored and regarded during mechanical design and manufacturing process. Still, small vibrations due to external disturbances will not have a big effect on the solar yield from the mechanical point of view.

Figure 4b shows the power change with respect to the horizontal change in position of the receiver. Since, the mirror setup is symmetric with respect to the receiver

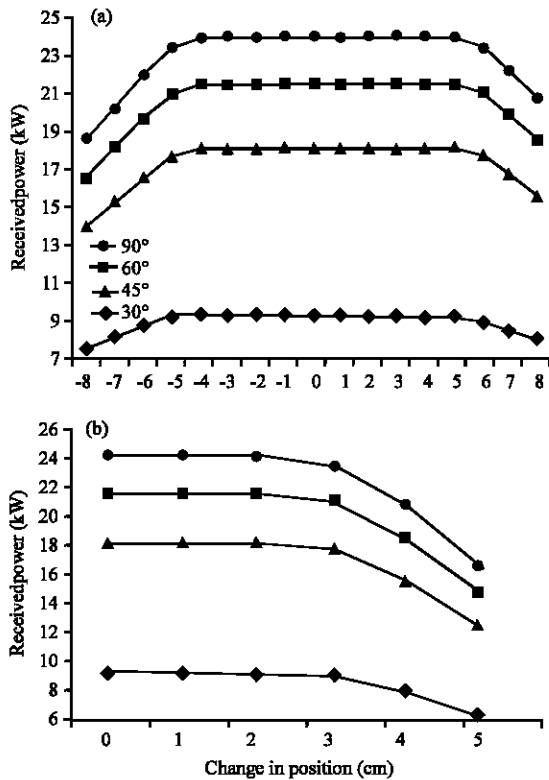


Fig. 4 a, b: Solar power vs. vertical and horizontal change in position of receiver, respectively

tower, the change in position is considered only for one direction. The initial design position of the receiver is taken as origin. Then simulation is performed by horizontally moving the receiver to one side with 1cm intervals until the received power at one of the solar altitude angles drops to zero. Same procedure is repeated with four solar altitude angles. As a result horizontal change in position of receiver should be <3 cm to prevent power loss more than 10%.

The conclusion for Fig. 3 can also be deduced, here, that the receiver is not very sensitive to horizontal changes of position. Still the receiver is more sensitive to position changes in horizontal direction than vertical position changes. The significant drop in the solar yield is observed after 2 cm where for vertical change of position it started at 4 cm downwards and 5 cm upwards.

Figure 5a shows the change in received power by tilting the mirror strips in certain intervals and decreasing the inclination angle of mirror strips. The initial design position of the mirror strips are taken as origin. Then simulation is performed by decreasing the inclination angle of mirror strips with 0.1° intervals until the received

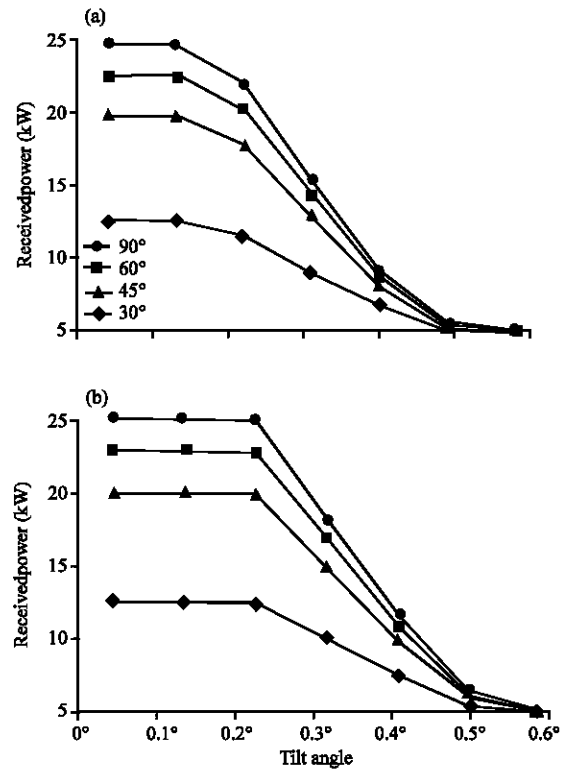


Fig. 5a, b: Solar power vs. angular smaller and greater inclination angles of mirror strips, respectively

power drops to zero. Same procedure is repeated with four solar altitude angles. As a result tilting the mirrors for smaller inclination angle should be less than 0.1° to prevent power loss more than 10%.

A possible solution to this problem can be to install a second stage reflector above the receiver tube facing downward towards the primary mirrors. Beside that the thermal drawbacks of the second stage reflector should be analyzed, since, the redirected sunrays will be hitting upper part of the receiver tube. In a DSG system if the receiver tube is not filled with water and there is two phase flow in the pipe, the gaseous part of the flow will be heated. That might cause structural defects on the receiver because of local overheated areas.

Figure 5b shows the change in received power by tilting the mirror strips in certain intervals and increasing the inclination angle of mirror strips. The initial design positions of the mirror strips are taken as origin. Then simulation is performed by increasing the inclination angle of mirror strips with 0.1° intervals until the received power drops to zero. Same procedure is repeated with four solar altitude angles. As a result tilting the mirrors for greater inclination angle should be <0.2° to prevent power loss

more than 10%. Within this 0.2° error the mirror is still aiming mostly at the receiver. Beyond that, a dramatic drop is observed and the mirrors are defocused.

The conclusion for Fig. 5a can also be deduced here with an addition that the system is somewhat less sensitive for greater inclination angles. In this case, the missing rays are not hitting the receiver by going away from downside of it. Making the second stage reflector longer at the edges and introducing a little bit more curvature might lessen the effect of this problem.

Making a stronger and more stable mechanical platform structure remains also as a possible solution for all of the four aforementioned error types.

Figure 6 shows the change in received power when there is an error in tracking the azimuth angle of the sun. Sun being perfectly tracked was chosen as origin. Then simulation is performed by introducing an error to the azimuth angle tracking in 0.1° intervals until the received power drops to zero. Same procedure is repeated with four solar altitude angles. As a result tracking error should be <0.3° to prevent power loss more than 10%.

Sun tracking is a vital factor in every CSP system to obtain a high efficiency. Due to reflection geometry, the azimuth tracking error doesn't play a role on the solar yield when the solar altitude angle is 90°. Still for lower solar altitude angles, the effect gets more powerful. The preciseness of a tracking system is most of the time directly related with its costs. This figure provides an understanding for the level of preciseness required in such a LFR system.

Table 2 and 3 summarize the figures presented in this chapter. If the parameter changes are kept within the limits stated in Table 1, the loss in received power should be <1%. The factors are listed in decreasing order of importance. The inclination error of the mirrors has the greatest effect on the solar yield where the allowable error interval lies between 0.1° and 0.2° for a solar power yield loss <1%.

Most of the energy loss due to sensitivity errors takes place at the outer mirrors. Main reason for this is that their inclination angles are relatively greater than the inner mirror which are near to the receiver and the reflected sun rays follows a shorter distance. These differences make the outer mirrors more vulnerable to external disturbances. Another notable point is that azimuth tracking error is the only error type that is affected by the solar altitude angle. Azimuth tracking error is also less effective by higher solar altitude angles.

A quick look will reveal that the angular changes affect the power yield at most. Even an increase of

Table 2: Maximum allowable parameter change to achieve < 1% change in the reflected power

Affected factor	Error	Solar altitude angle received power on the receiver (W)			
		90°	60°	45°	30°
Reference received power					
Mirrors	-	24053	21500	18097	9267
θ smaller	0.1°	24064	21502	18100	9221
θ greater	0.2°	23788	21256	17846	9118
Platform					
Azimuth tracking, (λ)	0.2°	24053	21466	18106	9040
Receiver					
Horizontal, (h _r)	±2 cm	24053	21491	18091	9218
Vertical, (f)	-4 cm	24053	21491	18091	9312
	+5 cm	24049	21491	18091	9091

Table 3: Maximum allowable parameter change to achieve less than 33% change in the reflected power

Affected factor	Error	Solar altitude angle received power on the receiver (W)			
		90°	60°	45°	30°
Reference received power					
-	24053	21500	18097	9267	
Mirrors					
θ smaller	0.2°	20818	18600	15657	8009
θ greater	0.3°	16093	14381	12055	6259
Platform					
Azimuth tracking, (λ)	0.3°	24053	21459	18088	6879
Receiver					
Horizontal, h _r	±4 cm	20603	18418	15496	7773
Vertical, (f)	-8 cm	18431	16457	13860	7451
	+8 cm	20687	18464	15551	7927

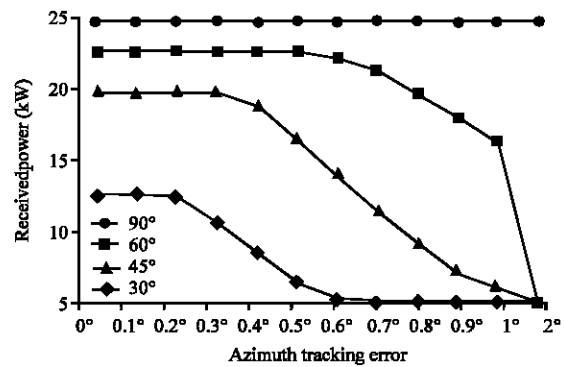


Fig. 6: Solar power vs. error of azimuth angle tracking

inclination angle errors by 0.1° causes a solar yield loss of up to 33%. The azimuth angle tracking has a comparable importance as well. One can conclude that angular stability has the uttermost importance to provide a stable solar yield from a LFR module.

The changes in position of the receiver in vertical or horizontal direction are not affecting the solar yield as much as the angular errors. This indicates that the LFR



Fig. 7: Optical characterization of the LFR mirror at RAKRIC

module is not much affected by the environmental effects on the receiver tower such as wind and external vibrations, assuming that no angular disturbance is existent (Fig. 7).

CONCLUSION

Linear Fresnel reflector is a promising technology despite of its relative lower efficiency in comparison to other CSP technologies. Its low cost and easy manufacturability provides powerful advantages.

In this study, an equation model to generate a LFR mirror setup is built. Based on this equation model a LFR mirror setup is calculated and the calculated module's functionality is validated by in ray tracing simulation. Developing a mirror setup without blocking and shading has been accomplished.

After proving that the calculated mirror setup functions correctly, the existing LFR module of the RAK Research Center and the calculated LFR module are compared via. ray tracing simulation. The similar characteristics and solar yields of the modules allowed continuing with the sensitivity analysis on the existing module. This shift from calculated to existing module was necessary in order to enable a faster start of further experiments for future researches. Sensitivity analysis revealed that the solar yield is highly dependent on the angular accuracy of both mirror installation and azimuth tracking. The stability of receiver tower structure also has significant effects but not as crucially as angular stability. It has been also proved that the LFR system should be unaffected by small vibrations due to external disturbances such as wind, unless no angular error is introduced. In accordance to these before manufacturing the LFR module, manufacturing and installation error

margins and preciseness should be considered thoroughly and measures should be taken, accordingly.

RECOMMENDATIONS

For future research, before continuing with experimental research it is suggested to conduct a research on thermal behavior of the receiver tube and the fluid in it. Only then the results can be used to build a reliable experimental setup and a realistic overall efficiency for the whole LFR module can be calculated. Since, research and engineering are iterative processes also in advance this study might need to be updated in light of these future results. This research led to the optical characterization of the LFR mirror at the RAK Research and Innovation Center.

ACKNOWLEDGEMENTS

The researchers greatly acknowledge the help and technical support given by Lambda Research Corporation and providing ray-tracing simulation software Trace Pro Version 7.0 for this research. Also many thanks to the RAK Research and Innovation Center, previously known as CSEM-UAE Institute for hosting and supporting this research.

REFERENCES

- Crabtree, G.W. and N.S. Lewis, 2007. Solar energy conversion. *Phys. Today*, 60: 37-42.
- Garg, H.P., 2009. *Fundamentals of solar energy*. Centre for Energy Studies, New Delhi, India.
- Gharbi, N.E., H. Derbal, S. Bouaichaoui and N. Said, 2011. A comparative study between parabolic trough collector and linear Fresnel reflector technologies. *Energy Procedia*, 6: 565-572.
- Haberle, A., C. Zahler, H. Lerchenmuller, M. Mertins and C. Wittwer *et al.*, 2002. The solarmundo line focussing fresnel collector: Optical and thermal performance and cost calculations. *Proceedings of the International Symposium on Concentrated Solar Power and Chemical Energy Technologies*, September 4-6, 2002, SolarPACES, Zurich, Switzerland, pp: 1-11.
- Larbi, A.B., M. Godin and J. Lucas, 2000. Analysis of two models of (3D) Fresnel collectors operating in the fixed-aperture mode with a tracking absorber. *Sol. Energy*, 69: 1-14.
- Mertins, M., 2009. [Technical and Economic Analysis of Horizontal Fresnel Collectors]. University of Karlsruhe, Karlsruhe, Germany, (In German).

- Mills, D., 2004. Advances in solar thermal electricity technology. *Sol. Energy*, 76: 19-31.
- Pujol, R., V. Marinez, A. Moia and H. Schweiger, 2006. Analysis of stationary Fresnel like linear concentrator with tracking absorber. *Proceedings of the Solar Paces 13th Symposium on Concentrating Solar power and Chemical Energy Technologies*, June 20-23, 2006, International Energy Agency, Seville, Spain, ISBN:84-7834-519-1, pp: 1-5.
- Sootha, G.D. and B.S. Negi, 1994. A comparative study of optical designs and solar flux concentrating characteristics of a linear Fresnel reflector solar concentrator with tubular absorber. *Sol. Energy Mater. Cells*, 32: 169-186.
- Walter, R.S. and W.B. George, 2005. Inclination of the ecliptic. Night and day, inc, Houston, Texas. http://www.cso.caltech.edu/outreach/log/NIGHT_DAY/inclination.htm.

Generic Approach for Classifying Spam Mails by Machine Learning Techniques

¹Banumathy Rajesh and ²Shanmugasundaram Hariharan

¹Research and Development Centre, Bharathiar University, 641046 Coimbatore, India

²Department of Information Technology, Vel Tech Multi Tech, Chennai, India

Abstract: Email communication is one of the fastest means of information sharing and has become successful among online users. This predominant success has made web users to generate anonymous contents which are called as spam. Research on identifying these fraudulent information has been a major research issue in recent years and continues to be a major threat. Spam occurs in the information in textual form, short messages and images. Variety of methods exists to ensure security like Naive Bayes, machine learning, Genetic algorithm and others. Machine learning techniques now days used to automatically filter the spam e-Mail in a very successful rate. Classifying the emails as genuine or vice versa is a major research concern. This study attempts to provide a study on this context and there by provide a framework for improving the security. Descriptions of the algorithms are presented and the comparison of their performance on the SpamAssassin spam corpus is presented.

Key words: Email, spam, machine learning, blacklist, Naive Bayes, classification

INTRODUCTION

Information users depend heavily on email's system as one of the major sources of communication. Due to tremendous growth in e-Content and communication there has been a great deal of interest of late in the problem of automatically detecting and filtering out unsolicited commercial e-Mail messages which is commonly referred to as spam. The internet has brought about fundamental changes in the way peoples generate and exchange media information. A variety of approaches have been discussed in this context identify fraudulent information (Lai, 2007). To reduce the spamicity, several algorithms are in existence. Machine learning approach is more efficient than knowledge engineering approach, it does not require specifying any rules (Guzella and Caminhas, 2009). A set of training samples these samples is a set of pre classified e-Mail messages. These unsolicited bulk electronic mails (spam e-Mail) were expressed in different formats and have become one of the most serious problems in internet era. Internet Service Providers (ISP), business firms and general end users the global rate of spam in email traffic in 2014 was 60% internet security threat (Corporation, 2015). Effective image spam detection is of importance in many different domain applications (Soranamageswari and Meena, 2010). Different kinds of features have been used in existing image spam filters including e-Mail header features, image metadata, text-based features and visual-based features. The success of email spam identity have been possible with multimodal textual spam filtering for mobile devices using dendritic cell algorithm (El-Alfy and AlHasan, 2016), Naive

Bayes approach (Almeida *et al.*, 2011). Its importance and usage are continuously growing despite the evolution of mobile applications, social networks, etc. A large set of personal emails is used for the purpose of folder and subject classifications. Algorithms were developed to perform clustering and classification for this large text collection. Classification based on N-Gram is shown to be the best for such large text collection especially as text is bi-language (i.e., with English and Arabic content) (Alsmadi and Alhami, 2015).

Academia and industry have shown their concern to accurately detect and effectively control web spam, resulting in a good number of anti-spam techniques currently available. Though there exists many solutions available for the spam mails, each method has its own significance. Moreover, many spam mail incidents are also observed in today's electronic mail system by the Messaging Anti-Abuse Working Group (MAAWG) found that (IDGCI., 2010).

Emails are used on both the personal and professional levels. They can be considered as official documents in communication among users. Email's data mining and analysis can be conducted for several purposes such as: spam detection and classification, subject classification, etc. Email is a convenient means of communication throughout the entire world today. The increased popularity of email spam in both text and images requires a real-time protection mechanism for the media flow. The effective communication is used by large number of users generating hundred of billion of messages every day. It has several advantages viz., less expensive, high speed, reliable and large data transfer

when compared with other mode of transfers. Tools exists for filtering spam contents, SpamAssassin (ASF., 2016) mail filtering tool is one such tool. SpamAssassin uses a large manually-generated feature set and a simple perceptron classifier with hand-tuned weights to select ham (non-spam) messages and discard spam. Several set of spam words can be filtered easily which could improve the accuracy of the filtering process. The methods available to control spam include anti-spam laws, email protocol exchanges, challenge based systems and response filters. The email originates from various sources. Figure 1 presents the origin of spam mails.

The spam is mixed information which is unwanted to the user. This could be the data originating from multiple sources, information which is irrelevant to the context,

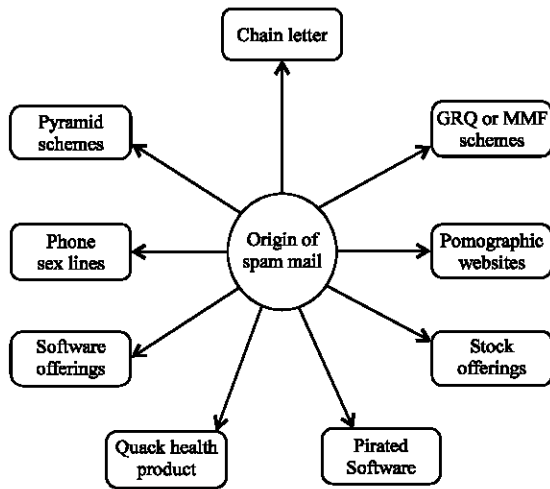


Fig. 1: Origin of spam mails

representation of data in unusual formats, etc. Figure 2 shows a sample image with spam content. Spam datasets are shown in Table 1 which falls into two categories as email or text. Openly accessible datasets are available for investigation and we extract data from SpamAssassin for experimental.

Literature review: Researches on detecting unsolicited message images (image spams) have become one of the most serious problems for Internet Service Providers (ISPs), business firms and general end users. The researchers have presented a novel system called RoBoTs (Robust BoosTrap based spam detector) to achieve a robust image spam filter. The system is developed based on multiple visual properties extracted from different levels of granularity, capturing discriminative contents to achieve effectiveness (Shen *et al.*, 2015). Also the research focuses on to develop a learning frame research to effectively integrate random forest and Linear Discriminative Analysis (LDA). Supervised machine learning algorithms is adopted for extracting subjective information from text documents online. This has facilitated in accurate spam classification process using public datasets used for training and test.

Over the last years, internet spam content has spread enormously inside web sites mainly due to the

Table 1: Spam datasets and its characteristics

Dataset	Focus	Nature	Spam	Legitimate
SpamAssassin	Email	Preprocessed	1813	2788
TREC 2005	Email	Raw	1323	1390
SpamBase	Email	Raw	1323	276
Spam collection	SMS	Raw	747	4827



Fig. 2: Sample content with spam

emergence of new web technologies, online sharing of resources and information. A vast approach uses ensemble learning paradigms for sentiment analysis to enhance the predictive performance (Onan *et al.*, 2016). A paradigm integrating multi objective, optimization based weighted voting scheme is used. An appropriate weight value is assigned to existing classifiers to each class based on the predictive performance of classification algorithms. The ensemble method based on static classifier selection involves majority voting error and forward search as well as a multi objective differential evolution algorithm.

Some illustrative schemes incorporates Bayesian logistic regression, Naive Bayes, linear discriminant analysis, logistic regression and support vector machines as base learners whose performance in terms of precision and recall values determines weight adjustment. Our experimental analysis of classification tasks including sentiment analysis, software defect prediction, credit risk modeling, spam filtering, and semantic mapping, suggests that the proposed classification scheme can predict better than conventional ensemble learning methods such as AdaBoost, bagging, random subspace and majority voting.

The successful integration of different algorithms for web spam classification is still a challenge. In this context, a present study introduces WSF2, a novel web spam filtering framework specifically designed to take advantage of multiple classification schemes and algorithms. The approach encodes the life cycle of a case-based reasoning system with appropriate knowledge and different parameters to ensure continuous improvement in filtering with passage of time leading to good precision. The evaluation of the dynamic model leads to improved efficiency with set of experiments involving a publicly available corpus. The approach using well known classifiers and ensemble approaches were used which revealed that WSF2 performed well, being able to take advantage of each classifier and to achieve a better performance when compared to other alternatives (Heydari *et al.*, 2016).

Earlier approaches were limited by the adaptive nature of unsolicited email spam. To alleviate this research, email detection system was introduced by some researchers. Particle Swarm Optimization (PSO) was implemented to improve the random detector generation in the Negative Selection Algorithm (NSA). The algorithm generates detectors in the random detector generation phase of the negative selection algorithm. The combined NSA-PSO uses a Local Outlier Factor (LOF) as the fitness function

for the detector generation. The detector generation process is terminated when the expected spam coverage is reached. A distance measure and a threshold value are employed to enhance the distinctiveness between the non-spam and spam detectors after the detector generation. The implementation and evaluation of the models are analyzed. The results show that the accuracy of the proposed NSA-PSO Model is better than the accuracy of the standard NSA Model. The proposed model with the best accuracy is further used to differentiate between spam and non-spam in a network that is developed based on a client-server network for spam detection (Idris *et al.*, 2015).

Existing reports clearly indicate that volume of spam over instant messaging and SMS is dramatically increasing year by year. It represents a challenging problem for traditional filtering methods nowadays, since, such messages are usually fairly short and normally rife with slangs, idioms, symbols and acronyms that make even tokenization a difficult task. In this scenario, the approach proposes a method to normalize and expand original short and messy text messages in order to acquire better attributes and enhance the classification performance. Preprocessing is done based on lexicographic and semantic dictionaries along with state-of-the-art techniques for semantic analysis and context detection. This technique is used to normalize terms and create new attributes in order to change and expand original text samples aiming to alleviate factors that can degrade the algorithms performance, such as redundancies and inconsistencies (Almeida *et al.*, 2016).

MATERIALS AND METHODS

Framework for email filtering systems: The framework for email filtering system is presented in this study. Figure 3 presents the framework of our email intrusion detection system. The sequence of steps involved in the proposed system is shown in Fig. 4. The proposed e-Mail Personalization system will be able to prioritize the e-Mail messages into different categories. The performance is evaluated using precision and recall using Eq. 1 and 2. Table 2 presents the spam datasets used for experiments and the characteristics exhibited by the spam datasets.

Table 2: Spam datasets and its characteristics

Message collection	Training set	Testing set
Ham message	842	432
Spam message	736	521

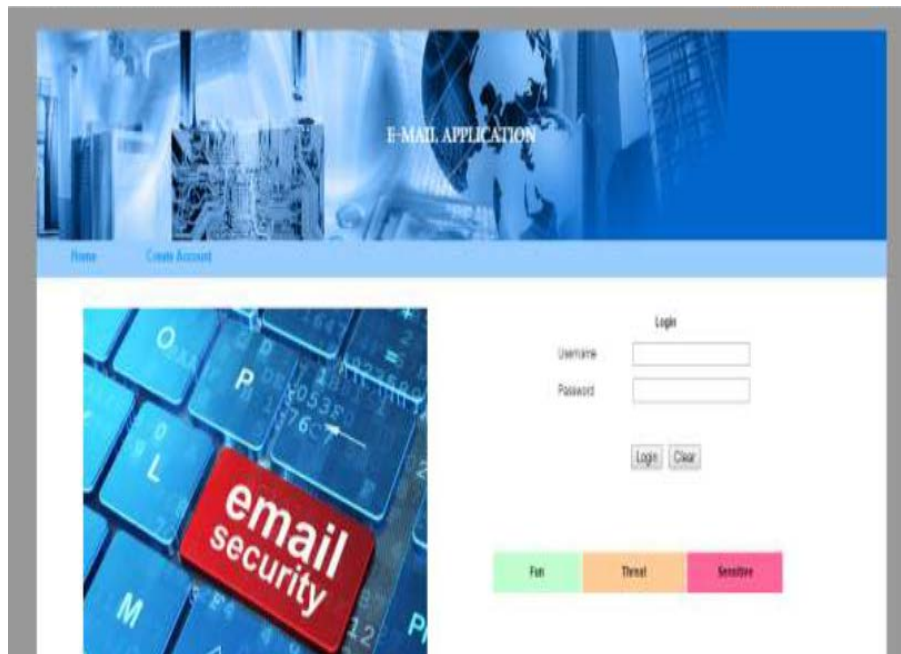


Fig. 3: Proposed framework for spam identification

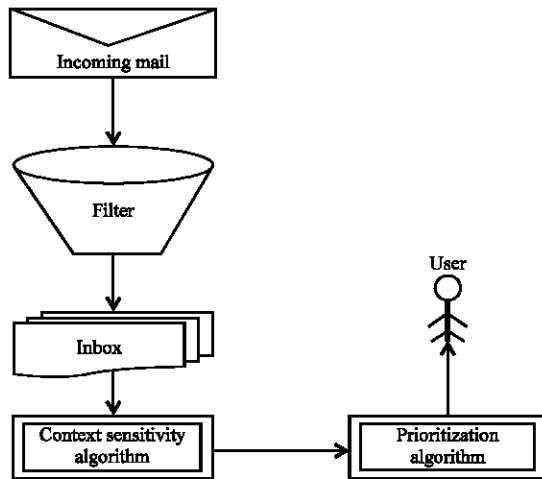


Fig. 4: Architecture for spam detection

RESULT AND DISCUSSION

The results from Table 3 shows that, the SVM approach is dominant and yields better results in the context spam identification. The result shows significant improvements in terms of precision and recall measures:

$$\text{Recall} = \frac{\text{No. of mails correctly classified}}{\text{Total No. of messages}} \quad (1)$$

$$\text{Accuracy} = \frac{\text{No. of mails correctly categorized}}{\text{Total No. of emails}} \quad (2)$$

Table 3: Performance using different classifiers

Methods	Recall	Precision	Accuracy
SVM	0.98	0.99	99.46
KNN	0.95	0.93	96.90

CONCLUSION

The study has presented a study on email filtering using SVM approach. The research shows some improvements as compared to KNN approach improvements in the existing schemes. Experimental illustration have significance using SpamAssassin and TREC 2005 datasets for emails which are raw contents. The proposed approach investigates on the significance of the preprocessing role to reduce the spam mails. Currently research is in progress on to identify image spam.

ACKNOWLEDGEMENTS

The researcher would like to express our sincere thanks to anonymous reviewers for their insightful suggestions. The researcher indeed would extend their acknowledgement for the support by all well wishers, faculty members and others who helped out in achieving this research task to get completed.

REFERENCES

ASF., 2016. Apache spam assassin. Apache Software Foundation, Forest Hill, Maryland, USA. <http://www.spamassassin.org>.

- Almeida, T.A., J. Almeida and A. Yamakami, 2011. Spam filtering: How the dimensionality reduction affects the accuracy of Naive Bayes classifiers. *J. Internet Serv. Appl.*, 1: 183-200.
- Almeida, T.A., T.P. Silva, I. Santos and J.M.G. Hidalgo, 2016. Text normalization and semantic indexing to enhance instant messaging and SMS spam filtering. *Knowledge Based Syst.*, 108: 25-32.
- Alsmadi, I. and I. Alhami, 2015. Clustering and classification of email contents. *J. King Saud Univ. Comput. Inf. Sci.*, 27: 46-57.
- Corporation, S., 2015. Internet security threat report. Symantec, Mountain View, California, USA. https://www4.symantec.com/mktginfo/whitepaper/ISTR/21347932_GAinternet-security-threat-report-volume-20-2015-social_v2.pdf
- El-Alfy, E.S.M. and A.A. AlHasan, 2016. Spam filtering framework for multimodal mobile communication based on dendritic cell algorithm. *Future Generation Comput. Syst.*, 64: 98-107.
- Guzella, T.S. and W.M. Caminhas, 2009. A review of machine learning approaches to spam filtering. *Expert Syst. Appl.*, 36: 10206-10222.
- Heydari, A., M. Tavakoli and N. Salim, 2016. Detection of fake opinions using time series. *Expert Syst. Appl.*, 58: 83-92.
- IDGCI., 2010. Sign up for network world newsletters. IDG Communications Inc. Framingham, Massachusetts. <http://www.networkworld.com/newsletters/>
- Idris, I., A. Selamat, N.T. Nguyen, S. Omatu and O. Krejcar *et al.*, 2015. A combined negative selection algorithm-particle swarm optimization for an email spam detection system. *Eng. Appl. Artif. Intell.*, 39: 33-44.
- Lai, C.C., 2007. An empirical study of three machine learning methods for spam filtering. *Knowledge-Based Syst.*, 20: 249-254.
- Onan, A., S. Korukoglu and H. Bulut, 2016. A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification. *Expert Syst. Appl.*, 62: 1-16.
- Shen, J., R.H. Deng, Z. Cheng, L. Nie and S. Yan, 2015. On robust image spam filtering via comprehensive visual modeling. *Pattern Recognit.*, 48: 3227-3238.
- Soranamageswari, M. and C. Meena, 2010. Statistical feature extraction for classification of image spam using artificial neural networks. *Proceedings of the 2010 2nd International Conference on Machine Learning and Computing (ICMLC)*, February 9-11, 2010, IEEE, Bangalore, India, ISBN:978-1-4244-6007-6, pp: 101-105.

Design of a Mechatronic Device for Measuring Stump Stresses on Transfemoral Amputees

Edilberto Mejia Ruda, Oscar F. Avile and Mauricio F. Mauledoux
Mechatronics Engineering Program, Faculty of Engineering,
Militar Nueva Granada University, Bogota, Colombia

Abstract: This study is devoted to the design of process for a low-cost mechatronic device intended for measuring the lower limb amputee stresses. This development is based on a biomechanical analysis and the force data obtained in a gait laboratory whereof a force and stress analysis over the stump-socket interface could be done. Later, gait critical conditions are simulated for acquiring the theoretical results needed to do a sensors selection (Flexiforce-Teckscan®). Once the selection is done, a sensors mesh is presented as a measuring device as well as the electronics for signal conditioning and the software for results analysis and presentation. This device is calibrated and validated by means. Finally, the device assembly is done with commercial prosthetic elements in order to proceed with the critical conditions tests and to get the experimental data. The data is compared with the theoretical results as the designed device validating stage.

Key words: Socket, stump, stress, biomechanics, interface, assembly

INTRODUCTION

The rehabilitation process after an amputation is very important for victims and physicians play a key role. During rehab, patients use prosthetic devices and the socket is the part of them that allows transferring forces of the leg stump to the rest of the device. Therefore, this proposal shows a tool for diagnosis, through a device that measures the pressure on the leg socket-stump interface in order to improve rehab methods and support the best diagnosis for physicians and patients.

The gait could be described as a dynamic system where the reaction force and the contact surface change according to the patient cadence. During gait, from the back of heel to the lifting of forefoot there are forces widely higher than corporal weight and also generate a foot plantar pressure more elevated than the one at standing on natural and relaxed position (Lacuesta *et al.*, 2005).

Several prosthetics researches have been developed on models which describe the behavior of socket stump through pressure analysis. Distributed pressure identification on the socket stump interface was proposed for people with lower limb amputation (above knee) they developed a socket model through Finite Element Analysis (FEA) (Tanaka *et al.*, 1997). The stress

identification was formulated for minimizing the elastic energy the most plausible stress was sought based on watched deformation.

A pressure measurement device was designed for the socket of 48 trans-tibial amputees, studying the behavior of the silicone as interface between stump and socket, analyzing statistically and concluding that pressure exerted by the stump over socket depends on the alignment made by physicians which allow defining the pressures on the front and the back of the residual member. The equipment used for pressure measurement was F-scan sensors from Tekscan. The sensor network is composed by four arrays distributed in the socket which carried out a biomechanical analysis that determined that the anterior, posterior and lateral surfaces showed higher pressure (Dumbleton and Buis, 2009).

Dynamic of the interfase: The devices used to measure pressure work with the mechanical concept of forces and moments. In this case of study, forces are produced by a prosthetic device with a rigid element (e.g., prosthetic foot with the ground where reaction forces and moments are generated by contact with the ground). This forces are reflected on the interface and therefore on the person. The dynamic event analysis is based on Newton's second law, with the calculation of forces and moments. By Jia *et al.* (2004), Fig. 1 shows the gait force and moment diagram at

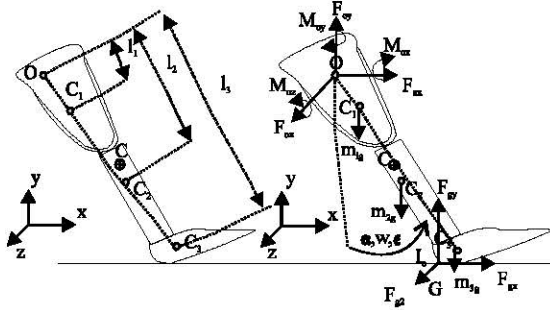


Fig. 1: Diagram of forces and moments

the back of heel phase on 3D Cartesian planes (x, y, z). The knee rotation as the sum of moments with respect to the origin O (Eq. 1):

$$M_{oz} - m_1 g_{i1} \sin b - m_2 g_{i2} \sin b - m_3 g_{i3} \sin b + F_{gx} y_G + F_{gy} y_G = I_o f \tag{1}$$

Where:

- b = The angular displacement on sagittal plane
- m₁, -m₃ = The stump, socket-tube and prosthetic foot masses, respectively
- l₁-l₃ = The distances from center of mass to origin

The sum of x and y moments regarding O (Eq. 2 and 3):

$$M_{ix} + F_{gz} y_G + F_{gy} z_G = 0 \tag{2}$$

$$M_{iy} + F_{gz} x_G + F_{gx} z_G = 0 \tag{3}$$

Equation 4-6 show the sum of forces on the different coordinated axis (x, y, z):

$$F_{ox} + F_{gx} = (m_1 + m_2 + m_3)(r_f \cos b - r_{p2} \sin b) \tag{4}$$

$$F_{oy} + F_{gy} - (m_1 + m_2 + m_3)g = (m_1 + m_2 + m_3)(r_f \sin b - r_{p2} \cos b) \tag{5}$$

$$F_{oz} + F_{gz} = 0 \tag{6}$$

Variable p is angular velocity, f is angular acceleration and r is the distance from the origin to the center of mass of entire model (Uriostegui *et al.*, 2007).

MATERIALS AND METHODS

Sensor network design: Based on Eq. 4-6, it was determined that the maximum force exerted on the interface was 26.1 N and corresponds to the force (Jia *et al.*, 2004) F_{ax}. Besides, the previous work results

established that gait pressures of a transfemoral amputee correspond to the 32% of the amputee bodyweight per cm. Hence, the proposal is to use the Flexiforce A301 sensor (120 N of nominal pressure), considering that the measuring area (0.2827 cm²) implies 9.04% of bodyweight. This sensor works as a variable resistance on an electric circuit, so that, its value is higher than 5 MΩ when there is no load applied and it decreases when applying (Mejia *et al.*, 2010).

A network sensor was designed to allow measuring pressure in the socket. This design has 5 force sensors which established a hotspot for estimate the pressure on the grid made of a sheet of acetate, due to its flexibility and low thickness, Fig. 2a shows dimensions of the grid.

The electronics for signal conditioning was based on manufacturer's recommendations which states that an operational amplifier must use at inverter configuration by adjusting the resistance R₂ to determine the gain configuration. Figure 2b shows the electronic network sensor circuit. A simulation was made to determine R₂ as 85 kΩ and thus, the sensor voltage output as (0-8 V).

$$(\sum y_u - n)^2 / \sigma^2 \cdot x_u^2 \tag{7}$$

An experiment was designed to assess the measurement error, based on statistical analysis, according to sensors hysteresis. This allows calibrating each sensor using ji-square method, Eq. 7 (Box *et al.*, 2008).

Test bench: A machine for structural testing of prosthetic devices was performed, following the international standard ISO 10328 which determines the critical load conditions which lower limb prostheses must be submitted. A pneumatic system with closed loop control system for testing was implemented, according to the standard requirements. The standard establishes different test arrangements which recreate the configurations in which the maximum pressure occurs (Jimenez *et al.*, 2012).

In this research, a user interface was designed for performing the tests according to standard where there was a heel-support static test setup, allowing reproducing a critical moment in the amputee gait. In order to understand the test, it was defined an amputee with 80 Kg of weight.

According to Lacuesta *et al.* (2005) there is a value of 120% of weight during gait. Therefore, the pneumatic actuator force was established to a maximum value equivalent to 960 N (100% is 80 kg of weight). Being this one of the parameters for the design of the control system of the test bench.

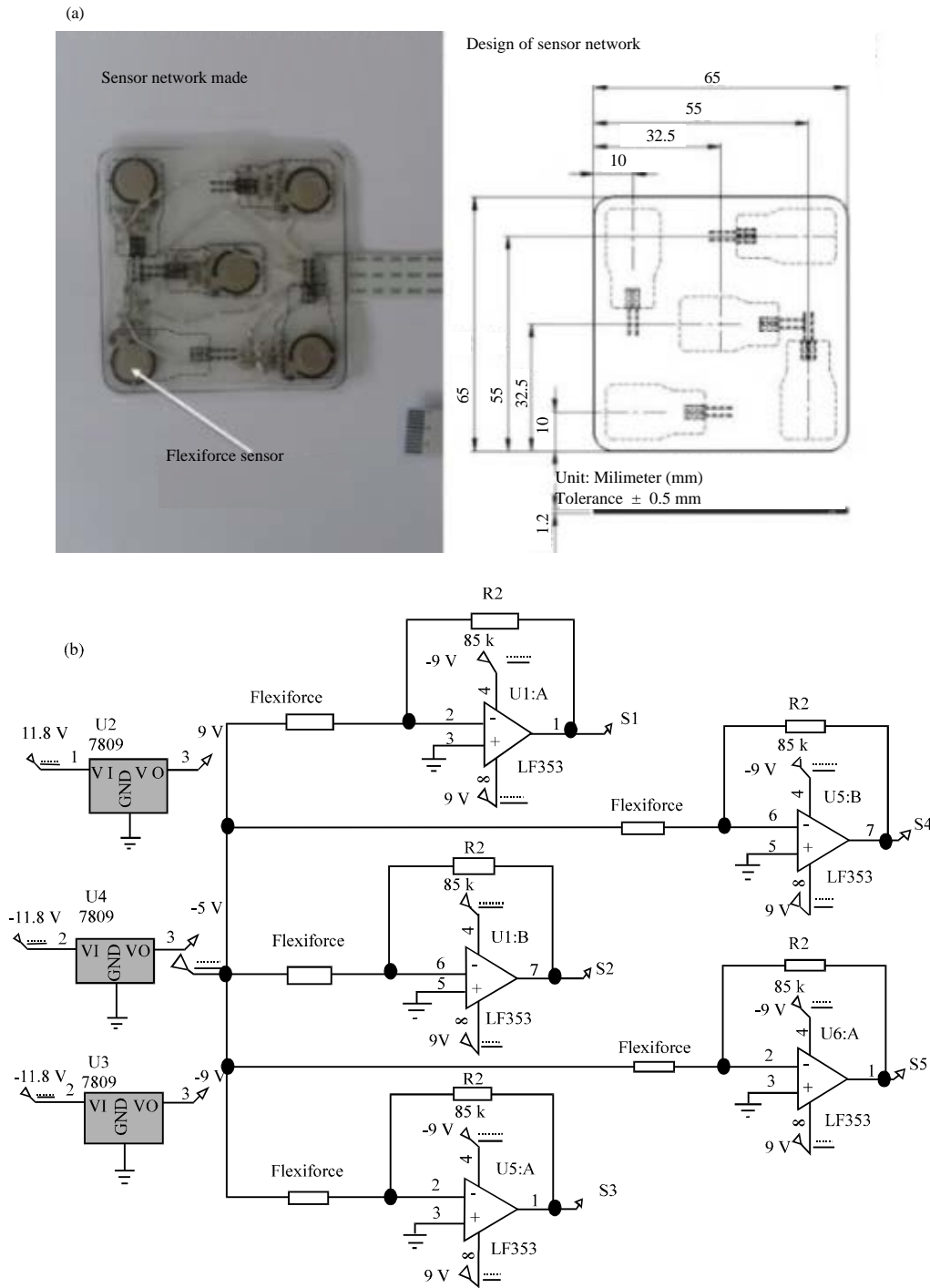


Fig. 2: Sensor scheme and electronics: a) Design and sensor network made and b) Electronic network sensor circuit

The close loop control was design base on advance control theory where was implemented a sliding mode control with differential flatness, considering the pressure in the pneumatic actuator as the nonlinear variable to controller.

For this variable, it's necessary realize an identification the plant for complete modelling test bench in two parts for simplification the process design controller. The first part was an open loop identification where the input is a step to the system is observed that the actuator does not

pause in a position nor in a pressure intermediate on the contrary, according to the magnitude of the step arrives faster to final position (mechanical limit) and the pressure of feeding of 600 kPa by the previous thing, it is not possible to realize an identification process in open loop of the system because the dynamics are only stabilized by mechanical limits (Monje, 2006).

To perform an identification process that reproduces the dynamics of the system, a pre-stabilization process is performed where it is identified in closed loop using a PID controller by auto-tuning (Bueno, 2011).

The second part is closed loop identification where a PID is proposed with a tuning of the constants in order to achieve an intermediate pressure value, initially neglecting the establishment time, the steady state error and the type of response. The control loop is closed by feedback of the pressure difference in the actuator chambers to evaluate the error according to the desired pressure reference in Fig. 3. The data obtained for the identification of the plant are observed in Fig. 4.

The pre-stabilized system permit the identification with MATLAB® ident tool which provides a parametric model from the experimental input and output data of the system (off-line) to design an advanced control for the system identified as the real system, so that, the control signal can be used for both plants which will have a similar dynamic.

For the selection of the appropriate identification model, several parametric models of 4th and 5th order (according to the theoretical model of Jimenez *et al.* (2012)) were evaluated that consider both measurement noise and load disturbances, comparing the transient response, the relation of the output and the model of the waste. This type of perturbations and considering that its structure is not complex as the system, we use ARX process models which are autoregressive with exogenous input and are based on the method of least squares optimization (Rodriguez and Bordons, 2005). The representation of states of the linearized and identified model is shown in Eq. 8:

$$\dot{\tilde{x}} = \begin{bmatrix} -530.6 \times 10^4 & -8.4 \times 10^4 & -1.1 \times 10^6 & -3.8 \times 10^6 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u \quad (8)$$

$$\tilde{y} = [0.83e4 \quad 30.16e4 \quad -1.08e6 \quad 3.90e6] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

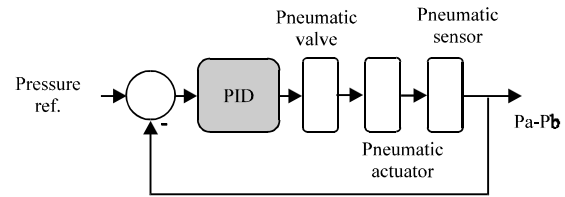


Fig. 3: Closed loop identification scheme

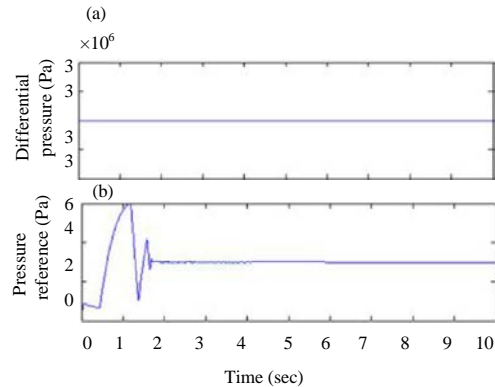


Fig. 4: a) Input and b) Output of the identified model

where, the states of the identified system are obtained which do not have a direct relation with the states of the real system allowing the output to describe the same dynamics.

Once the identified model has been obtained and the parameters of the amputee, the controller for the pressure difference and areas of actuator (Piston), according to Eq. 9 is designed.

$$F_0 = F_A - F_B = P_A A_A - P_B A_B \quad (9)$$

where, F_0 is the output force, A_A , F_A and P_A are the area of the piston, force and pressure in chamber 'A' and A_B , F_B and P_B are the area of the piston, force and pressure in the camera 'B', respectively.

The advanced control is made up of two techniques, the first is differential flatness where according to Morillo, the concept of differential flatness indicates that the states and inputs of a system can be expressed as functions of a set of differentially independent outputs, so that, problems such as stability, transition between points and trajectory tracking are easier to solve because they are not subject to any dynamics.

Therefore, a pressure controller is proposed by sliding modes in the space of the flat outputs whose reference variable is designed by means of a Bezier interpolation polynomial that allows to smooth the transition between desired values, so as to reduce sudden peaks in the control signal (Flores *et al.*, 2011). This polynomial describes a curve composed of a proportional relationship

between the initial and final points of the transition (y_1, y_2) and a Spline polynomial which passes through all the control points that define the shape of the curve and are created in the order of Polynomial defined by Eq. 10 and 11 (Pina, 2001):

$$P(t) = a_1 t^{n+1} + a_2 t^{n+2} + \dots + a_{n-1} t^{2n+1} \quad (10)$$

$$F^0 = y_2 + (y_2 - y_1) p\left(\frac{t}{T}\right) \quad (11)$$

Where:

- P(t) = The spline polynomial
- n = The order of the system
- t = The time
- T = The transition period
- a_i = Constants that are calculated according to the order of the system
- F = The controller reference

According to Ramirez and Agrawal (2004) and Conde *et al.* (2009), if a SISO system is controllable, it is differentially flat and its flat output is obtained from the last line of multiplication of the inverse matrix of controllability by the state vector. In this way, we obtain the flat output and its n-1 derivatives for the linear system of Eq. 12:

$$\begin{aligned} F = x_4; \dot{F} = x_3; \ddot{F} = x_2; \dddot{F} = x_1; F = \\ A_{1,1}x_1 + A_{1,2}x_2 + A_{1,3}x_3 + A_{1,4}x_4 \end{aligned} \quad (12)$$

For sliding mode control, the sliding surface based on the integral reconstruction of non-measured state variables is proposed as shown in Eq. 13, where it is also observed that a second integral term is added to compensate for the error in steady state:

$$\begin{aligned} \sigma = \ddot{\ddot{F}} + k_4 \ddot{\ddot{F}} + k_3 \dot{\ddot{F}} + k_2 \ddot{F} + k_1 \dot{F} + k_0 \int \int (F - F^0) \\ + k_1 \int (F - F^0) + k_0 \int \int (F - F^0) \end{aligned} \quad (13)$$

where, σ leads to a dynamics of the fifth order, so, when a characteristic polynomial of Routh-Hurwitz is proposed, a system of compatible equations is considered and the constants k_i are determined. From the resulting equation and from equation of mass flow of piston, we obtain the equivalent controller as the means to restrict the system dynamics to the sliding surface (Eq. 14):

$$\begin{aligned} u \in q = - (A_{1,1} + k_4) \ddot{\ddot{F}} - (A_{1,2} + k_3) \dot{\ddot{F}} - \\ (A_{1,3} + k_2) \ddot{F} - A_{1,4} \dot{F} - k_1 (F - F^*) + \int (F - F^*) \end{aligned} \quad (14)$$

In addition, the sliding surface is globally attractive by the continuous approximation by Ramirez (1993) for the

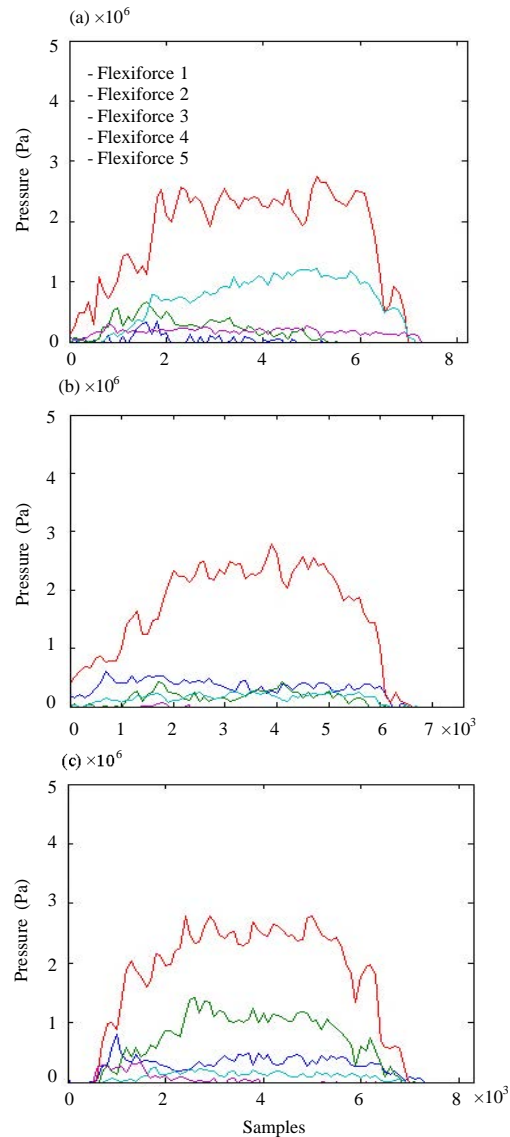


Fig. 5: Pressure chart of device: a) Test 1; b) Test 2 and c) Test 3

discontinuous dynamics of the controller, i.e., it is forced to comply with the dynamics of the equivalent control by Eq. 15 which depends on a constant 'L' tunable:

$$u_N = -L \text{sgn}(\sigma) \quad (15)$$

Finally, it will be assumed that, the control signal supplied to the pneumatic system and to the reference system is the sum of the equivalent control and the attractive control. With the above the test of the device to measure pressure was realized and thus obtain the results to determine if it works the design raised in this project, the test assembly shown in Fig. 5 was performed.



Fig. 6: Assembly on the test machine

RESULTS AND DISCUSSION

As a result, the pressure was recorded with the prosthetic device in the test bench with 3 static tests performed over the stump anterior surface and 960 N of applied force. For each test a socket for transfemoral amputee was used. Based on the foregoing an artificial stump was fabricated with human skin characteristics according to the state of the art (ballistic gel). Figure 6 shows the values recorded by the network which shows that central sensor has the greatest pressure.

The maximum pressure is generated in the anterior region in a 1 cm² area of pressure sensing at a maximum value of 32% of body weight. Therefore, it was established that pressure is equal to 2.51 MPa with an error of 0.12. Performing the calculation recorded by the device designed, it was established that the area of pressure is equal to 0.2827 cm² and the perceived weight percentage of the sensor network in this area is equivalent to 9.04% of bodyweight. The value obtained was 2.53 MPa with an approximate error of 0.13. The comparison between the results of the studies reported in the literature and the device designed, determine that the device is consistent and their measurements are accurate.

CONCLUSION

Added value to the device as designed is its low cost, also serve to support the medical diagnosis in terms of the pressures on the interface. It shall identify possible errors in the manufacture of the socket or pathologies of the amputation. These anomalies in the stump-socket interface can cause skin damage or interfere with the

rehabilitation of an amputee patient. It was also determined that the device, its physical properties will not affect the progress of an amputee when used.

IMPLEMENTATIONS

The implementation of a controller with advanced control and identification techniques allows manipulating and solving systems where the behavior of the system is non-linear. Also, laboratory test of the device for pressure measurement under conditions simulating the progress of a transfemoral amputee was provided, giving an approximation of the actual behavior in patients of the pathology described above.

ACKNOWLEDGEMENT

Thanks to Doctor J. P. Borrero of the Military Hospital and Militar Nueva Granada University, who supported the research.

REFERENCES

- Box, G.E., J.S. Hunter and W.G. Hunter, 2008. [Statistics for Researchers Design, Innovation and Discovery]. 2nd Edn., Barcelona Publisher, New Braunfels, Texas, ISBN:978-84-291-5044-5, Pages: 629 (In Spanish).
- Bueno, A.M., 2011. [Experimental identification of systems]. Master Thesis, University of Alicante, Alicante, Spain. (In Spanish)
- Conde, E.C., F.B. Carbajal, A.B. Ortega and H.M. Azua, 2009. Sliding mode and generalized PI control of vehicle active suspensions. Proceedings of the IEEE Joint Conference on Control Applications (CCA) and Intelligent Control (ISIC), July 8-10, 2009, IEEE, St. Petersburg, Russia, ISBN:978-1-4244-4601-8, pp: 1726-1731.
- Dumbleton, T. and A. Buis, 2009. Dynamic interface pressure distributions of two transtibial prosthetic socket concepts. *J. Rehabil. Res. Dev.*, 46: 405-415.
- Flores, J.L., A.A. Garcia and A.O. Molina, 2011. [Soft start for a CD engine via. a CD-CD drive converter (In Spanish)]. *Eng. Res. Technol.*, 12: 137-148.
- Jia, X., M. Zhang and W.C. Lee, 2004. Load transfer mechanics between trans-tibial prosthetic socket and residual limb-dynamic effects. *J. Biomech.*, 37: 1371-1377.

- Jimenez, S., O. Caldas, E. Mejia, J.C. Hernandez and O. Aviles, 2012. Modeling and control of destructive test equipment for lower limb prosthesis. Proceedings of the 1st International Conference on Advanced Mechatronics, Design and Manufacturing Technology AMDM, September 5-7, 2012, Technological University of Pereira, Pereira, Colombia, pp: 463-468.
- Lacuesta, J.S., J. Prat and J.V. Hoyos, 2005. [Biomechanics of the Normal and Pathological Human March]. Instituto de Biomecanica de Valencia, Valencia, Spain, ISBN:9788495448125, Pages: 444 (In Spanish).
- Mejia, E., J. Mejia, J. Delgado, O. Aviles and D. Amaya, 2010. [Transducer device for measuring force and speed in a task of scalpel cutting]. *Revista EIA*, 14: 147-160, (In Spanish).
- Monje, C.A., 2006. Design methods of fractional order controllers for industrial applications. Ph.D Thesis, University of Extremadura, Badajoz, Spain.
- Ramirez, H.S., 1993. A dynamical variable structure control strategy in asymptotic output tracking problems. *IEEE. Trans. Autom. Control*, 38: 615-620.
- Ramirez, S. and S.H. Agrawal, 2004. *Differentially Flat Systems*. 1st Edn., Marcel Dekker, New York, USA., ISBN:0-8247-5470-0, Pages: 422.
- Rodriguez, D. and C. Bordons, 2005. [Control engineering notes: Analysis and control of systems in state space-Identification of systems]. Master Thesis, University of Sevilla, Seville, Spain.
- Tanaka, M., Y. Akazawa, A. Nakagawa and I. Kitayama, 1997. Identification of pressure distribution at the socket interface of an above-knee prosthesis. *Adv. Eng. Software*, 28: 379-384.
- Uriostegui, I.Q., A.I. Perez, A. Alessi, G. Rodriguez and L. Nunez *et al.*, 2007. [Analysis of Pressures Exercised Inside the Socket During the Cycle of the March in Patients with Transfemoral Amputation]. In: 4th Latin American Congress on Biomedical Engineering and Bioengineering Solutions for Latin America Health, Muller-Karger, C., S. Wong and A. Cruz (Eds.). Springer, Berlin, Germany, ISBN:978-3-540-74470-2, pp: 774-778 (In Spanish).

Thermal Performance Characteristics of Three Sides 60° Inclined Wire Roughened Solar Air Heater

Ravi Kumar, L. Prasad and B.N. Prasad
Department of Mechanical Engineering, National Institute of Technology,
831014 Jamshedpur, India

Abstract: Three sides 60° wire roughened solar air heater together with on side 60° wire roughened solar air heater has been investigated. Effect of roughness and flow parameter, p/e , e/D and Re on thermal performance of the collectors has been represented. Three sides 60° wire roughened solar air heater has been found to be thermally more efficient by an amount of about 14-46% than that of one side roughened one in the range of parameters investigated.

key words: Artificial roughness, angle of attack, collector efficiency factor, collector heat removal factor, parameter, investigated

INTRODUCTION

The renewable energy sources playing a vital role to fulfil the required energy demand in industrialization and economical progress. Plenty of research works in the area of solar air heaters where enhancement of the efficiency of solar air heater by using different types of solar air heaters and various roughness elements and alignments are available in literature.

Generally thermal efficiency of solar air heaters is poor due to low convective heat transfer coefficient between the absorber plate and air flowing through the flow duct. The reason behind this is the formation of laminar viscous sub layer, adjacent to the wall which resists heat transfer. Thus, to enhance the heat transfer, provision of artificial roughness on the flow sides of solar air heater is introduced.

Use of artificial roughness on heat transferring surface creates artificial turbulency in the viscosity of the wall. Artificially roughened solar air heater is considered better to increase the heat transfer coefficient, since, artificial roughness elements break the laminar-sub-layer and reduce the thermal resistance. But simultaneously it increases the frictional losses in solar duct. Therefore, the requirement of turbulence to be created using artificial roughness should remain in the region very close to the heat transferring surface, i.e. within the laminar-sub-layer thickness only to reduce the friction loss. This can be possible by using the height of roughness elements to be small in comparison with the duct dimension (Patil *et al.*, 2012), limited to the laminar-sub-layer thickness of hydrodynamic boundary layer (Prasad and Saini, 1991).

Roughness geometries have been classified in four categories for solar air heater (Bhushan and Singh, 2011). Heat transfer and friction factor in artificially roughened solar air heater duct having these four types of roughness geometries have been found experimentally by various researchers (Prasad and Saini, 1988; Gupta *et al.*, 1993; Saini and Saini, 1997; Jaurkar *et al.*, 2006; Karmare and Tikekar, 2007; Singh *et al.*, 2011; Lanjewar *et al.*, 2011; Bhusan and Singh, 2011). Also, the thermal performance and heat transfer coefficient were reported by Chamoli *et al.* (2012a, b), Gawande *et al.* (2014), Saurav and Bartaria (2013), Shakya *et al.* (2013), Saini *et al.* (2007), Yadav and Bhagoria (2013, 2014a, b), Chabane *et al.* (2014), Prasad *et al.* (2015) and Adeyemo and Adeoye (2008). Nusselt number and friction factor correlations were developed by using these experimental data. The thermal performance in double flow solar air heater with aluminium cans has been investigated by (Ozgen *et al.*, 2009). But all these works remained limited to only one top side artificial roughness. Recently the research has been extended and experimentally investigated in three sides roughened solar air heaters, i.e., roughness used in top side as well as both side walls (Prasad *et al.*, 2014, 2015; Behura *et al.*, 2016). It has been found that the three sides roughened solar air heater is more effective and efficient than the one side roughened solar air heater. The roughness has been also used at different angle of attack and 60° angle of attack has been found most effective to enhance the thermal performance of only top side roughened solar air heaters by Behura *et al.* (2016). This study reports the results of

experimentally collected data in actual outdoor conditions on thermal performance of three sides 60° inclined wire roughened solar air heaters.

MATERIALS AND METHODS

Investigation: Figure 1 shows the schematic diagram of a typical three sides artificially roughened solar air heater duct model having three sides covered by glass. The duct height is 25 mm and the width is 200 mm, the glass covers are 4 mm thick ordinary glass. The absorber plates are galvanised-iron sheets of 20 standard wire gauge. Fluid temperature measured by digital thermometer of least count 0.1°C and the plate temperatures were measured by the thermocouples of 28 standard wire gauge. Pyranometer is used to measure the solar radiations. The pressure drop is measured along the test study by the help of the multi tube manometer.

The roughness element height is selected using the equations under used in (Prasad and Saini,1991) because the roughness height should fall within the laminar sub-layer thickness of hydrodynamic boundary layer and it won't protrude beyond it:

$$\frac{\partial'}{D} = \frac{58}{Re^{\frac{7}{8}}} \quad (1)$$

$$\frac{\partial'}{D} = \frac{140.185}{Re} \quad (2)$$

The data were collected simultaneously on both three sides and one side roughened solar air heaters. The experimental values of the heat transfer coefficient for the both one side and three side 60° inclined wire roughened solar air heaters have been found by Eq. 3 as given:

$$C_p (t_o - t_i) = hA_c (\bar{t}_p - t_f) \quad (3)$$

Which may be further used to obtain the values of Nusselt number for the both solar air heaters by Eq. (4) under:

$$Nu = \frac{hD}{k} \quad (4)$$

Artificial roughness have been provided on the fluid flow side of the absorber plate normal to the fluid flow direction at varying values of relative roughness pitch, p/e in the range of 10-20, relative roughness height, e/D in the range of 0.0180-0.0248. Flow Reynolds number, Re , varied in between 2000-16000 and angle of attack, α is 60°.

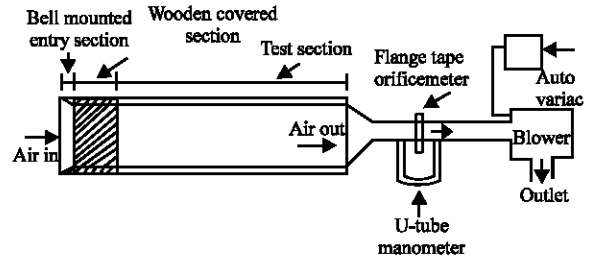


Fig. 1: Schematic diagram of three sides artificial roughened and glass covered solar air heater

RESULTS AND DISCUSSION

Thermal performance has been represented on the basis of outlet air temperature for solar air heaters, operating without recycling of air (Gupta *et al.*, 1997). Equation 5 has formed the basis of representation of thermal efficiency:

$$(th = F_R (\tau\alpha) - F_R U_L (t_o - t_\alpha)) / I \quad (5)$$

This could be written as Eq. 6 based on fluid outlet temperature:

$$(th = F_o (\tau\alpha) - F_o U_L (t_o - t_\alpha)) / I \quad (6)$$

For inlet temperature, t_i , to be the same as ambient temperature, t_α , Eq. 6 takes the form as of Eq. 7 under:

$$(th = F_o (\tau\alpha) - F_o U_L (t_o - t_i)) / I \quad (7)$$

And the instantaneous thermal efficiency given by Eq. 8:

$$(th = \dot{m} C_p (t_o - t_i)) / I \quad (8)$$

Mass flow rate is the strong function which is highly recommended in order to obtain reasonably reliable information with regard to the thermal performance of solar air heaters. Each mass flow rate results in an efficiency curve (Gupta and Garg, 1967; ASHARE Standards, 1977; Reddy *et al.*, 1980; Biondi *et al.*, 1988). Fig. 2 shows the typical performance characteristics of one side and three sides 60° inclined wire roughened solar air heaters at various mass flow rates and for the values of p/e , equal to 10 and e/D , equal to 0.0248.

The lines \dot{m}_1 - \dot{m}_5 shown in Fig. 1 are the efficiency curves for each mass flow rate based on Eq. 8. The upper points on these curves correspond to the three sides 60° inclined wire roughened collectors and the

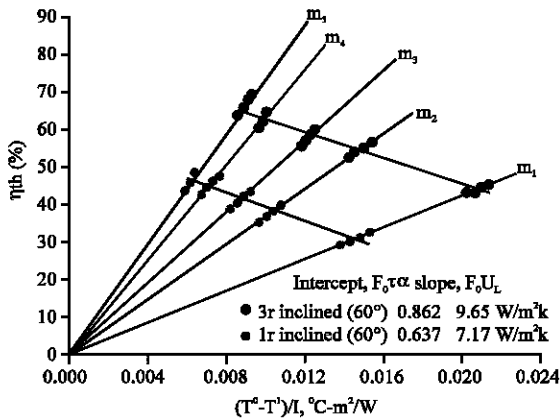


Fig. 2: Performance characteristic of solar air heaters

lower ones correspond to one side 60° inclined wire roughened collectors. The respective values of the intercept and slope have been found out and shown in this Fig. 1.

The values of the thermal performance parameters, F_R , and F' have been worked out by utilizing the following Eq. 9-11, written under (Bernier and Plett, 1988):

$$F_R(\tau\alpha) = F_o(\tau\alpha) \left[\frac{\dot{m}C_p / A_c}{\dot{m}C_p / A_c + F_o U_L} \right] \quad (9)$$

$$F_R U_L = F_o U_L \left[\frac{\dot{m}C_p / A_c}{\dot{m}C_p / A_c + F_o U_L} \right] \quad (10)$$

$$F' = GC_p \left[\ln \left(\frac{F_o U_L}{F_R U_L} \right) / U_L \right] \quad (11)$$

The respective values of the collector performance parameters F_R and F' have been plotted in Fig. 3-6 to see the effect of the roughness and flow parameters on F_R and F' in three sides and one side 60° inclined roughened solar air heaters. It could be seen from these Fig. 3-6 that the values of both the parameters F_R and F' is higher for three sides 60° inclined wire roughened solar air heater than those of one side 60° inclined wire roughened solar air heater by an amount in the range of 14-46%. It could also be attributed from these figures that the rate of increase in the values of F_R and F' is more at higher mass flow rates. As also, increasing value of the roughness parameter, p/e , results in decreasing values of both F_R and F' whereas increasing values of the roughness parameter, e/D , results in increasing values of both F_R and F' .

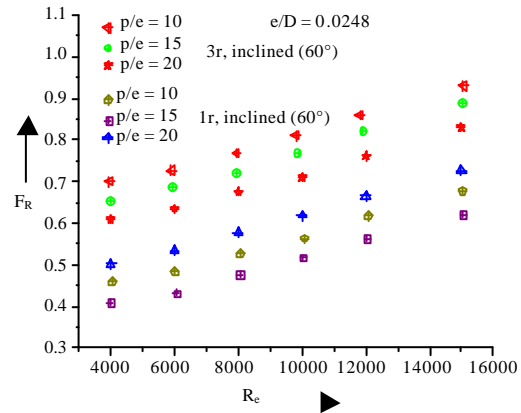


Fig. 3: Effect of p/e on collector heat removal factor in solar air heaters

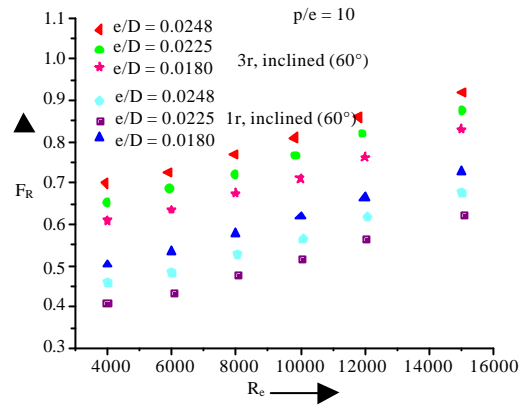


Fig. 4: Effect of e/D on collector heat removal factor in solar air heaters

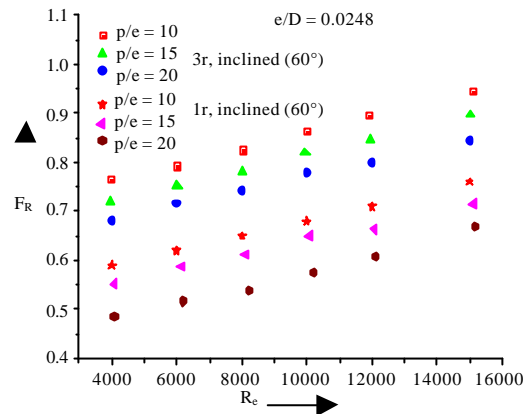


Fig. 5: Effect of p/e on collector efficiency factor in solar air heaters

Which further have been used to give the values of $F_R(\tau\alpha)$ and $F_R U_L$ with the values obtained for the collector performance parameter $F_R(\tau\alpha)$ and $F_R U_L$ for the respective solar air heaters and based on the conventional thermal

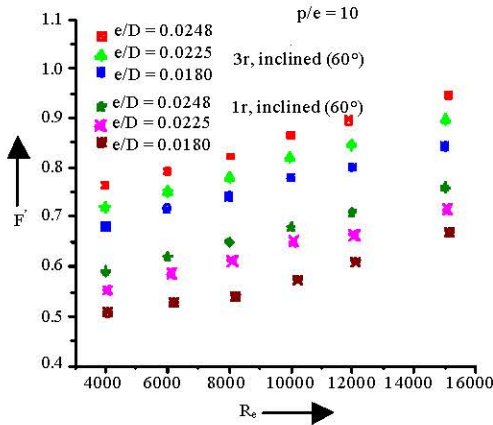


Fig. 6: Effect of e/D on collector efficiency factor in solar air heaters

performance equations the following thermal performance Eq. 12 and 13, for the respective solar air heaters have been obtained as under:

$$\eta_{th} (3r, inclined) = 0.892(\tau\alpha) - \left[0.892U_L \left(\frac{t_i - t_a}{I} \right) \right] \quad (12)$$

$$\eta_{th} (1r, inclined) = 0.678(\tau\alpha) - \left[0.678U_L \left(\frac{t_i - t_a}{I} \right) \right] \quad (13)$$

CONCLUSION

Three sides 60° inclined wire roughened solar air heater perform better than one side 60° inclined wire roughened solar air heater under the same operating conditions of roughness and flow parameters.

The values of the heat removal factor, F_R and plate efficiency factor, F' have been found to increase in the range of 14-46% over than of one side 60° wire roughened solar air heaters. The conventional thermal performance equations for three sides 60° wire roughened and one side 60° wire roughened solar air heaters have been derived as under:

$$\eta_{th} (3r, inclined) = 0.892(\tau\alpha) - \left[0.892U_L \left(\frac{t_i - t_a}{I} \right) \right] \quad (14)$$

$$\eta_{th} (1r, inclined) = 0.678(\tau\alpha) - \left[0.678U_L \left(\frac{t_i - t_a}{I} \right) \right] \quad (15)$$

NOMENCLATURE

- A_c = Collector area (m²)
- C_p = Specific heat at constant pressure of air (kJ/kg K)
- D = Hydraulic diameter of solar air heater duct (m)

- e/D = Relative roughness height
- F' = Collector efficiency factor
- F_R = Collector heat removal factor
- F_o = Collector heat removal factor based on outlet temperature (t_o)
- G = Mass flow rate per unit collector area (kg/sm²)
- h = Convective heat transfer coefficient (W/m²K)
- I = Intensity of solar radiation (W/m²)
- k = Thermal conductivity of collector material (W/m K)
- L = Collector Length (m)
- l = Length of the test section (m)
- \dot{m} = Mass flow rate (kg/sec)
- Nu = Nusselt number
- p/e = relative roughness pitch
- Re = Reynolds number
- t_a = Ambient temperature (°C)
- t_o = Outlet air temperature (°C)
- t_i = Inlet air temperature (°C)
- \bar{t}_p = Average plate temperature (°C)
- \bar{t}_f = Average fluid temperature (°C)
- U_L = Overall heat transfer coefficient (W/m² K)
- α = Angle of attack, degrees
- $\tau\alpha$ = Transmittance-absorptance product
- η_{th} = Collector thermal efficiency
- δ' = Laminar sub-layer thickness (m)

SUFFIX

- 1r = Top side inclined roughened solar air heater
- 3r = Three sides inclined roughened solar air heater

REFERENCES

ASHARE Standards, 1977. Methods of testing to determine the thermal performance of solar collectors. ASHARE Standards, Atlanta, Georgia.

Adeyemo, S.B. and O.F. Adeoye, 2008. Comparative performance of two discrete solar collectors. J. Eng. Appl. Sci., 3: 228-232.

Behura, A.K., B.N. Prasad and L. Prasad, 2016. Heat transfer, friction factor and thermal performance of three sides artificially roughened solar air heaters. Sol. Energy, 130: 46-59.

Bernier, M.A. and E.G. Plett, 1988. Thermal performance representation and testing of air solar collectors. J. Sol. Energy Eng., 110: 74-81.

Bhushan, B. and R. Singh, 2011. Nusselt number and friction factor correlations for solar air heater duct having artificially roughened absorber plate. Sol. Energy, 85: 1109-1118.

Biondi, P., L. Cicala and G. Farina, 1988. Performance analysis of solar air heaters of conventional design. Sol. Energy, 41: 101-107.

Chabane, F., N. Moumami and S. Benramache, 2014. Experimental study of heat transfer and thermal performance with longitudinal fins of solar air heater. J. Adv. Res., 5: 183-192.

Chamoli, S., N.S. Thakur and J.S. Saini, 2012a. A review of turbulence promoters used in solar thermal systems. Renewable Sustainable Energy Rev., 16: 3154-3175.

- Chamoli, S., R. Chauhan, N.S. Thakur and J.S. Saini, 2012b. A review of the performance of double pass solar air heater. *Renewable Sustainable Energy Rev.*, 16: 48-492.
- Gawande, V.B., A.S. Dhoble and D.B. Zodpe, 2014. Effect of roughness geometries on heat transfer enhancement in solar thermal systems: A review. *Renewable Sustainable Energy Rev.*, 32: 347-378.
- Gupta, C.L. and H.P. Garg, 1967. Performance studies on solar air heaters. *Sol. Energy*, 11: 25-31.
- Gupta, D., S.C. Solanki and J.S. Saini, 1993. Heat and fluid flow in rectangular solar air heater ducts having transverse rib roughness on absorber plates. *Sol. Energy*, 51: 31-37.
- Gupta, D., S.C. Solanki and J.S. Saini, 1997. Thermohydraulic performance of solar air heaters with roughened absorber plates. *Sol. Energy*, 61: 33-42.
- Jaurker, A.R., J.S. Saini and B.K. Gandhi, 2006. Heat transfer and friction characteristics of rectangular solar air heater duct using rib-grooved artificial roughness. *Sol. Energy*, 80: 895-907.
- Karmare, S.V. and A.N. Tikekar, 2007. Heat transfer and friction factor correlation for artificially roughened duct with metal grit ribs. *Intl. J. Heat Mass Transfer*, 50: 4342-4351.
- Lanjewar, A., J.L. Bhagoria and R.M. Sarviya, 2011. Heat transfer and friction in solar air heater duct with W-shaped rib roughness on absorber plate. *Energy*, 36: 4531-4541.
- Ozgen, F., M. Esen and H. Esen, 2009. Experimental investigation of thermal performance of a double-flow solar air heater having Aluminium cans. *Renewable Energy*, 34: 2391-2398.
- Patil, A.K., J.S. Saini and K. Kumar, 2012. A comprehensive review on roughness geometries and investigation techniques used in artificially roughened solar air heaters. *Intl. J. Renewable Energy Res.*, 2: 1-15.
- Prasad, B.N. and J.S. Saini, 1991. Optimal thermohydraulic performance of artificially roughened solar air heaters. *Sol. Energy*, 47: 91-96.
- Prasad, B.N., A. Kumar and K.D.P. Singh, 2015. Optimization of thermo hydraulic performance in three sides artificially roughened solar air heaters. *Sol. Energy*, 111: 313-319.
- Prasad, B.N., A.K. Behura and L. Prasad, 2014. Fluid flow and heat transfer analysis for heat transfer enhancement in three sided artificially roughened solar air heater. *Sol. Energy*, 105: 27-35.
- Prasad, B.N. and J.S. Saini, 1988. Effect of artificial roughness on heat transfer and friction factor in a solar air heater. *Sol. Energy*, 41: 555-560.
- Reddy, T.A. and C.L. Gupta, 1980. Generating application design data for solar air heating systems. *Sol. Energy*, 25: 527-530.
- Saini, R.P. and J.S. Saini, 1997. Heat transfer and friction factor correlations for artificially roughened ducts with expanded metal mesh as roughness element. *Intl. J. Heat Mass Transfer*, 40: 973-986.
- Saini, R.P. and S.K. Singal, 2007. A review on roughness geometry used in solar air heaters. *Sol. Energy*, 81: 1340-1350.
- Saurav, S. and V.N. Bartaria, 2013. Heat transfer and thermal efficiency of solar air heater having artificial roughness: A review. *Intl. J. Renewable Energy Res.*, 3: 498-508.
- Shakya, U., R.P. Saini and M.K. Singhal, 2013. A review on artificial roughness geometry for enhancement of heat transfer and friction characteristic on roughened duct of solar air heater. *Intl. J. Emerg. Technol. Adv. Eng.*, 3: 279-287.
- Singh, S., S. Chander and J.S. Saini, 2011. Heat transfer and friction factor correlations of solar air heater ducts artificially roughened with discrete V-down ribs. *Energy*, 36: 5053-5064.
- Yadav, A.S. and J.L. Bhagoria, 2013. A CFD (Computational Fluid Dynamics) based heat transfer and fluid flow analysis of a solar air heater provided with circular transverse wire rib roughness on the absorber plate. *Energy*, 55: 1127-1142.
- Yadav, A.S. and J.L. Bhagoria, 2014b. A CFD based thermo-hydraulic performance analysis of an artificially roughened solar air heater having equilateral triangular sectioned rib roughness on the absorber plate. *Intl. J. Heat Mass Transfer*, 70: 1016-1039.
- Yadav, A.S. and J.L. Bhagoria, 2014a. A numerical investigation of square sectioned transverse rib roughened solar air heater. *Intl. J. Thermal Sci.*, 79: 111-131.

Gravitational Search Algorithm for Effective Selection of Sensitive Association Rules

¹P. Gayathiri and ²B. Poorna

¹Department of Computer Science, Bharathiar University, 641046 Coimbatore, Tamil Nadu, India

²SSS Jain College for Women, T. Nagar, Chennai, Tamil Nadu, India

Abstract: Association rule mining is the process of identifying the frequent items and associative rules in a market basket data analysis for large set of transactional databases. Association rules are employed in different data mining applications including web mining, intrusion detection and bioinformatics. In recent years, it has been seen tremendous advances in the ability to perform effective association rule mining. It causes the need of sensitive rule selection to enhance the privacy preservation of data transactions. A new technique called Gravitational Search based Sensitive Rule Selection (GS-SRS) is proposed to select the sensitive rules is to be hidden for improving the privacy preservation of transactional database. The GS-SRS technique is introduced to select the sensitive rules from the derived association rule through conditional probability. The sensitive rules contain the sensitive information of transactional database. Sensitive rules are identified for many applications. One of the applications of sensitive rule identification is to preserve the privacy of an organization or an individual by hiding these rules. The GS-SRS technique initially generates association rules through identifying the frequent items by using conditional probability-based association rules and support count and confidence value. Next, GS-SRS technique used gravitational search algorithm that lists the cohesive and non-cohesive items for the given transactional database in shorter time than the conventional means of feature selection using rough set technique. The association rule containing more cohesive items is selected as sensitive rule. A threshold value is used to select sensitive rules with the convergence of cohesive items and divergence of the non-cohesive items. Finally, the sensitive rules selected are hidden for preserving the privacy of transactional database. The experiments have been carried out on transaction database using four data sets and compared with state of art existing techniques. The experiment results show that the proposed GS-SRS technique is able to improve the accuracy of privacy preservation with minimum execution time when compared to state-of-the-art works.

Key words: Association rule mining, gravitational search, sensitive rule selection, conditional probability, cohesive items, non cohesive items

INTRODUCTION

In data mining, association rule mining technique identifies the frequent items and associative rules for a large size of transactional databases. The conditional probability of the transactional data items are evaluated to present the associative rule which indicates the scenario of the buying habits and products in demand. Privacy preservation of transactional data has been considered as an important research problem in recent applications. However, disclosing the sensitive associative rule of the transactional data item may expose the confidentiality and privacy of the organizations and individuals. Identifying and selecting sensitive association rule of a transactional database is a crucial process for preserving the privacy of sensitive information.

Association rule hiding technique hides the sensitive association rules generated from the transactional data items. The privacy preservation with data mining needs to ensure hiding sensitive information. Hiding sensitive rules should be made with minimal side effects. Sensitive association rule selection play major role in effectiveness of association rule hiding technique. But the association rules hiding technique indirectly generates some data items which are not exist in original database and hide certain data items which are not sensitive which in turn affects the privacy of rules and affects the utility of the data mining results. This problem is addressed by developing the new technique called Gravitational Search based Sensitive Rule Selection (GS-SRS) to improve the selection of sensitive rules to enhance the Association Rule Hiding (ARH) technique using Gravitational Search

Algorithm (GSA). This research is an optimization process for selection of sensitive association rules from the possible set of association rule of a transactional database. The gravitational algorithm determines how appropriate association rules for hiding have been selected.

In this study, a Gravitational Search based Sensitive Rule Selection (GS-SRS) technique is designed to select the sensitive rules is to be hid and thereby preserving the privacy of sensitive information's in business transactions. The identification of frequent items based on support count which results in the generation of association rules. The GS-SRS technique is a sensitive rule selection-based technique that efficiently reduces the execution time for sensitive rule selection, the rule are governs the organization's and individual's privacy. GS-SRS also incorporates gravitation search algorithm to select the sensitive rules from derived association rule, enabling reduced memory consumption for measuring cohesive and non-cohesive items. After that, the sensitive rules selected are hid with aim of preserving the privacy of transactional database. Experimental results shows that, the gravitation search based technique can improve accuracy of privacy preservation over existing selection methods. The propose algorithm select optimum number of association rules which has threat to privacy of data owners.

Literature review: In recent trends, numerous researches have been done in association rule hiding for privacy preserving data mining. Group Incremental Approach using Rough Set (GIA-RS) (Liang *et al.*, 2014) is used as a group incremental feature selection algorithm to identify the new feature subset in a short time when multiple objectives were added to a decision table. Incremental feature selection algorithm was based on the information entropy and it dealt with effective and efficient mechanism. The feature selection algorithm generated more feasible subset in a short period of time with increasing number of data. However, the rules extracted from dynamic dataset need to be updated in time to improve the selection of sensitive rules.

Locality Sensitive Hiding for Privacy Preserving (LSH-PP) (Karapiperis and Verykios, 2015) with a homomorphic matching technique identifies the candidate record pairs. The matching of pairs was implemented by a basic protocol which performs simple distance computations. Matching technique is mostly used for privacy-preserving record linkage. The performances of the distance-preserving properties were highly correlated due to anonymization format. The parameters of blocking

system in LSH-PP were selected in such a way to attain highest possible accuracy there by significantly reducing the possible running time. However, it failed to produce accurate results because of the anonymization format, failing to preserve the privacy of the individual.

Association rule hiding methodologies aim at sanitizing the original database (Gkoulalas-Divanis and Verykios, 2010). This methodology makes all the sensitive rules disappear from the sanitized database when the database is mined under the same or higher levels of support and confidence as the original database. Then all the non-sensitive rules that were mined from the original database should also be mined from its sanitized counterpart at the same or higher levels of confidence and support. At last no false rules also known as ghost rules should be produced when the sanitized database is mined at the same or higher levels of confidence and support. A false rule is an association rule that was not among the rules mined from the original database.

Sensitive association rule selection for association rule hiding has received considerable attention in recent years, especially in the context of transaction databases. For example, the work employed by Ghinita *et al.* (2011) Locality Sensitive Hashing (LSH) was applied to reduce the data utility and execution time by proposing two novel anonymization methods for sparse high dimensional data. However, the time complexity involved in anonymization method remains unaddressed. By Abbasi *et al.* (2011), it is shown that by applying Feature Relation Networks, using rule-based multivariate text feature selection, features were selected in a more computationally efficient manner. Despite, efficient in terms of computational effort, redundancy and relevancy remained unsolved. To address this issue by Maji and Pal (2010), fuzzy approximation was used during rule extraction that concentrated on non-redundant and relevant features. Similarly, another method addressed by Pervaiz *et al.* (2014) that used accuracy constrained privacy preservation mechanism aiming at improving the imprecise bounds.

Association rule hiding of a transaction database has been identified as a significant problem and probabilistic data management received a lot of attentions to deal with uncertain data. An optimized Monte Carlo algorithm (Serra and Spezzano, 2015) drastically minimized the number of iterations in probabilistic databases, however, little concentration were made towards attribute redundancy. To minimize the attribute redundancy, fuzzy rough set (Zhao *et al.*, 2010) designed a rule-based classifier model that improved the sensitivity of rules being generated. However, focused remained unsolved in

differentiating the data between positive and negative samples. By Zhou *et al.* (2010), Probabilistic Latent Semantic Analysis (PLSA) was designed using minimum information from the user and made efficient use of the positive and unlabeled data by applying iterative algorithm. Since, a better performance was achieved through PLSA and side effects was not discussed. Hiding Missing Artificial Utility (HMAU) (Lin *et al.*, 2014) algorithm was applied to reduce the number of side effects and number of deleted transactions. Another method using randomized response technique (Sun *et al.*, 2014) was applied to preserve the personalized privacy in frequent itemset mining.

Data mining techniques are used to mine useful information and knowledge from several databases. By Lin *et al.* (2014a, b), GA-based privacy preserving utility mining method was introduced to hide sensitive items. Yet, another GA-based approach was introduced by Lin *et al.* (2014a, b) for hiding sensitive information. However with the introduction of smart phones and web services, privacy of metadata remained a major concern. To address this issue by De Montjoye *et al.* (2014), a personal metadata management framework was introduced with the objective of reducing the execution time and dynamically protecting the personal data. Another Multi Objective Optimization (MOO) was designed by Cheng *et al.* (2015) that discovered useful relationships from shared data. This in turn minimized the confidence of sensitive rules.

Data publishing is extensively applied in the field of information sharing and scientific research and providing security for user's privacy. Privacy protection method for multiple sensitive attributes was introduced by Yi and Shi (2015) to improve the sensitive rules. Though improvement was observed in sensitive rule generation, privacy preservation in distributed environment remained a major breakthrough. A novel algorithm for privacy preserving in data mining in distributed environment was applied by Ouda *et al.* (2015) using Elliptic Curve Cryptography (ECC) and Diffie Hellman key exchange. The application of ECC ensured privacy preservation and also minimized the performance time. A comprehensive review on privacy preserving data mining was studied by Aldeen *et al.* (2015). Though several methods were presented to hide sensitive association rules, the modification in database made certain amount of side effects. Genetic algorithm to preserve privacy was applied by Shah and Asghar (2014) to preserve confidential information and counter side effects of lost rules. By Verykios *et al.* (2004), two fundamental approaches were designed with the aim of protecting sensitive rules from

disclosure. Besides, it presented three strategies and five algorithms for hiding a group of association rules which was characterized as sensitive.

A novel technique was designed by Gkoulalas-Divanis *et al.* (2010) for privacy preserving mining of association rules from outsourced transaction a database. But privacy vulnerabilities were remained unaddressed. A Hybrid Partial hiding algorithm (HPH) was presented by Zhu and Li (2013) for improving the privacy preservation of association rule mining. Privacy preserving association rule mining was developed by Keshavamurthy *et al.* (2013) for distributed databases with the application of genetic algorithm. A novel method was intended by Tassa (2014) for secure mining of association rules in horizontally distributed databases. However, computational cost was higher. An accuracy-constrained privacy-preserving access control framework was designed by Pervaiz *et al.* (2014) to achieve privacy requirements in relational data. The association rule hiding technique was intended by Ponde and Jagade (2014) for hiding the sensitive data or information in transaction database.

Based on the aforementioned techniques and methods in this research a gravitational search based sensitive rule selection technique is introduced with the objective of improving sensitive rules selection to enhance the business transactions.

MATERIALS AND METHODS

Design of gravitational search based sensitive rule selection: The Gravitational Search based Sensitive Rule Selection (GS-SRS) technique is designed to select the sensitive rules is to be hidden and to enhance the privacy preservation of sensitive information in transactional database. This study presents two models to select sensitive rules to enhance association rule hiding and to enhance the privacy preservation of business transactions. The first model generates the association rules with support count and confidence value using conditional probability-based association rule generation. On the other hand, the second model selects sensitive rules from generated rules using gravitational search based sensitive rule selection. The elaborate description of these two models is presented in the forthcoming sections.

Conditional probability-based association rule generation: Association rule mining, Ghinita *et al.* (2011) discovers items frequently occurring in a transactional database with the objective of producing significant

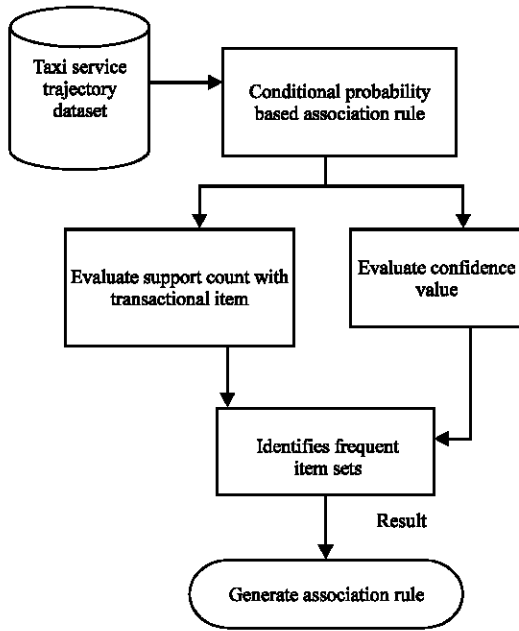


Fig. 1: Process of conditional probability based association rule generation

association rules that hold for the data. In GS-SRS, a transactional database using association rule mining identifies frequent items with support count and generates association rules with confidence value based on conditional probability. To preserve the privacy of an individual or an organization, sensitive rules are identified in the transactional database. In this work, Conditional probability-based association rule generation is used to identify the frequent items in transactional database for efficient association rule generation. Figure 1 shows the process involved in conditional probability-based association rule generation.

As shown in Fig. 1, conditional probability-based association rule generation process initially takes taxi service trajectory dataset as input and then evaluates the support count and confidence value with help of transactional item sets in the given data set. The frequent item sets are identified using evaluated support count values with the objective of generating the association rules. The GS-SRS technique uses conditional probability measure to estimate the support and confidence value and therefore identifies the frequent items. The conditional probability is a measure of probability of support and confidence value (for an itemset) given that another support and confidence (for another itemset) has occurred. Finally, it generates the association rule based on the identified frequent item sets with the aid of evaluated confidence value.

Let us consider a dataset, DS, a set of rules R over DS as well as a Sensitive Rule SR. The objective of conditional probability-based association rule generation is to identify a dataset DS such that when mining DS for generation of rules using same parameters as those used in the mining of DS, only the non-sensitive rules in R-SR are derived. The selection of sensitive rule influences the analytical strategies which governs the effectiveness of association rule hiding technique.

Let, $I = i_1, i_2, \dots, i_n$ denotes the set of items, for transaction $T = t_1, t_2, \dots, t_n$ where $T \in I$ with transaction id represented by TID, then the transaction T contains P, a set of items in P, if the following condition is said to be satisfied:

$$P \subseteq T \tag{1}$$

The association rule is an implication and is formulated as given as:

$$P \rightarrow Q, \text{ where } P, Q \subseteq I \text{ and } P \cap Q = \emptyset \tag{2}$$

From Eq. 2, P and Q denotes the items in the set of items I. The strength of the rule is measured using confidence and support value based on conditional probability. The support measures the frequency of the rule whereas the confidence measures the strength of the relation between item sets. The support of an association rule or probability of joint P and Q are mathematically formulated as given as:

$$\text{Sup}(P \rightarrow Q) = \left(\frac{\text{Support count of } P \cup Q}{\text{Total number of records in DS}} \right) \tag{3}$$

From Eq. 3, the support Sup is measured using the fraction of records that contain $P \cup Q$ to the total number of records in the database DS is said to be conditional probability (that involves both P and Q). On the other hand, the confidence value is measured in such a way that the transactions that contain P also contain Q and are mathematically formulated as given as:

$$\text{Conf}(P \rightarrow Q) = \left(\frac{\text{Support count of } P \cup Q}{\text{Support count of } P} \right) \tag{4}$$

By using Eq. 3 and 4, all the rules that satisfy the user-specified minimum support ‘minsup’ and minimum confidence minconf are retrieved. This in turn helps for efficient generation of association rules in a significant manner. The algorithmic description of conditional probability based association rule generation is shown.

Algorithm 1; Conditional probability based association rule generation:

Input: dataset 'DS' Items 'I = i₁, i₂, ..., i_n' Transaction 'T = t₁, t₂, ... t_n', minimum support 'minsup = 1', minimum confidence 'minconf = 0'

Output: efficient generation of frequent items (i.e.association rule generation)

- Step 1: Begin
- Step 2: For each dataset 'DS'
- Step 3: For each Transaction 'T'
- Step 4: Sort all rules according to support value using Eq. 3
- Step 5: If (Sup>minsup) then
- Step 6: return sup
- Step 7: End if
- Step 8: Sort all rules according to confidence value using Eq. 4
- Step 9: If (conf>minconf) then
- Step 10: return conf
- Step 11: End if
- Step 12: End for
- Step 13: End for
- Step 14: End

As shown in Algorithm 1, the conditional probability-based association rule generation algorithm consisting of three steps. For each dataset and transaction, all rules are first sorted according to the support value. Secondly, all rules are then sorted according to the confidence value in order to obtain the derived association rule through conditional probability. Finally, comparison of support and confidence value is made with the user-defined minimum support and minimum confidence to measure the strength of the rule. This in turn assists for efficient generation of association rule which resulting in minimized number of associative rules being generated.

Gravitational search algorithm for sensitive association rule selection:

In GS-SRS technique, Gravitational Search Algorithm (GSA) is introduced to select the sensitive rules is to be hided from the derived association rule for enhancing the privacy preservation of transactional database. The idea of GSA came from the Newtonian laws of gravitation and motion (Rashedi *et al.*, 2009) which says that all objects move due to the attraction with each other by gravitational forces. Therefore, objects with heavier mass have stronger attraction and move slower than the objects with relatively smaller mass.

Based on this fact, this work extended it by designing the gravitational search where data items with more cohesiveness has greater impact on the transactions rules being generated (i.e., mass interactions) than non-cohesive data items (i.e.. distance). Due to its fast convergence rate, the gravitational search is conducted to list the cohesive and non-cohesive items in each of the association rule generated for the given transactional

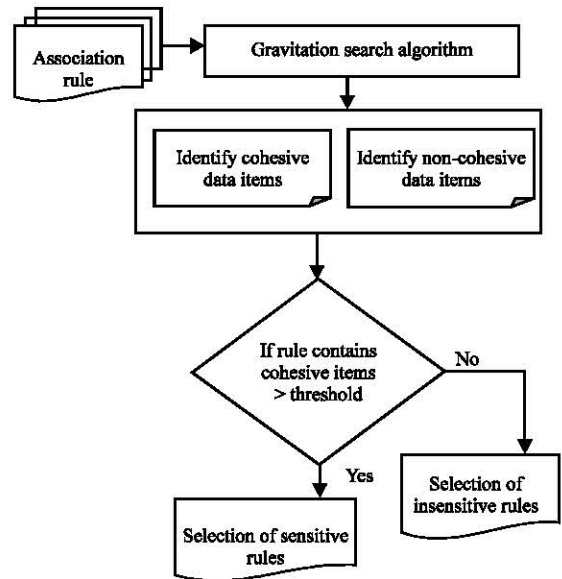


Fig. 2: Process of gravitation search model for selection of sensitive rules

cohesive items in turn is selected as sensitive rule is to be hided for improving privacy preservation of transactional database. The threshold is set for the selection of sensitive rules with the convergence of cohesive items and divergence of the non-cohesive items. Figure 2 shows the process of gravitation search model for sensitive rules selection.

As shown in Fig. 2, gravitation search algorithm initially takes association rule as input. Then, this algorithm identifies cohesive data items and non cohesive data items (i.e., association rule). If association rule contains more cohesive items, then it selected as the sensitive rules otherwise it selects as non-sensitive rules. This efficient separation of cohesive and non cohesive data items helps to reduce the execution time for selecting sensitive rules. The gravitational search algorithm is based on the law of gravity is formulated as given as:

$$F = G * \left(\frac{M1 * M2}{R^2} \right) \tag{5}$$

From Eq. 5, Fdenotes the gravitational force, G is gravitational constant, M1 and M2 are the mass of the first and second data items and R is the distance between two data items. In a transactional database with N items, the position of the ith item is defined as:

$$X_i = (X_1^d, \dots, X_1^d, \dots, X_1^n) \text{ for } i=1,2, \dots, n \tag{6}$$

where present the position of ith item in the dth dimension and N is dimension of the search space. At time t, a force acts on mass ifrom mass, j and this force F_{ij}^d is defined as:

$$F_{ij}^d = G(t) \frac{m_{pi}(t) * m_{pj}(t)}{R_{ij} + \epsilon} \quad (7)$$

Where:

- M_{pi} = Active gravitational mass of item i
- M_{pj} = Passive gravitational mass of item j
- $G(t)$ = Gravitational constant at time t, ϵ is a small constant
- $R_{ij}(t)$ = Euclidian distance between two items i and j, respectively

Then, the total force acting on mass, i in the dth dimension at time t is defined as:

$$F_i^d(t) = \sum_{j \in K_{Best}, j \neq i}^{N} \text{frand}_j F_{ij}^d(t) \quad (8)$$

where frand_j is a random number in the interval [0.0, 1.0], K_{best} is the set of first K items withthe best fitness value. The acceleration related to mass Pin time t in the dth dimension is given as follows:

$$a_i^d = \frac{F_i^d(t)}{M_{ii}(t)} \quad (9)$$

From Eq. 9, M_{ii} is inertial mass of ith item. The velocity of an item could be calculated as a fraction of its current velocity added to its acceleration. Position and velocity of agent is calculated as follows:

$$V_i^d(t+1) = \text{frand}_i V_i^d(t) + a_i^d(t) \quad (10)$$

$$X_i^d(t+1) = X_i^d(t) + V_i^d(t+1) \quad (11)$$

where frand is a uniform random variable in the interval [0.0, 1.0]. Let, n number of rules be generated through conditional probability which is denoted as:

$$R_i = \gamma_1, \gamma_2, \dots, \gamma_n \quad (12)$$

Based on the rules being generated, the gravitation search model is introduced to select the sensitive rules. In order to evaluate the sensitive rules, best fitness and worst fitness are evaluated using maximization problem which is mathematically formulated as:

$$\text{Best}(R_j) = \max \text{fit}(R_j) \quad (13)$$

$$\text{Worst}(R_j) = \min \text{fit}(R_j) \quad (14)$$

From Eq. 13 and 14, $\text{Best}(R_j)$ denotes the best fitness at iteration j and $\text{Worst}(R_j)$ denotes the worst fitness at iteration j, respectively. Followed by this, the gravitational search is performed to identify the cohesive and non-cohesive items in each of the association rule generated for given transactional database.

The fitness function provided in this algorithm based on high support value of the rule serves as a basic criterion in extracting association rules. On the other hand, support value of the rule is considered to be non sensitive and not interesting for the user. The threshold T is set for the selecting the sensitive rules with the convergence of cohesive items and divergence of the non-cohesive items using gravitation constant. The gravitation constant is measured for each iteration j is formulated as:

$$GR(j) = Te \left(\frac{-j}{n} \right) \quad (15)$$

From Eq. 15, the value of threshold T is initialized at the start of selection of sensitive rules with n representing the total number of iterations. Finally, using the gravitation constant, the rule comprising more cohesive items influences the privacy of an individual are selected as sensitive rule. It is formulated as given below:

$$CI = \left(\frac{\text{fit}(R_j) - \text{Worst}(R_j)}{\text{Best}(R_j) - \text{Worst}(R_j)} \right) \quad (16)$$

$$NCI = \left(\frac{\text{fit}(R_j) - \text{Best}(R_j)}{\text{Best}(R_j) - \text{Worst}(R_j)} \right) \quad (17)$$

From Eq. 16 and 17, the cohesive and non-cohesive items are differentiated with the rule comprising more cohesive items are selected as the sensitive rule. With the separation of cohesive and non-cohesive items, execution time for selecting sensitive rule is minimized in a significant manner. The algorithmic process of Gravitation Search based Sensitive Rule Selection is shown in below:

Algorithm 2; Gravitational search based sensitive rule selection:

Input: Frequent Item (F1), Set (item 1, item 2, ..., item n), Association Rules (I)

Output: Sensitive Rule ($\text{Best}(R_j)$), Non Sensitive Rule ($\text{Worst}(R_j)$), Occurrence Frequency (OF), Relative Frequency (RF), Threshold (TH), Cohesive Item (CI), Non Cohesive Item (NCI)

Step 1: Begin

Step 2: Initialize TH for RF in the range from 0.0-1.0

```

Step 3: For each FI
Step 4:   IF ((TH >=0.0) and (TH <=1.0))
Step 5:   Select FI as CI for corresponding Ri and Sort Descending
CI by Fitness
Step 6:   Select all Ri with CI as Best (Ri)
Step 7:   else
Step 8:   Select FI as NCI for corresponding Ri
Step 9:   Select all Ri with NCI as Worst (Ri)
Step 10:  End if
Step 11:  If TH <=0.0 and no rule is sensitive
Step 12:  Re-initialize TH with range value decremented to 0.01
Step 13:  Process step 1 to step 7
Step 14:  Else
Step 15:  List all the Best (Ri) and Worst (Ri)
Step 16:  End if
Step 17:  For ∀FI ∈ Item do
Step 18:  Calculate Acceleration-OF(Item [t+1],GI)
Step 19:  Calculate Velocity-RF (Item [t+1], GI)
Step 20:  If Frand() < S, (V i i d (t+1) then
Step 21:  Exchange (X i i d (t+1), t++)
Step 22:  Sensitive Rule -> Best (Ri) DiscoveredBestRules
Step 23:  End if
Step 24:  End for
Step 25:  End for
Step 26:  End
    
```

As shown in Algorithm 2, the threshold variance is initially set in the range of 0-1. The value of threshold variance changes during frequent item pruning while evaluating the relative frequency to identify the sensitive and non-sensitive items. When the threshold value lies between 0 and 1, frequent items are selected as cohesive items to the corresponding association rules and it is also chosen as the sensitive rule. When the condition is not satisfied, frequent items are selected as non-cohesive items to the corresponding associative rules, referred to as the non-sensitive rules. When the threshold value is less than or equal to zero, there is no occurrence of sensitive rules. So, the threshold value gets reinitialized and the process gets repeated. Finally, association rules that are selected as sensitive rule is hidden in order to improve the privacy preservation of sensitive information's in transactional database.

Experimental setup: Motivated by the work in Group Incremental Approach using Rough Set (GIA-RS) (Liang *et al.*, 2014) and Locality Sensitive Hiding for Privacy Preserving (LSH-PP) (Karapiperis and Verykios, 2015) that deals with dynamically increasing dataset and has undergone an anonymization transformation, a Gravitational Search Based Sensitive Rule Selection (GS-SRS) technique is introduced to improve the selection of sensitive rules to enhance the business transactions using MATLAB tool. The objective of the following experiments is to show the effectiveness and efficiency of the proposed GS-SRS technique. Three datasets namely Taxi service trajectory, Tic-tac-toe and shuttle datasets are used in measuring the efficiency of GS-SRS technique. Their description is provided in the following sections.

Dataset descriptions

Taxi service trajectory dataset: The performance evaluation of GS-SRS technique is performed using Taxi Service Trajectory dataset extracted from UCI repository. The dataset includes entire taxi service trajectory evaluation dataset comprising of 9 attributes (trip-ID, call-type origin-call origin-stand, taxi-ID, timestamp, day-type, missing-data and polyline) with the aid of MATLAB. This dataset has been chosen because it gives a clear picture that helps in evaluating the trajectories performed by all the 442 taxis running in the city of Porto in Portugal from the view of the company where sensitive rules are hidden and the attributes (i.e., characteristics) are displayed to the customer.

Tic-tac-toe dataset: The Tic-tac-toe database encodes the complete set of possible board configurations at the end of tic-tac-toe games where “x” is assumed to have played first. The target concept is “winning for “x” (i.e., true when “x” has one of 8 possible ways to create a “three in-a-row”).

Shuttle dataset: The shuttle dataset comprises of 9 attributes. All the nine attributes are numerical. The examples in the original dataset were in time order and this time order is more suitable for classification. However, this was not deemed relevant for Stat Log purposes, so, the order of the examples in the original dataset was randomised and a portion of the original dataset is removed for validation purposes. All the experiments have been carried out on a personal computer with Windows 7, Inter(R) Core (TM) i7-2600 CPU (2.66 GHz) and 4.00 GB memory. Experiment is conducted on the factors such as number of associative rules, selected sensitive rules, data utility rate, execution time for selecting sensitive rule and size of the transaction database. The results of the metrics of GS-SRS technique is compared against with the existing methods such as Group Incremental Approach using Rough Set (GIA-RS) (Liang *et al.*, 2014) and Locality Sensitive Hiding for Privacy Preserving (LSH-PP) (Karapiperis and Verykios, 2015).

Online retail data set: The online retail data set is a transnational data set that includes the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail. This data set consists of 8 attributes namely InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID and Country.

RESULTS AND DISCUSSION

The Gravitational Search based Sensitive Rule Selection (GS-SRS) technique is compared against the

Table 1: Tabulation for accuracy of privacy preservation using taxi service, shuttle and online retail dataset

No. of association rules	Accuracy of privacy preservation (%)								
	Taxi service trajectory dataset			Shuttle dataset			Online retail data set		
	GS-SRS	GIA-RS	LSH-PP	GS-SRS	GIA-RS	LSH-PP	GS-SRS	GIA-RS	LSH-PP
20	81.53	76.95	71.90	78.22	70.12	66.85	85.62	79.92	73.68
40	83.65	78.16	72.86	79.56	73.25	69.52	87.25	82.36	75.91
60	84.95	81.25	74.62	81.26	74.97	71.62	88.31	84.25	77.63
80	86.29	82.93	77.26	84.62	77.54	72.96	89.96	85.96	78.92
100	87.62	85.36	78.92	85.92	78.92	74.62	91.23	86.65	81.06
120	89.63	86.97	80.26	87.32	80.20	77.25	93.65	89.36	84.26
140	90.05	88.11	82.37	88.13	81.86	80.23	96.15	90.85	86.92

Table 2: Tabulation for execution time using taxi service trajectory and shuttle and online retail dataset

No. of association rules	Execution time (msec)								
	Taxi service trajectory dataset			Shuttle dataset			Online retail dataset		
	GS-SRS	GIA-RS	LSH-PP	GS-SRS	GIA-RS	LSH-PP	GS-SRS	GIA-RS	LSH-PP
20	12.6	15.1	18.8	10.2	13.5	16.7	8.2	11.6	14.9
40	17.1	18.5	23.6	12.6	16.9	20.3	10.5	13.9	17.2
60	20.5	24.9	27.2	18.6	21.3	26.4	16.4	18.3	22.8
80	25.9	30.7	33.8	21.5	28.5	31.7	19.6	22.5	26.3
100	29.3	33.2	39.6	23.6	31.7	36.2	22.2	27.6	30.4
120	32.4	35.8	42.7	27.1	33.4	40.1	25.9	30.8	33.7
140	36.7	40.4	48.3	32.8	38.9	44.8	30.5	32.7	35.8

existing Group Incremental Approach using Rough Set (GIA-RS) (Liang *et al.*, 2014) and Locality Sensitive Hiding for Privacy Preserving (LSH-PP) (Karapiperis and Verykios, 2015) using MATLAB tool.

Case study: The results in Table 1 and 2 show that, the impact of accuracy obtained and execution time for selecting sensitive rules have a profound influence on the final rule hiding. This part employs Frequency Item (FI) and Cohesive Items (CI) with gravitational search to conduct a brief case study on the data set taxi service trajectory, shuttle and Online Retail dataset, respectively.

Case scenario 1; Impact of accuracy for privacy preservation:

The accuracy of privacy preservation measures the ratio of number of sensitive rules that are correctly hidden and number of non sensitive rules that are correctly exposed to the total number of association rules generated. The accuracy of privacy preservation (A) measured in terms of percentages (%) and mathematically formulated as follows:

$$\frac{\text{(Number sensitive rules that are correctly hidden + Number non sensitive rules that are correctly exposed)}}{\text{(number rules gene)}} \quad (18)$$

From Eq. 18, the privacy preservation accuracy is measured. While the accuracy of privacy preservation is higher, the method is said to be more efficient.

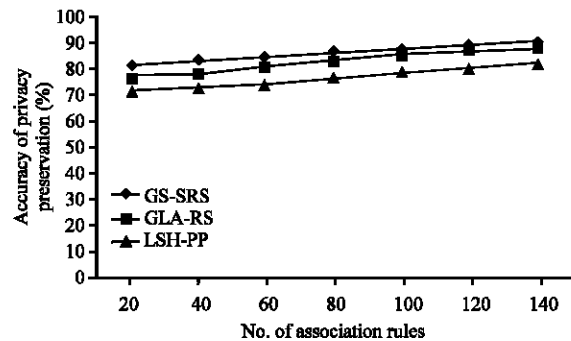


Fig. 3: Measure of accuracy of privacy preservation using taxi service trajectory dataset

Table 1 shows the accuracy of privacy preservation results is obtained with respect to different number of association rules in the range 20-140 using three datasets namely taxi service, shuttle and online retail data set. From the table value, it is illustrative that the accuracy of privacy preservation using proposed GS-SRS technique is higher as compared to other existing works using all three datasets.

Figure 3-5 show the impact of privacy preservation accuracy results is obtained versus different number of association rules using three datasets namely taxi service, shuttle and online retail data set. As exposed in Fig. 3-5, proposed GS-SRS technique is provides better accuracy for privacy preservation while compared to existing GIA-RS (Liang *et al.*, 2014) LSH-PP (Karapiperis and Verykios, 2015). This is due to application of gravitational

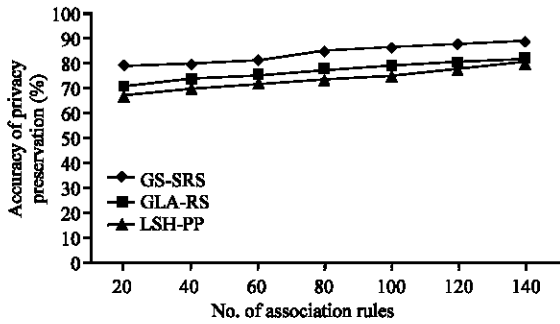


Fig. 4: Measure of accuracy of privacy preservation using Shuttle dataset

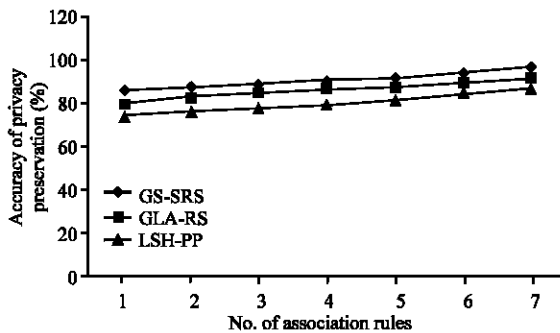


Fig. 5: Measure of accuracy of privacy preservation using online retail dataset

search based sensitive rule selection algorithm. By using this algorithmic process, proposed GS-SRS technique selects the sensitive rules to be hidden and thereby efficiently hides the sensitive information about the transactional databases in order to preserve the privacy. This in turn helps for improving the accuracy of privacy preservation in an effective manner. Therefore, proposed GS-SRS technique increases the accuracy of privacy preservation by 4% as compared to GIA-RS (Liang *et al.*, 2014) and 11% as compared to LSH-PP (Karapiperis and Verykios, 2015) when using Taxi Service Trajectory dataset. Further, proposed GS-SRS technique improves the accuracy of privacy preservation by 8% as compared to GIA-RS (Liang *et al.*, 2014) and 12% as compared to LSH-PP (Karapiperis and Verykios, 2015) when using shuttle dataset. For applying the online retail dataset, proposed GS-SRS technique improves the accuracy of privacy preservation by 5% as compared to GIA-RS (Liang *et al.*, 2014) and 12% as compared to LSH-PP (Karapiperis and Verykios, 2015), respectively.

Case scenario 2; Impact of execution time for selecting sensitive rule: The second case scenario considered for rule hiding is the execution time for selecting sensitive

rule. To measure the execution time, two factors are considered, namely, number of associative rules and the time required to extract the cohesive item. To this, 20- 140 association rules were considered at different intervals and seven iterations performed to measure the execution time. The comparison of execution time for selecting sensitive rule is presented in Table 2 with respect to the number of associative rules generated for the given transaction database (taxi service trajectory and shuttle) at varied time intervals. As shown in table, that it is feasible to perform rule hiding based on the experimental settings of associative rules selected to determine the execution time.

The execution time determines the time required for association rule generation and sensitive rules selection. The execution time is measured in terms of milliseconds (msec) and mathematically formulated as:

$$ET = n * \text{time} \left(\begin{array}{l} \text{Association rule generation +} \\ \text{Sensitive rule selection} \end{array} \right) \quad (19)$$

From Eq. 19, the execution time ET for association rule generation and sensitive rules selection is obtained. To investigate upon the factors impacts the results of execution time, the technique apply the Gravitation Search algorithm to further check the results for the given transactional database. In Table 2, it employ the gravitation search algorithm to arrive at the execution time. With increase in the number of associative rules, the execution time for selecting sensitive rule is also increased though not observed to be linear. This is because of the different types and nature of the transaction, collected at different time intervals based on the GPS data stream that invariably varies according to the timestamp.

To ascertain the performance of the execution time for selecting sensitive rule, comparison is made between two other existing methods Group Incremental Approach using Rough Set (GIA-RS) (Liang *et al.*, 2014) and Locality Sensitive Hiding for Privacy Preserving (LSH-PP) (Karapiperis and Verykios, 2015) applying the taxi service trajectory and shuttle and online retail dataset.

Figure 5-8 show a measure of execution time for sensitive rule selection using taxi service trajectory, shuttle dataset and online retail dataset. From Fig. 5-8, it is found that the execution time for selecting sensitive rule is less using the proposed GS-SRS technique when compared to the two other existing methods. Also by applying shuttle dataset, it showed better performance when compared to the taxi service trajectory. With the application of gravitation search algorithm, the cohesive and non-cohesive data items are measured for extracting the sensitive rules. This separation of cohesive and non-cohesive items minimizes the execution time while

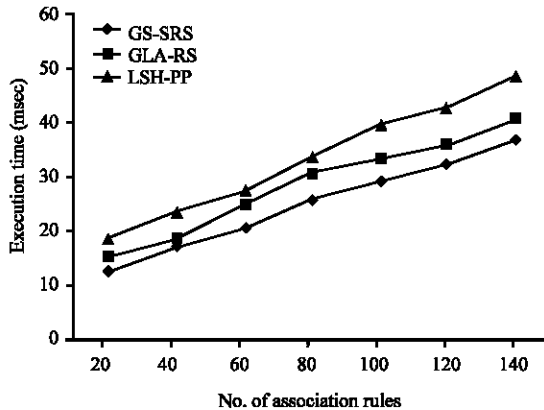


Fig. 6: Measure of execution time using taxi service trajectory dataset

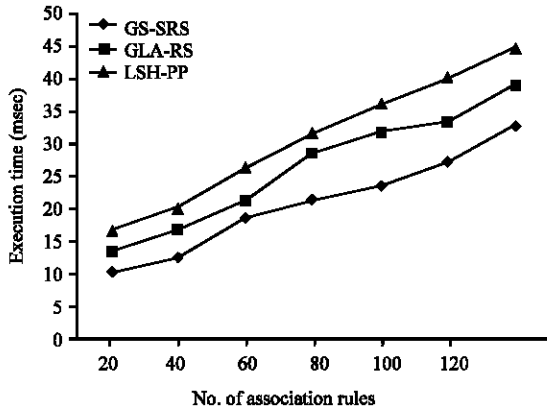


Fig. 7: Measure of execution time using shuttle dataset

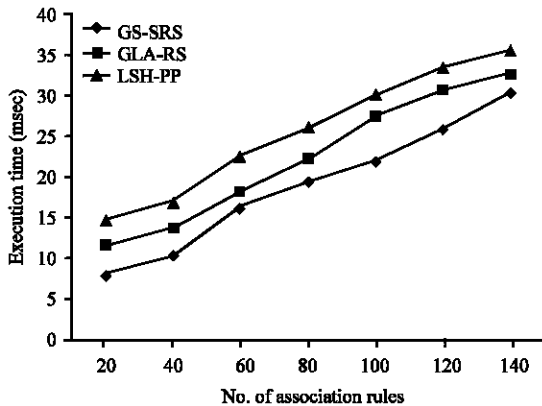


Fig. 8: Measure of execution time using online retail dataset

selecting sensitive rules using GS-SRS by 15% as compared to GIA-RS (Liang *et al.*, 2014) and 36% as compared to LSH-PP (Karapiperis and Verykios, 2015) while using taxi service trajectory dataset. Besides by

applying the gravitational constant, best fitness and worst fitness are evaluated. Followed by this, the gravitational search performed to identify the cohesive and non-cohesive items for the given transactional database which further reduces the execution time for selecting sensitive rule using GS-SRS technique by 27% as compared to GIA-RS (Liang *et al.*, 2014) and 50% as compared to LSH-PP (Karapiperis and Verykios, 2015) while Shuttle dataset. Besides for applying the online retail dataset, proposed GS-SRS technique reduces the execution time by 22% as compared to GIA-RS (Liang *et al.*, 2014) and 43% as compared to LSH-PP (Karapiperis and Verykios, 2015), respectively.

Case scenario 3; Impact of memory consumption: Finally, the third case considered is the memory consumption. To measure the memory consumption, three variables namely, the memory consumption for deriving sensitive rules, the memory for cohesive and memory for non-cohesive are used. Convergence characteristics for the measure of memory consumption for 14 iterations are taken into consideration and compared with two other methods.

With two iterations considered, the memory consumption for cohesive items was found to be 1 MB (2 MB using GIA-RS and 2 MB using LSH-PP) and with that for obtaining non-cohesive items to be 1 MB (2 MB using GIA-RS and 2 MB using LSH-PP) whereas the memory required for obtaining sensitive rules is found to be 1 MB using GS-SRS technique, GIA-RS and LSH-PP, respectively. Memory consumption for cohesive and non-cohesive items is mathematically formulated as given as:

$$M = M(SR) - [M(CI) + M(NCI)] \tag{20}$$

The memory consumption is measured in terms of MegaBytes (MB). While the memory consumption for achieving privacy preservation is lower, the method is said to be more efficient (Table 3).

The targeting results of memory consumption for cohesive and non-cohesive item generation using GS-SRS technique is compared with two state-of-the-art methods (Liang *et al.*, 2014; Karapiperis and Verykios, 2015) in Fig. 9-11 is presented for visual comparison based on the number of iterations. This technique differs from the GIA-RS (Liang *et al.*, 2014) and LSH-PP (Karapiperis and Verykios, 2015) in that, the technique have incorporated gravitational search in that cohesive and non-cohesive items are listed for the given transaction database in a much shorter period of time. The threshold is set for the selecting the sensitive rules with the convergence of cohesive items and divergence of the non-cohesive items using gravitation constant. As a result, the memory consumption for cohesive and non-cohesive items

Table 3 Tabulation for memory consumption using taxi service trajectory and tic-tac-toe dataset and online retail dataset

No. of iterations	Memory consumption (MB)								
	Taxi service trajectory dataset			Shuttle dataset dataset			Online Retail dataset		
	GS-SRS	GIA-RS	LSH-PP	GS-SRS	GIA-RS	LSH-PP	GS-SRS	GIA-RS	LSH-PP
2	4.8	5.5	2.5	3.6	4.2	1.9	2.4	3.5	3.5
4	4.1	5.5	6.2	3.1	4.2	4.8	2.2	3.2	4.0
6	4.8	5.8	6.5	3.3	4.4	5.0	2.7	3.9	4.9
8	4.2	5.6	6.3	3.6	4.7	5.3	3.1	4.3	5.2
10	5.3	6.7	7.4	3.9	5.0	5.6	3.9	4.7	5.4
12	5.5	6.9	7.6	4.3	5.4	6.0	4.3	5.1	5.7
14	5.0	6.0	6.7	4.5	5.6	6.2	4.5	5.3	5.9

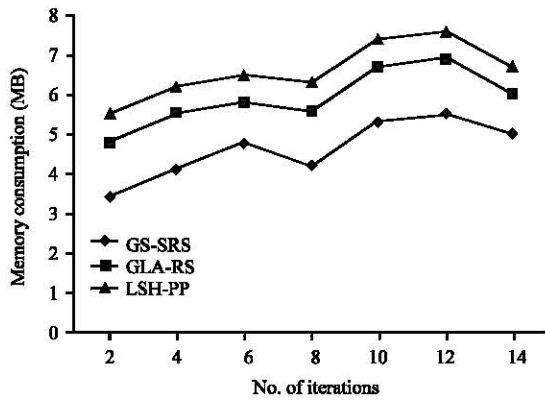


Fig. 9: Measure of memory consumption using taxi service trajectory dataset

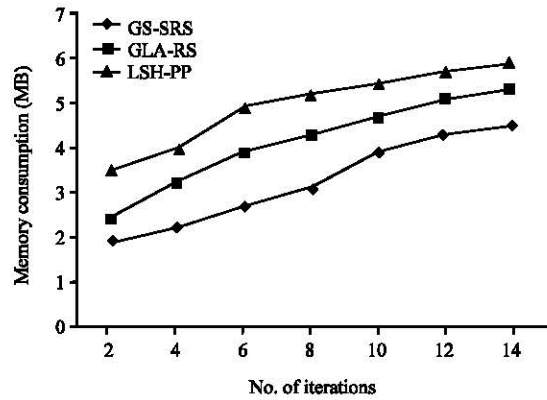


Fig. 11: Measure of memory consumption using online retail dataset

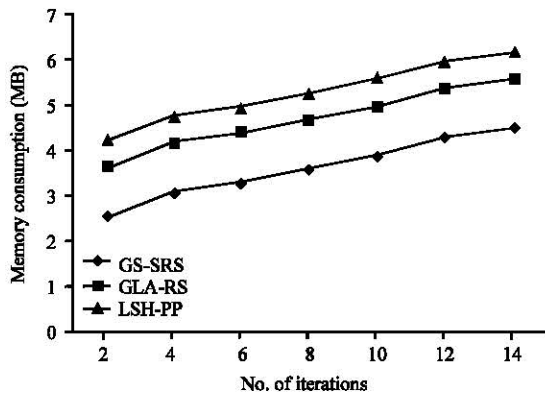


Fig. 10: Measure of memory consumption using Shuttle dataset dataset

generations using GS-SRS technique is decreased by 28% as compared to GIA-RS (Liang *et al.*, 2014) and 44% as compared to LSH-PP (Karapiperis and Verykios, 2015) while using taxi service trajectory dataset. Though nine attributes were used in Tic-tac-toe and Taxi Service Trajectory dataset with samples greater than Taxi service, the memory consumed is lesser because of a single factor considered (win for x). Therefore, the memory consumed using GS-SRS when applied with tic-tac-toe dataset is

33% reduced when compared to GIA-RS and 51% reduced when compared to LSH-PP, respectively. Besides for applying the online retail dataset, proposed GS-SRS technique reduces the memory consumption by 30% as compared to GIA-RS (Liang *et al.*, 2014) and 60% as compared to LSH-PP (Karapiperis and Verykios, 2015), respectively.

CONCLUSION

The Gravitational Search based Sensitive Rule Selection algorithm (GS-SRS) presented in this resrach is concluded as an effective sensitive rule selection technique for preserving the privacy of transactional database. The GS-SRS technique improves the selected sensitive rules that enhance the performance of association rule hiding technique for privacy preservation. The goal of gravitational search based sensitive rule selection is to select the sensitive association rules generated from frequent item sets, extracted from the transactional database to enhance the privacy of business transactions with support count and confidence value. This GS-SRS technique first designed a conditional probability-based association rule selection that measures the support and confidence value based on

the conditional probability which resulting in generation of association rules. In addition with the derived associative rule, Gravitational Search based Sensitive Rule Selection algorithm (GS-SRS) is designed to list the cohesive and non-cohesive items in each of the association rule generated and therefore reduces the execution time for selecting sensitive rules. Finally, the gravitation search model efficiently selects the sensitive rules for hiding and thereby preserving the privacy of transactional data base which resulting in improved accuracy of privacy preservation. The efficiency of GS-SRS technique is test with the metrics such as accuracy of privacy preservation, execution time and memory consumption using three datasets namely taxi trajectory, tic-tac-toe dataset, shuttle and online retail dataset. With the experiments conducted, it is observed that the GS-SRS technique provided more accurate results as compared to state-of-the-art works. The experimental results demonstrate that GS-SRS technique is provides better performance with an improvement of accuracy of privacy preservation and reduction of execution time when compared to the state-of-the-art works.

REFERENCES

- Abbasi, A., S. France, Z. Zhang and H. Chen, 2011. Selecting attributes for sentiment classification using feature relation networks. *IEEE. Trans. Knowl. Data Eng.*, 23: 447-462.
- Aldeen, Y.A.A.S., M. Salleh and M.A. Razzaque, 2015. A comprehensive review on privacy preserving data mining. *SpringerPlus*, 4: 694-730.
- Cheng, P., C.W. Lin and J.S. Pan, 2015. Use HypE to hide association rules by adding items. *PloS One*, 10: 1-19.
- De Montjoye, Y.A., E. Shmueli, S.S. Wang and A.S. Pentland, 2014. Openpds: Protecting the privacy of metadata through safeanswers. *PloS One*, 9: 1-9.
- Ghinita, G., P. Kalnis and Y. Tao, 2011. Anonymous publication of sensitive transactional data. *IEEE. Trans. Knowl. Data Eng.*, 23: 161-174.
- Gkoulalas-Divanis, A. and V.S. Verykios, 2010. *Association Rule Hiding for Data Mining*. Springer, New York, USA., ISBN:978-1-4419-6568-4, Pages: 149.
- Karapiperis, D. and V.S. Verykios, 2015. An LSH-based blocking approach with a homomorphic matching technique for privacy-preserving record linkage. *IEEE. Trans. Knowl. Data Eng.*, 27: 909-921.
- Keshavamurthy, B.N., A.M. Khan and D. Toshniwal, 2013. Privacy preserving association rule mining over distributed databases using genetic algorithm. *Neural Comput. Appl.*, 22: 351-364.
- Liang, J., F. Wang, C. Dang and Y. Qian, 2014. A group incremental approach to feature selection applying rough set technique. *IEEE. Trans. Knowl. Data Eng.*, 26: 294-308.
- Lin, C.W., B. Zhang, K.T. Yang and T.P. Hong, 2014b. Efficiently hiding sensitive itemsets with transaction deletion based on genetic algorithms. *Sci. World J.*, 2014: 1-14.
- Lin, C.W., T.P. Hong and H.C. Hsu, 2014a. Reducing side effects of hiding sensitive itemsets in privacy preserving data mining. *Sci. World J.*, 2014: 1-13.
- Zin, C.W., T.P. Hong, J.W. Wong, G.C. Lan and W.Y. Lin, 2014. A GA-based approach to hide sensitive high utility itemsets. *Sci. World J.*, 2014: 1-13.
- Maji, P. and S.K. Pal, 2010. Feature selection using F-information measures in fuzzy approximation spaces. *IEEE. Trans. Knowl. Data Eng.*, 22: 854-867.
- Ouda, M., S. Salem, I. Ali and E.S. Saad, 2015. Privacy-preserving data mining in homogeneous collaborative clustering. *Intl. Arab J. Inf. Technol.*, 12: 604-612.
- Pervaiz, Z., W.G. Aref, A. Ghafoor and N. Prabhu, 2014. Accuracy-constrained privacy-preserving access control mechanism for relational data. *IEEE. Trans. Knowl. Data Eng.*, 26: 795-807.
- Ponde, M.P.R. and S.M. Jagade, 2014. Privacy preserving by hiding association rule mining from transaction database. *Intl. Organisation Sci. Res. J.*, 16: 25-31.
- Rashedi, E., H. Nezamabadi-Pour and S. Saryazdi, 2009. GSA: A gravitational search algorithm. *Inf. Sci.*, 179: 2232-2248.
- Serra, E. and F. Spezzano, 2015. An effective GPU-based approach to probabilistic query confidence computation. *IEEE. Trans. Knowl. Data Eng.*, 27: 17-31.
- Shah, R.A. and S. Asghar, 2014. Privacy preserving in association rules using a genetic algorithm. *Turk. J. Electr. Eng. Comput. Sci.*, 22: 434-450.
- Sun, C., Y. Fu, J. Zhou and H. Gao, 2014. Personalized privacy-preserving frequent itemset mining using randomized response. *Sci. World J.*, 2014: 1-11.
- Tassa, T., 2014. Secure mining of association rules in horizontally distributed databases. *IEEE. Trans. Knowl. Data Eng.*, 26: 970-983.

- Verykios, V.S., A.K. Elmagarmid, E. Bertino, Y. Saygin and E. Dasseni, 2004. Association rule hiding. *IEEE Trans. Knowl. Data Eng.*, 16: 434-447.
- Yi, T. and M. Shi, 2015. Privacy protection method for multiple sensitive attributes based on strong rule. *Math. Prob. Eng.*, 2015: 1-15.
- Zhao, S., E.C. Tsang, D. Chen and X. Wang, 2010. Building a rule-based classifier: A fuzzy-rough set approach. *IEEE Trans. Knowl. Data Eng.*, 22: 624-638.
- Zhou, K., G.R. Xue, Q. Yang and Y. Yu, 2010. Learning with positive and unlabeled examples using topic-sensitive PLSA. *IEEE Trans. Knowl. Data Eng.*, 22: 46-48.
- Zhu, J. and Z. Li, 2015. Privacy Preserving Association Rule Mining Algorithm Based on Hybrid Partial Hiding Strategy. In: *LISS 2013*, Zhang, R., Z. Zhang, K. Liu and J. Zhang (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-40660-7, pp: 1065-1070.

DA Conceptual Security Sub-Model for Enhancing Human-Agent Collaboration Based on Generic Nodal Approach

¹Khudhair Abbas Mohammed, ^{1,2}Muamer N. Mohammed,
¹Mazlina Abdul Majidmohd and ³Mohd Sharifuddin Ahmad
¹Faculty of Computer Systems and Software Engineering,

²IBM Center of Excellence, University Malaysia Pahang, 26300 Kuantan, Pahang, Malaysia

³College of Graduate Studies, University Tenaga Nasional, Kajang, Selangor, Malaysia

Abstract: Security aspects are serious issues for multi-agent systems that should be taken into consideration when designing the systems. The significant effects and outcomes of the software agent technology make its application more and more popular in several fields. The rapid development of the agent technology promotes its ability to be applied in many different areas such as transportation, e-Business and healthcare. However as with any other systems, the agent technology is also vulnerable to security threats and consequently, there is a need to equip the system's functionalities with some security defense system. A number of methodologies are providing methods and techniques for constructing security for multi-agent systems. A survey of the current state-of-the-art of security issues for MAS is the aim of this study. Techniques and models for securing multi-agent systems are presented in this study which is categorized according to the concepts and models regarding agent's and system's functionalities. Additionally, possible threats and attacks on multi-agent systems are defined and considered with agent's roles and communications concerns. Moreover, up to date security solutions are elaborated at both levels: agent and the system levels. Finally, an analysis of existing work problems and the challenges of future research for the security of multi-agents are presented.

Key words: Software agents, multi-agent system, security aspects, nodal approach, techniques, transportation

INTRODUCTION

The traditional methods techniques of developing software systems are no more applied, especially when the facilitations of agents are expanded arises. Gradually, facilitations of agents improved several devices to become more cognitive and intelligent within its environment. Such devices are developed with software agents that can autonomously decide and accomplish its aim to satisfy user's objectives. The software agent technology is a prevalent paradigm that is increasingly popular in its application within organizations. However, for an agent to achieve its intended goal, it should have the ability to communicate and interact with other parties which could be agents and/or humans.

Simply put agents are considered as a group of computational entities that can perform tasks on behalf of their owners (Ramchurn *et al.*, 2016). An agent has the influence to collaborate with other agents to achieve its goal. Collaboration is one of the key aspects of agent's activities. It provides the necessary actions when an agent performs complex tasks. These collaborations use

some mechanisms and actions to execute a secure performance such as message exchanges. Message exchanges are autonomously performed without human intervention to resolve problems with other agents in the same system. A typical agent architecture is designed with Beliefs, Desires and Intentions (BDI), processing instructions that enable the agent to autonomously perform actions in its lifecycle.

Collaboration systems allow workers to solve a problem that is beyond their capabilities or knowledge (Majid *et al.*, 2016). Due to the distributed nature of heterogeneous systems, the objectives of a MAS are achieved when a variety of agents communicate and collaborate with each other (Mohammed *et al.*, 2014). The interactions made with other agent creates a virtual society that fulfills the reason for its existence. Such virtual society allows the MAS to implement the complexities of real world issues. Some complex systems comprise of multiple agents which interact with each other to achieve some goals. Multi-agent systems provide designers with more flexibility by creating agents with many functions (Hanna and Richards, 2012).

Collaboration, communication and intelligence are the critical functions with which agents should be designed. Such agent functions support the design of complex systems and making it closer to a real-world human society (Majid, 2011).

Multi-agent systems are widely used in complex applications such as industrial, commercial, governmental, military, entertainment and healthcare applications. Each agent has special abilities to perform tasks and it is specialized with certain capabilities to assist humans in their daily work. An agent is beneficial to humans when the agent assists to reduce the task's complexity. In many circumstances interfaces are provided to handle communication and task's exchanges between humans and agents. Agents are currently applied in both small applications such as email filters and personal assistants as well as open and complex applications such as air traffic control, military demining, logistic planning, financial portfolio management, among others (Shih *et al.*, 2013).

The particular capabilities of MAS are expanding when internet services are provided due to the availability of wide-area networks. Consequently, MAS including the autonomous systems, improve the choice for building complex and adaptive software applications agents in distributed MAS are usually interconnected and collaborate in utilizing the systems information and resources for goal achievement. Additionally, distributed MAS offer significant benefits to the system such as enhanced decision-making, suggest solutions and separating expertise.

However, in order to develop multi-agent systems that use the internet to interconnect the agents, security aspects should be implemented. MAS rely on collaboration between humans and agents which make it more vulnerable to interact with strangers. Systems in heterogeneous environments in which agents reside are open to security threats. A multi-agent system is similar to other systems when using the network, threaten by malicious actions and other security problems. The security problems in these environments which must not be neglected include malicious actions, blocking of resources, breach of information integrity and breach of private data. Additionally there are several vulnerabilities of MAS when connected to an open-network (Adameit *et al.*, 2010). Malicious entities in the environment of open-network may cause problems by rendering agents to misbehave or vulnerable to attack. By exploiting an agent's social ability, these entities applied in heterogeneous systems could force the agent to misbehave and redirect its result to the hacker's destination.

One of the techniques that attackers use to deceive is to use a masquerade agent that acts similarly to a real agent in a certain environment and steal valuable data and perform malicious actions. Attackers can also persuade an agent to perform mischievous tasks by observing data that belongs to other agents. For example in financial systems in which MAS are used to support team decision-making such systems may be inheritably unsafe. Such systems demonstrate a failure resolve security problems when a huge amount of data or online banking transactions are wrongly transferred by attackers. Furthermore, numerous insecure actions may occur and put the organization in a serious danger. In order to avoid these security threats organizations (such as healthcare administration and e-business) that depend on MAS, must install and deploy security systems within the organization.

The main purpose of this study is to identify the use of security systems in MAS and summarize the literature in related works.

Literature review: Several attempts have been made to identify the vulnerabilities in the Multi-Agent Systems (MAS). Security in computing is a platform that provides solutions which can contribute to resolving these vulnerabilities. Consequently, it is critical and important to protect data from being tampered or hacked by malicious actions. The term computer security does not have a unique definition by researchers. Security in MAS is quite susceptible to threats compared with other systems that make use of centralized performance. Many components of MAS are decentralized and deployed in different areas. Hence, it essential to promote multi-agent systems deployed with security aspects.

Researchers have proposed a diversity of techniques and methods to ensure that multi-agent systems are secured from attackers. Yue *et al.* (2009) propose a security model to prevent a colluded truncation attack which focuses on a sent message and modifies it. This attack comprises of two malicious agents which connect to a sender agent and conspire to change its message. Another technique presented by Becker-Asano and Wachsmuth (2010) in which an inference system is provided. The assistance of inference system is to tell and estimate the users about a malignant part and can be detected by the whole system. Even though their system is not established for MAS, it can be used to keep MAS safe from malevolent parts. Backer provides his framework based on the secrecy of policy language analysis.

Clark *et al.* (2010) propose a framework for a particular Service Level Agreement (SLA) that involves methods to dynamically protect and maintain reliable violation

monitoring policies. Clark uses a web service for an SLA specification which he implements in agentscape. The reason for utilizing the web services is to test and establish an agreement among participants in a centralized and decentralized monitoring. Clark framework offers usefulness of results when it is applied in decentralized systems.

Nowadays, many devices that use networks to join and share data need smart utilization to safely prevail in several areas. Definitely, security perspective considers a smarter utilization due to software vulnerabilities that cause huge problems (e.g., e-Commerce field). Security applications continuously improve system's devices with data theft invasion detection, fire detection, personal health protection, etc. Nevertheless, there are deficiencies to fully secure network devices and equip them with flexibility to serve users. Therefore, several systems have been proposed to enhance and construct security services by providing monitoring sensors, controlling parameters and robots. Moradian (2013) in the study entitled "Security of e-Commerce software systems" suggest using intelligent agents to develop business processes. The researchers show in their study the facilitation, implementation and design of the agent technology to support engineers during the development process. The proposed system enacts various security services in an e-Commerce system which assist engineers to monitor decisions and activities; Search for security measures and mechanisms; Perform checks and provide advice and feedbacks.

MATERIALS AND METHODS

Security service provision for mas: Numerous multi-agent based systems are being developed with practical applications such as the multi-agent based marketplace (Macal and North, 2014). However, in such systems often the developers tend to overlook important security features. This only leads to loss of confidence in such systems. Standard mechanisms for specifying security in multi-agent systems must be developed.

Multi-agent systems with practical applications are developed in a variety of domains (Majid *et al.*, 2009). Unfortunately, researchers often fail to include a security level in their systems, depending solely on the network security schemes. Security cannot be ignored due to its importance in developing systems. While basic security is important, gradually, confidentiality and authentication will be weaknesses to such systems. In order to solve this matter, the developer must design multi-agent systems with security in mind.

So far, there is no standard security model for MASs, although, various security models have been proposed in

the literature. Poslad *et al.* (2002) propose the asset security model in which security is defined as a set of safeguards that help protect the assets. They define the communication service, the name service and the directory service as the core MAS assets and then describe threats and safeguards for each.

Mouratidis *et al.* (2003) propose security concepts which enable the tropos methodology to model security concerns throughout the entire MAS development process. Tropos (Bresciani *et al.*, 2004) has been proposed as an agent-oriented software engineering methodology but it does not consider security.

Hence, Bresciani *et al.* (2004) add security to tropos methodology by employing various security concepts such as security constraint and security dependency, security entities such as secure goal and secure task and a security reference diagram. As an initial step, they introduce an algorithm that identifies and break security bottlenecks to reduce the complexity and criticality of MASs. They have extended Tropos to propose the secure Tropos Model (Mouratidis and Giorgini, 2007).

Potential threats and attacks: In the past decade, researchers have attempted to work on and develop agent-oriented software engineering methodologies. In fact, Gaia, one of these methodologies, provides various approaches in modeling multi-agent systems. But the development stages of such methodologies neglect the security aspects in design which few researchers seek to handle (Sawa *et al.*, 2016). In a typical organization, systems that use MAS can be crippled by cybercriminals which could ruin the business operation of the organization. Accordingly, agent behavior at some stage of operation should evolve, so that, its patterns of strength and capabilities should reduce the system's risks. In cybercrimes, deception shapes an intricate information and costs an organization both financially and administratively. These ruin the organization's reputation and image which is the objective of those criminals. Eventually, the threats mentioned above motivate the needs for security that should be considered early in the development stage.

Human vulnerabilities incite attacker's to hack systems in organizations which are hard to protect. Numerous researchers have identified and improved significant issues regarding the development of security approaches and system design. Therefore, security approaches incorporate both the design and implementation part of system functionalities. Information assets are protected once security is established in the development of MAS. The information assets and security include confidentiality integrity and availability (Mouratidis and Giorgini, 2007). Building a secure MAS

increases the degree of information confidentiality integrity and availability. Sawa *et al.* (2016) introduce an approach using natural language processing techniques for two purposes: first is to detect a social engineering attack, Second is to identify suspicious comments fabricated by attackers.

The researchers focuses on social engineering attacks which is divided into private information and command requests. Sawa *et al.* (2016) approach applies questions to the listener by performing tasks. The speaker is not authorized to perform any of these tasks before it uses a technique to detect questions and commands and extract their likely topics and finally determine which one is malicious. This approach employs the benefits of text dialogues in order to evolve its applicability and functionality on many attack vectors. However, a prime limitation to this approach is applicability because the openness of such approach refuses many traditional security solutions.

Although, applicable approaches in MASs award them an attractive behavior for various new applications, new problems arise, among which security is a key. Most of these applications are used in the real world, the need for security increases. Consequently, researchers apply their approaches to analyze several attacks and provide users and system designers with obvious and high detection technique (Mohammed *et al.*, 2014). Fortunately, there are several researchers who are working to address the potential gaps in the security of multi-agent systems and limit its vulnerability.

For example, Casey *et al.* (2016) focus on cyber security which relies on the private and asymmetric information which is significant to agent utilities. Furthermore, his approach provides an agent behavior decision when it links with an advanced model using signaling method. In addition, the nature attributes of MASs that utilizes an agent to perform its tasks in miserable systems where the traditional security techniques prevail, confer attacker to distort system data while new risks are appeared (Bijani and Robertson, 2014).

Security threats and risks enclose MASs and minimize its perfect performances and influence its outcomes. Consequently, these weaknesses need to study and detect to decrease the attacker's deception appetite. The vital features of multi-agent systems such as autonomy, openness and independence of multi-agent systems increase the value of being threatened, especially when connected to the human. Agents capabilities of interactions with humans and/or other agents enlarge the security difficulties in MASs which is an assumptive factor to keep the data of such systems safe. Phishing,

spoofing, sybil, elicitation and webpage compromises (Hornok *et al.*, 2014; Gonzalez *et al.*, 2008) are some examples of vulnerabilities of MASs used by attackers via. social engineering for their diabolical scheme. Consequently, social engineering is considered as an important and significant subject that should be intensely analyzed by cyber security designers.

Threats and attacks in multi-agent systems are not fully studied and understood by researchers. Since, MASs are gradually implemented in many distributed systems, it is imperative that designing such systems should entail a comprehensive and detailed analysis of security requirements. For instance, system's information is considered as prime risks due to insider threats and attacks. Subsequently, security in multi-agent systems is extremely fundamental and should be implemented for two reasons: firstly in order to defend systems against threats and attacks, Secondly, its design and analysis in the methodology phase which is ignored.

RESULTS AND DISCUSSION

Integration of security aspects for gna: Since, Multi-Agent Systems (MAS) have grown in popularity, secure communication is of paramount importance. A crucial limitation to MAS is security issues because of the openness of the system. Consequently, a strong mechanism with which agents in our model (GNA) can secure and defend themselves against attacks on other agents and humans needs to be developed and integrated with the main model (Mohammed *et al.*, 2012). GNA is a generic nodal approach in which several nodes can be created, for example, human-human node, human-agent node and agent-agent node. Node refers to interdependence entity where a human can collaborate with one or more than one agent via. mediator agent that act as a consultant for its human counterpart.

Basically, number of agents is restricted according to the assignment of mediator agent and delegated tasks that must be achieved. In Fig. 1, (ϕ) represents a human's set of functions and (λ) represents an agent's set of functions. GNA enhanced with security table for each node in the system. A security table store all the information of the node include: node name, agent ID and a secret key which play a role of verification and authentication when nodes interactions occur.

In such a model, appropriate security techniques need to be established by providing the following points:

- A mechanism with which agents in the generic nodal structure can secure and defend themselves against attacks on other agents and/or humans

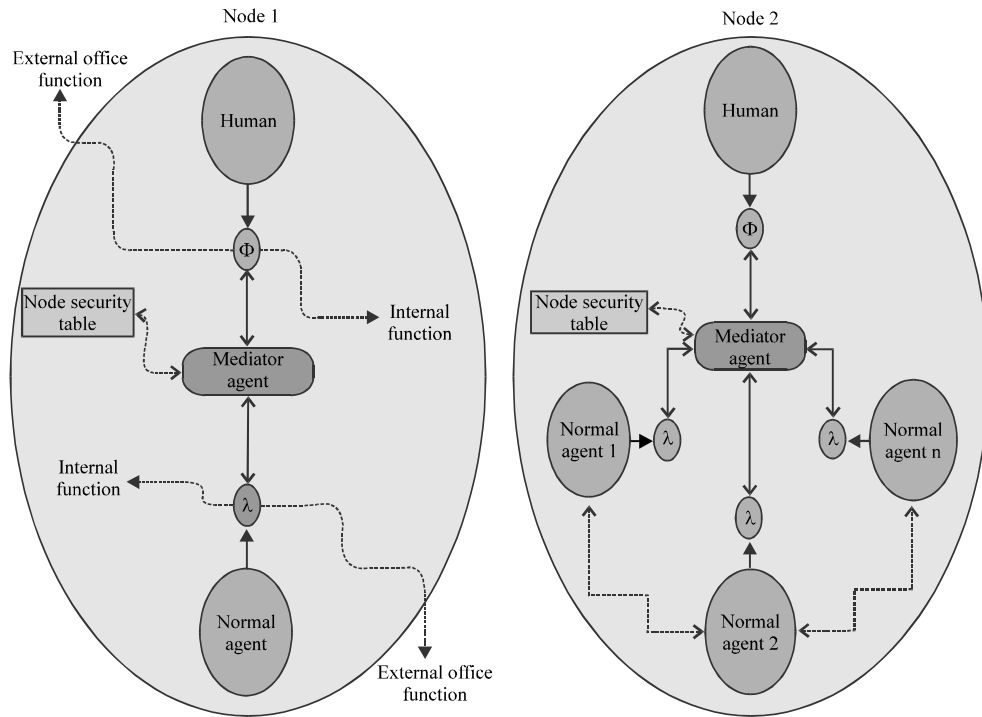


Fig. 1: Nodal approach concept

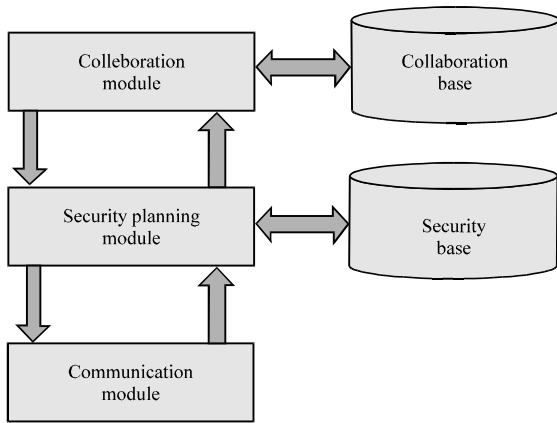


Fig. 2: The architecture for the GNA

- A security infrastructure that is applicable for the model securing agent’s actions and communication in general
- Implement security methods for agents in a MAS environment and/or distributed systems and enhance the individual agent’s abilities during collaboration

The architecture for the nodes comprises of modules which are shown in Fig. 2. The collaboration module illustrates the cooperative stage between human and agent.

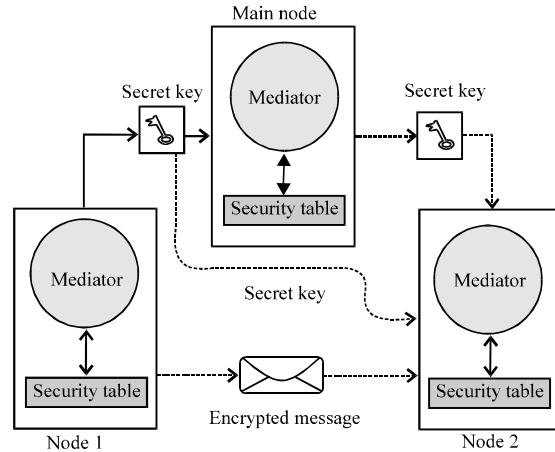


Fig. 3: Security planning module

Communication module provides protocols in which an agent is capability to exchange secured information including cooperative requests and feedback information. The security planning module receives security-resolving requests and decides whether to deal with the request according to the security policy base. Security planning module is detailed further in Fig. 3 which shows how message exchanged between nodes.

Security planning module described in Fig. 3 shows the message exchange among nodes in the GNA.

Specifically when node 1 sends a secret message to node 2 seniors occur before node 2 can open it and read it. Firstly, node 1 sends an encrypted message and to node 2. Simultaneously, node 1 sends also a random secret key to both node 2 and main node to be authenticated latter by node 2.

Note that a verbal node means practically the mediator agent inside the node which has the capabilities and authorities to check and retrieve the information with security table. Finally, node 2 need to send authentication for the main node includes sender node id and its secret key. The main node checks this information if the secret key of node 2 and node 1 are matched then an authentication sends to node 2 and now it can open the received message.

Security in multi-agent systems: Nowadays, a network is considered as a prevalent platform for organizations and people to communicate. The Internet is a manifestation of this network. In fact, netizens deal and communicate with each other via. the internet. However, the internet is highly subjected to be a target of attackers. Intrusion to websites is one of an example of an attack that has rapidly increased and deters some people and organizations from using the internet.

Generally, the current security solutions focus on preventing authorization and authentication access at various systems levels (Wardell *et al.*, 2016). Along with this security concern, a great deal of research have been done recently to secure their systems. Researchers in this field attempt to protect the systems from threats that might be caused from inside or outside the organization. Hence, these threats may cause damage to the internet assets as well as to the information.

To overcome these problems, researchers have attempted to provide better techniques by studying multi-agent systems and how the techniques can enhance and improve the security aspects. In particular, a typical security approach in MAS utilizes the software agent abilities and its intelligent, cooperative and autonomous characteristics as well as its collaboration with other agents.

For agent security, the origin of information must be authenticated to protect agents from improperly accessing or communicating with malicious agents (Jansen and Karygiannis, 2000). Additionally, an individual agent should share and permit access to its resources and internal status only with authenticated and authorized agents. Otherwise, the insecure communications among agents as well as between humans and agents could occur. Jung *et al* (2012) show in his study a number of security problems which are caused by various actors.

They point out some security causes that impair the entire system such as verification of information, unauthorized access intrusions to MAS, attacks from mobile agents and malicious agents.

Chae *et al* (2016) propose a model for information interoperability systems with security based on multi-agent. They present an agent mutual authentication method in which certificates and session keys are combined. Their MASs proposed for the national R&D information interoperability by providing session keys and mutual authentication method to overcome the security vulnerabilities in MASs. A mutual authentication method is conducted based on certification method where each agent synchronized session keys which are used to propose a method of securing channel.

Researchers have applied the MAS which provides interoperability within and across agent-based applications. They are studying the MAS approach and how the agent-based systems could be improved with better techniques to enhance the security of the systems. Security processes such as encryption, decryption, verification and digital signature are supported when the agents cooperate and communicate together.

CONCLUSION

Lack of comprehensive security techniques to multi-agent systems present a huge drawback. The challenge in multi-agent systems is to provide security mechanisms which do not affect the characteristics (such as collaboration intelligence, dynamic, cooperative problem solving and efficiency) of agents and its performances. Despite the complexities of multi-agent systems, many researchers have attempted to resolve the security problems and proposed various security approaches for MASs. Additionally, they have attempted to improve MAS facilities with security processes to deal with open issues such as secrecy of agent execution. In this study, we have categorized the existing security issues that are applicable to MAS. Among the diversity of security issues, we have devoted our study to review in more detail on threats and attacks in multi-agent systems and how the current security techniques protect message exchanges executions. However, new and aggressive attacks will continue to be released and new technologies will be developed to defeat them.

ACKNOWLEDGEMENTS

This study is supported and financed under Postgraduate Research Scheme Grand (PGRS 160319) from University Malaysia Pahang (UMP).

REFERENCES

- Adameit, S., T. Betz, L. Cabac, F. Hars and M. Hewelt *et al.*, 2010. Modelling Distributed Network Security in a Petri Net-and Agent-Based Approach. In: Multiagent System Technologies MATES, Springer, Berlin, Germany, ISBN: 13-978-3-642-16177-3, pp: 209-220.
- Becker-Asano, C. and I. Wachsmuth, 2010. Affective computing with primary and secondary emotions in a virtual human. *Auton. Agents MultiAgent Syst.*, 20: 32-49.
- Bijani, S. and D. Robertson, 2014. A review of attacks and security approaches in open multi-agent systems. *Artif. Intell. Rev.*, 42: 607-636.
- Bresciani, P., A. Perini, P. Giorgini, F. Giunchiglia and J. Mylopoulos, 2004. Tropos: An agent-oriented software development methodology. *Auton. Agents MultiAgent Syst.*, 8: 203-236.
- Casey, W., J.A. Morales, E. Wright, Q. Zhu and B. Mishra, 2016. Compliance signaling games: Toward modeling the deterrence of insider threats. *Comput. Math. Organiz. Theory*, 22: 318-349.
- Chae, C.J., K.N. Choi and K. Choi, 2016. Information interoperability system using multi-agent with security. *Wirel. Pers. Commun.*, 89: 819-832.
- Clark, K.P., M.E. Warnier, F.M. Brazier and T.B. Quillinan, 2010. Secure monitoring of service level agreements. Proceedings of the International Conference on Availability, Reliability and Security (ARES'10), February 15-18, 2010, IEEE, Krakow, Poland, ISBN:978-0-7695-3965-2, pp: 454-461.
- Gonzalez, M.C., C.A. Hidalgo and A.L. Barabasi, 2008. Understanding individual human mobility patterns. *Nature*, 453: 779-782.
- Hanna, N. and D. Richards, 2012. A Collaborative Agent Architecture with Human-Agent Communication Model. In: Cognitive Agents for Virtual Environments, Dignum, F., C. Brom, K. Hindriks, M. Beer and D. Richards (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-36443-3, pp: 70-88.
- Hornok, S., D. Kovats, T. Csorgo, M.L. Meli and E. Gonczi *et al.*, 2014. Birds as potential reservoirs of tick-borne pathogens: First evidence of bacteraemia with *Rickettsia helvetica*. *Parasites Vectors*, 7: 1-7.
- Jansen, W. and T. Karygiannis, 2000. Privilege management of mobile agents. Proceedings of the 23rd National Conference on Information Systems Security, October 16-19, 2000, NIST, Baltimore, Maryland, pp: 362-370.
- Jung, Y., M. Kim, A. Masoumzadeh and J.B. Joshi, 2012. A survey of security issue in multi-agent systems. *Artif. Intell. Rev.*, 37: 239-260.
- Macal, C. and M. North, 2014. Introductory tutorial: Agent-based modeling and simulation. Proceedings of the Conference on Winter Simulation (WSC'14), December 7-10, 2014, IEEE, Savannah, Georgia, pp: 6-20.
- Majid, M.A., 2011. Human behaviour modelling: An investigation using traditional discrete event and combined discrete event and agent-based simulation. Ph.D Thesis, University of Nottingham, Nottingham, England, UK.
- Majid, M.A., M. Fakhreldin and K.Z. Zamli, 2016. An enhanced simulation model for complex human pedestrian movement system using hybrid discrete event and agent based simulation. *Intl. Inf. Inst. Tokyo Inf.*, 19: 4213-4218.
- Majid, M.A., U. Aickelin and P.O. Siebers, 2009. Comparing simulation output accuracy of discrete event and agent based models: A quantitative approach. Proceedings of the Conference on Summer Computer Simulation (SCSC'09), July 13-16, 2009, ACM, Istanbul, Turkey, pp: 177-184.
- Mohammed, K.A., M.S. Ahmad, S.A. Mostafa and F.M.A.M. Sharifuddin, 2012. A nodal approach to modeling human-agents collaboration. *Intl. J. Comput. Appl. Found. Comput. Sci.*, 43: 33-40.
- Mohammed, K.A., S.A. Mostafa, M.S. Ahmad and M.A. Mahmoud, 2014. A qualitative analysis of human-agent functions for collaborative multi-agent system. Proceedings of the International Conference on Information Technology and Multimedia (ICIMU'14), November 18-20, 2014, IEEE, Putrajaya, Malaysia, ISBN:978-1-4799-5424-7, pp: 244-249.
- Moradian, E., 2013. Security of E-Commerce Software Systems. In: Agent and Multi-Agent Systems in Distributed Systems- Digital Economy and E-Commerce, Hakansson, A. and R. Hartung (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-35207-2, pp: 95-103.
- Mouratidis, H. and P. Giorgini, 2007. Secure tropos: A security-oriented extension of the tropos methodology. *Intl. J. Software Eng. Knowl. Eng.*, 17: 285-309.
- Mouratidis, H., P. Giorgini and G. Manson, 2003. An Ontology for Modelling Security: The Tropos Approach. In: Knowledge-Based and Intelligent Information and Engineering Systems, Palade, V., R.J. Howlett and L. Jain (Eds.). Springer, Berlin, Germany, ISBN:978-3-540-40803-1, pp: 1387-1394.

- Poslad, S., P. Charlton and M. Calisti, 2002. Specifying Standard Security Mechanisms in Multi-Agent Systems. In: Deception, Fraud and Trust in Agent Societies, Falcone, R., S. Barber, L. Korba and M. Singh (Eds.). Springer, Berlin, Germany, ISBN: 978-3-540-00988-7, pp: 163-176.
- Ramchurn, S.D., F. Wu, W. Jiang, J.E. Fischer and S. Reece *et al.*, 2016. Human-agent collaboration for disaster response. *Auton. Agents MultiAgent Syst.*, 30: 82-111.
- Sawa, Y., R. Bhakta, I.G. Harris and C. Hadnagy, 2016. Detection of social engineering attacks through natural language processing of conversations. Proceedings of the IEEE 10th International Conference on Semantic Computing (ICSC'16), February 4-6, 2016, IEEE, Laguna Hills, California, USA., ISBN:978-1-5090-0662-5, pp: 262-265.
- Shih, D.H., H.S. Chiang, D.C. Yen and S.C. Huang, 2013. An intelligent embedded system for malicious email filtering. *Comput. Stand. Interfaces*, 35: 557-565.
- Wardell, D.C., R.F. Mills, G.L. Peterson and M.E. Oxley, 2016. A method for revealing and addressing security vulnerabilities in cyber-physical systems by modeling malicious agent interactions with formal verification. *Procedia Comput. Sci.*, 95: 24-31.
- Yue, X., X. Qiu, Y. Ji and C. Zhang, 2009. P2P attack taxonomy and relationship analysis. Proceedings of the 11th International Conference on Advanced Communication Technology (ICACT'09) Vol. 2, February 15-18, 2009, IEEE, Phoenix Park, South Korea, ISBN: 978-89-5519-138-7, pp: 1207-1210.

New Linear Block Method for Third Order Ordinary Differential Equations Block Methods

Oluwaseun Adeyeye and Zurni Omar

Department of Mathematics, School of Quantitative Sciences, Universiti Utara Malaysia,
 Sintok, Kedah, Malaysia

Abstract: The diversity of adopting different step by step approach when developing block methods of the form $A^0 Y_{n+k} = A^1 Y_{n+k} + B^1 Y'_{n+k} + D^1 Y''_{n+k} + h^3(C^0 Y'''_{n+k} + C^1 Y'''_{n+k})$ is quite rigorous. Hence, this study presents an approach called Linear Block Method (LBM) capable of producing directly any k steplength value of block methods for solving third order ordinary differential equations. The LBM algorithm is validated by recovering certain existing k -step block methods in literature. Likewise, the computational complexity of the LBM algorithm is presented.

Key words: Linear block method, third order, ordinary differential equations, computational complexity, recovering, certain

INTRODUCTION

Lambert (1973) stated three approaches for determining the coefficients of linear multistep methods, namely, numerical integration, interpolation and Taylor series approach. These approaches have been modified for the transformation of the discrete linear multistep methods into block methods. These block methods consist of a family of schemes that simultaneously evaluate the solution of the differential equation under consideration at different grid points (Fatunla, 1995).

The transformation of discrete schemes to block methods using numerical integration approach have been explored by Omar (2004), Majid *et al.* (2013) and Phang *et al.* (2011), amongst many others.

Other researcher who have adopted the interpolation approach for developing these block methods include Jator and Li (2009), Badmus and Yahaya (2009) and Adesanya *et al.* (2014). Likewise, the Taylor series expansion has not been left out as it has been adopted in the research by Li (2008) and Chen and Li (2012), however, the methods developed were still restricted to development of the discrete schemes alone and further subject to boundary conditions. The direct solution of third order ordinary differential Eq. 1 of the form:

$$y''' = f(x, y, y', y'') \quad (1)$$

have been considered by several studies such as Adesanya *et al.* (2014), Mohammed and Adeniyi (2014), Adoghe (2014) and Omar and Kuboye (2015) amongst others. Hence, the introduction of the Linear Block Method (LBM) will be of good advantage for researcher when developing block methods of the form:

$$A^0 Y_{n+k} = A^1 Y_{n+k} + B^1 Y'_{n+k} + D^1 Y''_{n+k} + h^3(C^0 Y'''_{n+k} + C^1 Y'''_{n+k}) \quad (2)$$

for solving third order ordinary differential equations. Thus, the aim of this study.

MATERIALS AND METHODS

New linear block method for k -step block methods: For $y''' = f(x, y, y', y'')$. For block methods of the form Eq. 2 where, $Y_{n+k} = (y_{n+1}, y_{n+2}, \dots, y_{n+k})$ and $Y^{(i)}_{n+k} = (y^{(i)}_{n+1}, y^{(i)}_{n+2}, \dots, y^{(i)}_{n+k})$, the following LBM algorithm is given:

$$y_{n+\xi} = \sum_{i=0}^2 \frac{(\xi h)^i}{i!} y_n^{(i)} + \sum_{i=0}^k \phi_i f_{n+i}, \quad \xi = 1, 2, \dots, k \quad (3)$$

with derivatives:

$$y_{n+\xi}^{(a)} = \sum_{i=0}^{3-(a+1)} \frac{(\xi h)^i}{i!} y_n^{(i+a)} + \sum_{i=0}^k \omega_{ia} f_{n+i}, \quad a = 1_{(\xi=1, 2, \dots, k)}, 2_{(\xi=1, 2, \dots, k)} \quad (4)$$

$\phi_i = A^{-1} B$ and $\omega_{ia} = A^{-1} D$ where:

$$A = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & h & 2h & \dots & kh \\ 0 & \frac{(h)^2}{2!} & \frac{(2h)^2}{2!} & \dots & \frac{(kh)^2}{2!} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \frac{h^k}{k!} & \frac{(2h)^k}{k!} & \dots & \frac{(kh)^k}{k!} \end{pmatrix}; B = \begin{pmatrix} \frac{(\xi h)^3}{3!} \\ \frac{(\xi h)^4}{4!} \\ \frac{(\xi h)^5}{5!} \\ \dots \\ \frac{(\xi h)^{(3+k)}}{(3+k)!} \end{pmatrix}; D = \begin{pmatrix} \frac{(\xi h)^{(3+a)}}{(3-a)!} \\ \frac{(\xi h)^{(3-a)+1}}{((3-a)+1)!} \\ \frac{(\xi h)^{(3-a)+2}}{((3-a)+2)!} \\ \dots \\ \frac{(\xi h)^{(3-a)+k}}{((3-a)+k)!} \end{pmatrix}$$

For y_{n+1} :

$$\begin{pmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & h & 2h & 3h & 4h \\ 0 & \frac{h^2}{2!} & \frac{(2h)^2}{2!} & \frac{(3h)^2}{2!} & \frac{(4h)^2}{2!} \\ 0 & \frac{h^3}{3!} & \frac{(2h)^3}{3!} & \frac{(3h)^3}{3!} & \frac{(4h)^3}{3!} \\ 0 & \frac{h^4}{4!} & \frac{(2h)^4}{4!} & \frac{(3h)^4}{4!} & \frac{(4h)^4}{4!} \end{pmatrix}^{-1} \begin{pmatrix} \frac{h^3}{3!} \\ \frac{h^4}{4!} \\ \frac{h^5}{5!} \\ \frac{h^6}{6!} \\ \frac{h^7}{7!} \end{pmatrix} = \begin{pmatrix} \frac{113h^3}{1120} \\ \frac{107h^3}{1008} \\ \frac{103h^3}{1680} \\ \frac{43h^3}{1680} \\ \frac{47h^3}{10080} \end{pmatrix}$$

However, the following proposition needs to be noted.

Proposition 2.1: There exists only one block form for every k-step block method.

Verification of the LBM algorithm: To verify this algorithm, we develop the 4-step block method using the LBM algorithm and then compare the output to the derived k = 4 block method in literature:

$$A^0 Y_{n+k} = A^1 Y_{n-k} + B^1 Y'_{n-k} + D^1 Y''_{n-k} + h^3 (C^0 Y_{n+k}'' + C^1 Y_{n-k}''), k = 4$$

First, the LBM algorithm is expanded together with the expression for the derivatives:

$$\begin{aligned} y_{n+1} &= y_n + hy'_n + \frac{h^2}{2!} y''_n + (\phi_0 f_n + \phi_1 f_{n+1} + \phi_2 f_{n+2} + \phi_3 f_{n+3} + \phi_4 f_{n+4}) \\ y_{n+2} &= y_n + 2hy'_n + \frac{(2h)^2}{2!} y''_n + (\phi_0 f_n + \phi_1 f_{n+1} + \phi_2 f_{n+2} + \phi_3 f_{n+3} + \phi_4 f_{n+4}) \\ y_{n+3} &= y_n + 3hy'_n + \frac{(3h)^2}{2!} y''_n + (\phi_0 f_n + \phi_1 f_{n+1} + \phi_2 f_{n+2} + \phi_3 f_{n+3} + \phi_4 f_{n+4}) \\ y_{n+4} &= y_n + 4hy'_n + \frac{(4h)^2}{2!} y''_n + (\phi_0 f_n + \phi_1 f_{n+1} + \phi_2 f_{n+2} + \phi_3 f_{n+3} + \phi_4 f_{n+4}) \end{aligned} \tag{5}$$

with derivatives:

$$\begin{aligned} y'_{n+1} &= y'_n + hy''_n + (\omega_{01} f_n + \omega_{11} f_{n+1} + \omega_{21} f_{n+2} + \omega_{31} f_{n+3} + \omega_{41} f_{n+4}) \\ y'_{n+2} &= y'_n + 2hy''_n + (\omega_{01} f_n + \omega_{11} f_{n+1} + \omega_{21} f_{n+2} + \omega_{31} f_{n+3} + \omega_{41} f_{n+4}) \\ y'_{n+3} &= y'_n + 3hy''_n + (\omega_{01} f_n + \omega_{11} f_{n+1} + \omega_{21} f_{n+2} + \omega_{31} f_{n+3} + \omega_{41} f_{n+4}) \\ y'_{n+4} &= y'_n + 4hy''_n + (\omega_{01} f_n + \omega_{11} f_{n+1} + \omega_{21} f_{n+2} + \omega_{31} f_{n+3} + \omega_{41} f_{n+4}) \\ y''_{n+1} &= y''_n + (\omega_{02} f_n + \omega_{12} f_{n+1} + \omega_{22} f_{n+2} + \omega_{32} f_{n+3} + \omega_{42} f_{n+4}) \\ y''_{n+2} &= y''_n + (\omega_{02} f_n + \omega_{12} f_{n+1} + \omega_{22} f_{n+2} + \omega_{32} f_{n+3} + \omega_{42} f_{n+4}) \\ y''_{n+3} &= y''_n + (\omega_{02} f_n + \omega_{12} f_{n+1} + \omega_{22} f_{n+2} + \omega_{32} f_{n+3} + \omega_{42} f_{n+4}) \\ y''_{n+4} &= y''_n + (\omega_{02} f_n + \omega_{12} f_{n+1} + \omega_{22} f_{n+2} + \omega_{32} f_{n+3} + \omega_{42} f_{n+4}) \end{aligned}$$

(6)

Likewise for y_{n+2} :

$$\begin{pmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} = \begin{pmatrix} \frac{331h^3}{630} \\ \frac{332h^3}{315} \\ \frac{8h^3}{21} \\ \frac{52h^3}{315} \\ \frac{19h^3}{630} \end{pmatrix} y_{n+3}, \begin{pmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} = \begin{pmatrix} \frac{1431h^3}{1120} \\ \frac{1863h^3}{560} \\ \frac{243h^3}{560} \\ \frac{45h^3}{112} \\ \frac{81h^3}{1120} \end{pmatrix} y_{n+4}, \begin{pmatrix} \omega_{01} \\ \omega_{11} \\ \omega_{21} \\ \omega_{31} \\ \omega_{41} \end{pmatrix} = \begin{pmatrix} \frac{248h^3}{105} \\ \frac{2176h^3}{315} \\ \frac{32h^3}{105} \\ \frac{128h^3}{105} \\ \frac{8h^3}{63} \end{pmatrix} y'_{n+1}, \begin{pmatrix} \omega_{01} \\ \omega_{11} \\ \omega_{21} \\ \omega_{31} \\ \omega_{41} \end{pmatrix} = \begin{pmatrix} \frac{367h^3}{1440} \\ \frac{3h^2}{8} \\ \frac{47h^2}{240} \\ \frac{29h^2}{360} \\ \frac{7h^2}{480} \end{pmatrix} y'_{n+2}, \begin{pmatrix} \omega_{01} \\ \omega_{11} \\ \omega_{21} \\ \omega_{31} \\ \omega_{41} \end{pmatrix} = \begin{pmatrix} \frac{53h^2}{90} \\ \frac{8h^2}{5} \\ \frac{h^2}{3} \\ \frac{8h^2}{45} \\ \frac{h^2}{30} \end{pmatrix}$$

$$\begin{pmatrix} \omega_{01} \\ \omega_{11} \\ \omega_{21} \\ \omega_{31} \\ \omega_{41} \end{pmatrix} = \begin{pmatrix} \frac{147h^2}{160} \\ \frac{117h^2}{40} \\ \frac{27h^2}{80} \\ \frac{3h^2}{8} \\ \frac{9h^2}{160} \end{pmatrix} y'_{n+3}, \begin{pmatrix} \omega_{01} \\ \omega_{11} \\ \omega_{21} \\ \omega_{31} \\ \omega_{41} \end{pmatrix} = \begin{pmatrix} \frac{56h^2}{45} \\ \frac{64h^2}{15} \\ \frac{16h^2}{15} \\ \frac{64h^2}{45} \\ 0 \end{pmatrix} y'_{n+4}, \begin{pmatrix} \omega_{02} \\ \omega_{12} \\ \omega_{22} \\ \omega_{32} \\ \omega_{42} \end{pmatrix} = \begin{pmatrix} \frac{251h}{720} \\ \frac{323h}{360} \\ \frac{11h}{30} \\ \frac{53h}{360} \\ \frac{19h}{720} \end{pmatrix}$$

$$y_{n+2}^* \begin{pmatrix} \omega_{b2} \\ \omega_{22} \\ \omega_{32} \\ \omega_{42} \end{pmatrix} = \begin{pmatrix} \frac{29h}{90} \\ \frac{62h}{45} \\ \frac{4h}{15} \\ \frac{2h}{45} \\ \frac{h}{90} \end{pmatrix} y_{n+3}^* \begin{pmatrix} \omega_{b2} \\ \omega_{22} \\ \omega_{32} \\ \omega_{42} \end{pmatrix} = \begin{pmatrix} \frac{27h}{80} \\ \frac{51h}{40} \\ \frac{9h}{10} \\ \frac{21h}{40} \\ \frac{3h}{80} \end{pmatrix} y_{n+4}^* \begin{pmatrix} \omega_{b2} \\ \omega_{22} \\ \omega_{32} \\ \omega_{42} \end{pmatrix} = \begin{pmatrix} \frac{14h}{45} \\ \frac{64h}{45} \\ \frac{8h}{15} \\ \frac{64h}{45} \\ \frac{14h}{45} \end{pmatrix}$$

Combining these results give the block form:

$$\begin{aligned} y_{n+1} &= y_n + hy_n' + \frac{h^2}{2} y_n'' + \frac{h^3}{30240} (3051f_n + 3210f_{n+1} - 1854f_{n+2} + 774f_{n+3} - 141f_{n+4}) \\ y_{n+2} &= y_n + 2hy_n' + 2h^2 y_n'' + \frac{h^3}{1890} (993f_n + 1992f_{n+1} - 720f_{n+2} + 312f_{n+3} - 57f_{n+4}) \\ y_{n+3} &= y_n + 3hy_n' + \frac{9h^2}{2} y_n'' + \frac{h^3}{3360} (4293f_n + 11178f_{n+1} - 1458f_{n+2} + 1350f_{n+3} - 243f_{n+4}) \\ y_{n+4} &= y_n + 4hy_n' + 8h^2 y_n'' + \frac{h^3}{2205} (5208f_n + 15232f_{n+1} - 672f_{n+2} + 2688f_{n+3} - 280f_{n+4}) \\ y_{n+1}' &= y_n' + hy_n'' + \frac{h^2}{1440} (367f_n + 540f_{n+1} - 282f_{n+2} + 116f_{n+3} - 21f_{n+4}) \\ y_{n+2}' &= y_n' + 2hy_n'' + \frac{h^2}{90} (53f_n + 144f_{n+1} - 30f_{n+2} + 16f_{n+3} - 3f_{n+4}) \\ y_{n+3}' &= y_n' + 3hy_n'' + \frac{h^2}{160} (147f_n + 468f_{n+1} + 54f_{n+2} + 60f_{n+3} - 9f_{n+4}) \\ y_{n+4}' &= y_n' + 4hy_n'' + \frac{h^2}{45} (56f_n + 192f_{n+1} + 48f_{n+2} + 64f_{n+3}) \\ y_{n+1}'' &= y_n'' + \frac{h}{720} (251f_n + 646f_{n+1} - 264f_{n+2} + 106f_{n+3} - 19f_{n+4}) \\ y_{n+2}'' &= y_n'' + \frac{h}{90} (29f_n + 124f_{n+1} - 24f_{n+2} + 4f_{n+3} - f_{n+4}) \\ y_{n+3}'' &= y_n'' + \frac{h}{80} (27f_n + 102f_{n+1} - 72f_{n+2} + 42f_{n+3} - 3f_{n+4}) \\ y_{n+4}'' &= y_n'' + \frac{h}{45} (14f_n + 64f_{n+1} - 24f_{n+2} + 64f_{n+3} - 14f_{n+4}) \end{aligned} \tag{7}$$

This block expression corresponds with the coefficients presented by Adesanya *et al.* (2012) and Osa and Olaoluwa (2015) where the block method was used as a block predictor. Hence, this verifies the proposed LBM algorithm for developing block methods for solving third order ordinary differential equations.

RESULTS AND DISCUSSION

Computational complexity of the new linear block method for developing four-step block method: The development of the four-step block method using the new LBM algorithm involved obtaining the coefficients for the block corrector schemes and its corresponding derivatives at grid points x_{n+1} , x_{n+2} , x_{n+3} and x_{n+4} .

LBM algorithm:

Step 1: Evaluate:

$$y_{n+\xi} = \sum_{i=0}^2 \frac{(\xi h)^i}{i!} y_n^{(i)} + \sum_{i=0}^k \phi_i f_{n+i}, \xi = 1, 2, \dots, k$$

with derivatives:

$$y_{n+\xi}^{(a)} = \sum_{i=0}^{3(a+1)} \frac{(\xi h)^i}{i!} y_n^{(i+a)} + \sum_{i=0}^k \omega_{ia} f_{n+i}, a = 1_{(\xi=1, 2, \dots, k)}, 2_{(\xi=1, 2, \dots, k)}$$

$\phi = A^{-1} B$ and $\omega_a = A^{-1} D$ where

$$A = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & h & 2h & \dots & kh \\ 0 & \frac{h^2}{2!} & \frac{(2h)^2}{2!} & \dots & \frac{(kh)^2}{2!} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \frac{h^k}{k!} & \frac{(2h)^k}{k!} & \dots & \frac{(kh)^k}{k!} \end{pmatrix}, B = \begin{pmatrix} \frac{(\xi h)^3}{3!} \\ \frac{(\xi h)^4}{4!} \\ \frac{(\xi h)^5}{5!} \\ \dots \\ \frac{(\xi h)^{(3+k)}}{(3+k)!} \end{pmatrix}, D = \begin{pmatrix} \frac{(\xi h)^{(3-a)}}{(3-a)!} \\ \frac{(\xi h)^{(3-a)+1}}{((3-a)+1)!} \\ \frac{(\xi h)^{(3-a)+2}}{((3-a)+2)!} \\ \dots \\ \frac{(\xi h)^{(3-a)+k}}{((3-a)+k)!} \end{pmatrix}$$

and $k = 4$
Step 2: Stop

The two major mathematical operations required are matrix inverse and matrix multiplication. Recall that the computational complexity of taking the inverse of an $n \times n$ matrix is $O(n^3)$ while the computational complexity of the matrix multiplication of one $n \times m$ matrix with one $n \times p$ matrix is $O(nmp)$. Hence, the computational complexity of developing the four step block method using the LBM algorithm is obtained from:

$$3k \left[\left[O((k+1)^3) + O((k+1)^2) \right] \right] \tag{8}$$

as $O((k+1)^3)$.

CONCLUSION

This study has presented a new linear block method for developing any k-step block method for solving

third order ordinary differential equations. The algorithm is seen to be valid as verified with previously developed $k = 4$ block methods in literature. Hence, this algorithm is suitable for adoption when developing block methods of this kind as it bypasses the rigour attached to the step by step approach used by previous studies. The computational complexity of the method is also presented, so that, new algorithms developed in the future can be compared to this algorithm in terms of computational complexity. Likewise, this present research can be extended to using this algorithm to develop suitable k -step block methods for solving both initial and boundary value problems of third order ordinary differential equations.

REFERENCES

- Adesanya, A.O., B. Abdulqadri and Y.S. Ibrahim, 2014. Hybrid one step block method for the solution of third order initial value problems of ordinary differential equations. *Intl. J. Appl. Math. Comput.*, 6: 10-16.
- Adesanya, A.O., M.O. Udo and A.M. Alkali, 2012. A new block-predictor corrector algorithm for the solution of $y''' = f(x, y, y')$. *Am. J. Comput. Math.*, 2: 341-344.
- Adoghe, L.O., 2014. A four-point fully implicit method for the numerical integration of third-order ordinary differential equations. *Intl J. Phys. Sci.*, 9: 7-12.
- Badmus, A.M. and Y.A. Yahaya, 2009. An accurate uniform order 6 block method for direct solution of general second order ordinary differential equations. *Pac. J. Sci. Technol.*, 10: 248-254.
- Chen, J. and C. Li, 2012. High accuracy finite difference schemes for linear fourth order boundary value problem and derivatives. *J. Inf. Comput. Sci.*, 9: 2751-2759.
- Fatunla, S.O., 1995. A class of block methods for second order IVPs. *Intl. J. Comput. Math.*, 55: 119-133.
- Jator, S.N. and J. Li, 2009. Solving two-point boundary value problems by a family of linear multistep methods. *Neural Parallel Sci. Computations*, 17: 135-146.
- Lambert, J.D., 1973. *Computational Methods in Ordinary Differential Equations*. John Wiley & Sons, Hoboken, New Jersey, USA., ISBN:9780471511946, Pages: 278.
- Li, C.J., 2008. A kind of multistep finite difference methods for arbitrary order linear boundary value problems. *Appl. Math. Comput.*, 196: 858-865.
- Majid, Z.A., M.M. Hasni and N. Senu, 2013. Solving second order linear Dirichlet and Neumann boundary value problems by block method. *Intl. J. Appl. Math.*, 43: 71-76.
- Mohammed, U. and R.B. Adeniyi, 2014. A three step implicit hybrid linear multistep method for the solution of third order ordinary differential equations. *General Math. Notes*, 25: 62-74.
- Omar, Z. and J.O. Kuboye, 2015. Derivation of block methods for solving second order ordinary differential equations directly using direct integration and collocation approaches. *Ind. J. Sci. Technol.*, 8: 1-4.
- Omar, Z., 2004. Developing parallel 3-point implicit block method for solving second order ordinary differential equations directly. *Intl. J. Manage. Stud.*, 11: 91-103.
- Osa, A.L. and O.E. Olaoluwa, 2015. Hybrid and non-hybrid implicit schemes for solving third order odes using block method as predictors. *Math. Theory Modell.*, 5: 10-26.
- Phang, P.S., Z.A. Majid and M. Suleiman, 2011. Solving nonlinear two point boundary value problem using two step direct method. *J. Qual. Meas. Anal.*, 7: 129-140.

Basic Speech Feature Based Emotional Speech Analysis for Indian Native Language

¹K.M. Shiva Prasad and ²G.N. Kodanda Ramaiah

¹Department of Electronics Engineering, Jain University, Bangalore, India

²Department of ECE, Kuppam Engineering College (KEC), Kuppam, India

Abstract: Speech is known to be the most convenient form of communication between individuals. This study deals with the modus operandi for estimation of basics acoustic features as also of the function of the vocal tract area, direct from the acoustic of speech wave form. Linear Predictive Coding (LPC) is used for estimation of the function of vocal tract area. LPC is based on autoregressive modelling of the speech signals. The method is an effective technique for speech analysis. It is predominant in speech research work for locating and estimating basic acoustic features and vocal tract area function and also to represent the speech for low bit rate transmission and application in storage area. The principal objective of emotional speech analysis/recognition is to identify the different basic emotional states (primary emotions) and to categorize them under positive (non-negative) and negative emotions. Emotion analysis finds use as a tool for improvement of the quality of service found in many speech supported application like call centres. It also helps in interfaces for human computer application and others.

Key words: Emotions, LPC, MATLAB, speech analysis, Telugu, emotional speech, supported application

INTRODUCTION

Speech which is the most convenient form of communication between individuals is not just a sequence of some sounds in a steady state with abrupt changes from one to another or some signals to be ignored after listening. It actually conveys information of both linguistic and non-linguistic types and also messages of sources of multiple levels of knowledge. Speech signals are illustrative of information bearing ones that emerge as functions of a single independent variable like chronological. Speech does not imply just information signals but more than that it is wave known for complexity as the acoustic output flowing from the speaker's effort.

Sound considered typical during the articulation of a phoneme is referred to as phone. Most known languages have anything between 20-40 phonemes. They provide a range of alphabets which help us to get acquainted with many words, words have syllables as their components. Syllables are sequences of phonemes (Prasad *et al.*, 2015a, b). Sounds coming from speech are produced when air is exhaled from the lungs. They are modulated and get shaped by the vibrations of glottal cords and the resonance of the vocal tract during the process of air being pushed through the lips and the nose. Speech analysis is referred to as feature extraction of speech (He *et al.*, 2011).

Feature extraction is the process in which speech wave gets converted into a known form of parametric representation which is meant for further speech processing which in turn finds applications in many areas. Parameters so obtained from significant cues in acoustics. Raw speech data is unintelligible. It has to be reduced to manageable quantity. Information has then to extract. This process is of vital importance in understanding and interpreting the speech signal. A thorough and meaningful analysis of speech features and the linking there of to perception still remain challenging tasks (Kumar *et al.*, 2015a-c).

Speech is known for its information galore activity. It exploits frequency modulated amplitude. Such frequency and time modulated carriers convey information on words, identity of the speaker, his accent and style of speech as also the emotion with which he is changed. Examples of such carriers include resonance movements, harmonics, noise, pitch of voice, intonation and duration. The entire gamut of information is conveyed within the traditional telephone bandwidth of 4 kHz. Anything above this indicates audio quality and sensation (Prasad *et al.*, 2015a, b).

Scope of the work: Languages of a large number are in use the world over. Each language is identified through speech. A speech can get recognition through speech but identification of his speech on the basis of known and

accepted parameters still remains a complex issue with a number of digital speech processing techniques involved. Speech is caused by excitation of many organs. It consists of several complex and also simple resonant frequencies called formants. Measurements of formant frequencies is a *sin qua non* for locating speech utterances and quality of voice speech signals help identification of acoustic features and emotion which determine speaker recognition. The spectrogram technique is implemented and extraction of acoustic features is done for speeches in Telugu language samples of which have been provided by IIT-Kgp (Altun and Polat, 2009; Yu *et al.*, 2001).

Objective of the study: To generate the spectrogram and shape of vocal tract for different vowels that find place in a speakers speech through samples are taken. The researchers use the mat lab software for analysis of the formant frequencies and also samples of vocal tract shape speech. The spectrogram is used for extracting the formant frequencies and the pitch (Sheikhan *et al.*, 2013).

MATERIALS AND METHODS

Speech database: IIT-Kgp SESC has been selected as the database for the purpose of analysis. It helps the study on speech emotion recognition. The proposed speech data base is the pioneering effort in the direction of telugu an Indian language. It is directed for analysis of emotions common in conversations of day-to-day occurrence. analysis of changed emotions in areas of text, session variability and gender is possible due to the corpus being extensive. The corpus or database known as

simulated emotional speech corpus has been developed by the Indian Institute of Technology (Kharagpur) (Sheikhan *et al.*, 2013).

Ten professional artistes (5 males and 5 females) from All India Radio, Vijayawada form the population for the study. They were chosen for the recording of the corpus. The artistes are persons with rich experience known for their ability to freely express emotions from neutral sentences. They were in the age group 25-40 with professional experience of 8-12 years. These are significant details of the artistes. The 15 sentences in 8 varied emotions were delivered to each artiste in one session. 10 such sessions formed the core. Thus, the total number of utterances worked out to 12,000 (15 sentences×8 emotions×10 artistes×10 sessions) with 1500 utterances for each emotion.

Each signal was sampled at 16 kHz and represented as a 16 bit number the recording was done on alternate days to ensure the variability that is inevitable. Each artiste was asked to speak at a stretch all the sentences in a specific emotion. These ensured coherence among the artistes for each type of emotion. A single microphone was used for the entire database at the same location (He *et al.*, 2011).

Problem statement: The corpus relating to speeches in Telugu language considered in this study is combination of vowels, yogavahakas and consonants. Consisting of voiced and unvoiced sounds. Different sentences in the language were used vis-a-vis basic acoustic features (formant frequencies and pitch more specifically) (Pao *et al.*, 2008; Schüller *et al.*, 2004) (Fig 1).

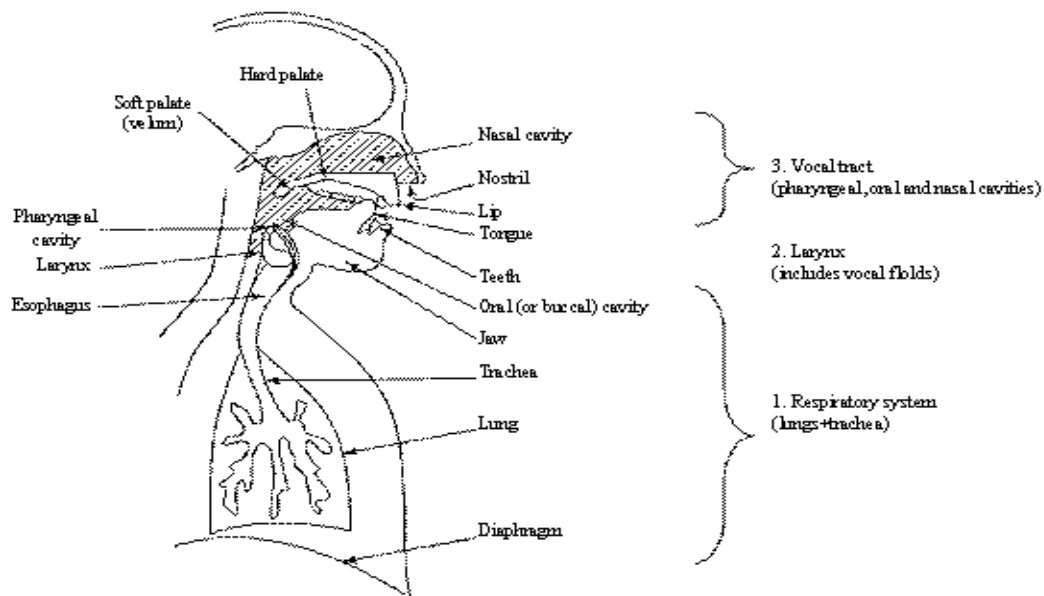


Fig. 1: Anatomy of speech production

Human-speech production mechanism: Figure 1 shows the anatomy of the human speech production systems, composed of three regions viz., larynx, vocal tract and respiratory system. This is capable of producing a variety of vibrations and spectral temporal composition which forms different speech sounds. Kumar *et al.* (2015a-c) the act of speech production begins with exhaling the inhaled air from the lungs (glottal volume). The air will sound like a random noise without any information in the absence of subsequent modulations. These are done through the manner and frequency of closing and opening of the glottal folds. Speech production is fused directly with the combination of volume and sinus cavity ending with the opening of the mouth (Prasad *et al.*, 2015a, b)

Basic acoustic features of speech: Formant frequency, pitch and intensity form the basic features in speech.

Formants: Formants are normal resonant frequencies (f_1 - f_4). They depend on the position and the manner of articulation. These formants are generally considered important but acceptable speech quality requires higher formants. Formant frequencies appear as dark bands in a spectrogram. The mode of execution in the vocal cavity decides the resonating frequencies. Formants find use in speech recognition and speaker verification. They describe the vocal tract resonance. They are quantitative characteristics of the vocal tract, the position of which depends on its shape and physical dimensions. The name has been proposed by researchers in view of the resonant frequencies tending to form the over-all spectrum of the speech signal. Formant frequencies describe the shape of the vocal tract during the production of an emotional speech (Schuller *et al.*, 2004).

Pitch: This is also known as the repetition rate of opening and closing of vocal folds, being the basic frequency of vibration of the vocal folds referred to as fundamental frequency (f_0). Pitch is useful in the classification of speech as voiced and unvoiced. It is low for the former and absent in the latter. Auto correlation is the commonly used method for measuring pitch (Schuller *et al.*, 2004).

RESULTS AND DISCUSSION

LPC based speech analysis: Analysis of speech signals is done through estimation of speech parameters (like formants) and eliminating their effect from speech signal. It estimates the intensity and frequency of the remaining buzz. LPC Model for speech processing is shown in Fig. 2-4. The process of renewal of resonant or formant frequencies is referred to as inverse filtering while

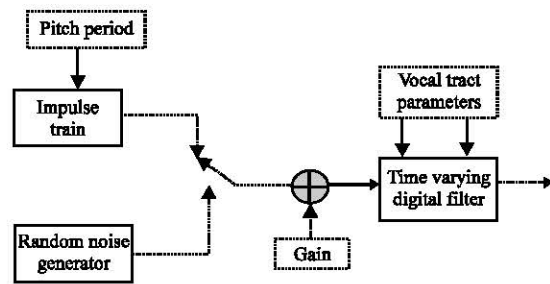


Fig. 2: LPC Model

the signal remaining after elimination is referred to as the residue signal. Speech segments of short duration called frames are used by LPC, the duration being 30 msec. Time variation in the digital system used in LPC is represented by the following transfer function (Kumar *et al.*, 2015a-c)

How emotion is created: Emotion originates from neural organization of the human brain. Generally, human brain processes the human emotion through chemical reaction. Emotion is created through a chemical reaction whose origin is not established in the nervous system due to arised situation. A cognitive state is created by number of situations and by a thinking process which relates those situations and draws some kind of conclusion as emotion (Yang and Luggler, 2010).

Emotional types: Generally their exists large types of emotions in real life namely: hot anger, cold anger, panic, fear, anxiety, despair, sadness, happiness, boredom, shame, pride, disgust and contempt. The commonly considered emotional states are anger, happiness, fear, sadness, disgust, boredom and neutral.

The basic emotions which are more primitive and universal than others are neutral, anger, fear and sadness. Neutral: it is an emotional state where emotionally lacking is noticeable. it is moderately negative, calm and weak (El Ayadi *et al.*, 2011).

Compassion: It is a emotion showing concern for the sufferings of other people (El Ayadi *et al.*, 2011).

Disgust: It is an emotional state or feeling that something is unpleasant, offensive or unacceptable (El Ayadi *et al.*, 2011).

Anger: It is a feeling or emotion showing extreme displeasure, very negative, very strong and very excited. it is an emotional state with high arousal levels which is characterised by the tense voice with faster speech rate. utterance duration, shorter inter word silence. There are two types, namely cold anger and hot anger. Hot anger is

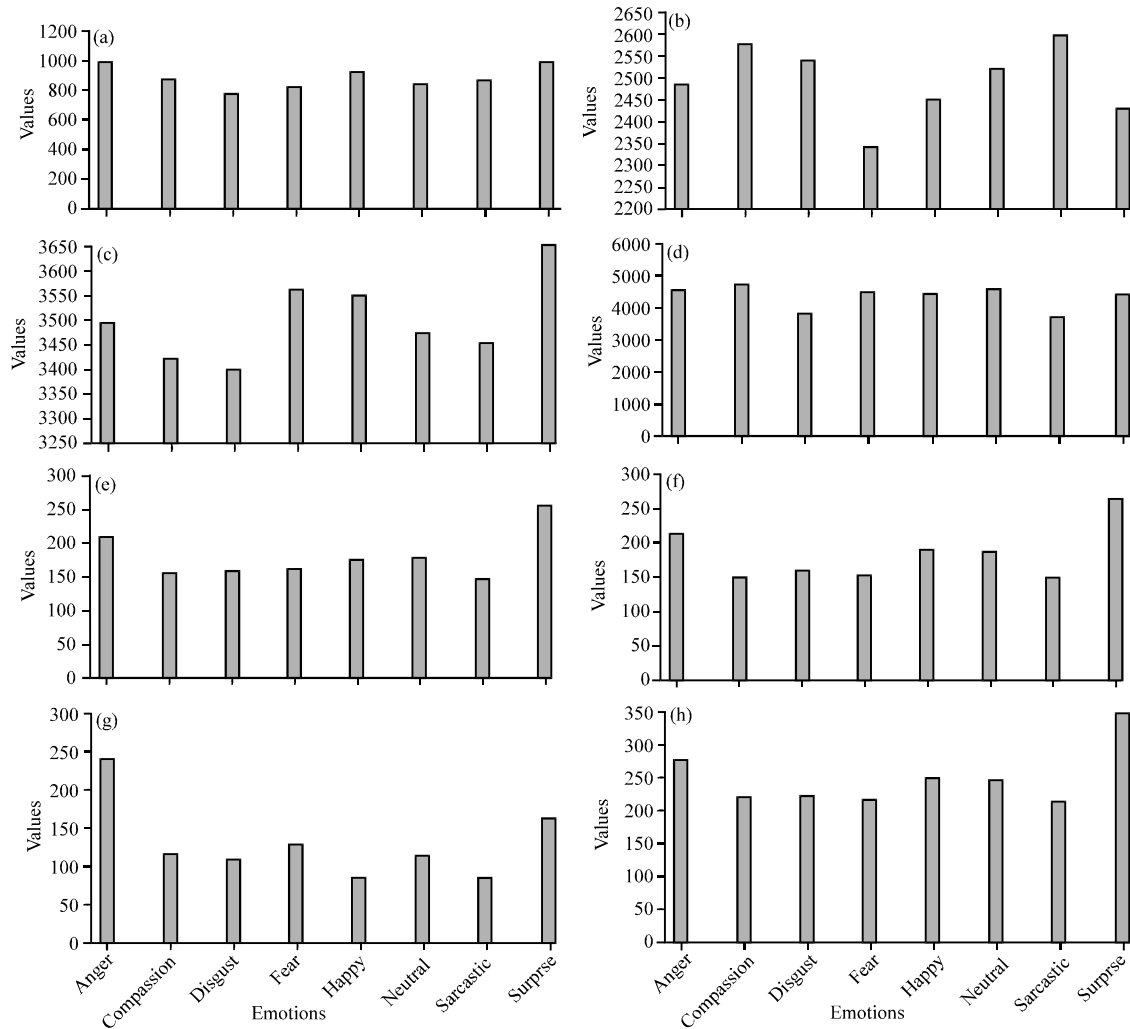


Fig 3: Comparative graph of 8 different emotions for sentence S₁ “Talli tandrulannu gowravinchavalanu”: a) F1 (Hz); b) F2 (Hz); c) F3 (Hz); d) F4 (Hz); e) Mean (Hz); f) Median (Hz); g) Minimum (Hz) and h) Maximum (Hz)

more intense in quality, characterised by the increased articulation rate. Cold anger is characterised by increase in mean F_0 , increase in high frequency energy and increase in mean intensity (Ayadi *et al.*, 2007)

Fear: It is an unpleasant emotion caused by threat of danger, afraid of others. it is characterised by increase both in mean F_0 , F_0 range and increase in energy and also articulation rate (Ayadi *et al.*, 2007).

It is an emotional state with high arousal levels which is characterised by the tense voice with faster speechrate, moderately positive, high F_0 and broader pitch range yields higher heart rate and blood pressure (Ayadi *et al.*, 2007).

Sarcastic: It is an emotional state of hurting or mocking someone. (use of words which says the opposite of what you mean) (Ayadi *et al.*, 2007).

Surprise: It is an emotional state describes the feeling of mild astonishment or shock caused by something unexpected. characterized by high values of glottal velocity (Ververidis and Kotropoulos, 2006) (Table 1).

The detailed analysis of speech features and their relationship to human perception is a challenging task in speech processing. Speech conveys both linguistic as well as non-linguistic information, there exist many more methods of speech analysis each having their own merits and demerits for required application. But there is still no single method to be considered distinctly the best method for speech analysis or recognition. Speech analysis or

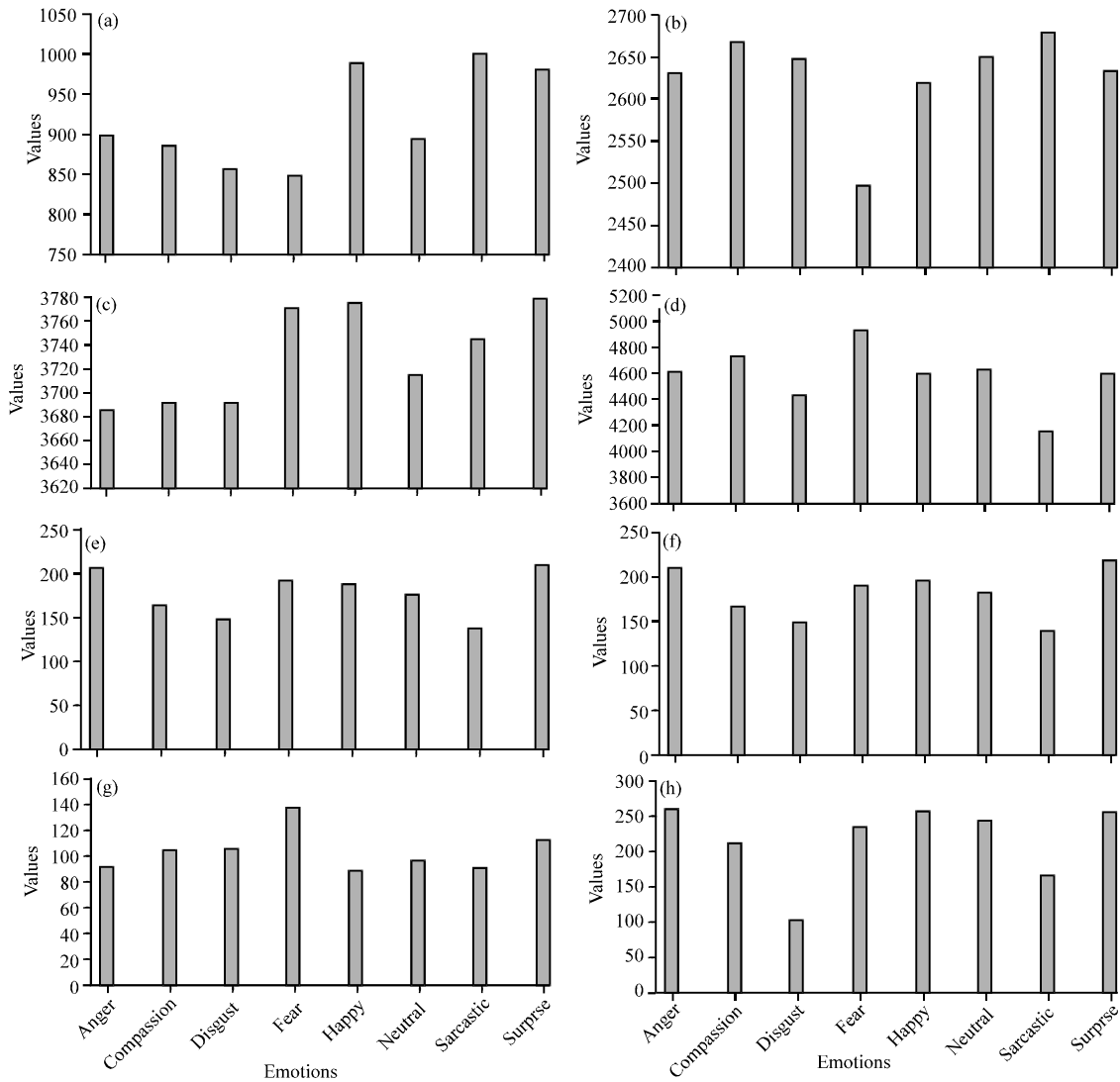


Fig 4: Comparative graph of 8 different emotions for sentence S_2 “meekosam chalachepadnundi chustannam”: a); F1 (Hz); b) F2 (Hz); c) F3 (Hz); d) F4 (Hz); e) Mean (Hz); f) Median (Hz); g) Minimum (Hz) and (h) Maximum (Hz)

Table 1: Formants and various pitch listing

Parameters		Formant frequency (Hz)				Pitch (Hz)			
Emotion	Sentence	F ₁	F ₂	F ₃	F ₄	Mean	Median	Mini.	Maxi.
Anger	S ₁	985.80	2490.98	3481.02	4908.81	210.56	218.890	240.680	262.980
	S ₂	898.52	2630.33	3679.18	4670.41	206.26	212.450	92.605	263.400
Compassion	S ₁	871.49	2587.55	3412.74	5057.56	156.41	153.780	117.170	210.430
	S ₂	885.24	2667.82	3684.55	4789.48	164.76	168.780	105.590	216.670
Disgust	S ₁	779.58	2548.61	3391.28	4094.20	159.24	164.577	109.250	213.020
	S ₂	856.75	2647.78	3682.87	4484.87	147.02	150.710	106.000	105.200
Fear	S ₁	833.15	2345.75	3543.55	4844.31	161.78	156.040	129.060	206.430
	S ₂	847.39	2497.54	3755.75	5007.88	191.11	192.300	138.370	238.045
Happy	S ₁	919.49	2456.95	3532.75	4762.24	176.97	193.420	87.140	237.260
	S ₂	988.28	2619.61	3758.28	4654.19	187.43	198.370	88.210	260.920
Neutral	S ₁	831.92	2527.00	3460.61	4943.05	179.59	191.890	114.150	235.080
	S ₂	894.46	2649.90	3703.92	4691.82	175.48	183.730	96.130	246.290
Sarcastic	S ₁	860.57	2607.21	3443.53	4009.25	147.69	151.790	84.030	202.830
	S ₂	1000.31	2678.59	3731.25	4186.03	137.99	141.350	90.011	167.620
Surprise	S ₁	1003.92	2435.33	3631.83	4740.87	255.86	267.560	162.840	331.190
	S ₂	979.69	2632.28	3761.98	4652.02	210.21	221.170	112.187	260.175

speech signal front ends helps in extracting the speech features. There are many number of features can be used for purpose of emotion analysis/detection. The selection of feature set are depends on application requirement. It is very difficult to identify the best among available. The analysis/recognition performance of emotional speech using prosodic (Acoustic) features is not found to be appreciable but recognition can be improved by combining prosodic with spectral features. The emotional analysis/recognition can be extended to databases recorded in other native languages having wide range of emotional expression by taking number of speakers, different sex, different age group as subject to find the degree of universality of emotional features (Table 1).

CONCLUSION

In this research we investigate the acoustic properties of speech corpus associated with 4 primary emotions namely neutral, happy, sad and angry intentionally expressed by an untrained actors. The speech associated with anger and happiness are characterised by longer utterance duration (time), higher pitch and energy values with wide range. However, it is observed that slightly higher pitch in anger, fear compared to neutral speech. It is noticed that we cannot draw any conclusion on the variability of mean formant frequencies as a function of emotion, as they vary depending on which formant is considered. First formant frequencies range is has higher values in angry emotion when compared to other emotions. Angry and happy have higher mean F_0 values and greater variations compared to that of neutral speech. The mean of F_0 for neutral, sad, angry and happy are as listed in the table. It is confirmed that angry speech have higher F_0 values and greater variations compared to that of neutral and happy speech. Individual mean F_0 values for each emotion category is as shown in the table.

REFERENCES

Altun, H. and G. Polat, 2009. Boosting selection of speech related features to improve performance of multi-class SVMs in emotion detection. *Expert Syst. Appl.*, 36: 8197-8203.

Ayadi, M.M.E., M.S. Kamel and F. Karray, 2007. Speech emotion recognition using Gaussian mixture vector autoregressive models. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) Vol. 4, April 15-20, 2007, IEEE, Honolulu, Hawaii, USA., ISBN: 1-4244-0727-3, pp: IV-957-IV-960.*

El Ayadi, M., M.S. Kamel and F. Karray, 2011. Survey on speech emotion recognition: Features, classification schemes and databases. *Pattern Recognit.*, 44: 572-587.

He, L., M. Lech, N.C. Maddage and N.B. Allen, 2011. Study of empirical mode decomposition and spectral analysis for stress and emotion classification in natural speech. *Biomed. Signal Process. Control*, 6: 139-146.

Kumar, A., G.K. Ramaiah and M.B. Manjunatha, 2015a. Speaker based vocal tract shape estimation for kannada vowels. *Proceedings of the International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), January 24-25, 2015, IEEE, Visakhapatnam, India, ISBN:978-1-4799-7676-8, pp: 1-6.*

Kumar, A., M.B. Manjunatha and G.K. Ramaiah, 2015b. Vocal tract shape estimation of vowels & CVVC for diversified Indian English speakers. *Proceedings of the International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), January 24-25, 2015, IEEE, Visakhapatnam, India, ISBN:978-1-4799-7676-8, pp: 1-7.*

Kumar, C.A., K.M.S. Prasad, M.B. Manjunatha and G.N.K. Ramaiah, 2015. Basic acoustic features analysis of vowels and C-V-C of Indian English language. *ITSI. Trans. Electr. Electron. Eng.*, 3: 20-23.

Pao, T., Y. Chen, J. Yeh and Y. Chang, 2008. Emotion recognition and evaluation of mandarin speech using weighted D-KNN classification. *Intl. Innov. Comput. Info. Control*, 4: 1695-1709.

Prasad, K.M.S., C.A. Kumar, M.B. Manjunatha and G.N.K. Ramaiah, 2015a. Gender based acoustic features and spectrogram analysis for kannada phonetics. *ITSI. Trans. Electr. Electron. Eng.*, 3: 16-19.

Prasad, S., A. Kumar and K. Ramaiah, 2015b. Various front end tools for digital speech processing. *Proceedings of the 2nd International Conference on Computing for Sustainable Global Development (INDIACom), March 11-13, 2015, IEEE, New Delhi, India, ISBN:978-9-3805-4415-1, pp: 905-911.*

Schuller, B., G. Rigoll and M. Lang, 2004. Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Volume 1, May 17-21, 2004, Montreal, Canada, pp: I-577-I-580.*

- Sheikhan, M., M. Bejani and D. Gharavian, 2013. Modular neural-SVM scheme for speech emotion recognition using ANOVA feature selection method. *Neural Comput. Appl.*, 23: 215-227.
- Ververidis, D. and C. Kotropoulos, 2006. Fast sequential floating forward selection applied to emotional speech features estimated on DES and SUSAS data collections. *Proceedings of the 14th European Conference on Signal Processing*, September 4-8, 2006, IEEE, Florence, Italy, pp: 1-5.
- Yang, B. and M. Lugger, 2010. Emotion recognition from speech signals using new harmony features. *Signal Process.*, 90: 1415-1423.
- Yu, F., E. Chang, Y.Q. Xu and H.Y. Shum, 2001. Emotion Detection from Speech to Enrich Multimedia Content. In: *Advances in Multimedia Information Processing-PCM 2001*, Shum, H.Y., M. Liao and S.F. Chang (Eds.). Springer, Berlin, Germany, ISBN:978-3-540-42680-6, pp: 550-557.

Implementation of Upgraded Single-Switch High Boost Converters with Infused Transformer Voltage Incremental Cell

D. Selvabharathi, K. Selvakumar, R. Palanisamy, D. Karthikeyan and K. Vijayakumar
Department of Electrical and Electronics Engineering (EEE),
Sri Ramaswamy Memorial (SRM) University, Chennai, India

Abstract: In this study, a voltage incremental circuit is used to give a universally derivatives on developed high step-up converters for high voltage multiplication circuits. N number of one-switch high step-up converters with infused transformer voltage multiplier cell gives us with many advantages, consisting of very high voltage conversion ratio, reduced voltage stress, effective diode reverse-recovery scheme, includes soft-switching operation. The turn's ratio of the infused transformer gives the adjustability to extend the voltage gain. Keeping the active clamp counterpart as a reference, one IGBT reduces the difficulties of the circuit, thereby increasing reliability and stability. Here, the circuit is analyzed with both higher level and lower resonant conditions with the provided key parameters. Incorporation of Pi-filter (LC filter) reduces the ripple content in output. Ultimately, the experimentally obtained results for a 600 W 46-480 V designed prototype are furnished in contributions in this study.

Key words: Built in transformer, high step-up, voltage multiplier cell, effective, scheme, India

INTRODUCTION

The usual boost converter does the function of multiplying voltage. The conventional boost converter achieves the voltage multiplication through the ideal CCM mode by varying the duty cycle (Do, 2010) but however, the increase in voltage is possible only 4 times the supply voltage shown in Fig. 1.

There are some applications where a 10 fold increase in voltage is required. We can give an example related to the automobile applications which uses the very high intensity light usually need to convert 20 V on battery back up to 300 V at steady state nominal operation, even to 350 V during the start-up stage (Garcia *et al.*, 2003). Normal 50-60 V supply systems like batteries are used in the transport applications but likely it has to boost up more than ten times of its output in order to meet the intensity need. Furthermore, the output voltage of the individual Photovoltaic (PV) cell is generally lower than 70 V (Shenkman *et al.*, 2004; Erickson and Maksimovic, 2001). However, the AC voltage at the grid side is 220 V for single-phase local utility in the country which also calls for high step-up and high-efficiency converters to realize the integrated PV modules:

$$\text{Power} = \text{Voltage} * \text{Current}$$

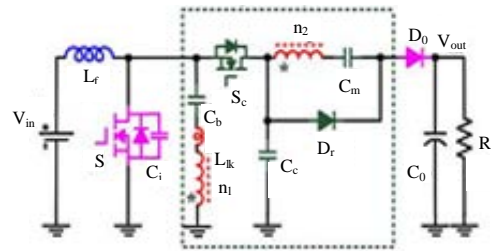


Fig. 1: Voltage multiplier cell

Assuming power = 200 W and voltage as 50 V, the current will be 4 A. Now, if a Voltage of 100 V is needed then the current reduces to 2 A in order to maintain the same 200 W power. It is clearly visible that on increasing the voltage profile, the current profile has to be reduced in order to maintain the same power rating.

Now, the question arises as to why current profile is not changed instead of voltage profile? It is because a voltage multiplication is needed, here in motors with the increase in load it draws more current, i.e. with increase in load the current levels increase. So, in motors the current profile is increasing but what is needed is increase in voltage profile. The best example for increase in voltage profile is a car's headlight where an increase in voltage in seen as illumination.

Also, the conventional boost converter uses multiple switches to achieve the voltage gain. This results in additional switching losses thereby reducing efficiency. Switch operating condition and the diode reverse-recovery condition are plays a vital role in the high output voltage loads. A Zero Voltage Switching (ZVS) Zero Current Switching (ZCS) boost converter with voltage multiplier circuits are developed to minimize the switching losses (Santhoshi *et al.*, 2016). The major limitations of the converters published by Vijayalakshmi *et al.* (2016) and Puviarasi and Dhanasekaran (2015) is that these topologies cannot provide another controllable and user free concepts except the change in duty pulse to obtain the large and wide voltage conversion that limitations are overcome by the new topology derived.

The novel idea revels from the new topology is used to overcome the above difficulties related to switching losses by means of new soft switching methods and an improved pulse width modulation techniques. With the help of this ZVS soft switching techniques the smooth operation of the circuits can be achieved due to this the stress across each switch can be balanced in the other hand we are using ZCS techniques, so, the switching stress at the point of circuit off condition can be reduced. The voltage multiplication by single cell can be achieved by voltage multiplier circuits. Either it may be of voltage doublers, tripler quadruples. This can be achieved by fly back capacitor circuits. In this regards, we have to large number of capacitors in turn one single switch to achieve.

MATERIALS AND METHODS

Topology derivation and circuit operational analysis: The stimulation of proposed ZVS high step-up built-in transformer based converter by Bharathi and Sasikumar (2016) is shown in Fig. 2. The boosting in voltage multiplier cell is composed of a transformer an active clamp switch S_c , a clamp capacitor C_c , a DC block capacitor C_b , a switched capacitor C_m and a regenerative diode D_r . The switches S and S_c work with the un symmetrical reverse operation to regulate the output C_c . ZVS soft-switching performance is achieved for the switches S and S_c due to the additional parallel capacitor C_s . The input current ripple is small due to its CCM operation to remove the large input electrolytic capacitors. As a result, this built-in transformer based converter is a best converter for the high step-up and more efficiency transfer systems.

From the above concept the capacitor is charged to the maximum and clamps to the maximum towards the of

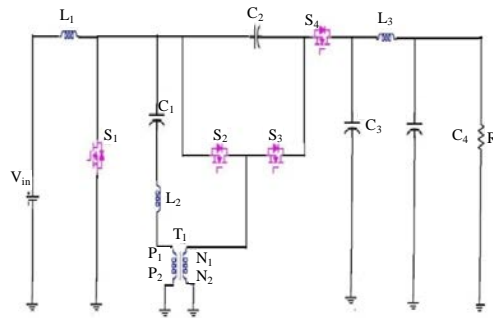


Fig. 2: Proposed high step-up converter

input voltage and it discharges to load with the limited switching time leads to the low switching frequency. The improved single-switch high step-up converter is introduced. This improved version has some other clearly advantageous performance which will be discussed in detail in the following sections. In the dashed block, a voltage multiplier cell is inserted to extend the voltage gain and reduce the switch voltage stress.

Here, we are using a infused transformer with different types of primary and secondary turns. With respect to the turns conditions of primary and secondary winding the boosting level are improved infuse to achieve the above incremental condition over resonance frequency and below resonance frequency conditions are used open switch uses the over resonance condition and closed switch uses the below resonance conditions. Due to the above category the leakage inductance value can be reduced at the point of transformer.

Part 1 [t₀, t₁]: When switch t₁ is in on condition, the semiconductor S will be in the running state and the voltage transfer towards the filter inductor L_f is V_{in}. The clamp diode and the output rectifier Do are developed by the update mathematical analysis Eq. 1-3:

$$V_{ts} = V_{cm} - V_{cc} \tag{1}$$

$$L_{lk} \frac{di_{LK}(t)}{dt} = V_{tp} - V_{cb}(t) \tag{2}$$

$$i_s(t) = i_{if}(t) + i_{Lk}(t) + \frac{i_{LK}(t)}{N} \tag{3}$$

Part 2 [t₁, t₂]: In the point of t₁, the I of the leakage inductance L_{lk} resonates (Wu *et al.*, 2008) towards the zero and the regenerative diode Dr turns OFF automatically, this gives the information about the reverse-recovery losses. The inductor I is still developing gradually due to the input voltage.

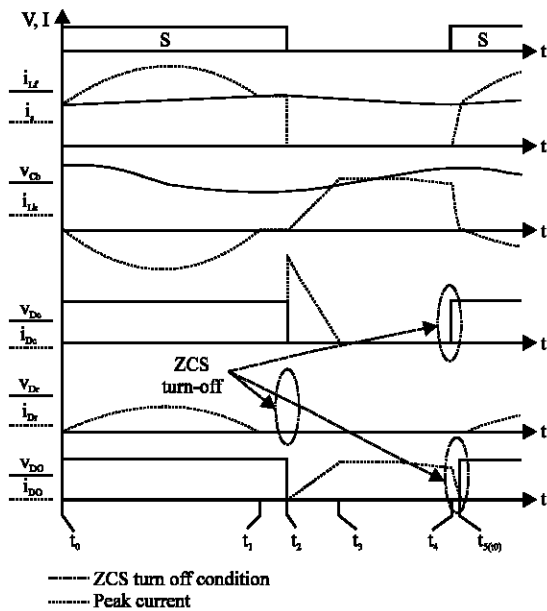


Fig. 3: Stage of switching output

Part 3 [t₂, t₃]: At t₂, the switch S turns OFF and then the clamp diode Dc and the output rectifier Do turn ON to gives out by means of the current equation given as:

$$i_D(t) = \frac{i_{LK}(t)}{N} \quad (4)$$

The operation of each stage of switching output and its modes of operation are shown in Fig. 3 and 4.

Voltage and current stress analysis of power semiconductor switches: The voltage stresses of the semiconductor S and the clamping diode Dc are equal to the voltage on the clamping capacitor Cc:

$$V_s = V_{dc} = V_{cc} = \frac{V_{in}}{1-D} = \frac{V_{out}}{N+2} \quad (5)$$

From Eq. 6, it can be given as that the voltage stress across the switch is decreased to the maximum as the turn's ratio move towards the maximum which makes high performance MOSFETs with low RDS ON available to increase the circuit actions in a better way (Vijayalakshmi *et al.*, 2016). The maximum switching voltage is only half of the high output voltage. The voltage stress of the regenerative diode Dr is the same as that of the output diode Do which is equal to the output voltage minus the clamp capacitor voltage:

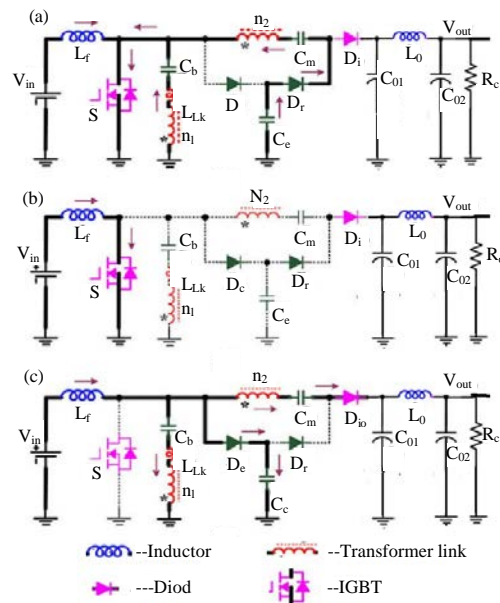


Fig. 4: a-c) Mode of operation output

$$V_{Do} = V_{Dm} = V_{out} - V_{co} = \frac{N+1}{1-D} V_{in} = \frac{N+1}{N+2} V_{out} \quad (6)$$

As a result, the average current of the diodes, Dr and Do is the same. In order to simplify the analysis, the switch turn-ON period is assumed to be half of the resonant period which is given as DTs ~ Tr /2. The peak current of the diodes Dc, Dr and Do can be derived by Eq. 7 and 8:

$$I_{Dc_{peak}} \approx I_{in} \approx \frac{(N+2)}{(1-D)} I_{out} \quad (7)$$

$$I_{Dr_{peak}} \approx \frac{\pi T}{T} I_{out} \approx \frac{\pi}{2D} I_{out} \quad (8)$$

$$I_{Do_{peak}} \approx \frac{I_{Lf}}{N+1} \approx \frac{I_{in}}{N+1} \approx \frac{(N+2)}{(N+1)(1-D)} I_{out} \quad (9)$$

And the switch peak current and RMS current are calculated by Eq. 9-11:

$$I_{s_{peak}} \approx I_{in} + \frac{\pi T_s}{T_r} I_{out} \approx \frac{(N+2-\pi)D + \pi}{2(1-D)D} I_{out} \quad (10)$$

$$\frac{1}{T_s} \int_0^{T_r} \left[\left(\frac{N+1\pi T_s I_{out}}{T_r} \right) \sin(Wrt) + I_{in} \right] \quad (11)$$

From the static analysis operation, it can be routed from the clamp diode current is reduced to zero before it

turns OFF that gives an output related to no reverse-recovery problem for the clamp diode. For the output diode, its turn-OFF current dragging rate is controlled by the inductance leakage of the inbuilt-transformer which is sated by Eq. 12:

$$\frac{di_{Do}(t)}{dt} = \frac{(N+1)V_{out}}{N(N+2)L_{lk}} \quad (12)$$

For the condition of regeneration, there are two different cases lies based on the operation modes. In O_{rf} operation, the current across the regenerative diode is resonated towards zero before it turns OFF without any recovery problem. In B_{rf} mode, its turn-OFF current falling rate is given by Eq. 13:

$$\frac{di_{dr}(t)}{dt} = \frac{V_{out}}{(N+2)L_{kk}} \quad (13)$$

From Eq. 13, it can be coated that the reverse-recovery problem can be sifted effectively by shifting a small leakage inductance as the turn's ratio increases, which can decreases the electromagnetic interference noise and develop the circuit efficiency.

RESULTS AND DISCUSSION

Circuit performance comparison and evaluation: In a nutshell, the next solution is to introduce the built-in transformer where its change in turns ratio is used to control and convert the output voltage. Added to it the switched capacitor technique can be added with the coupled inductor or in built-transformer to derive improved high step-up converters (Puviarasi and Dhanasekaran, 2015) which can further increase the voltage gain and reduction in switching voltage stress.

Related to the nowadays publication of journals, the coupled inductor-based high step-up converters can be related to the summation of the coupled inductor and switched capacitor techniques (Devikar *et al.*, 2016; Bharathi and Sasikumar, 2016) which can be derived by the voltage doubler, tripler and quadruple and by using this circuits adding in a cascaded manner more boosting voltage can be attached with the concepts of reduced switching losses, the switch voltage stress can be reduced correspondingly. Addition to, it the leakage energy can be processed by employing active filter scheme which cannot only suppress the possible turn-OFF voltage spikes on the power MOSFETs but also provide ZVS or ZCS soft-switching operation.

Consequently, a trade-off should be made to optimize the converter performance (Devikar *et al.*, 2016) based on the application specifications. In the recent trend, renewable energy plays a vital role in the production of electricity. In this regards, PV plays a major role but the power transfer from the PV need a voltage boosting circuit, so that, it can feed with a constant rated voltage towards the load, the conventional multiplier circuits are very bulky with separate transformer and large number of switch's in turn the cost control is going to be higher in this proposed system all the drawbacks are overcome by using the built in transformer here, the basic concepts of inductor charging and discharging are used by the way the voltage doubler, tripler quadruple concepts are achieved by means of a single switch.

Design analysis of key parameters

Turns ratio selection in built in transformer: The determination of duty cycle is decided by the turn's ratio of the transformer and it determine the switching and voltage rate, the stress across each switch is designed by the turns ratio only given by Eq. 14:

$$N = \frac{V_{out}}{V_{in} (1-D)^2} \quad (14)$$

Design of leakage inductance parameter present in built in transformer: Leakage inductance relation is given and the diode current falling rate is derived by Eq. 15:

$$\frac{di_{Do}(t)}{dt} = \frac{(N+1)V_{out}}{N(N+2)L_{Lk}} \quad (15)$$

Selection of input inductor towards filter: The input filter inductor L_f is designed to make that the input current ripple is approximately 20% of the average input current. This is derived by Eq. 16:

$$L_f = \frac{V_{in}D}{\Delta I_{L_f}fs} \quad (16)$$

Due to the built-in voltage multiplier cell of the proposed converter (Wu *et al.*, 2008), the switch duty cycle can be optimized to reduce the input current ripple which can prolong the usage life of the PV arrays or the fuel cells.

Power device selection: The voltage and current rating of the load decides the power supply circuits. It may be varies with raise in load vice versa.

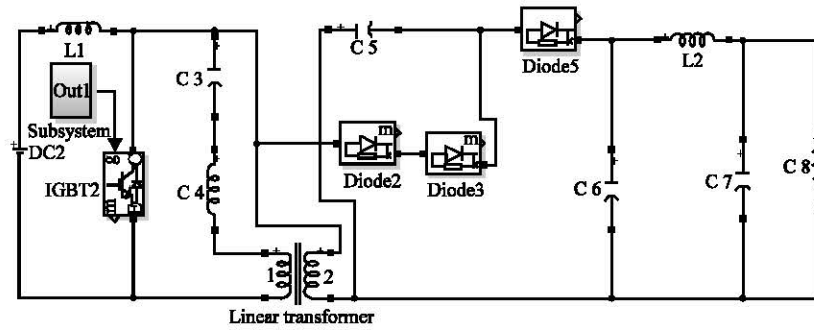


Fig. 5: Voltage multiplier cell converter simulation diagram

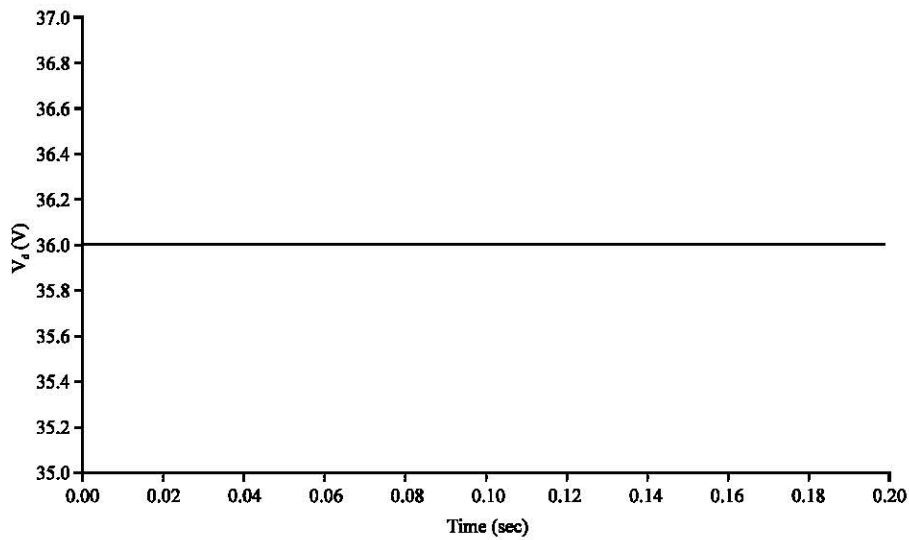


Fig. 6: DC input voltage waveform

Design of capacitor: Consideration for the capacitor design. The relationship between the voltage ripple and the output power can be derived by Eq. 17:

$$C \geq \frac{P_{out}}{V_{out} \Delta V_{cF}} \quad (17)$$

where, Δv_c is the maximum applicable voltage ripple on the capacitor C_c or C_m . The different operating of the circuits results decides the value of blocking capacitor.

Simulation of voltage multiplier cell: The resonant converter circuit with voltage multiplier cell diagrams is shown in Fig. 5 and the corresponding output results are obtained below. Figure 6 represents the DC input voltage waveform. Figure 7 represents the wave form V_G , I_D and I_{DS} across each switch for proposed system. Figure 8 represents the DC output voltage waveform (380 V).

Figure 9 represents the power output wave form (500 W). Figure 10 represents the DC output current wave form (1.1 A).

Efficiency of the proposed converter: The measured efficiency of the proposed converter is plotted and the maximum efficiency of 96.6% on the other side the of conventional circuit efficiency of 95.6% at 500 W max loads are obtained with $C_b = 6.9 \mu F$. The resulted efficiency comparison with different input voltages and $C_b = 6.9 \mu F$ is given in Fig. 11.

It can be seen that even the input voltages have a large variations shown in Fig. 12. Within a large load range the conversion efficiency can be kept up to a high level. The maximum efficiency is still over 96% even though the input voltage is reduced to 30 V.

Hardware prototype: By using the various components with the specified parameters cited in Table 1. The

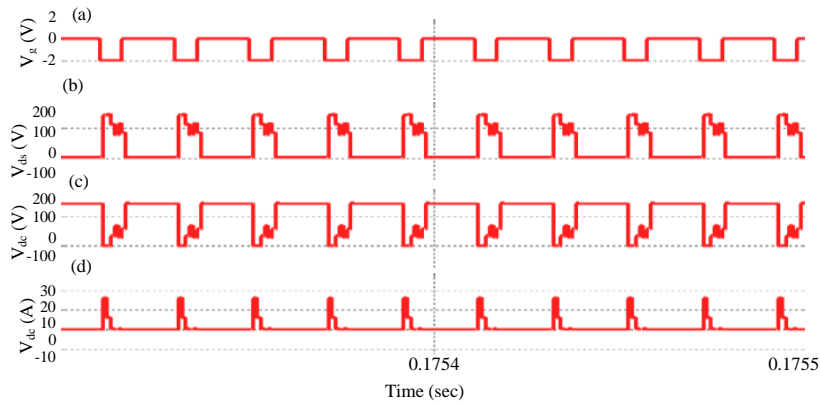


Fig. 7: a-d) Wave form VG, ID and IDS across each switches for proposed system

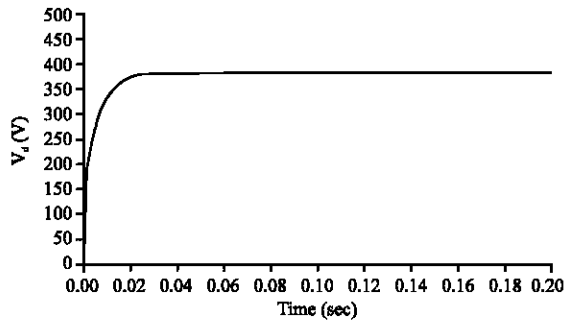


Fig. 8: DC output voltage waveform (380 V)

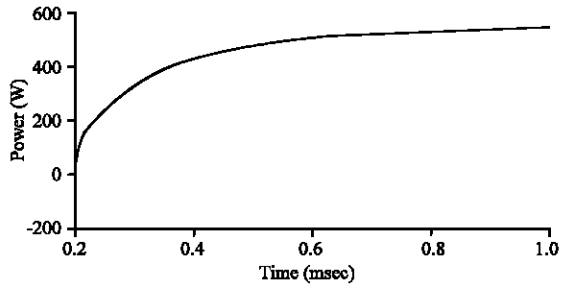


Fig. 9: Power output waveform (500W)

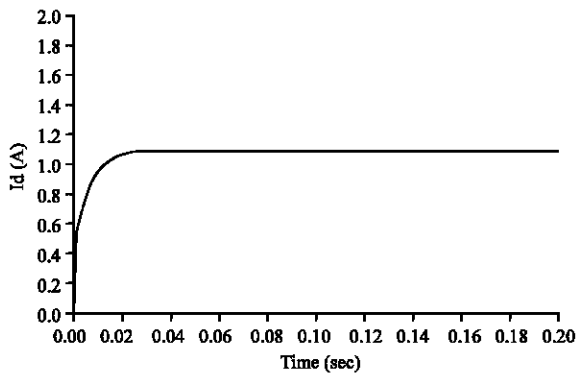


Fig. 10: DC output current waveform

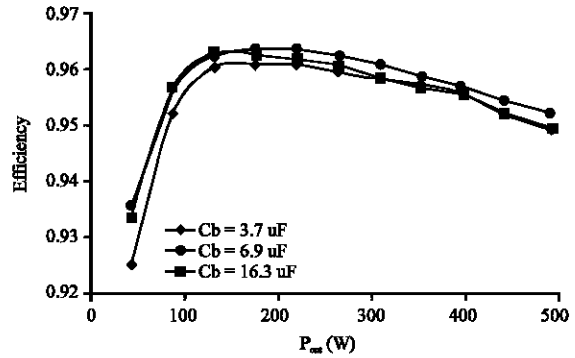


Fig. 11: Efficiency comparison for C, LC and Pi filter

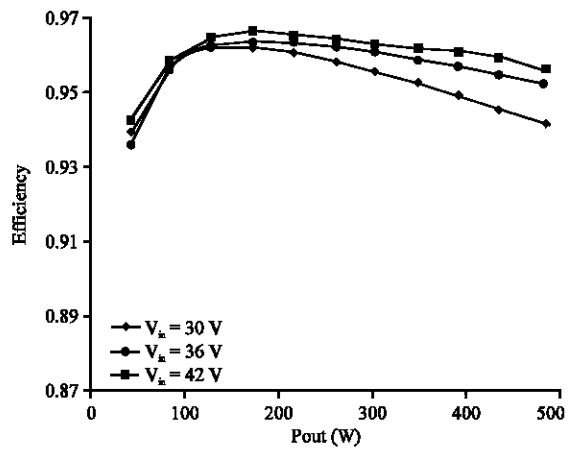


Fig. 12: Efficiency comparison for different input voltage Cb = 6.9 μ F

prototype hardware model was developed and its was executed with practical condition of input voltage and the boosting voltage N times was achieved and is proved. The details of hardware prototype are shown in Fig. 13-15.

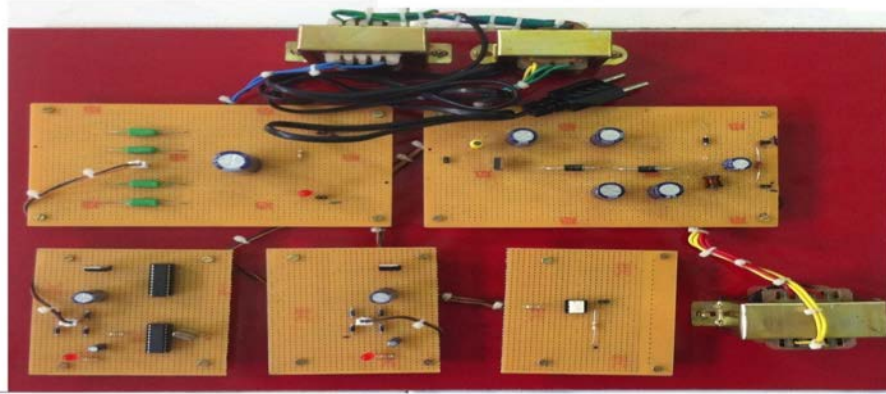


Fig. 13: Prototype



Fig. 14: Input voltage-prototype result



Fig. 15: Output voltage-prototype result

Table 1: Components and parameters of prototype

Components	Parameters
Vin input voltage	23V
Vout (output volatge)	241V
Pout maximum output power	500 W
Fs (switching frequency)	100 kHz
Lf (input filter inductor)	100 μ Henry
Turns ratio	17/7
Lm magnetizing inductor	260 μ Henry
LLK leakage inductor	1.6 μ Henry
S power MOSFET	IRFP4227
DC clamp diode	MUR820
Dr and Do diode	MURI560
Cc clamp capacitor	2.2 μ F
Cb block capacitor	3.7,6.9 or 16.3 μ F
Cm switched capacitor	1 μ F
Co1 (output capacitor 1)	100 μ F
Co1 (output capacitor 2)	100 μ F
Lo (output inductor)	57 μ Henry
Ro (load resistor)	350 Ω

CONCLUSION

The infused transformer-less single cell boost converter was stimulated using ZCS resonant techniques with the help of MATLAB tool and the results are compared with the various types of filter such as C, PI and LC filter and the efficiency comparison was done with various profiles like voltage, current and power including the amount of THD and finally conclude that the PI filter sounds good in the comparison of all profile and the comparison results leads in the development of hard ware of one switch voltage builder with PI filter. The result output proves the above statement.

REFERENCES

Bharathi, K. and M. Sasikumar, 2016. Voltage compensation of smart grid using bidirectional intelligent semiconductor transformer and PV cell. *Indian J. Sci. Technol.*, 9: 1-8.

Devikar, R.N., D.V. Patil and V. Chandraprakash, 2016. A soft computing approach to improve the network performance. *Indian J. Sci. Technol.*, 9: 1-8.

Do, H.L., 2010. A soft-switching DC-DC converter with high voltage gain. *IEEE. Trans. Power Electron.*, 25: 1193-1200.

Erickson, R.W. and D. Maksimovic, 2001. *Fundamentals of Power Electronics*. 2nd Edn., Wolters Kluwer, New York, USA., ISBN:0-7923-7270-0, Pages: 883.

Garcia, O., J.A. Cobos, R. Prieto, P. Alou and J. Uceda, 2003. Single phase power factor correction: A survey. *IEEE. Trans. Power Electron.*, 18: 749-755.

Prasanthi, G., V. Sumalatha and A. Sudheer, 2010. Optimum design of boost converter for charging the battery by non-conventional energy source. *Indian J. Sci. Technol.*, 3: 724-726.

Puviarasi, R. and D. Dhanasekaran, 2015. Interleaved boost converter fed DC machine with zero voltage switching and PWM technique. *Indian J. Sci. Technol.*, 8: 376-382.

Santhoshi, B.K., K.M. Sundaram, M. Sivasubramanian and S. Akila, 2016. A novel multiport bidirectional dual active bridge DC-DC converter for renewable power generation systems. *Indian J. Sci. Technol.*, 9: 1-4.

Shenkman, A., Y. Berkovich and B. Axelrod, 2004. Novel AC-DC and DC-DC converters with a diode-capacitor multiplier. *IEEE. Trans. Aerosp. Electron. Syst.*, 40: 1286-1293.

Vijayalakshmi, M., R. Ramaprabha and G. Ezhilarasan, 2016. Design of auxiliary resonant boost converter for flywheel based photovoltaic fed microgrid. *Indian J. Sci. Technol.*, 9: 1-6.

Wu, T.F., Y.S. Lai, J.C. Hung and Y.M. Chen, 2008. Boost converter with coupled inductors and buck-boost type of active clamp. *IEEE. Trans. Ind. Electron.*, 55: 154-162.

Analysis of Vocal Tract Shape Variability for Different Conditions in Indian English Vowels

¹Anil Kumar Chandrashekar and ²M.B. Manjunatha

¹Department of Electronics Engineering, Jain University, Bangalore, India

²Akshaya Institute of Technology (AIT), Tumkur, India

Abstract: The study presents vocal tract shape estimation for vowels. Autoregressive model of vocal tract has been used herein for different recording conditions. Vocal tract shapes are known for phonetic distinction and speaker exclusiveness among vowels found by individual speakers ingrained in them. The researchers propose to use this autoregressive model for utterances made by speakers from south India taken on different occasions and in various recording conditions. Samples of speeches have been recorded assuming ice cold water and having an instant check of its effect and again later after 5 min. Vocal tract shapes so obtained are compared to those recorded under normal conditions. Vowel utterances made by speakers have been recorded 30 times, variations in the resulting vocal tract for different conditions on intra speaker basis. The researchers propose to investigate those and also, the variability as a parameter for the speech recognition process. The entire process has been done using MATLAB.

Key words: Auto regressive model, digital signal processing, dynamic model, LPC, vocal tract, process

INTRODUCTION

That speech is the most accepted and convenient means of communication is well known and recognized. The narrow concept of speech is that it is just a sequence of sounds punctuated by abrupt changes happening from one to another or some signals that are ignored and go into oblivion soon after utterance. It is much more than that. It is a unique signal that conveys information of linguistic and non linguistic type. Such information conveyed by speech is for knowledge of multiple levels speech signals typify (typify) information bearing that come up as a function of a single independent variable like time. Speech is not just an information signal. It is something beyond that it is actually a complex wave and acoustic output arising as a result of the speaker's effort (Kumar *et al.*, 2015a-c).

Speech analysis is synonymous with feature extraction of speech. Speech sounds are sensations of air pressure variations produced by exhaled air and later modulated and shaped by vibration of glottal cords and the resonance of vocal tract during the time air is pushed out through the lips and the nose. Speech is signal with information galore exploring frequency modulated amplitude modulated and time-modulated carriers

(example: resonance movements, harmonics, noise, pitch, power, duration). The objective is to convey information on words, speaker identity, accent, speech style and emotion. The entire gamut of information is basically conveyed in the large of the traditional telephone bandwidth of 4 kHz. Speech energy 4 kHz reflects audio quality and sensation (Prasad *et al.*, 2015a, b).

Speech production: Speech is meant for communication. It has the distinct feature as a signal that carries a message or information. It is known to be an acoustic waveform that carries a message or information. It is known to be an acoustic waveform that carries temporal information from the speaker to the listener. Efficiency underlies acoustic transmission and reception of any speech. But this is applicable only for transmission over a short distance. There is a spread of radiated acoustic energy at frequencies that are used by the vocal tract and ear. But this gets reduced in intensity rapidly. Even on occasions when the source is able to produce substantial volume of acoustic power there is a support of only a fraction there of by the medium without any distortion while the rest of the it gets squandered in air dust particles molecular disturbance. It is also, resulting in getting over aero molecular viscosity. The ambient

acoustic noise places a restriction or limit on the sensitivity of the ear. Physiological noises to play this role in and around the ear drum. Voluntary, formalized motions of the respiratory and masticators apparatus have the speech as the acoustic end product. The closed loop has the ability to develop, control, maintain and correct it. Acoustic feedback of the hearing mechanisms and the kinesthetic feedback of the speech musculature too have a role here. The central neurons system organizes and coordinates information from the senses which is then used for directing the function as also for delivering the descend, linguistically dependant, vocal articulator motion and acoustic speed.

Problem statement: The vowel speech database considered in our study is by considering all five vowels of English at three different conditions. Database is created by us at normal recording room environment, vowels (a, e, i, o, u) to be analyzed with respect to vocal tract shape variability.

The speech communication pathway: Figure 1 gives a simplified view of speech communication route starting from the speech and reaching the listener. An idea or a concept originates at the linguistic level of communication in the speakers mind. It is the stimulated longitudinal acoustic wave propagation in air starting from the speaker and terminating with the listener.

Pressure changes within the vocal tract are caused by the vocal tract and vocal cord movement. These are seem more specifically at the lips, initiating the sound wave that is known for propagation in space, this propagation activity occurs through space as a sequence of compression and rarefaction of air dust molecules. As a result, temporal pressure variations are noticed at the listener's exterior ear which is funnel shaped collecting this acoustic energy efficiently. Later it manages to carry the media vibration ultimately to the final vibration sensor, the ear drum set in the interior ear (Prasad *et al.*, 2015a, b).

Variation in pressure experienced at the speakers lips, speakers lips causes sound. This sound propagates with channel losses, resulting in pressure variations at the listeners outer ear. The eventual vibrations in the ear-drum induces electric signals which move along the sensory nerves to the brain. To the extent of the listeners perception the brain decodes these electrical signals known for the sensitivity to language. Later, it filters these signals in a recognized pattern which becomes known as a language speech perception and hearing.

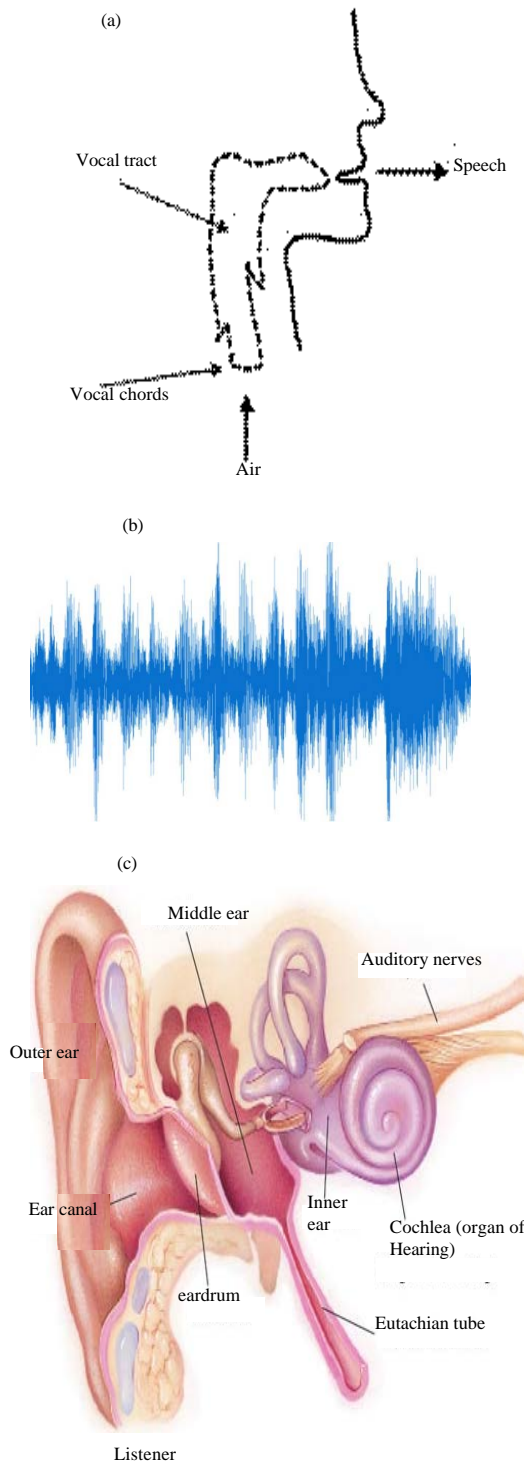


Fig. 1: Speech communication pathway: a) Speaker; b) Speech signal and c) Listener

Vocal tract shape: It is possible to model vocal tract as an acoustic tube resonance referred to as formants and anti-resonances and as a chain of cylinders of cross sectional area with varying features. Alteration in the form of the acoustic tube is caused by movement of articulators of vocal tract resulting in changes in frequency response. Poles are caused by resonances while anti-resonances arise from zeroes of frequency response. Kumar *et al.* (2015a-c) articulators are structures in the vocal tract which moves in the production of sounds of varying decibels. The important articulators are the lips and the tongue. They can be classified as direct and indirect the later are the jaw and mandible helping the positioning of the tongue and the lips for a variety of sounds such positioning of the tongue creates a number of vocal tract shapes that are needed for the production of speech sounds. The tract length for an adult male is generally 17 cm running from glottis to lips, the cross sectional area being up to 20 cm² (Prasad *et al.*, 2015a, b).

Voiced sound: Vocal fold vibrations along with positioned articulators generates the basic sound which is referred to as “voiced sound”. Voiced sound in singing is significantly different from voiced sound in speech.

Resonance: Voiced sound, frequency selected, modified and amplified by the vocal tract resonators (the throat, mouth cavity and nasal passages) and helped by articulators to produce a person’s recognizable voice.

Unvoiced: The basic sound generated by vocal fold is apart and not vibrations with positioned articulators are called “Unvoiced sound”. In the event of the vocal cords getting tensed up and closed, air flow experiences obstruction and air pressure builds up behind the constriction. This highly compressed air passes through the constriction in the vocal tract, becoming turbulent and producing what is known as unvoiced sound.

Articulation: The vocal tract articulators (the tongue, the jaws, the cheek, soft palate, the lips and the hyoid bone) have the function of modifying the voiced/unvoiced sound, they produce recognizable words helped by articulation.

Speech signals: Excitement of a fixed vocal tract produce vowels with quasi-periodic pulses of air, forced through the vibrating vocal cords. A quasi-periodic puff of air flow is the source, acting through vibrating vocal folds at a

definite basic frequency. The term “quasi” is used considered with perfect periodicity never being achieved. The term “periodic” is used henceforth. This shape of vocal-tract length from glottis to lips determines the resonant frequencies of the tract there by defining the produced sound.

MATERIALS AND METHODS

Concatenated tube model: The widely used model for speech production is based on the assumption of the capacity of the vocal tract to represent as a concatenation of small cylindrical tubes (Fig. 2 and 3). Independent variation of the cross sectional area of any tube for stimulating the changing shape of the vocal tract is seen. A change in the length of any tubular segment is effected for the purpose of reflecting the movements of articulators which include the lips, the jaws, the cheeks, the tongue and the hyoid bone. This variations in the shape and length of the tract at different points along its length, ultimately causes production of different sounds (Shah and Pandey, 2006).

Digital Signal Processing (DSP) techniques find use in modeling, use of the speech signal converted into its discrete time sequence on the assumption of all cylindrical segments being small are shown in Fig. 2 as of equal length.

Linear Predictive Coding (LPC): Linear predictive analysis is among the most powerful and widely used speech analysis techniques. This method has emerged as a predominate technique for estimating the basic speech parameters, e.g., pitch, formants, spectra, vocal tract area function. Render this method as highly important ability to provide accurate estimates of the speech parameters and in relative speed of computation. (Kodandaramaiah *et al.*, 2010) .The ability of approximation seen by a speech sample as a linear combination of past speech samples is the basic concept that explores linear prediction analysis. The determination of a unique set of predictor coefficients is done by reducing the sum of the squared differences (over a finite

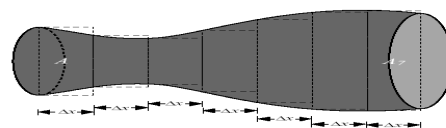


Fig. 2: Concatenated Tube Model

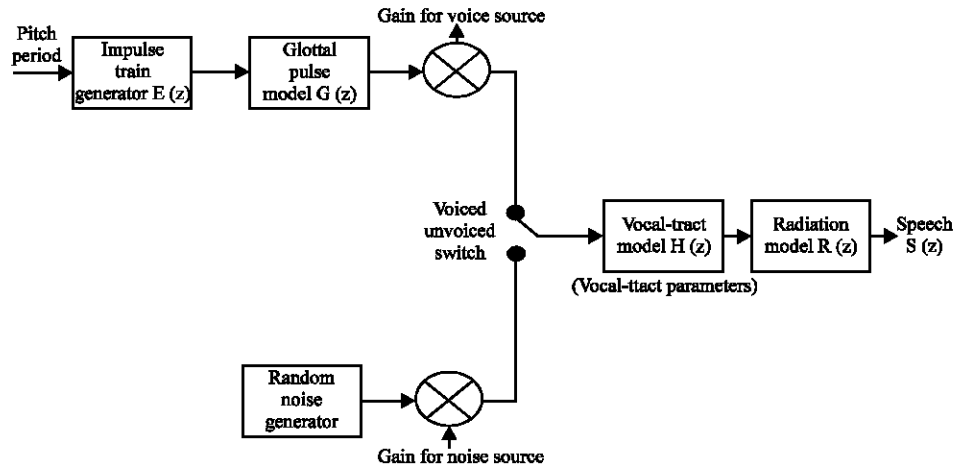


Fig. 3: Simplified model for speech production

interval) between the actual speech and the linearly predicted ones to the minimum. The predictor coefficients are the weighted coefficients used in the linear combination (Prasad *et al.*, 2015a, b; Kodandaramaiah *et al.*, 2010). Figure 3 shows the simplified model of speech production for providing idea of linear prediction is intimately linked to the basic speech synthesis model in which the sampled speech signal was shown as modelled as the output of a linear, slowly time-varying system excited by either quasi-periodic impulses (during voiced speech) or random noise (during unvoiced speech). The linear prediction method provides a robust reliable and accurate method for estimating the parameters that characterize the linear, time-varying system an all pole system function describes the linear system over intervals of short duration. A time varying digital filter with known steady state system function represents the composite spectrum effects of radiation, vocal tract and glottal excitation.

Speech database: The 30 samples of 30 subjects (male speakers aged about 18-25) at different times and at different conditions, i.e., normal sample by the instant of consuming ice cold water and with time lapse of 5 min were recorded using table top mic make of i-ball Model No. M27 with sensitivity -58 ± 3 dB, frequency response of 100-16 kHz with sampling frequency of 22 and 100 Hz in MATLAB and the samples were normalized to have amplitude normalization by considering English vowels by Indian English speakers (a, e, i, o, u) (Kumar *et al.*, 2015a-c).

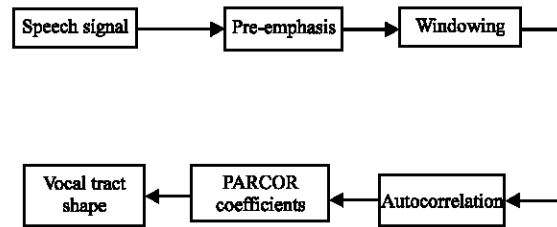


Fig. 4: Block diagram for vocal tract shape calculation

Block diagram for vocal tract shape calculation:

Figure 4 shows the various steps involved in the estimation of the vocal tract shape by providing speech signal as input and with various co-efficient calculation and its description as follows.

Pre-emphasis: The area function obtained through use of reflection coefficients cannot be the area function of the human vocal tract. In the event of the pre-emphasis being used prior to linear predictive analysis for removal of defects arising out of glottal pulse and radiation, the area function that emerge as results are often similar to vocal tract configurations that find use in human speech. The speech production model discloses the speech undergoing a special tilt of -6 dB/octave. The use of pre-emphasis in displaying the spectrogram if speech signals is rather common (Khodai-Joopari *et al.*, 2004).

Window technique: The window function $w(k-n)$ is a real window sequence that finds use in isolation of isolate the portion of the input sequence analyzed at a particular time index k . Many window functions have been generated to

improve upon the basic rectangular window design such as hamming, hanning, bartlett, blackman, kaiser, etc. each having different specification with regards to its frequency response. Laprie and Mathieu (1998) In this study, LP analysis was performed on frames weighted with the hamming window. This window, $w(n)$ was chosen as it provides a good balance between its main lobe width and side lobe attenuation.

The hamming window is also deemed to be adequate in determining the accuracy for approximating the transfer function of the vocal tract. This is a crucial aspect when calculating reflection coefficients for quantization purposes (Black, 1998).

Autocorrelation method: The most widely used method to linear predictive analysis is called the autocorrelation method. It is the most popular method of short-term LP analysis. This method provides the most computationally efficient manner in determining the LP parameters with guaranteed stability. It avails of the Toeplitz property possessed by the autocorrelation matrix.

PARCOR coefficients: PARCOR coefficients are bounded ± 1 . This has been seen earlier. This features has given them the facility of being the attractive parameter for quantization. With a set of PARCOR coefficients, it is first possible to use them at step 2 of the Levinson Durbin algorithm followed by the getting of an algorithm that can help conversion which can be computed from a given set of predictor coefficients by working backward through the Levinson Durbin algorithm (Kuc *et al.*, 1985).

Algorithm; Intra speaker vocal tract shape algorithm:

1. Using LPC (AR-Auto Regressive model) forward and backward and speech analysis determining the shape of vocal tract variability of the subject bringing out the vocal tract shape variability of an individual subject
2. Study of variability of the above vocal tract shape among 30 different subjects to high light and identify Intra speaker variability
3. The above identified can be adopted for personal identification similar to signature. It can also be named as vocal tract signature of an individual (Black, 1988)
4. Find the worst (with least variability) and the best maximum variability for each speaker (subject)
5. Finding the averages in Time for best and worst patterns for the ensure of 30 subjects (Paige and Zue, 1970)
6. Plot the resultant worst pattern and resultant best pattern for a subject for the vowel

RESULTS AND DISUSSION

Figure 5 shows the speech signal for vowel /e/ under normal recording condition, Fig. 6 depicts the speech signal vowel /e/ in consumption of ice cold water and Fig. 7 represents speech sample vowel /e/ by after

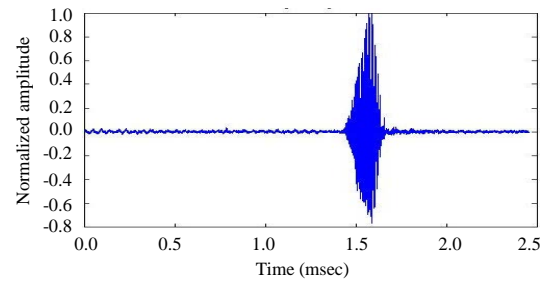


Fig. 5: Speech signal for normal sample /e/; Original signal

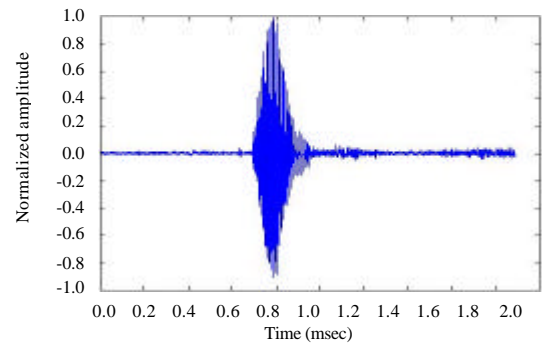


Fig. 6: Speech signal for sample /e/ after consumption of ice water; Original signal

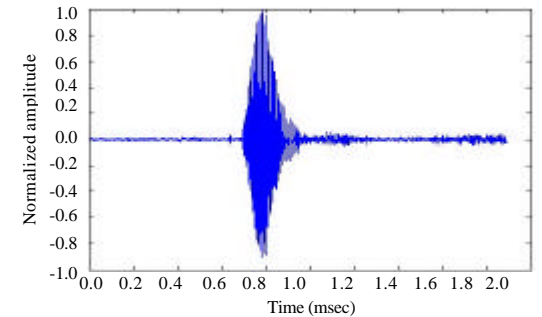


Fig. 7: Speech signal for speech sample /e/ after 5 min of consumption of ice cold water; Original signal

five min of consumption of ice cold water in the similar passion the speech signals are plotted for other vowels like a, i, o and u, under all the three recording conditions as mentioned above and the variations has been observed simultaneously.

Few noticeable variations can be observed while interpreting the same in the dynamic modelling of the vocal tract as plotted in the Fig. 8a vocal tract model of /e/ Fig. 8b vocal tract model of /a/ Fig. 8c vocal tract model of /i/ Fig. 8d vocal tract model of /o/ Fig. 8e vocal tract model of /u/.

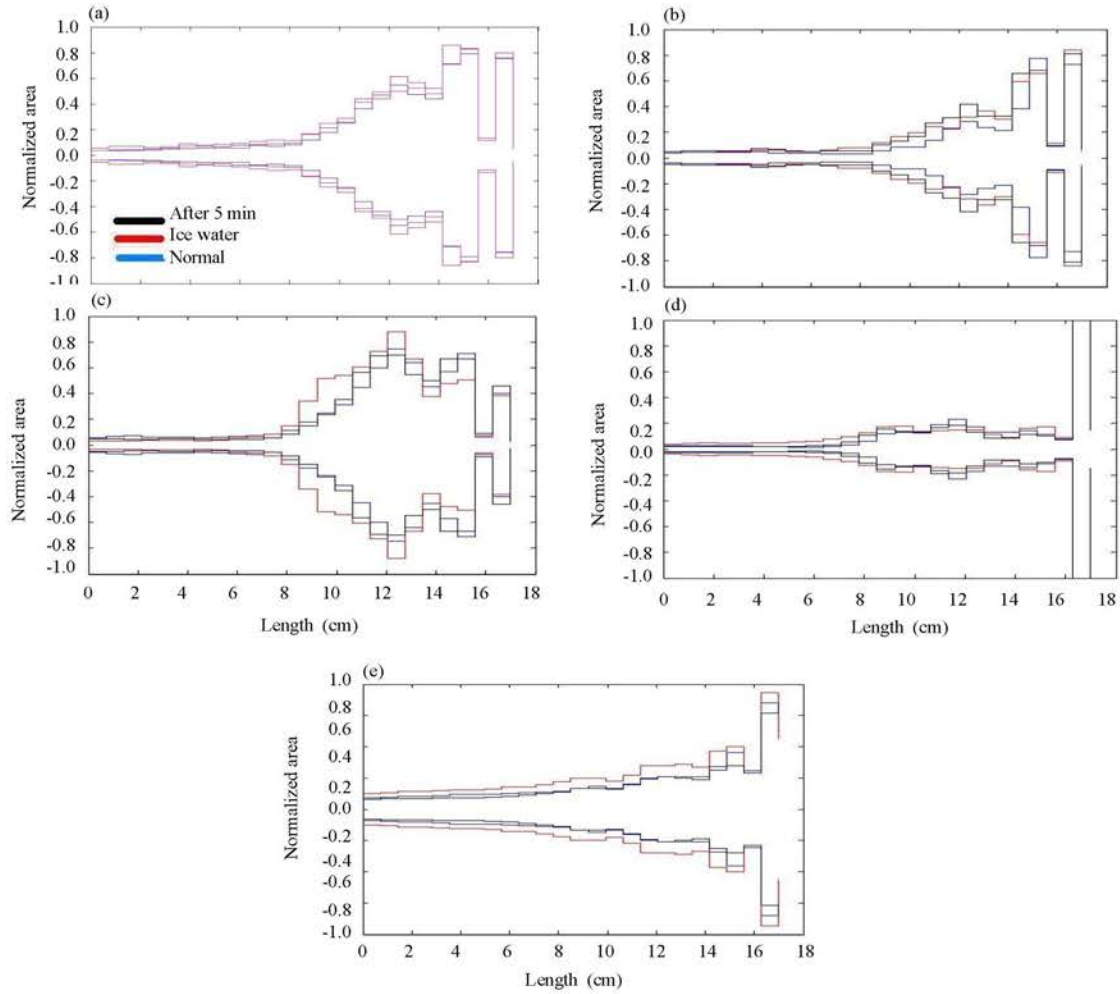


Fig. 8: a-e) Dynamic vocal tract model for vowel /e/; Dynamic vocal tract model for vowel /a/

As observed from the dynamic model in Fig. 8a the significant variability in vocal tract is observed in the region of 14-16 cm for the vowel /e/. From Fig. 8b, vocal tract will swing dominently in the area labelled 8-12 cm for the vowel /a/. In Fig. 8c, model represented for the vowel /i/ shows the clear variations for the length 10-14 cm. In regard of Fig. 8d, dynamic model of vowel /o/ the variability in the vocal tract shape observed between 4-10 cm and lastly Fig. 8e gives the modeling of vowel /u/ by showing the deviation of model in the full length till 17 cm.

CONCLUSION

Using LPC analysis of speech, reflection coefficients are utilized for the estimation of vocal tract. Vocal tract

length values are obtained for the vowels (a, e, i, o, u) for individual speakers. The speech input signal is separated into frames each frame length into 30 msec with an overlapping of 10 msec frame length the order of LPC filter is 25. Sampling rate of the speech signal is 22,100 Hz. Each frame consists of 663 samples. By using LPC method calculating the reflection coefficients. The incidence of the largest reflection of coefficients is seen in areas where the relative changes in vocal tract are the largest. Measurement of the area function of an individual taken as different occasions and in different contexts as also, the variability of the vocal tract shapes. The measurement is done on intra and inter speaker basis model getting vocal tract shape if the vowels for every speech. Vocal tract shape can be used as personal passwords or signature for verification and identification.

REFERENCES

- Black, N.D., 1988. Application of vocal tract shapes to vowel production. Proceedings of the IEEE Annual International Conference on Engineering in Medicine and Biology Society, November 4-7, 1988, IEEE, New Orleans, Louisiana, pp: 1535-1536.
- Khodai-Joopari, M., F. Clermont and M. Barlow, 2004. Speaker variability on a continuum of spectral sub-bands from 297-speakers non-contemporaneous cepstra of Japanese vowels. Proceedings of the 10th Australian International Conference on Speech Science and Technology, December 8-10, 2004, Macquarie University, Sydney, New South Wales, pp: 504-509.
- Kodandaramaiah, G.N., M.N. Giriprasad and M. Mukundarao, 2010. Implementation of LPC based vocal tract shape estimation for vowels. Technol. World Q. J., 5: 97-102.
- Kuc, R., F. Tuteur and J. Vaisnys, 1985. Determining vocal tract shape by applying dynamic constraints. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'85) Vol. 10, April 26-29, 1985, IEEE, Tampa, Florida, USA., pp: 1101-1104.
- Kumar, A., G.K. Ramaiah and M.B. Manjunatha, 2015a. Speaker based vocal tract shape estimation for kannada vowels. Proceedings of the International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), January 24-25, 2015, IEEE, Visakhapatnam, India, ISBN:978-1-4799-7676-8, pp: 1-6.
- Kumar, A., M.B. Manjunatha and G.K. Ramaiah, 2015b. Vocal tract shape estimation of vowels & CVVC for diversified Indian English speakers. Proceedings of the International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), January 24-25, 2015, IEEE, Visakhapatnam, India, ISBN:978-1-4799-7676-8, pp: 1-7.
- Kumar, C.A., K.M.S. Prasad, M.B. Manjunatha and G.N.K. Ramaiah, 2015c. Basic acoustic features analysis of vowels and C-V-C of Indian english language. ITSI. Trans. Electr. Electron. Eng., 3: 20-23.
- Laprie, Y. and B. Mathieu, 1998. A variational approach for estimating vocal tract shapes from the speech signal. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing Vol. 2, May 15, 1998, IEEE, Seattle, Washington, USA., pp: 929-932.
- Paige, A. and V. Zue, 1970. Calculation of vocal tract length. IEEE. Trans. Audio Electroacoust., 18: 268-270.
- Prasad, K.M.S., C.A. Kumar, M.B. Manjunatha and G.N.K. Ramaiah, 2015a. Gender based acoustic features and spectrogram analysis for kannada phonetics. ITSI. Trans. Electr. Electron. Eng., 3: 16-19.
- Prasad, S., A. Kumar and K. Ramaiah, 2015b. Various front end tools for digital speech processing. Proceedings of the 2nd International Conference on Computing for Sustainable Global Development (INDIACom), March 11-13, 2015, IEEE, New Delhi, India, ISBN:978-9-3805-4415-1, pp: 905-911.
- Shah, M.S. and P.C. Pandey, 2006. Estimation of vocal tract shape for VCV syllables for a speech training aid. Proceedings of the IEEE-EMBS 27th Annual International Conference on Engineering in Medicine and Biology Society, January 17-18, 2006, IEEE, Shanghai, China, pp: 6642-6645.

Internet of Things-Security and Trust in e-Business

¹Ali Shawket Thiab, ¹Zeratul Izzah Mohd. Yusoh and ²Abdul Samad Bin Shibghatullah

¹Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia, Melaka, Malaysia

²Faculty of Business and Information Science (FoBIS), UCSI University,
Kuala Lumpur (SouthWing), Malaysia

Abstract: The rapid evolution of computing and communication technologies and their standardizations have made the boom in e-Commerce possible. Lowering of the cost of operation, increase in the speed of transactions and easy global reach to customers and vendors have been the reasons for the overwhelming popularity of this new way of commerce. Success of an e-Commerce rests on many factors. e-Commerce is widely considered the buying and selling of products over the internet but any transaction that is completed solely through electronic measures can be considered e-Commerce. e-Commerce is subdivided into three categories: Business to Business or B2B (Cisco), Business to Consumer or B2C (Amazon) and Consumer to Consumer or C2C (eBay) also called electronic commerce. One of the important contributors is trust. Trust is something that an e-Commerce must strive to achieve over a period of time. Acquiring customer trust depends on many things that an e-commerce controls. However, customer's trust as such is not under the control of the e-Commerce. Then e-Commerce security is viewed as an engineering management problem and a life cycle approach is put forward. How the e-Commerce systems can be made secure using the life cycle approach is outlined.

Key words: e-Commerce security, security issues, internet consumers, trust, categories, electronic measures

INTRODUCTION

Nowadays as people are gradually aware of the contributions that reverse logistics has made to environment and economy; Enterprises are paying more attention to the promotion of reverse logistics. However, reverse logistics is different from forward logistics: the difficulties in application have aroused wide attention. The reason that leads to the delay in application of reverse logistics is heatedly debated and the academy field holds different views about it but it is an acknowledged fact that collecting reverse logistics information is difficult. The reason is that reverse logistics involves uncertainty of all attributes in products and this demands diverse reverse logistics processes. If we ignore the uncertainty and use the same process to treat all products, then we will waste many recyclable, valuable products, thus, reducing the profits created by reverse logistics in a large scale. To make up for it, we need to know the information of different products in the whole close-loop supply chain promptly and accurately. After Daugherty and Jalal Ashayeri confirmed that information system and reverse logistics performance was positively correlated through empirical study in 2002 and 2005, relevant research is more than active (Ashayeri and

Tuzkaya, 2011; Daugherty *et al.*, 2005). Thus, in implementing reverse logistics, we must absorb reverse management information system that can both record product life cycles vertically and record product attributes horizontally, helping enterprises improve decision management and efficiency of business operations.

During the last decade, Internet of Things (IoT) approached our lives silently and gradually, thanks to the availability of wireless communication systems (e.g., RFID, WiFi, 4G, IEEE 802.15.x) which have been increasingly employed as technology driver for crucial smart monitoring and control applications (Atzori *et al.*, 2010; Miorandi *et al.*, 2012; Palattella *et al.*, 2013). Nowadays, the concept of IoT is many-folded, it embraces many different technologies, services and standards and it is widely perceived as the angular stone of the ICT market in the next 10 years at least (Emmerson, 2010; Hersent *et al.*, 2011).

Internet of Things (IoT) also known as internet of objects connects billions of objects which include buildings, air conditioners, coffee machines washers, cars, air planes, animals and people as well. According to Cisco the IoT connects things and people on an unprecedented scale; Cisco predicts that, although, so far in 2015 more than 99% of things in the physical world are not

connected by 2020 the number of internet connected devices and objects will reach 50 billion. With the mixing of physical world and information world together. The future technology can be predicted that the communication is not going to be people communicating to people; It's not going to be people accessing information. It's going to be all about using machines to talk to other machines on their behalf. We are moving towards a new era of ubiquity in terms of technology, we are entering the internet of things era in which new forms of communication among human and things and between things themselves will be recognized (Tan and Wang, 2010).

One of the biggest breakthroughs of the internet of things is transforming the physical world and information world together. Sensors play a very vital role in bridging the gap between the physical world and information world. Sensors collect data from their environment, generating information and raising awareness about context. So that, the change of their environment can be monitored at a glance and the corresponding things can make some responses when needed (Conner, 2010). According to Stankovic (2014) if our vision is correct regarding the future technology, many IoT applications will be based on a deployed sensing, actuation and communication platform (connecting a network of things). In these deployments, it is common for the devices to know their locations have synchronized clocks, know their neighbour devices when cooperating and have a coherent set of parameter settings such as consistent sleep/wake-up schedules, planned power levels for communication and pair-wise security keys (Stankovic, 2014).

From a logical viewpoint an IoT system can be depicted as a collection of smart devices that interact on a collaborative basis to fulfil a common goal. At the technological floor, IoT deployments may adopt different processing and communication architectures, technologies and design methodologies, based on their target. For instance, the same IoT system could leverage the capabilities of a Wireless Sensor Network (WSN) that collects the environmental information in a given area and a set of smartphones on top of which monitoring applications run. In the middle, a standardized or proprietary middleware could be employed to ease the access to virtualized resources and services. The middleware in turn, might be implemented using cloud technologies, centralized overlays or peer to peer systems (Grieco *et al.*, 2014).

Of course, this high level of heterogeneity, coupled to the wide scale of IoT systems is expected to magnify security threats of the current internet which is being

increasingly used to let interact humans, machines and robots in any combination. More in details, traditional security countermeasures and privacy enforcement cannot be directly applied to IoT technologies due to their limited computing power, moreover, the high number of interconnected devices arises scalability issues. At the same time, to reach a full acceptance by users it is mandatory to define valid security, privacy and trust models suitable for the IoT application context (Roman *et al.*, 2013; Anderson and Rainie, 2014). With reference to security, data anonymity, confidentiality and integrity need to be guaranteed as well as authentication and authorization mechanisms in order to prevent unauthorized users (i.e., humans and devices) to access the system. Whereas concerning privacy requirement, both data protection and users personal information confidentiality have to be ensured, since, devices may manage sensitive information (e.g., user habits). Finally, trust is a fundamental issue since, the IoT environment is characterized by different devices which have to process and handle the data in compliance with user needs and rights.

Advantages of e-Commerce: The advantages of e-Commerce for business entities can be summarized thus, e-Commerce can increase sales and decrease costs. A firm can use e-Commerce to reach narrow market segments that are widely scattered geographically. The internet and the web are particularly useful in creating virtual communities that become ideal target markets. A virtual community is a gathering of people who share a common interest but instead of this gathering occurring in the physical world, it takes place on the internet. Just as e-Commerce increases sales opportunities for the seller, it increases purchasing opportunities for the buyer. Businesses can use e-Commerce in their purchasing processes to identify new suppliers and business partners. Negotiating price and delivery terms is easier in e-Commerce because the web can provide competitive bid information very efficiently (Shelanski, 2013).

e-Commerce increases the speed and accuracy with which businesses can exchange information which reduces costs on both sides of transactions. e-Commerce provides buyers with a wider range of choices than traditional commerce because they can consider many different products and services from a wider variety of sellers. The benefits of e-Commerce also extend to the general welfare of society. Electronic payments of tax refunds, public retirement and welfare support cost less to issue and arrive securely and quickly when transmitted via the Internet. Furthermore, electronic payments can be

easier to audit and monitor than payments made by check which can help protect against fraud and theft losses. e-Commerce can make products and services available in remote areas. For example, distance education is making it possible for people to learn skills and earn degrees no matter where they live or what hours of the day they have available for study.

Disadvantages of e-Commerce: e-Commerce also has its disadvantages. It is difficult to conduct a few businesses electronically. For example, perishable foods and high-cost items such as jewellery or antiques may be impossible to adequately inspect from a remote location, regardless of the technologies that are devised in the future. However, most of the disadvantages of e-Commerce today are due to the newness and rapidly developing pace of the underlying technologies. Return on investment numbers is difficult to compute for investments in e-Commerce because the costs and benefits are hard to quantify. Costs which are a function of technology, can change dramatically during even short-lived e-Commerce implementation projects because the underlying technologies change rapidly (Costa, 2016).

In addition to technology issues, many businesses face cultural and legal impediments to e-Commerce. Some consumers are still somewhat fearful of sending their credit card numbers over the Internet. The legal environment in which e-Commerce is conducted is full of unclear and conflicting laws. In many cases, government regulators have not kept up with technologies. As more businesses and individuals find the benefits of e-Commerce compelling, many of these technology and culture-related disadvantages will disappear. Another important issue is security. Transactions between buyers and sellers in e-Commerce include requests for information, quotation of prices, placement of orders and payment and after sales services. The high degree of confidence needed in the authenticity, confidentiality and timely delivery of such transactions can be difficult to maintain where they are exchanged over the internet. The interception of transactions and in particular credit card details, during transmission over the internet is often a major obstacle to public confidence in e-Commerce.

MATERIALS AND METHODS

Security threats to e-Commerce: e-Commerce security requirements can be studied by examining the overall process, beginning with the consumer and ending with the commerce server. Considering each logical link in the "Commerce chain", the assets that must be protected to

ensure secure e-Commerce include client computers, the messages travelling on the communication channel and the web and commerce servers including any hardware attached to the servers. While telecommunications are certainly one of the major assets to be protected, the telecommunications links are not the only concern in computer and e-Commerce security. For instance, if the telecommunications links were made secure but no security measures were implemented for either client computers or commerce and web-servers, then no communications security would exist at all.

Client threats: Until the introduction of executable web content, Web pages were mainly static. Coded in HTML, static pages could do little more than display content and provide links to related pages with additional information. However, the widespread use of active content has changed this perception (Bossomaier and Hope, 2015).

Active content: Active content refers to programs that reembedded transparently in web pages and that cause action to occur. Active content can display moving graphics, download and play audio or implement web-based spread sheet programs (Kadhim and Al-Taie, 2013). Active content is used in e-Commerce to place items one wishes to purchase into a shopping cart and to compute the total invoice amount including sales tax, handling and shipping costs. The best known active content forms are Java applets, ActiveX controls, JavaScript and VBScript. Since, active content modules are embedded in web pages, they can be completely transparent to anyone browsing a page containing them. Anyone can embed malicious active content in web pages. This delivery technique, called a trojan horse, immediately begins executing and taking actions that cause harm. Embedding active content to web pages involved in e-Commerce introduces several security risks. Malicious programs delivered quietly via web pages could reveal credit card numbers, usernames and passwords that are frequently stored in special files called cookies. Because the internet is stateless and cannot remember a response from one web page view to another, cookies help solve the problem of remembering customer order information or usernames or passwords. Malicious active content delivered by means of cookies can reveal the contents of client-side files or even destroy files stored on client computers.

Malicious codes: Computer viruses, worms and trojan horses are examples of malicious code. A trojan horse is a program which performs a useful function but performs an unexpected action as well. Virus is a code segment

which replicates by attaching copies to existing executable. A worm is a program which replicates itself and causes execution of the new copy. These can create havoc on the client side (Nieles *et al.*, 2017).

Server-side masquerading: Masquerading lures a victim into believing that the entity with which it is communicating is a different entity. For example, if a user tries to log into a computer across the internet but instead reaches another computer that claims to be the desired one, the user has been spoofed. This may be a passive attack (in which the user does not attempt to authenticate the recipient but merely accesses it) but it is usually an active attack (Sharma and Misra, 2014).

Communication channel threats: The internet serves as the electronic chain linking a consumer (client) to an e-Commerce resource (commerce server). Messages on the internet travel a random path from a source node to a destination node. The message passes through a number of intermediate computers on the network before reaching the final destination. It is impossible to guarantee that every computer on the internet through which messages pass is safe, secure and non-hostile.

Confidentiality threats: Confidentiality is the prevention of unauthorized information disclosure. Breaching confidentiality on the internet is not difficult. Suppose one logs onto a website, say www.anybiz.com, that contains a form with text boxes for name, address and e-Mail address. When one fills out those text boxes and clicks the submit button, the information is sent to the web-server for processing. One popular method of transmitting data to a web-server is to collect the text box responses and place them at the end of the target server's URL. The captured data and the HTTP request to send the data to the server is then sent. Now, suppose the user changes his mind, decides not to wait for a response from the [anybiz.com](http://www.anybiz.com) server and jumps to another website instead, say www.somecompany.com. The server [somecompany.com](http://www.somecompany.com) may choose to collect web demographics and log the URL from which the user just came (www.anybiz.com). By doing this, [somecompany.com](http://www.somecompany.com) has breached confidentiality by recording the secret information the user has just entered (Geistfeld, 2016).

Integrity threats: An integrity threat exists when an unauthorized party can alter a message stream of information. Unprotected banking transactions are subject to integrity violations. Cyber vandalism is an example of an integrity violation. Cyber vandalism is the electronic

defacing of an existing website page. Masquerading or spoofing, pretending to be someone you are not or representing a website as an original when it really is a fake was one means of creating havoc on websites. Using a security hole in a Domain Name Server (DNS), perpetrators can substitute the address of their website in place of the real one to spoof website visitors. Integrity threats can alter vital financial, medical or military information. It can have very serious consequences for businesses and people (Voeller, 2014).

Availability threats: The purpose of availability threats, also known as delay or denial threats is to disrupt normal computer processing or to deny processing entirely. For example if the processing speed of a single ATM machine transaction slows from one or 2-30 sec, users will abandon ATM machines entirely. Similarly, slowing any internet service will drive customers to competitor's web or commerce sites (Loukas, 2015).

Server threats: The server is the third link in the client-internet-server trio embodying the e-Commerce path between the user and a commerce server. Servers have vulnerabilities that can be exploited by anyone determined to cause destruction or to illegally acquire information.

Web-server threats: Web-server software is designed to deliver web pages by responding to HTTP requests. While web-server software is not inherently high-risk, it has been designed with web service and convenience as the main design goal. The more complex the software is, the higher the probability that it contains coding errors (bugs) and security holes, security weaknesses that provide openings through which evildoers can enter (Gillman *et al.*, 2015).

Commerce server threats: The commerce server, along with the web-server, responds to requests from web browsers through the HTTP protocol and CGI scripts. Several pieces of software comprise the commerce server software suite including an FTP server, a mail server, a remote login server and operating systems on host machines. Each of this software can have security holes and bugs (Ansari, 2015).

Database threats: e-Commerce systems store user data and retrieve product information from databases connected to the web-server. Besides product information, databases connected to the web contain valuable and private information that could irreparably damage a company if it were disclosed or altered. Some

databases store username/password pairs in a non-secure way. If someone obtains user authentication information, then he or she can masquerade as a legitimate database user and reveal private and costly information (Sharma and Misra, 2014).

Identification of trust: The phase of electronic payment (e-Payment) is confidential when all phases of the process are capable to satisfy the needs of participants and their security expectations. A fundamental prerequisite must be that all participants ought to have absolute trust in the system that they participate. The contraction of trust in an electronic payment system must take into consideration: data, identities and role behaviour. The adoption of e-Commerce must consider trust and risk as important determinants of adoption behaviour. Trust has been defined as “The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor irrespective of the ability to monitor or control that other party” (Mayer *et al.*, 1995). Trust requires a party to make a rational decision based on knowledge of possible rewards for trusting and not trusting. Trust enables higher gains while distrust avoids potential loss. “Generally an entity can be said to “Trust” a second entity when it (the first entity) makes the assumption that the second entity will behave exactly as the first entity expects” (Tsiakis and Sthephanides, 2005). The purpose of modelling trust is to establish a secure way to describe the decision of commerce process. A trusted environment is characterized by:

- The fact that all entities are uniquely identifiable
- That there is a minimum number of a priori trusted entities and
- That these entities have unquestionable trust to other participating entities

To design for trust, it is necessary to determine if and under what conditions trust mechanisms are brittle. Security architecture presumes that a trust model defines the trusted relationships between all involved components. Trust services are operated by sovereign organizations that are designed to protect consumers. Merchants concede to the organization’s trust standards (these standards cover areas such as privacy of personal information, return policies and security policies, etc.) in order to bind to legal obligations. Trust and trustworthiness are fundamental for every security solution. The needs for these trust aspects and the means that are used to implement it, affect the security mechanism of any commercial system. But we must

distinct trust form trustworthiness. Trust is an act of a trustor in which an entity places trust in some object (trust emanates from the entity). In contrast, trustworthiness is a characteristic of someone or something that is the object of trust. Trust is the enabling of confidence that something will or will not occur in a predictable or promised manner. The enabling of confidence is supported by identification, authentication, accountability, authorization and availability (Andert *et al.*, 2002).

Electronic payment (e-Payment) phase: Electronic payments have been reported to be the ultimate test of security and trust in e-Business environment. The notion of payment is an inborn part in any commercial transaction. The electronic payment (e-Payment) systems do two things in particular emulate existing payment frameworks from the real world or schematize new ways to execute payment transactions. Adoption of payment mechanisms and electronic money as other forms of payment depends upon trust in the security and reliability of the system and control of the particular transaction. The electronic transaction process takes place via the internet between three participants:

- Client e every user of the internet (client) can be considered as a potential customer. It is therefore imperative to establish mechanisms, to certificate trust and security
- Merchant e the typical merchant is the entity that needs to sell his goods (products or services) to the clients. In order to achieve this it has to secure transaction processes, so that, all participants are willing to act in a transactio
- Bank e the action of bank is familiar of every financial organization to validate and authorize transactions

In a commercial context, a payment process involves a payer, a merchant and a bank. In general, the entities transacting in a payment system are appointed by the specific commercial relationship which by itself may depend on series of conditions.

Literature review: Alan and William (Smith and Rupp, 2003) have stated that one of the most important issues in e-Lending is security. In the study, there was a statement from Richard Biell of Tower Group: “It’s one thing to submit a credit card number online to buy a product. It’s quite another thing to put your entire personal dossier online and hope that no one intercepts it, particularly if you’re not familiar with the lender”. It stated that a borrower will not proceed to exchanging personal

information without a sufficient level of confidence and the impact is the customers will not peruse other products and services without being familiar with the vendor and the process. In order to improve the trust among the customers, the suggestions given are keeping the customers informed through web presences and shared database where the study was not included.

Mukherjee and Nath (2007) stated that trust and commitment are the central tenets in building successful long-term relationships in the online retailing context. This study aims to re-examine the Commitment-Trust Theory (CTT) of relationship marketing in the online retailing context. The electronic hypermedia environment posed new challenges for relationship retailing where it is in the interest of retailers to establish and maintain long-term bonds with customers. This new marketing medium and channel is now an integral part of the multi-channel strategy for most retailers. The limitation the study has not discovered is fraud may increase, since, customers might not received the goods they had ordered and sometimes the goods are not reached to the customers.

Eben (2003) presented and reviewed the impact of information security for e-Business with emphasize on the security threats and potential losses that could arise from those vulnerabilities. e-Business security is analyzed as consisting of 6 dimensions: confidentiality, integrity, availability, legitimate use, auditing and non-repudiation. This study has proposed that designing a comprehensive and systematic security policy is a need for implementing e-Business security. The main thesis of this study was that e-Business security can only be effective if it is regarded as part of an overall corporate information security risk management policy. The resarchar proposed a six-stage security management strategy in this study. The limitations of this study are that implementing the framework were lack of building blocks in place and it was very costly for a company to developing software regarding to the increasing of the security.

Winch and Joyce (2006) mentioned that the dynamic nature in building and losing trust in Business to Consumer (B2C) e-Business in an effort to better understand the casual nature of trust. The purpose of this study is not to present yet another model but to suggest how to move from the information and knowledge those models provide into a better understanding of the problem of trust in B2C. Past models are largely descriptive and static in nature. This research helps to give a new understanding of trust building and maintenance as a dynamic process within what is in significant part, a closed-loop system. The study has therefore, taken the stock-flow diagramming approach from business dynamics modeling to reflect the structure of the trust

building systems. This emphasized that the management of system levels such as trust has to be through the control of the in and outflows if a company needs to build trust it has to work through the flows resulting from consumer's beliefs about how and whether problems might arise. The limitations of this study are that it did not mention that how can they reduce actual risks in the company and its process and in e-Transactions itself and how can they ensure perceived risks close to reality.

Hagi (2005) discovered that e-Business applications provide critical linkage between customers, suppliers and partners. Enterprises today have realized that their success will necessitate considerable attention to the security and privacy of their application software and particularly their e-Business applications. Vulnerabilities and threats related to e-Business application programs can be seen as occurring at all the different levels of the application system. All of these vulnerabilities and threats result in loss of confidentiality, integrity and authenticity. This study highlighted the methods and recommendations that can be utilized by organizations working through the complex maze of application security. This study has discovered that it is necessary to create security culture where employees, contractors and partners are educated on security policies, specific processes why security is important and what behavior is expected from them through training. The limitation of this study is that lack of explanation and discussion about how costly the training is since, it requires a number of experts and other resources in conducting training. Only the large corporations with sufficient resources are able to do that. Furthermore, it is hardly to ensure that their business partners will send their employees to go for training, since, it involved a large amount of cost.

Davidson (2001) proposed that having business on the internet offers potentially opportunities for increasing efficiency and reducing costs but it also offers unlimited risks. The much greater access to data will attract hackers and criminals. The resarchar identified the importance of business security will bring benefits to e-Business. The examples of security systems are virtual private database which provides a set of tools to enforce fine-gained access control within the database and oracle label security which is useful for hosting environments in which access to information can be formalized. The limitation of this article is that resarchar has not explained details about the impact of security in e-Business.

So and Sculli (2002) stated that in relation to trust and internet technologies, consumers are concern about the privacy and security. Trust must exist at all levels for the maintenance of cooperation, fundamental for any exchange and most routine of everyday interaction.

Nowadays, in uncertain and uncontrollable future, trust has become an important factor in needed in buyer-seller relationship in order to facilitate the transactions. Besides that trust is a critical factor in business is not enough, company must also build customer's trust because customers will attempt to evaluate supplier's trustworthiness before committing to business transactions. It's included appearance, performance and reputation. Combination all of these factors, e-Business will able to gain the higher return and customers will have satisfaction during a business transaction. The limitation of this study is that it did not state about the problems may occur when applying trust in business transactions.

Lord *et al.* (2002) discussed that e-Business is essential to reduce cost resulting from less overhead, greater economies of scales and increased of efficiency. The resarchars examined that the internet provides greater access to data, not only to legitimate users but also to hackers, disgruntled employees, criminals and corporate spies. But making business information accessible via. the internet increases the number of users who may be able to access the information. The internet creates challenges in terms of scalability of security mechanisms, management of those mechanisms and the need to make them standard and interoperable. It is essential to have security for e-Business but one of the limitations the resarchars did not focused is it needs highly cost to develop and maintain it. Besides, it is costly to hire an expert to develop and manage it especially if the company is only a small or medium enterprise.

Srinivasan (2004) stated that success of an e-Business rests on many factors. He defined trust is something that an e-Business must strive to achieve over a period of time. The resarchar suggested some contributing factors for gaining customer trust are: appeal of the Web site, product or service offerings, branding, quality of service and trusted seals. Trust can be viewed from many angles such as transaction, information content, product, technology and institution. This study analyses the role of trust from the transaction perspective and highlights the things that an e-Business could do for building customer trust. Factors contributing to trust are not easy to measure. It is developed over time. People trust a business based on their own past experience as well as by third party recommendations. He concluded that the e-Businesses are accessible from anywhere at any time, there are additional impediments in building and maintaining trust. The limitation of this study is that it did not focus much on how a company can enhance the security when doing online commerce.

Velmurugan (2009) analyzed the attitude of both e-Business and perspective of e-Partners toward the risks which were born by using the Internet as the communication channel. The risks were highly correlated and shape the trust of an e-Customer towards an e-Business. Starting from the organization's attitude towards risks, a number of criteria that influence the customer's trust are discussed. He has mentioned in the criteria of the characteristics which influence and shape the trust of an organization that: "the older the site, the higher the trust". But the question is if the site has offered products or services where the customers have high desired to own even though it is still new, the trust and the security of the website might be the least criteria for the customers to gain advantage, especially if the product offered is cheap. The limitation of the study is that, it did not mention that the sites those are able to offer a very useful and advanced product can also easily gain the trust from the customers. Sometimes, people will think that the site which is able to offer good product, that particular site should not have much of critical security problems, since, it needs powerful security tool or software to prevent other companies from imitate their products.

RESULTS AND DISCUSSION

Lack of trust: It is known that e-Commerce has becoming one of the most important elements in running business successfully particularly for small business. Managers have realized on how e-Commerce can benefits their business and thus, many organizations have started their business online. Internet is becoming more and more powerful tools for a firm to run their business as internet helps to connect the stakeholders from all over the world which has improved the efficiency of an organization. The customers who are buying their products online will face another risky problem on dealing with the vendors where the vendors cannot be observed and sometimes they are unknown. Besides, customers not only have to receive the products which are unacceptable in quality but they have to face the risks of not receiving the items they had ordered at all after they made their payment through online. Moreover, personal information and credit card numbers of customers might disclosed to other people during or after their online trades when the vendors are incapable to protect the customer's data. Due to the reason stated above, it in impossible for customers to do their online trades safely and because of those cases happened before, it caused the customers to lost their confident in buying through online and it does make customers lost their trust towards online trading.

Solution: To solve this major problem of trusting in e-Business, a large amount of trust will be very important to prove to the users that the certain website which the users are accessing are trust worthy. The trust makes a website in a good condition and all the users who are using it will feel comfortable dealing any type of business with them. As the issue of trust occurs in everything in the internet and e-Business, it is very important that the value of trust is big and it is able to gain the trust and the loyalty of the users and making sure that they stay loyal to the particular website. This trust can be achieved by taking few steps which can help gain this trust. The steps are making sure that all private data about the user are kept safe, well maintained and kept up-dated. By doing this, we can prevent the lack of trust a user's puts towards the particular website. When we have enough trust amongst the users, an e-Biz website can perform well and at full speed and makes sure that all the information are kept safe and making sure that no hackers go through the website and steal users information.

Unaware of how e-Business: Transactions take place e-Business has made an organization to run their business more efficient where the customers can get their products just in time. Through e-Business like by using extranet, the suppliers are able to access to an organization's database on when the wan to receive what they have ordered and that enable them to make the products available to the particular company just in time. Besides with the cooperation from supplier's a company can try to have their products sent to client's house on the time. But there are still some of the customers who are unaware on how e-Business transaction takes place. Some of the customers have the perception of buying things in shops is much more faster than buying through online as they can get the products immediately where they thought that buying through online might have to wait for another days in order to get the products arrived to their house. They do not realize that a close integration between the suppliers and the company can put the clients in the way that they still can receive their order on the time they want.

Solution: The issue of being unaware of not knowing about how online transaction works can lead to a lot of issues as stated above. This problem is a threat to users who don't seem to understand that their money can be blacked out if hackers hacks their account. To prevent this problem, the e-Biz has to make sure that they have enhanced versions of security and good transaction system for the users to cash in their money. By making this transaction system secure, we can no longer be afraid

of hackers. In the meantime, e-Biz needs to teach the users about the online transaction system. By teaching them how the system functions, users can learn and they too can be aware of the system's processes. All this learning will alert the users to be more caution on their online transaction. e-Biz websites with online transaction systems should list down all the possibilities of doing online transaction and all the misuses possible.

Degree of confidentiality involves: There is uncountable of ways that an e-Business set up will be attacking by hackers, crackers and disgruntled insiders. These have immediately decreased the degree of confidentiality of clients towards e-Business. According to Eben (2003) confidentiality is defined as making information accessible to authorized users and prevents the access from unauthorized users towards information. The problem of degree of confidentiality is always taken place in health centre. The recent surveys reported by Georgetown University Institute for Health Care Research and Policy contain statistics regarding to the people's concern for the confidentiality, there were about 63% of internet "health seekers" and 60% of all internet users oppose the ideas of keeping medical records online, even with a secure, password protected site since they worried that other people will see the detail of their records. Majority of internet users are worried about others finding out about their online activities where 89% fear that internet companies might sell or give away and 85% fear that insurance companies might change their coverage after finding out what online information they accessed.

Solution: Organizations have to keep secured of client's personal information and stored in a way that it is unable to be accessed by other unauthorized users. In order to build the confidence of users, the system has to be secured and tight, so that, no bugs or viruses could enter the website or the transaction system This prevention of this system can enlarge the safety and can ensure the safety of the users using the system and this confidentiality is essential because of the fear towards forgery and hacking. Confidential are privacy of data and safety of an individual's property and assets. In e-Biz, hackers always are on the wait for any loophole to enter the system and hack information about the user's confidential folders. In these folders, there may be a lot of important information which can be useful and harmful to others. By using this powerful system, hackers will find it difficult to hack the system and due to this users will have more confident on the e-Biz website and they will certainly cash in their money and deal their businesses thru the internet.

CONCLUSION

The development and improvement of technologies have brought successful towards e-Business. High technologies have attracted people misuse the technologies such as hackers and cybercrime which they can access to e-Business privacy easily. Thus, e-Business companies should build trust and using security during the business transaction. To provide value to the customers through service and goods provided, research found that companies should build up trust and security to protect their customers. Benefits of application trust and security include improved customer service, build customers trust, avoid the misuse of technologies, protect customer's privacy and maintain the company's reputation. In order to create an effective infrastructure for securing e-Business, it requires a comprehensive development of several elements including laws, policies, industry self-regulation, technical standards and law enforcement. These elements may provide positive environment and infrastructure to support the growth of e-Business and relation with customers. Therefore, governments and businesses need to work together to improve consumer trust and security are attempt to increase transactional efficiency and effectiveness in all aspects of the design, production, marketing and sales of products or services for existing and developing good relation through the utilization of current and emerging electronic technologies which will gain the more confidence in e-Business. Additionally, the government itself needs to re-examine existing regulations to ensure protection for the e-Business.

REFERENCES

- Anderson, J.Q. and H. Rainie, 2014. The internet of things will thrive by 2025. Pew Research Center, Washington, USA.
- Andert, D., R. Wakefield and J. Weise, 2002. Trust modeling for security architecture development. Master Thesis, Stanford University, Stanford, California.
- Ansari, J.A., 2015. Web Penetration Testing with Kali Linux. 2nd Edn., Packt Publishing, Birmingham, England, ISBN:9781783988525, Pages: 312.
- Ashayeri, J. and G. Tuzkaya, 2011. Design of demand driven return supply chain for high-tech products. *J. Ind. Eng. Manage.*, 4: 481-503.
- Atzori, L., A. Iera and G. Morabito, 2010. The internet of things: A survey. *Comput. Networks*, 54: 2787-2805.
- Bossomaier, T. and B.A. Hope, 2015. Online GIS and Spatial Metadata. 2nd Edn., CRC Press, Boca Raton, Florida, USA., ISBN:9781482220162, Pages: 414.
- Conner, M., 2010. Sensors empower the internet of things. *Electr. Des. News*, 55: 32-37.
- Costa, E.D., 2016. Global E-Commerce Strategies for Small Businesses. MIT Press, Cambridge, Massachusetts, USA., Pages: 198.
- Daugherty, P.J., R.G. Richey, S.E. Genchev and H. Chen, 2005. Reverse logistics: Superior performance through focused resource commitments to information technology. *Transp. Res. Part E. Logist. Rev.*, 41: 77-92.
- Davidson, M.A., 2001. Database security for E-business. Oracle Corporation, Redwood City, California, USA.
- Eben, O., 2003. A systematic approach to E-business security. Master Thesis, University of New Brunswick, Fredericton, Canada.
- Emmerson, B., 2010. M2M: The internet of 50 billion devices. *WinWin Mag.*, 1: 19-22.
- Geistfeld, M.A., 2016. Protecting confidential information entrusted to others in business transactions: Data breaches, identity theft and tort liability. *DePaul L. Rev.*, 66: 385-721.
- Gillman, D., Y. Lin, B. Maggs and R.K. Sitaraman, 2015. Protecting websites from attack with secure delivery networks. *Comput.*, 48: 26-34.
- Grieco, L.A., M.B. Alaya, T. Monteil and K. Drira, 2014. Architecting information centric ETSI-M2M systems. Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), March 24-28, 2014, IEEE, Budapest, Hungary, ISBN: 978-1-4799-2737-1, pp: 211-214.
- Hagi, S., 2005. Engineering E-business applications for security. IBM Canada Ltd, Ottawa, Canada,.
- Hersent, O., D. Boswarthick and O. Elloumi, 2011. The Internet of Things: Key Applications and Protocols. John Wiley & Sons, Hoboken, New Jersey.
- Kadhim, A.M. and M.Z. Al-Taie, 2013. Factors disrupting a successful implementation of E-commerce in Iraq. *J. Econ.*, 1: 1-14.
- Lord, P., A. Mary and B. Kristy, 2002. Managing E-business security challenges. Oracle Corporation, Redwood City, California.
- Loukas, G., 2015. Cyber-physical Attacks: A Growing Invisible Threat. Elsevier, Amsterdam, Netherlands, ISBN:9780128014639, Pages: 270.
- Mayer, R.C., J.H. Davis and F.D. Schoorman, 1995. An integrative model of organizational trust. *Acad. Manage. Rev.*, 20: 709-734.
- Miorandi, D., S. Sicari, F. De Pellegrini and I. Chlamtac, 2012. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 10: 1497-1516.

- Mukherjee, A. and P. Nath, 2007. Role of electronic trust in online retailing: A re-examination of the commitment-trust theory. *Eur. J. Market.*, 41: 1173-1202.
- Nieles, M., K.L. Dempsey and V.Y. Pillitteri, 2017. An introduction to information security. MIT Thesis, National Institute of Science and Technology, Berhampur, India.
- Palattella, M.R., N. Accettura, X. Vilajosana, T. Watteyne and L.A. Grieco *et al.*, 2013. Standardized protocol stack for the internet of (important) things. *IEEE. Commun. Surv. Tutorials*, 15: 1389-1406.
- Roman, R., J. Zhou and J. Lopez, 2013. On the features and challenges of security and privacy in distributed internet of things. *Comput. Netw.*, 57: 2266-2279.
- Sharma, A. and P. Misra, 2014. A security framework for E-business applications. *Intl. J. Comput. Appl.*, 102: 30-34.
- Shelanski, H.A., 2013. Information, innovation and competition policy for the internet. *Univ. Pennsylvania Law Rev.*, 161: 1663-1705.
- Smith, A.D. and W.T. Rupp, 2003. E-lending: Foundations of financial and consumer marketing in an information intensive society. *J. E. Bus. Inf. Technol.*, 3: 5-19.
- So, M.W. and D. Sculli, 2002. The role of trust, quality, value and risk in conducting e-Business. *Ind. Manage. Data Syst.*, 102: 503-512.
- Srinivasan, S., 2004. Role of trust in E-business success. *Inf. Manage. Comput. Secur.*, 12: 66-72.
- Stankovic, J.A., 2014. Research directions for the internet of things. *IEEE. Internet Things J.*, 1: 3-9.
- Tan, L. and N. Wang, 2010. Future internet: The internet of things. *Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering*, August 20-22, 2010, Chengdu, China, pp: V5-376-V5-380.
- Tsiakis, T. and G. Sthephanides, 2005. The concept of security and trust in electronic payments. *Comput. Secur.*, 24: 10-15.
- Velmurugan, M.S., 2009. Security and trust in E-business: Problems and prospects. *Intl. J. Electron. Bus. Manage.*, 7: 151-158.
- Voeller, J.G., 2014. *Cyber Security*. John Wiley & Sons, Hoboken, New Jersey, USA.,
- Winch, G. and P. Joyce, 2006. Exploring the dynamics of building, and losing, consumer trust in B2C E-business. *Intl. J. Retail Distrib. Manage.*, 34: 541-555.

Validate Geospatial Indicators for Assessing Community Resilience Capacities to Floods; A System-Performance-Based Approach

¹Chethika Abenayake, ²Mikami Yoshiki and ³Matsuda Yoko

¹Department of Information Science and Control Engineering, Graduate School of Engineering,

²Department of Management and Information Systems Science,

³Disaster Prevention Systems Group, Department of Civil and Environmental Engineering,
Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, 940-2188 Niigata, Japan

Abstract: Community resilience assessment methodologies play a vital role in decision-making for building community's resilience to disasters. Established resilience indicators can gain decision-maker's confidence on assessment methods. Hence, this study attempts to assess the validity and adequacy of a set of geospatial indicators extracted from existing methodologies. First, the study extracts 29 geospatial indicators to assess community resilience to floods and applies them for 23 divisions in Colombo, Sri Lanka. Second, this study plots the affected population data by division into system performance curves concerning the flood occurred in May, 2016. Third, the study develops system performance curve-based measures to quantify the three capacities of community resilience: transformative capacity, absorptive capacity and recovery capacity. Then the study statistically tests the association of geospatial indicators with each of the system-performance measures. The study has obtained spatial data for mapping geospatial indicators from national databases and the affected community data from the records of disaster management officers. Findings revealed 16 indicators having significant association with system-performance measures and the results discuss the ambiguities and cohesive nature of indicators regarding different capacities. This study has established a pilot initiative to validate a set of geospatial indicators where many of the practicing community assessment tools have little or no geospatial indicators. Furthermore, this study applies a set of system performance curve-based measures to externally validate the community resilience capacities. Therefore, the study contributes with comprehensions to make community resilience assessments tools more powerful in guiding communities towards resilience.

Key words: Disasters, resilience capacities, methods, socio-ecological systems, guiding communities, assessments tools

INTRODUCTION

Every year, thousands of people around the world struggle to confront floods. A flood is a hydro-meteorological hazard that has accounted for 47% of all weather-related disasters (1995-2015) affecting 2.3 billion people in the world during the past decade (UNISDR., 2016). Out of all hazards, floods pose the widely distributed natural risk to life today. Floods often inundate clusters of human settlements, making it a community crisis that calls for attention at local and regional geographies. As a global response, "Making cities and human settlements inclusive, safe, resilient and sustainable" has become a goal of the Sendai Framework for disaster risk reduction 2015-2030 and the adopted New

Urban Agenda 2030 (USAID, 2016). This global commitment emphasizes to mainstream urban development and disaster risk reduction programs towards building resilient communities. Directing these initiatives to cater the needs of the most vulnerable and the least resilient communities is a sustainable development challenge.

Resilience assessment is a policy and planning tool that has been developed to facilitate decision-making on resilience-building initiatives (Abenayake *et al.*, 2016). "Resilience measures (i.e., assessments) are helpful in identifying disaster risk, taking productive steps toward its reduction and getting stakeholders together to build capacity to prepare for respond to, recover from and more successfully adapt to threats" (Cutter, 2016). Despite

Corresponding Author: Chethika Abenayake, Department of Information Science and Control Engineering,
Graduate School of Engineering, Nagaoka University of Technology, 1603-1 Kamitomioka,
Nagaoka, 940-2188 Niigata, Japan

several utilities, “community resilience assessment tools have not yet been appropriately integrated into planning and policy-making process” (Sharifi and Yamagata, 2016). One of the key reasons is a lack of consensus and confidence in the assessment tools among practitioners. “Although, community resilience to disasters is still an emerging field and index developers often describe their products as frameworks or baseline assessments, there is little utility unless they can be confidently used to inform decision-makers” (Bakkensen *et al.*, 2017). In such context, this study has focused on two properties of resilience indicators, validity and adequacy, that are indispensable prerequisites in building decision-maker’s confidence in community resilience assessment methodologies.

Validation “assesses the explanatory power of an index using real world observations and can estimate the ability of an index to explain a variety of disaster losses, there by giving confidence in index’s ability and performance to end users” (Bakkensen *et al.*, 2017). Furthermore, Validation performs a vital role in identifying the relative importance of indicators (Burton, 2015; Cai *et al.*, 2016) and clarifying which indicator/s should be used in each decision (Bakkensen *et al.*, 2017). The limited studies on validation have revealed that “some variables were more strongly associated with actual recovery than others and thus were better proxies of resilience” (Parsons *et al.*, 2016). “The use of logical plausibility is presently most common in disaster resilience assessment because causal validation specifying the association between an indicator and disaster resilience or vulnerability is only recently attracting research focus (Parsons *et al.*, 2016). Hence, even though validation is a major step in the process of creating composite indices, it is rarely performed in the context of disaster resilience studies (Bakkensen *et al.*, 2017; Burton, 2015; Cai *et al.*, 2016). External validation of community resilience indicators has posed a challenge primarily because community resilience is not a directly observable phenomenon and the validation of resilience indicators requires the use of proxies (Tate, 2012; Cai *et al.*, 2016). “Currently, there is no commonly recognized independent proxy data used in the validation of resilience assessment” (Cai *et al.*, 2016). Few cross-disciplinary studies have developed proxies based on the system performance curve concerning the evidence of previous hazard events. This study focuses on system performance-based measurements and has attempted to utilize them for validating community resilience indicators.

Endorsing the adequacy of resilience assessment tools to represent community’s capability plays a vital

role in the credibility of resilience assessment indicators. Sharifi and Yamgatha (2016)’s recent study emphasizes that “there is a paucity of studies that evaluate community resilience assessment tools in terms of their suitability of guiding communities towards resilience” Ealcin *et al.* (2011) have evaluated the ability of community resilience assessment frameworks in the United States to account for four resilience actions: plan, absorb, recover and adapt (Larkin *et al.*, 2015). Sharifi and Yamagata have also assessed these four actions concerning 36 community resilience assessment tools that practice in several different countries (Sharifi and Yamagata, 2016). Findings of those studies revealed that many of the assessment methodologies have mediocre performance in adequately covering all 4 types of actions that drive community’s resilience. The studies mentioned above have initiated a policy dialog on some potential opportunities for improving community resilience assessment methodologies. This study attempts to constructively contribute to this line of studies primarily with two advancements. First, current recommendations are based on theoretically-validated heuristic evaluations. The study performs a statistically-supported, external validation based on empirical evidence. Second, the studies mentioned above have focused on community assessment methodologies whereas this study is focused on a set of geospatial indicators that has been filtered from several existing methods.

The overall objective of this study is to test the validity of geospatial indicators in assessing community resilience to floods and to investigate the adequacy of geospatial indicators for evaluating the three capacities of resilience. The three capacities are persistence capacity, absorptive capacity and recovery capacity. These capacities correspond to the plan, absorb and recover actions discussed in previous studies. Nevertheless, due to the deficiency of empirical data, this study does not discuss community’s ability to perform long-term adaptation from the 4 actions that have been examined in previous studies. The scope of this study is limited to floods and particularly focused on geospatial indicators that are input variables of existing composite resilience indices. This study provides a set of validated geospatial indicators along with their capability to consider distinct types of capacities that are required to make communities resilient.

MATERIALS AND METHODS

Colombo, Sri Lanka as the case study: Sri Lanka, being a tropical Island country is highly susceptible to the adverse consequences of hydro-meteorological

disasters (Ranjan and Abenayake, 2014). Flood is the most frequent natural hazard in Sri Lanka. The low-pressure system occurred in the Indian Ocean on May 2016 caused torrential rainfall across Sri Lanka. Kelani basin which is one of the main river basins in Sri Lanka, received 350 mm of total rainfall within three consecutive days from the 15-17th of May 2016. Flood was 6-12 feet in height and the damage was recorded as the highest number of flood-affected population over last 6 decades (DMC., 2016). Per the situation report issued by Disaster Management Center of Sri Lanka, over 200,000 people who reside in Colombo were affected by this flood (DMC., 2016). Property and livelihood losses were also significant because Colombo is the national capital that hubs commercial and economic infrastructure. This study was conducted in the lower drainage basin of the Kelani River including 23 DS divisions that belong to the Colombo Metropolitan Region, Sri Lanka. The Divisional Secretariat (DS) division is a local-government level, administrative unit in Sri Lanka and there are 329 DS divisions in the country. Figure 1 shows the selected study area including 23 DS divisions.

People residing in 20 DS divisions were evacuated to 140 nearby welfare canters during the flood. The remaining three DS divisions (Moratuwa, Dehiwala and Ja-Ela) had no people evacuated primarily because flood height has been lower due to elevation and exposure was limited to a small percent of the area.

Selection of geospatial indicators to assess community resilience to floods: Geospatial indicators represent terrestrial activities and processes derived from geospatial analysis that are widely applicable in decision-making science. The study has extracted geospatial indicators from practicing methodologies and proposed frameworks. The initial attempt was to investigate from electronic databases including Google Scholar, MEDLINE through PubMed and Scopus with no limitation on article type and date. The search strategy was to initially perform machine extraction by keywords and then to manually screen the extracted articles. Screening criteria were having processed by geospatial analysis, applicability at the



Fig. 1: Map of study area

regional scale, relevance to floods and availability of data. The first search term ‘geospatial indicator and resilience’ applied for title, abstract and keywords yet could not extract a valid result. The next search terms attempted were ‘spatial indicator and resilience’, ‘place indicator and resilience’ and ‘location indicator and resilience’. Many of the extracted indicators were found to have some possibility to geo-visualize if computed with spatial data. However, the manual screening was particularly aimed at the indicators that can be derived from the geospatial analysis. The limited application of geospatial analysis in assessing community resilience to floods shrank the extracted results to 52 indicators. Except for the works of Cutter *et al.* and Kotze and Reyers, many of the manually filtered indicators have basic algebraic processing of spatial data including ratio and density functions (Cutter *et al.*, 2008a, b; Kotzee and Reyers, 2016). Many of the extracted indicators are listed in the Paolo Cimellaro’s comprehensive literature review

on extant indicators to assess community resilience to disasters (Cimellaro, 2016). The list of 52 indicators was further filtered into 34 indicators by focusing on the relevance to floods and then into 25 considering the data availability. This study has made minor modifications to four of the extracted geospatial indicators while considering geospatial properties. Justifications of those minor modifications were presumed logically and yet to be tested. Therefore, the validation test considered both extracted indicators and modified versions. Table 1 presents the set of 29 geospatial indicators to be tested as independent variables to assess community resilience to floods.

Table 2 contains the information about data acquisition for computing geospatial indicators for the case study area.

The set of 29 indicators (Table 1) were geospatially processed for the case study area in Colombo using a GIS (Geographic Information System) Software.

Table 1: The selected geospatial indicators for theoretical validation

ID	Indicators	Direction	Justification	Data (code) *
1	Percent land area that is a wetland, swamp, marsh and mangrove	+	Barnes <i>et al.</i> and Klein <i>et al.</i> (2003)	A
2	Rapid urban population growth (Percentage increase of urban population density)	-	(Centre for Science Economics and the Environment in 2002)	G, B
3	Percent deep permeable soil per ward	+	Kotzee and Reyers (2016)	E, J
4	Percent fire, police, emergency relief services and temporary shelters outside of hazard zones	+	(US Indian Ocean Tsunami Warning System Program 2007)	A, F
5	Percent of building infrastructure not in flood inundation zones	+	Geis and Kutzmark (1995)	A, F
6	Percent of government offices outside of flood inundation zones	+	Cimellaro (2016)	A, F
7	Percent of commercial establishments outside of high hazard zones (flood, surge)	+	(US Indian Ocean Tsunami Warning System Program in 2007)	A, F
8	Population living in high-intensity urban areas/ population density	-	Cutter <i>et al.</i> (2008a)	A, G
9	Percent land area that does not contain erodible slopes	+	Cutter <i>et al.</i> (2008a)	C, E
10	Percent land area not in an inundation zone (100 years)	+	Cutter <i>et al.</i> (2008a)	E
11	Percent land area that does not contain impervious surfaces	+	Cutter <i>et al.</i> (2008b)	A, I
12	Percent land area with no forest and rangeland decline	+	Geis and Kutzmark, Cimellaro (2016)	A
13	Percent land area with no wetland decline	+	Cutter <i>et al.</i> (2008a, b)	A
14	Percent area that has changed into urban areas (by urban classification)	-	Cutter <i>et al.</i> , (2008a)	A, B, H
15	Percent land area that is high-intensity urban development (80% or more impervious surface)	-	Cutter <i>et al.</i> (2008a)	A, I
16	Percent land area of developed open spaces	+	Geis and Kutzmark, Cimellaro (2016)	A
17	Principal arterial miles	+	Cutter <i>et al.</i> (2008a)	A
18	Hospitals per square mile	+	Bruneau and Reinhorn (2007)	A
19	Schools (primary and secondary education) per square mile	+	Cutter <i>et al.</i> , (2008a)	A
20	Hotels and motels per square mile	+	Cutter <i>et al.</i> (2008) (US Indian Ocean Tsunami Warning System Program 2007; Centre for Science Economics and the Environment 2002)	A
21	Density of commercial infrastructure	+	Cutter <i>et al.</i> (2010)	A
22	Number of river miles	-	Allenby	A
23	Percent erodible soil per ward	-	Berke	A
24	Land use diversity (Proportion of land use categories per ward, multiplied by the natural logarithm. The resulting product is summed across wards and multiplied by -1)	-	Cutter <i>et al.</i> (2008a) and Kotzee and Reyers (2016)	J, E
		+	Kotzee and Reyers (2016)	A

Table 1: Continue

ID	Indicators	Direction	Justification	Data (code)*
25	Wetland diversity (Proportion of flood attenuating wetlands per ward, multiplied by the natural logarithm. The resulting product is summed across wards and multiplied by -1)	+	Kotzee and Reyers (2016)	A
26	Rapid urban growth (Percentage land cover change to urban areas from base year) [#]	-	Minor modifications made to indicator 2	A, B
27	Waterbodies density (Waterbody area/total land area) [#]	-	Minor modifications made to existing indicator 22	A
28	Access to hospital (Inverse of Euclidean distance to the hospitals) [#]	+	Spatial adjustments made to existing indicator 18	A
29	Movement potential (Inverse of Euclidian distance to road network) [#]	+	Spatial adjustments made to existing indicator 17	A

*Refer Table 2 for details; [#]Indicators modified/introduced by this research (26, 27, 28, 29)

Table 2: Data acquisition for computing geospatial indicators

Data type/Code*	Description	Years	Spatial scale	Source [#]
Map data				
A	Land use map	2014	1: 5000	Urban Transport System Development Project, Japan International Cooperation Agency, Japan
B	Topographic map	1984	1: 50,000	Survey Department, Sri Lanka
C	contour map	2012	1: 5,000	Tsunami Hazard Map Database, Coast Conservation and Resource Management
Department				
D	Rainfall isohyets	2007	1: 10,000	National Atlas, Survey Department of Sri Lanka
E	Soil map	2007	1: 10,000	National Atlas, Survey Department of Sri Lanka
F	Flood inundation map	2016	1: 30,000	Disaster Management Center, Sri Lanka
Tabular data				
G	Population	2012	GN Division	Population Census, 2012, Department of Census and Statistics, Sri Lanka
H	Land use classification	2013	National	Colombo Development Plan, 2013, Urban Development Authority, Sri Lanka
Classifications				
I	Floor krea ratio by land use	2013	Regional	Colombo Development Plan, 2013, Urban Development Authority, Sri Lanka
J	Soil hydraulic properties	1961	National	the National Soil Survey Published in Soil of Ceylon, 1961 by Soil Type Moormann and Panabokke, 1961

*ID to link Table 1; [#]The study has used published data

Formulation of system-performance measures to evaluate community resilience: Validation of indicators requires an independent set of outcome variables to surrogate community resilience. As mentioned in the introduction, community resilience is not a directly observable phenomenon. In order to overcome this inherent limitation practically, many studies have proposed to observe community resilience through the empirical evidence of population, housing and infrastructure system responses to hazards. Theoretically, resilience is measured concerning a desirable regime of function. Empirically, it is challenging to define when people cross such hypothetical status. In this study, the desirable regime of function has been referred to as the status that community has not been fallen into a situation that they cannot fulfill the basic needs. Accordingly, the desirable regime of function was attributed to the community's ability to survive without obtaining external assistance for food, shelter and clothing. Hence, the status that the community fails to withstand the desirable regime of function was related to the situation of temporarily falling

into welfare centers because of the flood. Similarly, bouncing off to the desirable regime of function was attributed to the situation of leaving the welfare center.

The number of population that stayed overnight in welfare centers was considered as the outcome variable. This includes people who self-evacuated in-advanced and those who were rescued during the flood. Daily data on the number of people that stayed overnight in welfare centers during the flood that occurred in May 2016 were initially collected from the Disaster Management Center, Sri Lanka. However, the data was not available for all consecutive days. Hence, the missing data was obtained by interviewing the disaster management officers in 23 DS divisions. The data was plotted into a system performance curve where 'number of people that stayed overnight in welfare centers' indicates the performance of community resilience to the flood event over a period. Onset date of the flood was the 15th of May and the residential population of each DS division was given as the initial performance level of the system. The time when no people

remained in welfare centers were considered as the point which the system returned to the desirable regime of function.

The system performance curve is widely employed to explain the disaster resilient behavior of socio-technical systems despite the limited attempts to apply in socio-ecological systems. The early research of Bruneau *et al.* (2003) have utilized the system performance curve to quantify resilience based on the ‘resilience triangle’ (Wang and Blackmore, 2009; Bocchini *et al.*, 2013). Later studies have developed this by using mathematical functions of indefinite integrals that measure either the area above performance curve (Bruneau, 2006; Bruneau and Reinhorn, 2007; Bruneau *et al.*, 2005) or the area under the performance curve (Cimellaro *et al.*, 2010a, b; Bocchini and Frangopol, 2012). A comprehensive overview of these measures is available elsewhere (Bocchini and Frangopol, 2010). In addition, the latest work by Mugume *et al.* (2015) has applied a mathematical function of indefinite integrals to quantify the resilience of storm water drainage systems to floods. This application has further improved the indefinite integrals-based function, normalizing the resilience levels by actual system performance. This improvement facilitates the comparison of sub-systems regarding their resilience performances. These studies have successfully quantified the resilience as an index that combines all resilience capacities into one measure. However, the decision of blending all different properties, actions and capacities into one measure has not been favorable for some applications (Bocchini, *et al.*, 2013). Such combined measure has less utility to test the adequacy of indicators to cover different capacities of resilience. States of system performance curve represent distinct types of system behavior such as planning and preparing to persist the perturbations, buffering the system degradation by absorbing shocks and recovering the system following the learning and adaptation Fig. 2. Hence, this study has formulated four measurements to quantify three selected states of system performance curve that correspond to three resilience capacities concerned. Persistence rate (P) of *i*th event:

$$P_i = \frac{1}{(P_j A_j)(t_f - t_0)} \quad (1)$$

Peak Failure (F) of *i*th event:

$$P_i = \frac{1}{(P_j A_j)(Q_s)} - Q_{mf} \quad (2)$$

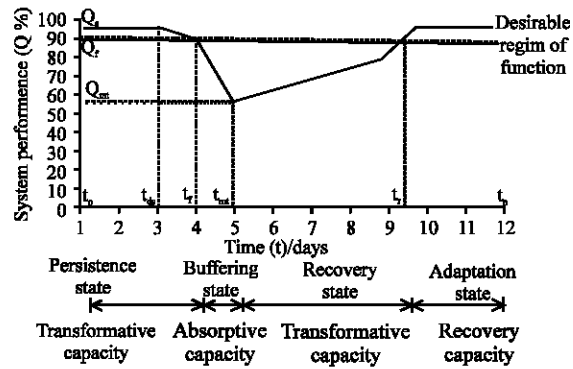


Fig. 2: A conceptual illustration of system performance curve

Degradation rate (B) of *i*th event:

$$B_i = \frac{1}{(P_j A_j)} \int_{t_r}^{t_{mf}} (100 - Q_i(t)) dt \quad (3)$$

Recovery rate (R) of *i*th event:

$$R_i = \frac{1}{(P_j A_j)} \int_{t_{mf}}^{t_r} (100 - Q_i(t)) dt \quad (4)$$

Where:

P = Total population of *j*th locality (DS division)

A = Percentage inundated area (A) of *j*th locality (DS Division)

$$A_j = \frac{FI_j}{L_j} \times 100 \quad (5)$$

where, FI represents flood inundated built-up are and L represents the total land extent of *j*th locality (DS Division).

In all four measures, the resilience has been normalized by population and the inundation area. The normalization facilitates the comparison by adjusting the differences of population size and percent area inundated among various localities to a notionally common scale. Hence, it indicates how a community system performs in a flood irrespective of the effect of the size of population and inundation area.

Persistence rate (Eq. 1) measures the duration that the community system with stands the disturbances while fulfilling the basic needs. A higher persistent rate indicates a higher level of community resilience. Persistence rate primarily expresses community’s preparedness as a result of long-term planning and adaptation. Hence, it partly captures the transformative

capacity of a given system. In this study, persistent rate has been measured based on the duration that residents can endure floods without leaving their houses.

The state when community cannot persist further and compelled to seek external assistance to fulfill their basic needs is 't_{ds}' where system starts degradation. When the degradation crosses the desirable regime of function 't_f', the system failure occurs. Peak failure (Eq. 2) and degradation rate (Eq. 3) measure the magnitude of the degradation. Peak failure and degradation rate increases if the system cannot buffer floods by absorbing the shock. Hence, these two measures are attributed to the absorption capacity of the system. In this study individuals seeking the assistance within the community members was attributed to the point 't_{ds}' and community seeking shelter at welfare centers was attributed to the point 't_f'. However, data collection was limited to the number of people seeking shelter at welfare centers. Therefore, 't_{ds}' was approximated to equal 't_f' and 'Q_s' (i.e., initial performance of the system) was approximated to equal 'Q_f' (i.e., system performance at the desirable regime of function).

Recovery rate (Eq. 4) measures the time taken to recover considering the correspondent system performance at each point of recovery. Decreased time taken to recover indicates a higher recovery capacity of the system. A higher recovery capacity indicates a higher community resilience. In this study, the point system based on the desirable regime of function 't_f' was attributed to the time when no people stayed overnight at the welfare center.

Accordingly, the persistent rate has a direct relationship with community resilience whereas other three measures have an inverse relationship with community resilience.

Framework of analysis: First, this study computed the resilience level of 23 DS divisions by 29 geospatial indicators separately. Second, the study plotted the system performance curves of each DS division with affected population data. Persistence rate, peak failure, degradation rate and recovery rate were computed for each of the DS divisions based on system performance curves. Next, the study tested the statistical association between geospatial indicators and system-performance measures. Association was tested by Spearman's correlation coefficient because the Shapiro-Wilk test and QQ plots of many of the indicators revealed a free-distribution with several outliers. A two-tailed test was conducted due to the difference in directions. In

interpreting the results, Spearman rank-order correlation coefficient (r_s) value equal or above 0.7 was considered a strong association and equal or above 0.5 was considered a moderate association. Coefficients (r_s) at confidence interval 0.01 were considered significant and 0.05 were considered moderately significant.

RESULTS AND DISCUSSION

Geospatial indicators that revealed an association with at least one of the system-performance measures were considered as valid for community resilience assessments. In overall, out of 29 geospatial indicators, 13 showed either significant or moderately significant correlation (Table 3).

'Rapid urban growth' recorded the highest correlation with degradation rate (r_s = 0.791, p<0.000), peak failure (r_s = 0.765, p<0.000) and recovery rate (r_s = 0.865, p = 0.000). 'Schools (primary and secondary education) per square mile recorded the highest correlation (r_s = 0.779, p<0.000) with persistence rate. Rapid urban growth concentrates physical development agglomerating buildings infrastructure and human activities. Inundation of such intensively urbanized locations can result in catastrophic failures due to many elements-at-risk within the system. Furthermore, rapid urban growth disrupts natural flood defense mechanisms of socio-ecological systems. For example, conversion of agricultural and other vegetative land uses into build-up areas reduces the infiltration, evaporation and increase the surface runoff, thereby weakening the absorptive capacity. Moreover, reclamation of water retention areas for urban development as in the case of Colombo, reduces the water retention and detention of ecosystems perturbing the recovery process. The second most associated indicator is 'schools per square mile'. The school is a community infrastructure which can be considered to represent the community's social well-being. Community systems that have access to education and social well-being are resourceful to anticipate floods, plan in advance and withstand disturbances. Per the above reasoning initial results indicate that geospatial indicators can meaningfully detect the environmental and physical influences over community resilience.

Ambiguity in the direction of association concerning the states of resilience: As Table 1 shows, existing literature has mentioned a possible direction when interpreting the influence of each spatial indicator on community resilience. Positive direction refers to a status where the

Table 3: Geospatial indicators revealed a significant association with outcome variable/sec

ID	Indicators	Outcome	Degradation rate	Peak failure	Recovery rate	Persistent rate
1	Percent land area that is a wetland, swamp, marsh and mangrove	r _s	0.617**	0.694**	0.669**	-
		Sig.	0.006	0.001	0.002	-
8	Population living in high intensity urban areas/ population density	r _s	0.569*	0.647**	0.583*	-
		Sig.	0.014	0.004	0.011	-
10	Percent land area not in an inundation zone (100 years)	r _s	0.461*	-	-	0.537**
		Sig.	0.031	-	-	0.008
14	Percent area that has changed into urban areas	r _s	-	-	-	0.742**
		Sig.	-	-	-	0.000
16	Percentage land area of developed open spaces	r _s	0.562*	0.520*	0.713**	0.570*
		Sig.	0.015	0.027	0.001	0.011
18	Hospitals per square mile	r _s	-	-	0.478*	0.678**
		Sig.	-	-	0.045	0.001
19	Schools (primary and secondary education) per square mile	r _s	-	-	-	0.779**
		Sig.	-	-	-	0.000
20	Hotels and motels per square mile	r _s	0.469*	-	0.525*	0.577**
		Sig.	0.050	-	0.025	0.010
21	Density of commercial infrastructure	r _s	0.474*	0.491*	0.490*	-
			0.047	0.039	0.039	-
26	Rapid urban growth (Percent land cover change to urban areas from base year)	r _s	0.791**	0.765**	0.865**	-
		Sig.	0.000	0.000	0.000	-
27	Waterbodies density	r _s	0.702**	0.686**	0.709**	-
		Sig.	0.001	0.002	0.001	-
28	Access to hospital	r _s	0.660**	0.557**	0.644**	0.561**
		Sig.	0.001	0.007	0.002	0.005
29	Movement potential	r _s	0.526*	0.453*	0.584**	0.699**
		Sig.	0.012	0.034	0.005	0.000

**Correlation is significant at the 0.01 level (2-tailed) *Correlation is significant at the 0.05 level (2-tailed)

Table 4: Ambiguity of indicators

ID	Indicators	Outcome	Persistence rate	Degradation rate	Peak failure	Recovery rate
10	Percent land area not in an inundation zone (100 years)	r _s	0.537**	0.461*	-	-
		Sig.	0.008	0.031	-	-
16	Percent land area of developed open spaces	r _s	0.570*	0.562*	0.520*	0.713**
		p	0.011	0.015	0.027	0.001
18	Hospitals per square mile	r _s	0.678**	-	-	0.478*
		Sig.	0.001	-	-	0.045
20	Hotels and motels per square mile	r _s	0.577**	0.469*	0.525*	-
		p	0.010	0.050	-	0.025
28	Access to hospital	r _s	0.561**	0.660**	0.557**	0.644**
		p	0.005	0.001	0.007	0.002
29	Movement Potential	r _s	0.699**	0.526*	0.453*	0.584**
		p	0.000	0.012	0.034	0.005

**Correlation is significant at the 0.01 level (2-tailed); *Correlation is significant at the 0.05 level (2-tailed)

given indicator has a direct relationship with community resilience and negative direction refers to inverse relationships. Results of this study revealed an ambiguity in the direction of six indicators when testing the association with different system-performance measures (Table 4). All six indicators are theoretically presumed to have a positive relationship with community resilience. As presumed, all of them revealed a positive association with the persistence rate. Nevertheless, this set of indicators also revealed a positive association with degradation rate, peak failure and recovery rate. Positive association with persistence rate indicates higher community resilience whereas the positive association with other three measures indicates lower community resilience.

In the cases of ‘percent land area not in an inundation zone’, ‘hotels and motels per square mile’ and

‘hospital per square mile’, the degree of ambiguity is not severe. The association with persistence rate is moderately strong and highly significant whereas the association with other measures are weak and less significant. Therefore, these indicators can be considered as maintaining a direct association with community resilience despite the minor internal inconsistency.

Percent land area of developed open spaces is often considered as a spatial feature indicating the urban resilience. Relative to other urban land uses, open areas infiltrate more, evaporate more and thereby runoff less. In case of Colombo, ‘percent land area of developed open spaces’ have revealed stronger and more significant association with recovery rate than the persistence rate. Detailed observations on Colombo case study noticed two possibilities that might have influenced the

results. First, many of these developed open spaces are located within the floodplain of the Kelani River. Floodplains lay at lower elevations closer to water bodies and are often subjected to higher flood heights. Soil hydraulic properties of flood plains facilitate water retention and detention holding for a longer time. On the above ground, it is logical for any land use on the floodplain to take a longer time to recover. Second, some parts of the flood plain in Colombo are highly densified including the vicinity of the developed open spaces. High-density development in floodplains increases the magnitude of damage making it difficult to recover once degraded. To support this reasoning, the study tested the relationship of percent land area of developed open spaces with elevation ($r_s = -0.675, p < 0.000$) and the population living in high-intensity urban areas ($r_s = 0.846, p < 0.000$). Accordingly, the ambiguity of this indicator can be interpreted as a result of multicollinearity with indicators that have inverse associations. Therefore, employing this indicator for assessing community resilience requires caution regarding the location and vicinity of such open spaces.

The real challenge of ambiguity could be noticed in ‘access to hospital’ and ‘movement potential’ because these two indicators revealed highly significant associations to both directions. There is a similarity between them regarding constituents. Access to hospitals is based on Euclidian distance to hospitals and movement potential is based on Euclidian distance to roads. The correlation between these two indicators is also highly significant and strong ($r_s = 0.949, p < 0.000$). However, there is no clarity as to whether such indicators represent resilience communities or non-resilient communities. Therefore, these two indicators should be avoided in resilience assessments despite the significant association. Overall ambiguous indicators require further investigations to elaborate them with causal relations, primarily because ambiguity can threaten the internal validity of resilience assessment.

Geospatial indicators by the resilience capacities: The study investigates the adequacy of geospatial indicators for assessing distinct capacities of community resilience. As mentioned previously, four system-performance measures were attributed to three capacities such as persistent rate to transformative capacity, the inverse of recovery rate to recovery capacity and inverse values of peak failure and degradation rate to absorptive capacity. The association of geospatial indicators with four system-performance measures infers their ability to represent the correspondent resilience capacities.

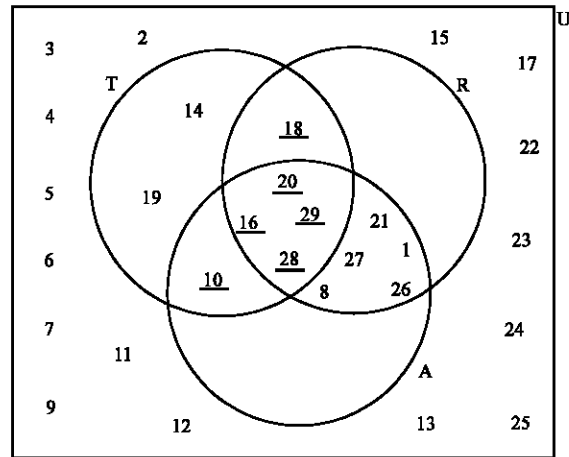


Fig. 3: Relationships of indicators with resilience capacities three overlapping sets in the Venn diagram illustrates how geospatial indicators are associated with resilience capacities. The 29 items in the venn diagram represent the set of geospatial indicators tested in this study. Set ‘A’ refers to the absorptive capacity, Set ‘R’ refers to the recovery capacity and set ‘T’ refers to the transformative capacity

This study tested the association of 29 geospatial indicators with community resilience to floods. Some indicators were only associated with one capacity while some of the others were only associated with either two or all three capacities. The Venn diagram provided in Fig. 3 illustrates the coherent relationships of all indicators with three resilience capacities. Accordingly, three overlapping sets in the Venn diagram represent three capacities of resilience. Each set contain indicators that reveal significant ($p < 0.05$) associations with the correspondent system-performance measures. Indicators which are ambiguous concerning the direction to different resilience capacities have been underlined in the Venn diagram.

Followings are the detailed inferences of the Venn diagram:

- A = {1, 8, 10, 16, 20, 21, 26, 27, 28, 29}
- R = {1, 8, 18, 16, 20, 21, 26, 27, 28, 29}
- T = {14, 18, 19, 16, 20, 28, 29}
- $(A \cap R \cap T)$ = {16, 20, 28, 29}
- $(T / (A \cap R))'$ = {14, 19}
- $(T / A \cup R)$ = {1, 8, 10, 14, 16, 18, 19, 20, 21, 26, 27, 28, 29}
- $(T \cup A \cup R)'$ = {2, 3, 4, 5, 6, 7, 9, 11, 12, 13, 15, 17, 22, 23, 24, 25}

Sets of absorptive capacity (A) and recovery capacity (R) contain ten indicators each following the eight indicators in the set of transformative capacity (T). Overall, geospatial indicators can represent all three resilience capacities. When comparing the relative component zones by capacities, only transformative capacity ($T/(A \cup R)$) contains indicators. 'Schools (primary and secondary education) per square mile' ($r_s = 0.783, p = 0.000$) and 'percent areas that has changed into urban' ($r_s = -.742^{**}, p = 0.000$) are uniquely to transformative capacity. In contrast, the unique indicators of other two capacities could not be distinguished.

If an indicator can represent all three capacities well, such indicators are better options for incorporating into assessment tools. If so, the assessment can perform efficiently with fewer data. However, any of the common indicators ($A \cap R \cap T$) cannot be confidently recommended due to ambiguity. Four indicators revealed non-ambiguous, significant associations with recovery and absorptive capacities. Rapid urban growth ($r_s = 0.791, r_s = 0.765, r_s = 0.865$ at $p < 0.01$) and water bodies density ($r_s = 0.702, r_s = 0.686, r_s = 0.709$ at $p < 0.02$) strongly and significantly associated with degradation rate, peak failure and recovery rate. 'Population living in high intensity urban areas' ($r_s = 0.569, r_s = 0.647, r_s = 0.583$ at $p < 0.01$) and 'density of commercial infrastructure' ($r_s = 0.474, r_s = 0.491, r_s = 0.490$ at $p < 0.05$) revealed moderate associations with the above. These four geospatial indicators well capture how high urban density which is due to the unplanned development in the case of Colombo, weakens community resilience making severe degradations and time-consuming restorations.

There were 13 indicators ($T \cup A \cup R$) that revealed significant associations with at least one resilience capacity. The rest of the 16 indicators has revealed no significant association ($(T \cup A \cup R)$). However, this validation test is not capable enough to nullify the utility of these indicators, primarily due to the limited scope of transformative capacity. The study tested the transformative capacity by persistence state of the system performance curve (Fig. 2). The persistent state covers only part of transformative capacity and the rest must be tested with the adaptation state. The study could not test the long-term adaptation due to data constraints. Therefore, at least some of these indicators ($(T \cup A \cup R)$) might show an association with the adaptation state.

Nevertheless, future studies can further validate the results mainly with three advancements. First, these findings are based on one critical flood event, therefore, the validity must be generalized after testing a series of flood events at different magnitudes. Second, the scope of the outcome variables in this study are limited to the function of fulfilling the basic needs but the overall

resilience of community can be captured by observing the other functions community systems and the other elements such as infrastructure resilience. Third, several geospatial indicators could not be tested because the study area is an urbanized region where some land uses including forests, grasslands and rangelands were not presented within the study area. Therefore, an expanded region including broader peripheries or agricultural region can further validate geospatial indicators.

CONCLUSION

This study consolidated a set of 29 geospatial indicators from existing community resilience assessment methodologies and tested their validity and adequacy to assess community resilience against floods. Per the literature review, geospatial indicators were very limited in existing assessment methodologies. Initial findings of the study listed 13 geospatial indicators that show significant associations to system-performance measures. The detailed analysis of this study could detect ambiguities regarding the association among distinct capacities. Decision-makers ought to be cautious of such ambiguities because it can weaken the internal validity of assessment and can misguide resilience-building actions.

Validated geospatial indicators demonstrated the capability to represent all three capacities of resilience. Minimum or no change to urban areas and school density has uniquely represented the transformative capacity of the socio-ecological system. High densities of water bodies, residential population infrastructure density and rapid urban growth mutually represented weakened absorptive and recovery capacities of the system. The dynamics of indicators at different states of system performance curve embodies the processes of community actions through life-cycle stages of community resilience. The aggregated resilience index should signify each of the state, ensuring that all types of capacities are adequately accounted in the assessment.

Statistically validated indicators can be employed in resilience assessment methodologies with high reliability. Overall, results clearly revealed that geospatial indicators can demonstrate the resilience processes and behaviors of socioecological systems hence, they can be utilized to measure the community resilience. Particularly, geospatial indicators well capture the effect of increasing vulnerability due to the intensive physical development and perturbed natural flood defense mechanisms. In an urbanizing world where flood damages are outnumbered, geospatial indicators can provide deep insights into resilience-building initiatives. Therefore, geospatial indicators can strongly be recommended in community resilience assessment tools.

SUGGESTIONS

Further studies on assessing the validity and adequacy of indicators can make the assessment process more scientific and comprehensive leading towards promising resilience-building initiatives.

REFERENCES

- Abenayake, C., M. Yoshiki, A. Marasinghe, Y. Takashi and I. Masahiro, 2016. Applicability of extra-local methods for assessing community resilience to disasters: A case of Sri Lanka. *J. Environ. Assess. Policy Manage.*, 18: 1-28.
- Bakkensen, L.A., C. Fox-Lent, L.K. Read and I. Linkov, 2017. Validating resilience and vulnerability indices in the context of natural disasters. *Risk Anal.*, 37: 982-1004.
- Bocchini, P. and D.M. Frangopol, 2010. Optimal resilience-and cost-based postdisaster intervention prioritization for bridges along a highway segment. *J. Bridge Eng.*, 17: 117-129.
- Bocchini, P. and D.M. Frangopol, 2012. Restoration of bridge networks after an earthquake: Multicriteria intervention optimization. *Earthquake Spectra*, 28: 426-455.
- Bocchini, P., D.M. Frangopol, T. Ummenhofer and T. Zinke, 2013. Resilience and sustainability of civil infrastructure: Toward a unified approach. *J. Infrast. Syst.*, Vol. 20,
- Bruneau, M. and A. Reinhorn, 2007. Exploring the concept of seismic resilience for acute care facilities. *Earthquake Spectra*, 23: 41-62.
- Bruneau, M., 2006. Enhancing the resilience of communities against extreme events from an earthquake engineering perspective. *J. Secur. Educ.*, 1: 159-167.
- Bruneau, M., A. Filiatrault, G. Lee, T.D. O'Rourke and A. Reinhorn *et al.*, 2005. White paper on the Sdr Grand challenges for disaster reduction. Multidisciplinary Center for Earthquake Engineering Research, USA.
- Bruneau, M., S.E. Chang, R.T. Eguchi, G.C. Lee and T.D. O'Rourke *et al.*, 2003. A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra*, 19: 733-752.
- Burton, C.G., 2015. A validation of metrics for community resilience to natural hazards and disasters using the recovery from Hurricane Katrina as a case study. *Ann. Assoc. Am. Geographers*, 105: 67-86.
- Cai, H., N.S.N. Lam, L. Zou, Y. Qiang and K. Li, 2016. Assessing community resilience to coastal hazards in the Lower Mississippi River basin. *Water*, 8: 1-18.
- Cimellaro, G.P., 2016. *Urban Resilience for Emergency Response and Recovery*. Springer, Berlin, Germany, ISBN:9783319306568, Pages: 522.
- Cimellaro, G.P., A.M. Reinhorn and M. Bruneau, 2010b. Framework for analytical quantification of disaster resilience. *Eng. Struct.*, 32: 3639-3649.
- Cimellaro, G.P., A.M. Reinhorn and M. Bruneau, 2010a. Seismic resilience of a hospital system. *Struct. Infrastruct. Eng.*, 6: 127-144.
- Cutter, S.L., 2016. The landscape of disaster resilience indicators in the USA. *Nat. Hazards*, 80: 741-758.
- Cutter, S.L., L. Barnes, M. Berry, C. Burton and E. Evans *et al.*, 2008a. A place-based model for understanding community resilience to natural disasters. *Global Environ. Change*, 18: 598-606.
- Cutter, S.L., L. Barnes, M. Berry, C. Burton and E. Evans *et al.*, 2008b. Community and regional resilience: Perspectives from hazards, disasters and emergency management; CARRI research report, 1. Community and Regional Resilience Institute, South Carolina. http://www.resilientus.org/wp-content/uploads/2013/03/FINAL_CUTTER_9-25-08_1223482309.pdf.
- DMC., 2016. Flood extent for Kelani ganga river basin and surrounding. Disaster Management Centre, Colombo, Sri Lanka. <http://www.dmc.gov.lk/IndexInundationMaps.htm>.
- Klein, R.J., R.J. Nicholls and F. Thomalla, 2003. Resilience to natural hazards: How useful is this concept?. *Global Environ. Change Part B Environ. Hazards*, 5: 35-45.
- Kotzee, I. and B. Reyers, 2016. Piloting a social-ecological index for measuring flood resilience: A composite index approach. *Ecol. Indic.*, 60: 45-53.
- Larkin, S., C. Fox-Lent, D.A. Eisenberg, B.D. Trump and S. Wallace *et al.*, 2015. Benchmarking agency and organizational practices in resilience decision making. *Environ. Syst. Decisions*, 35: 185-195.
- Mugume, S.N., D.E. Gomez, G. Fu, R. Farmani and D. Butler, 2015. A global analysis approach for investigating structural resilience in urban drainage systems. *Water Res.*, 81: 15-26.
- Parsons, M., S. Glavac, P. Hastings, G. Marshall and J. McGregor *et al.*, 2016. Top-down assessment of disaster resilience: A conceptual framework using coping and adaptive capacities. *Intl. J. Disaster Risk Reduct.*, 19: 1-11.
- Ranjan, E.S. and C.C. Abenayake, 2014. A study on community's perception on disaster resilience concept. *Procedia Econ. Finance*, 18: 88-94.
- Sharifi, A. and Y. Yamagata, 2016. On the suitability of assessment tools for guiding communities towards disaster resilience. *Intl. J. Disaster Risk Reduct.*, 18: 115-124.

- Tate, E., 2012. Social vulnerability indices: A comparative assessment using uncertainty and sensitivity analysis. *Nat. Hazard.*, 63: 325-347.
- UNISDR., 2016. The human cost of weather-related disasters 1995-2016. United Nations International Strategy for Disaster Risk Reduction, Geneva, Switzerland.
- USAID., 2007. How resilient is your coastal community?. United States Agency for International Development, Washington, USA.
- Wang, C.H. and J.M. Blackmore, 2009. Resilience concepts for water resource systems. *J. Water Resour. Plann. Manage.*, 135: 528-536.

The Creation of Thai Animation: The Analysis Through the System Theory

¹Samrit Bhukonglee, ¹Preechawut Apirating and ²Nattapong Yamcharoen

¹Faculty of Fine and Applied Arts, Khon Kaen University, Nai Muang Sub-District,
Muang District, 40002 Khon Kaen, Thailand

²Faculty of Communication Arts, Kasem Bundit University, 1761 Pattanakarn Road,
10250 Suanluang, Bangkok, Thailand

Abstract: This research is a qualitative research studying 8 full length Thai animations: Sudsakorn, Khan Kluay, The Life of Buddha, Nak, Khan Kluay 2, Echo Planet, Yak the Giant King and The Story of Mahajanaka. As for the research process, the data was collected from the direct practitioners through in-depth interviews. The data was then analyzed through system theory in order to acquire the history of Thai animations. From the research, it was found that there are three important inputs: artists, budget and technology. This research created 2 different types of production process of the animations: the 2 dimensions productions which result in 2 following animations: Sudsakorn and The Life of Buddha The 3 dimensions productions which result in 6 following animations: Khan Kluay, Nak, Khan Kluay 2, Echo Planet, Yak the Giant King and The Story of Mahajanaka. After the release of each animation, what followed were gross and many rewards. Khan Kluay was the most successful by winning over 79 million baht and won over 4 Thailand National Film Association Awards which include best movie script and best movie of 2006.

Key words: Animation, Thai animation, history of Thai animations, applied sciences, system theory, Sudsakorn, Mahajanaka

INTRODUCTION

Human have always full length to record the movement of the things around them, since, ancient time. Cave paintings which were made by pre-historic people shows images of 4 legged creatures, sometimes 5 or 6. Williams (2002) had noticed that the fact that these pre-historic men have more legs than reality was that they were trying to record the movements of those animals while they were running. Thousands of years after wards, men had discovered "Persistence of vision". This principle explained that human vision will remember any pictures that they saw for a couple of moments eventhough there were newer pictures that replaces the old ones which also related to Yeon-Hee Hwang and Hai Beom Kim have mentioned in their study that modern innovation technology is shifting toward human-oriented technology structure combining technology and emotion. So, the principle of persistence of vision made men realize the connection between the old pictures and the new pictures they saw. This principle turned into moving pictures. When still pictures were shown through human eyes continuously at a certain speed which led to the creation of animations later on.

In 1896, Thosmas Edison who were at the time experimenting with a project on creating moving pictures on films has invited James Stuart Blackton to joined his project. Blackton's job was to draw series of pictures that shows animal's movment little by little in each frame. Once, Blackton was done with the pictures, Edison would then took the pictures and put them together at a later process. In 1906, they released their work which was called "Humorous Phases of Funny Faces". That was the first time that drawings and pictures were put together which could be considered as the ancestors of animations (Williams, 2002).

Up until the 1910s, a short animation, Felix had become famous all over the world through simple black and white drawing which shows the charms of the character. "Felix" was the animation that brought Walt Disney to the animation industry which resulted in, 1928 when Mickey Mouse first appeared in "Steam Boat Willie" which was the first animation of the world with recored audio. That was the first development that Walt Disney had brought to the world of animation. Disney had created anothere miraculous thing in 1932 with an animation called "Flowers and Trees" which was the first coloured animation in the world. A year after that "Three

Little Pigs” came out. Later in 1937, Walt Disney released “Snow White and the seven Dwarfs” to the audiences. That was the first full length animation of the world. That animation had perfectly raised the level of acceptance for animations to be considered as art with an impressive story throughout the whole 83 min. The success in both the gross and the critics of Snow White and the Seven Dwarfs became the basic of many other animations by Walt Disney (Williams, 2002).

Up until the year 1995, Pixar Animation Studio had created a new history to the animation business through the creation of the first full length 3D animation called “Toy Story”. This kind of animation is different from the hand drawn animations by Disney. It was created through the rendering of the pictures through the calculation of computers. The production of this kind of animation was accepted and spread quickly throughout the world, including Thailand.

As for the animations in Thailand, it all started from the effort of Payut Ngaokrachang, who was considered as “The Father of Thai Animation” or “Thai Walt Disney”. Payut pioneered Thai animations from short animations as an experiment in 1955 call “A Miracle One Morning” (in Thai). Later on when he became more skilled in making animations, Payut had created the important history for Thai animation through the creation of “Sudsakorn”, the first full length animation of Thailand. It was a pity that no full length Thai animation was ever made ever, since, “Sudsakorn”, it wasn’t until 2006 or 27 years later, that Kantana Animation Studio created the second full length Thai animation called “Khan kluay”.

Khan Kluay was create using 3D process like the animations by Pixar. Khan Kluay by Kantana was considered very successful until many companies in Thailand turned their attention to the production of many full-length animations in Thailand. Since, 2006 when Khan Kluay was released to 2014, there were all together 7 full length animations that was shown in Thai movie theatre: Khan Kluay, The Life of Buddha, Nak, Khan Kluay 2, Echo Planet, Yak the Giant King and The Story of Mahajanaka. Including Sudsakorn, Thailand will have altogether 8 full length animations. However, throughout these 10 years, the histories of all the 8 animations were never recorded in detail. There was only public relation information which were not in-depth information with no contain no useful historical information for Thai animation business.

Therefore, the goal of this research was to bring up this topic of compiling the information to answer the question of how Thai animations were created, especially, after 2006 as well the questions to why full-length Thai animation came into continuous production after having disappeared for 27 years.

The objective of the study: To study the history of full length Thai animation.

The scope of the study: The animations that will be studied has to be full length feature films produced by Thai people and came into the theatre in Thailand. These consist of 8 following animations: Sudsakorn, Khan Kluay, Khan Kluay 2, The Life of Buddha, Nak, Yak the Giant King, Echo Planet and The Story of Mahajanaka.

MATERIALS AND METHODS

The history of Thai animations this time is the study of only 8 full length Thai animations which consist of Sudsakorn, Khan Kluay, The Life of Buddha, Nak, Khan Kluay 2, Echo Planet, Yak the Giant King and The Story of Mahajanaka. The collection of the data was done through the interviews of these practitioners:

- Khan Kluay, Interview, Kompin Kemgumnird, Director
- The Life of Buddha, Interview, Dr. Wallapa Pimthong Producer
- Nak, Interview, Natthaphong Ratanachoksirikul Director
- Khan Kluay 2, Interview, Taweelap Srivuthivong Director
- Echo Planet, Interview, Kompin Kemgumnird Director
- Yak the Giant, Interview, Punlop Sinjaroen, Assistant Director
- The Story of Mahajanaka, Interview, Nop Dharmavanich, Director

Apart from the in-depth interviews of the practitioners, this research also compiled the data from the documents due to the fact that Payut Ngaokrachang, the Director of Sudsakorn had passed away in 2010. Therefore, the collecting of the data about the animation needed research from the following documents:

A book called “Life, Payut, Animation” (in Thai) which is a biography of Payut’s life and work by Suttana (2006).

A book called “Yak the Giant King” which has the same name as the movie edited by Kunsachee Wongsawattananukul. The book compiled the interviews of the director, the crew and all the cast on Yak the Giant King.

Audio recording of an interview with Payut Ngaokrachang. The name of the interviewer and time of the interview is unknown. This recording is being kept in the Library and Mediatheque “Cherd. Songsri”, Salaya, Nakhon Pathom.

From the information from the research, the data will be analyzed through “System theory”. This theory was adapted by Thasanabanchong (2002a, b) to analyze Thai movie industry in the past. Therefore, this theory was used to study animations which can be seen from the overall Thai animation industry from the past to the present. This theory consists of 4 analytical aspects:

Input which consist of all important components that contribute to the production of an animation such as the budget and the knowledge in the production.

Process which is the production process of each animation. Output which is the unique characteristic of each animation that was created such as length and year.

Outcome which is the result that comes after the release of each movie. This research will analyze 2 aspects: the gross and the awards in this case, the research only focuses on Thailand National Film Association Awards. After the nalalysis of the data, the result will be presented through descriptive analysis.

Result of the study: From the in-depth interview with the directors, the assistent directors and the producers of all the 7 animations in combination with the research documents, the history of each animation are as follows:

Sudsakorn: Saneh Klaikluan was the first Thai person to successfully create “Running animation”. But the experiment itself was never developed into animations, since, it was never supported by the government. When Saneh passed away, the knowledge in producing moving pictures went away with him. Years later, younger animation artist like Payut Ngaokrachang made Saneh’s dream became a reality.

Payut learned from Saneh’s colleague that Saneh drew his animation on a transparent sheet. Payut decided to experiment by using the X-ray films from the hospital and wash them in lye in order to make the film transparent and draw animations on them. Through this experiment, Payut discovered and became confident that that was the way moving cartoon should be done. Payut decided to redo the experiment by buying transparent sheets and draw animations on them. This experiment ended up as the first Thai animation called “a miracle one morning” which was 20 min long. The film was shown at Chalermthai on July 5th, 1955. Payut created 2 more short animations later on called “the new adventure of hanuman” (in Thai) and “A Child and A Bear” (in Thai) which were shown in 1960 (Suttana, 2006). After that, Payut never create any more animation until 16 years later when he decided to do what he most desired in life, to create full length animation.

Payut started from asking for funding starting with police lieutenant general Bantheung Gumpanartsaenyakorn who was willing to take the risk and invest 3 million baht under the condition that Payut needed to be responsible for all the production process and the gross from the movie should be split to each of them every other time the movie was shown. After receiving enough funding, Payut started the project right away by searching for an appropriate story and eventually chose “Sudsakorn”, a character from Sunthornpoo’s main novel “Phra aphai mani” to be made into an animation. Due to the fact that Payut saw Sudsakorn as a child that was raised by a hermit. Sudsakorn therefore was a kid with morality who can be an example of Thai kids who would be watching the animation.

As for the production process, Payut was responsible for the main drawing that shows the movement throughout the story at night due to the fact that Payut still need to work at USIS during the day and his students will be colouring the drawings during the day. The drawings were then filmed using a 35 mm. film. Payut needed to draw over 66,000 draings alone until he permanently lost his right eyesight in 1979. Due to the length of 82 min, Sudsakorn was shown for the first time on 13 April 1979. That was the first full length animation of Thailand. Sudsakorn was able to make 2 million baht which was considered very high at the time (Suttana, 2006).

Khan Kluay: Khan Kluay was the first full length animation from Kantana Animation Studio Ltd. Eventhough Khan Kluay was the most famous animation, it had an origin as a television series.

Mr. Jareuk Kaljareuk, the owner of Kantana told the work team that the new serie for the copany needs to be about elephants. The script writers needed to do a lot of homework after receiving the assignment and they found one of the most interesting ideas, that was to tell a story of King Naresuan during the elephant duel. The writer team named his elephant Khan Kluay according to the character of the elephant that had a beautiful arch back similar to the Khan Kluay or the banana branch and used it as the title.

The script team had to tell the story of an elephant by making the elephant into a living and talking freely just likes human beings. Through this kind of story telling, Kantana found that there was no way that they can shoot using a real elephant while making them do what they needed them to do, so, the created an elephant using visual effects. That wasn’t possible in Thailand due to the

limitation of human resources and the technology requires doing it. The only option left for them at the time was to turn this tv serie into an animation.

Kantana started a search for the crew that would make this animation a reality. The first team that came in to over see the production was a team of Foreigners which consist of Tod Polson as the Director and visual developer, Ariel Prendergast as script writer and Aaron Sorensen as the Assisting Director. These three person oversee the Thai crews. Tod's team had created a trailer for the animation when Kantana saw the production by Tod's team, they decided to move Khan Kluay up from being a TV serie into a full-length movie to be shown in theatre. Unfortunately, after a year of production, Kantana wasn't impressed with the story line that Tod's team came up with. After the contract of the first team ran out, Kantana decided to not renew the contract. Khan Kluay, therefore, fell into the hand of a new, young blood, Thai director, Kompin Kemgumnird.

Kompin have had an experience working with Walt Disney as an in-betweener in 2 movies, Tarzan (1999) and Atlantis, The Lost Empire (2001). After having left Disney, Kompin had a chance to work with Blue Sky Studio in the company's first production of Ice Age (2002). It was here in Blue Sky that Kompin got promoted to an animator and started to learn about the production process of 3D animations. This knowledge and experience was what Kompin was using in directing Khan Kluay, the animation later on.

Khan Kluay was the first full length animation in Thailand that was produced using 3D softwares throughout the whole production. Khan Kluay was released to the movie theatre on May 18, 2006. It was the second full length animation in Thailand after Sudsakorn.

The Life of Buddha: This animation got started from a woman named Dr. Wallapa Pimthong who was responsible for the production of this animation. Before becoming an animation creator, Dr. Wallapa Pimthong had an opportunity to be a coordinator for a project called "Genius classroom" by the ministry of education. This project required a creation of an interactive media and an animation character as a medium to communicate to the learners. This had brought Dr. Wallapa Pimthong to get to know many animation creators and saw the influence of Thai animations towards the audiences, especially, children. She came up with an idea that if she wants to send a message to the society doing through animations would create an impact at a larger scale. This is the starting point that sparks Dr. Wallapa Pimthong to want to make animations (Pimthong, 2015).

Dr. Wallapa had invited Asst. Prof. Krismant Whattananarong, one of the committee members that involved with the quality control for the "Genius classroom" to sit in the Director chair for this movie and invited animators whom she used to work with to be part of the production of this animation. Most of the animators were from a company called Tai Wang. It just, so, happened that during the time that Dr. Wallapa started to produce the animation called The Life of Buddha, Tai Wang company was on its way of closing down. Many animators who still loved making animations decided to come work for Dr. Wallapa. This allowed the creative team to move on with their work without any obstacles. As for the story of the animation, there was none other than what Dr. Wallapa had always had on her mind, only the story of Lord Buddha due to the fact that she really had a strong belief in Buddhism.

After the production process gets started, the biggest obstacle for the production was the budget. No financial institution wanted to give a loan out to this company due to the fear that that the movie would lose profit and would not have money to pay back to the bank. The government themselves also didn't see the importance and provide no funding. However, Dr. Wallapa had to put her house her land and her car in the bank in order to take out a loan as budget for this Lord Buddha film.

"The Life of Buddha was created using 2D production through the drawings on several frames to show the movement of the characters. Each drawing will then scanned into digital files and later, colored using a computer program. The post process after this point was all done on computers until it was released as a movie. After 4 years of dedication and effort, The Life of Buddha, the animation was shown in movie theatres on December 5, 2007. It was the third full length animation of Thailand with the duration of 98 min.

Nak: Nak was the first and only animation produced by Beboydeg Ltd., company, established by Boyd Kosiyabong a well-known song writer in Thailand. Apart from music, animation was another thing that he loved. Boyd Kosiyabong got to meet with many cartoonists and set up a team called B-Boyd characters working on two magazines called Katch and Manga Katch as the two main projects. Due to financial problems later on, B-Boyd characters had to be discontinued these magazines. However, the cartoonist working at the company at the time still love drawing, Nuttaphong Ratanachoksirikul, the editor for the two magazines at the time had to struggle to find some work for his crew. He came up with an idea of making an animation as a television show. He invited his

friends who used to work on animations from other countries to come and help trained the team. B-Boy character was then changed to Beboydcg Company Ltd. And continuous produced animation shows for TV.

The team at the time had created a trailer for a full-length animation with the story of Japanese Super Hero team. They called this movie “Cyberger”. Eventually, they got an opportunity to present this trailer to Somsak Techaratanaprasert the owner of Sahamongkol Film International which is a big movie production company of Thailand. After watching the trailer Somsak allowed this team to continue producing the movie, except that it has to be about “Mae Nak”, the most famous love story of Thailand between man and ghost. If the movie isn’t about Mae Nak, Somsak will not allow the movie to be made. Beboydcg didn’t hesitate to grab this opportunity to produce the movie even though it wasn’t what they wanted to do.

After that the team started to write the script by having Nuttaphong Ratanachoksirikul as the director. In this animation, they were trying to redefine Mae Nak. They wanted to tell a story of Mae Nak in the angle that has never been talked about before. Eventually, the movie was just borrowing the character of Mae Nak from the Thai folklore in a whole new story line which talked about the love between Mae Nak and her kid, the devotion that a ghost like Mae Nak tried to give to human in order to maintain the peace of humanity.

This movie took about 6 years to make. It was released to the theatre on April 3, 2008. It was the 4th full-length animation of Thailand with the duration of 90 min.

Khan Kluay 2: The first Khan Kluay was released in 2006 and became really successful in both the gross and the awards. Kantana therefore approved of the second animation project under the name Khan Kluay 2. In this sequel, the story was about Khan Kluay after winning the elephant duel in the first one. Khan Kluay went to live in the palace with Chabakaew his beloved wife. However, Khan Kluay needed to go out to war very often that he didn’t have time to take care of his family. This conflict between the love of the country and his family was what the Khan Kluay 2 was trying to communicate to the audience.

For Khan Kluay 2 originally, Kantana wanted Kompin Kemgumnird to take over as the director but Kompin denied. He proposed Taweelap Srivuthivong to take over as the director of Khan Kluay 2 instead and Kantana didn’t oppose to that Kemgumnird. Taweelap wasn’t a stranger but one of the person who was also involved in the first Khan Kluay movie. Taweelap had been working

as an illustration artist until he got a chance to work as one of the crew in the first Khan Kluay movie as the Art Director. The main responsibility for Taweelap was to control the beauty of each frame and every detail that would go on screen. Apart from being the art director, he also contributed many suggestions and ideas to the Khan Kluay team Srivuthivong. Therefore, the fact that Kompin proposed Taweelap to take over the project wasn’t a coincidence but he was confident about Taweelap’s skill and was certain that he could lead Khan Kluay 2 to the same success as the first one.

Khan Kluay 2 took around 3 years to produce and was released on March 26, 2009. It was the second animation by Kantana but the 5th full-length animation film of Thailand.

Echo Planet: This movie was directed by Kompin Kemgumnird, the director from the first Khan Kluay movie. Due to the fact that he gave over Khan Kluay 2 project to Taweelap to direct which allowed him to have a lot of time on his hand to come up with the ideas to tell the story. After the production Khan Kluay 2, Kompin came back again with the script of a new animation that he lay the plot out himself in which he called “Echo Planet”.

Echo Planet was an effort by Kompin to present the picture of Thai rural area and the lives of the long neck Karen people who live far away from the technology and material world. Two Karen siblings, the main character saved Sam’s life by accident. Sam was the only son of the President of Capital State, the most powerful country of the West. In order to repay the debt, Sam decided to bring the two siblings to Capital City, the capital to Capital State. It was there that the director showed the clash between the “Old world” and the “New world” which was what he intended to present. At the end of the day what the movie wanted to show the audience was that the old way of life which was the way of nature was the best way to counter global warming and can sustainably heal our world.

Similar to the two animations released earlier by Kantana, Echo Planet took around 3 years to produce and was released on August 2, 2012. This was the 6th full-length animation of Thailand and the 3rd for Kantana with the duration of 90 min.

Yak the Giant King: This movie came from the personal reference of Prapas Cholsaranont. He was an animation lover and had dreamed that one day he would direct his own animation. Currently, Prapas Cholsaranont is working with Panya Nirunkul and had opened up a company called Workpoint Entertainment Ltd. and had produced

television shows. Even with his success from his various works in the entertainment business, he hasn't forgotten about his dream to create his own animation. One day he met a man who would later on make his dream a reality; Chaiporn Panichrutiwong, a young artist from Silpakorn University who has had experience with the world's renowned game company (Piampiyachart, 2012).

After returning to Thailand, Chaiporn went and worked with Vithita Animation Company Ltd. At the time, the company had brought "Pang Pond", a famous character from "Kai Hua Roh" magazine and turned it into an animation through 3D production process that he had learned while he was abroad which had become a production base for Vithita Company. Eventually, "Pang Pond, the Animation" was created in a 3D animation as a TV serie in 2002. After the team had created the first 3D animation for Thailand, Chaiporn moved to work with Prapas Cholsaranont at Workpoint and that was the time that Chaiporn needed to pioneer the first animation for Workpoint.

Once he got his right hand man, Prapas moved ahead with his project. He used a famous Thai novel and turned it into an animation. The movie that he picked was "Ramayana". It was a story that had been ingrained into the minds of the Thai people. But to tell the whole story of Ramayana, the team felt that it was difficult to just pull out some of the parts and tell it to the audiences in a limited amount of time. Especially, if the audiences don't have any background knowledge of the story who will never understand what was going on. With these limitations, Prapas Cholsaranont came up with the new idea of bringing Ramayana into a new setting. (The same kind of concept that was used in "Nak" by beboydcg).

For this movie, after starting writing the script in 2004, the movie took 8 years in production through 3D production process by using the same software as Khan Kluay by Kantana. The movie was released on October 4, 2012. It was the 6th full-length animation of Thailand with the duration of 100 min.

The Story of Mahajanaka: The Story of Mahajanaka was the only animation that came from the government's initiative which was the Software Industry Promotion Agency (Public Organization) or SIPA. This department is responsible for support and develops digital contents within Thailand. Therefore and support for the animation production fell into the hands of SIPA. However, SIPA was just a supporting unit, it doesn't have enough manpower to produce the actual content. Therefore, SIPA needed to find assistants in order to make this project a success. It was considered luck for SIPA that

Thailand already had a group of animation firms that came together called Thai Animation and Computer Graphids Association (TACGA). SIPA therefore, contacted TACGA to spear head the production of "The Story of Mahajanaka" by department will be providing the budget in the production. Eventhough it wasn't a huge budget, since, it was the animation to commomerade King Rama IX, TACGA was very honored to be part of this project.

The board of TACGA had a meeting and chose 3 representatives to be the director of this animation. The 3 directors were Kriengkrai Supornsahusrungsri, Thananchanok Suban Na Ayudhaya and Nop Dharmavanich. After TACGA came up with the directors, they discussed further about which company will be leading the production. The committee came to the same conclusion that producing a full-length animation was too big for 1 company to handle. This had created a new phenomenon that had never happened in Thailand which was the collaboration of several companies to produce one animation. TACGA had asked for volunteers within the associations and had acquired 15 companies. The 3 directors had a meeting to deligate the jobs between the companies and the deadline for the jobs. They also acted as advisors providing the feedback and suggestions until the work came out good.

The Story of Mahajanaka was a 3D animations and the 8th full-length animation of Thailand with the duration of 110 min. It was primiered on November 28, 2014. After that the animation was shown in all the theatre within the Major brand on November 29 and December 6, 2014, 2 times per day; 10 am and 3 pm in order for the Thai people to go see for free. It was also shown on television from the December 6-8, 2014.

Conclusion and the analysis of the system theory frame:

The information above was the background of all the 8 animations that was acquired through indepth interviews from the practitioners and the document research. If we were to analyze the information using the system theory, we would be able to see the big picture of the history of Thai animations clearly. According to the system theory, the analysis will be done in 4 points: input, process, output and outcome.

Input: The input within the system theory means that basic componenets that is required in the production of the animations. From the analysis of the 8 animations, it was found that the most important input or what we can call the "Creating point" of the Thai animations consist of 3 inputs: artist, budget and technology.

Artist: One animation would require tens of lives of people working in the team or even hundreds in some big companies. However, all the artist will need to follow the direction of the big artist called “The director”. Apart from directing the course for the movie, the director also has the power to make all decisive decision in many aspects of the production process. Therefore, this research had used the names of the directors as the representatives of the artist involved in the production of the 8 animations. There are altogether 10 following directors:

- Sudsakorn, Director, Payut, Ngaokrachang
- Khan Kluay, Director, Kompin, Kemgumnird
- The Life of Buddha, Director, Krismant Whattananarong
- Nak, Director, Natthaphong, Ratanachoksirikul
- Khan Kluay 2, Director, Taweelap, Srivuthivong
- Echo Planet, Director, Kompin, Kemgumnird
- Yak the Giant King, Director, Prapas
- Cholsaranont, Chaiporn, Panichrutiwong
- The Story of Mahajanaka, Director, Kriengkrai Supornsahusrungsri
- Thananchanok, Suban Na Ayudhaya, Nop, Dharmavanich

For these 10 artists, at one point graduated from Silpakorn Univeristy; Some from the Faculty of Architecture and some from technology study.

Budget: For the budget of each animation, the source of gross is different and can be categorized into 4 types: private investor, company investment, central budget and government’s funding according to the following detail:

Private investors: Are investors who use their own money to invest in the animations and also by depending on other source of funding more or less. There are 2 people that fits in this category:

- Pol. Gen. Banthoeng Kampanatsaenyakon, funded Sudsakorn
- Dr. Wallapa Pimthong, funded, The Life of Buddha

Company’s investments: Are the investment to the animations in the name of the companies: using the budget for all the production process. There are 2 companies that fit this category:

- Kantana Animation Studio Ltd., Funded, Khan Kluay, Khan Kluay
- 2 and Echo Planet, Sahamongkolfilm International Co., Ltd., Funded, Nak

Central budget: Is through fundraising from close friends. The money was then put in the central funding and used for the animation production. There is only one animation that fits into this category and that was: Yak the Giant King. The investors for this movie was the following: Workpoint Entertainment Company Ltd., Sahamongkol film International Co., Ltd, Superjeew Company Ltd., Chaiporn Panichrutiwong (personal money) and Prapas Cholsaranont (personal money).

Government’s funding: Was the funding provided by the government’s office to create the animation. The office was the office responsible for directly supporting creative work; That was the Software Industry Promotion Agency (Public Organization) or SIPA who support the budget for The Story of Mahajanaka.

Technology: As for all the 8 animations, the production tools used in producing the animations were all from other countries. Therefore, we can briefly conclude that Thai animations still mainly rely on the technology from other countries. However, from the information of this research, it was found that Payut Ngaokrachang was the only person who didn’t depend the technology from other countries but relying on the “Borrowed” method from different sources and endless experimenting until he found a way to make moving pictures. It can be concluded that the technology in animation production can be categorized into 2 categories: depending on foreign technology and the technology that Payut Ngaokrachang had adapted and created.

The three inputs which are artists, budget and technology are the important inputs that help create each of the animations. It was admirable that each animation was all from Thai artists which show that Thai people also have this kind of skill which can be competed with other countries. Apart from this, many investors in Thailand are also ready to invest in full-length animations even though it might require tens of millions of baht. But as for the technology in the production, Thailand still needs to depend on the software of from other countries in every production process.

Production process: After all the three inputs, the artists, the budget and the technology are all put together, the next stage is the production process. In all the 8 animations, the production process can be further categorized into 2 kinds: 2D animation production process and 3D animation production process.

2D animation production process: Is a traditional animation process which requires drawings of the



Fig. 1: The year when each of the animation were released chronologically

character frame by frame continuously in order to create the movement. There are only 2 animations that fit this category: Sudsakorn and The Life of Buddha. These 2 movies show only a few minor differences in the production process.

Sudsakorn was the only animations in Thailand that truly used the traditional animation process, drawing and coloring on the transparent sheets and record using a 35 mm film. The recording is done for each individual picture and put together to create moving pictures. This was the same technics that Walt Disney had been using. As for the life of Buddha, the drawings were also done for each fram but it was done on study and then scanned into digital files. The coloring part was done in the computer as well as other special technics and the post production.

3D animation production process: is a new technic for creating animations. The first animations that used this kind of technic were Toy story which came out in 1995. This production process is totally different from the traditional way. All the process happens in computers. Each still picture were then put together to create movements. Each of the frames came from the calculation of the background, surface and lighting which would be rendered into 1 still picture. Only the pre-production part wasn't done in commputers. All the other process fully relies on computer works. No pictures were hand drawn. There are 6 movies altogether that rely on this kind of production process: Khan Kluay, Nak, Khan Kluay 2, Echo Planet, Yak the Giant King and The Story of Mahajanaka.

Output: The combination of the 3 inputs: artists, budget and technology led to the production of animations. The production process also took years before the animations were completed at the end of the production process.

From "Sudsakorn", the first full-length animation that was aired in 1979 until "The Story of Mahajanaka" in 2014,

Table 1: The gross from the 8 animations

Movie	Year released	Estimated gross (Million baht)
Sudsakorn	1979	2
Khan Kluay	2006	79
Khan Kluay 2	2009	77
Echo Planet	2012	19
The Life of Buddha	2007	17
Nak	2008	10
Yak the giant	2012	58
The Story of Mahajanaka	2014	No admission fee no record of gross

Thailand has had altogether 8 following full-length movies Sudsakorn, Khan Kluay, The Life of Buddha, Nak, Khan Kluay 2, Echo Planet, Yak the Giant King and The Story of Mahajanaka. If all the animations were to be put on a timeline, it will produce the following timeline (Fig. 1).

Outcome: The production processes of each of the animation were difficult. After the production is done and the animations were already shown in theatre, one indicator that will determine the success is the gross. And here are the gross from the 8 animations (Table 1).

Apart form the gross, the awards are also another outcome from the production of each movie. This research will only consider the Thailand National Film Association Awards due to the fact that it's a national award which is the biggest and most formal award in Thailand. Here is the list for the animations:

Khan Kluay was nominated for 6 awards and won in 4 following awards: best music, best recording, bests screenplay and best picture. Khan Kluay 2 was monimated for 4 awards but didn't win in any award.

Echo Planet was nominated for 3 awards and won 2 following awards: best original song and best recording and sound mixing. The Life of Buddha was nominated for 3 awards and won 1 award: best original score.

Nak was nominated for 3 awards but didn't win any award. Yak the giant king was the movie with the most nominations similar to Khan Kluay which was 6 nominations but didn't win any award.

The Story of Nahajanaka wasn't nominated for any award due to the fact that it was produced in honor of King Rama IX and wasn't seeking any benefit or award.

In conclusion, if we look at the history of Thai animations through the system theory lens, since, the first animation, *Sudsakorn* in 1979, we would find that Thai animations was created through only 3 important inputs: artists, budget and technology. These three inputs were the most basic. If one of the inputs was missing an animation can never be made. When all the 3 inputs come together, it will lead to the production of animations. In Thailand, animations were only produced through 2 methods: 2 Dimension (2D) and 3 Dimensions (3D) which will result in the production of 2 types of animations, the 2D movies consist of *Sudsakorn* (1979) and *The Life of Buddha* (2007). There are 6 D animations: *Khan Kluay* (2006), *Nak* (2008), *Khan Kluay 2* (2009), *Echo Planet* (2012), *Yak the Giant King* (2012) and *The Story of Mahajanaka* (2014). The last part when looking through the lens of system theory is the outcome. This research focused on the gross and the awards. It was found that *Khan Kluay* was the most successful animation both in the gross and the awards. It made over 76 million baht and was nominated for 6 awards and won 4 awards which include big awards such as best screenplay and best picture of 2006.

RESULTS AND DISCUSSION

From the conclusion of the background of Thai animations, it was found that Thai animation was created on the 3 important bases: artists, budget and technology. Without any of the input, the animations will never be created. However, considering the year that each of the animation was released, it was found that *Sudsakorn* which was released in 1979, the other animations were all released post 2006. This phenomenon led to the question of why Thai animations were all created during this period of time. The basic answer to this question was that, since, animations were created based on these three inputs, we can fairly assume that all the three inputs, artists, budget and technology, all came together during that time. The next question would be what caused the three input to come together at that point in time in Thailand. This question forced us to think back to an important phenomenon in the history of animation when *Toy Story* was created in 1995.

This animation was the first 3D animation of the world created by Pixar Animation with John Lasseter as

the director. The released of toy story shook the world and totally changed the production process from the original process. Hand drawings and coloring was unnecessary. These new productions of 3D animations are all created by computers. Only the pre-production part required works on study. This kind of process helps safe time and man power which also mean that lower cost of production. Therefore, this 3D production process became popular in no time within Hollywood and spread throughout the world soon after. Certainly, Thailand was one of the countries that were influenced by this phenomenon.

The popularity of *Toy Story* pushed many company in America to bring in 3D animation production technology into their company and sold the technology in a form of softwares that people can buy and install in their computer. Production technology became a primary input with easy access. As for the artists who will be using these technologies, most of the Thai artists have had experience working in other countries, especially in America such as *Kompin Kemgumnird*, the Director of *Khan Kluay* who had prior working experience with Disney). He came back to work in Thailand, since, 1997. After these artists came back, they also brought back with them the knowledge and the experience that they required from their work abroad. After requiring the technology and the artists who will be using them, Thailand was in the ready stage, only lacking the budget. However, that wasn't the problem, since, this 3D animations requires less funding. Many investors in Thailand were willing to support. When the investors became involves, that completed the three basic inputs: artist, technology and budget. Thai animations were created from these three basic inputs. The first 3D animation was "Pang pond, the Animation" produced by *Vithita Animation Ltd.* and aired on channel 3 in 2002. Four years afterwards, "*Khan Kluay*", the first full-length 3D animations was released in Thailand in 2006. These three inputs, artists, budget and technology are not only the important inputs for Thailand. Even the first 3D animation mie such a "*Toy Story*" also needed to rely on these three inputs.

Pixar was started by a scientist named *Ed Catmull*. *Catmull* and his team at the University of Utah found a way to create surfaces with computer and were able create a 3D model. They were also able to fill in the surface and the lighting to the model. Since, he had always been a fan of walt disney animation since, he was a child, *Catmull* believed that the technology will be able to lead to creation of animations. *Catmull* and his team are

considered the first basic input to the creation of animations. Although, Catmull didn't have any artistic ability to create an animation, later on, he met John Lasseter.

John Lasseter was an animator from Walt Disney. He graduated from art school and had a dream to create animations using 3D production. The combination of Catmull's technology and his creativity, he was able to successfully create a 3D animation. "The Adventure of Aally and Andre B." was the first 3D animation by John Lasseter. This animation was created since, Catmull's team was still working with George Lucas; and later on with Steve Jobs and changed the name to Pixar. John created another short animation called "Luxo Jr." This animation was very successful and was nominated for best short animation at the Oscar. This success was what made Pixar a Pixar company that everyone knows of Thechaisrisutti.

The beauty and the content of Luxo Jr. made Walt Disney lower his pride and asked John Lasseter to come back and work for Disney again. John refused but instead asked Disney to come and invest with Pixar. After all the detail was discussed, Walt Disney decided to work with Pixar by providing the budget to create 3 full-length animations with Pixar handling all the production. This was the point when the first full-length animation was created with John Lasseter as the representative of the creative artist and Ed Catmull's team a representative for the production technology and Walt Disney as the representative for the budgeting.

CONCLUSION

The creation of Toy Story showed that the technology in the hands of Catmull would not be much of a benefit if there were no creative tools and the ability to produce and increase the ability by many folds. When the budget input was added as the last input that completed the 3 required inputs which are artists, budget and technology which are the same input as the creation of the Thai animation.

REFERENCES

- Piampiyachart, O., 2012. *Yak the Giant Prapas Cholsaranont*. Fullstop Publisher, Bangkok, Thailand, (In Thai).
- Suttana, U., 2006. *Life, Prayut, Animation*. SUM System Co. Ltd., Bangkok, Thailand, (In Thai).
- Thasanabanchong, C., 2002a. *Thai film standard development*. Ph.D Thesis, Chulalongkorn University, Bangkok, Thailand.
- Thasanabanchong, C., 2002b. *The creation of Thai movie standard*. Ph.D Thesis, Faculty of Communication Arts, Chulalongkorn University, Bangkok, Thailand. (In Thai)
- Williams, R., 2002. *The Animator's Survival Kit: A Manual of Methods, Principles and Formulas for Classical, Computer, Games, Stop Motion and Internet Animators*. 2nd Edn., Faber and Faber, Singapore, ISBN:9780571212682, Pages: 352.

Analysis of Motor Speed Three Phase Induction Using Harriot and Smith Method

¹Ahmad Faizal, ¹Harman, ²Rado Yendra and ³Ahmad Fudholi

¹Department of Electrical Engineering,

²Department of Mathematics, Faculty of Science and Technology,

Universitas Islam Negeri Sultan Syarif Kasim (UIN Suska), 28293 Pekanbaru, Riau, Indonesia

³Solar Energy Research Institute, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Abstract: Induction motor is the driving tool needed by the industry but has the drawback that is difficult speed control for induction motor analysis many experimental methods that can be used to describe the plant that resulted in the control engineer must choose the right model at the time of open loop, hence, need to be analyzed on when close loop with hybrid SMC and PID. Based on the simulation shows the Harriot and Smith method when given the controller does not have a significant difference when reaching the desired setpoint, the Harriot Model has a faster settling time of 2.864 sec and the Smith Model has the ability to overcome the more robust interference that is overshoot in the first area of 0.02%, the second 0.03% and third 0.07% without steady state error. In this case on the control with the Harriot and Smith Models depending on the design of the controller.

Key words: Induction motors, rotational speed (RPM), SMC hybrid, PID, Harriot and Smith Models, design

INTRODUCTION

Induction motor is the most motor found in industry when compared with other motor. This is due to the advantages of induction motor. Among its supremacy is its simple construction part, its mechanical robustness, its relatively cheap price and does not require complex maintenance and can also be directly connected to the Alternating Current (AC) resources (Utoro and Pramudijanto, 2014).

But in addition to the advantages of induction motor, not apart from the weaknesses it has. Among the disadvantages is the difficulty of adjusting the speed of this motor because the characteristics of the motor Induction itself is not linear, consequently unable to maintain the speed when the load increases. In order to get a good performance and in accordance with the setpoint provided on the motor induction when performing the operation required a controller that is able to make the motor induction becomes more robust against the disruption given (Nofendra, 2009).

To analyze the behavior of the system is done by simulation but the plant must be changed into mathematical form, one way to get the transfer function is by experimental method. Many forms of experimental methods are used to describe actual system behavior but

the weakness of the experimental method has different accuracy levels for each method in an open loop situation, so, it is necessary to test and get the exact method that works with the actual plant. In this case each method must have an equivalent capability in describing a plant therefore it is necessary to analyze on each method used by providing the right controller and a strong controller or able to follow the plant and able to overcome Interference at the time of controlling the speed of three phase induction motor.

Although, the method for describing the plant is different at the time of the open loop but needs to be done close loop analysis by using a sturdy controller. One of the controllers that is able to overcome the parameter uncertainty is Sliding Mode Control (SMC). This controller has the advantage that is strong and able to work on non-linear system that is a system that has uncertainty model or parameter (Nofendra, 2009).

The SMC controller has a disadvantage that lies in uninterrupted chattering that can lead to overshoots, one controlling handler capable of eliminating overshoot is a PID controller. In this research the method used latzel and Smith method with hybrid controller SMC-PID is expected to overcome the experimental method and the disturbance given.

MATERIALS AND METHODS

In this research, there are several stages which writer do in experimental method analysis at speed of induction motor by using hybrid SMC and PID controller which is done using MATLAB R2013a application.

Three fasa induction motor modeling: This induced induction pad is tested in an open loop and then the result of the plant response read by the rotary encoder is displayed on the computer through the data acquisition process with Mitsubishi PLC (Utoro and Pramudijanto, 2014). After that one sample was applied to an experimental model based on an international research paper written by Jean-Jacques and Li (1991). The results of identification as follows:

The Harriot method (Utoro and Pramudijanto, 2014): In Fig. 1, known input signal $X_{ss} = 1000$ rpm and output response $Y_{ss} = 999.366548$, so that, the value obtained $K = Y_{ss}/X_{ss} = 0.991366548$. Searchable value τ_{dt} , $\tau_{H1} + \tau_{H2}$ on equals:

$$\tau_{H1} + \tau_{H2} = \frac{t_{73}}{1.3} \tag{1}$$

Based on Fig. 2-5, we get t_{73} value. By trial and error method is 4.8136998s:

$$\tau_{H1} + \tau_{H2} = \frac{4.8136998s}{1.3} = 3.702846s \tag{2}$$

Substituted into the equation and obtained:

$$t_i = \frac{3.702846s}{2} = 1.851423 \tag{3}$$

So, we get the value of y_i from t_i in Fig. 1 using trial and error method that is $y_i = 290$ and can be determined $y_i/y_{ss} = 0.292$. Furthermore, from Harriot curve got the value of the equation $\tau_{H1}/\tau_{H1} + \tau_{H2}$ is 0.7199999. And substituted the value of $\tau_{H1}/\tau_{H1} + \tau_{H2}$, so that, the value obtained τ_{H1} and τ_{H2} , so that, the transfer function is obtained:

$$G_H(S) = \frac{0.991366548}{2.76415201s^2 + 3.702846s + 1} \tag{4}$$

Smith's method (Utoro and Pramudijanto, 2014): Given input signal $X_{ss} = 1000$ rpm and output response $Y_{ss} = 999.366548$. So, we get the value of $K = Y_{ss}/X_{ss} = 0.991366548$ which will be included in Eq. 14. Figure 2 obtained:

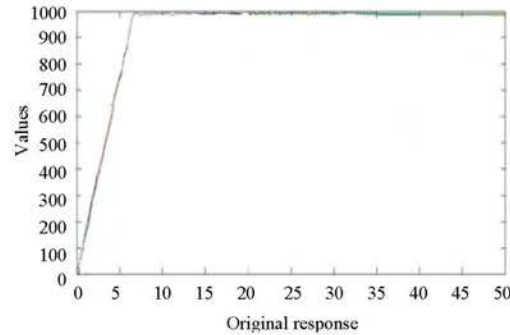


Fig. 1: Step response of three phase induction motor at time t_i and t_{73} (Utoro and Pramudijanto, 2014)

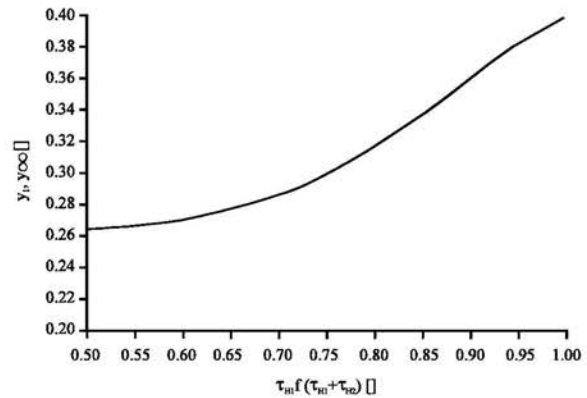


Fig. 2: Harriot curve (Ali, 2004)

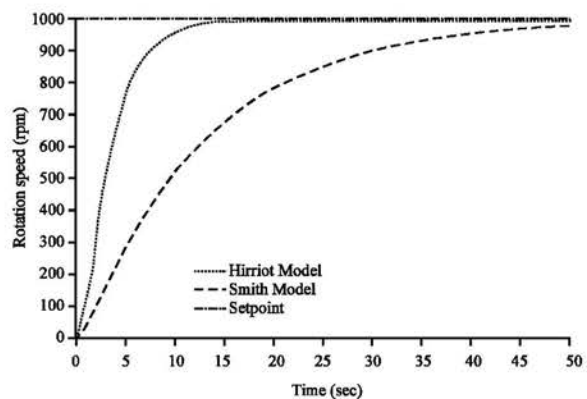


Fig. 3: Response identifies the model approach with Harriot and Smith methods

$$\frac{t_{20}}{t_{60}} = \frac{1.23}{4.22} = 0.291s \tag{5}$$

In Fig. 6, the Smith curve 1, $0.291 = 1.94 \zeta$ and got the value. And in Fig. 7, the Smith curve 1 is obtained: $t_{60}/\tau = 3.14$, sehingga $\tau = 4.22/1.23 = 4.43$:

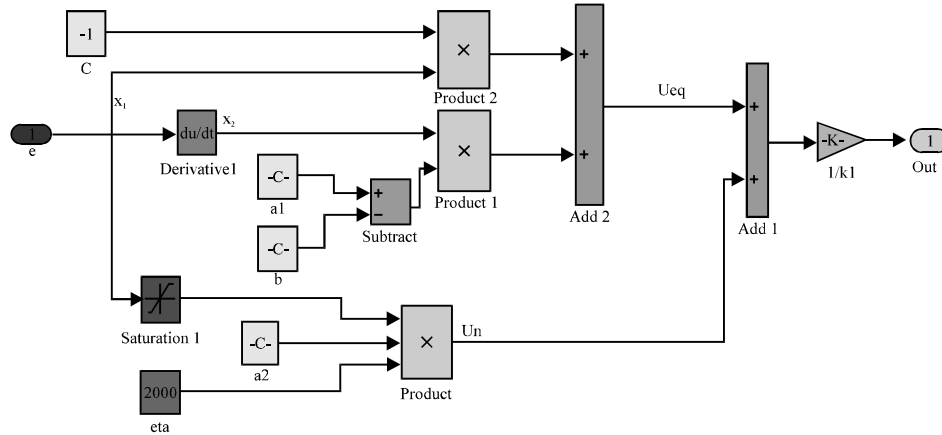


Fig. 4: SMC control block

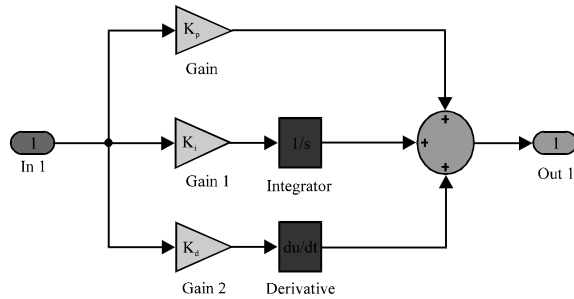


Fig. 5: PID control block

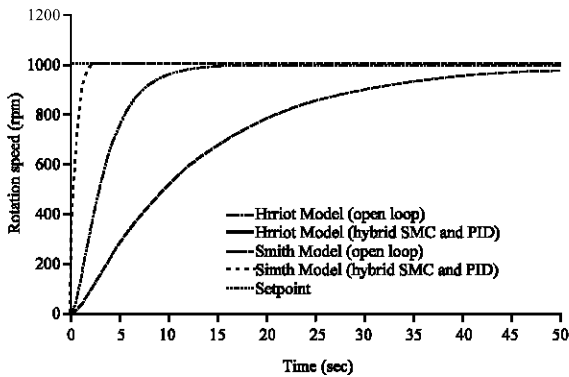


Fig. 6: Comparison of methods in achieving the desired setpoint

$$\tau_{SM1} = \tau\zeta - \tau\sqrt{(\zeta^2 + 1)}$$

$$\tau_{SM1} = (3.43 \times 1.94) + 3.34\sqrt{(1.94^2 + 1)} = 6.6542 + 5.70 = 12.3542s \quad (6)$$

$$\tau_{SM2} = \tau\zeta - \tau\sqrt{(\zeta^2 + 1)}$$

$$\tau_{SM2} = (3.43 \times 1.94) + 3.34\sqrt{(1.94^2 + 1)} = 6.6542 + 5.70 = 12.3542s \quad (7)$$

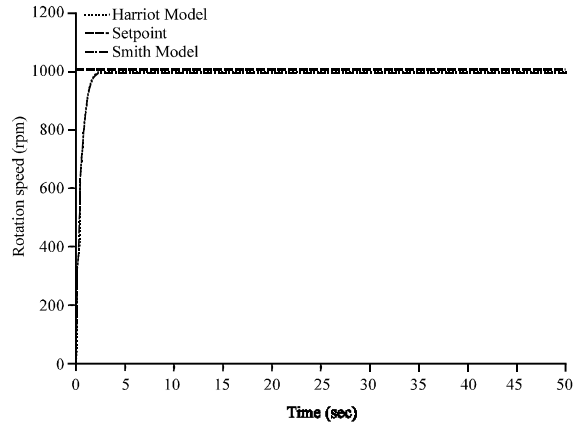


Fig. 7: Comparison of models in overcoming control signal interference

Table 2: Comparison of model approach method

Methods	Mathematical model	ISE
Harriot	0.991366548/(2.7641s+3.7028s+1)	2.253
Smith	0.991366548/(11.7325s ² +13.3s+1)	7.194

So, the following transfer function is obtained:

$$G_{SM}(S) = \frac{0.991366548}{(11.7325s^2 + 13.3s + 1)} \quad (8)$$

Table 1 got the identification response model in Fig. 3.

Design of SMC controller for SMC controllers: Transfer function at minimal load will be exemplified as fixed variable to make it easier in designing the controller, so that, the function of conductor of plant is determined by the following equation:

$$G(S) = \frac{Y(s)}{U(s)} = \frac{K}{as^2 + bs + 1} \quad (9)$$

Where:

$$\begin{aligned} k &= 0.991367 \\ a &= 2.644152 \\ b &= 3.702846 \\ c &= 1 \end{aligned}$$

The transfer function can be presented in the form of a differential equation, assuming the initial value is 0:

$$a\ddot{y}+b\dot{y}+cy = Ku \rightarrow a\ddot{y} = -b\dot{y}-cy+Ku$$

Defined:

$$x_1 = y \rightarrow \dot{x}_1 = \dot{y} = x_2 \quad (10)$$

$$x_2 = \dot{y} \rightarrow \dot{x}_2 = \ddot{y} \quad (11)$$

and found:

$$\ddot{y} = -\frac{b}{a}x_2 - \frac{c}{a}x_1 + \frac{k}{a}u \quad (12)$$

Then the error signal is taken as state variable:

$$x_1 = e \rightarrow \dot{x}_1 = \dot{e} = x_2 \quad (13)$$

$$x_2 = \dot{x}_1 = \dot{e} \quad (14)$$

The equation on the error signal is expressed:

$$e = r-y \quad (15)$$

And got:

$$\dot{x}_1 = \dot{e} = r-\dot{y} \rightarrow \dot{x}_1 = r-x_2 \quad (16)$$

Because the system is a regulator then obtained:

$$\begin{aligned} \dot{x}_1 &= x_2 = r-\dot{y} \rightarrow \dot{y} = x_2 \\ \dot{x}_2 &= \ddot{y} = r-\ddot{y} \rightarrow \dot{y} = \dot{x}_2 \end{aligned} \quad (17)$$

Substitute Eq. 17 and 18 in Eq. 13, so that:

$$\dot{x}_2 = -\frac{b}{a}x_2 - \frac{c}{a}(r-x_1) + \frac{k}{a}u \quad (18)$$

$$\dot{x}_1 = -\frac{b}{a}x_2 - \frac{c}{a}x_1 + \frac{k}{a}u \quad (19)$$

So, we get satate-space equation:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{c}{a} & -\frac{b}{a} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k}{a} \end{bmatrix} \quad (20)$$

Defined a glide surface:

$$\begin{aligned} \dot{\sigma}_s &= 0 \\ S(\dot{x}_1 + \dot{x}_2) &= 0 \end{aligned}$$

Then, we can find equivalence control signal with assumption $U_n = 0$:

$$\begin{aligned} x_2 + \left(-\frac{b}{a}x_2 - \frac{c}{a}x_1 + \frac{k}{a}U \right) &= 0 \\ -\frac{c}{a}x_1 + \left(1 - \frac{b}{a} \right)x_2 - \frac{k}{a}U_{eq} &= 0 \quad (21) \\ U_{eq} &= \frac{-cx_1 + (a-b)x_2}{k} \end{aligned}$$

Having found the equivalent control signal can be searched natural control signal:

$$\begin{aligned} \dot{\sigma}_s &= x_2 + \left(-\frac{b}{a}x_2 - \frac{c}{a}x_1 + \frac{k}{a}U \right) \\ \dot{\sigma}_s &= x_2 + \left(-\frac{b}{a}x_2 - \frac{c}{a}x_1 + \frac{k}{a}(U_{eq} + U_n) \right) \quad (22) \\ \dot{\sigma}_s &= -\frac{k}{a}U_n \end{aligned}$$

Based on the stability requirements lyapunov in Eq. 19 then selected:

$$\begin{aligned} \dot{\sigma}_s &= -\frac{k}{a}U_n \\ -\eta \cdot \text{sign}(\sigma) &= -\frac{k}{a}U_n \quad (23) \\ U_n &= \frac{a}{k} \eta \cdot \text{sat}(\sigma) \end{aligned}$$

where, $\eta > 0$ (constant positive). Thus, Eq. 18 and 20 are converted to block simulink as follows. In Fig. 4, it is the result of modeling design on SMC control using Simulink MATLAB R2013a.

Design of PID controller: Once designed the SMC controller is then designed PID controller which will be combined. Based on the general form of PID controller in the following Eq. 5:

$$u(t) = K_p e(t) + K_i \int_0^t e(t) dt + K_d \frac{de}{dt} \quad (24)$$

Then the block diagram of the PID controller is shown in Fig. 5. In Fig. 5, it is a common form of modeling in PID control that is in Eq. 21 which is converted into blocks using Simulink MATLAB R2013a.

RESULTS AND DISCUSSION

Results and analyzes were based on response time (Ogata, 1996).

Comparison of models in achieving setpoints using hybrid SMC and PID: After getting the model response in the state of the open loop then do the control in a close loop on Harriot and Smith Models. On the issue of tuning control parameters are given with different values on each model. The response result with hybrid controller SMC and PID is shown in Fig. 6.

In Fig. 6, the response analysis shows the Harriot and Smith models are able to reach the desired setpoint quickly and have no significant difference when reaching the setpoint in a close loop state or when given a hybrid controller SMC and PID with the timing setting of Harriot 2.864 sec and steady state error 0% while on the Smith Model has a set time of 2.89 sec and a steady state error 0%.

Comparison of models with interference: After Harriot and Smith Models were tested using SMC and PID hybrid controllers, each model was then given a disturbance of control signals in the first region at 13-15 sec with one reinforcement, the second at 20-22 sec with twice the reinforcement and a third area at 35-37 sec with three reinforcements. The simulation results shown in Fig. 7.

Based on Fig. 7, when disturbed in the first and second areas the Harriot Model has an overshoot on the first region of 0.11%, both 0.34% and the third 0.56% while in the Smith Model having overshoot on the first 0.02%, the second 0.03% and the third 0.07% without steady state error.

CONCLUSION

Based on time response analysis obtained: in the simulation results show the Harriot and Smith methods when given the hybrid SMC and PID controllers have no significant difference when reaching the desired setpoint. At the time of simulation Harriot Model have faster time that is time of 2.864 sec setting and 0% steady state error but Smith Model is more solid in overcoming overshoot interference at first area 0.02%, second 0.03% and third 0.07% without steady error state.

Overall experimental method with Harriot and Smith Model can be used in describing a plant well in controlling the speed of the induction motor, in this case on the control of Harriot and Smith's three phase induction motors depends on the design of the controller.

REFERENCES

- Ali, M., 2004. Learning PID control system design with MATLAB software. *J. Educ.*, 1: 1-8.
- Jean-Jacques, E.S. and W. Li, 1991. *Applied Nonlinear Control*. Prentice Hall, Upper Saddle River, New Jersey, USA., ISBN:9780130408907, Pages: 459.
- Nofendra, R., 2009. Design of slide mode controller for induction motor. Master Thesis, Andalas University of Padang, Padang, Indonesia.
- Ogata, K., 1996. *Automatic Control Technique*. 2nd Edn., Erland, Jakarta, Indonesia,.
- Utoro, M.R. and J. Pramudijanto, 2014. [Design and implementation of sliding mode controller on three-phase induction motor speedsettings (In Indonesian)]. *J. Eng. ITS.*, 3: F84-F89.

The Effect of Leadership Style on Job Satisfaction: A Case Study Analysis on KIMA Company in Makassar Industrial Area, Indonesia

¹Salmiaty Taty, ¹Ahmad Musseng, ²Eka Suhartini and ³Bayu Taufiq Possumah
¹STIE YPUP Makassar, Makassar, Indonesia
²UIN Alauddin Makassar, Makassar, Indonesia
³Universiti Malaysia Terengganu, Terengganu, Malaysia

Abstract: This study aims to explain the influence of leadership on employee job satisfaction. The analysis was performed using Structural Equation Modeling (SEM) with alpha 5% to test the significance of research variables. Data obtained from permanent employees of PT KIMA are determined as samples based on cluster random sampling technique. Data collection also uses questionnaires (as the main instrument), interviews, observations and documentation. The result of analysis (at the 5% level of significance) shows that leadership has a significant effect on job satisfaction in PT KIMA Sul sel.

Key words: Leadership style, alpha, job satisfaction, human resources development, data obtained, questionnaires

INTRODUCTION

Growth and development of a work organization cannot be separated from the leadership of the organization in maintaining the satisfaction and performance of employees to remain loyal. The aspects of leadership, satisfaction and performance are understood by the management that the three are closely related in determining the achievement of organizational goals. According to Alimuddin (2002), Miftah (2007) and Hadari and Martini (2004) an organization needs effective leaders who have the ability to influence the performance of subordinates to achieve organizational goals. The existence of influence between leadership and job satisfaction is evidenced by Seno *et al.* (2011), Fatima *et al.* (2011) and Ghafoor *et al.* (2011) when their research proving that leadership has a strong influence on job satisfaction. Meanwhile, the influence of leadership on employee performance, Totok (2013) and Gary (2007) states that if the leader is able to apply the right leadership, the employee then will satisfied and ultimately affect its performance goes better.

Leadership style is the norm of behavior used by a person when he trying to influence the behavior of others (Suranta, 2002). Leadership style is a leader who able to improve employee performance and appropriate to the situation and condition of the company, flexible to adapt the environment of his subordinates. Therefore, when the

employees satisfied with the leadership style of the company where they work, the employee's performance will increase. The attitude and leadership behavior are very influence on employee performance in the company. Hence, the purpose of this study is to test and analyze the influence of leadership variables on employee's job satisfaction of KIMA Company in Sul-sel. The results of this research are expected to give consideration for the company's management in its policy, decision and business development program.

Theoretical framework

Job satisfaction: Job satisfaction is an individual orientation that affects the role in work. Job satisfaction is the judgment, feelings or attitudes of a person or employee towards his work and relationships with the work environment, type of work, compensation, social relations at work and others. Job satisfaction is the fulfillment of some wants and needs through work or work activities. Each individual has a different level of job satisfaction in accordance with the value system that applies to them. Job satisfaction is a pleasant emotional attitude and job devotion. There are five determinants of job satisfaction or called Job Descriptive Index (JDI) according to Luthans and Spector in Robbins, such as the job itself, salary, promotion opportunity, supervisor and co-workers. While according to Mangkunegara (2000) some of the variables

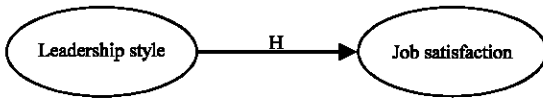


Fig. 1: The structural equation model

related to job satisfaction are: turnover, absenteeism (absenteeism) level, age, job level and organization size.

Leadership style: Leadership is a process of influencing others to achieve certain goals (Hughes *et al.*, 2002). In the business framework a manager is a leader when he able to influence subordinates, co-workers or even their bosses to manage their efforts toward the achievement of organizational goals. The process of motivating, influencing and directing subordinates is not an easy task for leaders. Not easy, since, each subordinate has different characteristics. Thus, in order a leader can be easy to motivate, influence and direct subordinates, the leaders in applying leadership style must first understand who subordinates, understand the strengths and weaknesses of subordinates and understand how to harness the power of subordinate to compensate for the their weakness. The right leadership style will cause a person's motivation to perform. According to Hadari (2008), the style of leadership defined as the behavior or manner chosen and used by the leader in influencing the thoughts, feelings, attitudes and behavior of members or subordinates of the organization. Refer to the situation leadership style according to Sondang (2002) there are five types: the autocratic leader, militaristic type, paternalistic leader, a charismatic leader and the democratic leaders. While according to Stephen (2008) the leadership styles are charismatic style, transactional style, transformational style and visionary style. To achieve company goals a leader must have the ability to motivate, influence and direct his employees. The leader must have a leadership style that is appropriate to the situation and condition of the company in managing its employees and will encourage employees to work more eager in carrying out duties and obligations. The better the leadership style of a leader, the better employee performance will increase. Hence, the theoretical framework of this research is shown in Fig. 1.

MATERIALS AND METHODS

Data analysis method: This study aims to examine the influence of leadership on job satisfaction by using a quantitative approach. The data needed to perform

hypothesis testing analysis is obtained from the field survey at the research site. The data collected by using questionnaires and interview instruments to the respondents. Respondent perception data then assessed by using Likert scale, therefore, the perception can be analyzed quantitatively. The technical data analysis using Structural Equation Model (SEM) which will be connected between exogenous and endogenous variables. To ensure the model is made correctly and the results are valid, a testing process is required. The structural equation model for exogenous and endogeneous seen as in Fig. 2.

Validity and reliability tests are also performed to ensure that each question is classified on each of the predefined variables. Questionnaires that have been compiled, tested by giving a question to a group of respondents with the aim to know the extent to which the measuring instrument has the validity and reliability. Valid and reliable instruments are the main requirement to obtain valid and reliable results. The validity of a measurement scale can be defined as "the extent to which the difference between the observed scale scores shows the actual difference between the object/respondent on the measured characteristics and not because of a systematic or random error". The validity of an item can be determined by comparing the correlation index of product moment with its probability value if the items are not correlated significantly at 5% significant level, then the instrument is declared failed. To test the validity, then correlation formula of product moment is used.

Leadership has a direct significant positive effect on job satisfaction with $p = 0.000 < 0.10$ with coefficient value of 0.359. This positive coefficient sign indicates that the existence of good leadership will make employees satisfied with his work. Employees who work in the Industrial Area of Makassar Sul-sel feel fit with the leadership because the leaders have good communication and provide instructions in accordance with what is desired by the leader. The employees feel happy to follow the wishes of leadership. Since, the leaders continually provide improvement suggestions in wise manner, the employees feel satisfied over the leadership behavior. Instead leaders feel happy to employees because they remain obedient to the wishes of leadership. The value of regression coefficients between leadership style variables and job satisfaction as endogenous variables gives the highest result compared to other relationships, reinforces the finding that leadership style in KIMA company has succeeded in creating job satisfaction although the type of leadership style applied is based on the result orientation style.

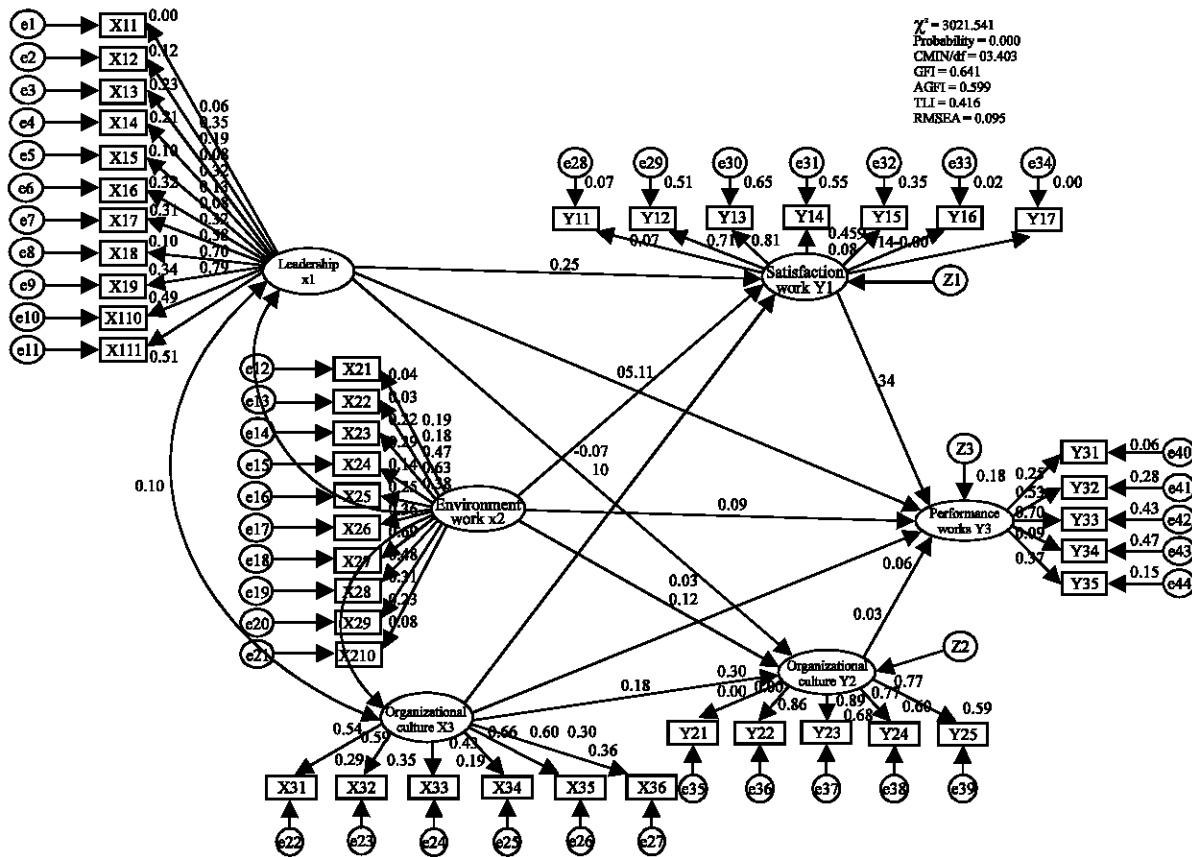


Fig. 2: The measuring instrument

RESULTS AND DISCUSSION

The result of SEM analysis presented in Table 1 above also produces coefficient of direct influence of leadership variable on employee work satisfaction equal to 0.359. The coefficient of direct influence is marked positive which means that the applied leadership model will have a positive effect on employee job satisfaction. To answer the problem formulation and hypothesis can be observed from the results of path goal analysis in Table 1. The table shows that leadership (X1) has a positive and significant effect on job satisfaction (Y1). This result is proved by the value of t count (critical ratio) is less than the table (4.700>1.96) and the probability value or p = 0.000 value is smaller than 0.10. The results of this study reinforced by the theory of Miller *et al.* showed that leadership style has a positive relationship to employee job satisfaction. The dominant indicators in leadership style variables are directive behaviors with leader statement items always provide specific guidance to worker to coordinate work, instruct worker to set time and to follow the rules in their work procedures. The dominant indicator in job satisfaction is the opportunity to move forward with the statement item, the company will give promotion to the successful employees.

Table 1: Loading factor correlation between job satisfaction and leadership

Correlation between variables	Est	SE	CR	p-value	Ket
Job satisfaction_Y1 <- Leadership_X1	0.370	0.079	4.700	***	Sig.

Job satisfaction of KIMA employees is achieved because the leaders have competence in performing their leadership duties, have the ability to make good planning and decision making and consequently in applying the applicable work regulations indicated by giving sanction for employees who violate the discipline/regulation applicable. Conditions based on personal qualities and administrative actions by the employer lead to employee job satisfaction and employees feel that supervision and assessment of employees has been done fairly, assignments by the leadership have been made on the basis of the competence of each employee as well as the policies and regulations Enacted firmly and fairly by the leadership of KIMA company. Conditions in which the leadership emphasizes the importance of togetherness in completing the task and always create a harmonious relationship with its employees also has created employee job satisfaction on team work management in the working group of KIMA company. These findings support the results of research by Gary (2005) and Pierce and

Newstrom (2006) which suggest that the behavior of a leader can influence satisfaction from employees and support the leadership theory which can generate job satisfaction outcomes and also support the research results by Challagalla and Shervani (2006) which states that leadership has a significant effect on employee job satisfaction.

These findings also reinforce the research finding by Gruenberg which shows that the close and mutual relationship between colleagues and supervisors is very important and have a strong correlation with job satisfaction and type of work. One of the factors causing job dissatisfaction is the nature of the supervisor who is unwilling to hear the workers complaints and not willing to help if necessary. This is evidenced by Blakely in which workers who receive awards from supervisors are more satisfied than those assessed by themselves, but overly restrictive supervision will lead to lower satisfaction levels.

Basically, the finding of this study is accordance with the theory forwarded by Siagian, Miftah (2007) and Robbins. Moreover, Stonner and Freeman unveiled a leadership approach as developed by Hersey and Blanchard which is outlining how leaders should adjust their leadership style in response to a desire to succeed in their work, the experience, abilities and willing of employees.

From descriptive statistics side the analysis also indicate that the role of leadership participation style is still a dominant indicator to form leadership variables as seen from high mean values compared with other indicators. This indication proves that the role of leader who can be a role model is expected to improve work performance because the employee will always supports the policy of leader and this will be better because it is supported by situation where employees are involved and feel appreciated, until arises sense of ownership where employees willing to work with sincerity and maximum.

Based on the results of interviews on some employees and the facts in the area of research indicate that employees who work in the Industrial Area of Makassar (KIMA company) feel that with listening leadership style, rely on improvements in performance and leadership that always gives trust to his employee is a type of leadership that can bring job satisfaction employees. This fact confirms that although, theoretically, Path Goal's leadership theory which developed by Evan and House is believed by the researcher to be fully applicable to Makassar Industrial Zone, it is only the principle of achievement oriented that is adopted perfectly in this study. Empirical facts in this study that employees do not like dictated by managers because they

feel that leadership participation model is more effective and employees are eager to be involved in decision making. With achievement orientation leadership model, employees will feel more happy and can lead to the realization of job satisfaction.

The findings of this study are as well as in line with the research conducted by Mine Sancar which states that leadership behaviors of school principals in relation to teacher job satisfaction in North Cyprus, procedia social behavioral sciences (leadership behavior toward the principle of school relations with teacher's satisfaction in North Cyprus). Factors studied were school principals, leadership, behaviors, consideration, initiation of structures and job satisfaction with the number of respondents 599 from population amounted to 2200 people. The method used is multiple regression and the most influential variable is job satisfaction.

The differences of research findings occurred due to several factors including the different research area, where Brahmasari and Suprayetno conduct research on PT.Pei Hai Internanational Wiratama Indonesia with a population of 1737 employees and 325 samples of people with various and different the character of respondents and indicator variables. The same leadership style indicator was used, where the indicator of satisfaction variables is the compensation, working conditions, administrative system and policy in opportunity to develop. In other side, this research is done on manufacturing business in Makassar Industrial area with the population of 10.319 and 450 respondents. The indicators of leadership style used the same but the indicators of satisfaction used are the job description, job actualization, supervision, opportunity to advance level, salary and compensation

CONCLUSION

The finding of this research shows that leadership style has positive and significant effect to employee job satisfaction, means that the result of the implementation of managerial activity of leadership is positive in organization because the higher implementation of leadership managerial activity will affect the continuous improvement of employee performance. The development theory model as reflected in this research is the answer to the problem formulation. Leadership and human resource development are relevant to the issues on how to improve job satisfaction and employee performance. Thus, the relationship of influence between research variables is a simplification model of the complexity of phenomenon of job satisfaction and employee performance. Hence, this research model can be regarded as one of problem solving models in a management or human resources research.

REFERENCES

- Alimuddin, 2002. [Effect of leadership style on performance of employees of regional supervisory board of Makassar city]. Master Thesis, Gadjah Mada University, Yogyakarta, Indonesia. (In Indonesian)
- Challagalla, N.G. and A.T. Shervani, 2006. Dimensional and type of supervisory control: Effect on salesperson performance and satisfaction. *J. Marketing*, 60: 89-105.
- Fatima, B., U. Ahmad and N. Asvir, 2011. Effect of transformational leadership on employees' job satisfaction and organizational commitment in banking sector of Lahore (Pakistan). *Intl. J. Bus. Social Sci.*, 2: 261-267.
- Gary, Y., 2007. [Leadership in Organizations]. 5th Edn., PT. Gramedia, Jakarta, Indonesia, (In Indonesian).
- Ghafoor, A., T.M. Qureshi, M.A. Khan and S.T. Hijazi, 2011. Transformational leadership, employee engagement and performance: Mediating effect of psychological ownership. *Afr. J. Bus. Manage.*, 5: 7391-7403.
- Hadari, N. and H.M. Martini, 2004. [Effective Leadership]. Universitas Gadjah Mada, Yogyakarta, Indonesia, (In Indonesian).
- Hadari, N., 2008. [Human Resource Planning for a Competitive Profit Organization]. UGM Press, Yogyakarta, Indonesia, (In Indonesian).
- Hughes, R.L., R.C. Ginnett and G.J. Curphy, 2002. *Leadership: Enhancing the Lessons of Experience*. McGraw-Hill, New York, USA., ISBN:9780071123600, Pages: 604.
- Mangkunegara, 2000. [Human Resource Management Company]. PT. Ghalia Indonesia Printing, Bogor, Indonesia, (In Indonesian).
- Miftah, T., 2007. [Organizational Performance: Basic Concepts and their Applications]. Rajawali Press, Jakarta, Indonesia, (In Indonesian).
- Pierce, J.L. and J.W. Newstrom, 2006. *Leaders and the Leadership Process: Readings, Self Assessments and Applications*. McGraw-Hill, New York, USA.,
- Seno, A., A.T. Eka, Idrus and Djumahir, 2011. The influence of leadership style, organizational culture, motivation, program of training on performance and job satisfaction of employees at PT. Telkom TBK Pekanbaru. *J. Appl. Manage.*, 9: 1-1.
- Sondang, P.S., 2002. [Human Resource Management]. Bumi Aksara, Jakarta, Indonesia, (In Indonesian).
- Stephen, P.R., 2008. [Organizational Behavior]. 10th Edn., PT Indeks Publisher, Klaten, Indonesia, (In Indonesian).
- Suranta, S., 2002. [The impact of employee motivation on the relation between leadership style with employee performance of business companies]. *Empirical*, 15: 116-138.
- Totok, S., 2013. [Effect of leadership style on employee performance at KPRI Bima Jaya Pasuruan (In Indonesian)]. *J. Appl. Manage.*, Vol. 7.

Intelligent Parking System using Circle Hough Transform

¹Ajay Kamath and ²R. Swarnalatha

¹Department of EIE,

²Department of EEE, BITS-Pilani, Dubai Campus, Dubai, UAE

Abstract: Image processing is the science of extracting vital information from an image like edges, color distribution, etc. This information can be used as for analysis and also as a feedback to control many processes. This study gives an insight of how image processing can be used to control crowding of vehicles in a parking lot and compares this system to current smart systems already in place in terms of efficiency and feasibility. A new method to achieve the aforementioned goals involves a slight alteration to the appearance of parking spaces, thus, streamlining the algorithm for a faster detection without a heavy load on computing power. The different algorithms that are the focal point for this study are: canny edge detection and circle hough transform.

Key words: Image processing, manipulation, hough transform, circle, parking system, computing

INTRODUCTION

Let us begin with comparing the application at hand as an analogy of a control system with a feedback. This study starts with this basic idea as a foundation and then builds up on this idea to solve a practical challenge. (Davies, 2000).

The basic elements of process control involves a desired set point, a controller to correct errors in result and a feedback loop that provides the current output to the controller. The feedback is for comparison of the current output with the desired set point. The controller then gives the control signal to an actuator, so as to influence the output and correct the error (Davies, 2000; Arora *et al.*, 2008).

Image processing is the science of manipulation of digital pictures and analysis of these digital pictures to obtain numeric results for further computations. Image processing is a field with vast number of applications ranging from artificial intelligence, robotics to security systems.

A feedback loop comprises of a sensor to sense the variable in question and a transducer converts it to a usable signal which is understandable by the controller. These sensors are usually designed with a specific goal in mind and often their use cannot be expanded to other purposes. This is where image processing comes into the picture. Modern day digital cameras are equipments designed to capture digital images which are an array of numeric data. This data can be processed to provide information for vital feedbacks. Here, cameras are acting as a sensor for the required feedback with the added advantage of expansion to other purposes.

Literature review

Data and information from images: An image initially begins as a capture of an array of data from the environment. Digital images are stored as a two dimensional array of data. Each element of the array is called pixel. Displaying these pixels in the intended two dimensional order is what gives us a final image we can see and understand. The size of the aforementioned array is what gives the size of the image. For example, an image with a resolution of 1920×1080 is basically a two dimensional array of pixels with 1920 rows and 1080 columns (Lew, 2001).

Now this image that has been captured, doesn't have any information that the computer can understand. Thus, this image undergoes image processing and then the algorithms return values that can be used by the computer for further calculations. These values returned that can be used is termed as the information extracted from the array of data (Lew, 2001). This is the basic theory that comes in handy for cameras being used for artificial intelligence and automation.

Edge detection algorithms: Edge detection is a major section of innovation in the field of image processing. As the name suggests, edge detection involves detection of the boundaries in an image, so as to distinguish and analyze the objects that can be seen in an image. These algorithms work on the principle of marking discontinuities in the colors and output levels of the image while the algorithm does an overall scan of the image (Kaur and Kaur, 2016).

Edge detection is widely used to feed information extracted from processed images to neural networks

engines to help develop artificial intelligence and to develop an artificial understanding of the world around us to computers. It works by marking the areas of discontinuities of in brightness and colors in the image.

The most common edge detection algorithm in use is canny edge detection (Kaur and Kaur, 2016). It is a multi-stage algorithm which involves the following steps.

Application of a Gaussian filter: This step serves two purposes. Firstly, it reduces the overall noise present in the image. Edge detection algorithms are very sensitive to noise in the image. There is a possibility that the small specks that appear as noise can be detected as tiny objects resulting in erroneous edge detection. Secondly, it gives the sharp object edges a gradient fading towards the background which is the basic principle on which canny detection works (Cope and Rockett, 2000).

Finding intensity gradient of the image: This step passes the image through a sobel kernel which is a preliminary detection of the edges. The basic logic behind this is that the gradients caused by the gaussian filter is always literally perpendicular to the edge of the object.

Non-maximum suppression: This stage removes the unwanted pixels that may have a gradient and was marked as an object.

Hysteresis thresholding: This passes the image through two threshold values that gives sharper edges resulting in a high clarity result.

Hough transform: Hough transform is a process for extracting features which is used in digital image processing and image analysis. This process is used to find instances of objects within a specific class of shapes using a voting procedure (Wang and Chen, 2009).

The classical hough transform was used to mark and position lines in the image. But now it is being used to mark shapes like circles and ellipses. This brings us to Circle Hough Transform (CHT). CHT is a variation of hough transform used to detect circular objects within a digital image. The circles are detected by the aforementioned “Voting” procedure in hough parameter space and then picking out the local maxima in an accumulator matrix. On a cartesian plane, a circle is represented by the following equation:

$$(x-a)^2+(y-b)^2 = r^2$$

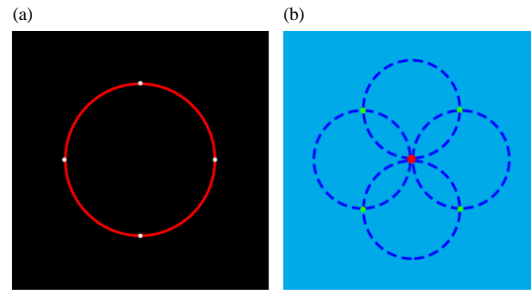


Fig. 1: Example of hough transform being performed on an image: a) Original and b) Hough parameter space

Where:

x, y = Points on the cartesian plane

a, b = Coordinates of the center of the circle

‘r’ = The radius of the circle

The concept of circle hough transform revolves around the above mentioned equation. An additional feature used in circle hough transform is edge detection (Chai *et al.*, 2014).

Take a sample image with a circle in Fig. 1a. Consider 4 points, marked on the circle in white as shown on the circle. Note that the radius of the circle is assumed to be known for this example. For each of the marked points (x, y) on the shown circle, define another circle with the marked points as the centers and radius r. This forms the hough parameter space as Fig. 1 b.

This is where the accumulator matrix comes into play. The accumulator matrix is used to track all the points in the Hough parameter space where an intersection occurs. From here, the local maxima point can be found (marked in red) in the Hough parameter space. The position (a, b) of this marked maxima is the center of the original circle (Chai *et al.*, 2014).

This algorithm can then be extrapolated and used to find multiple circles within a radius range. This extrapolation of hough transform is what will be used in the application of image processing to control crowding in a parking lot.

MATERIALS AND METHODS

Overview: One everyday example of automation where image processing can be applied is to control crowding of vehicles in a parking lot (Banerjee *et al.*, 2011).

Currently, every mall in Dubai and many around the world have systematic parking monitoring systems in place. These systems help the motorists entering those respective parkings to be able to find parking spots for

their vehicles with ease. It helps them to avoid areas if the system returns that the area has no current free spaces.

In more advanced systems such as the ones in place in the city center malls and also in Dubai Mall, an ultrasonic sensor above each parking space which detects the presence of a sizable object in that respective space, in this case, a car. Now in the event of a car already present, it will signify this using an LED which turns from green (when empty) to red (when occupied). Before entering the row of parkings, a motorists can take a quick scan of the row and try to spot a green light before advancing into said row. An added feature is that the simple boolean data from all sensors are sent to a central system for monitoring and thus that will display the number of parkings open in that parking lot, thus, giving an overall density of the crowd in the lot (Sun *et al.*, 2015).

The disadvantages of the above-mentioned system is that there is a large amount of complexities in the installation of such a system. Considering the capacities of a large mall like Dubai Mall with over thousands of parking spaces, installing a sensor over each parking space is a huge and time intensive process.

The solution for this disadvantage is to use image processing and cameras to monitor the parking lot. This reduces the complexity of installation as one camera can serve to monitor a whole area of the parking lot. The salient feature here being that most parking lots of commercial complexes have cameras in place for security and surveillance. It will be just a matter of calibrating these cameras, acquiring these images and posting them to the program which will process the image and return the number of spots available to incoming motorists.

The advantage of this system is that it serves 2 purposes: one, as a convenience for incoming motorists to find parking spots easily and also, it doubles as a security system for surveillance. Both of which are necessary for any commercial complex in today's day and age.

The solution to the above mentioned drawback is the usage of cameras mounted and calibrated in a way to recognize the difference between empty and occupied parking spaces (Banerjee *et al.*, 2011).

Setup and calibration: A rig of a parking lot is prepared as shown in Fig. 2. A green circle is placed within the central area in a way such that a vehicle occupying the spot in mention will eclipse the circle from the view of the camera monitoring the lot (Jung, 2014). The green circle thus, acts as a marker as to whether the spot is occupied or vacant. The parking lot rig that has been designed has a total of

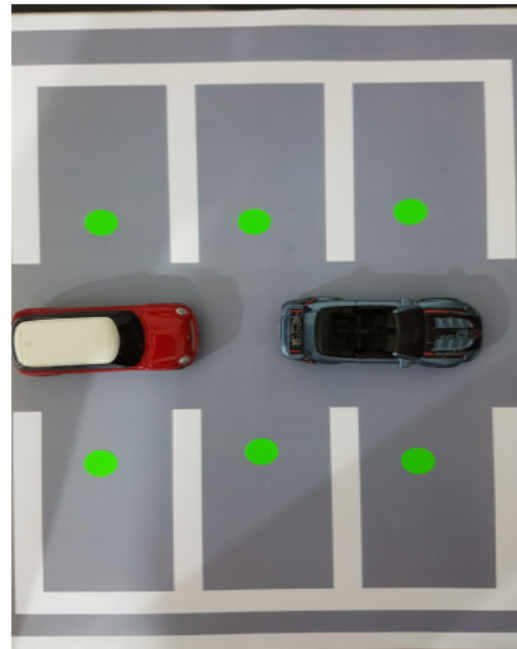


Fig. 2: A miniature rig of a parking lot

6 parking space as shown in Fig. 2. Now, once it passes through the circle hough transform algorithm, the output image marks all the green circles visible, counts the number of green circles, thus, returning the number of parkings available. This is where circle hough transform comes into the picture.

Currently, for the algorithm to work at full potential, the camera has to be placed at an almost vertical, i.e., 90° with respect to the plane of the parking lot. This vertical angle helps in easy detection of the circle. At an oblique point of view for the camera, the circle will appear as an ellipse which may cause the algorithm to not be able to detect the green marker circles on the parking spaces.

A further extension of this algorithm could be to expand the circle hough transform algorithm to detect ellipses to allow for an easier calibration of the camera as a sensor in this environment.

Application of circle hough transform: The green circle acts as a marker as to whether the spot is occupied or vacant. Simply enough, if a parking space is occupied, the green circle is not visible.

The parking lot in mention has a total of 6 parking spaces. Now, once it passes through the algorithm, the output image marks all the green circles visible, counts the number of green circles, thus, returning the number of parkings available.

This code marks the green circles and stores a matrix of the respective circle center coordinates in the variable 'Centers'. Also it stores their respective radii in the variable 'radii'. Figure 3 shows the code window in MATLAB.

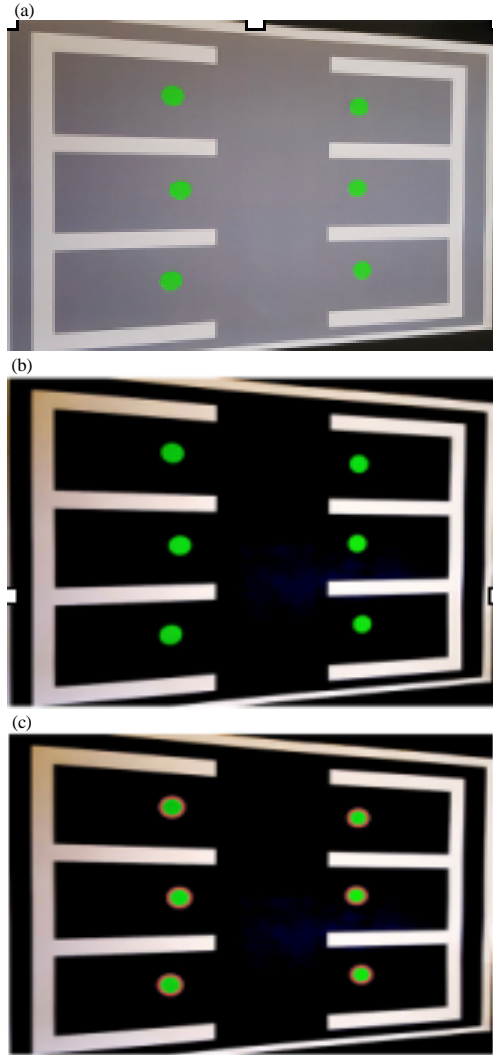


Fig. 3: a) The initial image of empty parking lot; b) The image with a higher contrast and c) The green circles marked in red

RESULTS AND DISCUSSION

Completely empty parking lot: This uses an image with a completely empty parking lot and passed to the algorithm. This step was used as a calibration step, so that, given a certain camera distance from the surface and camera angle, we can find the correct radius range for which the algorithm can find the green circle markers.

Note: The image contrast had to be drastically increased, so that, when the image is converted to grayscale, the background tarmac is darker than the bright green circle. This causes a sharper edge for the circle which can be detected faster and more accurately. Figure 3 shows the different phases of the image being processed (Fig. 4).

Partially occupied parking lot: This iteration simulates a situation wherein 3 out of the 6 parking spaces are occupied. Once the image of the parking lot is passed to the algorithm, it returns an image where the green markers in the available spots are marked in red and also returns the correct number of parking spaces available to park in. Figure 5 shows the image of the partially occupied parking lot being processed to show a successful end result (Fig. 6).

Both the iterations of the experiment have shown successful results. This proves to be a novel idea with a minimum amount of software maintenance to control crowding in a parking lot.

In the first iteration, "Completely empty parking lot": The rig was made with 6 parking slots and none of them occupied. The script returned an accurate result of 6 parking spaces being empty as seen in the following image of the command window.

In the second iteration, "Partially occupied parking lot": 3 spaces out of the 6 on the rig were occupied by vehicles. The script once again returned the accurate no. of spaces being empty, i.e., 3 parking spaces empty.

Improvisation of the algorithm: During the initial few tests and trials of the experiment. The process would be to capture the image and directly run it through the

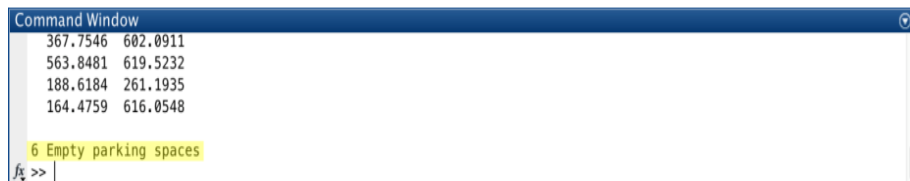


Fig. 4: MATLAB command window showing the number of parking spaces available in the empty parking lot

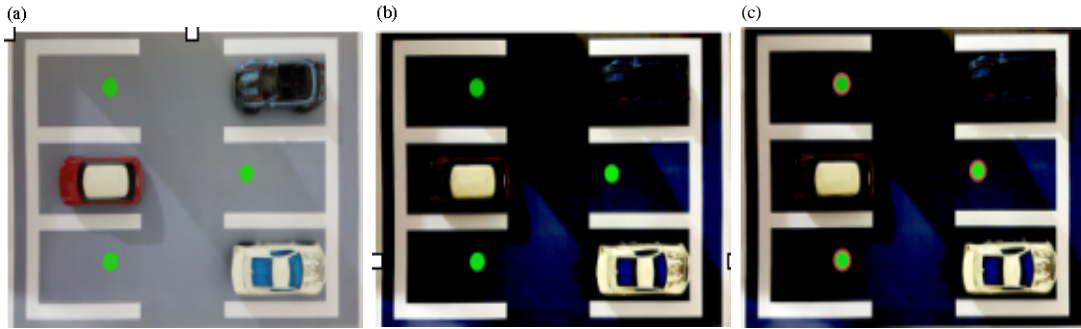


Fig. 5: a) The initial image; b) The image with a higher contrast and c) The green circles on the empty spaces are marked in red

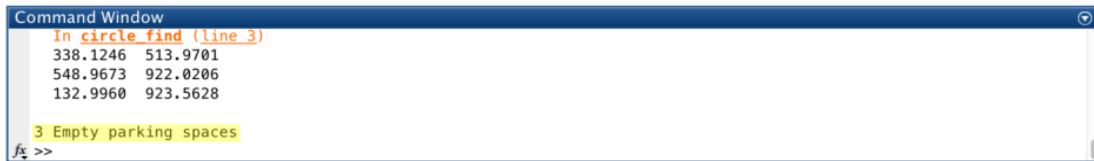


Fig. 6: MATLAB command window showing the correct number of parking spaces available in a partially occupied parking lot

algorithm. The problem with those trials was that in a grayscale image, there wasn't a clear distinction between the circular marker and the grey tarmac. This yielded results whose accuracy was very dependent on the ambient light. In the event of any error in angle of light and/or intensity of light the accuracy of the edge detection would be lowered significantly.

As you have observed in Fig. 3 and 5, in stage b of both images, the image goes through a high contrast filter, so as to give a significant distinction between the marker and the tarmac once the image is passed on from the RGB to grayscale domain for the circle hough transform algorithm. This produced very accurate results as observed in the iterations of the experiment and these results were independent of the ambient light as the high contrast filter created a distinction in itself and brought only bright circular objects into focus.

CONCLUSION

The most optimum method for controlling crowding in a parking lot is with circle hough transform. This is where the parking lot is designed with a bright green circle marker on each parking space which gets covered by a vehicle in that respective parking space. Circle hough transform method is the most optimum method because the logic is easy to parse and it is the least complex to

implement. There is a much thinner margin of error with this method because of the fact that the algorithm is very modular.

For further development of this system, it can be extended to a system where multiple cameras can monitor multiple areas of a larger parking lot and the data from all these cameras can be consolidated at one central computing system. This computing system can then deploy numeric results to LED displays at parking lot entrances to help incoming motorists to find parking spaces faster. From a practical aspect, the cameras capturing images should be placed at an angle inside indoor parking lots which do not provide much vertical space for a camera to be completely vertical above the plane of the parking lot. In such a scenario, from the view point of the camera, the circular marker will appear as an ellipse. Thus, circle hough transform can be modified to detect these ellipses.

ACKNOWLEDGEMENTS

On the outset, we would like to express our gratitude to Prof. R.N. Saha, Director of BITS-Pilani, Dubai Campus and Dr. Abdul Rajak, HOD of EEE Department. We would also like to thank Mr. Avinash Kamath who helped to refresh web development skills which has been used extensively in this project to give a realization to theoretical algorithms on edge detection using Javascript.

REFERENCES

- Arora, S., J. Acharya, A. Verma and P.K. Panigrahi, 2008. Multilevel thresholding for image segmentation through a fast statistical recursive algorithm. *Pattern Recognit. Lett.*, 29: 119-125.
- Banerjee, S., P. Choudekar and M.K. Muju, 2011. Real time car parking system using image processing. *Proceedings of the 3rd International Conference on Electronics Computer Technology (ICECT'11) Vol. 2, April 8-10, 2011, IEEE, Kanyakumari, India, ISBN:978-1-4244-8678-6, pp: 99-103.*
- Chai, Y., S.J. Wei and X.C. Li, 2014. The multi-scale hough transform lane detection method based on the algorithm of Otsu and Canny. *Adv. Mater. Res.*, 1042: 126-130.
- Cope, R.K. and P.I. Rockett, 2000. Efficacy of gaussian smoothing in canny edge detector. *Electron. Lett.*, 36: 1615-1617.
- Davies, E.R., 2000. Image processing and its applications. *Electron. Commun. Eng. J.*, 12: 194-194.
- Jung, H.G., 2014. Semi-automatic parking slot marking recognition for intelligent parking assist systems. *J. Eng.*, 1: 1-8.
- Kaur, P. and B. Kaur, 2016. 2-D geometric shape recognition using canny edge detection technique. *Proceedings of the 3rd International Conference on Computing for Sustainable Global Development (INDIACom'16), March 16-18, 2016, IEEE, New Delhi, India, ISBN:978-1-4673-9417-8, pp: 3161-3164.*
- Lew, M.S., 2001. *Principles of Visual Information Retrieval.* Springer, London, UK., ISBN:1-85233-381-2, Pages: 359.
- Sun, W., N. Shibata, M. Kenmotsu, K. Yasumoto and M. Ito, 2015. A method for navigating cars in multilevel parking facility. *J. Inf. Process.*, 23: 488-496.
- Wang, L. and L.Q. Chen, 2009. Concentric circle detection based on chord mid-point hough transform. *J. Comput. Appl.*, 7: 1937-1939.

Luktung Mohlum Song: Creative Development Model for Economy

¹Chinda Kaensombat, ¹Souneth Phothisane and ²Sitthisak Champadaeng

¹Faculty of Fine and Applied Arts, Khon Kaen University, Nai Muang Sub-District,
Muang District, 40002 Khon Kaen, Thailand

²Faculty of Cultural Science, Mahasarakham University, Khamriang,
Kantarawichai, 44150 Mahasarakham, Thailand

Abstract: This study is part of the qualitative dissertation using the data from documentation research, observations and interviews. The content of this study is a presentation on the current status of Luktung Mohlum which is one of the important terms within this research. The result of the study shows that Luktung Mohlum in its current status is going through an era of adjustment due to modern communicating technologies which is the main factor to its broadcast on today's online society. This had caused the music creators to manage their broadcast channels, distribution channel and channels to collect royalties. Currently, the direction of Luktung Mohlum music is the development on the creativity of the music, presenting a more up-to-date content in responding to the need of the society while considering the values that the consumers hold to ensure the survival of the music itself. However, one important aspect that is still ingrained in Luktung Mohlum music is that it is still rooted in Isaan culture using Isaan dialect in the lyrics and the "waad" (conversational style of singing) and the use of Isaan traditional music instruments along with contemporary musical instruments which is considered a legacy in the cultural wisdom, poetry, music and had created a beauty in arts and culture.

Key words: Luktung Mohlum song, current status, Isaan culture, Isaan music, arts and culture, dissertation, communicating

INTRODUCTION

Background and importance of the research: Mohlum is local art that is the most popular and most outstanding of Isaan, since, the past to the present. The music has evolved itself until it became Luktung Mohlum. "Rum Tone" is one component that got integrated into the Mohlum performance. Historically, during the Second World War (1938-1944) under General Por Phibunsongkhram as the prime minister, Western Nationalism ideology was supported as a way to make sure the is safe by supporting Rum Tone. It has become a national policy. Rum Tone was turned into a record and reach its popularity. Later on, Rum Tone was rearranged into a new form, incorporating Isaan language as a lyric and evolved itself into Luktung Isaan music (Plangwan and Plangwan, 2014). "Waad" was added to the Luktung music and eventually became Luktung Mohlum music.

Different era for Luktung Mohlum: Luktung Mohlum music started during the year 1967-1982. The factors that

caused the changes in Isaan culture started from the basic structure in communication and transportation. Because of this reason, Mohlum Klon artists came into existence (Suwannawong, 2010). Soon after, Luktung Mohlum became popular due to the radio stations. This era was the opening era for the Luktung Mohlum. The Luktung Mohlum artists during this period were Angkhana Khunchai singing "Isaan Lum Plern", Chabaphai Namwai singing "Lum Plern Chern Yim", Supap Daoden singing "Lum Plern Salab Toey", Phailin Pornphiboon singing "Lum Klom Toong" and Hongthong Dao-Udon singing "Hongthong Khanong Lum".

Flourish era for Luktung Mohlum music was during 1983-1997. An important business was created in the Luktung Mohlum industry. From producing records going into a commercial music industry in a capitalist system. There were competition between this type of music against the other types such as contemporary Thai music which has stronger base than Luktung Mohlum. This had forced the Luktung Mohlum business to have to adjust itself by becoming more modern (Rachburi, 1994). During this era, many Luktung Mohlum artist became well-known

such as Pornsak Songsaeng singing “Teoy Sao Jun Kung Kobe”, Chalermphol Malakham singing “Tam Jai Teud Nong”, Satit Thongchan singing “Pak Grod Jai Kidteung”, Pimpa Pornsiri singing “Namta Mia Sa-U”, Duenphen Amnuayporn singing “Sao Na Kaad Rak”, and Siriporn Amphaiphong singing “Bow Rak See Dum”.

The regressing era for the Luktung Mohlum music (1998-2007). This is the era of music copyrights were violated which had caused many music company to shut down. This had brought about many changes due to the changes in the information aspect. It was the internet communication though social network that had taken over the audience society. Smartphones (Anonymous, 2007) which had incorporate the capacities found in computers cause the regression for the Luktung Mohlum music. It was a crisis which required struggle to be able to survive. The artist who were able to survive were the ones who were popular and whom the company can launch a full album for. For new artists, they were required to make “single”. Full album wasn’t an option which had reduce their opportunity of making a living.

Modern technology, new kind of communications, and the changes in the economy, social and political system had created a new form of society and these processes had also created small broken pieces of the culture. This includes the changes in the experience concerning space and time and new kind of experience, identity and culture. These conditions had defined the social and economy base as well as the cultural roots off Isaan. The economic system shifted from producing for self-reliance to producing for the selling. The tradition and culture as well as the worldview of the audiences and changed. The status and the role of Luktung Mohlum in the midst of social changes caused Mohlum artists who, originally were supposed to be passing on the local knowledge and traditional culture to have to adjust themselves. From Mohlum Pheun and Mohlum Klom to Mohlum Pleun, Mohlum Sing and eventually Luktung Mohlum (Chanthabut and Chuntalu, 2004).

The new society after the Industrial revolution that connects the world through technology or so-called the digital age, had created stakes, benefits and other things in the arts and cultural aspects, economy and politics. All these changes came from the reasons that globalization affect all of today’s aspect which includes the music industry. There is a need to study the current status of Luktung Mohlum with the hope that one day, Luktung Mohlum will come back with the music adjusted for survival.

Research objectives: To study the current status of Luktung Mohlum.

MATERIALS AND METHODS

The goal of this research was to study the current status of Luktung Mohlum music using secondary source of data from the related documents and research and the primary source of data from the population selected for this research and field visit within Thailand.

This study was conducted with a population of 14 people who have the knowledge, understanding and are related to the Luktung Mohlum music which includes. The 5 experts governmental experts who are directly related to the study of Luktung Mohlum music such as academics, local wise men and leading people in the Mohlum performances industry. The 6 practitioners which include the president of the music company, artists, song writers, chorus editors, musicians and producers.

The 3 related people who are in in direct contact with the Mohlum music which are producers and music company CEOs. The tools used for collecting the data are surveys, observations and interviews.

RESULTS AND DISCUSSION

The current status of Luktung Mohlum music from 2008-2016 is in its adjusting era. In this current status, the business needs to adjust reduce their investments from when compared with the past. This change affect people within the industry starting from the artists, the creators, or the people working behind the scene. Today, Luktung Mohlum music also needs to adjust its market strategies at the same time. Apart from the fact that every company is focusing on digital downloading, they also make profits from the copyrights. Additionally, they also bring back the Luktung Mohlum artists who already have their fan base which makes it easier than building up a new artist since they already have fans and are still considered popular artists. Former Luktung Mohlum artists also remake their own music. These phenomenon can be concluded as follows:

The characteristics of current Luktung Mohlum artist: These are the characteristics of the current Luktung Mohlum artists:

Older Luktung Mohlum artists were brought back with the remake of their music and new music that isn’t necessary Luktung Mohlum in order to create the novelty amongst the audiences. This kind of things are well accepted such as the song “Mua Rai Ja Poh” by Duenphen Amnuayporn.

Taking the current popular artist to sing old popular songs or making a remake of the songs by different artists. This proved to be rather popular such as an album called “Tam Hoy Isaan”.

Different kind of music artists such as rock singers singing a song together with older Luktung Mohlum singer, singing together in a contemporary music style with unique composition and chorus and created a sense of features and exotic for the audiences. This proves to boost the popularity really rapidly such as the song “Kid Hod” by Siriporn Amphaiphong and Bodyslam.

Taking old popular songs and recreate the song for the original Luktung singer to feature with older Luktung Mohlum singer by adjusting some part of the lyrics and add additional “klon”. The chorus were all rearranged to show the assimilation of the local Isaan music and contemporary music which increases the song popularity such as the song “Cha Ting Nong Noi” by Chai Muangsing and Jintara Poonlarp.

The current characteristics of the song composition:

Today, the music and lyrics are composed in the following manners. Remake of the original music and lyrics sung by new artists such as “Tam Hoi Isaan” by Grammy Gold music company, under Grammy. The composition of the music and lyrics through the integration of different types of music such as rock music integrated with Luktung Mohlum music in order to create a featuring and the feeling of exotic for the audiences such as the song “Kid Hod” by Siriporn Amphaiphong and Bodyslam.

The melody or “waad” from “lum pleun”, “teuy”, “lumpaen”, “lum deun kalasin” and “waad lum deun khon kaen” are integrated into Luktung music. Some parts of the music and lyrics were changed and newly composed “klon” added to the song such as “Cha Ting Nong Noi” by Chai Muangsing and Jintara Poonlarp.

Currently, many new blood music composers had emerged due to the fact that it is the era of change, a technology and digital age and the age of communication which led business of musical industry to commercial song downloading system. This forced the song writers to have to improve their musical form and lyrics to be more up-to-date and improve the technology in the music production to respond to the need of society.

Chorus arrangement in the current Luktung Mohlum music:

The current status of the chorus arrangement in the Luktung Mohlum music consist of the integration between rock music and the old Luktung Mohlum music which resulted in a contemporary kind of music creating the featuring and the feeling of exotic to for the audiences while rapidly increasing the popularity such as the song “Kid Hod” which is an empirical example.

The current status of the Luktung Mohlum music: It was found that in terms of the recording of the current songs, a local instruments such as “kan”, “saw”, “pin”,

“ponglang”, “wote” or other contemporary musical instruments such as saxophone, organ and guitar, were used in the solo. Some of the instruments were inserted to as a counter point to the vocals where the musician can improvise. Many synthetic sounds were also added to make the music more modern. Most music composers today have at least the basic knowledge in how to use computers since many need to have home studio to make records, make music scores, music demo and using the computer to fix the bad points before the real recording process happen in the studio.

The current characteristics of the Luktung Mohlum music production:

The current process of music production needs to consider the creativity of the music in order to increase its value. There needs to be at least 5 components as follows:

- Music creation
- Production
- Manufacturing
- Distribution and application management
- Consumption

Every component, every group, every relationship, every connection, all the way to the educational system that helped polish the skills through organizations and other art supported institutions. The evaluation within those institution needs to happen to determine the next round of budget. There needs to be lawyers taking care of the copyrights, accountants to take care of the income and expenses, marketing team to analyze and plan. There needs to be IT people to take care of the computer systems to manage the information of concerning the music and other aspect of the company. The point is to create a database to prepare for future recording. The characteristics of the Luktung Mohlum production is in its adjusting era. The company had brought back the famous Luktung Mohlum artists that already have a fan base which will make it easier for the music to be sold comparing to building up new artist. These famous artists were brought back to do a remake of their music. These kind of phenomenon can be observed as follows. The phenomenon of the Luktung Mohlum artist name “Duenphen Amnuayporn” who already have her own fan base who was brought back to have her music reproduced as well as making new music. She also sing new style of songs to create the feeling of exotic for the audience. One of her popular case is a song called “Mua Rai Ja Por”.

The phenomenon of Luktung Mohlum singer “Siriporn Amphaiphong” who was brought back to sing the songs that was made popular in the past by the original singers such as the song “Tam Hoi Isaan” which became quite popular.

The phenomenon of the Luktung Mohlum artist “Siriporn Amphaiphong Featuring Body Slam” in a song called “Kid Hod”. This method of bringing artists from different kind of music style such as Luktung Mohlum artist and a rock artist is the first phenomenon for Thai music industry. Bringing artists such as Siriporn Amphaiphong who is a successful artist with lots of fan and Bodyslam which is a well-known band together was very successful because of the music composition style and the chorus arrangement in a contemporary music style and created a sense of erotic for the listeners. It was also an integration of local Isaan music culture and a modern music culture that caused this attempt to be successful.

The phenomenon of Luktung Mohlum like “Jintara Phoonlab featuring with an actor/Luktung music composer Chai Muangsing” in a song “Cha-Ting-Nong-Noi” after being adapted and some parts of the lyrics change as well as adding the “klon” into the song which makes it unique and exotic. The chorus was also rearranged in the Luktung Mohlum song which is an integration of local Isaan music culture and a modern music culture that caused this attempt to be successful.

Currently, there is a quick advance in technology. Things that used to be difficult became easy things for the consumers. The things that used to provide income for the music company became expenses due to the music player equipment of the people in this digital age that plays digital files. CDs that used to be produced disappeared from the market. Consumers developed that behavior in consuming free media from smartphones and computers. Neither the audio nor video files are generating income for the music companies who needs to find ways to survive by focusing on the big producer markets for support such as selling tickets for a concert that consist of multiple artists using the artists to promote a product or performing concerts. The companies also make money from collecting copyrights fee from karaoke bars, pubs, restaurants, TVs, radios and any other ways that allows them to make income. Although, consumers are still demanding good quality of music but they purchase the music less and less. From this phenomenon, the way to survive for the mid well-known Luktung Mohlum artists is launching a single to maintain their reputation while gaining continuous work. Many music companies and song writers shifted themselves to do other things. The Luktung Mohlum industry today is still in the adjusting period due to the time that keeps changing. Many artists and music companies still have the determination to produce continuous good quality music by holding on to the hope that the music industry

will get better after the adjustment period and the changing behavior of the consumers. The producers need to find ways to develop the Luktung Mohlum music to increase its value which can be listed out as the followings.

Improvement of the artists the CEO and the producers needs to find and select the screening and artist development department as their first priority. They need to have an screen department that will help improve the quality of the artists both in the vocal department and the uniqueness of the artist’s vocal for the performance and the Mohlum. They need to create their uniqueness in their costumes, dance, their determination, their patience, and their pride as Luktung Mohlum artists.

The direction on development in composing Luktung Mohlum song to create the uniqueness for the listeners needs both the CEO and the producers to setup an screen department for the song writers in order to get the songs that meet the needs of the companies as well as increasing the value. Therefore, the songs composed need to address the characters of the artists. The songs have to be composed in the right direction that can create the uniqueness for the characters. In addition, the content of the songs need to have the correct prosody, the “waad” part needs to be correctly composed and the concept of the songs need to be different and based on the positive perspective that kept up with the world incidents.

The chorus arrangement of the Luktung Mohlum song to achieve an outstanding characteristics and the feeling of exotics for the listeners requires the CEO and the producers to setup the screening department for the arrangers, since, the arrangement of the chorus is the adjustment of the music and the vocal of the artists to go in harmony. People who can arrange the chorus need to have good knowledge in music. Moreover, they have to have the knowledge in the melody and the “waad” as well as understanding the nature of every musical instruments chosen for the songs in order to bring each quality of the sounds and mix them together in a harmonious way. The most important thing is that chorus arrangers today need to have a home studio to prepare the music score or record a music demo in order to find the bad points in the songs before the actual recording.

The direction in developing the musicians for Luktung Mohlum music ipn order to achieve the exotictness for the listeners requires the CEO and the producers to setup a screening department as the first step. Since, the recording and the creation of music needs one main instrument for the solo by using a local instrument integrated with the contemporary musical instrument. Therefore, the musicians need to communicate their emotion to the listeners very clearly. There are many

technics in the performing as well as the knowledge in the melody or local music patterns. Producers needs to have a project to setup a workshop for using computers to compose music in order to provide the basic knowledge in the recording.

Therefore, the current production needs to consider the creativity for the music in order to create the value requires 5 components and related people, since, every component, every group and ever relationship needs to be connected in order to achieve the success.

CONCLUSION

The current situation of Luktung Mohlum music, from 2008-2016 was the era of adjustment due to the fact that the communication technology of social media system has been playing more roles in the consumption of information as well as song listening in today. The fact that human society communicate faster through different technologies such as the internet, phone and smart phones, the behavior of the consumers had also changed a lot compared to the past. This created new forms of entertainment which changes people's behavior in listening to music by listening to more mp3 songs or through the internet in stead of buying CDs or cassette tapes. This had caused the sale of both cassette tapes and CDs to reduce. Because of this many entrepreneur began to develop an interest in digital music more and more. This source of income had replaced the sales of cassette tapes and CDs. To maintain the customer base and their own market share which had caused the music industry to continue expanding in this digital age while causing the overall picture of the music industry smaller.

Once the music industry gets smaller, the budget for the production also gets smaller which initiate the trend of artists shifting to different companies. This is also the same with the people who are working on the background who weren't prepared for the changes. These factors directly affect the people who are working in the music industry since the producers need to screen for the artists that the listeners want to hear while the income is no longer from the sales. The music industry also adjust their own investment to be smaller than the past and when this change happens, it also affect the people who are working both front and back stage, the ever changing Luktung Mohlum artists and the Luktung Mohlum songs themselves which need to adjust their marketing strategies at the same time.

The producers today will need to consider the creativity of the music to create the added value of the music which will need 5 components and related people

which are the people involving in the music creation, production, manufacturing, distribution and application management and consumption.

Luktung Mohlum music today is going through the era of adjustment due to the communication technology of social media which plays more role in the consumption of information as well as the way people are listening to music today. The fact that human society now have faster communication through all the technological system and caused the behavior of the consumers to change compared to the past. We need to create and develop the songs in order to achieve the success by using the knowledge, the education, the creativity and the intellectual property to connect with the root of Isaan traditional culture while using the technology and new innovation to develop the Luktung Mohlum music to create the value. The word "value" here can be categorized into 2 aspects) the concrete aspect such as literature language, the melody, the arrangement, the chorus, the vocal and the performing) the abstract aspect; the beauty that comes from the lyrics, from the singing, the performing, the chorus arrangement, the expression of emotion and the value which means the value in Isaan culture, legacy of the cultural wisdom which ensure the income and the survival of the Luktung Mohlum artists.

RECOMMENDATIONS

The modern technology, the new kind of communication and the changes in the economic, social, and political system are creating a new kind of society. These processes are breaking the culture into smaller cultures which includes the free space and time and also new characteristics of experience, identity and culture. These conditions gave a new base to the socio-economy and culture which goes along with the research by Chanthabut and Chuntalu (2004) who studied the status and role of Mohlum amidst the trend of social change: A case study of Mohlum in Ubonratchathani Province and provide the explanation that from the production system that aimed at household consumption to the production system aimed at selling Isaan traditions and culture along with the changing world view of the listeners. Mohlum artists who used to be people who pass on the local wisdom and culture now have to adjust themselves in order to survive from Mohlum Peun to Mohlum Plern to Mohlum Sing and eventually Luktung Mohlum.

The music industry is adjusting in developing their music for survival which goes along with the structural functional theory that states that all societies have different functions and systems for survival. The parts and the subsystems within the society will work in

accordance to reach the final goal which is to survive. This also includes the ability to adapt or change under different circumstances to fit in with certain period of time according to the theory by Radcliffe and Radcliffe (1971).

SUGGESTIONS

The suggestion for further research is to study the development of the Luktung Mohlum music industry in an ASEAN level.

REFERENCES

- Anonymous, 2007. Management system of luktung music business. RS Public Company Limited, Bangkok, Thailand.
- Chanthabut, S. and T. Chuntalu, 2004. The status and roles of molan in ubolrachathani amid global social changes. Office of the National Culture Commission (ONCC), Bangkok, Thailand.
- Plangwan, N. and W. Plangwan, 2014. Interview. Academic Publishing, Cambridge, Massachusetts,.
- Rachburi, C., 1994. History of luktung music business from 1958-1982. Master Thesis, Thammasat University, Bangkok, Thailand.
- Radcliffe, B. and A.R. Radcliffe, 1971. Structure and Function in Primitive Society. Cohen & West, London, UK.,.
- Suwannawong, N., 2010. The development of luktung isaan music to increase economical value. Master Thesis, Mahasarakham University, Talat, Thailand.

Computationally Effective and Practically Aware Pareto-Based Multi-Objective Evolutionary Approach for Wireless Sensor Network Deployment

Mustafa Ali Hameed and Ravie Chandren Muniyandi
Research Center for Software Technology and Management (SOFTAM),
Faculty of Information Science and Technology,
Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia

Abstract: Wireless Sensor Network Deployment (WSND) is an active research topic. Different approaches have been effectively developed for WSND. Multi-Objective Evolutionary Algorithms (MOEAs) are regarded as powerful deployment methods because of their adaptive flexibility in effectively searching and providing numerous deployment options for the user. In this study, a computationally effective and practically aware Pareto-based multi-objective evolutionary approach was developed for WSND. On the one hand, the initialization of the population and crossover operation were modified to obtain solutions that meet the connectivity constraints and improve the computational aspect for producing the solutions. On the other hand, a constraint of the dead zone was added to make the deployment practically aware in presence of restricted areas in the Region of Interest (ROI). The approach of the current study was compared with that of Khalesian and Delavar by generating the values of the lifetime and coverage as the conflicting objectives of the deployment. Results showed that the developed approach outperforms the previous approach with respect to these objectives.

Key words: Deployment, Pareto optimization, multi-objective, restricted area, dead zone, lifetime evaluation, coverage zone evaluation

INTRODUCTION

Wireless Sensor Network (WSN) is a technology developed to address the growing need for the observation and control of environments. Originally, the growth of WSN applications is conducted by military applications. WSN are currently used in various fields, such as medicine, environmental risk monitoring, traffic control, disaster monitoring and industrial process control (Debnath *et al.*, 2016). The various types of sensors as well as the limited need for infrastructure, make WSNs capable of providing continuous measurements in a wide range of environments for a large variety of applications (Juul *et al.*, 2015).

A sensor is a device capable of sensing objects in an environment and transforming the data gathered by the network for processing. Each sensor has a low processing unit, small data storage, limited battery and limited coverage and communication ranges (Njoya *et al.*, 2016). Sensor nodes are distributed in different places and work together to communicate information gathered from a Region of Interest (ROI) through links that connect them wirelessly and send this information using multi-hop

communication to the base station called sink which sends the data to the user or to other networks (Abdollahzadeh and Navimipour, 2016).

WSN deployment is a major step in WSN design and is considered a solution for reducing the effects of sensor limitations (Yick *et al.*, 2008). Efficient deployment of the sensors is very important in improving coverage area and prolong lifetime of the network (Juul *et al.*, 2015; Tsai *et al.*, 2015). Connectivity is another critical issue in designing a sensor network for proper functioning regardless of diverse situations (Debnath *et al.*, 2016). One of the major challenges in sensor deployment is finding a trade-off among the conflicting objectives of the network, coverage and lifetime under certain connectivity constraints.

Multi-objective optimization has been used as an effective mathematical tool for addressing this type of optimization (Debnath *et al.*, 2016). Multi-objective considers several conflicting objectives simultaneously. In such a case, a set of alternatives with different trade-offs, called Pareto optimal solutions or non-dominated solutions generated instead of single optimal solution.

Corresponding Author: Mustafa Ali Hameed, Research Center for Software Technology and Management (SOFTAM), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia

In a practical environment, restricted areas such as rivers, lakes and unsafe areas can block the connectivity of sensor nodes. Connectivity blocking is an important challenge in WSN (Wahab *et al.*, 2016a, b).

Most of the existing studies on WSN assume that homogeneous sensors can be deployed anywhere (not regarding “restricted areas” that exist in practical situations). Khalesian and Delavar (2016) proposed the Constrained Pareto-based Multi-objective Evolutionary Approach (CPMEA) to maximize coverage and minimize energy consumption. They also designed a genetic-based operator which involves crossover and mutation for ensuring an effective search for optimal solutions. However, their work not aware for some areas in practical environments that are not suitable for sensor deployment. Another criticized aspect of the work by Khalesian and Delavar (2016) is their design of crossover operation from the perspective of computational complexity. In this study, we improve the CPMEA with the following objectives:

- Make CPMEA a Computationally Effective and Practically Aware Pareto-based Multi-objective Evolutionary Approach (CE-PA-PMEA)
- Evaluate CE-PA-PMEA and compare it with CPMEA

Literature review: Many researchers have proposed approaches to enhanced the sensor deployment problem for monitoring and surveillance in many environments and to improving networks and their functionalities (Syarif *et al.*, 2014).

The existence of restricted areas or obstacles in an ROI may degrade WSN functionality and these obstacles may cause loss of connectivity links between sensor nodes, thereby causing the network to break down. Wang and Ssu (2013), proposed a scheme for detecting obstacles in the ROI by identifying the size and location of these obstacles using radio units fitted in each sensor. The scheme recognizes the obstacles by marking the sensor nodes around the obstacle boundaries. Jourdan and de Weck (2004), proposed a multi-objective algorithm that is based on the genetic approach to optimize WSN deployment by providing Pareto-optimal (non-dominated) solutions to maximize the coverage and lifetime of the network. Jameii *et al.* (2015), proposed an algorithm that is based on Non-dominated Sorting Genetic Algorithm (NSGA-2) to optimize coverage, number of active nodes and energy consumption by putting some sensor nodes in sleep state. Liu and He (2014), proposed an approach that is based on ant colony optimization with a greedy mechanism to solve the problem of grid-based coverage

with low cost and connectivity maintenance. This approach can dynamically adjust the sensing and communication range to reduce energy consumption, thereby prolonging network lifetime. Zorlu and Sahingoz (2016), proposed a Genetic algorithm for sensor deployment optimization to maximize WSN coverage with the minimum number of homogeneous sensors. Syarif *et al.* (2014a, b), used a multi-objective approach for WSN deployment with some fixed obstacles. This approach which is based on NSGA-2, aims to optimize coverage and connectivity for WSN deployment by dividing the ROI into grid cells to identify the obstacles. They also proposed fitness and ranking functions for defining the best solution from Pareto Fronts (PFs). Sengupta *et al.* (2013), provided a method that depends on a multi-objective evolutionary algorithm and uses a decomposition approach to convert the problem of PF approximation into several single-objective optimization problems.

The deployment problem: In this study, the deployment problem is defined and the related objectives are introduced with mathematical formulations.

Problem definition: This research considers a two-dimensional rectangular environment ROI where K homogeneous sensors are deployed. This environment is subject to some restricted areas. The sensors have a pre-defined coverage zone for the sensing and a communication zone for connecting with one another. The sensors send data directly or via multi-hop communication to the single sink node (unlimited energy node), assuming that the sink is located at the center of the ROI. No movement is involved as the sensors are assumed stationary.

Problem formulation: The formulation of the wireless deployment problem as a multi-objective optimization is as follows:

$$F(x) = [f_1(x), f_2(x)] \rightarrow \text{Max}$$

- The first objective f_1 represents the coverage
- The second objective f_2 represents the lifetime

Decision variable x:

- x_i, y_i : represents the location of sensor node,
- $x = x_1, y_1, x_2, y_2, \dots, x_k, y_k$: represents the location of the sensor nodes as the decision variable

where, x represents the feasible solutions with problem constraints.

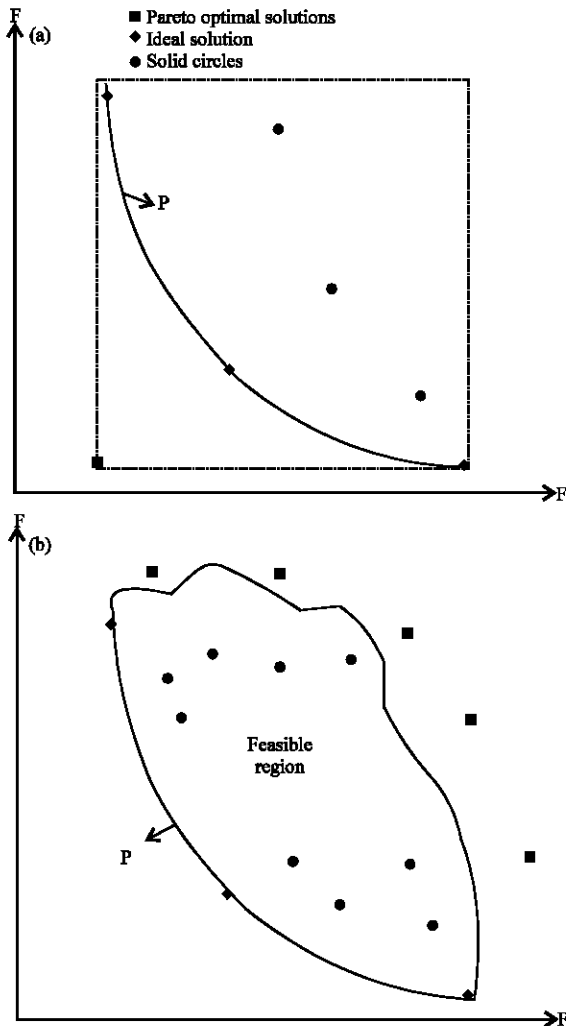


Fig. 1: The Pareto front of multi-objective problem MOP

A solution x^* dominates the other solution x' (if $(x^*) \geq f_i(x') \forall i \in \{1, 2, \dots, M\}$ and $f_i(x^*) > f_i(x') \exists i \in \{1, 2, \dots, M\}$ is denoted as $F(x^*) > F(x')$ where M is the number of the objectives in the problem.

The solution that is not dominated by any other solution in the objective space is an optimal (non-dominated) solution. The set of these optimal solutions is the Pareto Front (PF).

Figure 1a illustrates a PF with two objectives: the Pareto optimal solutions in the PF (marked with an asterisk) provide better values for the objective functions than any other solution in the objective space. The rectangle represents the ideal solution which provides the minimum objective values and is often considered unreachable. The solid circles represent the solutions that are dominated by at least one solution in the PF. In Fig. 1b, the bold curve indicates the PF. The solid circles

are the feasible solutions in the feasible region and the remaining solutions outside the feasible region (marked with a triangle) are considered infeasible (Jameii *et al.*, 2015). The conflicting objectives, we considered in WSN are as follows:

Coverage: The total coverage of the ROI is the ratio of the areas covered by all the sensors to the total area of the ROI:

$$Cover = \frac{\sum_{(x_G, y_G) \in ROI} g(x_G, y_G)}{N_G}$$

And:

$$g(x_g, y_g) = \begin{cases} 1 & \exists i \in \{1, \dots, k\}, \sqrt{\Delta x_i^2 + \Delta y_i^2} \leq R_s \\ 0 & \text{otherwise,} \end{cases}$$

where, $\Delta x_{iG} = x_i - x_G$, $\Delta y_{iG} = y_i - y_G$, N_G is the total number of grids in the ROI and R_s would be the sensing range of a sensor node if the ROI were divided into grid points where each point is covered by at least one node.

Lifetime: The lifetime is the ratio of time taken until one of the sensor nodes runs out of energy (failure) to the maximum network lifetime:

$$Life = \frac{\min\{T_{failure, i} \mid i=1, \dots, k\}}{T_{max}}$$

where, $\min\{T_{failure}\}$ is the minimum value of the failure time of the sensor nodes and represents the maximum number of sensing cycles before the energy runs out and T_{max} is the maximum possible number of sensing cycles and represents the maximum lifetime of the network.

MATERIALS AND METHODS

The proposed approach: This study proposes the CE-PA-PMEA for wireless sensor deployment as well as the operators designed and used for this purpose.

The WSN is modeled as a graph $G(k) = (Ver, Lnk)$ where Ver is the set of vertices (sensor nodes), Lnk is the set of links connecting the vertices and K is the number of sensor nodes. A link may exist between any two sensor nodes if the Euclidean distance between them is less than the communication range R_c . Theorem 1 of Khalesian and Delavar (2016) of graph theory is used for checking the connectivity of the produced graphs to ensure the required connectivity of the designed WSN. The general steps of the proposed algorithm are illustrated by a flowchart of Fig. 2.

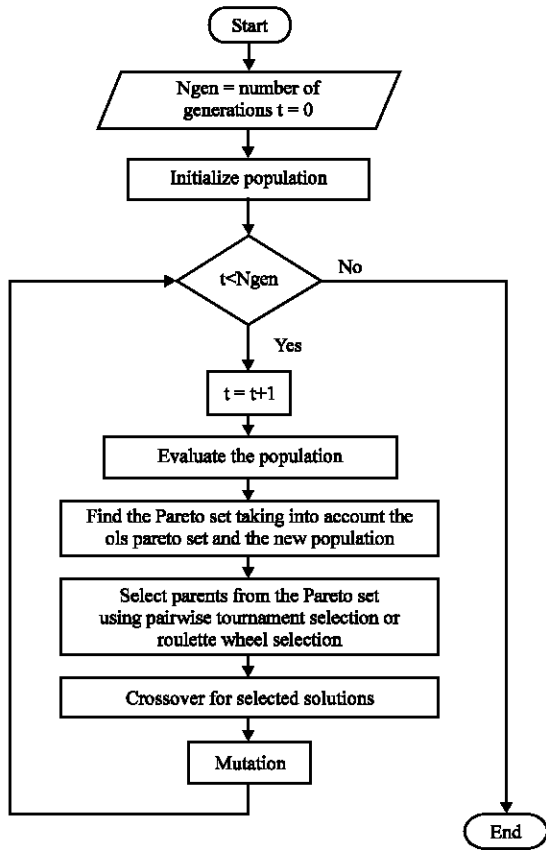


Fig. 2: The CE-PA-PMEA general steps

Theorem 1: Let G be the adjacency matrix of a connected graph and let $Y = [y_{ij}]$, $j = 1, \dots, k$ be the matrix $y = G + G^2, \dots, + G^{n+1}$. Then G is connected if and only if: $y_{ij} \neq 0$ for all distinct $i, j = 1, 2, \dots, k$.

Initialization: The proposed initialization process aims to produce an initial graph with one node which is the sink node. The first node is randomly generated in the environment while maintaining the avoidance constraint of the restricted areas in ROI by not generating nodes within the boundaries of this areas. Then, the next sensor is also randomly generated but with two constraints: the avoidance constraint of the restricted areas and being in the union of the communication zone of at least one of the preceding sensors until a graph with k nodes and $k-1$ edges is obtained. With this process, we do not need to check the connectivity of the graph and only need k adding operation thus, the time needed for initialization operation is less than that in the previous approach. This process is clarified by the flowchart in Fig. 3.

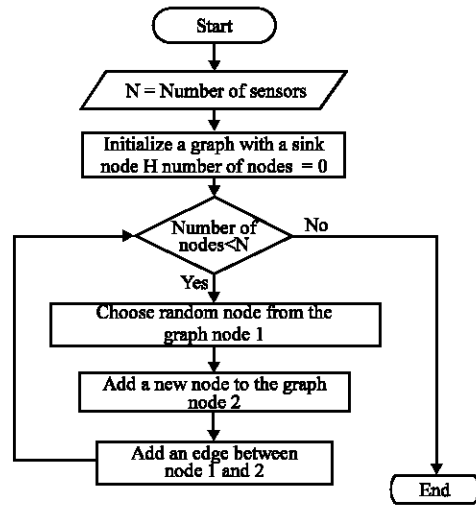


Fig. 3: The initialization steps

Evaluation of the population: After the initialization process, we choose promising solutions from the current population as parents for the next generation.

The values of the objectives should be calculated for the whole population and then the old Pareto solutions are combined with the population. Then, the Pareto set is determined and the two best solutions are selected in terms of Pareto dominance. The coverage and lifetime objectives are computed in study 2. Algorithm 1 shows the evaluation of the lifetime objective which calculated as (Khalasian and Delavar, 2016).

Algorithm 1; Lifetime evaluation:

Input: Number of sensor nodes (k), initial energy of each sensor ($E > 0$), path loss exponent ($\alpha \in [2, 6]$), transmission quality parameter (β), power amplifier energy consumption (amp) and minimum distance between nodes (d_{min})

Output: Lifetime

Step 1: Set $E_i(0) = E$ for all $i = 1, 2, \dots, k$

Step 2: Calculate the minimum transmit energy via. $P_{min} = \beta d_{min}^\alpha$

Step 3: Calculate the maximum possible number of sensing cycles via:

$$T_{max} = \frac{E}{P_{min} \text{ amp}}$$

Step 4: FOR $I = 1, 2, \dots, k$ do

Step 4.1: Calculate the shortest path from the sensor node to the sink

Step 4.2: Calculate the traffic load for the sensor node according to node level with the sink

Step 4.3: Calculate the distance from the sensor node to the next node

Step 4.4: Calculate the transmit energy $P_{ij} = \beta d_{ij}^\alpha$

Step 5: Calculate $T_{failure}$

$$T_{failure} = \min \left(\frac{E}{P \text{ amp}} \right)$$

Step 6: Set

$$\text{Lifetime} = \frac{T_{failure}}{T_{max}}$$

Determining the pareto solutions: For finding the Pareto set in the multi-objective optimization, we calculate the value of the objectives for whole population and combined with previous Pareto solutions. we follow who proposed an approach to determining the non-dominated solutions. This approach is summarized as follows:

- Sort all the solutions in decreasing order of their first objective function and create a list (O)
- Initialize a set S_1 and add the first element of list O to S_1
- For every solution O_i (other than the first solution) of list O, compare solution O_i with the solutions of S_1
 - If any element of set S_1 dominates O_i , delete O_i from the list
 - If O_i dominates any solution of set S_1 , delete that solution from S_1
 - If O_i is non-dominated by the elements in set S_1 , then update set $S_1 = S_1 \cup O_i$
- The Pareto solutions are in the resulting non-dominated set S_1

Selection operation: The well-known Pairwise Tournament (PT) and Roulette Wheel (RW) selection operators are used. We also use a selection operator equal to 0.5. If the random number generated is greater than 0.5, then the selection will depend on the coverage fitness values; otherwise, the selection will depend on the lifetime fitness values. Selecting parents by PT is explained by Algorithm 2 and selecting parents using RW is explained by Algorithm 3.

Algorithm 2; PT selection:

Input: The values of the objectives of the Pareto solutions and the number of tournament solutions k
 Output: The selection of the best solution as a parent
 Step 1: FOR each Pareto solution
 $x(j), j \in \{1, \dots, \text{number of pareto solutions}\}$
 $f_i(x), \forall i \in \{1, 2, \dots, M\}$, M is the number of objective functions
 IF r and > 0.5 Then
 objective Value (j) = $f_1(x), \dots$, (Coverage objective)
 ELSE
 objective Value (j) = $f_2(x) \dots$ (Lifetime objective)
 END IF
 END FOR
 Step 2: Select k random individuals from the input
 Step 3: Select the best individual of k with the highest objective value as a parent

Algorithm 3; RW selection:

Input: The values of the objectives of the Pareto solutions
 Output: The selection of the best solution as a parent
 Step 1: FOR each Pareto solution $x(j), j \in \{1, \dots, \text{number of Pareto solutions}\}$

$f_i(x), \forall i \in \{1, 2, \dots, M\}$, M is the number of objective functions

IF r and > 0.5 THEN

objective Value (j) = $f_1(x), \dots$, (Coverage objective)

ELSE

objective Value (j) = $f_2(x), \dots$, (Lifetime objective)

END IF

END FOR

Step 2: FOR each Pareto solution $x(j), j \in \{1, \dots, \text{number of Pareto solutions}\}$

$$\text{Probability ObjectiveValue}(j) = \frac{\text{Objective value (j)}}{\text{Sum (objective value)}}$$

END FOR

Step 3: Calculate the cumulative sum of the probability objective value

Step 4: Choose the first member of the cumulative sum where:

$\text{cumulativeSum} \geq \text{rand}$ where $\text{rand} \in [0, 1]$

Crossover operators: The crossover aims to produce new solutions for the next generation by selecting two solutions from the population; these solutions are called parents and exchange their information with each other with a rate probability (crossover rate) to generate new solutions called children (offsprings). The crossover operation is different from the previous approach. In the proposed approach, we perform the crossover in a light way. Two parents are combined to generate a new WSN with a graph of $2k-2$ edges after that a new graph with sink node is initialized and then a random node with its associated edges is selected from the combination and communicated to the new graph with respect to the connectivity constraint of the combined parents graph. we repeatedly select random nodes until a new graph with $k-1$ edges and k nodes which inherited the positions and their connections from their parents. The crossover process is applied to generate the two children shown in the flowchart in Fig. 4.

Mutation: The purpose of the mutation is to maintain diversity in the population by modified the location of some nodes accordingly to the connectivity constraints with some probability rate (mutation probability). The mutation implemented as (Khalesian and Delavar, 2016).

Termination criterion: When the maximum Number of Generations (N_G) reached, the pareto front returned as optimum solutions in the search space.

Practical awareness for dead zone areas: The previous approach does not process the problem of presence of restricted areas or geographical areas where sensors should not be localized. The network may be negatively affected if the sensors are placed in these areas, thereby causing the sensor nodes to fail and the network to break

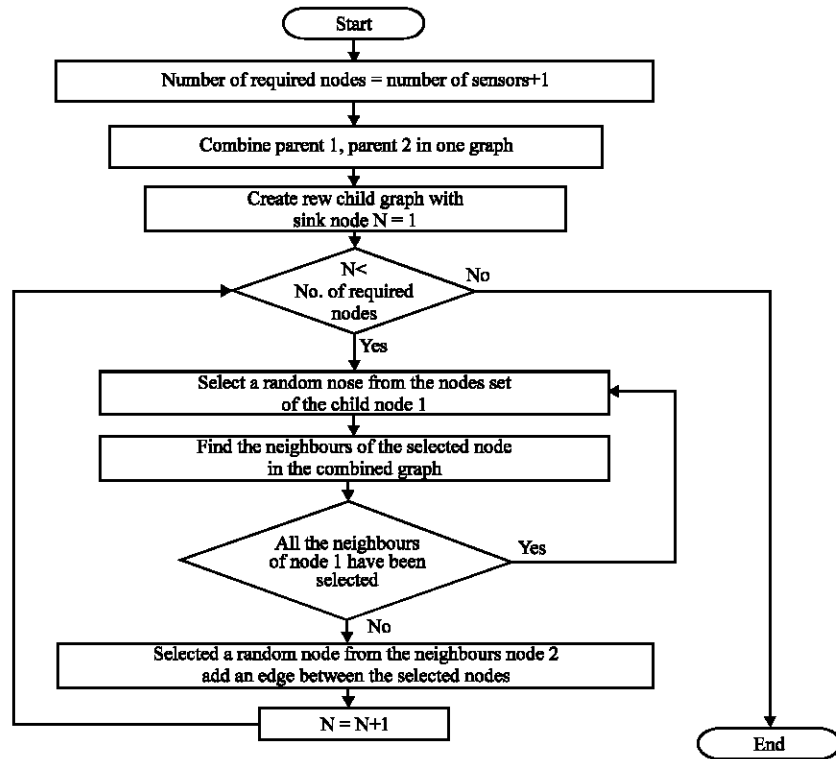


Fig. 4: Flowchart of crossover process

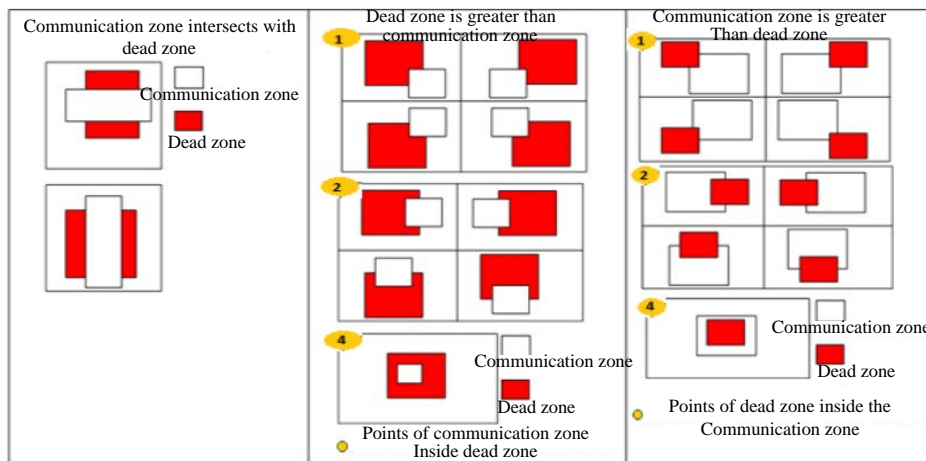


Fig. 5: The several states of crossing between the dead zone and the communicating zone

down. Thus, inside the environment, we assume the presence of a square-shaped area with specific boundaries we called it dead zone in which the sink node should not be placed. During the geometric solution of the initialization and crossover operation, the potential positions of the sensor nodes is tested.

Several states of crossing are founded between the dead zone and the communicating zone

because the position of the sensor nodes (x_i, y_i) should be generating inside a communication area (Fig. 5). These states are processed by redefining the communication area boundaries after exclusion of the dead zone area that defined as inequality constraints: $x_i < x_{r1}$; $x_i > x_{r2}$; $y_i < y_{r1}$; $y_i > y_{r2}$ where, x_{r1} , x_{r2} , y_{r1} , y_{r2} denoted as the boundaries of the dead zone.

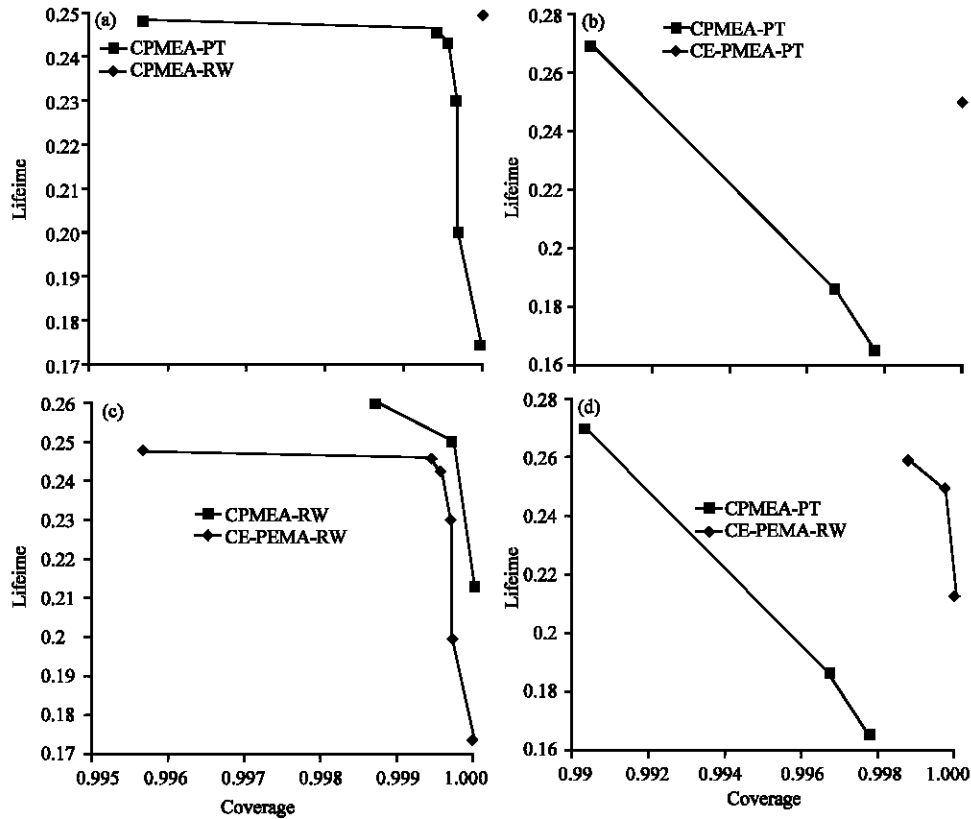


Fig. 6: a- d) PFs of the comparisons between the approaches without dead zone for P1

RESULTS AND DISCUSSION

The efficiency and effectiveness of the proposed evolutionary approach are evaluated by investigating its results and comparing them with those of CPMEA (Khalesian and Delavar, 2016).

We present the evaluation measures of the developed approach with its two selection operations, PT and RW, for both developments CE (Computationally Effective) and PA (Practically Aware) and the Comparison with the Previous Approach (CPMEA) is applied.

Two main scenarios (areas with and without a dead zone) are considered to validate the practical awareness of the developed approach. Moreover, the objectives of the network coverage and lifetime are calculated. The parameters settings considered during the operation of these approaches shown in Table 1.

No dead zones in the ROI: First, we compare the previous and our proposed approaches in the ROI without a restricted area.

CPMEA-PT vs. CPMEA-RW: The performances of the previous approach with PT and with RW are compared.

Table 1: Parameter settings

Parameter setting No.	No. of generations	No. of individuals	Crossover rate	Mutation rate
P1	250	200	0.1	0.1
P2	100	200	1.0	0.1

The PFs for the two approaches are generated for the evaluation. Figure 6a shows the PFs for the parameter P1. The PF generated by PT dominates that generated by RW for the both objectives. Figure 7a shows the Pfs for the parameter P2. The PF generated by PT dominates that generated by RW with respect to the lifetime.

CPMEA-PT vs. CE-PMEA-PT: The previous approach and our computationally efficient approach, both with PT are compared. The PFs for the two approaches are generated. Figure 6b shows the PFs for P1. The solution locations are optimal for the coverage with respect to CE-PMEA-PT and optimal for the lifetime with respect to CPMEA-PT. While Fig. 7b shows the CE-PMEA-PT is optimal for the two objectives for P2.

CPMEA-RW vs. CE-PMEA-RW: The previous approach and our computationally efficient approach, both with RW

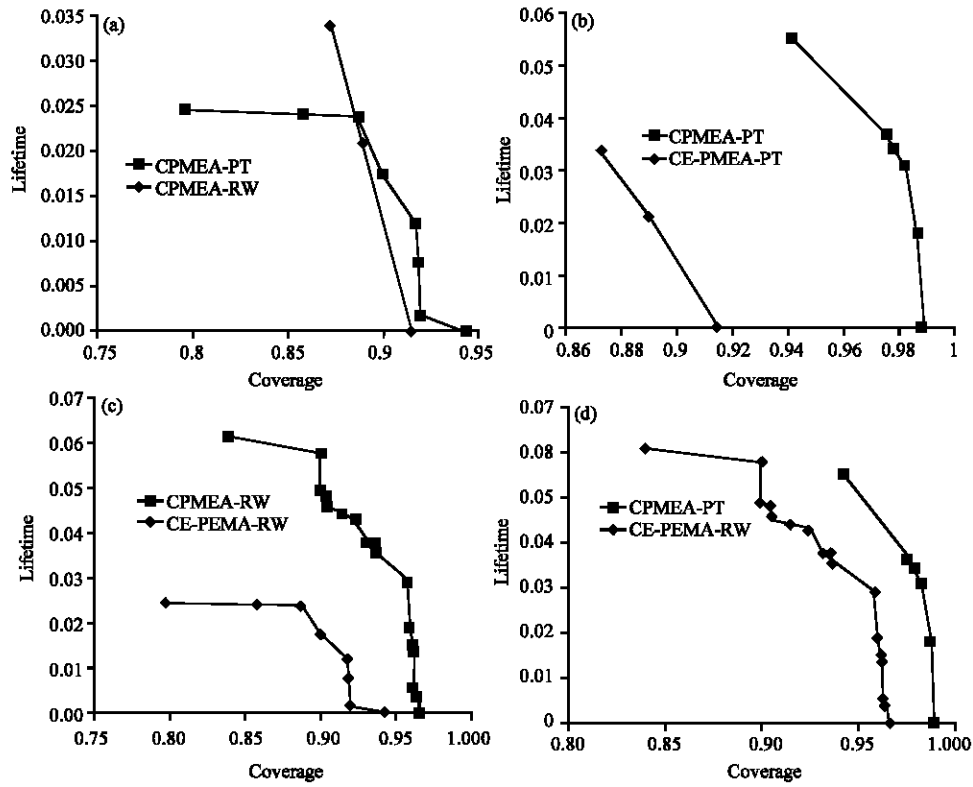


Fig. 7: a-d) PFs of the comparisons between the approaches without dead zone for P2

are compared. The PFs for the two approaches have been generated. Figure 6-c shows the Pfs for P1 and Fig. 7-c shows the Pfs for P2. Both of them shows that CE-PMEA-RW is more optimal for the lifetime and the coverage with respect to the previous approach.

CE-PMEA-PT vs. CE-PMEA-RW: We compare the performances of our computationally efficient approach with PT and with RW for checking which of them is superior on another. Figure 6d shows the PFs for P1. The PF generated by PT is more optimal for the lifetime than for the coverage. By contrast, the PF generated by RW is more optimal with respect to the coverage than to the lifetime. In Fig. 7d, the PF generated by PT is more optimal for the coverage than lifetime and there is no superiority of one of them over another.

Dead zones in ROI: This set of comparisons is conducted with the assumption of a retracted area intersects with the communication area that lead to the dead zones in ROI.

CPMEA-RW vs. CPMEA-PT: We implement the previous approach in the ROI with a dead zone in which the sensor nodes cannot be placed. The PFs for the two approaches

have been generated. Figure 8a shows the PFs for P1. PFs show that the previous approach fails to provide feasible solutions because of its lack of the practical awareness achieved by incorporating the constraints of preventing the dead zone. Figure 9-a shows the PFs for deferent parameter settings.

CPMEA-PT vs. CE-PA-PMEA-PT: The previous approach and our practically aware approach CE-PA-PMEA, both with PT are compared. Figure 8b shows the PFs for P1. We can see that the PF of the previous approach does not appear in this case, thereby indicating that the previous approach fails to provide a solution when a dead zone exists in ROI. Figure 8b show that the PF of CE-PA-PMEA-PT outperform the CPMEA-PT for the two objectives in with respect to P2.

CPMEA-RW vs. CE-PA-PMEA-RW: The previous approach and CE-PA-PMEA, both with RW are compared. The PFs of CPMEA-RW and CE-PA-PMEA-RW are seen in Fig. 8c. The PF for our approach dominates that of the previous approach for the lifetime and the coverage. Figure 9c also shows that CE-PA PMEA-RW outperform the CPMEA-RW for the two objectives.

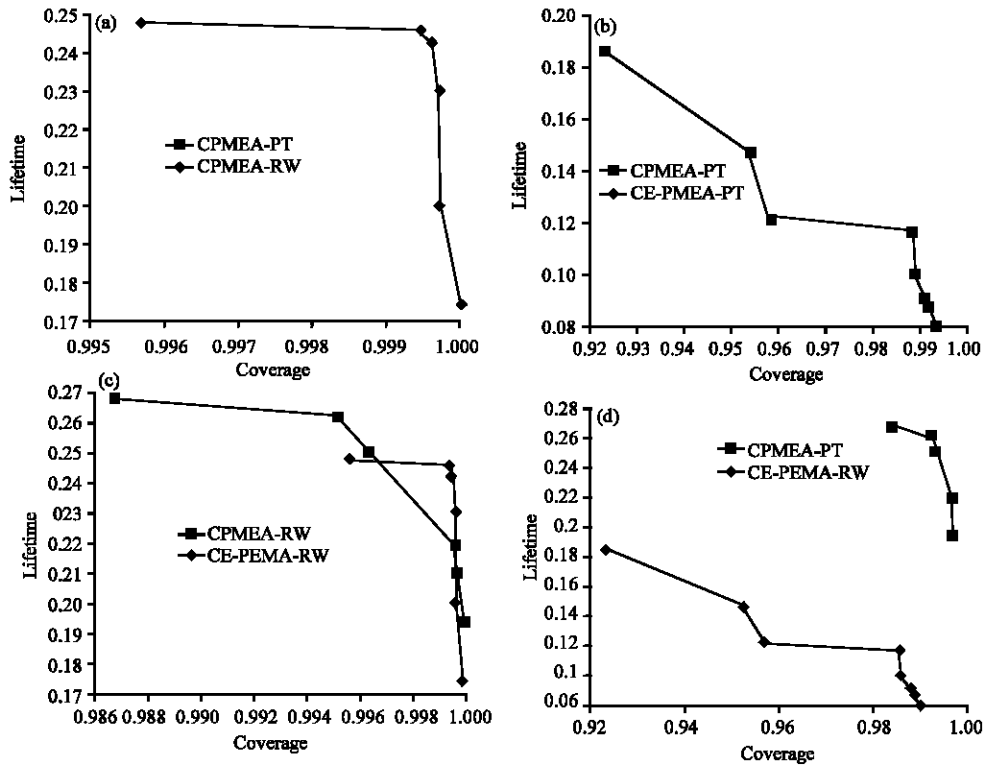


Fig. 8: a-d) PFs of the comparisons between the approaches with presence of dead zone area for P1

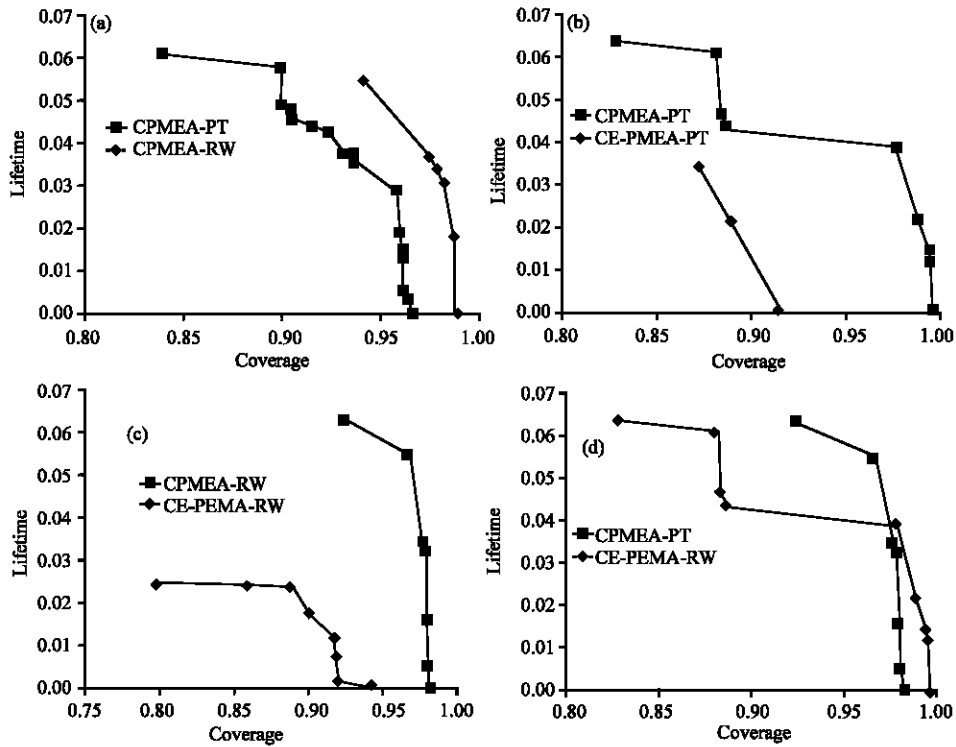


Fig. 9: a-d) PFs of the comparisons between the approaches with presence of dead zone area for P2

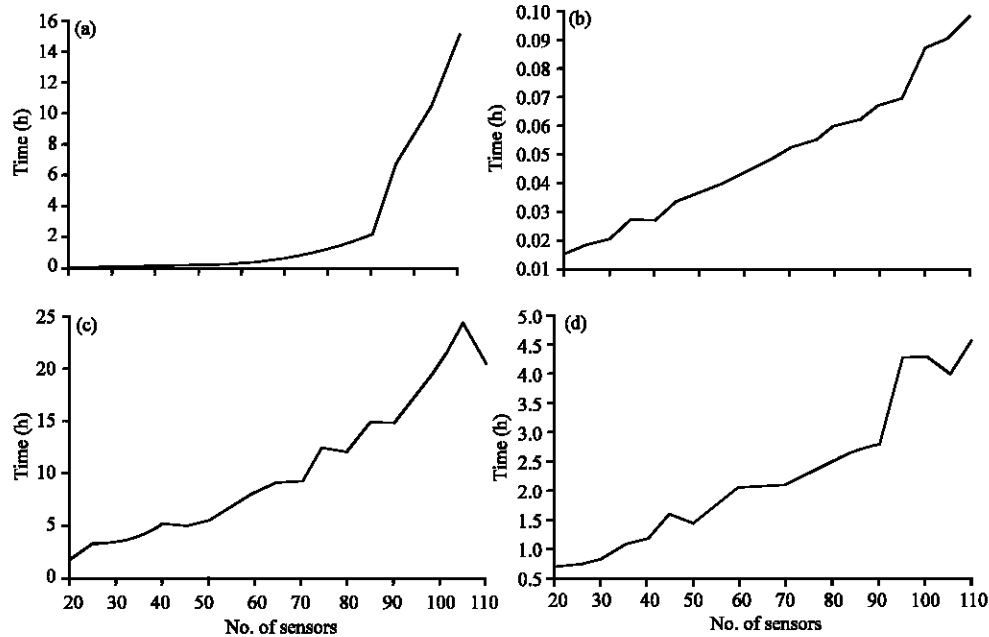


Fig. 10: a- d) The execution time for initialization and crossover for CPMEA and CE-PA-PMEA

CE-PA-PMEA-PT vs. CE-PA-PMEA-RW: Finally, we compare the performances of our approach CE-PA-PMEA with PT and with RW. Figure 8-d shows that PT has a better achievement with respect to lifetime and coverage than RW for P1. Also, Fig. 9d shows that with respect for P2.

Computational time comparisons: A comparison regarding the computational time of the initialization operation and the crossover operation for both previous and proposed approaches was implemented to check the computational complexity, Fig. 10 shows that execution time for the proposed approach outperform that of the previous approach, this prove that the proposed approach is computationally effective.

CONCLUSION

In this study, an improvement of the CPMEA by Khalesian and Delavar has been proposed to maximize coverage and minimize energy consumption that lead to prolong the lifetime of the network. The improvement is concentrated on two aspects: computational performance and practical awareness of the dead zone. The former has been implemented by modifying the initialization and crossover operations and the latter has been implemented by incorporating an additional constraint in the algorithm. Results show the superiority of our developed algorithm over CPMEA with respect to the PF and the two objectives of deployment, lifetime and coverage.

ACKNOWLEDGEMENT

I would like to thank all those whose assistance proved to be a milestone in the accomplishment of my end goal.

REFERENCES

Abdollahzadeh, S. and N.J. Navimipour, 2016. Deployment strategies in the wireless sensor network: A comprehensive review. *Comput. Commun.*, 91: 1-16.
 Debnath, S., A.K. Singh and A. Hossain, 2016. A comprehensive survey of coverage problem and efficient sensor deployment strategies in wireless sensor networks. *Indian J. Sci. Technol.*, 9: 1-6.
 Jameii, S.M., K. Faez and M. Dehghan, 2015. Multiobjective optimization for topology and coverage control in wireless sensor networks. *Intl. J. Distrib. Sens. Netw.*, 11: 363815-363825.
 Jourdan, D.B. and O.L. de Weck, 2004. Layout optimization for a wireless sensor network using a multi-objective genetic algorithm. *Proceedings of the 59th Vehicular Technology Conferences, Volume 5, May 17-19, 2004, Milan, Italy*, pp: 2466-2470.
 Juul, J.P., O. Green and R.H. Jacobsen, 2015. Deployment of wireless sensor networks in crop storages. *Wirel. Pers. Commun.*, 81: 1437-1454.

- Khalesian, M. and M.R. Delavar, 2016. Wireless sensors deployment optimization using a constrained Pareto-based multi-objective evolutionary approach. *Eng. Appl. Artif. Intell.*, 53: 126-139.
- Liu, X. and D. He, 2014. Ant colony optimization with greedy migration mechanism for node deployment in wireless sensor networks. *J. Netw. Comput. Appl.*, 39: 310-318.
- Njoya, A.N., W. Abdou, A. Dipanda and E. Tonye, 2016. Optimization of sensor deployment using multi-objective evolutionary algorithms. *J. Reliable Intell. Environ.*, 2: 209-220.
- Sengupta, S., S. Das, M.D. Nasir and B.K. Panigrahi, 2013. Multi-objective node deployment in WSNs: In search of an optimal trade-off among coverage, lifetime, energy consumption and connectivity. *Eng. Appl. Artif. Intell.*, 26: 405-416.
- Syarif, A., A. Abouaissa, L. Idoumghar, R.F. Sari and P. Lorenz, 2014. Performance analysis of evolutionary multi-objective based approach for deployment of wireless sensor network with the presence of fixed obstacles. *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, December 8-12, 2014, IEEE, Austin, Texas, USA., ISBN:978-1-4799-3512-3, pp: 1-6.
- Syarif, A., I. Benyahia, A. Abouaissa, L. Idoumghar and R.F. Sari *et al.*, 2014. Evolutionary multi-objective based approach for wireless sensor network deployment. *Proceedings of the IEEE International Conference on Communications (ICC)*, June 10-14, 2014, IEEE, Sydney, NSW, Australia, ISBN: 978-1-4799-2003-7, pp: 1831-1836.
- Tsai, C.W., P.W. Tsai, J.S. Pan and H.C. Chao, 2015. Metaheuristics for the deployment problem of WSN: A review. *Microprocess. Microsyst.*, 39: 1305-1317.
- Wahab, A., F.A. Mustika, R.B. Bahaweres, D. Setiawan and M. Alaydrus, 2016. Energy efficiency and loss of transmission data on wireless sensor network with obstacle. *Proceedings of the 10th International Conference on Telecommunication Systems Services and Applications (TSSA)*, October 6-7, 2016, IEEE, Denpasar, Indonesia, ISBN:978-1-5090-5171-7, pp: 1-6.
- Wang, W.T. and K.F. Ssu, 2013. Obstacle detection and estimation in wireless sensor networks. *Comput. Netw.*, 57: 858-868.
- Yick, J., B. Mukherjee and D. Ghosal, 2008. Wireless sensor network survey. *J. Comput. Network*, 52: 2292-2330.
- Zorlu, O. and O.K. Sahingoz, 2016. Increasing the coverage of homogeneous wireless sensor network by genetic algorithm based deployment. *Proceedings of the 6th International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, July 21-23, 2016, IEEE, Konya, Turkey, ISBN:978-1-4673-9610-3, pp: 109-114

Portfolio Selection Optimization Which Involves Minimum Transaction Lot and Transaction Cost Using Rank Dependent Expected Utility

Pasrun Adam and La Gubu
Department of Mathematics, Kampus Bumi Tridharma Anduonohu, Universitas Halu Oleo,
93232 Kendari, Indonesia

Abstract: The purpose of this research was to develop a model of a portfolio selection optimization problem involving transaction lot and transaction cost in accordance with conditions in the Indonesia Stock Exchange. In the development of this model, the rank dependent expected utility theory was used. From the simulation result using the daily data of stock price, it was found that the behaviour of the objective function $G(w)$ decreases, if parameter value a and b increase.

Key words: Portfolio selection, RDEU theory, quadratic utility function, transaction lot, transaction cost, decreases

INTRODUCTION

A lot represents a standard quantity of financial instruments established by a financial institution or stock exchange. In Indonesia, the institution that sets the lot of transaction is set by the Indonesia Stock Exchange. In stock trading, the transaction lot specifies the minimum number of shares that can be traded. So, when investors buy stock instruments, the unit of transaction used is the lot.

On the other hand, transaction cost is these that investors must pay when they buy or sell equity. Institutions that charge this transaction fee are brokers or financial intermediaries appointed by the government. The transaction cost will weigh on the investors, so, it will reduce their profits.

Mansini *et al.* (2015) divided the transaction cost structure in several types or models, namely fixed transaction cost, proportional transaction cost, convex piecewise linear cost, concave piecewise linear cost and linear cost with minimum charge. In addition, other researchers used different types of transaction costs, for example concave transaction cost (Xue *et al.*, 2006; Gao *et al.*, 2009). In the calculation of transaction costs, the use of these types of transaction costs gives a different impact on the magnitude of the profits of investors in investing.

Nowadays, the development of the model of portfolio selection problem has been done by many researchers. However, they are still rarely involving transaction lot or/and transaction cost in their models. In general, a

mean-variance model is widely used in the development of their models, for example (Xue *et al.*, 2006; Lin *et al.*, 2005).

This study aimed to build a model of portfolio selection optimization problem involving transaction lot and proportional transaction cost. The proportional transaction cost is involved in the model because it has been used on practical issues in Indonesia. The theory used in the development of this model is the theory of Rank-Dependent Expected Utility (RDEU) developed by Quiggin (1982, 1993). Thus, the main contribution of this study is the finding of a portfolio selection optimization model involving transaction lot and transaction cost that can be beneficial to investors in their portfolio diversification.

MATERIALS AND METHODS

In this research, the development of model of portfolio selection problem using the RDEU theory proposed by Quiggin (1982, 1993). This model uses three types of functions namely, utility function, cumulative distribution function and probability weight function. There are many types of utility functions that have been used by previous researchers to analyze the portfolio selection including: linear utility function (Cenci and Fillipine, 2006; Gao *et al.*, 2010), quadratic utility function (Rahadi *et al.*, 2015; Adam and Gubu, 2017; Koo *et al.*, 2016), isoelastic form utility function (Bahaji and Casta, 2016; Sims, 2015) and logarithmic utility function (Xuan and Yue, 2011; Barrachina *et al.*, 2012).

Suppose that financial markets have uncertain events in the probability space (Ω, \mathcal{f}, p) where, Ω is the set of universal events, \mathcal{f} is a field or set collection of events and P is a probability measure defined in subcollections of the set of events E of \mathcal{f} such that $0 \leq p(E) \leq 1$. Assume that X is a random variable defined in the probability space (Ω, \mathcal{f}, p) . This random variable has a cumulative distribution function $F_x(x) = \text{pr}(x \leq x)$, $x \in (-\infty, +\infty)$. The function $F_x(x)$ is an increasing function such that $F_x(-\infty) = 0$ and $F_x(+\infty) = 1$.

Furthermore, the RDEU theory according to Quiggin (1982, 1993) and had been referenced by Adam and Gubu (2017) is as follows. Suppose that the random variable X has the values $x_i (i = 1, 2, \dots, n)$ with $x_1 \leq x_2 \leq \dots, x_n \leq$ and has a probability measure p_1, p_2, \dots, p_n . The formulation of discrete form of the RDEU theory is:

$$\text{RDEU}(X) = \sum_{i=1}^n U(x_i) \left[g\left(\sum_{j=1}^i p_j\right) - g\left(\sum_{j=1}^{i-1} p_j\right) \right] \quad (1)$$

where, $U: (0, \infty) \rightarrow (-\infty, +\infty)$ is a utility function of investor preferences in the portfolio selection. The function $U(x)$ is an increasing function with $U(0) = 0$. The function $g: [0, 1] \rightarrow [0, 1]$ is called the probability weight function. This function is an increasing function with $g(0) = 0$ and $g(1) = 1$. Equation 1 can be expressed in the form of Lebesgue-Stieltjes integral:

$$\text{RDEU}(X) = \int_{\Omega} U(x) dg(F_x(x)) \quad (2)$$

If the cumulative distribution function $F_x(x)$ has a derivative $f(x) = F'_x(x)$ and the functions $U(x)$ and $g(x)$ are continuous, then Eq. 2 can be expressed in the form of Riemann integral:

$$\text{RDEU}(X) = \int_{-\infty}^{+\infty} U(x) g'(F_x(x)) f(x) dx \quad (3)$$

The function $f(x)$ is a density function of the random variable X . Next, let \succeq be a relation of investor preference to select an asset or a portfolio. Suppose that there are two random variables X and Y with their cumulative distribution functions are F_x and F_y . Then $X \succeq Y$ if and only if $\text{RDEU}(X) \geq \text{RDEU}(Y)$. In this case, the investor will choose X , if the $\text{RDEU}(X)$ value is greater than the $\text{RDEU}(Y)$ value or the investor can choose either X or Y if $\text{RDEU}(X) = \text{RDEU}(Y)$.

Adam and Gubu (2017) developed the RDEU Model in which the utility function is a quadratic function:

$$U(x) = x - \alpha x^2, 0 \leq x < \frac{1}{2\alpha}, 0 < \alpha < 1$$

And the probability weight function $g(x)$ is:

$$g(x) = \begin{cases} bx, & \text{if } 0 \leq x \leq \text{Pr}(X \leq v) \\ cx+d, & \text{if } \text{Pr}(X \leq v) < x \leq 1 \end{cases} \quad b > c > 0$$

The function $g(x)$ is a concave piecewise linear function and continuous on $[0, 1]$. Further, it is assumed that the random variable X is normally distributed with the probability function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), x \in (-\infty, +\infty)$$

where, $m = E(x)$ is the mean and σ is standard deviation of X . With this assumption and the monotonic, concavity and continuity properties of the function $g(x)$, it is obtained $0 < c < 1 < b < 2$. If the value of a is close to 0, Adam and Gubu (2017) had shown that the approximate value of $\text{RDEU}(x)$ is:

$$\text{RDEU}(X) = -\alpha\sigma^2 + \frac{(2-2b)(1-2\alpha m)\sigma}{\sqrt{2\pi}} + (m - \alpha m^2) \quad (4)$$

The investor portfolio has a weight vector $w = (w_1, w_2, \dots, w_n)$ with $\sum_{i=1}^n w_i = 1$. Furthermore, the returns of equity R_i , $(i = 1, 2, \dots, n)$ are mutually independent and normally distributed. Then, the random variable return of the portfolio $R_p = \sum_{i=1}^n w_i R_i$ is normally distributed which has the expected return $E(R_p) = \sum_{i=1}^n w_i E(R_i)$ and the standard deviation:

$$\sigma_p = \left(\sum_{i=1}^n w_i w_j \sigma_i^2\right)^{\frac{1}{2}}$$

Cenci and Filippini (2006) provided a formula for risk based on the RDEU theory, namely $\varphi(X) = U(E(X)) - \text{RDEU}(X)$. By using this formula for $X = R_p$ and Eq. 4, then the portfolio standard risk approximation value (Adam and Gubu, 2017) of the investor is:

$$\varphi(R_p) = \alpha\sigma_p^2 + \frac{(2b-2)(1-2\alpha E(R_p))\sigma_p}{\sqrt{2\pi}} \quad (5)$$

where, $0 < \alpha < 1$ and $1 < b < 2$. In the results study, Eq. 5 is used to develop a model of objective function of the portfolio selection optimization problem involving transaction lot and transaction cost and its

constraints. Furthermore, a simulation is performed to evaluate the behaviour of the objective function. The stock price data used in this simulation is daily stock price data from ten companies where the ten stocks of the companies are in the LQ45 stock group in the Indonesia Stock Exchange. The daily stock price data was selected from the closing price that spans from June, 20, 2016-June, 16, 2017.

RESULTS AND DISCUSSION

Suppose that $M(M_0 \leq M \leq M_1)$ represents the amount of money invested in the securities of the stock and let k_i is a number of shares of the securities $I(i = 1, 2, \dots, n)$. Thus, the investors have a portfolio. Suppose that n_i is a minimum lot of transaction purchased by the investor for securities i , then K_i is a multiple of $n_i (i = 1, 2, \dots, n)$. Next, let also o_i is a number of shares per lot for securities $(i = 1, 2, \dots, n)$. Thus, we obtain $k_i = n_i o_i, (i = 1, 2, \dots, n)$. If p_i is a price of each share, then the total purchase price of each securities i is:

$$M_i = k_i p_i = n_i o_i p_i, (i = 1, 2, \dots, n) \tag{6}$$

Based on Eq. 6, the total purchase price of n type of securities (or amount of initial wealth) of the investor is:

$$M = \sum_{i=1}^n M_i = \sum_{i=1}^n k_i p_i = \sum_{i=1}^n n_i o_i p_i \tag{7}$$

where, $M_0 \leq M \leq M_1, M_0$ is a lower bound of investment and M_1 is an upper bound of investment. Further, let $\alpha (0 < \alpha < 1)$ is the proportional transaction cost or fee to be paid by the investor to the broker for the purchase of the security. The amount of commission received by the broker is:

$$\alpha M = \sum_{i=1}^n \alpha k_i p_i = \sum_{i=1}^n \alpha n_i o_i p_i \tag{8}$$

The random variable of return of the securities i is R_i with the expected return $(R_i), (i = 1, 2, \dots, n)$. If the weight of each securities in the portfolio W is $w_i = k_i p_i / M = n_i o_i p_i / M (i = 1, 2, \dots, n)$ and R_p^{wic} is the return of portfolio W which involves the transaction lot and proportional transaction cost, then from Eq. 7 and 8 the expected return of portfolio W with weight vector $w = (w_1, w_2, \dots, w_n)$ is:

$$E(R_p^{wic}) = \sum_{i=1}^n w_i E(R_i) - \frac{\sum_{i=1}^n \alpha n_i o_i p_i}{M} = \sum_{i=1}^n \frac{n_i o_i p_i}{M} E(R_i) - \alpha \tag{9}$$

In accordance with investment theory and practical review, the investors usually want the return from their investment to be maximum, so, the value of $E(R_p^{wic})$ in Eq. 9 is maximum. The problem of portfolio selection optimization is the combination of Eq. 5 and 9 as follows:

$$\text{Min } \varphi(R_p) = \text{Min} \left(\alpha \sigma_p^2 + \frac{(2b-2)(1-2\alpha E(R_p)) \sigma_p}{\sqrt{2\pi}} \right) \tag{10}$$

$$\text{Maks } E(R_p^{wic}) = \text{Maks} \left(\sum_{i=1}^n \frac{n_i o_i p_i}{M} E(R_i) - \alpha \right) \tag{11}$$

$$M_0 \leq M \leq M_1 \tag{12}$$

$$\sum_{i=1}^n w_i = \sum_{i=1}^n \frac{n_i o_i p_i}{M} = 1 \tag{13}$$

$$w_i \leq 0, (i = 1, 2, \dots, n) \tag{14}$$

By using the concept of a multiobjective function, then the portfolio selection objective function which involves transaction lot and transaction cost is:

$$G(w) = -\beta \left(\alpha \sigma_p^2 + \frac{(2b-2)(1-2\alpha E(R_p)) \sigma_p}{\sqrt{2\pi}} \right) + (1-\beta) \left(\sum_{i=1}^n \frac{n_i o_i p_i}{M} E(R_i) - \alpha \right) \tag{15}$$

where, $0 < \alpha < 1, 1 < b < 2$ and $0 < \beta < 1$. The value of the objective function $G(w)$ at Eq. 11 must be maximum with constraints (Eq. 12-14). Variance σ_p^2 and expected return $E(R_p)$ at Eq. 11 are:

$$E(R_p) = \sum_{i=1}^n \frac{n_i o_i p_i}{M} E(R_i)$$

And:

$$\alpha_p^2 = \sum_{i=1}^n \sum_{j=1}^n \frac{n_i o_i p_i}{M} \frac{n_j o_j p_j}{M} \sigma_{ij}^2$$

The objective function in Eq. 11 has the assumption that the investors only make one purchase. The behaviour of the objective function $G(w)$ is evaluated by selecting the values of a and b for a vector weighted portfolio $w = (w_1, w_2, \dots, w_n)$ through a simulation. In this simulation the data used is daily stock price data from ten companies where the ten stock companies are listed in the Indonesia Stock

Table 1: The company names, stock code, purchase price and expected return

Company name	Stock code	Purchase price in rupiah currency (p _i)	Expected return E(R _i)
Astra Agro lestari Tbk	AALI	16000	-0.001050
Adhi Karya (Persero) Tbk	ADHI	16000	0.000879
Adaro Energy Tbk	ADRO	1750	-0.000729
AKR Corporindo Tbk	AKRA	6650	-0.000382
Aneka Tambang (Persero) Tbk	ANTM	810	-0.001016
Astra International Tbk	ASII	8050	0.001009
Alam Sutera Realty Tbk	ASRI	386	-0.001942
Bank Central Asia Tbk	BBCA	15425	0.001685
Bank Negara Indonesia (Persero) Tbk	BBNI	5950	0.001098
Bank Rakyat Indonesia (Persero) Tbk	BBRI	11950	0.002381

Table 2: The values of portfolio selection objective function

Values	Variables	G(w)
a = 0.1	b = 1.1	0.005324
a = 0.3	b = 1.3	0.005319
a = 0.5	b = 1.5	0.005314
a = 0.7	b = 1.7	0.005309
a = 0.9	b = 1.9	0.005304

Exchange. These stocks are included in the LQ45 stock group. The stock code, the purchase price and the stock return and the names of the ten companies are shown in Table 1.

The Indonesia Stock Exchange has determined the number of shares per lot in a transaction, namely 100 shares per lot, so $n_1 = n_2 = \dots = n_p = 100$. Transaction cost of purchase is 0.2%, so, $\alpha = 0.0002$. Suppose that the initial investment is IDR 10,000,000, so, $M = 10,000,000$. The result of data processing is shown in Table 2.

The objective function values $G(w)$ for some values a and b as shown in Table 2, it seems that the value of $G(w)$ is decreased if the values of a and b are greater towards its the upper bound. This is due to the declining value of $-\varphi(R_p)$.

The decreasing value of $-\varphi(R_p)$ in Eq. 11 is due to the rise in the value of $-\varphi(R_p)$ in Eq. 5 if values a and b rise (Adam and Gubu, 2017). Furthermore, the results of this study differ from those of Lin *et al.* (2005) and Gao *et al.* (2009). These differences are due to the fact that in the model built by them, the lot is used as a unit of portfolio vector elements. Meanwhile, in this study, the units of portfolio vector elements used are the number of shares.

CONCLUSION

This study aimed to develop a model of optimization problem of portfolio selection using the RDEU theory involving transaction lot and transaction cost. In the development of this model, it is assumed that the investor only makes one purchase. The transaction cost to be paid by the investor only at the time of purchase. The

simulation result shows that the bigger the value of a and b , the smaller the expected return value of the investor.

RECOMMENDATIONS

This study does not consider the cost of sale transaction in the model. In the next study, the model development can consider the cost of sale transaction. Because in fact, the investors can buy shares more than once in a certain time period, so in the next study, the payment of transaction cost more than once can also be considered in the development of the model.

ACKNOWLEDGEMENTS

This research article is our research results with contract No. 65/UN29.20/PPM/2017 which is supported by The Directorate of Research and Community Service, General Directorate of Research Strengthening and Development, Ministry of Research, Technology and Higher Education.

REFERENCES

- Adam, P. and L. Gubu, 2017. The portfolio standard risk model based on rank dependent expected utility model with quadratic utility function. *Global J. Pure Appl. Math.*, 13: 3801-3810.
- Bahaji, H. and J.F. Casta, 2016. Employee stock option-implied risk attitude under rank-dependent expected utility. *Econ. Modell.*, 52: 144-154.
- Barrachina, A., G. Rubio and A. Urbano, 2012. Multiplicity in financial equilibrium with portfolio constraints under the generalized logarithmic utility model. *Spanish Rev. Financial Econ.*, 10: 41-52.
- Cenci, M. and F. Filippini, 2006. Portfolio selection: A linear approach with dual expected utility. *Appl. Math. Comput.*, 179: 523-534.
- Gao, Y., B. Wang, X. Quan and J. Zhou, 2010. A new multi-objective portfolio optimization model based on dual expected utility. *Proceedings of the IEEE 5th International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA'10)*, September 23-26, 2010, IEEE, Changsha, China, ISBN:978-1-4244-6437-1, pp: 793-798.
- Gao, Y., Y. Sun and Y. Li, 2009. Mean-CVaR portfolio optimization problem under concave transaction costs and minimal transaction unit constraints. *Proceedings of the 2nd International Conference on Information and Computing Science (ICIC'09) Vol. 1*, May 21-22, 2009, IEEE, Manchester, England, UK., ISBN:978-0-7695-3634-7, pp: 79-82.

- Koo, J.L., S.R. Ahn, B.L. Koo, H.K. Koo and Y.H. Shin, 2016. Optimal consumption and portfolio selection with quadratic utility and a subsistence consumption constraint. *Stochastic Anal. Appl.*, 34: 165-177.
- Lin, D., X. Liang and M. Li, 2005. A Genetic Algorithm for Solving Portfolio Optimization Problems with Transaction Costs and Minimum Transaction Lots. In: *Advances in Natural Computation ICNC*, Wang L., K. Chen and Y.S. Ong (Eds.). Springer, Berlin, Germany, ISBN:978-3-540-28320-1, pp: 808-811.
- Mansini, R., W. Ogryczak and M.G. Speranca, 2015. *Linear and Mixed Integer Programming for Portfolio Optimization*. Springer, Berlin, Germany, ISBN:978-3-319-18481-4, Pages: 115.
- Quiggin, J., 1982. A theory of anticipated utility. *J. Econ. Behav. Organiz.*, 3: 323-343.
- Quiggin, J., 1993. *Generalized Expected Utility Theory the Rank Dependent Model*. Springer, Berlin, Germany.
- Rahadi, A.P., N.A. Rizal and B.A. Surya, 2015. Execute trading policies on optimal portfolio when stochastic volatility and inflation effect were considered. *J. Eng. Appl. Sci.*, 11: 1706-1713.
- Sims, T.S., 2015. Income taxation, wealth effects and uncertainty: Portfolio adjustments with isoelastic utility and discrete probability. *Econ. Lett.*, 135: 52-54.
- Xuan, L. and Y. Hu, 2011. The portfolio investment decisions about logarithmic utility. *Proceedings of the International Conference on Multimedia Technology (ICMT)*, July 26-28, 2011, IEEE, Hangzhou, China, ISBN:978-1-61284-771-9, pp: 3746-3749.
- Xue, H.G., C.X. Xu and Z.X. Feng, 2006. Mean-variance portfolio optimal problem under concave transaction cost. *Appl. Math. Comput.*, 174: 1-12.

Hybrid Content Based Image Retrieval System Using Exhaustive Feature Set Processing by Multi Objective Optimization

¹Baddeti Syam and ²Y. Srinivasa Rao

¹Mandava Institute of Engineering and Technology, Department of ECE,
521275 Jaggayyapet, Andhra Pradesh, India

²Department of Instrument Technology, AU College of Engineering,
Andhra University, Visakhapatnam, Andhra Pradesh, India

Abstract: In real world digital image processing, Content Based Image Retrieval (CBIR) is an emerging concept, it is an image retrieval framework based on some visual semantic features like color, texture and shape to search relevant images from multiple image sources. CBIR consist single query or multiple queries to semantic data or same object for multiple class labels with reference to multi query images. In retrieval of query image comparison from multiple image sources may cause optimization problem in image retrieval because of ambiguity in image search. To solve optimization problem for efficient query image retrieval from multiple image archives, in this study, we propose and develop hybrid content based image retrieval system. This system consist pareto optimized method (for removing non-dominated features) and Genetic Algorithm (GA) for effective content based image retrieval. Integration of contourlet features with color, texture and shape is considered to build exhaustive feature set for all images. In our approach Pareto optimal solution consist multi-objective selection functions to process image retrieval from multiple image sources for individual feature selection procedure with query image evaluation for image retrieval. This approach gives better performance than traditional approaches in query image retrieval from multiple image archives. Our experimental results show effectiveness of the proposed method in retrieving all images from image sources with improvement sorting the individual images in searching Pareto solutions in multi-objective optimization problem. Furthermore, this approach gives better accuracy with respect to precision, recall and time efficiency in image retrieval from multiple image sources based on extensive visual features.

Key words: Content based image retrieval, Genetic algorithm, Pareto solution, multi-objective selection functions, squared euclidean distance and visual features, better, image

INTRODUCTION

In earliest days, CBIR has been aggressive concept in machine learning and information or image retrieval applications. Several image retrieval approaches for multi query image selection have been introduced in previous literature. In most of the systems, image picture is correspond to same semantic image with same concept but the background is multiple for multiple sources but continuously have multiple image background with multiple objects in image with alternative angles between same class with multiple objects. The thought is that by using numerous questions of a similar protest, the execution of single-inquiry recovery can be made strides. Various query image retrieval techniques are sorted based on numerous single query processes (Hsiao *et al.*, 2015). A considerable lot of the strategies for single-semantic-various inquiry recovery include joining

the low-level elements from the query pictures to create a solitary arrived at the midpoint of inquiry. In this study, we consider the all the more difficult issue of discovering pictures that are applicable to various inquiries that speak to various picture semantics (Ramamurthy and Chandran, 2011). For this situation, the objective is to discover pictures containing applicable components from each furthermore, every query. Since, the inquiries relate to various semantics, alluring pictures will contain highlights from a few particular pictures, and won't really be firmly identified with any individual query. This makes the issue on a very basic level not quite the same as single query retrieveal and from single-semantic numerous inquiry recoveries. For this situation, the inquiry pictures won't have comparable low-level components and shaping an arrived at the midpoint of query is not as valuable. Since, pertinent pictures don't really have includes firmly lined up with a specific inquiry, a significant number of the

standard recovery procedures are not valuable in this unique situation. For instance, pack of-words sort approaches which may appear to be normal for this issue, require the objective picture to be firmly identified with a few of the inquiries. Another basic method is to include each inquiry each one in turn and normal the subsequent likenesses. This tends to deliver pictures firmly identified with one of the inquiries, however once in a while identified with at the same time. Numerous other multiple inquiry recovery calculations are planned particularly for the semantic single query problem to identify related images with just a single or a couple of the inquiries.

Various query recovery is identified with the meta search issue in software engineering. In meta search, the issue is to join indexed lists for a similar inquiry over multiple web search tools. This is like the single-semantic-multiple query issue as in each web index is issuing a similar inquiry (or semantic). Consequently, meta search calculations are not reasonable with regards to multiple query recovery with a few particular semantics.

So, in this study, we propose to develop a novel approach, i.e., hybrid content based image retrieval system which combines Pareto Optimization Algorithm (POA) and Genetic algorithm. In this approach, POA retrieve each image and rank them individually with their discriminative to the input query. Some of the techniques like SIFT and HoG has been proposed in different computer vision oriented applications can be utilized to define image dissimilarities. These dissimilarities explore query retrieval in image achieves, we utilize a quick positioning calculation called Manifold Efficient Ranking (MER) to calculate the ranking without consider total presented samples sets. MER can productively find the basic dimension of given image database and find the evaluation time of the conventional ranking system. MER effectively defines single query image retrieval to define various retrieval query issue.

The next stage in our hybrid approach is to utilize the positioning delivered by MER to make Pareto focuses,

which compare to non-similarities between a specimen and each inquiry. Pareto optimal points are called Pareto fronts to define non-dominated events Pareto front mainly focuses on non-dominated focuses and it is called origin in group image databases. The second Pareto front (profundity two) is acquired by expelling the principal Pareto front and finding the non-overwhelmed focuses among the rest of the Pareto front algorithm contains specimens for each user and those samples are exhausted. Pareto front sort the non-dominated class references.

Key distribution of Pareto front for query image from multiple image databases shown in Fig. 1 to recovery query images based on forest and mountain keywords. In that image are listed based on their positioning with front and connection from tail to other present in image retrieval.

However, as seen in Fig. 1, images (images 10, 11, 12) in center of front optimization contain related functions for both queries suitable for multi image retrieval problem. Our proposed approach is an well known approach to provide symmetric calculation with respect to query image retrieval from multiple image class retrieval. We also applicable our approach to automatic image annotation of large databases. The main contribution of our proposed approach is as follows: Propose the hybrid CBIR for query image retrieval from multiple image archives based on visual features of query image. Discuss about Pareto Optimization Method (POM) for image annotation in multiple image retrieval. For effective similarity measure between query image and data base image with extensive feature presentation is explained using genetic algorithm for query image retrieval from multiple image archives.

Literature review: Quick object retrieval using direct spatial matching paper (Syam and Rao, 2010) appears for the substantial scale question recovery Bag of visual-Words (BOW) demonstrates is prevalent however there is a disadvantage it disregards spatial data. Direct

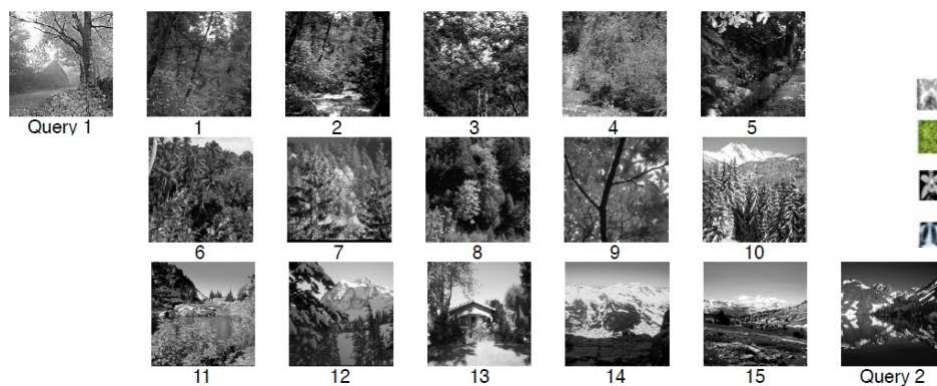


Fig. 1: Images located based on Pareto front method with middle of semantic query images

Spatial Matching (DSM) way to deal with discover the scale variety utilizing district sizes in which all component matches for evaluating geometric transformation. DSM is significantly snappier than RANSAC-based techniques and thorough specification approaches. Alogarithmic term recurrence converse archive recurrence weighting plan is acquainted with increase the execution of the base framework. The disadvantage is a high cost because of high recovery exactness. Non-overwhelmed arranging (Syam and Srinivasa Rao, 2012; Syam *et al.*, 2013; Arandjelovic and Zisserman, 2012; Binitha and Sathya, 2012; Martens *et al.*, 2011) is a critical issue in multi-target advancement. They have discovered that in the extensive specimen estimate constrain, the non-commanded fronts merge most likely to the level arrangement of a capacity that fulfills Hamilton-Jacobi Partial Differential Equations (PDE). To plan for a quick, possibly sub straight, surmised non-ruled arranging calculation PDE is utilized and the aftereffects of applying the calculation to genuine information (Rashedi *et al.*, 2013). They likewise give a quick numerical plan for tackling the PDE and utilize it to build up a quick, possibly sub straight, estimated non-ruled arranging calculation. This calculation is connected to genuine information from an oddity location issue and amazing exactness is accomplished by this calculation while fundamentally decreasing the computational multifaceted nature of non-dominated arranging. For huge information gushing issues, this calculation is utilized which include steady re arranging of substantial datasets upon the landing of new examples. For the straight arrangement of highlights, it is reasonable.

The expansive scale particular question picture recovery frameworks are to discover pictures that contain the inquiry protest in the picture database. For instance media, generation groups are keen on seeking pictures or video film to go with news reports and daily paper articles from the database. Get different pictures of the questioned protest utilizing Google picture seek. Demonstrate that various questions altogether enhance the framework to discover testing events of the questioned protest. The point is to recover all pictures containing a particular protest in a picture informational collection. This is an issue that has seen much improvement and accomplishment in the course of the most recent decade, the beginning stage of the inquiry has been a solitary question picture of the particular protest of intrigue. In this work two changes to the standard approach: to begin with, beginning stage the protest is content, as they are keen on testing informational indexes to discover known questions and second the informational index utilizing numerous picture inquiries and join the outcomes into a solitary positioned list. Tending to this issue has been one of the principle inquire about subjects in particular protest recovery look into with advancements in include

encoding to lighten Vector Quantization (VQ) misfortunes and in growth of the sack of visual word (BOW) portrayal to mitigate locator and descriptor drop out (and in addition, once more, VQ losses). All time the change of printed is impractical for every one of the pictures. In this way, tree-based lists are not attractive for high-dimensional pursuit issues. Another sort of ANN look calculation depends on the vector quantization for example, k-implies LSH (Martens *et al.*, 2011) and product quantization (AB) (Rashedi *et al.*, 2013). In AB, by isolating every datum into a few sub spaces and communicating information as far as repeating parts, the illustrative limit of AB, becomes forceful in the quantity of sub spaces. In these strategies, every datum point is spoken to by an element vector and the inquiry procedure is performed. Subsequently, the hunting procedure is tedious. As of late, the hashing based technique has been generally utilized for closeness seek and comparative applications as it permits the steady time look. A considerable measure of hashing techniques has been proposed these can be isolated into two fundamental classifications: information free strategies (Syam and Rao, 2010; Quellec *et al.*, 2010) and information subordinate strategies. Hypothetically, it is ensured that the first separations are asymptotically spared in the Hamming space inside wrinkling code length; thus, LSH related strategies more often than not require long codes to accomplish great exactness. In any case, long codes result in the impact likelihood of comparable focuses mapped to close codes diminishes as the code length increments. Subsequently, the LSH-related techniques, for the most part, build multi-tables to guarantee a sensible likelihood that an inquiry will crash into its close neighbors in no less than one table which prompts long question time and expands the memory occupation. As of late, numerous information subordinate techniques which concentrate on taking in hash capacities from the informational index, have been produced to take in more minimized codes.

The area connections between sub-locales and the relative vicinities between neighbors of every sub region are both basic (Ruiz-Vanoye *et al.*, 2012) for viable measurement decrease. Many hashing strategies have been created for neighborhood-protecting (Syam and Srinivasa Rao, 2012) or separation saving (Mahant *et al.*, 2012). The previous one sign or the area rankings and later to help and a complex with unbending steel pillar (Ruiz-Vanoye *et al.*, 2012) while as a rule, the ideal inserting of a complex needs some adaptability: some sub-areas ought to be privately extended or contracted to install the min to alow-dimensional space where the topology can be very much saved. In this manner, these hashing strategies can't safeguard the informational index topology well.

In complex learning region, the reality of the matter is that a complex can be totally described by giving the

relative or similar vicinities (Ruiz-Vanoye *et al.*, 2012), e.g., a first area is near a moment one yet a long way from a third one. Near in arrangement between separations, as in correspondences or rankings, get the job done to describe a complex for any embeddings. In addition, for some comparability seek issues, the relative rankings of results are more vital than their genuine similitude's to a question. In light of the above presentation, we can infer that more powerful Hamming installing's can be learned by consolidating the informational index topology with the learning procedure. Assist more, if the nearby topology is all around protected in Hamming space, the vagueness caused by positioning with Hamming separation can be better all mitigated.

Syam *et al.* (2013) proposes similarity measure in genetic algorithm for efficient image retrieval with exhaustive feature presentation in image retrieval conditions based on query image processing (Shivakumar and Amudha, 2012; Rao *et al.*, 2010). In this scenario when the query image is taken from image database and then Genetic Algorithm (GA) similarity measure is performed in between query image retrieval and database image retrieval. It is also performed different query image retrieval for medical image.

Basic utilization of Pareto front calculation aggressively (Subotic, 2011; Boonrong and Kaewkamnerdpong, 2011; Sudhir and Baboo, 2011) in the machine learning group. Complex multi-target advancement issues are illuminated by this strategy, where finding the main Pareto front is testing. A few productive calculations have been created for finding the skyline strategy. In our POA calculation, efficient and quick skyline calculations or quick non-commanded arranging can be utilized to discover every Pareto front.

MATERIALS AND METHODS

Image query retrieval from multiple image archives:

For effective image retrieval, images are extracted and pre-processed based on low dimensional feature representation instead of pixel values in directing the image indexing and image retrieval (Rao *et al.*, 2010; Hashni and Amudha, 2012; Padmaa and Binoona, 2016; Binoonal and Padmaa, 2016; Singla and Garg, 2014). Some of the feature extraction and similarity extraction techniques has been proposed on various computer vision and image processing oriented applications to define best query matching with different images. In this scenario ambiguity may appear in image query retrieval from multiple image archives, we present and use manifold ranking procedure, POA and GA for effective CBIR in (Velmurugan and Baboo, 2011; Syam and Srinivasa Rao,

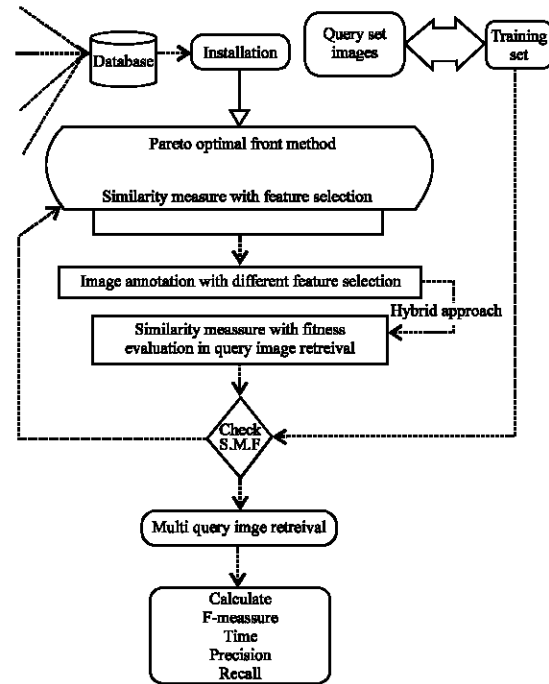


Fig. 2: Proposed implementation for multi query optimized image retrieval with Genetic algorithm

2012). Procedure of the Image query retrieval from multiple image archives using hybrid image retrieval systems may show in Fig. 2 with extensible feature presentation in image retrieval.

As discussed previous paragraph to obtain effective CBIR and verify sample query pair in image retrieval for image positioning in data base based on query image with respect to visual features (Zhong *et al.*, 2015; Calder *et al.*, 2014). Procedure for ranking manifold for image positioning with different frames as follows:

Manifold ranking procedure (MER): Let $\mathcal{N} = \{X_1, X_2, \dots, X_n\} \subset T^m$ to be feasible set of Pareto points and $d: \mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R}$ be a metric on \mathcal{N} such that call as squared euclidean distance define sequential attribute processing with following $y = \{y_1, y_2, \dots, y_n\}$ in which $y_i = 1$, if X_i is the query $y_i = 0$ otherwise. Let $r: \mathcal{X} \rightarrow \mathbb{R}$ is the ranking function to assign with image position score r_i to each point X_i . The entirely allocated a position of 1 and all other examples will be allocated more compact positions with range to the keyword along with image information. To create a graph on X , first sort the for all ranges between different examples in ascending order to link related Graph G is designed to explore image retrieval. Based on this procedure in manifold ranking approach associated with ranking vector r processed for image ranking is as follows:

$$O(r) = \sum_{i,j=1}^n w_{ij} \left| \frac{1}{\sqrt{D_{ij}}} r_i - \frac{1}{\sqrt{D_{ij}}} r_j \right|^2 + \mu \sum_{i=1}^n |r_i - y_i|^2 \quad (1)$$

where, D is the diagonal matrix with $D_{ij} = \sum_{s=1}^n w_{is}$ and $\mu > 0$ is sequential regularization parameter. The first phrase in the cost function is a level of smoothness phrase that causes close by factors have similar position ratings (Calder *et al.*, 2014; Christoudias *et al.*, 2012). The second phrase is a regularization term which causes the question to have a position near to 1 and all other examples to have positions as near to 0 as possible. Based on this procedure efficient manifold calculation is measured with different formation in final ranking is computed as follows:

$$r^* = (I_n - H^T (HH^T - \frac{1}{\alpha} I_d)^{-1} H) y \quad (2)$$

where, $H = ZD^{-1/2}$ and D is the diagonal matrix, notice construction of the anchor graph structure with matrix inversion to present non-ambiguity based sample image retrieval conditions with perspective measure of fast image retrieval.

Retrieval of query image from multiple classes of archives: Based on previous knowledge regarding relevance of keyword for efficient calculation with selection of starting image from multiple image archives in (Gonen and Alpaydin, 2011; Calder *et al.*, 2014) query image retrieval from multiple archives efficient manifold ranking to each data based image with query image. Main contribution of this study shows Pareto method for standard image retrieval. To return relevant tests to the client, we return tests as indicated by their Pareto front number, i.e., we return focuses on F_1 , to begin with, at that point F_2 , et cetera, until the point that adequately many pictures have been recovered. Inside a similar front, we return focuses in a specific request, e.g., for $T = 2$, from the center first (Calder *et al.*, 2015). With regards to this study, the center piece of each front will contain tests identified with all inquiries. Relevant input plans can likewise be utilized in our calculation to upgrade recovery execution. For efficient CBIR in query image retrieval from multiple archives apply Pareto optimal algorithm and GA in image retrieval after ranking the images with different visual semantic features.

Hybrid approach implementation for multiple query image: In this study, we deal with an application of hybrid approach based on multi objective process. Hybrid consists both Pareto optimization method and Genetic algorithm to obtain multi-objective image retrieval, for different set of datasets, single query measurement is not worked properly to efficient image retrieval. Moreover, a

combination of different actions is very challenging, because of the method for the combination may be not possible and sometimes those actions may be unsuitable. In this study, we propose Pareto optimal method to check optimization which helps to find the related objectives in information retrieval applications. We incorporate optimization approach with in GA algorithm implementation to extract multi query access from different data sets.

Pareto optimal method: Pareto optimality is an effective concept and applied in different domains like, Economic Sciences, Computer Sciences and Social Sciences. Here, we give brief review about Pareto optimality and define Pareto front. In general setting of a discrete multi-target enhancement issue, we have a limited set S of achievable arrangements and T criteria f_1, f_2, \dots, f_T for assessing the plausible arrangements. One conceivable objective is to discover x 2 S limiting all criteria at the same time. In many settings, this is an inconceivable errand. Many ways to deal with the multi-target streamlining issue diminish to joining all T criteria into one. At the point when this is finished with a direct mix, it is typically called linear scalarization. Selection of different weights in the straight mix yield distinctive minimizes. Without earlier learning of the relative significance of every foundation, one must utilize a matrix look over every single conceivable weight to distinguish an arrangement of achievable arrangements. To find optimal solution is a robust approach involves finding the optimality. $X \in S$ be the Pareto optimal solution defines ranking for objective selection of query image from different sources. More precisely, we say that x strictly dominates y if any $f_i(x) \leq f_i(y)$ for all i and $f_j(x) < f_j(y)$ for some i with different objects. A product $x \in S$ is Pareto-optimal if it is not strictly dominated by another product. Pareto optimal feasible alternatives are well known solutions to define scalability for straight formation with different front sides like F_1 and F_2 by eliminating irrelevant images based on manifold ranking procedure to arrange different images. More usually, the ith Pareto front side is determined by:

$$F_i = \text{ParetoOptimalSet} / S \setminus \left(\bigcup_{j=1}^{i-1} F_j \right) \quad (3)$$

If $x \in F_i$, we say that x is the Pareto depth of k, then we say that a Pareto front F_i is deeper than F_j if $i > j$.

Information retrieval using Pareto optimal: We present the Pareto front side means for the multiple-query information recovery issue. Believe that a dataset $X_N = \{X_1, \dots, X_N\}$ of data samples are available in data set.

Given a question q , the potential of recovery is to come back samples that are associated with the question. When several concerns are present, our strategy problems each question independently and then combines their results into one partly requested list of Pareto equivalent recovered items at subsequent Pareto absolute depths. For $T > 1$ which denotes T -records of queries by $\{q_1, q_2, \dots, q_T\}$ and the dissimilarity between q_i and j th in image database X_j , by $d_i(j)$. For efficient consistence, define $d_i \in R_+^N$ as a similarity vector between query q with all the samples in database. For example, given T queries, we define Pareto optimal point is as follows:

$$P_j = [d_1(j), d_2(j), \dots, d_T(j)] \in R_+^T, j \in \{1, 2, \dots, N\} \quad (4)$$

Each Pareto optimal point P_j correspond to the sample X_j from data point in data set χ_N . For convenience with set of all points with Pareto points by p , based on these criteria Pareto point P_i , effectively dominates another point P_j , if $d_l(i) \leq d_l(j)$ for all $l \in \{1, \dots, T\}$ and $d_l(i) < d_l(j)$ for some l . One can without much of a stretch see that if P_i overwhelms P_j , then X_i is nearer to each inquiry than X_j . Accordingly, the framework should return X_i before X_j . The key thought of our approach is to return tests comparing to which Pareto front they lie on, i.e., we restore the focuses from F_1 to start with and after that F_2 , et cetera until the point when an adequate number of pictures have been retrieved. After apply Pareto optimal point in image data sets, for multiple objective image query evaluation, we apply GA procedure to elaborate multi-objective optimization with image retrieval from different image sources.

Multi-objective optimization with (POA and GA): For efficient CBIR from different image sources based on benchmark utilization from above equation can be write as follows:

$$G = \left\{ \begin{array}{l} a \in R^n : l_i \leq a \leq \mu_i \quad \forall i = 1, \dots, m \\ b_i(a) \geq 0, \quad \forall i = 1, \dots, q \end{array} \right\} \quad (5)$$

Where:

- m = The number of objectives
- G = The searchable space process, $a = \{a_1, a_2, a_3, a_4, a_5, \dots, a_n\}$ is the set of m dimensions
- R = The set of real numbers
- R^n = m -dimensional hyper plane space with each label l_i and u_i are lower and upper limits of i^{th} decision variable

Hybrid CBIR technique simultaneously advances the all capacity and creates Pareto optimal arrangements. Pareto front is an arrangement of Pareto ideal (non-commanded) arrangements, being

viewed as ideal if no goal can be enhanced without giving up a less than one other target.

Hybrid CBIR for feature extraction: This study discuss about different feature extractions in multiple query objective optimization using our proposed approach. For efficient CBIR, we use four types of features like contourlet transform, color, texture and shape based on their properties required in image retrieval.

Finding multi-objective solution with GA procedure

Fitness function: The traditional way to deal with tackle a multi-target advancement issue is to appoint a weight W_i to each standardized objective work Z_i , so, the issue is changed over to a solitary target issue with a scalar target work as follows: $G(S)$. Therefore, different arrangements can be at the same time looked in a solitary run. What's more, weight vectors can be acclimated to advance assorted variety of the population.

Crossover and mutation: Among the diverse sorts of hybrids, we utilize the two point hybrid at a hybrid rate of C_R . In two point hybrid, two focuses, $C1$ and $C2$ are chosen in the parent chromosomes. At that point, $N_p/2$ children chromosomes are acquired by exchanging the genes in the middle of the two focuses $C1$ and $C2$ between the two parent chromosomes. Based on above consideration cross over points $C1$ and $C2$ calculated as follows:

$$c_1 = \frac{|D_k^{(i)}|}{3} \quad (6)$$

$$c_2 = c_1 + \frac{|D_k^{(i)}|}{2} \quad (7)$$

From that point forward, the Pareto arrangement of the child chromosomes are put away independently and the new lists comparing to $A_k^{(i)}$ are put away as $A_{nwk}^{(i)}$. In this manner change is performed on the chromosomes acquired after hybrid utilizing above Eq. 6 and 7. The transformation procedure replaces N_M number of genes from each chromosome with the new genes. The N_M quantities of quality are only genes that have greatest estimation of Pareto arrangement (as decided from the Eq. 6 and 7. The supplanted genes are the haphazardly created genes with no reiteration inside the chromosome. At that point, chromosomes which are chosen for hybrid operation and the chromosomes which are acquired from the change are joined, thus the populace pool is topped off with the N_p chromosomes. At that point, the procedure is rehashed iteratively until the point that it achieves a greatest cycle of X_{max} . After gene extraction performed in

cross over and mutation to satisfy maximum Pareto solution in chromosomes with best fitness value, best chromosomes are the gene properties of best fitness value. Our proposed multi-objective calculation procedure as shown in Algorithm 1.

Algorithm 1; Procedure of the proposed approach:

Input: Image Database
 Output: Effective Multi objective Image Retrieval
 Parameters
 DB = Database
 N×A = N- row size and A - column size
 F = fitness
 DM = distance matrix
 π = Instance,
 α = Size of population
 β = Rate of elitism
 γ = Rate of mutation
 δ = Number of iterations

Population size: 100 individuals (10×10)
 Step 1: For each image database DB, I = {i₁, i₂, i₃, ..., i_n} with M*N matrix to store images
 Step 2: Apply Pareto optimal method to generate α feasible solutions randomly for each For I = 1 to δ do and then select best optimal solutions
 Step 3: To check feature extraction for image retrieval in multi objective optimization apply GA. GA consists
 Cross-Over: For each for J = 1 to N do select random two solutions X_a and X_b to parent elements X_c; save these solutions for further use in calculation of fitness value
 Mutation: For each for J = 1 to N do select solution from stored image source X_c, then mutate each inner bit formation of X_c and then generate new solution X_c store with feasible solutions
 Step 4: Update each solution as per image retrieval, end each for loop for better solution
 Step 5: Return Better optimal query image results

After evaluating optimal solution of the query image, the precision, recall and F-measure values are generated for performance evaluation:

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of images retrieved}} \quad (8)$$

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of relevant images in database}} \quad (9)$$

$$\text{F score} = 2 \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

These measures are the important factors to guage the performance of the methods. Perfection is the portion of the appropriate pictures which has been recovered, assessments the completeness of the criteria. Remember is the portion of the appropriate pictures which has been recovered, assessments the precision of the algorithm.

RESULTS AND DISCUSSION

Implementation: Now we present experimental study to compare our proposed hybrid CBIR approach with several existing CBIR methods to multiple query image retrieval. Since, our proposed approach was created for the situation where each inquiry relates to an alternate semantic, we utilize multi-mark datasets in our tests. By multi-name, we imply that many pictures in the dataset have a place with more than one class. This enables us to gauge precise our calculation’s capacity to discover pictures that are similar to all queries. Our proposed approach implemented in working platform of MATLAB, and our approach accumulated with different type image datasets. We use a unity evaluate for unique processing the level of unity to the Pareto maximum and evaluate the remedy with other methods. Since, the test problem is regarded as the recovered, assessments the precision of images. The recommended technique is examined on the five different purpose picture data source of 500 varying pictures includes 5 groups as some different types of seeing stars and blossoms with 200 pictures for each dataset. The precision, recall, efficiency are measured for the example question pictures for each classification using above equations (Fig. 3-6).

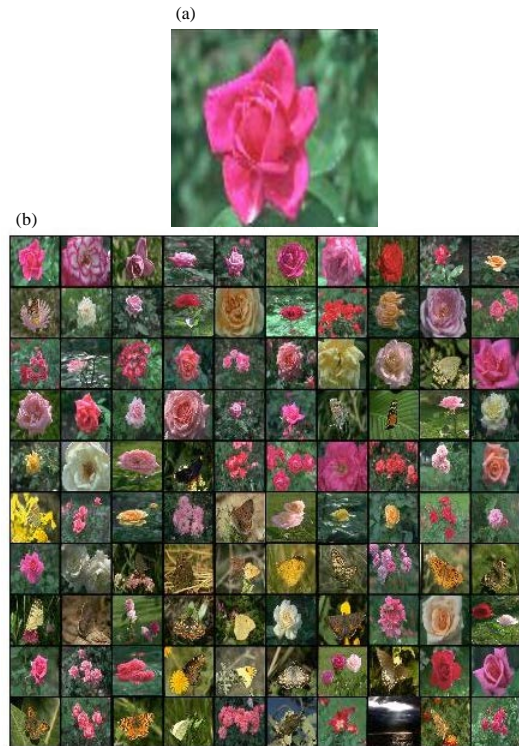


Fig. 3: Query image with resultant images based on color feature

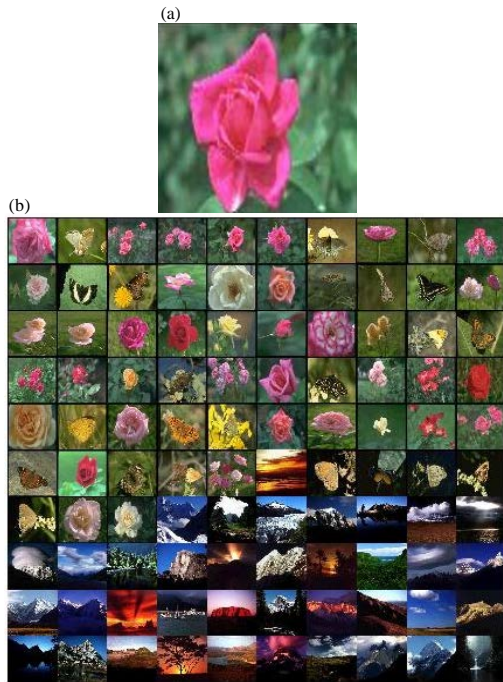


Fig. 4: Query image retrieval with respect to shape feature

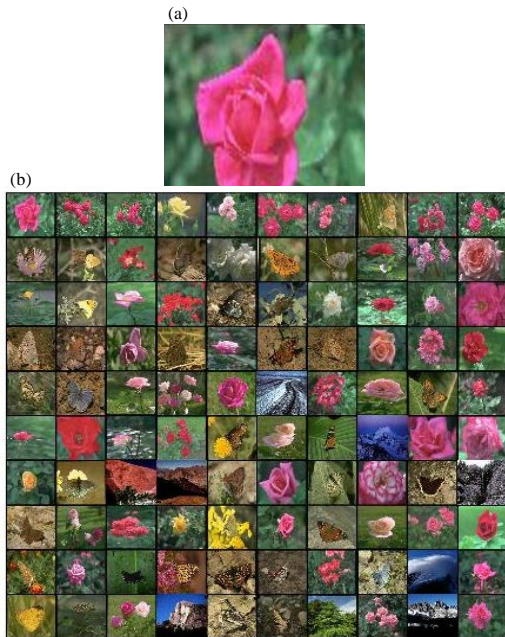


Fig. 5: Query image retrieval with respect to texture feature

Performance results: Performance results are calculated based on different features with query image checking in image databases. Testing results with different features as follows:

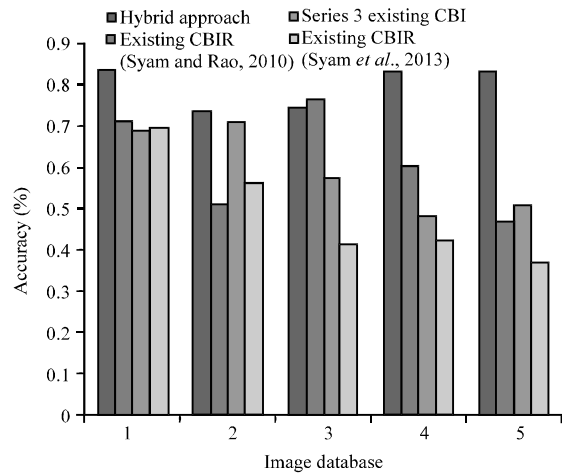


Fig. 6: Accuracy values with respect to different image databases

Table 1: Accuracy values comparison formation with different datasets

Image class	Accuracy			
	Hybrid approach	Existing previous CBIR (Syam and Rao, 2010)	Existing previous CBIR (Syam and Srinivasa Rao, 2012)	Existing previous CBIR image (Syam et al., 2013)
1	0.840	0.712	0.689	0.696
2	0.736	0.510	0.708	0.562
3	0.746	0.764	0.576	0.415
4	0.832	0.604	0.484	0.423
5	0.832	0.470	0.508	0.371

Table 2: Precision presentation values for different image databases.

Image achieves	Precision			
	Hybrid approach	Existing system (Syam and Rao, 2013)	Existing system (Syam and Srinivasa Rao, 2012)	Existing system (Syam et al., 2013)
Class 1	0.560	0.342	0.4245	0.352
Class 2	0.420	0.301	0.4150	0.401
Class 3	0.472	0.399	0.2800	0.338
Class 4	0.445	0.444	0.3920	0.345
Class 5	0.472	0.333	0.3350	0.345

Result related to different features

Results related to different measures: Table 1 shows our proposed accuracy values with respect to existing techniques with different value formats as follows:

Figure 6 shows the accuracy presentation of proposed approach with different image databases. Table 2 follows precision representative values with different images in optimization of image query retrieval.

Figure 7 shows precision results of proposed approach with different representation values. Table 3 shows recall effective image retrieval formation with relevant and irrelevant images from different image datasets.

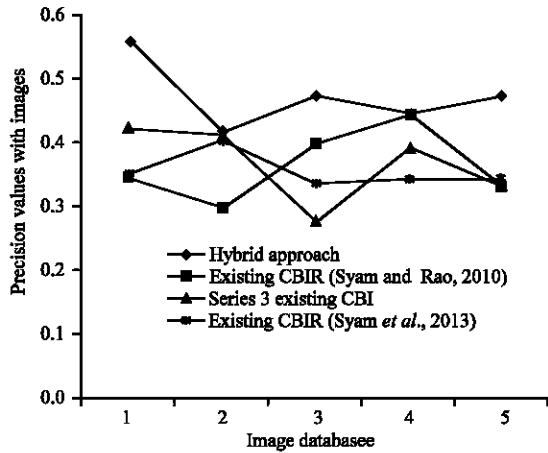


Fig. 7: Precision values with comparison of different image datasets.

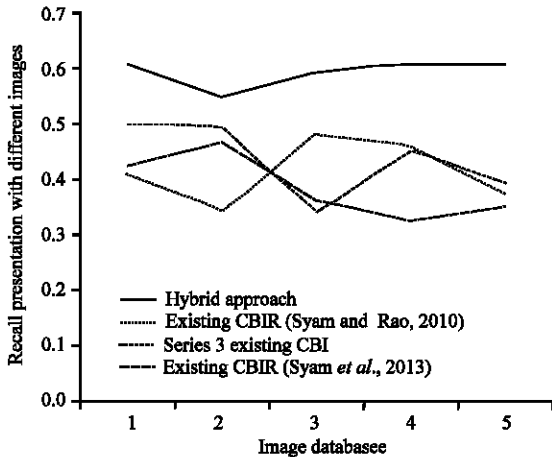


Fig. 8: Image retrieval results of proposed approach (Recall)

Table 3: Recall values with different image databases

Recall				
Databases image archive	Hybrid approach	Existing previous CBIR (Syam and Rao, 2010)	Existing previous CBIR (Srinivasa Rao, 2012)	Existing previous CBIR image (Syam et al., 2013)
1	0.610	0.412	0.501	0.427
2	0.552	0.346	0.496	0.468
3	0.598	0.486	0.346	0.367
4	0.610	0.462	0.454	0.329
5	0.610	0.376	0.396	0.356

Figure 8 shows effective image retrieval formations with different query images from different image datasets. In the all five different dataset recovery were done by querying five different pictures and for every dataset, we get perfection, remember and F-measure principles is

identified. The perfection and remember value evaluation between the suggested comprehensive methods functions removal centered on multi-objective marketing techniques demonstrates the efficiency of the suggested system is best over that of the current systems.

CONCLUSION

We have provided novel criteria for content-based query picture recovery from multiple image archives where the concerns all correspond to different picture semantics and the objective is to find images relevant to all concerns. This criterion can retrieve samples which are not quickly recovered by other multiple-query retrieval methods and any straight line scalarization technique. We have provided theoretical outcomes on asymptotic non-convexity of Pareto methodologies that show that the Pareto strategy is better than using straight line mixtures of position outcomes. Experimental studies on real-world datasets demonstrate the benefits of the suggested Pareto front side method.

REFERENCES

Arandjelovic, R. and A. Zisserman, 2012. Multiple queries for large scale specific object retrieval. Proceedings of the International Conference on British Machine Vision (BMVC'12), September 3-7, 2012, British Machine Vision Association (BMVA), Malvern, England, UK., pp: 1-11.

Binitha, S. and S.S. Sathya, 2012. A survey of bio inspired optimization algorithms. Int. J. Soft Comput. Eng., 2: 137-151.

Binoonal, N.T. and M. Padmaa, 2016. Multiple query image retrieval by topology preserving hashing with EM ranking. Intl. J. Adv. Res. Basic Eng. Sci. Technol., 2: 32-47.

Boonrong, P. and B. Kaewkamnerdpong, 2011. Canonical PSO based nanorobot control for blood vessel repair. World Acad. Sci. Eng. Technol., 58: 511-516.

Calder, J., S. Eshedoglu and A.O. Hero, 2014. A continuum limit for non-dominated sorting. Proceedings of the International Workshop on Information Theory and Applications (ITA'14), February 9-14, 2014, IEEE, San Diego, California, USA., ISBN:978-1-4799-3589-5, pp: 1-7.

Calder, J., S. Eshedoglu and A.O. Hero, 2015. A PDE-based approach to nondominated sorting. SIAM. J. Numer. Anal., 53: 82-104.

Christoudias, C., R. Urtasun and T. Darrell, 2012. Multi-view learning in the presence of view disagreement. Master Thesis, Cornell University, Ithaca, New York, USA.

- Gonen, M. and E. Alpaydin, 2011. Multiple kernel learning algorithms. *J. Mach. Learn. Res.*, 12: 2211-2268.
- Hashni, T. and T. Amudha, 2012. Relative study of CGS with ACO and BCO swarm intelligence techniques. *Int. Comput. Technol. Appl.*, 3: 1775-1781.
- Hsiao, K.J., J. Calder and A.O. Hero, 2015. Pareto-depth for multiple-query image retrieval. *IEEE. Trans. Image Process.*, 24: 583-594.
- Mahant, M., B. Choudhary, A. Kesharwani and K. Rathore, 2012. A profound survey on swarm intelligence. *Intl. J. Adv. Comput. Res.*, 2: 31-36.
- Martens, D., B. Baesens and T. Fawcett, 2011. Editorial Survey: Swarm intelligence for data mining. *Mach. Learn.*, 82: 1-42.
- Padmaa, M. and T.N. Binoona, 2016. Pareto depth with Em-ranking for multiple-query image retrieval. *Intl. J. Adv. Res. Trends Eng. Technol.*, 20: 51-55.
- Quelleg, G., M. Lamard, G. Cazuguel, B. Cochener and C. Roux, 2010. Wavelet optimization for content-based image retrieval in medical databases. *Med. Image Anal.*, 14: 227-241.
- Ramamurthy, B. and K.R. Chandran, 2011. CBMIR: Shape-based image retrieval using canny edge detection and K-means clustering algorithms for medical images. *Int. J. Eng. Sci. Technol.*, 3: 1870-1877.
- Rao, C., S.S. Kumar and B.C. Mohan, 2010. Content based image retrieval using exact legendre moments and support vector machine. *Intl. J. Multimedia Appl.*, 2: 69-79.
- Rashedi, E., H. Nezamabadi-pour and S. Saryazdi, 2013. A simultaneous feature adaptation and feature selection method for content-based image retrieval systems. *Knowl. Based Syst.*, 39: 85-94.
- Ruiz-Vanoye, J.A., O. Diaz-Parra, F. Cocon, A. Soto and D.L.M. Angeles *et al.*, 2012. Meta-heuristics algorithms based on the grouping of animals by social behavior for the traveling salesman problem. *Intl. J. Comb. Optim. Prob. Inf.*, 3: 104-123.
- Shivakumar, B.L. and T. Amudha, 2012. A novel nature-inspired algorithm to solve complex generalized assignment problems. *Intl. J. Res. Innovation Comput. Eng.*, 2: 280-284.
- Singla, A. and M. Garg, 2014. CBIR approach based on combined HSV auto correlogram color moments and Gabor wavelet. *Intl. J. Eng. Comput. Sci.*, 3: 9007-9012.
- Subotic, M., 2011. Artificial bee colony algorithm with multiple onlookers for constrained optimization problems. *Proceeding of the European Computing Conference, April 28-30, 2011, Paris, France*, pp: 251-256.
- Sudhir, R. and L.D.S.S. Baboo, 2011. An efficient CBIR technique with YUV color space and texture features. *Comput. Eng. Intell. Syst.*, 2: 78-84.
- Syam, B. and Y. Srinivasa Rao, 2012. An effective similarity measure via genetic algorithm for content-based image retrieval with extensive features. *Intl. J. Signal Imaging Syst. Eng.*, 5: 18-28.
- Syam, B. and Y.S. Rao, 2010. Integrating contourlet features with texture, color and spatial features for effective image retrieval. *Proceedings of the 2nd IEEE International Conference on Information Management and Engineering (ICIME'10), April 16-18, 2010, IEEE, Chengdu, China, ISBN:978-1-4244-5263-7*, pp: 289-293.
- Syam, B., J.S.R. Victor and Y.S. Rao, 2013. Efficient similarity measure via genetic algorithm for content based medical image retrieval with extensive features. *Proceedings of the 2013 International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s'13), March 22-23, 2013, IEEE, Kottayam, India, ISBN:978-1-4673-5089-1*, pp: 704-711.
- Velmurugan, K. and B.S. Baboo, 2011. Image retrieval using harris corners and histogram of oriented gradients. *Int. J. Comput. Applic.*, 24: 6-10.
- Zhong, Z., J. Zhu and S.C. Hoi, 2015. Fast object retrieval using direct spatial matching. *IEEE. Trans. Multimedia*, 17: 1391-1397.

Hybrid Content Based Image Retrieval System Using Exhaustive Feature Set Processing by Multi Objective Optimization

¹Baddeti Syam and ²Y. Srinivasa Rao

¹Mandava Institute of Engineering and Technology, Department of ECE,
521275 Jaggayyapet, Andhra Pradesh, India

²Department of Instrument Technology, AU College of Engineering,
Andhra University, Visakhapatnam, Andhra Pradesh, India

Abstract: In real world digital image processing, Content Based Image Retrieval (CBIR) is an emerging concept, it is an image retrieval framework based on some visual semantic features like color, texture and shape to search relevant images from multiple image sources. CBIR consist single query or multiple queries to semantic data or same object for multiple class labels with reference to multi query images. In retrieval of query image comparison from multiple image sources may cause optimization problem in image retrieval because of ambiguity in image search. To solve optimization problem for efficient query image retrieval from multiple image archives, in this study, we propose and develop hybrid content based image retrieval system. This system consist pareto optimized method (for removing non-dominated features) and Genetic Algorithm (GA) for effective content based image retrieval. Integration of contourlet features with color, texture and shape is considered to build exhaustive feature set for all images. In our approach Pareto optimal solution consist multi-objective selection functions to process image retrieval from multiple image sources for individual feature selection procedure with query image evaluation for image retrieval. This approach gives better performance than traditional approaches in query image retrieval from multiple image archives. Our experimental results show effectiveness of the proposed method in retrieving all images from image sources with improvement sorting the individual images in searching Pareto solutions in multi-objective optimization problem. Furthermore, this approach gives better accuracy with respect to precision, recall and time efficiency in image retrieval from multiple image sources based on extensive visual features.

Key words: Content based image retrieval, Genetic algorithm, Pareto solution, multi-objective selection functions, squared euclidean distance and visual features, better, image

INTRODUCTION

In earliest days, CBIR has been aggressive concept in machine learning and information or image retrieval applications. Several image retrieval approaches for multi query image selection have been introduced in previous literature. In most of the systems, image picture is correspond to same semantic image with same concept but the background is multiple for multiple sources but continuously have multiple image background with multiple objects in image with alternative angles between same class with multiple objects. The thought is that by using numerous questions of a similar protest, the execution of single-inquiry recovery can be made strides. Various query image retrieval techniques are sorted based on numerous single query processes (Hsiao *et al.*, 2015). A considerable lot of the strategies for single-semantic-various inquiry recovery include joining

the low-level elements from the query pictures to create a solitary arrived at the midpoint of inquiry. In this study, we consider the all the more difficult issue of discovering pictures that are applicable to various inquiries that speak to various picture semantics (Ramamurthy and Chandran, 2011). For this situation, the objective is to discover pictures containing applicable components from each furthermore, every query. Since, the inquiries relate to various semantics, alluring pictures will contain highlights from a few particular pictures, and won't really be firmly identified with any individual query. This makes the issue on a very basic level not quite the same as single query retrieveal and from single-semantic numerous inquiry recoveries. For this situation, the inquiry pictures won't have comparable low-level components and shaping an arrived at the midpoint of query is not as valuable. Since, pertinent pictures don't really have includes firmly lined up with a specific inquiry, a significant number of the

standard recovery procedures are not valuable in this unique situation. For instance, pack of-words sort approaches which may appear to be normal for this issue, require the objective picture to be firmly identified with a few of the inquiries. Another basic method is to include each inquiry each one in turn and normal the subsequent likenesses. This tends to deliver pictures firmly identified with one of the inquiries, however once in a while identified with at the same time. Numerous other multiple inquiry recovery calculations are planned particularly for the semantic single query problem to identify related images with just a single or a couple of the inquiries.

Various query recovery is identified with the meta search issue in software engineering. In meta search, the issue is to join indexed lists for a similar inquiry over multiple web search tools. This is like the single-semantic-multiple query issue as in each web index is issuing a similar inquiry (or semantic). Consequently, meta search calculations are not reasonable with regards to multiple query recovery with a few particular semantics.

So, in this study, we propose to develop a novel approach, i.e., hybrid content based image retrieval system which combines Pareto Optimization Algorithm (POA) and Genetic algorithm. In this approach, POA retrieve each image and rank them individually with their discriminative to the input query. Some of the techniques like SIFT and HoG has been proposed in different computer vision oriented applications can be utilized to define image dissimilarities. These dissimilarities explore query retrieval in image achieves, we utilize a quick positioning calculation called Manifold Efficient Ranking (MER) to calculate the ranking without consider total presented samples sets. MER can productively find the basic dimension of given image database and find the evaluation time of the conventional ranking system. MER effectively defines single query image retrieval to define various retrieval query issue.

The next stage in our hybrid approach is to utilize the positioning delivered by MER to make Pareto focuses,

which compare to non-similarities between a specimen and each inquiry. Pareto optimal points are called Pareto fronts to define non-dominated events Pareto front mainly focuses on non-dominated focuses and it is called origin in group image databases. The second Pareto front (profundity two) is acquired by expelling the principal Pareto front and finding the non-overwhelmed focuses among the rest of the Pareto front algorithm contains specimens for each user and those samples are exhausted. Pareto front sort the non-dominated class references.

Key distribution of Pareto front for query image from multiple image databases shown in Fig. 1 to recovery query images based on forest and mountain keywords. In that image are listed based on their positioning with front and connection from tail to other present in image retrieval.

However, as seen in Fig. 1, images (images 10, 11, 12) in center of front optimization contain related functions for both queries suitable for multi image retrieval problem. Our proposed approach is an well known approach to provide symmetric calculation with respect to query image retrieval from multiple image class retrieval. We also applicable our approach to automatic image annotation of large databases. The main contribution of our proposed approach is as follows: Propose the hybrid CBIR for query image retrieval from multiple image archives based on visual features of query image. Discuss about Pareto Optimization Method (POM) for image annotation in multiple image retrieval. For effective similarity measure between query image and data base image with extensive feature presentation is explained using genetic algorithm for query image retrieval from multiple image archives.

Literature review: Quick object retrieval using direct spatial matching paper (Syam and Rao, 2010) appears for the substantial scale question recovery Bag of visual-Words (BOW) demonstrates is prevalent however there is a disadvantage it disregards spatial data. Direct

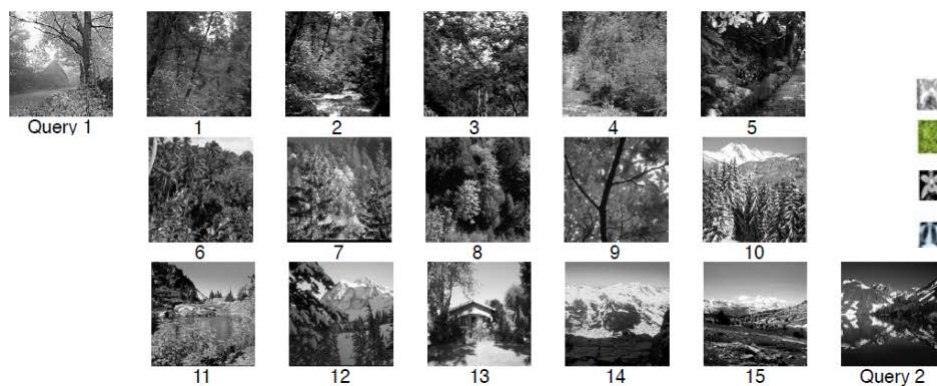


Fig. 1: Images located based on Pareto front method with middle of semantic query images

Spatial Matching (DSM) way to deal with discover the scale variety utilizing district sizes in which all component matches for evaluating geometric transformation. DSM is significantly snappier than RANSAC-based techniques and thorough specification approaches. Alogarithmic term recurrence converse archive recurrence weighting plan is acquainted with increase the execution of the base framework. The disadvantage is a high cost because of high recovery exactness. Non-overwhelmed arranging (Syam and Srinivasa Rao, 2012; Syam *et al.*, 2013; Arandjelovic and Zisserman, 2012; Binitha and Sathya, 2012; Martens *et al.*, 2011) is a critical issue in multi-target advancement. They have discovered that in the extensive specimen estimate constrain, the non-commanded fronts merge most likely to the level arrangement of a capacity that fulfills Hamilton-Jacobi Partial Differential Equations (PDE). To plan for a quick, possibly sub straight, surmised non-ruled arranging calculation PDE is utilized and the aftereffects of applying the calculation to genuine information (Rashedi *et al.*, 2013). They likewise give a quick numerical plan for tackling the PDE and utilize it to build up a quick, possibly sub straight, estimated non-ruled arranging calculation. This calculation is connected to genuine information from an oddity location issue and amazing exactness is accomplished by this calculation while fundamentally decreasing the computational multifaceted nature of non-dominated arranging. For huge information gushing issues, this calculation is utilized which include steady re arranging of substantial datasets upon the landing of new examples. For the straight arrangement of highlights, it is reasonable.

The expansive scale particular question picture recovery frameworks are to discover pictures that contain the inquiry protest in the picture database. For instance media, generation groups are keen on seeking pictures or video film to go with news reports and daily paper articles from the database. Get different pictures of the questioned protest utilizing Google picture seek. Demonstrate that various questions altogether enhance the framework to discover testing events of the questioned protest. The point is to recover all pictures containing a particular protest in a picture informational collection. This is an issue that has seen much improvement and accomplishment in the course of the most recent decade, the beginning stage of the inquiry has been a solitary question picture of the particular protest of intrigue. In this work two changes to the standard approach: to begin with, beginning stage the protest is content, as they are keen on testing informational indexes to discover known questions and second the informational index utilizing numerous picture inquiries and join the outcomes into a solitary positioned list. Tending to this issue has been one of the principle inquire about subjects in particular protest recovery look into with advancements in include

encoding to lighten Vector Quantization (VQ) misfortunes and in growth of the sack of visual word (BOW) portrayal to mitigate locator and descriptor drop out (and in addition, once more, VQ losses). All time the change of printed is impractical for every one of the pictures. In this way, tree-based lists are not attractive for high-dimensional pursuit issues. Another sort of ANN look calculation depends on the vector quantization for example, k-implies LSH (Martens *et al.*, 2011) and product quantization (AB) (Rashedi *et al.*, 2013). In AB, by isolating every datum into a few sub spaces and communicating information as far as repeating parts, the illustrative limit of AB, becomes forceful in the quantity of sub spaces. In these strategies, every datum point is spoken to by an element vector and the inquiry procedure is performed. Subsequently, the hunting procedure is tedious. As of late, the hashing based technique has been generally utilized for closeness seek and comparative applications as it permits the steady time look. A considerable measure of hashing techniques has been proposed these can be isolated into two fundamental classifications: information free strategies (Syam and Rao, 2010; Quellec *et al.*, 2010) and information subordinate strategies. Hypothetically, it is ensured that the first separations are asymptotically spared in the Hamming space inside wrinkling code length; thus, LSH related strategies more often than not require long codes to accomplish great exactness. In any case, long codes result in the impact likelihood of comparable focuses mapped to close codes diminishes as the code length increments. Subsequently, the LSH-related techniques, for the most part, build multi-tables to guarantee a sensible likelihood that an inquiry will crash into its close neighbors in no less than one table which prompts long question time and expands the memory occupation. As of late, numerous information subordinate techniques which concentrate on taking in hash capacities from the informational index, have been produced to take in more minimized codes.

The area connections between sub-locales and the relative vicinities between neighbors of every sub region are both basic (Ruiz-Vanoye *et al.*, 2012) for viable measurement decrease. Many hashing strategies have been created for neighborhood-protecting (Syam and Srinivasa Rao, 2012) or separation saving (Mahant *et al.*, 2012). The previous one sign or the area rankings and later to help and a complex with unbending steel pillar (Ruiz-Vanoye *et al.*, 2012) while as a rule, the ideal inserting of a complex needs some adaptability: some sub-areas ought to be privately extended or contracted to install the min to alow-dimensional space where the topology can be very much saved. In this manner, these hashing strategies can't safeguard the informational index topology well.

In complex learning region, the reality of the matter is that a complex can be totally described by giving the

relative or similar vicinities (Ruiz-Vanoye *et al.*, 2012), e.g., a first area is near a moment one yet a long way from a third one. Near in arrangement between separations, as in correspondences or rankings, get the job done to describe a complex for any embeddings. In addition, for some comparability seek issues, the relative rankings of results are more vital than their genuine similitude's to a question. In light of the above presentation, we can infer that more powerful Hamming installing's can be learned by consolidating the informational index topology with the learning procedure. Assist more, if the nearby topology is all around protected in Hamming space, the vagueness caused by positioning with Hamming separation can be better all mitigated.

Syam *et al.* (2013) proposes similarity measure in genetic algorithm for efficient image retrieval with exhaustive feature presentation in image retrieval conditions based on query image processing (Shivakumar and Amudha, 2012; Rao *et al.*, 2010). In this scenario when the query image is taken from image database and then Genetic Algorithm (GA) similarity measure is performed in between query image retrieval and database image retrieval. It is also performed different query image retrieval for medical image.

Basic utilization of Pareto front calculation aggressively (Subotic, 2011; Boonrong and Kaewkamnerdpong, 2011; Sudhir and Baboo, 2011) in the machine learning group. Complex multi-target advancement issues are illuminated by this strategy, where finding the main Pareto front is testing. A few productive calculations have been created for finding the skyline strategy. In our POA calculation, efficient and quick skyline calculations or quick non-commanded arranging can be utilized to discover every Pareto front.

MATERIALS AND METHODS

Image query retrieval from multiple image archives:

For effective image retrieval, images are extracted and pre-processed based on low dimensional feature representation instead of pixel values in directing the image indexing and image retrieval (Rao *et al.*, 2010; Hashni and Amudha, 2012; Padmaa and Binoona, 2016; Binoonal and Padmaa, 2016; Singla and Garg, 2014). Some of the feature extraction and similarity extraction techniques has been proposed on various computer vision and image processing oriented applications to define best query matching with different images. In this scenario ambiguity may appear in image query retrieval from multiple image archives, we present and use manifold ranking procedure, POA and GA for effective CBIR in (Velmurugan and Baboo, 2011; Syam and Srinivasa Rao,

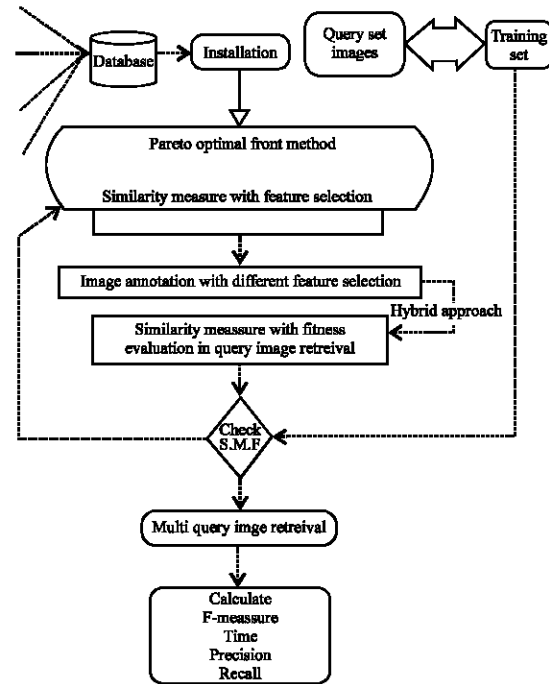


Fig. 2: Proposed implementation for multi query optimized image retrieval with Genetic algorithm

2012). Procedure of the Image query retrieval from multiple image archives using hybrid image retrieval systems may show in Fig. 2 with extensible feature presentation in image retrieval.

As discussed previous paragraph to obtain effective CBIR and verify sample query pair in image retrieval for image positioning in data base based on query image with respect to visual features (Zhong *et al.*, 2015; Calder *et al.*, 2014). Procedure for ranking manifold for image positioning with different frames as follows:

Manifold ranking procedure (MER): Let $\mathcal{N} = \{X_1, X_2, \dots, X_n\} \subset T^m$ to be feasible set of Pareto points and $d: \mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R}$ be a metric on \mathcal{N} such that call as squared euclidean distance define sequential attribute processing with following $y = \{y_1, y_2, \dots, y_n\}$ in which $y_i = 1$, if X_i is the query $y_i = 0$ otherwise. Let $r: \mathcal{X} \rightarrow \mathbb{R}$ is the ranking function to assign with image position score r_i to each point X_i . The entirely allocated a position of 1 and all other examples will be allocated more compact positions with range to the keyword along with image information. To create a graph on X , first sort the for all ranges between different examples in ascending order to link related Graph G is designed to explore image retrieval. Based on this procedure in manifold ranking approach associated with ranking vector r processed for image ranking is as follows:

$$O(r) = \sum_{i,j=1}^n w_{ij} \left| \frac{1}{\sqrt{D_{ij}}} r_i - \frac{1}{\sqrt{D_{ij}}} r_j \right|^2 + \mu \sum_{i=1}^n |r_i - y_i|^2 \quad (1)$$

where, D is the diagonal matrix with $D_{ij} = \sum_{j=1}^n w_{ij}$ and $\mu > 0$ is sequential regularization parameter. The first phrase in the cost function is a level of smoothness phrase that causes close by factors have similar position ratings (Calder *et al.*, 2014; Christoudias *et al.*, 2012). The second phrase is a regularization term which causes the question to have a position near to 1 and all other examples to have positions as near to 0 as possible. Based on this procedure efficient manifold calculation is measured with different formation in final ranking is computed as follows:

$$r^* = (I_n - H^T (HH^T - \frac{1}{\alpha} I_d)^{-1} H) y \quad (2)$$

where, $H = ZD^{-1/2}$ and D is the diagonal matrix, notice construction of the anchor graph structure with matrix inversion to present non-ambiguity based sample image retrieval conditions with perspective measure of fast image retrieval.

Retrieval of query image from multiple classes of archives: Based on previous knowledge regarding relevance of keyword for efficient calculation with selection of starting image from multiple image archives in (Gonen and Alpaydin, 2011; Calder *et al.*, 2014) query image retrieval from multiple archives efficient manifold ranking to each data based image with query image. Main contribution of this study shows Pareto method for standard image retrieval. To return relevant tests to the client, we return tests as indicated by their Pareto front number, i.e., we return focuses on F_1 , to begin with, at that point F_2 , et cetera, until the point that adequately many pictures have been recovered. Inside a similar front, we return focuses in a specific request, e.g., for $T = 2$, from the center first (Calder *et al.*, 2015). With regards to this study, the center piece of each front will contain tests identified with all inquiries. Relevant input plans can likewise be utilized in our calculation to upgrade recovery execution. For efficient CBIR in query image retrieval from multiple archives apply Pareto optimal algorithm and GA in image retrieval after ranking the images with different visual semantic features.

Hybrid approach implementation for multiple query image: In this study, we deal with an application of hybrid approach based on multi objective process. Hybrid consists both Pareto optimization method and Genetic algorithm to obtain multi-objective image retrieval, for different set of datasets, single query measurement is not worked properly to efficient image retrieval. Moreover, a

combination of different actions is very challenging, because of the method for the combination may be not possible and sometimes those actions may be unsuitable. In this study, we propose Pareto optimal method to check optimization which helps to find the related objectives in information retrieval applications. We incorporate optimization approach with in GA algorithm implementation to extract multi query access from different data sets.

Pareto optimal method: Pareto optimality is an effective concept and applied in different domains like, Economic Sciences, Computer Sciences and Social Sciences. Here, we give brief review about Pareto optimality and define Pareto front. In general setting of a discrete multi-target enhancement issue, we have a limited set S of achievable arrangements and T criteria f_1, f_2, \dots, f_T for assessing the plausible arrangements. One conceivable objective is to discover x 2 S limiting all criteria at the same time. In many settings, this is an inconceivable errand. Many ways to deal with the multi-target streamlining issue diminish to joining all T criteria into one. At the point when this is finished with a direct mix, it is typically called linear scalarization. Selection of different weights in the straight mix yield distinctive minimizes. Without earlier learning of the relative significance of every foundation, one must utilize a matrix look over every single conceivable weight to distinguish an arrangement of achievable arrangements. To find optimal solution is a robust approach involves finding the optimality. $X \in S$ be the Pareto optimal solution defines ranking for objective selection of query image from different sources. More precisely, we say that x strictly dominates y if any $f_i(x) \leq f_i(y)$ for all i and $f_j(x) < f_j(y)$ for some i with different objects. A product $x \in S$ is Pareto-optimal if it is not strictly dominated by another product. Pareto optimal feasible alternatives are well known solutions to define scalability for straight formation with different front sides like F_1 and F_2 by eliminating irrelevant images based on manifold ranking procedure to arrange different images. More usually, the ith Pareto front side is determined by:

$$F_i = \text{ParetoOptimalSet} / S \setminus \left(\bigcup_{j=1}^{i-1} F_j \right) \quad (3)$$

If $x \in F_i$, we say that x is the Pareto depth of k, then we say that a Pareto front F_i is deeper than F_j if $i > j$.

Information retrieval using Pareto optimal: We present the Pareto front side means for the multiple-query information recovery issue. Believe that a dataset $X_N = \{X_1, \dots, X_N\}$ of data samples are available in data set.

Given a question q , the potential of recovery is to come back samples that are associated with the question. When several concerns are present, our strategy problems each question independently and then combines their results into one partly requested list of Pareto equivalent recovered items at subsequent Pareto absolute depths. For $T > 1$ which denotes T -records of queries by $\{q_1, q_2, \dots, q_T\}$ and the dissimilarity between q_i and j th in image database X_j , by $d_i(j)$. For efficient consistence, define $d_i \in R_+^N$ as a similarity vector between query q with all the samples in database. For example, given T queries, we define Pareto optimal point is as follows:

$$P_j = [d_1(j), d_2(j), \dots, d_T(j)] \in R_+^T, j \in \{1, 2, \dots, N\} \quad (4)$$

Each Pareto optimal point P_j correspond to the sample X_j from data point in data set χ_N . For convenience with set of all points with Pareto points by p , based on these criteria Pareto point P_i , effectively dominates another point P_j , if $d_l(i) \leq d_l(j)$ for all $l \in \{1, \dots, T\}$ and $d_l(i) < d_l(j)$ for some l . One can without much of a stretch see that if P_i overwhelms P_j , then X_i is nearer to each inquiry than X_j . Accordingly, the framework should return X_i before X_j . The key thought of our approach is to return tests comparing to which Pareto front they lie on, i.e., we restore the focuses from F_1 to start with and after that F_2 , et cetera until the point when an adequate number of pictures have been retrieved. After apply Pareto optimal point in image data sets, for multiple objective image query evaluation, we apply GA procedure to elaborate multi-objective optimization with image retrieval from different image sources.

Multi-objective optimization with (POA and GA): For efficient CBIR from different image sources based on benchmark utilization from above equation can be write as follows:

$$G = \left\{ \begin{array}{l} a \in R^n : l_i \leq a \leq \mu_i \quad \forall i = 1, \dots, m \\ b_i(a) \geq 0, \quad \forall i = 1, \dots, q \end{array} \right\} \quad (5)$$

Where:

- m = The number of objectives
- G = The searchable space process, $a = \{a_1, a_2, a_3, a_4, a_5, \dots, a_n\}$ is the set of m dimensions
- R = The set of real numbers
- R^n = m -dimensional hyper plane space with each label l_i and u_i are lower and upper limits of i^{th} decision variable

Hybrid CBIR technique simultaneously advances the all capacity and creates Pareto optimal arrangements. Pareto front is an arrangement of Pareto ideal (non-commanded) arrangements, being

viewed as ideal if no goal can be enhanced without giving up a less than one other target.

Hybrid CBIR for feature extraction: This study discuss about different feature extractions in multiple query objective optimization using our proposed approach. For efficient CBIR, we use four types of features like contourlet transform, color, texture and shape based on their properties required in image retrieval.

Finding multi-objective solution with GA procedure

Fitness function: The traditional way to deal with tackle a multi-target advancement issue is to appoint a weight W_i to each standardized objective work Z , so, the issue is changed over to a solitary target issue with a scalar target work as follows: $G(S)$. Therefore, different arrangements can be at the same time looked in a solitary run. What's more, weight vectors can be acclimated to advance assorted variety of the population.

Crossover and mutation: Among the diverse sorts of hybrids, we utilize the two point hybrid at a hybrid rate of C_R . In two point hybrid, two focuses, $C1$ and $C2$ are chosen in the parent chromosomes. At that point, $N_p/2$ children chromosomes are acquired by exchanging the genes in the middle of the two focuses $C1$ and $C2$ between the two parent chromosomes. Based on above consideration cross over points $C1$ and $C2$ calculated as follows:

$$c_1 = \frac{|D_k^{(i)}|}{3} \quad (6)$$

$$c_2 = c_1 + \frac{|D_k^{(i)}|}{2} \quad (7)$$

From that point forward, the Pareto arrangement of the child chromosomes are put away independently and the new lists comparing to $A_k^{(i)}$ are put away as $A_{nwk}^{(i)}$. In this manner change is performed on the chromosomes acquired after hybrid utilizing above Eq. 6 and 7. The transformation procedure replaces N_M number of genes from each chromosome with the new genes. The N_M quantities of quality are only genes that have greatest estimation of Pareto arrangement (as decided from the Eq. 6 and 7. The supplanted genes are the haphazardly created genes with no reiteration inside the chromosome. At that point, chromosomes which are chosen for hybrid operation and the chromosomes which are acquired from the change are joined, thus the populace pool is topped off with the N_p chromosomes. At that point, the procedure is rehashed iteratively until the point that it achieves a greatest cycle of X_{max} . After gene extraction performed in

cross over and mutation to satisfy maximum Pareto solution in chromosomes with best fitness value, best chromosomes are the gene properties of best fitness value. Our proposed multi-objective calculation procedure as shown in Algorithm 1.

Algorithm 1; Procedure of the proposed approach:

Input: Image Database
 Output: Effective Multi objective Image Retrieval
 Parameters
 DB = Database
 N×A = N- row size and A - column size
 F = fitness
 DM = distance matrix
 π = Instance,
 α = Size of population
 β = Rate of elitism
 γ = Rate of mutation
 δ = Number of iterations

Population size: 100 individuals (10×10)
 Step 1: For each image database DB, $I = \{i_1, i_2, i_3, \dots, i_n\}$ with M*N matrix to store images
 Step 2: Apply Pareto optimal method to generate α feasible solutions randomly for each For I = 1 to δ do and then select best optimal solutions
 Step 3: To check feature extraction for image retrieval in multi objective optimization apply GA. GA consists
 Cross-Over: For each for J = 1 to N do select random two solutions Xa and Xb to parent elements Xc; save these solutions for further use in calculation of fitness value
 Mutation: For each for J = 1 to N do select solution from stored image source Xc, then mutate each inner bit formation of Xc and then generate new solution Xc' store with feasible solutions
 Step 4: Update each solution as per image retrieval, end each for loop for better solution
 Step 5: Return Better optimal query image results

After evaluating optimal solution of the query image, the precision, recall and F-measure values are generated for performance evaluation:

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of images retrieved}} \quad (8)$$

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of relevant images in database}} \quad (9)$$

$$\text{F score} = 2 \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

These measures are the important factors to guage the performance of the methods. Perfection is the portion of the appropriate pictures which has been recovered, assessments the completeness of the criteria. Remember is the portion of the appropriate pictures which has been recovered, assessments the precision of the algorithm.

RESULTS AND DISCUSSION

Implementation: Now we present experimental study to compare our proposed hybrid CBIR approach with several existing CBIR methods to multiple query image retrieval. Since, our proposed approach was created for the situation where each inquiry relates to an alternate semantic, we utilize multi-mark datasets in our tests. By multi-name, we imply that many pictures in the dataset have a place with more than one class. This enables us to gauge precise our calculation’s capacity to discover pictures that are similar to all queries. Our proposed approach implemented in working platform of MATLAB, and our approach accumulated with different type image datasets. We use a unity evaluate for unique processing the level of unity to the Pareto maximum and evaluate the remedy with other methods. Since, the test problem is regarded as the recovered, assessments the precision of images. The recommended technique is examined on the five different purpose picture data source of 500 varying pictures includes 5 groups as some different types of seeing stars and blossoms with 200 pictures for each dataset. The precision, recall, efficiency are measured for the example question pictures for each classification using above equations (Fig. 3-6).

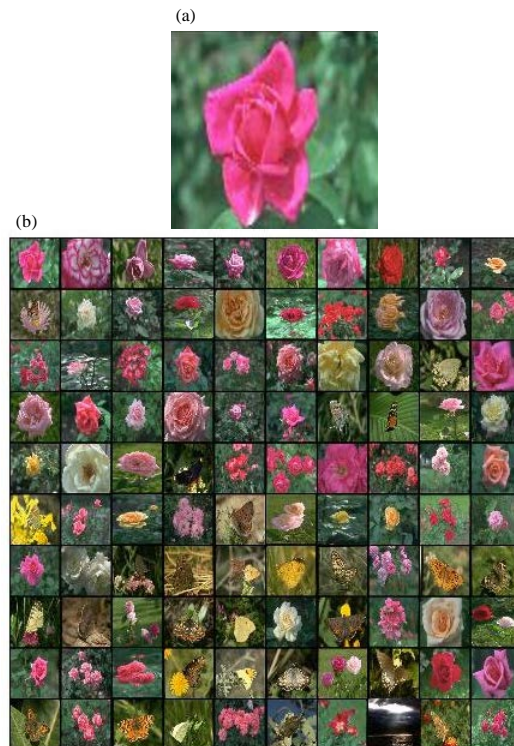


Fig. 3: Query image with resultant images based on color feature

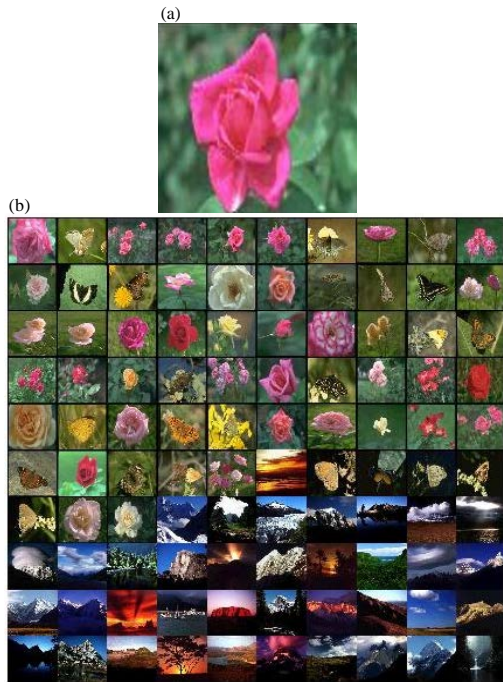


Fig. 4: Query image retrieval with respect to shape feature

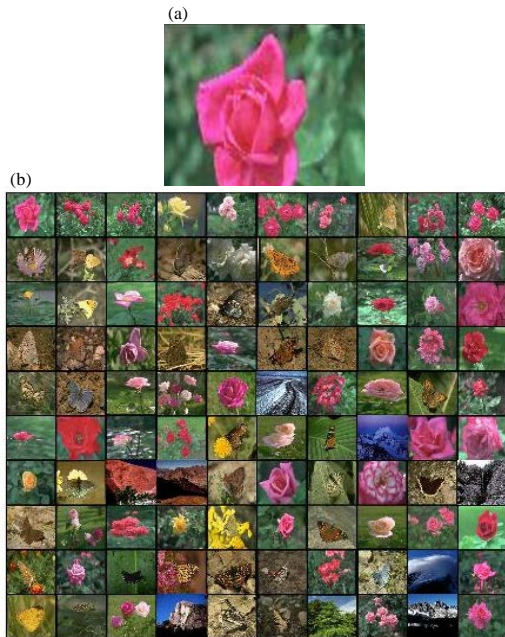


Fig. 5: Query image retrieval with respect to texture feature

Performance results: Performance results are calculated based on different features with query image checking in image databases. Testing results with different features as follows:

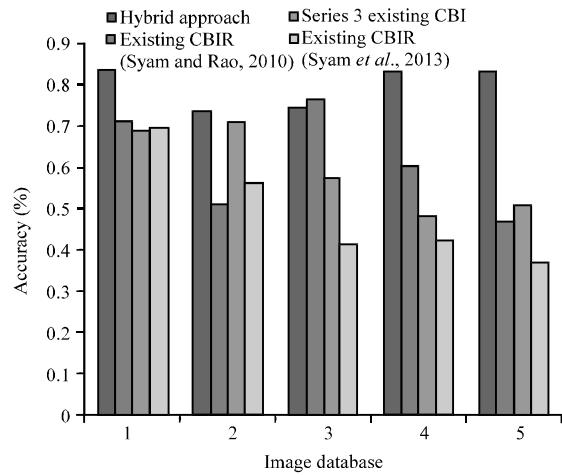


Fig. 6: Accuracy values with respect to different image databases

Table 1: Accuracy values comparison formation with different datasets

Image class	Accuracy			
	Hybrid approach	Existing previous CBIR (Syam and Rao, 2010)	Existing previous CBIR (Syam and Srinivasa Rao, 2012)	Existing previous CBIR image (Syam et al., 2013)
1	0.840	0.712	0.689	0.696
2	0.736	0.510	0.708	0.562
3	0.746	0.764	0.576	0.415
4	0.832	0.604	0.484	0.423
5	0.832	0.470	0.508	0.371

Table 2: Precision presentation values for different image databases.

Image achieves	Precision			
	Hybrid approach	Existing system (Syam and Rao, 2013)	Existing system (Syam and Srinivasa Rao, 2012)	Existing system (Syam et al., 2013)
Class 1	0.560	0.342	0.4245	0.352
Class 2	0.420	0.301	0.4150	0.401
Class 3	0.472	0.399	0.2800	0.338
Class 4	0.445	0.444	0.3920	0.345
Class 5	0.472	0.333	0.3350	0.345

Result related to different features

Results related to different measures: Table 1 shows our proposed accuracy values with respect to existing techniques with different value formats as follows:

Figure 6 shows the accuracy presentation of proposed approach with different image databases. Table 2 follows precision representative values with different images in optimization of image query retrieval.

Figure 7 shows precision results of proposed approach with different representation values. Table 3 shows recall effective image retrieval formation with relevant and irrelevant images from different image datasets.

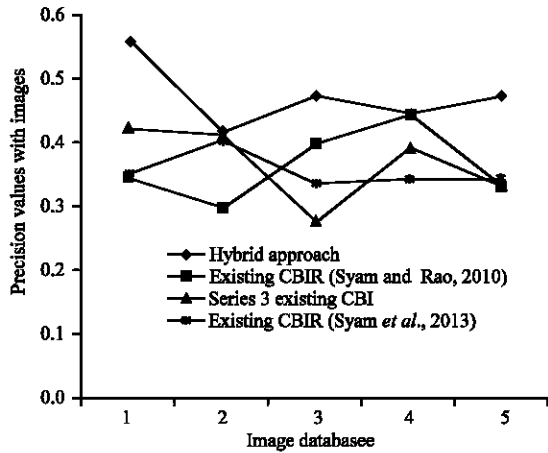


Fig. 7: Precision values with comparison of different image datasets.

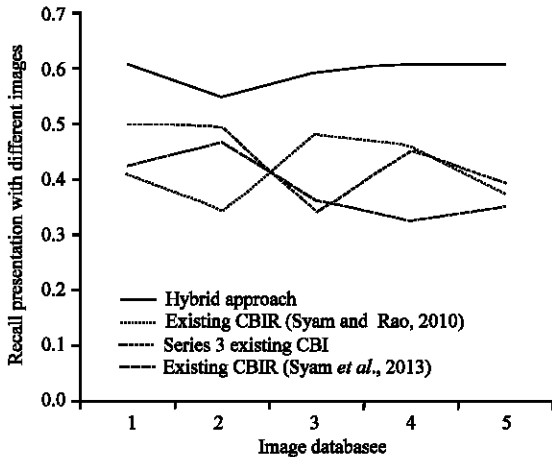


Fig. 8: Image retrieval results of proposed approach (Recall)

Table 3: Recall values with different image databases

Recall				
Databases image archive	Hybrid approach	Existing previous CBIR (Syam and Rao, 2010)	Existing previous CBIR (Srinivasa Rao, 2012)	Existing previous CBIR image (Syam et al., 2013)
1	0.610	0.412	0.501	0.427
2	0.552	0.346	0.496	0.468
3	0.598	0.486	0.346	0.367
4	0.610	0.462	0.454	0.329
5	0.610	0.376	0.396	0.356

Figure 8 shows effective image retrieval formations with different query images from different image datasets. In the all five different dataset recovery were done by querying five different pictures and for every dataset, we get perfection, remember and F-measure principles is

identified. The perfection and remember value evaluation between the suggested comprehensive methods functions removal centered on multi-objective marketing techniques demonstrates the efficiency of the suggested system is best over that of the current systems.

CONCLUSION

We have provided novel criteria for content-based query picture recovery from multiple image archives where the concerns all correspond to different picture semantics and the objective is to find images relevant to all concerns. This criterion can retrieve samples which are not quickly recovered by other multiple-query retrieval methods and any straight line scalarization technique. We have provided theoretical outcomes on asymptotic non-convexity of Pareto methodologies that show that the Pareto strategy is better than using straight line mixtures of position outcomes. Experimental studies on real-world datasets demonstrate the benefits of the suggested Pareto front side method.

REFERENCES

Arandjelovic, R. and A. Zisserman, 2012. Multiple queries for large scale specific object retrieval. Proceedings of the International Conference on British Machine Vision (BMVC'12), September 3-7, 2012, British Machine Vision Association (BMVA), Malvern, England, UK., pp: 1-11.

Binitha, S. and S.S. Sathya, 2012. A survey of bio inspired optimization algorithms. Int. J. Soft Comput. Eng., 2: 137-151.

Binoonal, N.T. and M. Padmaa, 2016. Multiple query image retrieval by topology preserving hashing with EM ranking. Intl. J. Adv. Res. Basic Eng. Sci. Technol., 2: 32-47.

Boonrong, P. and B. Kaewkamnerdpong, 2011. Canonical PSO based nanorobot control for blood vessel repair. World Acad. Sci. Eng. Technol., 58: 511-516.

Calder, J., S. Eshedoglu and A.O. Hero, 2014. A continuum limit for non-dominated sorting. Proceedings of the International Workshop on Information Theory and Applications (ITA'14), February 9-14, 2014, IEEE, San Diego, California, USA., ISBN:978-1-4799-3589-5, pp: 1-7.

Calder, J., S. Eshedoglu and A.O. Hero, 2015. A PDE-based approach to nondominated sorting. SIAM. J. Numer. Anal., 53: 82-104.

Christoudias, C., R. Urtasun and T. Darrell, 2012. Multi-view learning in the presence of view disagreement. Master Thesis, Cornell University, Ithaca, New York, USA.

- Gonen, M. and E. Alpaydin, 2011. Multiple kernel learning algorithms. *J. Mach. Learn. Res.*, 12: 2211-2268.
- Hashni, T. and T. Amudha, 2012. Relative study of CGS with ACO and BCO swarm intelligence techniques. *Int. Comput. Technol. Appl.*, 3: 1775-1781.
- Hsiao, K.J., J. Calder and A.O. Hero, 2015. Pareto-depth for multiple-query image retrieval. *IEEE. Trans. Image Process.*, 24: 583-594.
- Mahant, M., B. Choudhary, A. Kesharwani and K. Rathore, 2012. A profound survey on swarm intelligence. *Intl. J. Adv. Comput. Res.*, 2: 31-36.
- Martens, D., B. Baesens and T. Fawcett, 2011. Editorial Survey: Swarm intelligence for data mining. *Mach. Learn.*, 82: 1-42.
- Padmaa, M. and T.N. Binoona, 2016. Pareto depth with Em-ranking for multiple-query image retrieval. *Intl. J. Adv. Res. Trends Eng. Technol.*, 20: 51-55.
- Quellec, G., M. Lamard, G. Cazuguel, B. Cochener and C. Roux, 2010. Wavelet optimization for content-based image retrieval in medical databases. *Med. Image Anal.*, 14: 227-241.
- Ramamurthy, B. and K.R. Chandran, 2011. CBMIR: Shape-based image retrieval using canny edge detection and K-means clustering algorithms for medical images. *Int. J. Eng. Sci. Technol.*, 3: 1870-1877.
- Rao, C., S.S. Kumar and B.C. Mohan, 2010. Content based image retrieval using exact legendre moments and support vector machine. *Intl. J. Multimedia Appl.*, 2: 69-79.
- Rashedi, E., H. Nezamabadi-pour and S. Saryazdi, 2013. A simultaneous feature adaptation and feature selection method for content-based image retrieval systems. *Knowl. Based Syst.*, 39: 85-94.
- Ruiz-Vanoye, J.A., O. Diaz-Parra, F. Cocon, A. Soto and D.L.M. Angeles *et al.*, 2012. Meta-heuristics algorithms based on the grouping of animals by social behavior for the traveling salesman problem. *Intl. J. Comb. Optim. Prob. Inf.*, 3: 104-123.
- Shivakumar, B.L. and T. Amudha, 2012. A novel nature-inspired algorithm to solve complex generalized assignment problems. *Intl. J. Res. Innovation Comput. Eng.*, 2: 280-284.
- Singla, A. and M. Garg, 2014. CBIR approach based on combined HSV auto correlogram color moments and Gabor wavelet. *Intl. J. Eng. Comput. Sci.*, 3: 9007-9012.
- Subotic, M., 2011. Artificial bee colony algorithm with multiple onlookers for constrained optimization problems. *Proceeding of the European Computing Conference, April 28-30, 2011, Paris, France*, pp: 251-256.
- Sudhir, R. and L.D.S.S. Baboo, 2011. An efficient CBIR technique with YUV color space and texture features. *Comput. Eng. Intell. Syst.*, 2: 78-84.
- Syam, B. and Y. Srinivasa Rao, 2012. An effective similarity measure via genetic algorithm for content-based image retrieval with extensive features. *Intl. J. Signal Imaging Syst. Eng.*, 5: 18-28.
- Syam, B. and Y.S. Rao, 2010. Integrating contourlet features with texture, color and spatial features for effective image retrieval. *Proceedings of the 2nd IEEE International Conference on Information Management and Engineering (ICIME'10), April 16-18, 2010, IEEE, Chengdu, China, ISBN:978-1-4244-5263-7*, pp: 289-293.
- Syam, B., J.S.R. Victor and Y.S. Rao, 2013. Efficient similarity measure via genetic algorithm for content based medical image retrieval with extensive features. *Proceedings of the 2013 International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s'13), March 22-23, 2013, IEEE, Kottayam, India, ISBN:978-1-4673-5089-1*, pp: 704-711.
- Velmurugan, K. and B.S. Baboo, 2011. Image retrieval using harris corners and histogram of oriented gradients. *Int. J. Comput. Applic.*, 24: 6-10.
- Zhong, Z., J. Zhu and S.C. Hoi, 2015. Fast object retrieval using direct spatial matching. *IEEE. Trans. Multimedia*, 17: 1391-1397.

Secure Big Data Storage in Cloud and Data Computing Using Cryptographic Techniques

¹N. Madhusudhana Reddy, ²C. Nagaraju and ³A. Ananda Rao

¹Department of CSE, Jawaharlal Nehru Technological University Anantapur (JNTUA),
Andhra Pradesh (AP), India

²YSR Engineering College, Yogi Vemana University, Proddatur, Andhra Pradesh (AP), India

³Department of CSE Director of Academic and Planning, JNT University Anantapur,
Andhra Pradesh (AP), India

Abstract: Executing dispersed figuring empowers different routes for web-based organization offerings to address varying issues. Regardless, the data protection and assurance has transformed into a fundamental concern which limits numerous cloud submissions. One of the critical stresses in safety and assurance occurs by which that cloud executives have the possibility to accomplish and change the fragile data. This stress definitely extends customer's concern and abatements the acceptance of circulated processing in numerous fields, for instance, the money related business and regulatory associations. This study mainly concentrates on this problem and suggests an astute cryptographic technique where the cloud advantage admin can't particularly accomplish fragile data and meta cloud data storage Architecture for securing big data in cloud computing environment. The suggested approach partitions the record and freely inserts available dataset in the passed on cloud storage servers. An option is suggested to choose if the data groups require a split remembering the true objective to curtail the operation time. The suggested contrive Protection-Aware Efficient Distributed Storage (SA-EDS) illustrate which is essentially sustained by our suggested Secure Efficient Data Distributions (SED2) calculation. Our trial assessments have evaluated both safety and viability displays and the trial occurs represent that our approach can enough shield crucial dangers from fogs and requires with an suitable calculation instance.

Key words: Protection, cloud computing, delicate information, big data, distributed storage, calculation

INTRODUCTION

Dispersed figuring can portrayed as five properties, for instance, massive scalability, multi-inhabitation (shared resources), flexibility, pay only for used and self-provisioning of benefits (Purushothaman and Abburu, 2012; Schwarz and Miller, 2006). Dispersed registering enables customer to get to the remote servers encouraged on the web to store and process the data. Organization reproductions of cloud is orchestrated into 3 sorts (Wang *et al.*, 2009a, b, 2012) for instance, SaaS, PaaS, IaaS and particular sending sculpts are requested into private, public and hybrid. On account of the elevated openness of the cloud to all customers, appropriated figuring goes up against more noteworthy protection challenges (Murugesan and Sudheendran, 2013). These question are accumulated into two general classes as assurance issues confronted by cloud suppliers and protection issues confronted by customers.

As one of the tremendous progressions used as a piece of circulated processing, the passed on collection has engaged (Ma *et al.*, 2012; Bowers *et al.*, 2009) the mass remote data aggregation by methods for Storage-as-a-Service (SaaS) advantage illustrate. This cloud advantage show has completely transformed into a tasteful approach in tremendous data nearby the progression of web organizations and frameworks (Garg and Sharma, 2013; Zhou and Huang, 2012). Many cloud shippers have given engaging gathering organization offerings that give creature and versatile cloud-dependant storage locations for customers (Vijay and Reddy, 2013; Hendricks *et al.*, 2007). In any case, the protection issues occurred by the procedures on cloud region is so far a block of using SaaS for wanders. Many cloud customers stress over the sensitive data where the cloud heads have the passage (Gampala *et al.*, 2012). This issue mortifies existing use of SaaS in spite of the way that various prior investigates have watched out for this field.

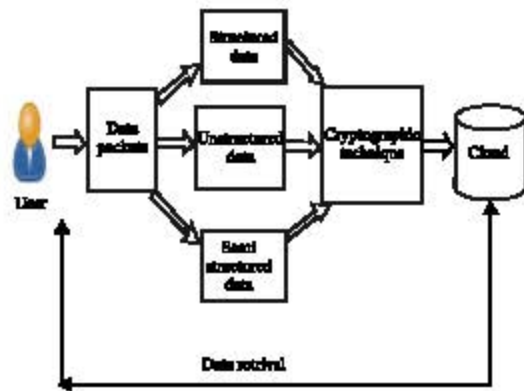


Fig. 1: Architecture of secure data storage in cloud

Additionally, Mass Distributed Storage (MDS) has investigated to increase the data aggregation measure starting late. The anomalous state presentations of the adaptable calculation are considered points of interest of completing MDS (Donald *et al.*, 2013a, b; Wang *et al.*, 2009 a, b). One point that necessities changes is to secure circled data collection in which the dangers begin from an arrangement of sides.

This research article contemplates on the issue of cloud executives misuse issues and efforts to keep away from cloud customer's data retrieved from cloud servers. The information is considered as organized, unstructured (Guha and Shrivastav, 2013; Donald *et al.*, 2013 a, b) or semi-organized information in which they are put away in the cloud. We recommend an insightful cryptographic measure, named Protection-Aware Efficient Distributed Storage (SA-EDS) exhibit that is planned to get a viable MDS advantage and furthermore strange state protection confirmations. Figure 1 delineates the working of SA-EDS (Wang *et al.*, 2009 a, b). As showed up in Fig. 1, customer's data are reviewed by an different technique in which open named-data packs frameworks are associated. This framework is essentially kept up by our suggested calculation. In addition, part information process is refined by the focal estimation, Secure Efficient Data Distributions (SED2) Algorithm which is intended to divide information recalling a definitive goal to shield precarious data from spilling on the cloud region utilizing scarcest expenses (Ma *et al.*, 2012; Ateniese *et al.*, 2007).

The delicate data recuperation needs an unscrambling methodology that is reinforced by our suggested calculation. The tremendousness of the suggested part as we give an flexible approach to manage those endeavors that intend to use SaaS (Raju *et al.*, 2011), however, require an anomalous state data collection protection, for instance, the cash related organization

industry. The guideline issue unwound by our suggested plot is keeping cloud supplier from direct accomplishing customer's one of a kind data (Garg and Sharma, 2013; Shah *et al.*, 2007). The essential duties of this article are twofold: We propose a narrative cryptographic technique for passing on mass scattered amassing by which customer's one of a kind data can't be direct come to by cloud heads. The recommended technique is a valuable cryptographic surmises for protecting destructive exercises happened on the cloud server. We propose a convincing information split part that does not pass on immense overheads (Purushotham an and Abburu, 2012; Bowers *et al.*, 2009) and similarly guarantees information retrievability.

Literature review: Zhou and Huang (2012) planned to ensure data responsibility for duplicates over the flowed amassing structure. They completed the PDP plan to hide different impersonations without performing encoding on each of duplicate liberally, giving accreditation that various copies of data are truly kept up. As of late, Garg and Sharm (2013) gave an examination on many existing courses of action on isolated data trustworthiness inspecting and argued their points of interest and hindrances under diverse blueprint circumstances of protected disseminated stockpiling organizations.

Singh *et al.* (2012) proposed the provably-protected arrangement to affirm the trustworthiness of data amassing in the cloud without downloading each one of the data. The producers utilized the RSA-based homomorphic Verifiable Tag (HVT) to convey a solitary tag by joining the piece marks. HVT besides engages a server to produce a verification and empowers the customer to verify about server limitations, despite the way that the customer won't not approach the pieces. Regardless, this process gets high server calculations or correspondence costs for the entire record on account of the usage of RSA numbering.

Shimpi *et al.* executed another kind of RDA methodology Proofs Of Accessibility (POA) to favor data respectability and to thwart data degradation with forward screw up modifying codes, remotely. The evaluator by then subjectively embeds the protected data blocks into pieces already trading towards the server. Regardless, the measure of demand in this framework is obligated to the measure of sentinel squares. The POA procedure acknowledged high calculational overhead on customer part which worked out as intended in perspective of playing out the blunder recuperation and information encryption shapes.

Ateniese *et al.* (2007) familiar a system with give adequacy and protection to the POR strategy in light of the BLS homomorphic confirmation method. The BLS technique empowers the evaluator to add up to the names into a settled size remembering the true objective to restrict the framework computational overhead, also, utilizes the Reed-Solomon policy to recoup the misunderstandings. Supporting dynamic information fortify satisfactorily is a crucial issue in most by far of the remote information evaluating systems in which the controller can intensely revive the outsourced data without recovering the outsourced record.

Wang *et al.* (2012) proposed a rank-dependent information to diagram a totally capable evaluating process. This procedure can also verify the uprightness of data squares of varied-sizes, yet, it can't affirm the reliability of individual piece. Most POR systems can't gainfully support dynamic data revive in light of the way that the server can't develop an association between the data squares and the mixed code-words.

The other weight of this method is that it can't adequately reinforce visit dynamic invigorate operations profitably due to the center point re-modifying issue and this will similarly realize elevated computational transparency on the analyst side. Additionally, bilinear coordinating calculation is more exorbitant than the logarithmic configuration which is used as a piece of our technique. The data inserted into the cloud is categorized as type 1 which is structured data and type 2 which is considered as semi-structured and unstructured data (Murugesan and Sudheendran, 2013). Here, two algorithms are used for performing encryption on both the types of data.

Algorithm 1; Types of data:

```

Selection of data (type)
If (type == 1)
{
  Structured data
  Data from SQL type data bases
  AESEncrypt()
  Send to Cloud
}
If (type == 2 or type == 3)
{
  Unstructured or semi structured data
  Data from Hbase or Casandra or MongDB. etc
  HomomorphicEncrypt()
}

```

Among the two types of data sets considered a encryption method is used to encrypt the data and then it will be stored in the cloud.

Algorithm 2; AES encryption:

```

AESEncryption ()
{
  Infer the arrangement of round keys form the figure key
  Introduce the state exhibit with the piece information (plaintext)
  Add the underlying round key to the beginning stste cluster
  Perform nine rounds of state control
  Duplicate the last state exhibit out as the scrambled information (ciphertext)
}

```

The AESEncrypt() method is used for only for organized information and after that the information will be safely put away in the cloud and it is dodged from misuse (Gampala *et al.*, 2012). If the information is utilized by any unapproved client then it must be decoded for understanding the information which can be done only by the authorized (Juels and Kaliski, 2007) person if the private key is provided.

Algorithm 3; Privacy key:

```

HomomorphicEncrypt ()
{
  Infer the arrangement of round keys form the figure key
  Introduce the stste exhibit eith the square information (plaintext)
  Add the underlying round key to the beginning stste cluster
  Perform nine rounds of state control
  Play out the tenth and last round of state control
  Duplicate the last stats cluster out as the encoded information (ciphertext)
  Step 1: Select any two prime numbers say p and q
  Step 2: compute the result of those two prime numbers
  Say N = p classified and Nis open.
  Step 3: select arbitrary number X and a root g of GF (p). Where g and X are littler then p
  Step 4: compute y = gx mod p. utilize this y for the encryption
  Step 5: encryption will be performed in following two stages:
  1. Select arbitrary whole number r and apply following homomorphic encryption
  E1 (M) = (M+r*p) mod N
  2. Select arbitrary whole number k and the encryption calculations are:
  Eg (M) = (a, b) = (gk mod p, yk E1 (M) mod p)
  Step 6: Decrypted calculation Dg() is M = b x (hatchet)- 1 (mod p)
}

```

The HomomorphicEncrypt() method is used for both semi and un structured data for performing encryption method and then it will be stored in cloud (Vijay and Reddy, 2013). The data which is encrypted with either of the methods are strongly secured from usage of data by unauthorized users.

MATERIALS AND METHODS

Secure storing of data into cloud: In cloud information collection structure, clients accumulate their information in cloud and never again have the information nearby. Thusly, the precision and transparency of the information records being secured on the scattered cloud servers must be ensured. The considerable issue is to adequately see any unapproved information adjustment and defilement, possibly as a result of server exchange off and also sporadic Byzantine disillusionments. Also, in the passed on circumstance at the point while such irregularities are effectively perceived, to determine the information held by server go of keeps in will be in like way of stunning criticalness, since, it can be the fundamental walk to quick recoup the cutoff messes up. To address these issues, our rule anticipate guaranteeing cloud information conglomeration is presented in this fragment.

The underlying fragment of a region is focused on a assessment of basic gadgets from convention speculation that is required in our arrangement for record spread transversely finished cloud servers.

Limit system design based on cloud computing:

Appropriated figuring virtualization development is used as a piece of the layout of limit system to finish high synchronization and high adjustment to interior disappointment under the condition of the consistency (Donald *et al.*, 2013a, b; Bowers *et al.*, 2009). The pro slave dispersal configuration is used as a piece of the diagram of data amassing to keep up a vital separation from the data hardship and damage caused by the power outage under the traditional aggregation development which grasps a lone strategy for limit structure. System uses allotted limit in different physical and support gathering contraption (Wang *et al.*, 2010, 2012). So, as to improve the protection and genuineness of data. Virtualized physical resources are facilitated into the pro center as the system organization center which is accountable for the organization and checking the step by step operation of the slave centers and notwithstanding ensure the run of the mill state of center points.

The expert center point sort out virtualization and flowed organization are progressed as the diagram thought to deal with the issues of pro center point organization bottleneck of the standard advancement. The pro center point is gone to by the lead recently visit, to a particular degree, relieve the passageway and organization of the pro center point inconvenience, keep up a vital separation from server disillusionments caused by consolidated access to the pro center which will incite

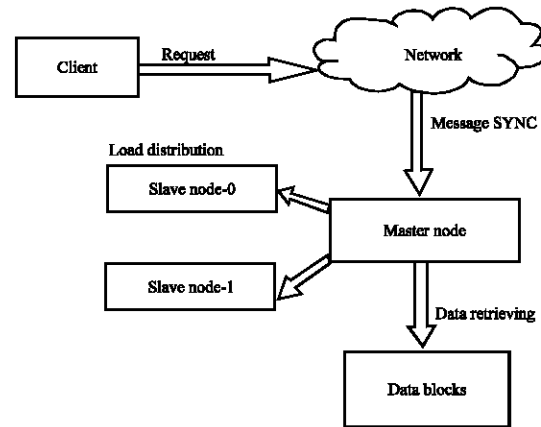


Fig. 2: Data storage method

the fold of whole system and furthermore to deal with the issue of the bottleneck of the whole structure operation, consequently improve the general capability. Various slave centers as the data amassing, finishes stack balanced scattering of set away data. The inconvenience of limit of different data sorts under the standard social database is handled and various of support amassing of data is spread into data center points (Raju *et al.*, 2011; Ma *et al.*, 2012). The loss of data is avoided on the begin of common system operation execution, the limit creation arrangement is showed up in Fig. 2.

Limit structure get the opportunity to stream begins with get the opportunity to request sent by client by then the message gets the chance to close expert center server through framework (Purushothaman and Abburu, 2012). After got and responded to customers request to examine and create, the center point will discover snippet of data in demonstrated slave center as showed by the convey to execute decided operation. Data is secured by allocate limit, accommodating in organization. With the usage of examination of discernment gadgets (Vijay and Reddy, 2013; Wang *et al.*, 2012), electronic report is made in the customer’s terminal illustrate which gives comfort in examination and decision. For whatever period of time that interfacing with the Internet client terminal can get to the structure. The terminal can be hardware, settled and mobile phones and introduced contraptions.

The steps involved in inserting the data into cloud environment is clearly explained in Fig. 2 in which based on client request the data is encoded using master node and ten it will be divided as data blocks and t hen stored in the cloud.

Protection-Alert Efficient Disseminated Storage (PA-EDS) illustrate:

Our suggested SA-EDS show for the most part contains two sections, to be particular

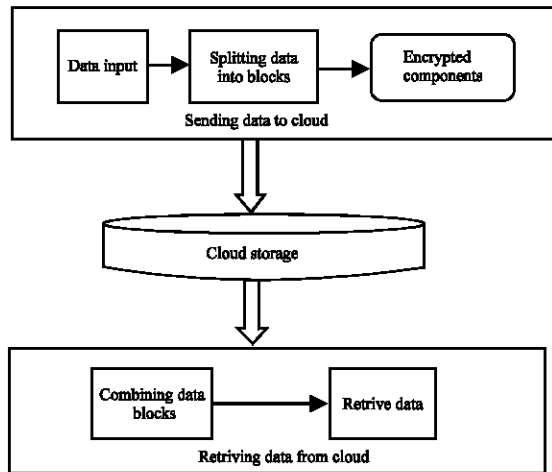


Fig. 3: Process of data stored and retrieved from cloud

Deterministic Process (DP) and Data Distributed Storage Process (D2SP). The imperative segment is intended to pick if the information bundles request an abnormal state assurance ensure. The other part is utilized to shield information from the unforeseen exercises occurred by cloud-side workers. It is an inside piece of our recommended represent.

Data Disseminated Storage Process (D2SP): Figure 3 addresses an irregular state work process configuration of D2SP in the suggested SA-EDS illustrate. The recommended Fig. 3 addresses the standard plan of our point of view. The two sorts of information shape two essential strategies amidst the information transmissions.

Besides, at the accessing data available from cloud segment information clients need to get information bundles from either of cloud suppliers. Achieving the first information requires a progression of methods subsequent to the information parcels are gotten from cloud sides. In the first place, the cor-reacting information bundles should be binded up to create the new information series. Later, clients will utilize the way to do two activities took after by the underneath arrange, XOR methods to the information string and include the Key information esteem after. The first information will be increased after this method is done.

Unsafe methodologies: The cloud server generally speaking accept a confidential part in cloud advantage sending methods with the ultimate objective that many cloud advantage show planners expect that the overseers

on the cloud-part are protected. Be that as it may, many threats are caused by the sudden be-haviors made by cloud administrators as opposed to the pernicious assaults. Much of the time, it transforms into a profound quality issue rather than a specialized issue, since cloud workers more often than not require the entrance to the information with the end goal of the information administration despite the fact that the exercises are limited by the directions. Then information are not secure in spite of the fact that encryptions are dependent. The data can be discharged in an extraordinary possibility if the malignant procedures are considered.

Plan goals: Our suggested framework plans to all the while accomplish a couple of focused exhibitions as takes after which can ensure the information protection required by specific information clients, for example, monetary specialists or inspecting experts: Keeping dangers from inside dangers: we hope to fulfill a bigger sum insurance data aggregation part information into different cloud servers with interior dangers can neither misuse the data nor recoup the information from the set away data into server. Data confirmation beside exterior dangers: The recommended scheme will shield data from the strikes concerned by the outside enemies. Data ought to be mixed in the midst of the conduction methodology. High viability data taking care of: Our structure will similarly keep up a key separation from high correspondence and calculation overhead remembering the true objective to drop down the torpidity. The succeeding segment portrays the principle algorithms utilized as a part of our suggested demonstrate.

Secure Efficient Data Divisions (SED2) algorithm: SED2 algorithm is intended to achieve the information handling previous transmission to cloud space. The yields incorporate two separate encoded information α and β . Executing SED2 calculation can skillfully shield the hazard models. In ARCT hazard illustrate, acknowledge that cloud delegates have the key and can access to the data on the server. Deficient data don't contain any information while the most important data won't be gained until the point that two areas are cooperated. Furthermore in MAT hazard illustrate, the foes have foundation data about the information and mean to mishandle the information. In this manner, our suggested plan can adequately shield both risk models in the hypothetical point of view. Simulated codes of SED2 is represented in Algorithm 4 and 5.

Algorithm 4; Codes of SED2:

Algorithm: Secure Efficient Data Distributions (SED2)algorithm.
 Requir: D, C
 Ensure: α, β
 1: Input D, C
 2: Initialize R= 0, α = 0, β = 0
 3: /*C is a random binary that is shorter than D*/
 4: Randomly generate a key K

 5: for \forall input data packets do
 6: if D == C&&C == 0 then
 7: DO R= D- C
 8: α = C_K
 9: β = R_K
 10: endif
 11: end for
 12: Output α, β

The main steps of Algorithm are given as follows:

Algorithm 5; SED2:

1. Input data packet D and C. Data C needs to be a non-empty set that is shorter than D. C should not be as same as D. Create and initialize a few data set, R, α and β ; assign 0 value to each of them
 2. Randomly generate a key K that is stored at the User's special register for the pupose of encryption and decryption. This is the crucial part for protecting privacy before the data sre sent out
 3. We calculate the value of R by (D-C), then execute two XOR operations to obtain the data value stored in the clouds. The data in the renote storage are denoted to α and β . We use the following formulas to obtain α and β : $\alpha = C_K$; $\beta = R_K$
 4. Output α and β and sparately story them in the different cloud servers

RESULTS AND DISCUSSION

This area demonstrateed a few test happens made in our execution evaluations. Figure 4 and 5 laid out an examination of the completing point among EDS and AES. We utilized the same measured information and dismembered the encrypted point in time employments. Figure 4 and 5displayed several outcomes that were made under setting 1-1 and 1-2. As exhibited by the lines appeared in Fig 4 and 5, our recommended plot had a shorter execution time than AES under both demonstrated situations. The unraveling time necessary is more broadened day and age under the two settings.

The number of iterations leves performed on the data sets for providing protection with encryption methodologies is clearly depicted in Fig. 4 which compares the levels with existing system.

The importance of the use of data for the wander is explored in this paper in the underlying stride. To satisfy the demand of huge data taking care of stage, data application arrange in Hadoop and data blend organize in data stockroom are progressed. Disseminated processing advancement is grasped in the arrangement of limit

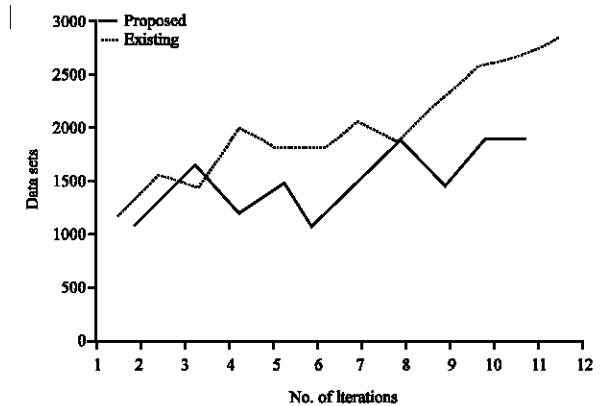


Fig. 4: Iteration level comparison with existing and suggested algorithms

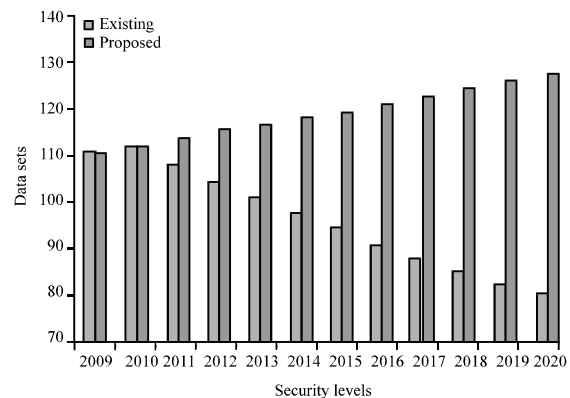


Fig. 5: Protection levels in existing and suggested systems

structure and advances the standard development in the expert center point server organization in coursed limit by contrasting the protection levels are shown in Fig. 5. Through separating and recognizing key developments of limit structure, for instance, archive square invigorate framework and accuse recovery instruments, achievable particular game plans away organization for colossal data are given.

CONCLUSION

In this study, we look at the issue of data protection in cloud data gathering which is fundamentally an appropriated aggregation structure. To finish the affirmations of cloud data respectability and openness and maintain the idea of reliable dispersed stockpiling advantage for customers, we propose a fruitful and versatile passed on contrive with unequivocal dynamic data reinforce including square revive, delete and include. By utilizing the time, calculation resources and associated

online weight of customers, we also give the continuation of the suggested guideline intend to help third party assessing where customers can safely choose the uprightness checking endeavors to outcast evaluators and be easy to use the circulated stockpiling organizations. Through distinct protection and expansive examination comes to fruition, we exhibit that our arrangement is exceedingly viable and solid to Byzantine dissatisfaction, toxic data change strike and altogether server plotting attacks.

SUGGESTIONS

In this study, we suggested meta cloud data storage architecture for securing big data in cloud computing environment. Guide reduce structure is used to find the amount of customers who were marked into the cloud server cultivate. Suggested framework guarantees the mapping of various data segments to each provider using meta cloud data storage interface. Despite the way this suggested approach requires high use effort, it gives critical information to circulated processing condition that can have high impact on the bleeding edge structures. Our future research is to grow the suggested meta cloud data storage architecture for steady planning of spouting data.

REFERENCES

Ateniese, G., R. Burns, R. Curtmola, J. Herring and L. Kissner *et al.*, 2007. Provable data possession at untrusted stores. Proceedings of the 14th ACM Conference on Computer and Communications Security, October 28-31, 2007, ACM, Whistler, Canada, ISBN:978-1-59593-703-2, pp: 598-609.

Bowers, K.D., A. Juels and A. Oprea, 2009. HAIL: A high-availability and integrity layer for cloud storage. Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS'09), November 09-13, 2009, ACM, Chicago, Illinois, ISBN:978-1-60558-894-0, pp: 187-198.

Donald, A.C., S.A. Oli and L. Arockiam, 2013. Mobile cloud protection issue and challenge: A perspective. Intl. J. Eng. Innovative Technol., 3: 401-406.

Donald, A.C., S.A. Oli and L. Arockiam, 2013. Mobile cloud security issues and challenges: A perspective. Intl. J. Electron. Inf. Technol., 3: 1-6.

Gampala, V., S. Inuganti and S. Muppidi, 2012. Data security in cloud computing with elliptic curve cryptography. Intl. J. Soft Comput. Eng., 2: 138-141.

Garg, P. and V. Sharma, 2013. Secure data storage in mobile cloud computing. Intl. J. Sci. Eng. Res., 4: 1154-1159.

Guha, V. and M. Shrivastava, 2013. Review of information authentication in mobile cloud over SaaS and PaaS layers. Intl. J. Adv. Comput. Res., 3: 116-121.

Hendricks, J., G.R. Ganger and M.K. Reiter, 2007. Verifying distributed erasure-coded data. Proceedings of the 26th Annual ACM Symposium on Principles of Distributed Computing, August 12-15, 2007, ACM, Portland, Oregon, ISBN:978-1-59593-616-5, pp: 139-146.

Juels, A. and B.S.Jr. Kaliski, 2007. PORs: Proofs of retrievability for large files. Proceedings of the 14th ACM Conference on Computer and Communications Security, October 29-November 2, 2007, ACM, New York, USA., ISBN:978-1-59593-703-2, pp: 584-597.

Ma, L., J. Sum and Y. Li, 2012. Comparing general paradigm on data secrecy protection for outsourced file in mobile cloud computing. J. Netw., 7: 1449-1455.

Murugesan, K. and S. Sudheendran, 2013. Ensuring user security and data integrity in multi-cloud. Intl. J. Soft Comput. Eng., 3: 355-358.

Purushothaman, D. and S. Abburu, 2012. An approach for data storage security in cloud computing. Intl. J. Comput. Sci. Issues, 9: 100-105.

Raju, V., R. Kumar and A. Raj, 2011. Techniques for efficiently ensuring data storage security in cloud computing. Intl. J. Comp. Tech. Appl., 2: 1717-1721.

Schwarz, T.S. and E.L. Miller, 2006. Store, forget and check: Using algebraic signatures to check remotely administered storage. Proceedings of the 26th IEEE International Conference on Distributed Computing Systems (ICDCS'06), July 4-7, 2006, IEEE, Lisboa, Portugal, pp: 12-12.

Shah, M.A., M. Baker, J.C. Mogul and R. Swaminathan, 2007. Auditing to keep online storage services honest. Proceedings of the 11th USENIX Workshop on Hot Topics in Operating Systems, May 7-9, 2007, USENIX Association, Berkeley, California, USA., pp: 1-6.

Singh, K., I. Kharbanda and N. Kaur, 2012. Protection issue occurs in cloud computing and their solution. Intl. J. Comput. Sci. Eng., 4: 945-949.

Vijay, G.R. and A.R.M. Reddy, 2013. Data protection in cloud based on trusted computing environment. Intl. J. Soft Comput. Eng., 3: 187-191.

- Wang, C., Q. Wang, K. Ren and W. Lou, 2010. Privacy-preserving public auditing for data storage security in cloud computing. Proceedings of the 2010 IEEE Conference on INFOCOM, March 14-19, 2010, IEEE, San Diego, California, pp: 1-9.
- Wang, C., Q. Wang, K. Ren and W. Lou, 2009a. Ensuring data storage security in cloud computing. Proceedings of the 2009 17th International Workshop on Quality of Service IWQoS09, July 13-15, 2009, IEEE, Charleston, South Carolina, USA., ISBN:978-1-4244-3875-4, pp: 1-9.
- Wang, C., Q. Wang, K. Ren, N. Cao and W. Lou, 2012. Toward secure and dependable storage services in cloud computing. IEEE. Trans. Serv. Comput., 5: 220-232.
- Wang, Q., C. Wang, J. Li, K. Ren and W. Lou, 2009b. Enabling public verifiability and data dynamics for storage security in cloud computing. In: Computer Security-ESORICS 2009, Backes, M. and P. Ning (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-04443-4, pp: 355-370.
- Zhou, Z. and D. Huang, 2012. Efficient and secure data storage operations for mobile cloud computing. Proceedings of the 8th International Conference on Network and Service Management (CNSM) and 2012 Workshop on Systems Virtualization Management (SVM), October 22-26, 2012, IEEE, Nevada, USA., ISBN:978-1-4673-3134-0, pp: 37-45.

Hardware Implementation and Testing of PAPR Reduction Using Order Bit Selector and Trellis Structure

¹K.M. Gayathri, ²S. Bhargavi and ¹N. Thangadurai

¹Department of Electronics and Communication Engineering,

School of Engineering and Technology, Jain University, Bangalore, India

²Department of Telecommunication Engineering, S.J.C.I.T, Chickballapur, India

Abstract: Orthogonal Frequency Division Multiplexing (OFDM) is one of the well known and widely using multicarrier modulation method in digital communication system. In OFDM, the important parameter to be considered before transmission is to minimize the Peak Average Power Ratio (PAPR). In this study, the design and implementation of OFDM system along with Fragmentary Control Transmit (FCT), order bit selector and trellis structure is used to reduce the PAPR. The modified OFDM system is implemented on an FPGA board. The implementation results show the study of RTL schematic and test bench of the design. The results obtained from the proposed system have been compared with the existing system. It has been observed that the proposed system considerably better than the existing system.

Key words: Order bit selector, fragmentary control transmit, trellis structure, OFDM, PAPR, FCT

INTRODUCTION

One of the multicarrier modulation technique used in wireless communication is orthogonal frequency modulation. It as has been accepted for several wireless standards because of its advantages such as large spectral efficiency, robustness against multipath fading and less complexity in implementation. OFDM is a good modulation scheme for high data rate communication system. However, the practical implementation of OFDM is limited to certain extent because of its high Peak Average Power Ratio (PAPR) (Han and Lee, 2005). Due to its high PAPR it demands for excess transmitter power amplifier which has a very good linear range. It also demands for increased complexity design of data converters. Due to increase of sub carrier the high PAPR is obtained which is not effective. The existing schemes include Active Constellation Extension (ACE) (Krongold and Jones, 2003), Tone Reservation (TR) (Chen *et al.*, 2011) and Tone Injection (TI) (Abdullah *et al.*, 2011; Gayathri *et al.*, 2017a, b) amplitude clipping and filtering (Wahab and Ain, 2010) and multiple signal representation techniques such as Iterative flipping algorithm (Tellado, 1999), Partial Transmit Sequence (PTS) (Muller and Huber, 1997; Cimini and Sollenberger, 2000), Selective Mapping (SLM) (Bauml *et al.*, 1996). These techniques achieve PAPR minimization at the cost of increase in transmit power of signal and Bit Error Rate (BER). The data loss and computational complexity also increases.

Peak average power ratio: OFDM will have many modulated sub carriers which lead to problem of high PAPR. The coherent summation of N modulated sub carriers, peak power is increased by N than the average power in the base band signal. Hence, the intricate wrap of the OFDM transmitted signal is written as:

$$S(t) = \frac{1}{\sqrt{Y}} \sum_{m=0}^{M-1} M_m e^{j2\pi m \Delta f t}, 0 \leq t \leq MT \quad (1)$$

Where:

j = It is given by $\sqrt{-1}$

Δf = The spacing of subcarrier

MT = It is given by the effective data block duration

PAPR is described as the proportion of maximum instantaneous power to its average power during OFDM symbol duration. The peak average power ratio of OFDM symbol can be represented as:

$$PAPR = \frac{\max_{0 \leq t \leq NT} [|x(t)|^2]}{1/NT \int_0^{NT} |x(t)|^2 dt} \quad (2)$$

where, $x(t)$ is input signal. The idea behind scaling down PAPR is to decrease the maximum power of signal $x(t)$. Since, predominance of the system demands the discrete-time signals, dealt is sampled by amplitude in most of the peak average power reduction techniques. The time domain L-times oversampled samples are

derived by an LN-point Inverse Fast Fourier Transform (IFFT) of the given data block by considering zero-padding.

FPGA implementation: A Field-programmable Gate Array (FPGA) is an integrated circuit design, after manufacturing it the designer will configure the design. The configuration of FPGA is by using a specified Hardware Description Language (HDL). Any logical function can be implemented using FPGAs. FPGA consists of programmable logic components such as logic blocks and a reconfigurable interconnect which connects the block together. It is like numerous logic gates that can be connected in different configurations. The simple and complex combinational functions are performed by configuring the logic blocks. In many of the FPGAs, the logic blocks consist of block of memory or memory elements (Lakkannavar, 2012; Reddy and Reddy, 2013). Along with digital functions, some FPGAs also have analog features. The slew rate and drive strength are the most common analog feature which allows the designer to set slow rates and faster rate on lightly and heavily loaded pins, respectively on high-speed channels. The data converters can be integrated along with signal blocks in some mixed signal FPGA (Kaur and Mehra, 2012; Farahani and Eshghi, 2007).

Figure 1 depicts the Atrix-7 FPGA board. The FPGA used for this project implementation is Artix-7. The Artix 7 family FPGA is fabricated on 28 nm high performance, low power product which is used for mainly bio medical instruments, military radios and some compact wireless system. One of the high end features of Artix-7 FPGAs is integration of advanced Analog Mixed Signal (AMS) technology. It has eight 7 segment display, 16 switches and 16 output LED. The board can be connected to computer using Jtag cable where the communication happens between the system and the Atrix 7 board.

PAPR reduction approaches: There are two main categories to reduce PAPR in OFDM signal namely signal distortion technique and signal scrambling technique.

Signal distortion techniques: One of the most effective and simple technique to minimize PAPR is amplitude clipping and filtering. The logic behind this is to set a threshold value, if the signal is above the threshold value then the amplitude peak is clipped off. If the threshold value is more than the samples clipped will be less which decreases the CCDF. If the CCDF decreases the clipping threshold increases and PAPR reduces. Clipping is a non linear process where the distortion is viewed as noise. Another approach to reduce PAPR is peak windowing technique. Here, the Gaussian shaped window is multiplied to a huge signal peak to minimize PAPR. The

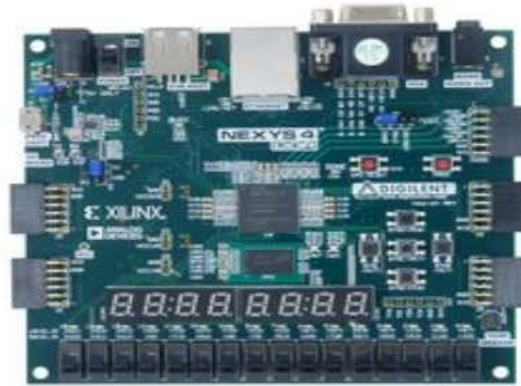


Fig. 1: Atrix-7 FPGA board

obtained spectrum is a convolution of indigenous OFDM and window spectrum. This technique scales down PAPR at a demand of raise in bit error rate.

Signal scrambling techniques: Linear codes are also used to reduce PAPR. In this method the proper codeword is chosen for transmission. Coding scheme uses a known data block with constellation modulation such as QPSK and PSK. The code word which is having high peak power is blocked from transmission. This method is useful only for limited number of sub carriers where PAPR reduction is achieved.

The main objective of tone injection approach is to expand the size of constellation and map the individual constellation to its analogues points. For every symbol in the data block there are many constellations which are mapped in extended constellation. The amount of scaling down of PAPR is dependent on the modified symbols in the data block. As this method uses the same frequency band for injected signal, it is more complicated when compared to tone reservation. Due to injected signal the power requirement and the implementation complexity increase. Tone reservation is one of the efficient methods used for minimization of PAPR. The fundamental concept is to reserve the tone both at transmitter and receiver. In tone reservation the main aim is to know the value of time domain signal with the help of convex optimization problem which can be determined using linear programming.

Active constellation technique is another method used for PAPR reduction. The outer signal constellation is expanded towards outside of the indigenous constellation, so, the PAPR is scaled down. This can be implemented in different modulation techniques such as QAM, QPSK and MPSK. The combination of additional signals can be used for peak cancellation. This method reduces the bit error rate and no side information is required for transmission. But it demands for a larger

constellation size in modulation technique. One of the probabilistic schemes used for PAPR reduction is selected mapping. This method creates a similar OFDM signals in time domain which are asymptotically independent. For these parallel input data the phase sequence are multiplied and then IFFT is applied for each data. Out of that many data series the one having a less PAPR is chosen and transmitted. Each data block is weighted with a unique phase sequences which results in ‘S’ different blocks of data. Along with these altered data block the one with less PAPR is selected for transmission along with the corresponding phase vector as side information. The reduction of PAPR relies on the design of phase factor and number of phase sequences.

In OFDM the powerful probabilistic based PAPR reduction methods is Partial Transmit Sequence (PTS). In this scheme the actual data X is splitted into N non-overlapping sub blocks. In each individual sub-block the sub carriers are weighted by a phase factor. The selection of phase factors make sure the PAPR is reduced. The flow of PTS algorithm is the OFDM sub carrier is split into M separate sub series. For every sub series the OFDM signal is generated by taking IFFT. The OFDM signal combined with weighted phase factor b_i . By using optimization algorithm the phase factors are generated. To retrieve the data at the receiver, receiver should have the generation scheme.

MATERIALS AND METHODS

Proposed algorithm: The binary data is the input to the OFDM system. The stream of binary data is given to the convolution encoder. The encoded samples are modulated using QAM modulator. After modulation the discrete cosine transform and wavelet transform is applied in order to minimize the PAPR of OFDM system. The output is then applied the conventional OFDM and PTS technique. The output from PTS technique is given to the AWGN channel and at the receiver end the inverse procedure of the transmitter will be done and finally the BER is calculated (Gayathri *et al.*, 2017a, b).

Order bit selector: In communication systems the simplest way of channel coding technique is convolution coding. By virtue of its high coding gain and performance it is frequently used in wireless communication systems. The convolution coding is suitable for noisy channel such as Additive White Gaussian Noise (AWGN). In order to control the error the channel encoding is used. The basic building block of the convolution encoder is a shift register. It requires less hardware and storage memory. The convolution encoder is mainly defined by 3 variables namely, n, k, L where ‘n’ is number of input bits,

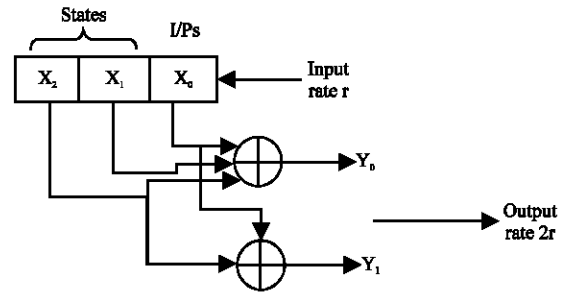


Fig. 2: Block diagram of convolution encoder

‘k’ is number of output bits and L is number of shift registers used as memory elements which are also called as constraint length. The code rate r is given by $r = k/n$. Figure 2 shows the sketch of convolution encoder for $(n, k, L) = (2, 1, 2)$.

To obtain output at the encoder the two previous input bits and one present input bit are used. The output Y_0 and Y_1 is generated by using modulo-2 addition which is denoted by:

$$\begin{aligned}
 Y_0 &= x_0 \oplus x_1 \otimes x_2 \\
 Y_1 &= x_0 \oplus x_2
 \end{aligned}
 \tag{3}$$

After convolution coding the order bit selector is used to select the QAM constellation such that the PAPR is reduced.

Trellis diagram with Viterbi decoder: At receiver for convolution encoder viterbi decoder is used. Viterbi decoder uses two basic operation synchronization and quantization. The synchronization is used to know the range limits of code word and symbol. The quantization is used to quantize the analog signal and converter to digital using quantization square. There are two types of quantization techniques used in viterbi decoder. They are:

- Hard decision
- Soft decision

Hard decision: The decoding process uses the trellis diagram and Hamming distance. It is quantized into one bit precision either 0 or 1. The Hamming distance is used to measure the distance between the expected data at the decoder and the data transferred from the encoder.

Soft decision: The information when transmitted over a Gaussian channel is decoded using probability decoding. It uses multi bit quantization for received bits. If there 3 or 4 bits of quantization the performance is better than the hard decision. The Euclidian distance is used to measure the distance between the bits. Figure 3 shows the internal

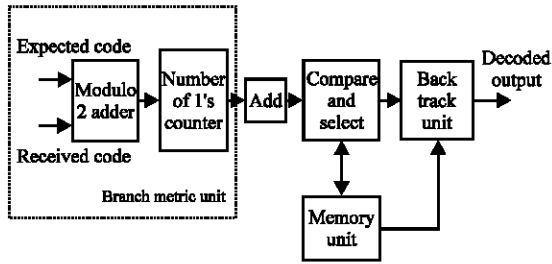


Fig. 3: Internal architecture of viterbi decoder

flow diagram of viterbi decoder. The functional blocks of viterbi decoder are branch metric unit, add compare and select unit, memory unit and back track unit. The steps involved in the viterbi decoder are as follows:

- The two parallel bits are inputs to the viterbi decoder
- The Hamming distance of the expected code and received code is calculated by using modulo 2 additions. Number of one's is counted to measure the distance
- The previous stage and present values are added and compared to select with the minimum path value to reach the next node
- The each stage calculation is stored in the memory for further processing
- The back track unit is used to compare and track the optimal path value and the corresponding output is produced

Trellis diagram: Figure 4 shows the general trellis diagram used in decoder. It has four rows of horizontal dots where each row depicts one state of encoder. The solid line joining the dots illustrate the input transition bit is one and the dotted line connecting the dots represent when the input transition bit is zero. To achieve better performance soft decision is used. The distance between the received codes and all possible codes are computed. Here, the Hamming distance is used to compute the distance. The computation of Hamming distance is easy, it counts how many bits are different from received code to the all possible codes. The output of the hamming distance can be 0, 1 or 2. At each unit of time the hamming distance is computed and it is called as branch metric. These values are stored and accumulated to compute the optimal path. The FPGA flow includes the following steps.

Design entry: To achieve the proposed system the HDL language such as verilog is used.

Synthesis: The proposed design is synthesized into a hardware circuit that includes the logic blocks.

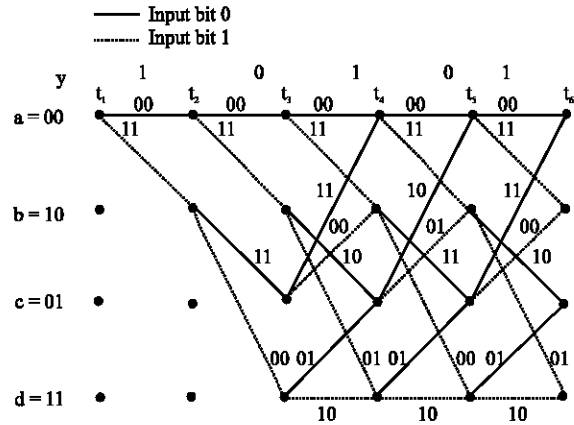


Fig. 4: Trellis structure

Functional simulation: The synthesized circuit is verified with its functional working.

Implementation: The logical blocks are placed into the actual FPGA with the help of netlist. It also chooses interconnecting wires in the chip to make the required connections.

Programming and configuration: The proposed circuit is programmed on to a FPGA chip by fusing the switches that composed of logical blocks and establish the required wiring connections.

RESULTS AND DISCUSSION

Quadrature Amplitude Modulation (QAM): Figure 5a shows the RTL schematic of QAM modulator and demodulator. The QAM is a combination of amplitude and phase shift keying. Here, the bits entering to the modulator is divided into two equal parts namely quadrature and in phase which is multiplied by cosine and sine values, respectively. These 2 modulated signals are combined at the source and then transmitted. At the receiver the two signals are separated and sampled at every time instant. Then the decision is made accordingly it is one or zero. For the different phase shift different mux is been used as shown in the Figure 5a. The input data is 8 bits for those bits the sine and cosine values are generate and multiplied in order to get the different phase shift. These two multiplied signals are summated and then transmitted.

Figure 5b depicts the simulation results of QAM modulation. Here, we can observe that the rst line should be always high in order to obtain the output. After 15.65 nsec the output is obtained. At every falling edge of the clock we can observe the output change.

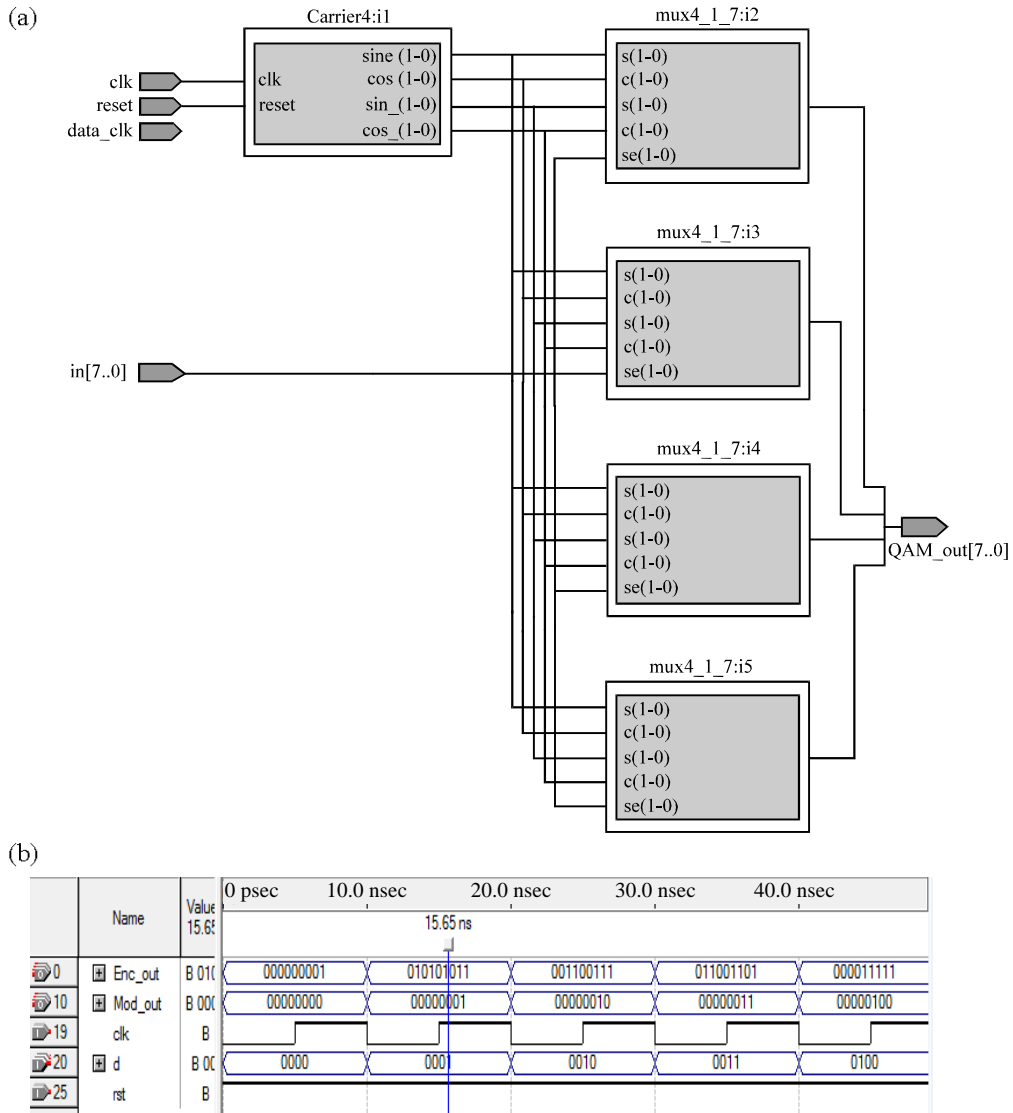


Fig. 5: a) RTL Schematic of quadrature amplitude modulation and b) Simulation results of quadrature amplitude modulation

Order bit selector: In the presented method, the data input is converted into binary stream. The binary stream is encoded by using the order bit selector technique as shown Fig. 6a. This example uses the rate 2/3 feed forward order bit selector depicted in the Fig. 6a.

OFDM: Figure 7b shows the simulation results of conventional OFDM. The random 5000 samples are taken at the input of the OFDM and the output of the OFDM is modulated and then transmitted. These 5000 samples which is entering the modulator is divided into 8 bits of data. The white colour oval in Fig. 7b shows the change in modulated output for every different 8 bits of the incoming input.

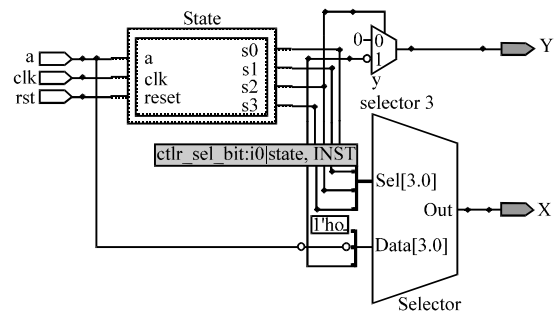


Fig. 6: A RTL schematic of order bit selector

Figure 7a shows the RTL schematic of conventional OFDM modulator. The input is modulated using QAM

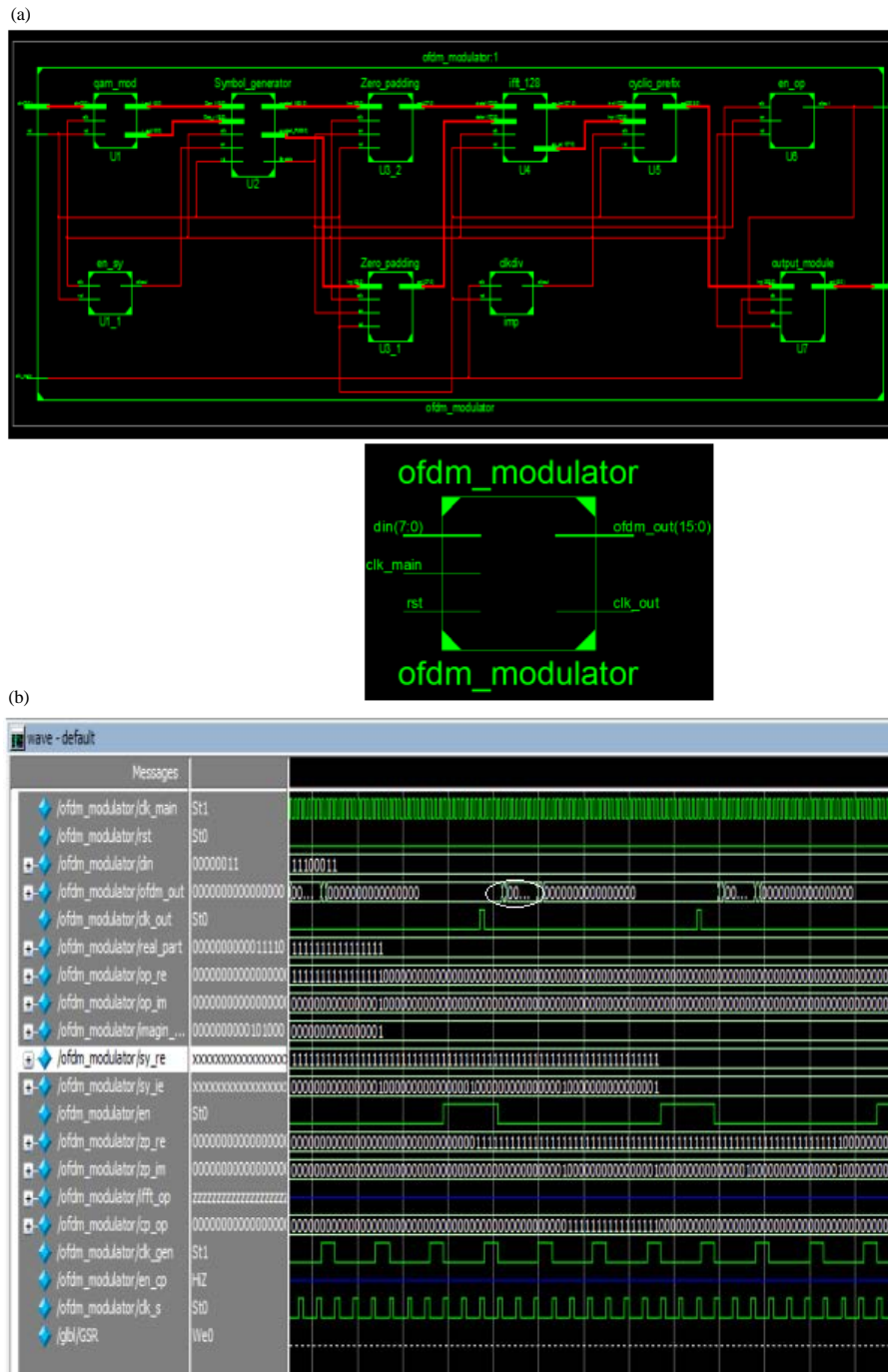


Fig. 7: a) RTL schematic and top view of OFDM system and b) Simulation result of OFDM system

modulator. After modulation the symbols are generated where each symbol are of 8 bits. The output of the symbol generator block is given as the input to the zero padding and then the 128 point IFFT is calculated. The output of IFFT is fed to cyclic prefix. In order to avoid ISI and ICI the cyclic prefix is done. Then finally the output module gives the conventional OFDM output.

Proposed OFDM system: Figure 8a depicts the simulation results of presented OFDM system. The random 5000 samples are taken at the input of the OFDM and the output of the OFDM is modulated and then transmitted. For the same conventional OFDM the IDCT and IWPT are combined and then the modulated signal is obtained.

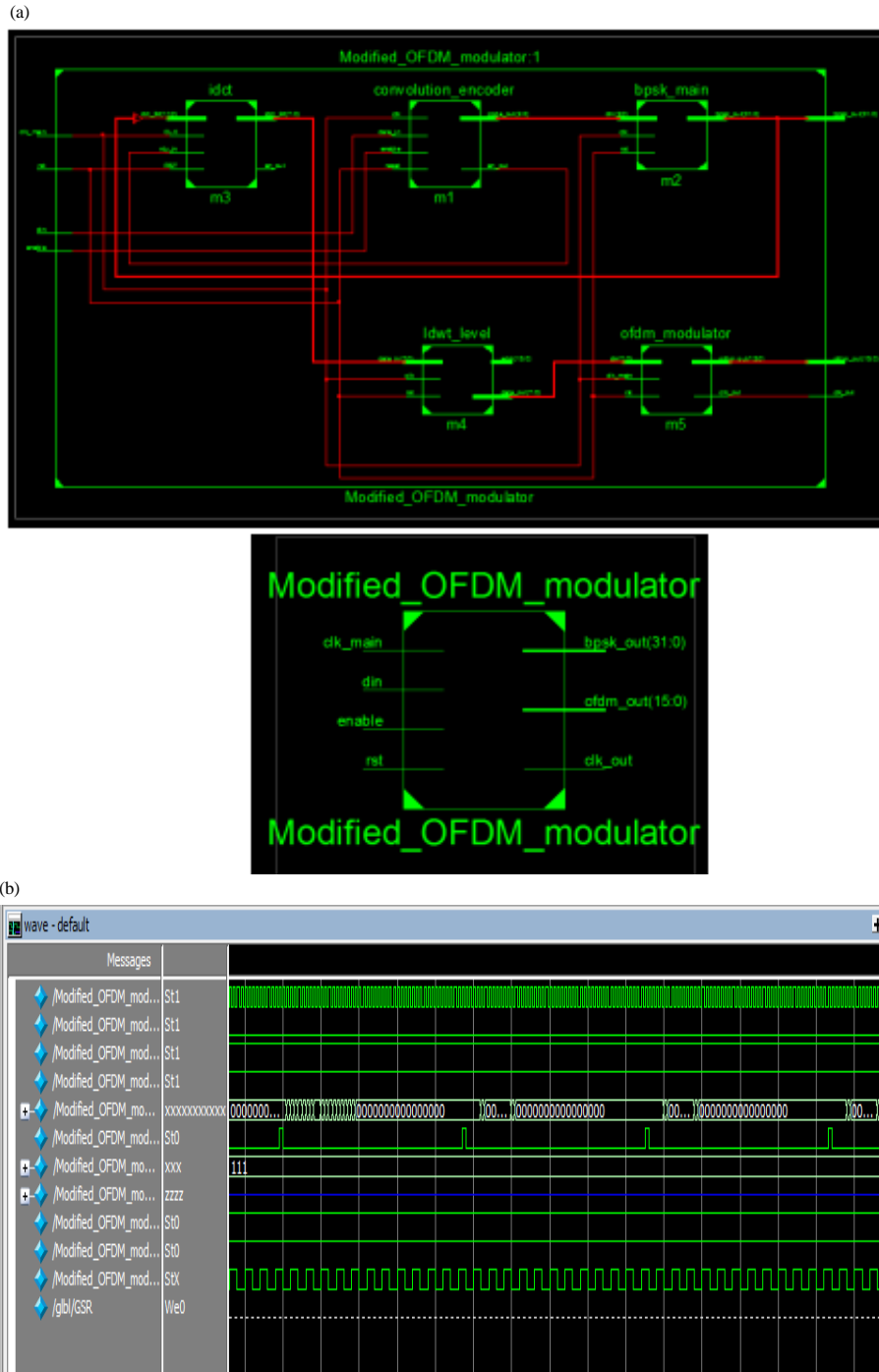


Fig. 8: a) RTL schematic and top view of proposed OFDM system and b) Simulation result of proposed OFDM system

Figure 8b shows the RTL schematic of modified OFDM modulator. The input is given to the convolution encoder. The output of convolution encoder is modulated using BPSK. The output of the modulator is given to the IDCT for auto correlation. The output of the IDCT is given to the wavelet transforms and then

the final output is taken after the conventional OFDM modulator. The white colour oval shape on the Fig. 8b shows the output of the modified OFDM for every change in the input. We can observe that after 16 clock cycles we can see the output of the OFDM modulator.

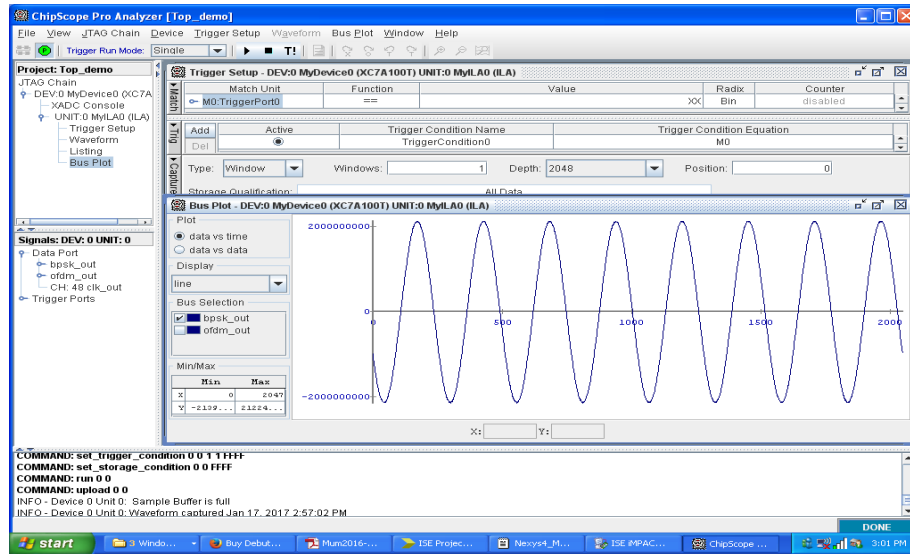


Fig. 9: Chipscope pro-result

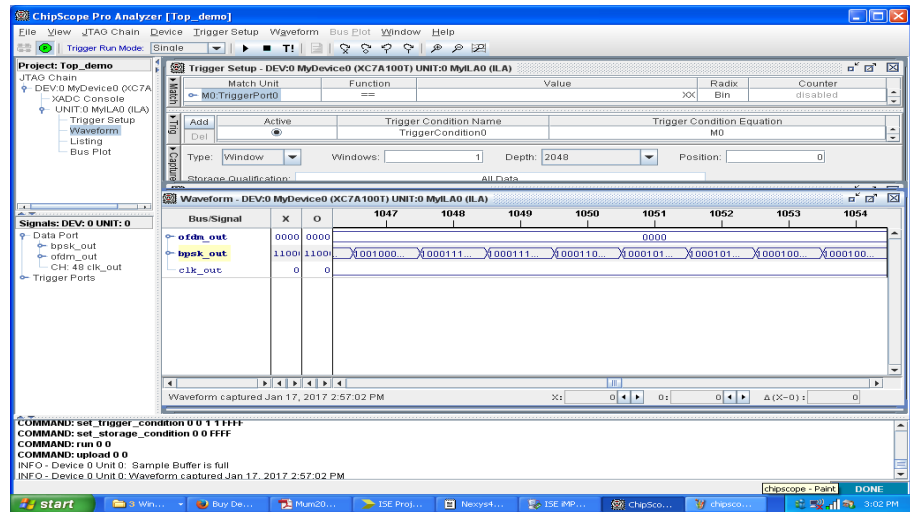


Fig. 10: OFDM result on chipscope pro

Chip scope is embedded software logical analyser for intermediate signal. The results of modulation technique on chip scope are shown in Fig. 9. Figure 10 shows the entire proposed design result on chip scope.

The FPGA board is connected though the JTAG cable. The entire code is converted into bitmap file and then it is dumped on to the selected FPGA board. Figure 11 shows the code dumped on to the FPGA successfully. Figure 11 shows the output on FPGA board. Once, the enable pin is high for the given input the outputs will be displayed on LED.

The area utilised for the proposed design is shown in Table 1 and 2. The estimated number of slice registers,

Table 1: Device utilization data

Device utilization summary (estimate values)

Logic utilization	Used	Available	Utilization(%)
Number of slice registers	1962	126800	1
Number of slice LUTs	6989	63400	11
Number of fully used LUT-FF pairs	1350	7601	17
Number of bonded IOBs	53	210	25
Number of BUFG/BUFGCTRLs	3	32	9
Number of DSP 48E1s	30	240	12

Table 2: Timing summary

Particulars	Values
Speed grade	-3
Minimum period	17.366 nsec
Minimum input arrival time before clock	1.298 nsec
Maximum output required time after clock	35.678 nsec
Maximum combinational path delay	No path found

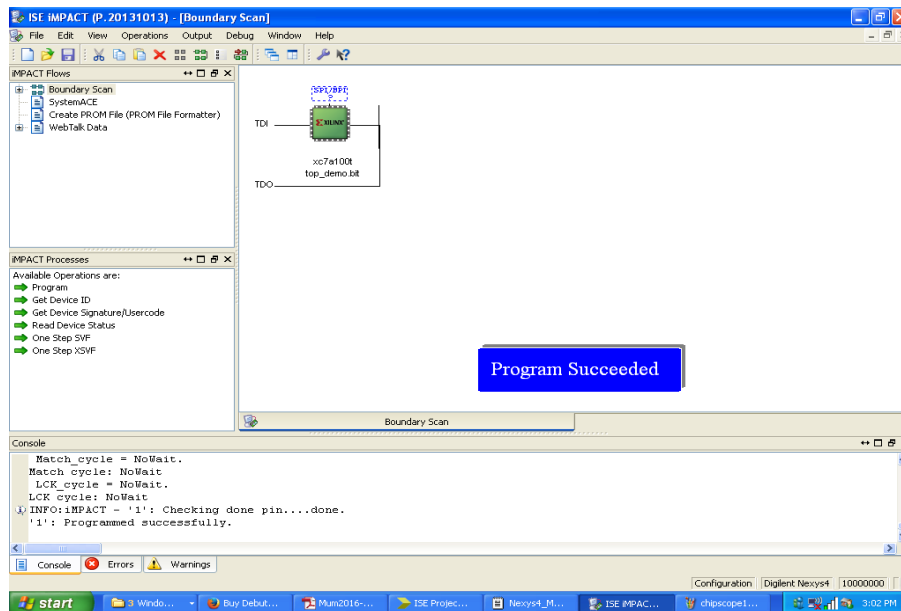


Fig. 11: Verilog code implemented on FPGA board

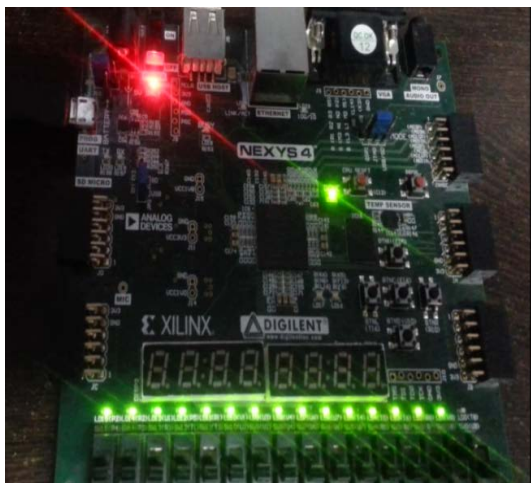


Fig. 12: Output on FPGA board

LUTs and IOBs are given in table. It shows that hardly 12% of the overall device is used for the proposed design. The timing report of the entire design is shown in Table 2. The speed grade of the selected FPGA is-3. The maximum operating frequency is 57.584 MHz. The clock to output delay is 35.678 nsec. In the proposed design there is no combinational path delay.

CONCLUSION

In digital communication system OFDM is an well organised multi carrier modulation method. In this study,

we have presented modified OFDM system for PAPR minimization. In most of the PAPR reduction techniques there is a small amount of reduction in PAPR but the computational complexity has increased along with the BER. One of the major contributions during the development of the proposed system is the integration of FCT, order bit selector and trellis structure. The overall performance of the system is improved and we can say that the presented system is efficient and better in terms of reduction of PAPR and improvement in BER. The proposed system is simulated, synthesized and implanted on to FPGA board. The hardware implementation of the design shows the complexity of the proposed design. It can be seen that the overall device utilization is almost 12.5% of the selected FPGA. It can be observed in the result that it has been able to achieve the reduction in the PAPR with reduced hardware complexity.

REFERENCES

- Abdullah, M.I., M.Z. Mahmud, M.S. Hossain and M.N. Islam, 2011. Comparative study of PAPR reduction techniques in OFDM. ARPN. J. Syst. Software, 1: 263-264.
- Bauml, R., R.F.H. Fischer and J.B. Huber, 1996. Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping. Elect. Lett., 32: 2056-2057.

- Chen, J.C., M.H. Chiu, Y.S. Yang and C.P. Li, 2011. A suboptimal tone reservation algorithm based on cross-entropy method for PAPR reduction in OFDM systems. *IEEE. Trans. Broadcast.*, 57: 752-756.
- Cimini, L.J. and N.R. Sollenberger, 2000. Peak-to-average power ratio reduction of an OFDM signal using partial transmit sequences. *IEEE Commun. Lett.*, 4: 86-88.
- Farahani, M.A. and M. Eshghi, 2007. Implementing a new architecture of wavelet packet transform on FPGA transform with tree structure. *Proceedings of the 8th WSEAS Conference on Acoustics and Music: Theory and Applications*, June 18-20, 2007, Vancouver, Canada, pp: 37-41.
- Gayathri, K.M., S. Bhargavi and N. Thangadurai, 2017. Fragmentary chronological transmission with order bit select to reduce PAPR in OFDM system. *Intl. J. Control Theor. Appl.*, 10: 309-325.
- Gayathri, K.M., S. Bhargavi and N. Thangadurai, 2017. Modified PAPR reduction scheme by integrating DCT and WPT with partial transmit sequence. *Indian J. Sci. Technol.*, 10: 1-6.
- Han, S.H. and J.H. Lee, 2005. An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. *IEEE Wireless Commun.*, 12: 56-65.
- Kaur, S. and R. Mehra, 2012. FPGA implementation of of DM transceiver using FFT algorithm. *Intl. J. Eng. Sci. Technol.*, 4: 1532-1537.
- Krongold, B.S. and D.L. Jones, 2003. PAR reduction in OFDM via active constellation extension. *IEEE. Trans. Broadcast.*, 49: 258-268.
- Lakkannavar, A.D.M., 2012. Design and implementation of OFDM using VHDL and FPGA. *Intl. J. Eng. Adv. Technol.*, 1: 2249-8958.
- Muller, S.H. and J.B. Huber, 1997. OFDM with reduced peak to average power ratio by optimum combination of partial transmit sequences. *Electron. Lett.*, 33: 368-369.
- Reddy, D.D. and S.K. Reddy, 2013. FPGA implementation of QAM transmitter and receiver. *Intl. J. Eng. Res. Appl.*, 3: 48-51.
- Tellado, J., 1999. Peak to average power ratio reduction for multicarrier modulation. Ph.D Thesis, University of Stanford, Stanford, California.
- Wahab, A.A.A. and M.F. Ain, 2010. Clipping and time domain statistical method for PAPR reduction in OFDM system. *Proceedings of the 9th International Conference on Environment and Electrical Engineering (EEEIC)*, May 16-19, 2010, IEEE, Prague, Czech Republic, ISBN:978-1-4244-5370-2, pp: 487-490.

An Enhanced Automatic Room Temperature Control System

Adamu Murtala Zungeru, Mmoloki Mangwala, Joseph Chuma and Baboloki Gaebolae
Department of Electrical, Computer and Telecommunication Engineering,
Botswana International University of Science and Technology,
Private Bag 16, Palapye, Botswana

Abstract: This study presents the design and implementation of an automatic room heater control system. An automatic room heater control system is a self-regulating temperature system which uses a set point or value to maintain the temperature of a room. This system allows the user to set a desired temperature which is then compared to the room temperature measured by a temperature sensor and with the help of a microcontroller, the system responds by turning on any of the two loads (cooler or a heater) automatically depending on the temperature difference. The cooler is triggered on when the room temperature is higher than the set temperature and the heater is triggered on when the room temperature is lower than the set temperature. The system was designed and simulated using Proteus 8, a circuit building software used for building electronics system. For coding the PIC microcontroller, micro-c compiler was used. A 5 V DC power supply was designed in order to provide a biasing voltage to most of the active devices used in the system design circuit. The DC power supply was designed and simulated using Multisim software. For testing and simulation of the circuit design in Proteus, the micro-c hex file was loaded on the Proteus schematic design. The system was simulated and working according to the design specifications.

Key words: Design, room heater, control system, multisim, proteus, simulation, cooler, temperature sensor, heater, microcontroller

INTRODUCTION

With the advancement of technology, automation has become part of our lives. The home is usually the most occupied place in any culture. Areas in the home that are usually occupied by people such as the living room and bedrooms need to be maintained within habitable temperature ranges. The human body has an optimum temperature of about 27°C. Temperatures that are higher or lower than the set temperature of the home body can result in damage to some body organs or tissues and eventual death. These issues become more pertinent in areas of the home that are occupied by infants. Adults could possibly find their way around “thermal discomforts” but infants may not. Other areas of the home that are used as storage areas for perishable food items also need to be thermally regulated in order to prevent accelerated decay of such items. This makes necessary the need for a temperature control system within the home.

The idea of programmed room heater control systems goes back in the eighteenth century and this thought was

first secured in Norman School, Oklahoma by an educator named Warren S. Johnson. Before that time, Janitors were compelled to go in every classroom to check the temperature of the classes and after that control the dampers in the S-basement in like manner. Johnson looked for an approach to end or possibly limit the classroom intrusions of the janitors and increment the solace level of the understudies. The automatic temperature control system was to meet this very need which prompt Warren S. Johnson stopping instructing and beginning his electric administration organization which was gone for outlining programmed control systems. Warren S. Johnson initially built up the pneumatic temperature control framework which took into account temperature control on a room by room premise in structures and homes. By the mid 20th century the automatic temperature control system creation ended up noticeably famous in enterprises and homes. As of late, a considerable measure of work is being finished by organizations in this field. A great deal of automatic room heater system business items are promptly accessible in the market and this includes devices such as AIRCONS (Johnson *et al.*, 2006).

Weather is forever varying and changes on short intervals and as a result, the external conditions always have an influence changes on the indoor conditions. The temperature control systems that are currently in use have limitations. One of these limitations is that the user has to adjust the system every time the external conditions change. This is very tiring and proves out not to be an effective way of controlling temperature of a room. Also, disabled people get to face a lot challenges when they want to operate temperature control system in their houses because this systems require them to use physical contact or some hand remote devices to operate them. To reduce the need to do this, a system that works automatically needs to be put in place.

This study presents an automatic room heater control system. This is an air-conditioning system which monitors the room temperature and controls the circulation of fresh air inside the room without human intervention. This design uses a microcontroller and a temperature sensor to monitor and control the temperature of a room. At first the user will have to set the system temperature to a desired value that he or she wants to maintain in that room. The temperature sensor will then sense measured surrounding temperature and communicates with the microcontroller. The micro controller reads the temperature every 10 sec and compares it with the desired value. If the measured value is less than the desired value, then the heater will be automatically be triggered on to warm up the temperature of the room until it returns back to the desired value and turns off. If the measured value is greater than the desired value, the cooler/fan will be turned on to cool the room temperature back to the normal set point and turns off once it is at that set point (Anonymous, 2017). By Bell (1992), self-programmable thermostat is presented. In a related work, researchers by Tate and Ries (1990), Poll (2006), Agarwal (2006), Nagata (1996), Fiedler and Landy (1959), Hedges (1947), Chengxiang *et al.* (2011) and Fu *et al.* (2010) highlighted on the benefits of using temperature adjustable and fan temperature control systems.

MATERIALS AND METHODS

The automatic room heater control system comprises of three main subsystem: power supply unit, the sensor unit and the control/switching unit as shown in the system block diagram in Fig. 1.

Power supply: The power supply system supply a 5 V DC and 12 VDC power supply to the other units which is

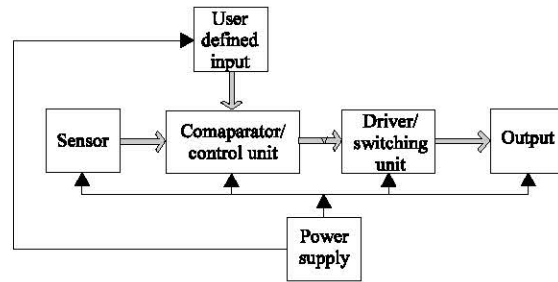


Fig. 1: Block diagram showing the major parts of the system

stepped down with a transformer (240-12 V). This power supply include a full wave rectifier (4 of 1N4001) which converts AC-DC power and a filter (capacitor) to smoothen the output from the rectifier. Voltage regulators were also included in the power supply in order supply regulated DC voltages to the other main units (LM7805 (5 V) and LM7812 (12 V)).

Sensing unit: This study of the system include a temperature sensor (LM 35). A temperature sensor is a device that is temperature sensitive and it responds to changes in temperature.

Control/switching unit: The control/switching unit houses the microcontroller which will receive temperature status from the sensor unit and ensures that it does not compromise the set value by initiating the correct sequence of action. This unit consist of microcontroller, two transistors and two relays to switch on and off a fan and or a heater. At first the user is prompted to input reference temperature that he or she wants to maintain in their room. The temperature sensor will then measure surrounding temperature and communicates with the microcontroller. The micro controller reads the temperature every 10 sec and compares it with the desired value. If the measured value is less than the desired value then the heater will be automatically be triggered on to warm up the temperature of the room until it returns back to the desired value and turns off. If the measured value is greater than the desired value, the cooler/fan will be turned on to cool the room temperature back to the normal set point and turns off once it's at that set point. The measured room temperature from the temperature sensor is analog in nature. The microcontroller has an in built Analog-to-Digital (A/D) converter which convert the analog signal into digital signal because the microcontroller is a digital device, it can only work with binary numbers.

Power supply: The power supply was designed considering the available resources while meeting the design specifications. Most of the components operates on 5 V DC while relays operating at 12 V were used, hence, the need to step down the normal power supply voltage from mains (Approx. 240 V AC) to a reasonably voltage that will have to be rectified (convert to DC) and further filter to remove unwanted pulsation. The 240V AC power was stepped down to 12 V AC (12 V RMS value wherein the peak value is around 17 V) as can be seen from the calculation that follows and the 17 V was further regulated using a voltage regulator (LM7805) to 5 V and (LM7812) to 12 V. A transformer of turn ratio of 20:1 was used for the purpose of stepping down the voltage and rectifier diodes (IN4001) were also used for rectification. Using the turn ratio:

$$\frac{N_p}{N_s} = \frac{V_p}{V_s}, \text{ i.e., } \frac{20}{1} = \frac{240}{V_s}$$

$$V_{RMS} = V_s = \frac{240V}{20} = 12 V \quad (1)$$

$$V_p = 12 \times \sqrt{2} = 16.9705 \approx 17 V \quad (2)$$

Assuming a ripple voltage of 20%:

$$dv = \frac{20}{100} \times 17 = 3.4 V \quad (3)$$

$$df = \frac{1}{2f} = \frac{1}{100} = 0.01$$

$$C_1 \frac{1 \times 0.01}{3.4} = 2.94 \times 10^{-3} F \quad (4)$$

A preferred value of 3300 μF was however employed for the filtering of the assumed ripples as the value is higher than the calculated value, hence will filter much more than expected. Figure 2 shows the designed power supply circuit.

The block diagram showing the major parts of the system circuit is as shown in Fig. 1. Figure 2 is the complete power supply circuit showing all the components and the results gotten from simulation. Figure 3 gave the logical operation of the automatic room heater control system (flow chart) (Fig. 4). Figure 1 consist of 6 different block each housing several components: transmitter and receiver subsystem. The sensor block consists of a temperature sensor (LM35), the user defined

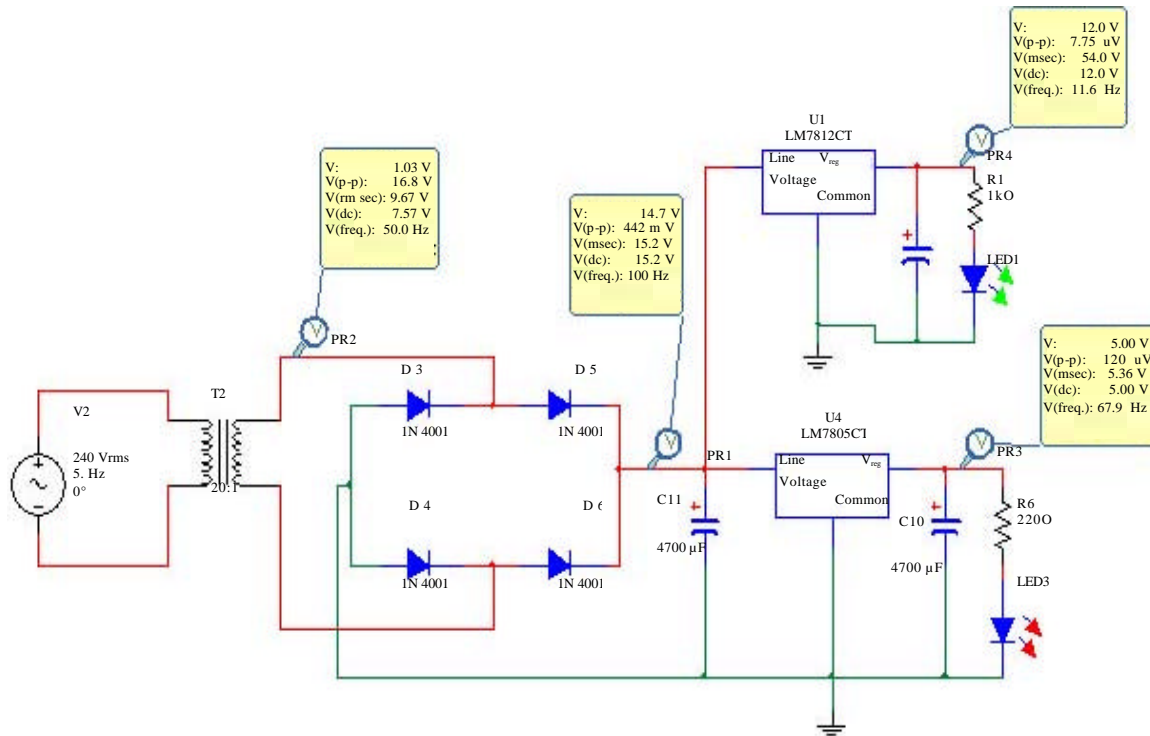


Fig. 2: Power supply circuit

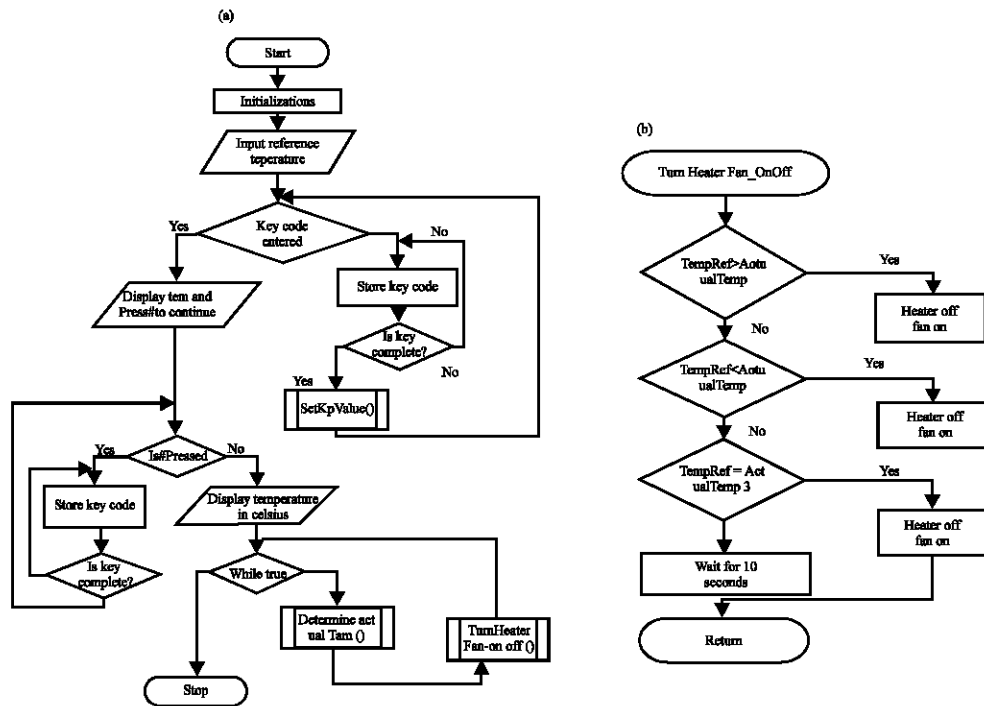


Fig. 3: System flow chart

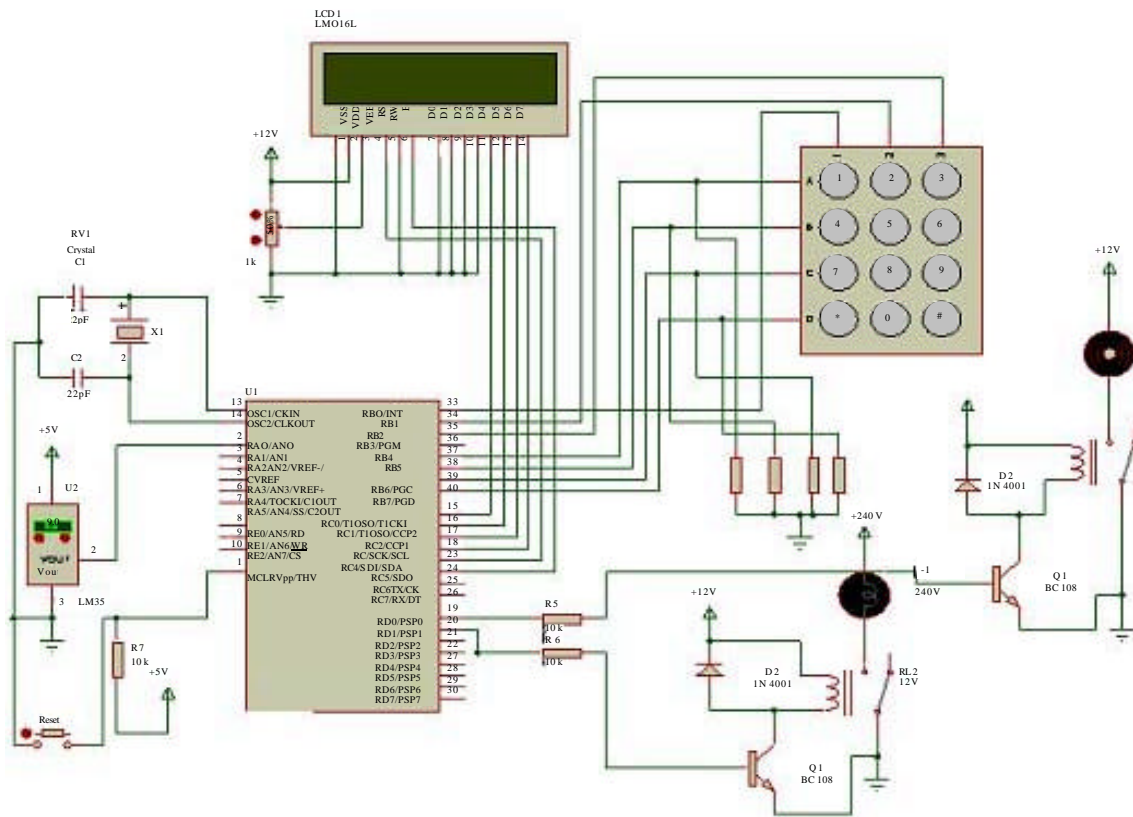


Fig. 4: Circuit diagram of automatic room heater control showing all components of the system

input consists of a keypad, the comparator/control unit is basically the heart of the system that consists of the microcontroller. Generally, the system circuit comprises of the PIC16F877A micro-controller, LM35 temperature sensor, LCD display, crystal oscillator, 4 by 3 keypad for display, 2 transistors for switching purpose, 2 relays also used to support the transistor in the switching effect, a bulb modelled as a heater and a Dc fan. The microcontroller is clocked by the crystal oscillator as it does not have an internal clock. Connected to the microcontroller is a temperature sensor LM 35 which measure the room temperature and give the value reading to the microcontroller for reading. The 2 loads of the microcontroller switched on and off by the relays. The relays are not directly connected to the microcontroller but rather transistors as switches are place in between the microcontroller and the relay to prevent the relay from damaging the microcontroller. The resistors connected in every component of the system are used to limit the amount of current passing to that particular component. The LCD is connected to the microcontroller for displaying the data feed into the microcontroller. The brightness of LCD is controlled by the variable resistor as seen in Fig. 4.

RESULTS AND DISCUSSION

Automatic room heater control system is designed, simulated and analysed in this research. From Fig. 2, it is seen that the calculated results agreed with the simulation results. From PR2 in Fig. 2, it is seen that though the value not exactly equal to the calculated result but approximately equal to the value. If we then compare the peak voltage of the simulation result, 16.8 V, the value is approximately equal to the calculated value of 17 V as can be seen in Eq. 2. As can be seen from Fig. 2, U1 and U4 gave +12 V and +5 V, respectively when deployed voltage regulators (LM7812 and LM7805).

Figure 5 shows the result of a user prompted to enter reference temperature. As can be seen, both RL1 and RL2 are disconnected from Lamp (L1) and the fan motor as loads. In Fig. 6, user entered 12 as the reference temperature which is higher than the room temperature 9.27°C as can be seen from temperature sensor (LM35). But in this case, the microcontroller had not sent any signal to both loads, since, the user has not press the hash key to enter the value 12. For Fig. 7-9, the room temperature measured by the TEMP sensor is 9.27°C and

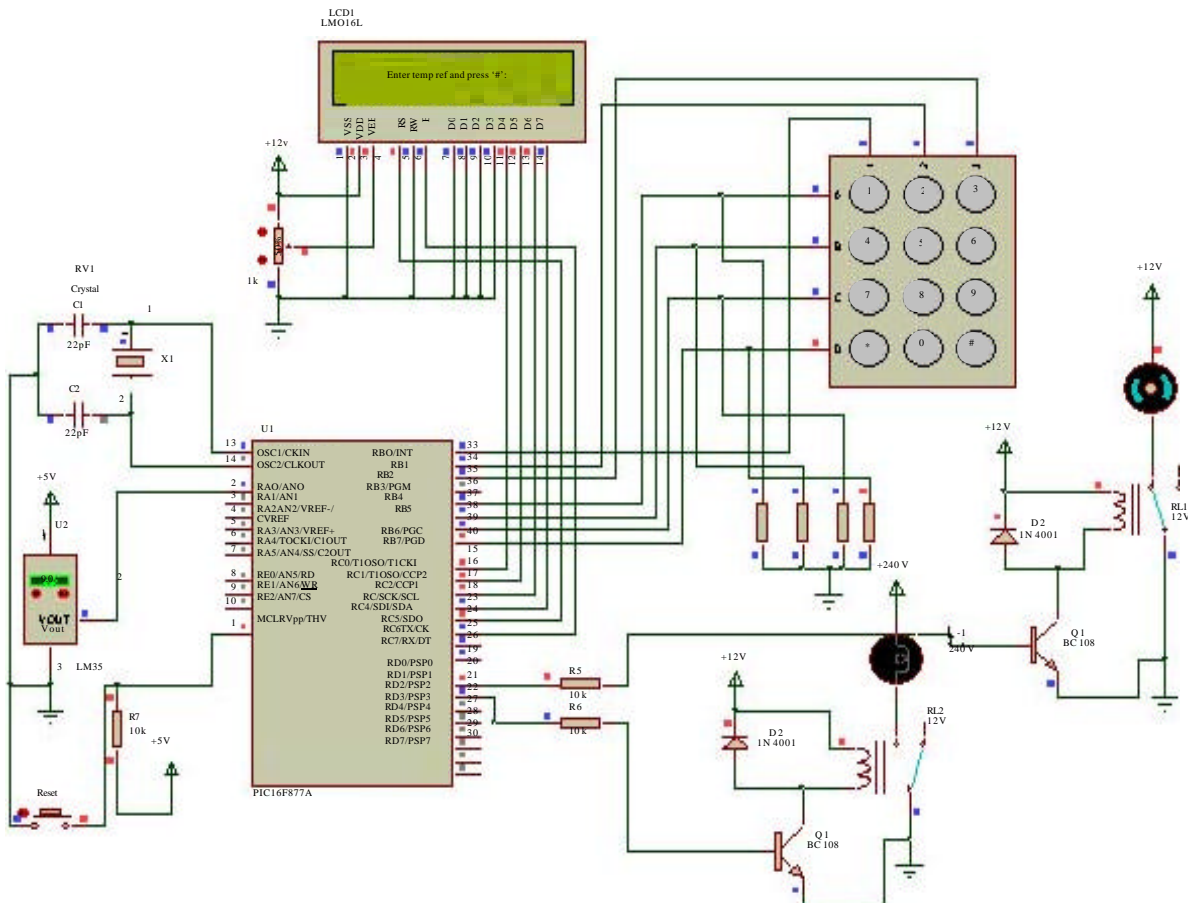


Fig. 5: Result of user prompted to enter a reference temperature

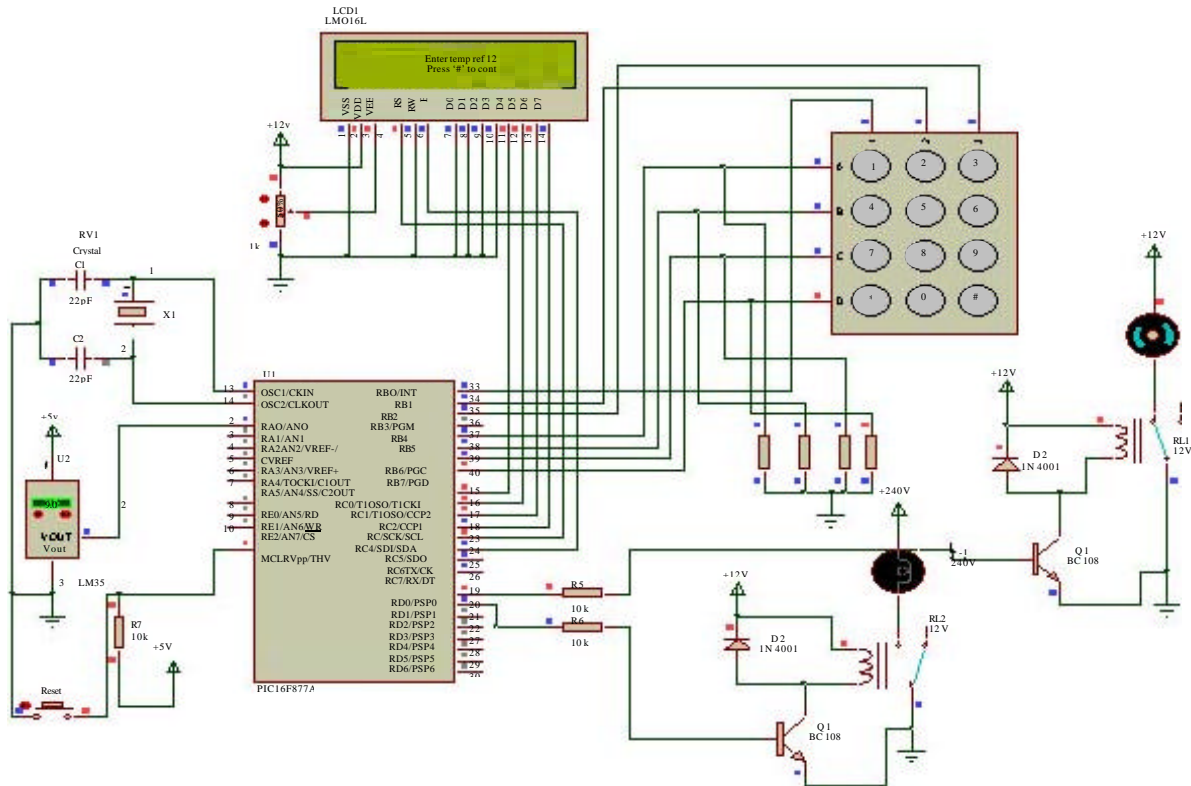


Fig. 6: Result of user entered 12 as the reference temperature

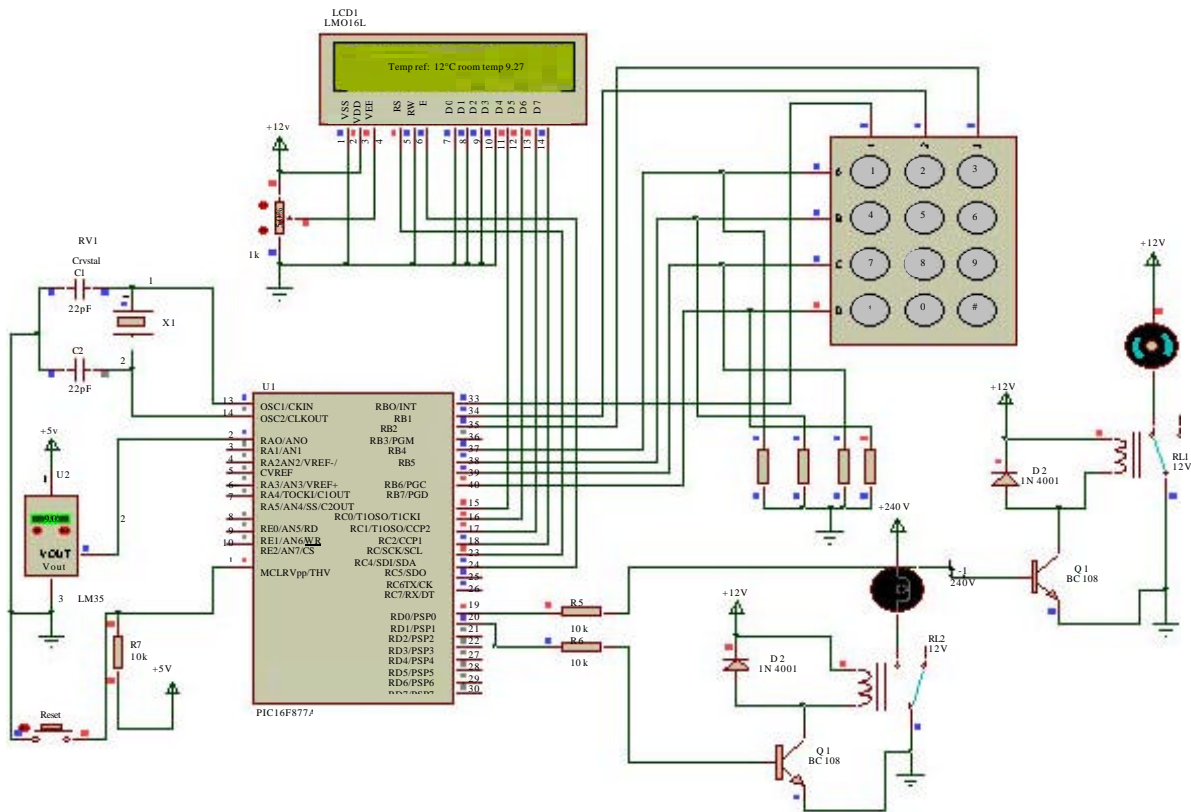


Fig. 7: Result of user entered 12°C as the reference temperature and Heater (L1) switched ON

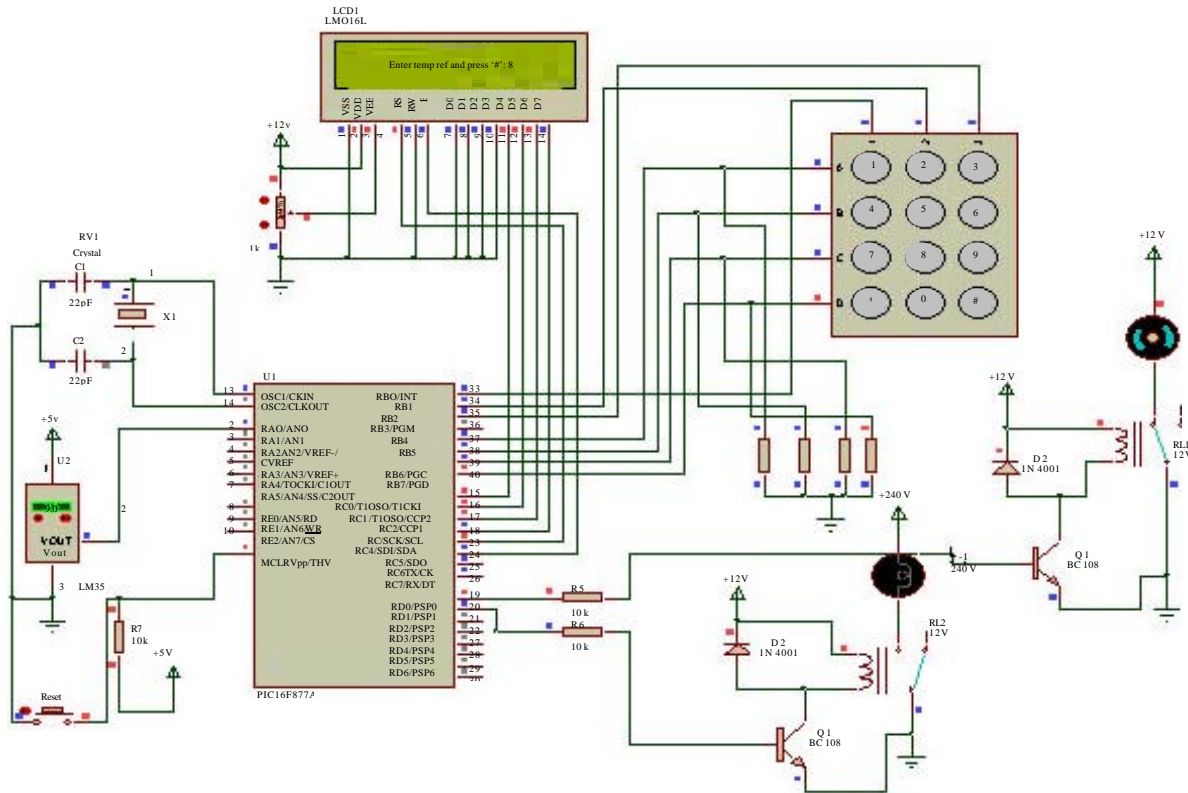


Fig. 8: Result of user entered 8°C as the REF temp, then entered # on the keyboard to continue

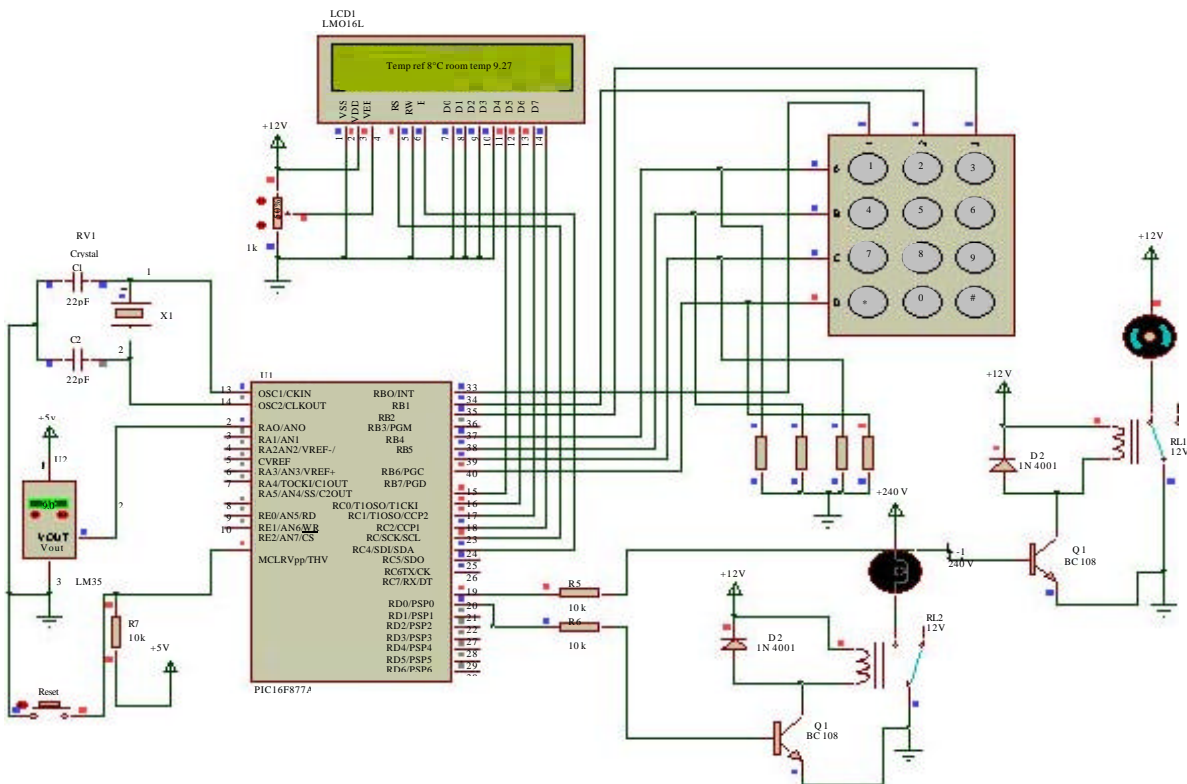


Fig. 9: Result of User entered 8°C as the reference temperature and Heater (L1) switched ON

the REF temp is 12°C. The microcontroller compared the two temperatures and switch the heater ON, since, the REF temp was higher than the room temperature this is when the user has pressed the hash key. As seen from Fig. 7, the bulb is ON as an indication that the heater has been triggered on.

Figure 8 shows the result when the user has entered 8°C reference temperature but the hash key has not been entered which means the microcontroller has not been instructed to compare results.

Also from Fig. 9, when 8°C and was entered as the REF temp and hash key pressed and the Room measured to be 9.27°C, the microcontroller again compares the 2 temperatures values and turn on the fan because the REF temp is lower than the room temp.

CONCLUSION

In this study, we present the design, simulation and analysis of an automatic room heater control system. The system uses PIC 16F877A microcontroller for the control unit and LM35 as the temperature sensor. The output was varied by setting the temperature at various levels and it was discovered that the bulb turn on and off when the system temperature either exceed or less than the predetermined temperature. The system is exceptionally helpful for people who are disabled. This system can be utilized as a part of industry and in addition in home. The system was designed using Proteus and Multisim Software. The system was simulated and working according to the design specifications.

REFERENCES

Agarwal, T., 2006. Arduino: Basics and design. ELPROCUS, USA.
Anonymous, 2017. Service champions heating and air conditioning. Service Champion, Pleasanton, California.

Bell, I., 1992. Self-programmable temperature control system for a heating and cooling system. U.S. Patent and Trademark Office, Washington, USA.
Chengxiang, L., Y. Zhenhua, W. Xu and L. Feng, 2011. Design of automatic temperature control system on laser diode of erbium-doped fiber source. Proceedings of the International Conference on Intelligent Computation Technology and Automation (ICICTA) Vol. 1, March 28-29, 2011, IEEE, Shenzhen, Guangdong, China, ISBN:978-1-61284-289-9, pp: 404-407.
Fiedler, G. and J. Landy, 1959. Multi-loop automatic temperature control system design for fluid dynamics facility having several long transport delays. IRE. Trans. Autom. Control, 4: 81-96.
Fu, T., X. Wang and G. Yang, 2010. Design of automatic-temperature-control circuit module in tunnel microwave heating system. Proceedings of the International Conference on Computational and Information Sciences (ICCIS), December 17-19, 2010, IEEE, Chengdu, China, ISBN:978-1-4244-8814-8, pp: 1216-1219.
Hedges, R.E., 1947. Automatic temperature control for transport airplanes. Trans. Am. Inst. Electr. Eng., 66: 1197-1202.
Johnson, J.E., P.F. Maccarini, D. Neuman and P.R. Stauffer, 2006. Automatic temperature controller for multielement array hyperthermia systems. IEEE. Trans. Biomed. Eng., 53: 1006-1015.
Nagata, M., 1996. Automatic temperature control system. U.S. Patent and Trademark Office, Washington, USA.
Poll, S.B., 2006. Automatic heater controller. US Patent and Trademark Office, Washington, USA.
Tate, B.G. and R.P. Ries, 1990. Wireless thermostat and room environment control system. US Patent and Trademark Office, Washington, USA.

Emotion based Audio Player Using Mind Wave EEG Sensor

G. Muneeswari

Department of Computer Science and Engineering,
Vel Tech Multi Tech Dr. Rangarajan and Dr. Sakunthala Engineering College, Chennai, India

Abstract: In this study, an attempt is made to develop an android based application that can play songs automatically, based on our emotional state. The application is connected via. Bluetooth to the mind wave Electroencephalograph (EEG) sensor which helps to determine the mood of a person based on the brainwave readings. Human brainwave consists of alpha, beta, gamma and theta waves. The EEG sensor can detect two states of mind active and sleepy. It determines the user's mood as active when the brain wave has excess beta waves and sleepy when the brain wave has excess alpha waves. Using this feature, it is possible to play a suitable song that is congruous to the user's frame of mind. The results obtained were analyzed using Naive Bayes classifier and this low cost implementation showed an overall accuracy of 97.5%.

Key words: Electroencephalograph (EEG), sensor, brainwave detection, alpha wave, beta wave, frame of mind

INTRODUCTION

Finding the right song to suit the human brain mood can be exasperating, especially if the playlist is too big. To avoid this unwanted frustration, it is possible to automatically select and play the favorite songs from the stored playlist with ease, based on the mood, instead of having to search for the songs manually. The smart headphone or audio player reads the user's brainwaves and picks the right tune for the user automatically. The smart headphone uses an Electroencephalograph (EEG) sensor to interpret our emotional state. The EEG sensor is synchronized with the application (Think2Play) that is created and installed in the user's android mobile.

In case of the human mood changes and if it wants a different song, the user has to blink rapidly. Doing, so, allows the EEG sensor to pick up a new set of signals and repeat the automatic song selection process again. Thus, information in the headphone is refreshed and the sensor scans the brain again, so that, the application can select a new song. By using the concept of signal processing along with the NeuroSky EEG sensor, numerous simple yet, effective smart Android applications can be developed. The sequence of actions performed by the application is as follows: the application is connected to the wireless NeuroSky smart headphones to detect any change in the brain waves. If a valid change is observed, the appropriate activity is performed based on the information received, the application can be programmed to behave in a certain manner.

Literature review: By Rebsamen *et al.* (2010) EEG based wheelchair application has been discussed to navigate in familiar environments. Motor imagery and direct brain-computer communication with an effective sensor

based application is implemented by Pfurtscheller and Neuper (2001). Silva (2014) work mainly focusses on Bluetooth communication and with the help of that a brain wave controlled robot has been proposed. Millan *et al.* (2002) found a local neural classifier for the recognition of EEG patterns associated to mental tasks. This is best suitable for the classification of mental illness.

Millan *et al.* (2010) discussed the various challenges related to combining brain computer interfaces and assistive technologies state-of-the-art. P300-based brain computer interface and a prototype of a chinese speller has been implemented by Su *et al.* (2008). This application uses hardware oriented brainwave simulation. A conceptual study of alpha and beta EEG brainwave signal classification technique proposed by Zainuddin *et al.* (2014) formed the basis for the development of the proposed audio player. A survey on EEG based brain-controlled mobile robots was elaborated in the research by Bi *et al.* (2013). Kale *et al.* (2014) has proved that EEG based signals are very effective in robot navigation control and future realistic application. Solanki (2015) implemented a sensor based brain wave controlled robot which could perform many obstacle detection operation. Alpha, beta wave etc., classification is elaborately discussed by Jeffrey Fannin.

Human mind regulates every-day activities by means of electric waves which are registered in the brain, emitting tiny electrochemical impulses of varied frequencies which can be registered by an Electroencephalogram (EEG sensor). These brainwave patterns are categorized into four types as shown in Table 1.

The waves are categorised based on the frequencies, as gamma, delta, theta, alpha and beta waves. The gamma waves are used to indicate if a person is alive. Thus, they

Table 1: Categories of brainwaves

Brain waves	Frequency range (Hz)	Mood
Delta wave	0.5-4	Deep relaxation or sleep
Theta wave	4-7	Light meditation
Alpha wave	7-13	Deep meditation
Beta wave	13-40	Attention or active state

will not be considered for this application. From Table 1, we can make out exactly what state of mind a person is said to be in based on the brainwaves detected. As depicted, delta waves indicate a person being sad or in deep sleep. Theta and alpha waves, both signify a state of meditation where theta waves imply light meditation and alpha waves indicate deep meditation. Finally, beta waves denote an active or attention state. Alpha and beta waves are the most commonly observed waves in most cases. Hence, these waves are often used in studies and application (Pfurtscheller and Neuper, 2001). The proposed application also uses the alpha and beta waves to detect either active or relaxed state of mind.

MATERIALS AND METHODS

Most often when a user feels like listening to music, they may not be able to find a suitable song because they may be too indecisive in selecting a song and simply give up altogether. At present, the apps that help the user to choose music are based on the most frequently heard songs or usage patterns. The application that is being used to select songs that are more in-tune with the user’s emotional state. In order to avoid this unnecessary hassle and make human life simpler this idea of implementing the smart headphone using the Think2Play android application has been proposed. To implement this solution it is important to understand the composition and classification of the brainwaves. Human brain is made up of billions of brain cells called neurons which use electricity to communicate with each other. The combination of millions of neurons sending signals at once produces an enormous amount of electrical activity in the brain. Using sensitive equipment (such as an EEG sensor) it is easy to measure the electric levels over areas of the scalp. The combination of the various electrical activities of the brain is commonly called a brainwave pattern or brainwave because of its cyclic, ‘wave-like’ nature.

Now, these brainwaves are relatively hard to detect, as the signals are very weak. Thus, a highly sensitive device must be used. We make use of the MindWave mobile headset which is designed and developed at NeuroSky technologies. NeuroSky-enabled solutions deliver unique insights into body and mind that can motivate people to constantly innovate and develop new applications. The company’s proprietary biosensor technologies provide foundation for analyzing

Table 2: Comparison of sensors

Metric	Min.	Max.	Mean	SD	Skewness	Shapiro-wilk normality test	
						W	p-values
NeuroSky MindWave							
Attention	0.05	74.79	19.05	12.95	1.2865	0.8780	0.1131
Meditation	0.31	49.79	25.35	17.75	-0.0150	0.9435	0.1405
Emotiv EPOC							
Attention	50.00	79.79	65.20	8.57	0.0028	0.8836	0.1878
Meditation	20.21	89.86	55.92	20.78	-0.0535	0.9394	0.1261

biometric data, using Electroencephalogram (EEG) and Electrocardiogram (ECG) in a way that’s never been practical before.

From the family of NeuroSky, the MindWave headsets are designed to be used by developers to easily and quickly work with complete EEG-monitoring products. The MindWave mobile headset turns your android phone into a brain activity monitor, using the Think2Play App. The headset safely measures brainwave signals and monitors the attention levels of individuals as they interact with a variety of different apps. This headset is useful for OEMs and developers building apps for health and wellness, education and entertainment. The comparison of NeuroSky MindWave and Emotiv EPOC sensors are displayed in Table 2.

The MindWave family consists of MindWave and MindWave mobile headsets. The MindWave is designed for PCs and Mac while the MindWave mobile is compatible with PCs, Mac and mobile devices like the iPhone, iPad and Android. If you want a mobile compatible device, check out the MindWave mobile. Both headsets share the following characteristics. The NeuroSky ThinkGear ASIC chip is priced to power mass adoption in health and wellness, educational and entertainment devices, popular EEG technology.

Think2Play application design: Using the smart headphones makes life easy. The workflow diagram is illustrated in Fig. 1. This smart system is made up of two parts: the Think2Play app installed in the user’s android mobile a device, NeuroSky MindWave mobile headset, consisting of the Electro-encephalograph (EEG) sensor that is placed on the user’s head and attached to the user’s headphones. This device scans the user’s brainwaves and interprets their emotional frame of mind.

Initially, once the user has downloaded the Think2Play application, they must tag the songs in their playlist as either being in the happy or sad category. Once this has been done for all the songs, we can begin using the actual application. Secondly, the EEG sensor from the MindWave mobile headset, scans the user’s

brainwaves and detects the greater state (attention/happy or meditation/sad). Finally, based on the detected state of mind, a suitable song is played.

Assumptions: For the smart headphone to work properly some conditions have to be met. The mobile device has to have the specific Think2Play App. installed in order to interpret the brainwaves.

The brain signals are very weak, thus, the EEG sensor has to be highly sensitive to capture the various signal effectively. The database should be regularly updated to include new songs and their signal tags. Since, the

electrodes are delicate we must make sure that the headphone is handled with proper care because damage to the electrodes can cause signal loss or distorted signal storage.

RESULTS AND DISCUSSION

The Think2Play application has been designed to run on an android operating system and developed using Android Studio 3.0. It has been tested to show an accuracy of 97.5% using Naive Bayes classifier. The application is tested on Samsung Galaxy Grand 2 Android mobile.

Figure 2 shows the initial user interface once the application is installed. It begins to pair via. Bluetooth, with the MindWave mobile headset in Fig. 3.

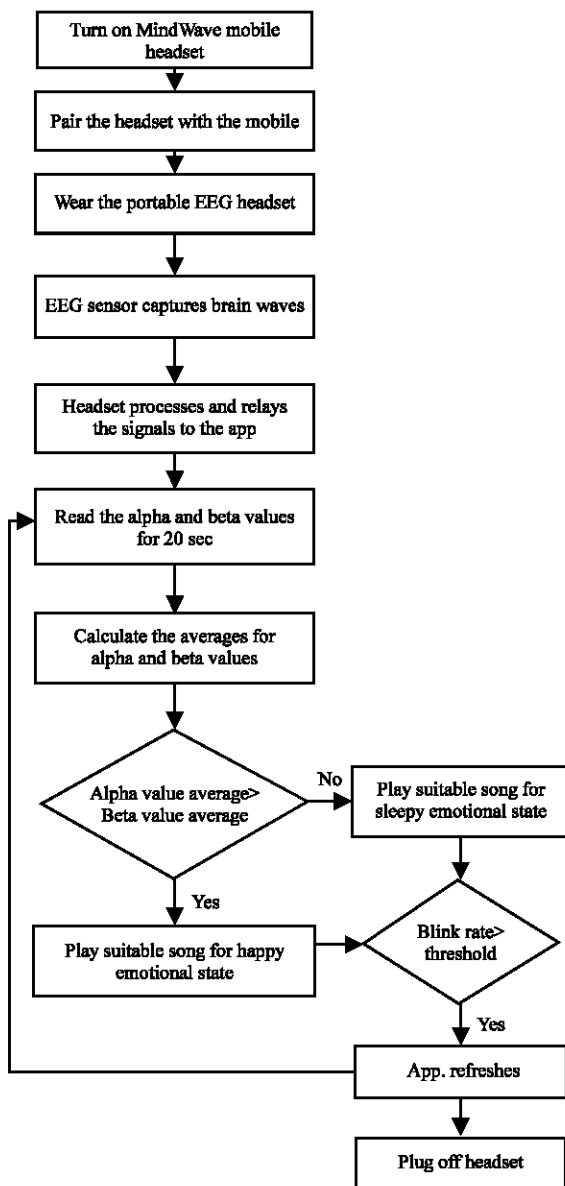


Fig. 1: System architecture-Think2Play

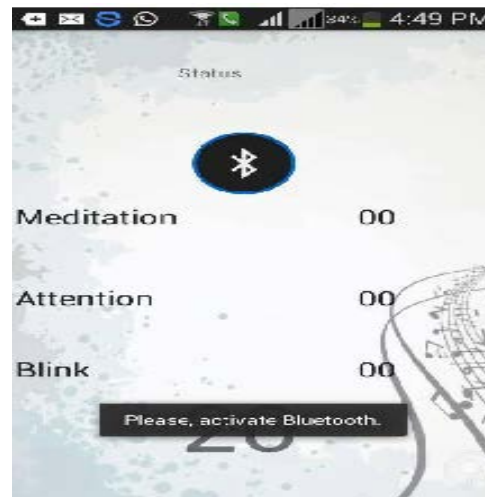


Fig. 2: Initial user interface



Fig. 3: Pairing using Bluetooth



Fig. 4: Timer initialization

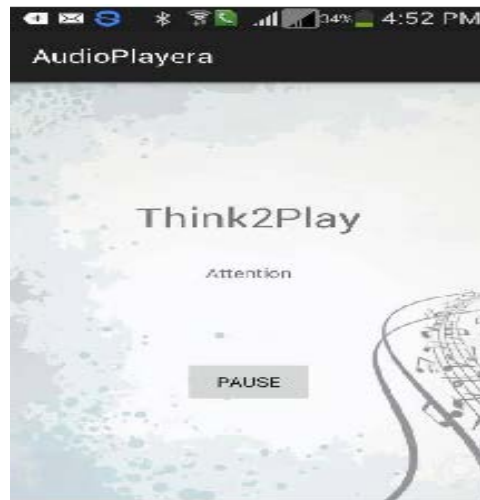


Fig. 6: Attention state

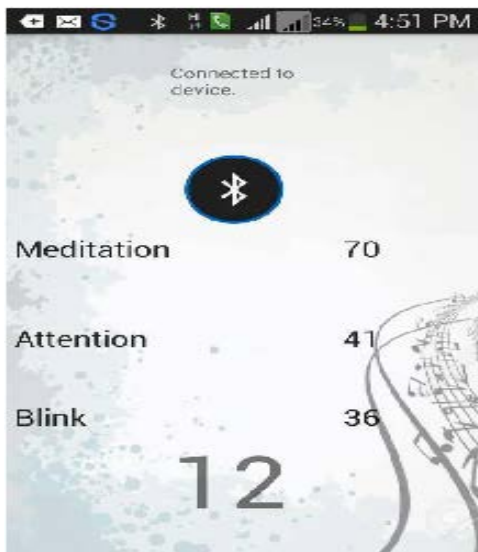


Fig. 5: All state of brain waves

Alpha waves, ranging from 7.5-12 Hz are slower and associated with relaxation and disengagement. Paring is established and the 20 sec timer is started for detecting the brainwaves in Fig. 4 and 5 show all states of initial brain waves.

Thinking of something peaceful with eyes closed should give an increase of alpha activity. Several studies have found a significantly rise in alpha power after smoking.

Finally, according to the monitored brain signals either happy songs for attention state (Fig. 6) or sad songs for meditation state (Fig. 7) are played.

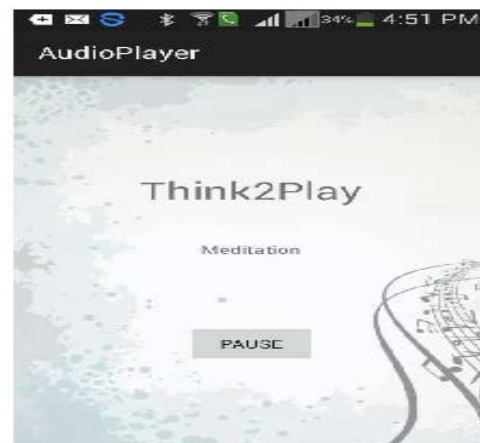


Fig. 7: Meditation state

Beta waves are in the frequency range of 12 and 30 Hz but are often divided into 1 and 2 to get a more specific range. The waves are small and fast, associated with focused concentration and best defined in central and frontal areas. When resisting or suppressing movement, or solving a math task, there is an increase of beta activity as per Naive Bayes classifier, the mean accuracy of 97.5% has been shown in Fig. 8 for attention state and meditation state is shown in Fig. 9.

Theta waves, ranging from 3.5-7.5 Hz are linked to inefficiency, day-dreaming and the very lowest waves of theta represent the fine line between being awake or in a sleep state. Theta arises from emotional stress, especially frustration or disappointment. It has also been associated with access to unconscious material, creative inspiration and deep meditation.

Figure 10 shows that the human brainwave is connected via the EEG sensor and attention, meditation and blink state is detected and after that it will show the attention values and blink values of brain wave sensor.

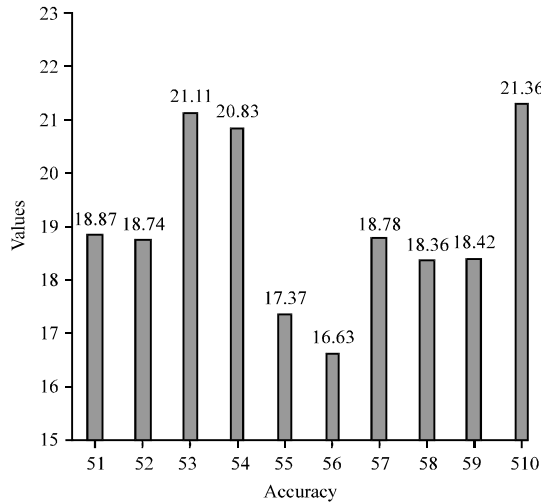


Fig. 8: Mean recognition accuracy of attention state

In Fig. 11 the attention and blink values analysis interms of graph will be generated and there are two signals generated in the graph namely the black color signal is for blinking level and the red color signal is for attention level.

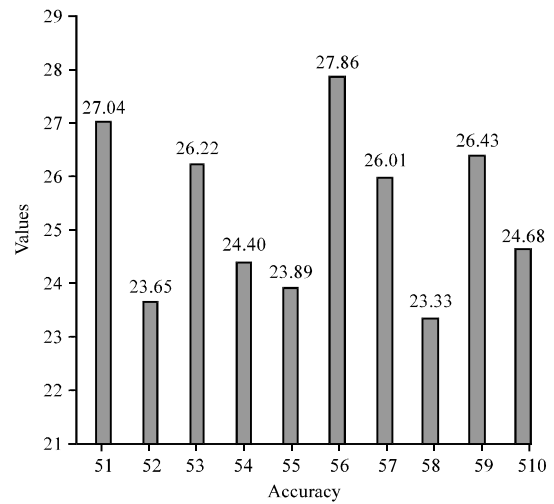


Fig. 9: Mean recognition accuracy of meditation state

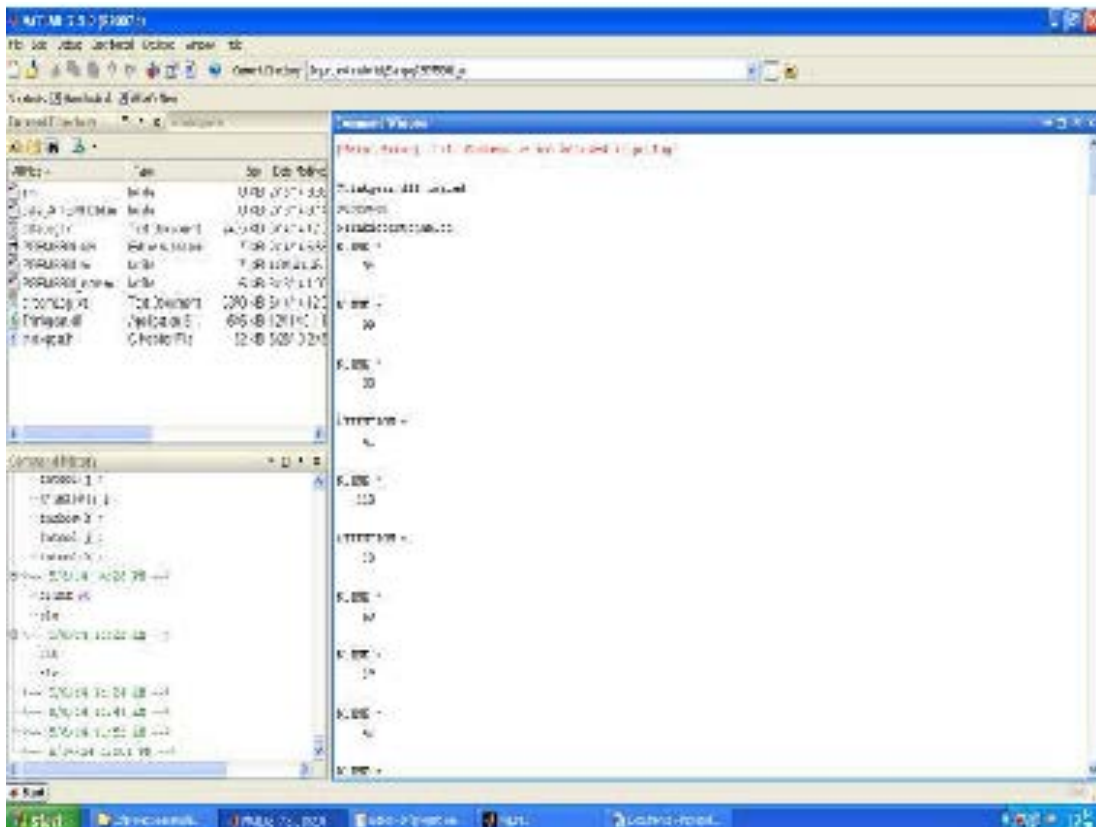


Fig. 10: Performance of EEG sensor with all states simulation

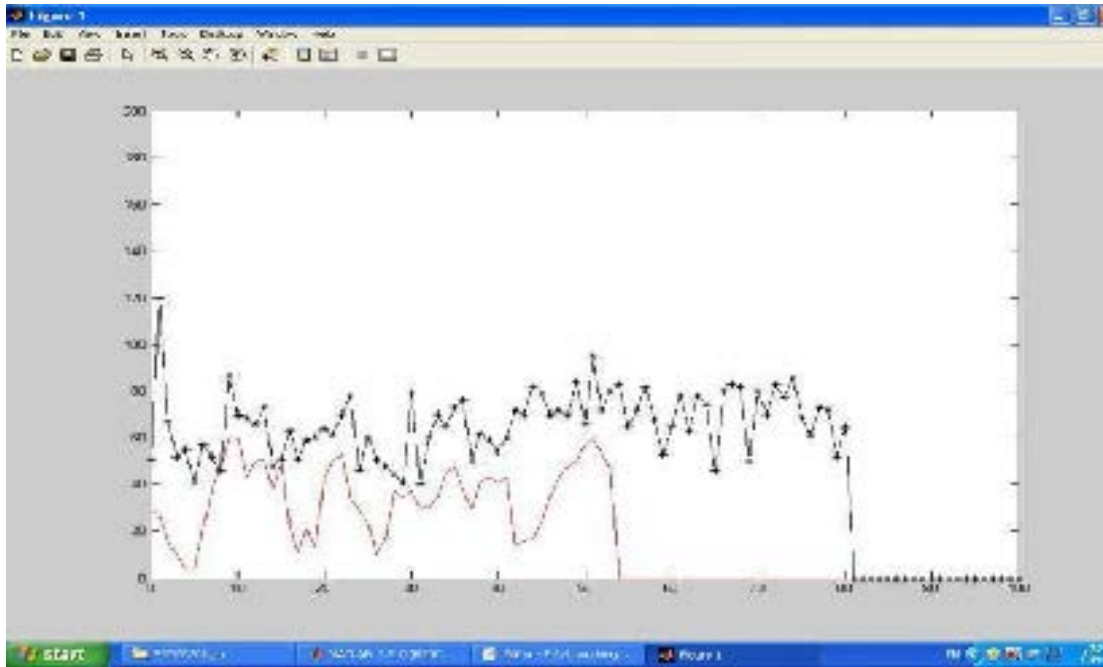


Fig. 11: Analysis of brain wave signals

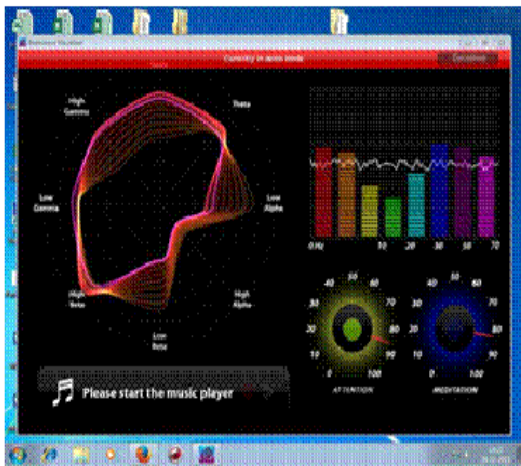


Fig. 12: Visualization of brain wave

The actual visualization of the brain wave signal for the implementation of the audio player indicates the blink, attention level and meditation levels are shown in Fig. 12.

CONCLUSION

The smart headphones developed can be used to make process of selecting a song more fun and easy with a guaranteed degree of accuracy. This small application

can pave the way for more innovative application development that can consider a wider array of emotions apart from just the happy and sad states such as, sleepy, depressed, stressed, etc. Also, the accuracy can also be improved by using different algorithms apart from the standard ones used in the NeuroSky MindWave mobile headset.

REFERENCES

- Bi, L., X.A. Fan and Y. Liu, 2013. EEG-based brain-controlled mobile robots: A survey. *IEEE Trans. Hum. Mach. Syst.*, 43: 161-176.
- Kale S.T., S.P. Mahajan, B.G. Rakshe and B. Nk, 2014. Robot navigation control through EEG based signals. *Intl. J. Eng. Comput. Sci.*, 3: 5109-5113.
- Millan, D.R.J., J. Mourino, M. Franze, F. Cincotti and M. Varsta *et al.*, 2002. A local neural classifier for the recognition of EEG patterns associated to mental tasks. *IEEE Trans. Neural Netw.*, 13: 678-686.
- Millan, J.D.R., R. Rupp, G.R. Muller-Putz, R. Murray-Smith and C. Giugliemma *et al.*, 2010. Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges. *Front. Neurosci.*, 4: 1-15.
- Pfurtscheller, G. and C. Neuper, 2001. Motor imagery and direct brain-computer communication. *Proc. IEEE.*, 89: 1123-1134.

- Rebsamen, B., C. Guan, H. Zhang, C. Wang and C. Teo *et al.*, 2010. A brain controlled wheelchair to navigate in familiar environments. *IEEE. Trans. Neural Syst. Rehabil. Eng.*, 18: 590-598.
- Silva, F.L.D., 2004. Functional localization of brain sources using EEG and/or MEG data: Volume conductor and source models. *Magn. Reson. Imaging*, 22: 1533-1538.
- Solanki, K.H., 2015. Brain wave controlled robot. *Intl. Res. J. Eng. Technol.*, 2: 609-612.
- Su, Y., B. Wu, W. Chen, J. Zhang and J. Jiang *et al.*, 2008. P300-based brain computer interface: Prototype of a Chinese speller. *J. Comput. Inf. Syst.*, 4: 1515-1522.
- Zainuddin, B.S., Z. Hussain and I.S. Isa, 2014. Alpha and beta EEG brainwave signal classification technique: A conceptual study. *Proceedings of the 2014 IEEE 10th International Colloquium on Signal Processing and its Applications (CSPA'14)*, March 7-9, 2014, IEEE, Kuala Lumpur, Malaysia, ISBN:978-1-4799-3090-6, pp: 233-237.

Early Alarm for Emergency Response Based on the Priority Associated with the Cooperative Awareness Messages in Vehicular Ad-hoc Network

Mohammed Saadoun Mohammed and Abdul Kareem A. Najem Alaloosy
College of Computer Science and Information Technology,
University of Al-Anbar, Anbar, Iraq

Abstract: Vehicular Ad hoc Network (VANET) is a promising technology for future smart vehicles systems and an essential component of Intelligent Emergency System (IES). The IES includes a wide range of modern technologies such as Global Positioning System (GPS), digital maps, video cameras, sensing devices and the wireless communication devices. It provides necessary information about the condition of the roads in time for drivers and traffic management systems to improve traffic efficiency, reduce traffic congestion, waiting times and fuel consumptions. Design and implement an IES which automatically controls the encryption of the Cooperative Awareness Messages (CAMs) according to the priority associated with CAMs exchanged between emergency vehicles and Road Side Units (RSUs). The CAMs sent from the emergency vehicles to RSUs be signing using a Secure Hash Algorithm (SHA-2) to distinguish them from normal messages issued from other vehicles. The IES uses the features extracted from the trace file that describes the normal and urgent behavior in the VANETs. The type and the number of features have an important role in increasing the classification accuracy rate and decreasing false alarms, especially False Negative Rate (FNR). In this study, the process of classification of urgent records by using (self-organizing map, feed-forward neural network and Elman neural network). The proposed system is based on a program written by MATLAB R2015a. Our selection used for design and programming the proposed system. These algorithms have already been employed to solve the problem because of its importance in saving time and effort as well as providing high results accuracy in quick time unlike other programming languages. The result is clear in overall system in each technique in SOM accuracy degree 99.5% and FNR 0% while FFNN accuracy degree 99.3% and FNR 0.84211% for number of features 16.

Key words: VANET, routing protocol, intelligent emergency system, secure hash function, RSU, employed

INTRODUCTION

Recent advances in wireless networks have led to introducing a new technology called Vehicular Ad hoc Networks (VANETs) that combines different types of networks such as Ad hoc network, cellular technology and wireless LAN (WLAN). VANET is an extension of Mobile Ad hoc Networks (MANETs) which can enhance road safety, essential emergency alerts, traffic management and infotainment facilities for drivers and passengers with increased efficiency of the transportation systems (Anonymous, 2009). VANETs communications can be provide three techniques; They are Inter-Vehicle Communication (IVC), Inter-Roadside Communication (IRC) and/or communication between Vehicles and fixed Roadside equipment (V2R) communication (Dass *et al.*, 2012). In VANETs, each vehicle acts as a communication node (sender and/or receiver) to exchange the

information either directly between vehicles as single-hop communication or vehicles be capable of retransmit messages in that way enabling multi-hop communication (Karagiannis *et al.*, 2011).

The technology that has been standardized for communication in VANET is DSRC (dedicated short-range communications) that includes wireless technologies like WiFi, IEEE 802.11, WIMAX, IEEE 802.15, Bluetooth, IRA and Zig Bee. In VANET environment, each vehicle has its built in device called On Board Unit (OBU)/radio interface that facilitates them to communicate with other vehicles and RSUs and also enabling short-range wireless ad hoc networks which receives safety messages such as sudden break warning from other vehicles, essential emergency alerts, etc. and non-safety message such as comforts and entertainment related information (Yang *et al.*, 2014). Figure 1 shows vehicles, RSUs and the occurrence of an accident. Once

Fig. 1: An example of the process of responding to cases of emergency on the road

an accident has occurred, CAMs and control data generated and communicated to the RSUs and other vehicles in that zone.

In our research, we designed a new emergency detection system by adding a sign to emergency messages to give priority for urgent case which allow the emergency vehicle to get to the incident scene faster and more safely. The EDS utilizes trace files generated from the Network simulator Version 2 (NS2).

Literature review: Traffic jams and controlling flow of vehicles are one of the most major problems in modern urban area which caused multiple road accidents in the city and loss of life due to the delay in the arrival of the ambulance to the accident site in time (Assila *et al.*, 2012). The above problems can be solved by taking advantages of recent advances in the field of wireless networks. Vehicles and equipment manufacturers have the opportunity of enhancing the surface transportation via. using the communication capabilities of the VANET to support Vehicle-to-Vehicle (V2V) communication and Vehicle-to-Road Side Unit (V2RSU) communication and offer an Intelligent Transportation System (ITS) to the drivers (Hung *et al.*, 2008; Gaur *et al.*, 2013). In real-time applications, emergency vehicles need access to important information such as warning messages and CAMs between vehicles and RSUs. Therefore, security and privacy becomes a very important issue in VANETs developments (Alheeti *et al.*, 2015).

Athavan *et al.* (2012) designed an intelligent transportation system for the ambulance which is controlled automatically the traffic signals based on the accident site and the hospital site according to the information stored for each node in VANETs to cross

ambulance smoothly traffic junctions to reach the hospital in time when an accident on the road. Buchenscheit *et al.* (2009) designed an emergency vehicle warning system that is using wireless communication to warn other vehicles or to preempt traffic lights for the purpose of the arrival of the emergency vehicle to the intended destination in the fastest time and reduce accident risks during emergency reaction trips. Alarm system in emergency vehicles gets detailed information from roadside infrastructure like a traffic light about congestion, speed and ambulance site to take the appropriate action and timing. Kabiri and Aghaei (2011) suggested a new approach of extracting suitable and static features from the trace files which describe the type of vehicles in the network. This approach is based on feature selection method and it applies PCA theory to determine network operating conditions. The different numbers and types of features had a direct impact on the accuracy of the IES. Raut *et al.* (2014) proposed an early alert system of dedicated vehicles based on the great circle algorithm and RSU deployment in vehicular networks to receive early alert messaging and identify the vehicles existing in the dangerous zone via. relative positioning. Discussion a set of security issues and challenges that threaten Vehicular Ad hoc Networks (VANETs) and focus on attacks that get the message itself rather than the car such that attacker dropping packets and it may contain important information for the recipient, attacker may send wrong information in the network. Samara *et al.* (2010) proposed suitable solutions for these challenges and problems to reach a satisfactory level for the driver and manufacturer to achieve safety of life and infotainment. Doijad *et al.* (2015) proposed a new security and safety system based on wireless technology to secure the communication between Vehicle-to-Vehicle (V-to-V) in VANETs. The proposed system is the development of communication module by using ATmega with transceiver and LCD.

MATERIALS AND METHODS

In our study, we propose an IES that is based on a dataset which was collected from a trace file that was generated utilizing NS2 to identify behavior of vehicles, whether normal or urgent case. The steps below explain the methodology:

Mobility model: We used two tools to create a real scenario of normal and/or urgent behavior in VANETs. These are Simulation of Urban Mobility Model (SUMO) and MOBility Vehicles (MOVE). The SUMO is widely known in the field of VANET simulations, it purveys

Fig. 2: Interconnection of MOVE, SUMO and NS2

efficient computation even in various sizes of scenarios and purvey a better way to effectively plan. Also, support operates and design intelligent transportation systems. The move generator is designed on SUMO (CCC., 2011). The output file of these tools represents a mobility file for normal and emergency vehicles in VANETs. These files are used as input to NS2 (Danquah and Altılar, 2014). Below is the logical flow and interconnection between the above three simulators that we based on to simulate vehicles, road network and finally the wireless sensor network as shown in Fig. 2.

In this study, we used the Manhattan urban mobility model to create a mobility and traffic scenario for vehicles. The reasons for selecting this type of mobility is flexibility and easily is selection of direction for vehicles which may be vertical or horizontal direction and is widely used in this research field.

Simulation environmental and parameters: We selected an efficient network simulator NS2 to evaluate and measure the performance of the proposed IES. The NS2 is designed to simulate different networks such as wired and wireless networks. A common problem that is exposed when simulating the VANETs with NS2 is a realistic mobility model and traffic model because the simulator is not designed specifically for VANETs. To overcome this problem, we used the SUMO and MOVE tools to create realistic mobility and traffic model for VANETs. A screenshot of NS-2 utilizing Network the Animator (NAM) trace file is shown in Fig. 3. Figure 3 shows Manhattan mobility model that consists of 9 RSUs and 140 vehicles (136 normal vehicles and 4 emergency vehicles).

One of the important issues in the simulation system is the initial parameters. Some parameters used in

Fig. 3: Screenshot of Simulation in NS2 NAM

Table 1: Simulator environmental and parameters

Parameters	Values
Simulation time	1000 (sec)
Number of nodes	140 vehicle
Number of traffic lights	5 signal
Number of edges	24 street
Number of RSUs	9 RSU
Type of traffic	Constant Bit Rate (CBR)
Topology	552×452 (m)
Transport protocol	UDP
Packet size	1000
Routing protocol	AODV
Channel type	Wireless
Queue length	50 packets
Number of road lanes	2
Radio propagation model	Two ray ground
MAC protocol	IEEE 802.11
Speed	60 (m/sec)
Interface queue type	Priority queue
Network interface type	Physical wireless
Mobility models	Manhattan mobility model

simulating the VANETs are: Constant Bit Rate (CBR) application that sends constant packets through the transport protocol such as (UDP), mobility model (Manhattan) and radio propagation model (two ray ground) in Table 1 (Samara *et al.*, 2010).

Feature sets and extraction a trace file and a network animator represent output files of the NS2. These two files are used for analyze and viewing network simulation of the behavior in VANETs. The trace file generated in NS2 is divided into three parts. These are basic trace, internet protocol trace and AODV (CCC., 2011). The type of event taking place at the node and can be one of the five types: Send (s), Receive (r), Drop (D), Forward (f) and Movement (M) (Zhou and Hass, 1999). The events of VANETs are described using trace file, it contains many different data features which are used for analysis. These features describe normal and urgent behavior in the VANETs. To increase the efficiency and the accuracy of the proposed system, the most effective features are extracted from the trace file. Many methods used in the field of VANETs to extracted effective features. We used the AWK language

Table 2: Features selection

Basic trace	IP trace	AODV trace
Packet ID payload size and type, source and destination MAC and ethernet	IP source and destination	Packet tagged, hop, counts, broadcast ID, destination IP with sequence number source IP with sequence number and priority

Table 3: FFNN parameters

Parameters	Values	Description
TrainParam.epochs	1000	Maximum number of epochs to train
TrainParam.goal	0	Performance goal
TrainParam.max_fail	6	Maximum validation failures
TrainParam.min_grad	1×10^{-6}	Minimum performance gradient
TrainParam.sigma	5×10^{-5}	Determine change in weight for second derivative approximation
TrainParam.lambda	5×10^{-7}	Parameter for regulatng the indefiniteness of the Hessian

Fig. 4: SOM structure

to extract features that capture the events of vehicles (Alheeti *et al.*, 2015). In this study, a new field (the priority) is added to the trace file to give a higher priority to emergency vehicles to navigate and smooth access to the intended destination in a timely manner. Priority value (0) given for emergency vehicles and (1) for normal vehicles. Table 2 shows the 16 selected features from whole features that were used in the proposed emergency system.

Intelligent emergency system: In our research, we designed an IES based on SOMs (unsupervised nets) and FFNN (supervised nets) to detect emergency vehicles in VANETs. SOM network consisting of input layer, the hidden layer and output layer, the hidden layer and output layer are compatible. An input layer consists of 16 neurons with the actual data and both hidden layer and output layer consists of 4 neurons as shown in Fig. 4. SOM trains a network with weight and bias learning rules with batch updates. Weights and biases updates occur at the end of an entire pass through the input data.

In our study, we used trial-and-error to select the best structure of FFNN employed in the proposed emergency system. The best structure of the FFNN is consist of three layers, the input layer consists of sixteen neurons, the hidden layer consists of two neuron and the output layer consists of three neurons as shown in Fig. 5.

The initial parameters play an important role in the performance of the FFNN that have a direct impact on the performance of detection. Table 3 shows the parameters of the training phase used in the FFNN. We design the simulation on system with an Intel core i5 processor (2.40 GHZ) and RAM memory (4GB).

Fig. 5: FFNN structure

Fig. 6: Architecture of IES

The proposed model of IES: The proposed system has five stages; Fig. 6 shows the overall architecture of the proposed IES, namely:

The first stage (Generate the mobility and the traffic model): At this step, we used two software programs to

generate the realistic scenarios that reflect vehicles behavior on roads. NS2 used the output files from SUMO and MOVE as input to generate a trace file that describes the normal and urgent behavior for vehicles in VANETs.

The second stage (data collection and pre-processing):

The features are extracted from the data in the trace file were generated in the previous step. The selected features were pre-processed to convert some letters and symbols to numeric values and to generate a uniform distribution to balance the different types of classes in collecting the data to increase the efficiency of the classification rate and normalization process to convert all the values of the features between zero and one.

The third stage (training and testing phase-SOM):

We trained and tested the SOM with the extracted features that describe normal and urgent behavior. In this stage, we obtained the detection rate, the error rate and we calculated four types of alarm.

The fourth stage (training and testing phase-FFNN):

We trained and tested the FFNN with significant data that was extracted in the second stage to check the efficiency of the proposed security system in the detection of emergency vehicles in comparison to normal vehicles.

The five stage (comparison):

In this stage, we compared the two proposed intelligent detection systems based on the three criteria, detection rate, the number of false alarms and error rate.

RESULTS AND DISCUSSION

The emergency system may be installed in three configurations: emergency vehicles, RSU or both on the emergency vehicles and RSU. In our study, we selected to install the system in RSU. The evaluation of implemented approaches and comparative results are based mainly on calculating the accuracy rate, the error rate as well as we need to calculate four types of alarms: True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR) and False Negative Rate (FNR). The most important factor is to measure the alarm rates, since, the issue is to find out the false alarms (especially false negative the most dangerous alarm). The accuracy and error rate are used as a performance metric to evaluate the proposed an IES. The accuracy and error rate of the system result should be calculated as follows:

$$\text{Accuracy rate(\%)} = \frac{N_T}{N_p} \times 100 \quad (1)$$

$$\text{Error rate(\%)} = \frac{N_F}{N_p} \times 100 \quad (2)$$

In addition, the measures will be calculated as follows:

$$\text{True positive rate(\%)} = \frac{TP}{TP+FN} \times 100 \quad (3)$$

$$\text{True negative rate(\%)} = \frac{TN}{TN+FP} \times 100 \quad (4)$$

$$\text{False negative rate(\%)} = \frac{FN}{FN+TP} \times 100 \quad (5)$$

$$\text{False positive rate(\%)} = \frac{FP}{FP+TN} \times 100 \quad (6)$$

In these equations, NP is the total number of patterns, NT is the number of correctly classified patterns, NF is the number of patterns classified as unknown, TP is the number of normal connection record classified as normal, TN is the number of urgent connection record classified as urgent, FP is the number of normal connection record classified as urgent, FN is the number of urgent connection record classified as normal.

Result of training and testing SOM neural network:

During the training and testing phase, we used the same dataset in both phases that was generated from the trace file to detect emergency vehicle in VANET. During the testing phase, calculate the total accuracy of an IES and error rate of the system. In this SOM neural network the accuracy obtained is (99.5%) and the error rate obtained is (0.5%). The experimental results obtained for TPR, FPR, TNR and FNR were explained in Table 4.

Train and test FFNN neural network:

During the training and testing phase, we used the same dataset in both phases to calculate the total accuracy of the EDS, true positive, false positive, true negative, false negative. The total accuracy of the training classification was (99.3%) and the error rate was (0.7%). The experimental results obtained for TPR, FPR, TNR and FNR were explained in Table 5.

The motivation of this research is to provide an intelligent security system that creates a safe environment for emergency vehicles. The methodology of the proposed IES was implemented in five phases: generating the mobility and traffic model, data collection and preprocessing phase, training and testing for the SOM, training and testing for the FFNN and comparing the

results that were generated in the two types of IES. When we compare the two types of IES, we can observe that the IES was based on the SOM was more

Table 4: Alarms rate

Alarm types	Accuracy(%)
True positive	100
True negative	84
False negative	0
False positive	16

Table 5: Alarms rate

Alarm types	Accuracy(%)
True positive	99.1579
True negative	100
False negative	0.84211
False positive	0

Fig. 7: Performance comparison

effective and efficient in detecting emergency vehicles with a low false negative alarm rate than the IES based on the FFNN. The comparison performance between the SOM and FFNN is as shown in Fig. 7.

The error rate for the IES based on the SOM was 0.5%. In this system, the alarm rate fluctuated between 84 and 100% with good and efficient accuracy. On the other hand, the average false negative alarm rate was 0% which is an excellent indicator of the results. At the same time, the error rate for the IES based on the FFNN was 0.7%. The alarm rate fluctuated between 100% and 99.1579% with excellent and efficient accuracy. On the other hand, the average false negative alarm rate was low at about 0.84211% which is a good indicator of the results.

We could improve the emergency system by using SHA-2 to encrypt the messages exchanged between emergency vehicles and RSUs as shown in Fig. 8.

We could improve the detection rate by adding priority feature to the trace file that creates flexibility in selecting the system that is more efficiently with different conditions. In addition, in our proposal, we selected the significant features based on the previous study (Alheeti *et al.*, 2015). All these factors make the proposed security system more efficient in securing the external communication systems for emergency vehicles.

Fig. 8: Sign the CAM using SHA-2

CONCLUSION

Security and safety are a serious issue and a crucial requirement for emergency vehicles. The network is exposed to various types of attacks that have a direct impact on the development and deployment of emergency vehicles in that zone. Our proposed IES can provide reliability for drivers by encrypting CAMs sent from the emergency vehicles to RSUs using a hash function to distinguish them from normal messages issued from other vehicles in order to ensure the protection of the network from any breach or fraud by normal vehicles to be emergency vehicles so we have achieved security and authentication. In other words, emergency vehicles without security cannot achieve their task in providing comfort and safety while in operation. In our study, we designed an intelligent security system to secure external communication and ease of movement for emergency vehicles in VANETs environment to reach the accident site in time to save lives. The IES has been designed for training and testing with important features by using SOM and FFNN. This system deals with two system scenarios: emergency and normal that have been created on the NS2. Our system is to analysis the behavior of each mobility vehicle in the VANETs to identify if it is an emergency vehicle or normal vehicle. The number and the type of features have a vital direct role to increase the efficiency and the accuracy of the detection system. Further attentions should be given to use a fuzzy data set to reduce the rate of error, the number of false alarms and may be used to represent integrity to dataset with fuzzy concepts.

REFERENCES

- Alheeti, K.M.A., A. Gruebler and K.D. McDonald-Maier, 2015. On the detection of grey hole and rushing attacks in self-driving vehicular networks. Proceedings of the 7th IEEE Conference on Computer Science and Electronic Engineering (CEECE'15), September 24-25, 2015, IEEE, Colchester, England, ISBN:978-1-4673-9481-9, pp: 231-236.
- Anonymous, 2009. Intelligent transportation systems. Ministry of Transportation, Toronto, Canada.
- Assila, A., I. Jabri and A. Ltifi, 2012. Secure architecture dedicated for vanet alarm messages authentication through semantic verification. Proceedings of the 6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT'12), March 21-24, 2012, IEEE, Sousse, Tunisia, ISBN:978-1-4673-1657-6, pp: 652-657.
- Athavan, K., G. Balasubramanian, S. Jagadeeshwaran and N. Dinesh, 2012. Automatic ambulance rescue system. Proceedings of the 2nd International Conference on Advanced Computing and Communication Technologies (ACCT'12), January 7-8, 2012, IEEE, Rohtak, India, ISBN: 978-1-4673-0471-9, pp: 190-195.
- Buchenscheit, A., F. Schaub, F. Kargl and M. Weber, 2009. A VANET-based emergency vehicle warning system. Proceedings of the 2009 IEEE Conference on Vehicular Networking (VNC'09), October 28-30, 2009, IEEE, Tokyo, Japan, ISBN:978-1-4244-5685-7, pp: 1-8.
- CCC., 2011. The handbook for vehicle-to-X cooperative systems simulation. Car 2 Car Communication Consortium, Europe.
- Danquah, W. and D. Altılar, 2014. Hybrist mobility model: A novel hybrid mobility model for VANET simulations. Intl. J. Comput. Appl., 86: 15-21.
- Dass, R., R. Sangwan and I. Girdhar, 2012. Vehicular Ad Hoc networks. Intl. J. Adv. Trends Comput. Sci. Eng., 1: 121-129.
- Doijad, G., P. Kamble and M. Kshirsagar, 2015. Design and implementation of secure vehicle to vehicle formatted communication. Intl. J. Adv. Res. Comput. Sci. Software Eng., 5: 84-88.
- Gaur, S.K., S.K. Tyagi and P. Singh, 2013. VANET system for vehicular security applications. Intl. J. Soft Comput. Eng., 2: 279-282.
- Hung, C.C., H. Chan and E.H.K. Wu, 2008. Mobility pattern aware routing for heterogeneous vehicular networks. Proceedings of the IEEE Wireless Communications and Networking (WCNC'08), March 31- April 3, 2008, IEEE, Las Vegas, Nevada, ISBN:978-1-4244-1997-5, pp: 2200-2205.
- Kabiri, P. and M. Aghaei, 2011. Feature analysis for intrusion detection in mobile Ad-hoc networks. Intl. J. Netw. Secur., 12: 42-49.
- Karagiannis, G., O. Altintas, E. Ekici, G. Heijenk and B. Jarupan *et al.*, 2011. Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions. IEEE. Commun. Surv. Tutorials, 13: 584-616.
- Raut, Y.C., V.N. Katkar and S.M. Sarode, 2014. Early alert system using relative positioning in vehicular ad-hoc network Proceedings of the 2014 Annual International Conference on Emerging Research Areas: Magnetics Machines and Drives (AICERA-iCMMMD'14), July 24-26, 2014, IEEE, Kottayam, India, ISBN:978-1-4799-5203-8, pp: 1-8.

- Samara, G., W.A. Al-Salihy and R. Sures, 2010. Security issues and challenges of Vehicular Ad Hoc Networks (VANET). Proceedings of the 4th International Conference on New Trends in Information Science and Service Science (NISS'10), May 11-13, 2010, IEEE, Gyeongju, South Korea, ISBN:978-1-4244-6982-6, pp: 393-398.
- Yang, Q., L. Wang, W. Xia, Y. Wu and L. Shen, 2014. Development of on-board unit in vehicular ad-hoc network for highways. Proceedings of the International Conference on Connected Vehicles and Expo (ICCVE'14), November 3-7, 2014, IEEE, Vienna, Austria, pp: 457-462.
- Zhou, L. and Z. Hass, 1999. Security ad hoc networks. IEEE. Netw. Mag., 13: 24-30.

Success Factors for Utilizing e-Learning Systems in Higher Education Institutions

Khaldoon Hussein Aliway and Nurhizam Safie
Research Centre for Software Technology and Management (SOFTAM),
Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM),
43600 Bangi, Malaysia

Abstract: e-Learning is important in developing countries in order to chart new directions in education, especially in Iraq's higher education sector. It has been started and implemented for educational associations. Most of the world has embraced the concept of open education via the internet or by enabling students to access information and knowledge resources from anywhere and anytime. e-Learning offers many advantages to institutions which can use this system to grow the level of education for different kinds of students including normal students and those with disabilities. However, the e-Learning system faces many challenges in developing countries such as Iraq because of the lack of understanding and weaknesses in experience and skills among the students as well as among many institutions. Also, there is a lack of customization to facilitate the development of ICT. e-Learning system can be useful for students if the critical success factors associated with the utilization of ICT e-Learning systems are determined.

Key words: e-Learning, success factors, ICT, higher education, utilization, Iraq

INTRODUCTION

e-Learning is explained as an educational way of providing learning to stakeholders through courses and programs employing new technology on the internet. In other words, the stakeholders such as students and teachers can utilize this platform to improve learning skills at anytime and anywhere by using desktop computers, laptops, tablets and smartphones, all of which are powered by internet technologies (Garrido *et al.*, 2016). In addition, e-Learning is defined as any platform technology that is able to develop effective students through the open education concept of e-Learning which has proved to be a game-changer in the field of education, particularly higher education in this new millennium. (Aparicio *et al.*, 2016).

In tandem with the growth and rapid advances in technology and Internet requests in this century, ICT is an important component that makes e-Learning more beneficial and accessible. Furthermore, with the use of internet technology, e-Learning systems allow education programs to be planned without having to be constrained by time and place (Fahad *et al.*, 2013; Ahmed *et al.*, 2016). This novel approach to virtual teaching and learning has been widely accepted as significantly enhances higher education and knowledge acquisition (Chen and Tseng, 2012).

The world after the technology revolution and emergence of internet-based higher education since late in the first decade of the 21st century has been utilizing e-Learning in several establishments of higher education for the purpose of increasing the number of students as well as enhancing their skills and experiences (Fahad *et al.*, 2013).

Recently, some researchers have designed models for promoting e-Learning and considered the implementation of an e-Learning system with specific spaces suggested but these models have a weakness in considering the students as the main focus (Pata, 2009).

As such, ICT is important to improve and develop the utilization of the e-Learning system. Infrastructure is prepared to be compatible with any modern technology and this technology is more flexible when applied in high education. Furthermore, this technology must be useful and easy to use among both students and teachers.

According to a UNESCO report in 2011, Iraqi higher education has been limited in using e-Learning based on information and communication technology. In addition, UNESCO has been initiated a different approach to improve ICT infrastructure for students that would embrace a transformation in the quality of e-Learning system based on ICT.

Unfortunately, e-Learning is still a new experience in Iraqi higher education because Iraqi universities have

been the last to adopt e-Learning in the Middle East. Furthermore, until 2009, only 1% of Iraqi students had Internet access. Thus, Iraq lags behind other countries in terms of benefitting from the revolution of e-Learning based on ICT in higher education. Therefore, The Ministry of Higher Education has made concerted efforts to raise the status of Iraq universities. However, strategically, Iraqi higher education has seen limited progress in the adoption of e-Learning (Al-Azawei and Lundqvist, 2015a, b).

Therefore, high lighting the requirements in order to implement e-Learning system based on ICT in Iraqi universities is important to fill the gap in the literature by determining the critical success factors and specifying the major barriers that have a significant effect on adoption of the e-Learning system. In the context of the above-mentioned scenario, the researchers discuss in this study, the e-Learning system based on ICT in Iraqi universities in order to identify the success factors and have a clear understanding to ensure the successful implementation of the e-Learning project in future.

According to Safie and Aljunid (2013), the study high lights the use of e-Learning system which is of great importance as e-Learning system plays a pivotal role in the developing different organizations in education and healthcare among others which is essential in developing countries and in in line with the ideals of the Millennium development goals. Furthermore, it has been shown that e-Learning is time saving as well as cost-effective.

MATERIALS AND METHODS

In this study, the literature was reviewed to collect secondary date related to e-Learning systems in order to discuss ways of adopting a conceptual model to show the relationship between factors that affect the utilization of an e-Learning system in higher education.

The current issues in some developing countries such as Iraq, Libya, etc. are handicapped regarding the application of the appropriate resources and the required experience in terms of e-Learning (Fahad *et al.*, 2013).

Although, the ministry of higher education in Iraq has taken steps to promote teaching and learning through e-Learning in Iraq (Al-Azawei and Lundqvist, 2015a, b), problems persist in coping with the rapidly changing challenges that require innovation.

Furthermore, UNESCO has agreed that e-Learning is an essential tool which needs to be adopted and adapted for implementation in most education sectors for Iraq (Basha *et al.*, 2013).

This study aims to determine the critical success factors of e-Learning systems based on ICT in Iraqi public

Table 1: Frequency of factors

Factors	References
IT infrastructure	Alsabawy <i>et al.</i> (2013, 2016) and Jabbouri <i>et al.</i> (2016)
System quality	Hassanzadeh <i>et al.</i> (2012), Bhuasiri <i>et al.</i> (2012) and Liaw (2008)
Multimedia instruction	Liaw (2008)
Service quality	Hassanzadeh <i>et al.</i> (2012)
Ease of use	Al-Azawei and Lundqvist (2015a, b), Basha <i>et al.</i> (2013), Aparicio <i>et al.</i> (2016), Abdullah <i>et al.</i> (2016), Chen and Tseng (2012) and Wu and Chen (2017)
Usefulness	Chen and Tseng (2012), Wu and Chen (2017), Alsabawy <i>et al.</i> (2013, 2016), Liaw (2008), Bhuasiri <i>et al.</i> (2012), Aparicio <i>et al.</i> (2016) Abdullah <i>et al.</i> (2016), Sun <i>et al.</i> (2008), Chen and Tseng (2012), Islam (2013), Al-Gahtani (2016) and Lee <i>et al.</i> (2009)
Openness	Wu and Chen (2017)
Self-efficacy	Liaw (2008), Al-Azawei and Lundqvist (2015a, b), Sun <i>et al.</i> (2008), Lim <i>et al.</i> (2007), Chen and Tseng (2012), Tarhini <i>et al.</i> (2013) and Islam (2013)
User satisfaction	Alsabawy <i>et al.</i> (2013), Hassanzadeh <i>et al.</i> (2012), Liaw (2008), Al-Azawei and Lundqvist (2015a, b), Sun <i>et al.</i> (2008) Wu <i>et al.</i> (2010), Lin (2012), Chow and Shi (2014) and Paechter <i>et al.</i> (2010)

universities. In addition, the e-Learning systems should be supported and both students and lecturers should be encouraged to use these technologies. The various relevant factors are summarized and presented in the table according to the references in Table 1.

Success factors: As we can see, the majority of the researches focuses on the usefulness factor while some factors have low frequency as shown in Fig. 1. Therefore, it is recommended to focus on the rarely considered factors in this research area in order to ensure a successful utilization of e-Learning system.

After reviewing the literature for the success determinants, the researchers recommend the following factors which can draw a successful the utilization of e-Learning system in higher education institutions.

IT infrastructure: An IT infrastructure defines that all new technologies should be considered to make the use of the e-Learning system more flexible and integrated. In addition, IT infrastructure for stakeholders should include both technical and management skills which are important to give the best results. The capability of IT infrastructure should be exploited in using e-Learning system to achieve a better service for the stakeholders in terms of accuracy and confidence. Therefore, the necessary infrastructure and appropriate human resources are necessary for the successful implementation of modern technology (Gorla *et al.*, 2010; Kadhum and Hasan, 2017).

The IT infrastructure includes all the IT technologies such as internet bandwidth, network equipment and a

Fig. 1: The percentage of frequent factors

sufficient number of computers (Raouf *et al.*, 2012). IT infrastructure is one of the key factors in information systems and is a prerequisite for the successful implementation of any modern technology which means that the universities must pay attention to establish an adequate IT infrastructure to achieve competitive performance (Alsabawy *et al.*, 2016).

The IT infrastructure is a crucial factor in the implementation of the service activities in education and in the context of e-Learning systems, it forms the basis of computer technology, communications and the essentials of the information system within the technical scope to improve education. With an adequately established IT infrastructure an organization can deliver reliable services through the central information system. To achieve competitive advantage higher education can adopt a method to classify the IT resources such as hardware, software, networking and communications, human resources and databases (Jabbouri *et al.*, 2016).

According to Alsabawy *et al.* (2013, 2016) and Jabbouri *et al.* (2016), this factor has a significant effect on the utilization of the e-Learning system.

System quality: Today, one of the biggest problems facing e-Learning system is system quality that can facilitate the successful implementation of technology in universities. In fact, system quality is a multi-dimensional measure and it is important to determine how system quality can help the higher education sector in developing

countries to build and improve strategies. The significance of system quality for organizational performance has been highlighted (Gorla *et al.*, 2010).

The benefits of this factor in e-Learning system are reflected it assists in functions and end-user simplification in education operation. e-Learning system is important for the delivery of learning outcomes because students interact more in an e-Learning environment than in traditional face-to-face instructions. System quality affects a student's belief in e-Learning performance characteristics (Bhuasiri *et al.*, 2012). Furthermore, it is measured based on the way the system is run, ease of use, accessibility, flexibility, information quality, portability, integration and significance (DeLone and McLean, 2003).

According to Alsabawy *et al.* (2016), this can be considered as a significant factor in using e-Learning system and system quality is the important factor in evaluating the success of the e-Learning system.

In addition, the researcher (Alsabawy *et al.*, 2013) high lighted this factor which determines continuous implementation of the e-Learning systems.

Multimedia instruction: Multimedia instruction include voice and graphics information such as picture and text, so, learning can be more interesting and easy and effective for students (Tabbers *et al.*, 2001).

According to Liaw (2008), multimedia instruction assist stakeholders such as students and lecturers to increase their skills and experience in voice or video media by converting the difficult pedagogical materials and making them easier to understand. Furthermore, this factor would help students to use e-Learning system for other acquired concepts and enable them to apply conceptual knowledge using the flexibility of e-Learning systems.

Service quality: Service quality is known as the scale of service delivered by the information system service. This factor involves supplying the stakeholders, parallels to their anticipation in terms of expression, accuracy, realization, understanding, confirmation and consideration. These concepts of service quality are reflected in the stakeholder's reactions to the ability of the system to meet their expectations, through experts in information systems which give advice freely. Furthermore, providing services to learners when they are needed builds confidence (Gorla *et al.*, 2010).

On the other hand, this factor implies providing a range of services to lecturers and students reliably, so, they are aware of all the benefits of the system. This can have a significant effect on using e-Learning system. Some researchers think that this factor of service quality

is a division of system quality but it has been changing the function of information systems in recent years and has become a separate factor (Hassanzadeh *et al.*, 2012). According to Alsabawy *et al.* (2016), service quality has an effect on utilization e-Learning system in the universities. Previous studies showed that universities and other organizations pay more attention to improve service quality. According to Sun *et al.* (2008), service quality will positively influence utilization of e-Learning systems.

Ease of use: Ease of use is defined as “the degree to which an individual believes that using a particular system would be free of physical and mental effort” (Davis, 1989). Furthermore, Technology Acceptance Model and TAM2 explain the significant effect of ease-of-use when using a new technology or system (Venkatesh and Davis, 2000). It is directly associated with using e-Learning because students, more specifically in training courses are reluctant to continue using an e-Learning system if they face difficulties in utilizing it which could induce them to drop a course or to change to another learning environment (Al-Azawei and Lundqvist, 2015a, b).

As such, it is crucial to show how easy and effective a system is in terms of use and system effectiveness which would encourage understanding and acceptance of modern technology among students (Lim *et al.*, 2007).

According to Bhuasiri *et al.* (2012), prospective stakeholders and learner anticipation in using e-Learning system can be made easier and more compressive. In otherwords, the ease-of-use has a significant effect on utilization of e-Learning systems (Sun *et al.*, 2008) (Al-Azawei and Lundqvist, 2015a, b).

Ease-of-use in contrast, refers to “the degree to which a person believes that using a particular system would be free of effort”. This definition refers to “ease” “freedom from difficulty or great effort”. Effort is a finite resource that a person may allocate to various activities for which he or she is responsible (Davis, 1989).

In addition, the UTAUT Model states that the “effort expectancy” construct can be important in determining user approval of e-Learning systems and worry about ease-of-use may become non-important through extended implementation. Thus, ease-of-use can be expected to be significant in the early stages of utilizing e-Learning systems (Marchewka *et al.*, 2007).

Usefulness: Davis (1989) defined usefulness as “the degree to which a person believes that using a particular

system would enhance his or her job performance” (Davis, 1989). According to Davis *et al.* (1989), usefulness is considered to be the main construct in the original representation of the Technology Acceptance Model (TAM).

According to Bhuasiri *et al.* (2012), this factor measures the degree to which a stakeholder trusts that using e-Learning system will enhance achievement. In addition, usefulness shows how the students use the e-Learning system to enhance their research and execution and with expectations of more benefits in the future (Gorla *et al.*, 2010). Furthermore, this component can support using e-Learning system in higher education (Sun *et al.*, 2008; Chen and Tseng, 2012). In addition, this factor is considered important to measure the success of using e-Learning system and acceptance of any new technology (Alsabawy *et al.*, 2016).

In the e-Learning system, usefulness has been much and frequently used by researchers such as Arbaugh (2000). Usefulness or benefits of using technology help to deliver educational materials and promote the behavior of students toward their educational skills. Thus, it would help students to learn and be educated through the internet in the future.

Openness: Today, modern universities focus on a mix of technology and teaching materials to make education more visible and accessible to deliver resources anytime and anywhere which enables education to be more flexible. The idea is to encourage students to be more motivated toward changing traditional practices in learning and embrace new technologies in education (Wu and Chen, 2017; Safie, 2004).

Self-efficacy: Self-efficacy has been described as students cognitive trust that significantly affects their behavior when utilizing e-Learning system and the stakeholder has the ability to implement e-Learning with more confidence in achieving successful utilization (Wu *et al.*, 2010).

e-Learning has recently been used and has gained popularity in higher education in developing countries which means that the self-efficacy factor is important in helping students to utilize the new technology that is inherent in the e-Learning system (Bhuasiri *et al.*, 2012; Wu *et al.*, 2010; Samsuri *et al.*, 2014).

Self-efficacy is reflected in student’s desire to play a private functional part. It is an estimation of the effect and potential for success before implementing the new system. Students with rising self-efficacy can be confident in performing the e-Learning system activities

and developing their skills. Much research has explained the impact of self-efficacy on using new technology. Furthermore, it has been reported that particular self-efficacies can impact learning outcomes when students are using e-Learning system (Joo *et al.*, 2000; Lewis, 2002; Sun *et al.*, 2008).

According to Wang and Newlin (2002), investigation has shown that students tend to use the factor of self-efficacy to adopt network-based learning which significantly affects the utilization of the e-Learning system. This study defined this factor as the student's ability to evaluate the viability of implementing activities required in the e-Learning system.

User satisfaction: User satisfaction is shown in the way students think about using the e-Learning system mainly to evaluate the success in implementing modern technology. Furthermore, this is one of the critical factors with a direct effect on utilizing the e-Learning system (Hassanzadeh *et al.*, 2012) and it is not surprising that many earlier researches have focused on user satisfaction (Hassanzadeh *et al.*, 2012).

According to Min *et al.* (2008) it has been found that user satisfaction is one of the critical factors that influence the utilization of e-Learning because there are significantly different user skills involved such as transmission, input and output, features which endear students to the e-Learning system. It can therefore be concluded that for a system like e-Learning to be accepted and used it must offer user satisfaction. According to Wixom and Todd (2005) it has been shown that the various design attributes explicitly enumerated such as accuracy and reliability contribute significantly to user satisfaction and the utilization of the e-Learning system. Furthermore, TAM and user satisfaction have developed in parallel study streams and the two approaches can be integrated to help in building a conceptual model incorporating the design and system characteristics to forecast the success of utilizing e-Learning system.

By DeLone and McLean (1992) the researcher shows that user satisfaction is one of the significant factors that can estimate the success of utilizing e-Learning because some users have stopped their e-Learning following a first attempt.

According to Roca *et al.* (2006), the user satisfaction factor directly affects learner impact and has significant impact on the organization. Furthermore, this factor is important as it affects e-Learning usage besides having the strongest, direct personal impact. In light of the many benefits of the user-satisfaction factor, it can be used by

universities to target resources more efficiently, achieve improvements in user satisfaction and increase the probability of users reusing the e-Learning system.

RESULTS AND DISCUSSION

e-Learning has reached a new level of importance for decision makers in higher education institutions. It offers a new direction in developing countries like Iraq to be used in teaching and learning, so, implementing and using this system has become essential for higher education in the country. Use of e-Learning can help students through a system of distance education using the internet alone or to support the student's activities wherever and whenever needed, developing the interaction of class students beyond the class without the need for face-to-face meetings or to support semi present teaching activities.

Although, there has been acceptance of e-Learning as an important teaching-learning platform to our best knowledge there have not been many studies that have evaluated the determining factors of successful e-Learning implementation.

This study provides related studies in the field of online education that have presented critical success factors influencing utilization of the e-Learning system based on ICT implementation, although, the majority of the available research has focused on technological and educational aspects which represent the operational level in the organization. The aim of the study is appropriate because it highlight multiple organizational factors that can lead to successful project implementation of e-Learning in institutes of higher education.

An approach to understanding e-Learning based on ICT should consider the determinants that should be effectively addressed if a project is to succeed. To the best of our knowledge here have been very few studies that have identified e-Learning based on ICT critical success factors.

A research focused specifically on the critical determinants highlights the significant elements that must be taken into consideration when implementing the e-Learning system in higher education. The literature review, we conducted has gathered relevant information from multiple studies related to this research focus on utilizing e-Learning in higher education.

Figure 2 explains the conceptual model of the present study. The researcher suggest these factors in order to help the decision makers in Iraqi Universities to successfully adopt e-Learning system based on ICT

Fig. 2: Conceptual model

technology by understanding the effect of these factors on successful adoption. It is suggested that future research on this topic be empirically investigated.

CONCLUSION

This study aims to investigate the success factors to utilize ICT e-Learning systems in Iraqi universities in order to make students more satisfied when using this system. The study used a literature survey to gather related information and identify relevant factors and their effect on student satisfaction with using e-Learning system. The outcome revealed that nine factors should be highlighted when utilizing e-Learning system based on ICT in Iraqi public universities.

REFERENCES

- Abdullah, F., R. Ward and E. Ahmed, 2016. Investigating the influence of the most commonly used external variables of TAM on students Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) of E-portfolios. *Comput. Hum. Behav.*, 63: 75-90.
- Ahmed, F., H. Zainuddin and S. Zahraa, 2016. Barriers of adopting E-learning in developing countries. *J. Eng. Appl. Sci.*, 11: 2815-2820.
- Al-Azawei, A. and K. Lundqvist, 2015a. Development of a universal design for learning-based teaching collaborative system (UDL-BTCS) to support accessible learning. *Proceedings of the 7th Conference on Computer Science and Electronic Engineering (CEECE'15)*, September 24-25, 2015, IEEE, Colchester, England, UK., ISBN:978-1-4673-9480-2, pp: 129-134.
- Al-Azawei, A. and K. Lundqvist, 2015b. Learner differences in perceived satisfaction of an online learning: An extension to the technology acceptance model in an Arabic sample. *Electron. J. E. Learn.*, 13: 408-426.
- Al-Gahtani, S.S., 2016. Empirical investigation of E-learning acceptance and assimilation: A structural equation model. *Appl. Comput. Inf.*, 12: 27-50.
- Alsabawy, A.Y., A. Cater-Steel and J. Soar, 2013. IT infrastructure services as a requirement for E-learning system success. *Comput. Educ.*, 69: 431-451.
- Alsabawy, A.Y., A. Cater-Steel and J. Soar, 2016. Determinants of perceived usefulness of E-learning systems. *Comput. Hum. Behav.*, 64: 843-858.
- Aparicio, M., F. Bacao and T. Oliveira, 2016. Cultural impacts on E-learning systems success. *Internet Higher Educ.*, 31: 58-70.
- Arbaugh, J.B., 2000. How classroom environment and student engagement affect learning in internet-based MBA courses. *Bus. Commun. Q.*, 63:9-26.
- Basha, A.D., I.N. Umar and M. Abbas, 2013. A sight on the use of ICT resources for E-learning in the Iraqi Higher Education Institutions (IHEL): A framework for the use of ICT in E-learning. *Proceedings of the 4th International Conference on Intelligent Systems Modelling and Simulation (ISMS)*, January 29-31, 2013, IEEE, Bangkok, Thailand, ISBN:978-1-4673-5653-4, pp: 327-330.
- Bhuasiri, W., O. Xaymoungkhoun, H. Zo, J.J. Rho and A.P. Ciganek, 2012. Critical success factors for e-Learning in developing countries: A comparative analysis between ICT experts and faculty. *Comput. Educ.*, 58: 843-855.
- Chen, H.R. and H.F. Tseng, 2012. Factors that influence acceptance of web-based e-learning systems for the in-service education of junior high school teachers in Taiwan. *Eval. Prog. Plann.*, 35: 398-406.
- Chow, W.S. and S. Shi, 2014. Investigating students satisfaction and continuance intention toward E-learning: An extension of the expectation-confirmation model. *Procedia Social Behav. Sci.*, 141: 1145-1149.
- Davis, F.D., 1989. Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS. Q.*, 29: 319-340.
- Davis, F.D., R.P. Bagozzi and P.R. Warshaw, 1989. User acceptance of computer technology: A comparison of two theoretical models. *Manage. Sci.*, 35: 982-1003.
- DeLone, W.D. and E.R. McLean, 2003. The DeLone and McLean model of information systems success: A ten-year update. *J. Manage. Inform. Syst.*, 19: 9-30.
- DeLone, W.H. and E.R. McLean, 1992. Information systems success: The quest for the dependent variable. *Inform. Syst. Res.*, 3: 60-95.
- Fahad, A., Z. Hassan and Z. Salman, 2013. A study towards using E-learning based information technology resources for the development of students thinking skills in higher education. *Proceedings of the 4th International Conference on Intelligent Systems Modelling and Simulation (ISMS'13)*, January 29-31, 2013, IEEE, Bangkok, Thailand, ISBN:978-1-4673-5653-4, pp: 331-334.
- Garrido, A., L. Morales and I. Serina, 2016. On the use of case-based planning for E-learning personalization. *Expert Syst. Appl.*, 60: 1-15.

- Gorla, N., T.M. Somers and B. Wong, 2010. Organizational impact of system quality, information quality and service quality. *J. Strategic Inf. Syst.*, 19: 207-228.
- Hassanzadeh, A., F. Kanaani and S. Elahi, 2012. A model for measuring E-learning systems success in universities. *Expert Syst. Appl.*, 39: 10959-10966.
- Islam, A.N., 2013. Investigating E-learning system usage outcomes in the university context. *Comput. Educ.*, 69: 387-399.
- Jabbouri, N.I., R. Siron, I. Zahari and M. Khalid, 2016. Impact of information technology infrastructure on innovation performance: An empirical study on private universities in Iraq. *Procedia Econ. Finance*, 39: 861-869.
- Joo, Y.J., M. Bong and H.J. Choi, 2000. Self-efficacy for self-regulated learning, academic self-efficacy and internet self-efficacy in web-based instruction. *Educ. Technol. Res. Dev.*, 48: 5-17.
- Kadhum, A.M. and M.K. Hasan, 2017. Assessing the determinants of cloud computing services for utilizing health information systems: A case study. *Intl. J. Adv. Sci. Eng. Inf. Technol.*, 7: 503-510.
- Lee, B.C., J.O. Yoon and I. Lee, 2009. Learner's acceptance of e-learning in South Korea: Theories and results. *Comput. Educ.*, 53: 1320-1329.
- Lewis, C., 2002. Driving factors for e-learning: An organisational perspective. *Perspect. Policy Pract. Higher Educ.*, 6: 50-54.
- Liaw, S.S., 2008. Investigating students perceived satisfaction, behavioral intention and effectiveness of E-learning: A case study of the blackboard system. *Comput. Educ.*, 51: 864-873.
- Lim, H., S.G. Lee and K. Nam, 2007. Validating E-learning factors affecting training effectiveness. *Intl. J. Inf. Manage.*, 27: 22-35.
- Lin, W.S., 2012. Perceived fit and satisfaction on web learning performance: IS continuance intention and task-technology fit perspectives. *Intl. J. Hum. Comput. Stud.*, 70: 498-507.
- Marchewka, J.T., C. Liu and K. Kostiwa, 2007. An application of the UTAUT model for understanding student perceptions using course management software. *Commun. IIMA*, 7: 93-104.
- Min, Q., S. Ji and G. Qu, 2008. Mobile commerce user acceptance study in China: A revised UTAUT model. *Tsinghua Sci. Technol.*, 13: 257-264.
- Paechter, M., B. Maier and D. Macher, 2010. Students expectations of and experiences in E-learning: Their relation to learning achievements and course satisfaction. *Comput. Educ.*, 54: 222-229.
- Pata, K., 2009. Modeling spaces for self-directed learning at university courses. *J. Educ. Technol. Soc.*, 12: 23-43.
- Raouf, J.B., I.S. Naser and B.K. Jassim, 2012. Determinants of E-learning implementation success in the Iraqi MoHE. *Eng. Tech. J.*, 30: 659-671.
- Roca, J.C., C.M. Chiu and F.J. Martinez, 2006. Understanding e-learning continuance intention: An extension of the technology acceptance model. *Int. J. Hum. Comput. Stud.*, 64: 683-696.
- Safie, N. and S. Aljunid, 2013. E-learning initiative capacity building for healthcare workforce of developing countries. *J. Comput. Sci.*, 9: 583-591.
- Safie, N., 2004. The use of Short Messaging System (SMS) as a supplementary learning tool in Open University Malaysia (OUM). *Proceedings of the 18th Annual Conference on Association of Asian Open Universities (AAOU'04) Vol. 27, November 27-30, 2004, Shanghai Open University, Shanghai, China*, pp: 1-11.
- Samsuri, N.N., F.A. Nadzri and K.B.M. Rom, 2014. A study on the student's perspective on the effectiveness of using E-learning. *Procedia Social Behav. Sci.*, 123: 139-144.
- Sun, P., R. Tsai, G. Finger, Y. Chen and D. Yeh, 2008. What drives a successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction. *Comput. Educ.*, 50: 1183-1202.
- Tabbers, H.K., R.L. Martens and J.J.V. Merrienboer, 2001. The modality effect in multimedia instructions. *Proc. Cognit. Sci. Soc.*, 23: 1-10.
- Tarhini, A., K. Hone and X. Liu, 2013. User acceptance towards web-based learning systems: Investigating the role of social, organizational and individual factors in European higher education. *Procedia Comput. Sci.*, 17: 189-197.
- Venkatesh, V. and F.D. Davis, 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Manage. Sci.*, 46: 186-204.
- Wang, A.Y. and M.H. Newlin, 2002. Predictors of web-student performance: The role of self-efficacy and reasons for taking an on-line class. *Comput. Hum. Behav.*, 18: 151-163.
- Wixom, B.H. and P.A. Todd, 2005. A theoretical integration of user satisfaction and technology acceptance. *Inform. Syst. Res.*, 16: 85-102.
- Wu, B. and X. Chen, 2017. Continuance intention to use MOOCs: Integrating the Technology Acceptance Model (TAM) and Task Technology Fit (TTF) model. *Comput. Hum. Behav.*, 67: 221-232.
- Wu, J.H., R.D. Tennyson and T.L. Hsia, 2010. A study of student satisfaction in a blended E-learning system environment. *Comput. Educ.*, 55: 155-164.

The Development of National Spatial Data Infrastructure (NSDI) in Jordan-Challenges and Awareness

¹Ibrahim Gharaibeh, ¹Mohd Sanusi S. Ahamad and ²Bassam Saleh Malkawi
¹School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia,
14300 Nibong Tebal, P. Pinang, Malaysia
²Faculty of Engineering, Isra University, Amman, Jordan

Abstract: The aims of this study to examine available and missing Jordan NSDI components because the developments of current status of Jordan NSDI is not available. Then explore the level to which Jordan is prepared to accomplish an NSDI and identify main challenges. This study adopted a modified state of play assessment framework which is one of the multi-view assessment frameworks, gathering data was by questionnaire surveys and interview from 29 of stakeholders as most organizations that have a GIS division and are registered under the JNSDI stakeholder categories. Jordan has no organizational framework and there is an institutional void in leadership responsibility in regard to GIS, no metadata, no data model, poor documentation and there is no national strategy for GIS. Lack of awareness on the importance of JSDI by the organizations prevents the government from supporting JSDI through distinctive budget allocation, although, Jordan has a rich human resource base and educational infrastructures. On the other hand, there are many advantages to create Jordan NSDI like barring the duplication in the spatial data, eliminating the redundancy of the spatial data, raising the collaboration among organizations and different departments. Consequently, only cooperation between the public sector, private sector, academia and non-government organizations can achieve a successful NSDI system. Assessment of the current status of NSDI is the starting point of planning to build the principles of Jordan NSDI to promote the creation of Jordan NSDI. NSDIs is not identical in the world and it is important for Jordan to develop national a strategy by itself. The process is to prove the possible benefits of such initiatives and to make awareness and better understanding between the decision-makers, users and producers involved in the SDI activities of the Jordan to move forward.

Key words: Spatial data infrastructures, challenges for SDI, SDI components, Jordan NSDI, SDI development, activities, creation

INTRODUCTION

Spatial Data Infrastructure (SDI) is a term covering the entire spatial data domains (Groot and McLaughlin, 2000). At present, there are more than 120 states that have developed their national SDI (Crompvoets *et al.*, 2004). According to the office of management and budget in the USA, the NSDI confirms that spatial data from multiple agencies (Federal, state, local, academia and the private sector) are existing and easily integrated to enhance the understanding of our cultural and physical world (OMB., 2002). Meanwhile, the Federal Geographic Data Committee (FGDC) has defined NSDI as an umbrella that covers policies (institutional frameworks), standards, organizational procedures and technologies in the usage, management and production of spatial data.

The Canadian Spatial Data Infrastructure (CSDI) comprises of the technology, policies (Institutional frameworks) and standards that make access to geo-spatial data bases available on the internet. The SDI notion is generally used to avoid discrepancies within spatial datasets and to eliminate redundant data production addendum to other famed spatial data problems (Loenen and Kok, 2017). An SDI is also a process to prepare and enlighten a technique for sharing and developing spatial datasets. As a result, data owners communicate with the current technologies to realize the varied degrees of political and organizational goals in an effective way (Chan *et al.*, 2001). Coleman and McLaughlin (1998) includes the policies (institutional frameworks), technologies, standards and human resources (people) needed for the efficient collection,

management, access, delivery and employment of geo-spatial data in an global community. Therefore, the National Spatial Data Infrastructure (NSDI) will consist of the standards, technology, policies (institutional frameworks) criteria and people (Partnership) to promote geospatial data sharing throughout all levels of government, public sector, private sectors and academia (FGDC., 1997). Spatial Data Infrastructure (SDI) can also act as a framework that help the sharing, assortment, accessibility and swapping of spatial data between users among the spatial community (Crompvoets *et al.*, 2004). In this regard, NSDIs are being executed by many nations to better management and utilization of spatial datasets (Rajabifard *et al.*, 2003). Consequently, SDI provides a foundation for spatial data discovery, evaluation and application for all users and providers with all levels of national governments, academia, civil societies and NGOs. Furthermore, the NSDI of a nation can be used for the network survey of coordinates, transportation (railway networks and road), communication facilities, electricity supply, cultivation, fishing line, forestry nations, tourism and planning of services (Okuku *et al.*, 2014).

Meanwhile, SDIs have also been defined in a different way by different researchers at different period. This point out the multifaceted model of SDIs (Man, 2006) and as agreed by Grus (2010) SDI assessment still problematic because of its dynamic, multifaceted and complex nature and its ambiguously defined objectives. The ability to meet user’s needs and to deliver services and tools within the spatial information community has gone far beyond the capacity of single organizations, especially when more value-added and integrated spatial data are required for more complex analysis. As a result, organizations move towards the sharing of spatial datasets and collecting and integrating spatial data from different sources (Akinyemi and Uwayezu, 2011). Therefore, the goals of NSDI are defined through the legislative framework, political initiatives and by applying the management and funding models (Hecimovic *et al.*, 2014).

Based on the argument above, this study aims to define the status of development of the National Spatial Data Infrastructure (NSDI) in Jordan in order to examine data interoperability issues and to define the challenges of NSDI in a developing country like Jordan.

MATERIALS AND METHODS

Components of SDI: Rajabifard and Williamson (2003) breaks SDI into five core components, i.e., people, access networks, standards, policy and data (Fig. 1). However, there are certain organizations and researchers that have

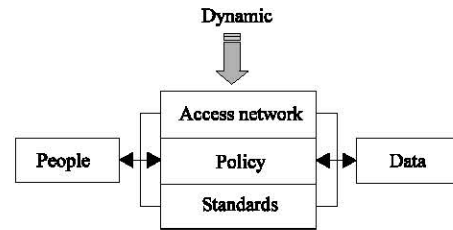


Fig. 1: Interrelation among SDI components adopted from rajabifard

built their own NSDI models depending on their vision, goals and the national basic requirements and priorities. For example, the FGDC created an NSDI with 6 components, partnerships (people) portal/clearing house, framework, metadata, Geo-data and standards) (Masser, 1997; Tosta, 1995). In addition, the Australia New Zealand Land Information Council (ANZLIC) has created an NSDI model that comprise of four components. The components are consist of technical standards, institutional framework (policy) clearing house network (access network) and fundamental datasets (ANZLIC., 1998).

Data (fundamental datasets): The aim of this component is to create datasets that offer unique geo-referenced environment to guarantee an easy conversion of data between organizations. Moreover, the existence of such datasets will minimize the efforts in collecting and managing the data and prevent duplication in fields such as imagery, geodetic control, elevation, transportation (road, railway) cadastre, hydrography and governmental units (Tosta, 1995).

Standards: These are the common and repeated rules, conditions, guidelines or characteristics for the data and related processes, technology and organization. To be successful in its objectives each national spatial data infrastructure needs standards for aspects of data (dictionary, reference system, quality, models, metadata) and data transfer (ANZLIC., 1998). There are two kinds of international standardization organizations both of them working in the field of standardizing the digital geographic information, the first is the ISO/TC211 which is a standard technical committee and its part of the International Organization for Standardization (ISO). the second organization is Open GIS Consortium (OGC), Its role is similar to the first organization. Consequently, both organization are responsible for creating international standards in geographic and geomatics data.

Access network (clearing house): The clearing house is an electronic service which provides access to metadata

and stored spatial data from the distributed data sources or from a number of servers connected with each other. The data can be accessed using the internet depending on processes of searching and querying in the metadata (Tosta, 1994; Shariff *et al.*, 2011). In this light, the clearinghouse is not centralized warehouse or a database of geo-information in fact a distributed network.

People (partnership): Partnership is the collaboration between stakeholders who are interested in the process of NSDI. In this consideration, relationships are built by an NSDI to raise the ability to share information to improve and maintain standard datasets in a spatial data community, guaranteeing data accuracy, maintaining access security. Such collaboration is mostly depends on the people relationship as good partnership greatly raise the performance of any NSDI Model (Rajabifard and Williamson, 2001).

Institutional frameworks (policy): This refers to the coordination of the many organizations involved in the development and maintenance of the NSDI to avoid the duplication of efforts. This advantage cannot be achieved when there is no coordination and without good arrangements among organizations (governance, data privacy and security, sharing of data, cost retrieval).

The status of development of Jordan National Spatial Data Infrastructure (JNSDI): This research study deals with improving the National Spatial Data Infrastructure (NSDI) in Jordan. The aim is to develop an improvement strategy towards the approach and the process of building the Jordan National Geo-Spatial data base (JNGIS). The objectives are to determine the most appropriate component for the Jordan Spatial Data Infrastructure (JSDI) and examine the requirements of standards and policies for data and technology.

In addition, the research analyses the strengths and weaknesses of the current map projection, i.e., Cassini and the Jordan Transverse Mercator (JTM) towards proposing a common system for the whole of Jordan. This questionnaire helps in determining the current status of NGIS in Jordan specifically in searching for data standards, policies and procedures currently being practised by government and private agencies. It is anticipated that the respondents will provide the much-needed remedies in making this research successful.

The stakeholder survey of JSDI: A stakeholder's survey was conducted through interviews and questionnaires surveys on 29 correspondents of the Jordan National Spatial Data Infrastructure (JNSDI). The stakeholders

Table 1: List of sampled JNSDI stakeholders

Stakeholder category	Number of sample
Academic sector	3
Public sector	99
Private sector	34
NGO	1
Total	137

come from a various classis like academia, NGOs, public sector and private sector. These data were accomplished with findings from the books, researches and displayed studies. We adopted a modified state of play assessment framework for this study which is one of the multi-view assessment frameworks by Grus *et al.* (2007). The number of stakeholders was appropriate and more representative as most stakeholder's organization has a GIS division and are registered under the JNSDI stakeholder.

The heads of the GIS departments interviewed are involved in JNDI activities. The study was conducted by doing interviews with GIS managers in the GIS agencies, distribute the questionnaire to search for details and best practices of JNSDI. The questionnaire has 93 questions and is categorised according to the SDI components. The questions were divided into 6 thematic groups:

- Organization information
- Institutional, organizational and strategic status
- Software, hardware and network
- GIS data and layers
- Human resources
- Funding

The questionnaire was distributed to organizations in the government, private and academic sectors as well as Non-Governmental Organizations (NGO) as shown in Table 1 to make the findings representative of all concerned stakeholders.

Summary of stakeholders institutional involvement in JSDI

The Royal Jordanian Geographic Centre (RJGC): The Royal Jordanian Geographic Centre (RJGC) was established in 1975 as a national agency responsible for all surveys related fields, including aerial, space surveying and map production of various types and scales. The RJGC mandate is to provide national surveying and mapping services. In this regard, it has established the Jordan geodetic network which consists of 2619 geodetic points and 14 doppler points. The RJGC cooperated with French experts in 1975 to apply a new projection for Jordanian maps called the Jordan Transverse Mercator (JTM). RJGC has more than 17 GIS users.

The Greater Amman Municipality (GAM): The population of greater Amman area in 2015 was estimated at 4,000,000 which represents 42% of the total population in Jordan. Amman governorate is divided into 27 districts and the geographic information directorate was initially launched in 2007 at greater amman municipality called the Geographic Information Department. GAM utilizes the Palestine grid coordinate system for their maps. The GIS department produces maps of various scales (1/2,500, 1/50,000). On the other hand, GAM does not have a legal mandate or an approved policy for SDI however it provides free public access to its web based GIS services. In addition, GAM has a quality assurance section and has obtained ISO9001 certification in 2008. Currently, GAM has more than 40 GIS users.

Department of Land and Survey (DLS): The Department of Lands and Survey (DLS) is the agency responsible for both national level land registration and cadastral survey. Prior to year 2000, DLS relied on paper based maps but its transformation to digital maps was completed through a 6 year project during the period of 1995-2000. Through scanning and vectorization all hard copy maps were transformed into digital form and rigorous quality assurance measures were implemented to account for the poor accuracy of paper based maps. The process includes scanning of the paper maps, examining the scanned images, digitizing the maps, auditing digitized images, transforming maps into vectors and lastly auditing vectors. DLS has more than 30 GIS user and similar to GAM, they provide free public access to its web based GIS services.

Yarmouk water company: Jordan Yarmouk Water Company (YWC) was established the North of Jordan in 2010. Yarmouk water company pursue for the provision of services in the wastewater and water sectors with good quality of services provided in northern government of Jordan, Irbid, Mafrq, Ajloun and Jerash. YWC implement initiatives aiming to help the state's treasury by reducing the burden on. provides its services in water and wastewater sectors for northern government of Jordan, the customer base around 301,000 for water and around 101,000 for wastewater. YWC has a GIS division with 20 GIS users.

Orange Jordan: Orange Jordan is a Jordanian public mobile telephone network company and is the operator of the mobile communications license granted by Jordan telecom. The company was first registered in 1999 with the aim to build a mobile communications network to cater for the Jordanian public. Orange started using GIS for asset

management and network planning purposes such as to keep track of communication towers location. GIS is also used to visualize internal and external data, such as demographics and market trends as well as to provide geographical data for the company to make smarter decisions on cables, towers and service coverage. It was known as mobile com until it was rebranded in 2007 under the orange brand with 19 users in the GIS division.

InfoGraph: Infograph is the leading GIS and remote sensing software supplier and GIS application provider in Jordan. It is the only authorized distributor of the Environmental Systems Research Institute (ESRI) in USA. ESRI is the world's leading GIS Software company and its flagship software product, ArcGIS is one the most commonly used software worldwide, including in Jordan. InfoGraph offices are located in Amman, Jordan and the company employs a staff of professional GIS specialists, ESRI-certified trainers and technical support and administrative personnel. Although, infograph was registered in Jordan in 1991, the experience of their senior staff in GIS and other computer-based applications ranges from 10 to more than 25 years. InfoGraph has significantly contributed to the overall use and effectiveness of GIS applications in addition to the GIS and remote sensing projects in Jordan. They also offer consultation and project development services in a number of GIS application areas. These include map production systems, cadastral and land information systems, public utility system (water, electricity and telecommunications), urban planning and electronic map publishing. Infograph has 17 GIS users.

RESULTS AND DISCUSSION

Organizational information: The sets of questionnaires were distributed to 29 organizations with a combined number of 289 GIS users as shown in Table 2. It is notable that Non-Governmental Organizations (NGO) represent only 1% because only one NGO was identified as a stakeholder and it has only two GIS users. Meanwhile, for the academic sector, there are two universities teaching geomatic technology having 6 academic persons dealing with GIS. In this regard all samples represent more than 90% from GIS stakeholder.

Institutional, organizational and strategic status: With regards to the question on the national initiative in Geographic Information System (GIS) in Jordan, the majority of the respondents 90% answered no. This indicates that there are no institutional collaboration frameworks.

Table 2: Organizational breakdown of stakeholder

Stakeholder category	Number of sample	Percentage	Number of organizations	GIS user
Academic sector	3	2	2	8
Public sector	99	72	20	202
Private sector	34	25	6	75
NGO	1	1	1	2
Total	137	100	29	287

Table 3: Organizations with GIS policy

Organizations	Types
Greater Amman Municipality (GAM)	Public sector
Jordan Investment Commission (JIC)	Public sector
Ministry of Information Technology and Communication (MoICT)	Public sector
Petra Development and Tourism Region Authority (PDTRA)	Public sector
Yarmouk water company (north governorate water company)	Private Sector

The organizations that confirmed having GIS national initiative are: Al-Balqaa University, Royal Society for the Conservation of Nature (RSCN), Ministry of Information Technology and Communication (MoICT) and the Royal Jordanian Geographic Centre (RJGC).

Organizations responsible for GIS: A majority of the respondents 93% answered no (Fig. 2) except three organizations, i.e., the Royal Society for the Conservation of Nature (RSCN) Ministry of Information Technology and Communication (MoICT) and Royal Jordanian Geographic Centre (RJGC).

Organization having policy regarding GIS: Only 5 out of the 29 organizations confirmed that they have a policy regarding GIS. This represents 17% of the surveyed organizations. The absence of a policy has negative consequences on SDI development (Fig. 3). The organizations which confirmed that they have a GIS policy are listed in Table 3.

Organizations having Spatial Data Infrastructure (SDI)/GI/GIS legal mandates: Only 7% of the organization surveyed confirmed that they have a legal SDI/GIS/GI mandate. This indicates that most of the organizations have no legal mandate. The only organization that confirmed to have a legal mandate SDI/GIS/GI was the Royal Jordanian Geographic Centre (RJGC). However, the RJGC did not provide law provision pertaining to neither such mandate nor its contents. Meanwhile, for the question, whether the country has a National Spatial Data Infrastructure Law/Policy/Directive, 100% of the respondents answered no. This indicates that there is no legislative framework for SDI.

The challenges facing the organization in using SDI/GI/GIS: In identifying the challenges that affect the

Table 4: Challenges facing the organizations in using SDI/GI/GIS

Challenges	Frequency
Funding limitation	77
Difference in projection system	70
Lack of GIS strategy	56
Lack of national interoperability	53
Lack of expert	37
Absence of regulation	19
Weak cooperation	17
Incompatibility of data (lack of standards)	16
The base data accuracy still not reliable (lands map from DLS)	15
Availability of digital data	15
Need of coordination (institutional arrangements)	13
Lack of awareness	10
Brain drain	10
Lack of advanced training	9
Need of legal aspects:	8
Misunderstanding of GIS benefits	8
Lack of national commitment	8
Lack of private sector partnerships	7
Long term benefits	7
Availability of metadata	6
Rapid development of the software, hardware and networking problem	5
The department far away from capital	2
Lack of hardware component	1

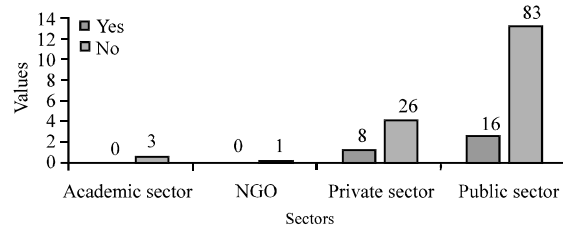


Fig. 2: Percentage organizations having policy regarding GIS

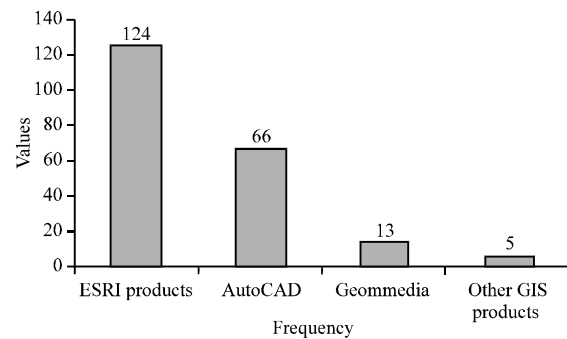


Fig. 3: GIS/CAD packages used by organizations

progress of using GIS systems in Jordan, a prioritized list of challenges was surveyed and the result is shown in Table 4. It can be observed that the main challenge is funding limitation, followed by the differences in projection system.

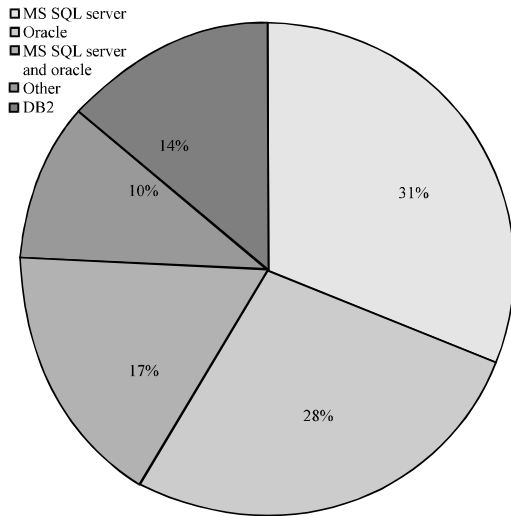


Fig. 4: Database platforms used in organizations to store GIS data

The subsequent question asked whether, it is better to create a new organization for National Geographic Information System (NGIS) or to transfer the responsibilities to one of existing organizations. A majority of the respondents 91% answered yes for the need to create a new organization for the (NGIS). Only 9% of the respondents answered NO and want to transfer the responsibility to one of the existing organizations. The organizations that wanted to transfer the responsibility to the existing organizations are:

- Royal Jordanian Geographic Centre (RJGC)
- Jordan investment commission (JIC)
- The Royal Society for the Conservation of Nature (RSCN)

Software, hardware and the network: On the questions pertaining to the GIS/CAD packages used in the organization, it was found that most organizations use ESRI products (Fig. 3) with the exception of 3 organizations that use different software are:

- Municipality of Irbid (Geomedia Software)
- Ministry of Municipal Affairs (MoMA Software)
- Department of Antiquities (Open Source Software)

With regards to the database platform (s) used by the agencies for GIS data storage, MS-SQL server is the most popular relational database engine. Oracle is also common while some organizations used MS Access, DB2 and other DBs (Fig. 4). Moreover, 100% of the organizations use the windows operating system to run GIS for integration of data.

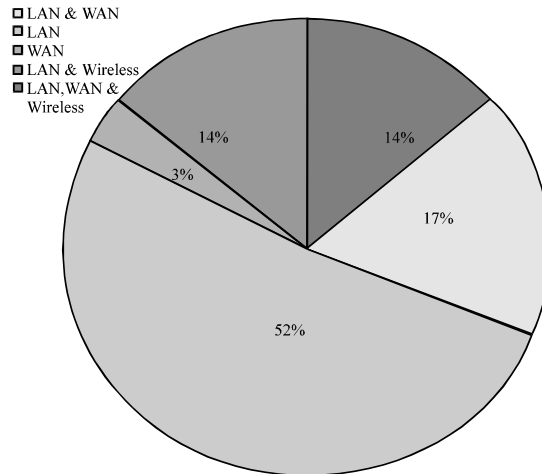


Fig. 5: Types of organizations networking; wireless (0%) other (0)%

In terms of organizational networking, Fig. 5 indicates that 52% of the organizations have a LAN network, 17% LAN and WAN, 4% LAN, WAN and wireless, 14% LAN and wireless and 3% WAN. In this light, the problem of technology deficiency does not exist. Apart from that, more than half of the organizations are connected to the National Information Technology Centre (NITC) ensuring a Secure Government Network (SGN) for sharing data.

The usage of the projection or coordinated system: The survey result in Table 5 show variations in the coordinate systems used thus, indicating no interoperability between organizations because differences in the spatial reference systems used.

GIS data and layers

Metadata: Metadata provides documentation of existing internal geospatial data resources within an organization. The survey has shown that 69% of the corresponding organization do not have metadata. Consequently, the lack of metadata will derail SDI development. Organizations that have their own metadata are Table 6.

Common data standards: It was also found that 76% of organizations do not have a common data standard while 24% have a common data standard. The lack of standards on spatial datasets, metadata, data transfer, software and hardware might cause negative consequences on SDI development and on the compatibility of data. In addition, 59% of organizations share digital geographic data with other organizations.

Table 5: Spatial reference systems used in Jordan (Cassini, Jordan transverse Mercator JTM or others)

Organizations	Projection
Al-Balqaa University	Both
Department of Lands and Survey (DLS)	Both
Jordan Investmentn Commission (JIC)	Both
Jordan Water Company (Miyahuna)	Both
Ministry of Agriculture (MoA)	Both
Ministry of Environment (moenv)	Both
Ministry of Municipal Affairs (MoMA)	Both
Al al-Bayt University	Both (JTM, CASS) and other
The Royal Society for the Conservation of Nature (RSCN)	Both (JTM, CASS) and other
Infograh Company	Both (JTM, CASS) and other
National Electricity Power Company (NEPCO)	Both (JTM, CASS) and other
Aqaba Water Company (AWC)	Both (JTM, CASS) and other
Department of antiquities	Both (JTM, CASS) and other
Department of Statistics (DoS)	Both (JTM, CASS) and other
Ministry of Public Works and Housing (MoPWH)	Both (JTM, CASS) and other
Natural Resources Authority (NRA)	Both (JTM, CASS) and other
Water Authority of Jordan (WAJ)	Both (JTM, CASS) and other
Orange telecommunications	CASS and other
Royal Jordanian Geographic Centre (RJGC)	JTM and other
Irbid District Electricity Company (IDECO)	Only CASSINI
Yarmouk water company (North Governorate water company)	Only CASSINI
Greater Amman Municipality (GAM)	Only CASSINI
Land Transport Regulatory Commission (LTRC)	Only CASSINI
Municipality of Irbid	Only CASSINI
Municipality of Zarqa	Only CASSINI
Aqaba Development Company (ADC)	Only JTM
Aqaba Special Economic Zone Authority (ASEZA)	Only JTM
Ministry of Information Technology and Communication (MoICT)	Only JTM
Petra Development and Tourism Region Authority (PDTRA)	Only JTM

This effort can push GIS projects to success when partnerships are strong enough and need more cooperation between organizations. Likewise, the majority of the respondents (86%) claimed that they did not use naming standards and these indicate the spatial data exchanged between different organizations will not be consistent in terms of meaning.

In this study as well, 79% of the organizations believe that the geo-spatial data produced by Department of Land and Survey can be used as base maps for Jordan. Only 6 organizations did not believe they could be used due to poor quality and they did not have building features, contours and land marks.

Human resources: With regards to the number of GIS users in the respective organizations (GAM) has the largest number (40 users), Department of Land and Survey (30 users). The distribution of GIS users is shown in Fig. 6.

Table 6: How the organizations rank themselves in GIS

Ranks	Organizations
High	Department of Lands and Survey (DLS)
	Yarmouk Water Company (North Governorate Water Company)
	Royal Jordanian Geographic Centre (RJGC)
	Greater Amman Municipality (GAM)
	Jordan Water Company (Miyahuna)
	Infograh company
	Aqaba Development Company (ADC)
	Orange Telecommunications
	Department of Antiquities
Medium	Irbid District Electricity Company (IDECO)
	Municipality of Irbid
	Department of Statistics (DoS)
	Natural Resources Authority (NRA)
	Aqaba Special Economic Zone Authority (ASEZA)
	Aqaba Water Company (AWC)
	Water Authority of Jordan (WAJ)
	Ministry of Municipal Affairs (MoMA)
	Ministry of Information Technology and Communication (MoICT)
	Petra Development and Tourism Region Authority (PDTRA)
Low	Municipality of Zarqa
	Ministry of environment (moenv)
	National Electricity Power Company (NEPCO)
	Ministry of Public Works and Housing (MoPWH)
	Al-Balqaa University
	Land Transport Regulatory Commission (LTRC)
	Al al-Bayt University
	The Royal Society for the Conservation of Nature (RSCN)
	Jordan Investmentn Commission (JIC)
	Ministry of Agriculture (MoA)

Table 7: Organizations that have there wn metadata

Organizations	Sectors
The Royal Society for the Conservation of Nature (RSCN)	NGO
Department of Antiquities	Public
Al-Balqaa University	Academic
Orange telecommunications	Private
Aqaba Development Company (ADC)	Private
Aqaba Water Company (AWC)	Public
Infograh company	Private
National Electricity Power Company (NEPCO)	Private
Greater Amman Municipality (GAM)	Public

In terms of retention strategies, the majority of organizations 52% have not done anything, 23% of the organizations offer better training, 13% offers financial incentives and 12% offers higher salaries. The 100% of respondents also agreed that there is a need for a national GIS professional association to coordinate with official bodies in the Kingdom.

In addition, the majority of the organizations 72% do not employ GIS experts while 28% of the respondents (8 organizations) employed outside experts. Lastly when enquired about the rating level of their organization in GIS, 31% of respondents (9 organizations) rated high level, 35% of respondents (10 organizations) rated medium while 31% of respondents (10 organizations) acknowledged to having low level of organization in GIS. The details of ranking level are provided in Table 7.

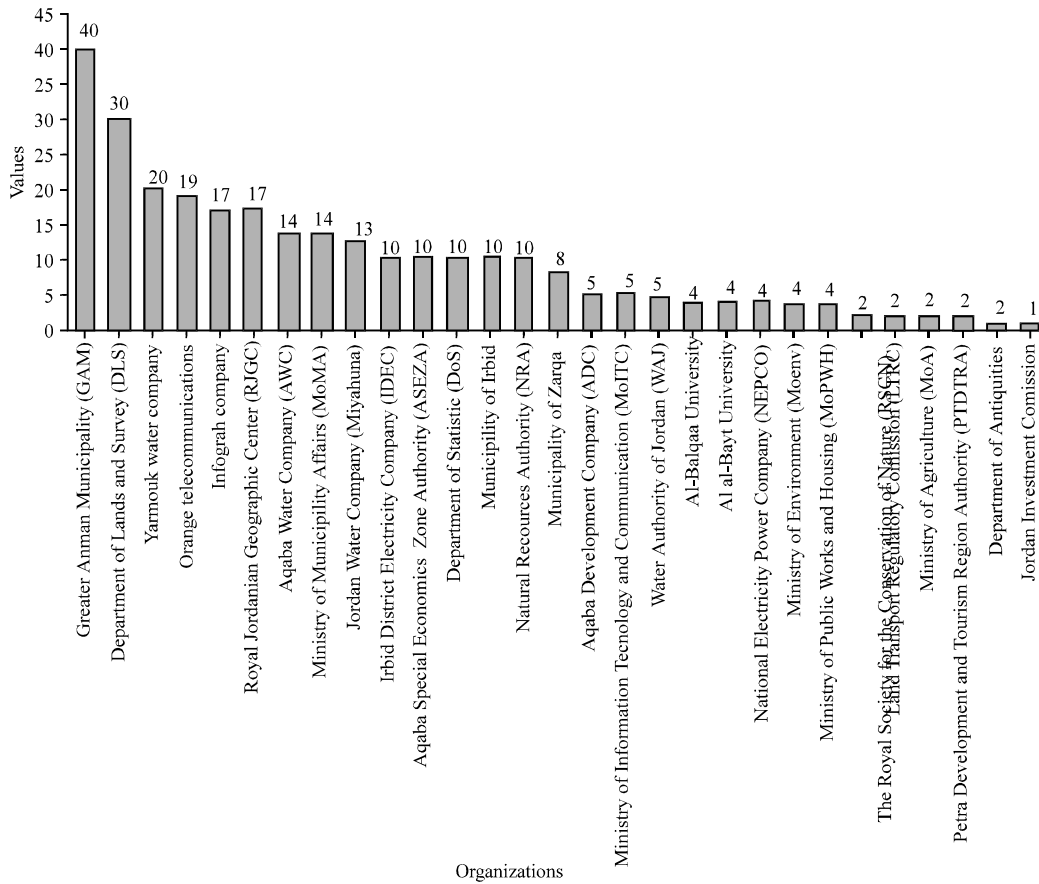


Fig. 6: Number of GIS users within the respective organizations

Funding: In the funding of the national GIS activities, 45% of respondents (13 organizations) received funds for GIS via. their annual budget (from the government) 28% of respondents (8 organizations) via. public and private funding, 10% of respondents (3 organizations) via. other methods of financing, 7% of respondents (2 organizations) received funds for GIS via. government solely and others, 7% of respondents (2 organizations) received funds via. private sector and 3% of respondents (1 organizations) received funds via. private sector and other methods of financing (Fig. 7).

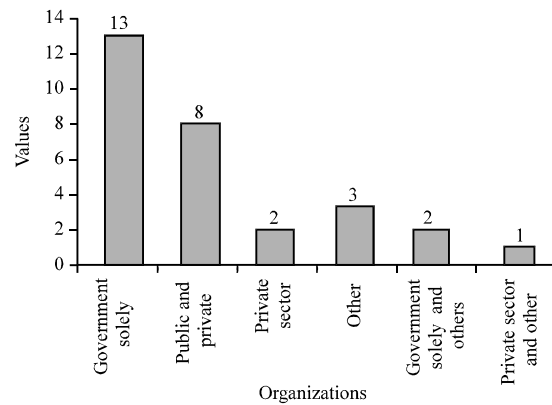


Fig. 7: The funding of the national GIS activities

Relating to the budget support in building the Jordan Spatial Data Infrastructure (JSDI) all of the 29 organizations have no funds to make available for building JSDI. Therefore, the benefit of JSDI should be strong enough to motivate any organizations in different levels to support such a project. When asked about government allocation in supporting JSDI, 90% of respondents (26 organizations) did not request a government allocation to support JSDI. This is because there is no national strategy for GIS in most surveyed organizations. Furthermore, some of the organizations resisted the JSDI project when there is no evidence of

short term benefits because JSDI projects requires some specific time in showing their benefits. Therefore, the lack of funding is a prominent challenge across all organizations. It hinders their ability to procure software, hardware, base maps, advanced training and enlarge their human resources retention challenge.

CONCLUSION

It can be comprehended that there is no institutional and organizational framework in Jordan that has the leadership responsibility regarding GIS. Most organizations prefer to create an independent national GIS council that has the authority, accountability and legal mandate to manage all GIS works throughout the kingdom of Jordan. Hence, coordination and cooperation must play a key role as there are no organizations that can establish nor improve the NSDI by itself. Only cooperation between, the private sector, public sector, academia, non-government organizations and individual users can provide a clear vision of the NSDI improvement.

Other significant problems that can hinder the success of JSDI include the variations in spatial reference systems, non-uniform standards in most GIS organizations to serve the national interests and no interoperability associated with the variations in feature semantics in spatial databases.

The issue of funding deficiency is also crucial across all organizations that put limits on the ability to procure software, hardware, base maps, advanced training and to deal with human resources retention challenge. Most developed countries with advanced SDI spend considerable amounts of funds and time to correct inconsistencies in data formats. Lastly, the lack of awareness on the importance of JSDI by the organizations prevents the government from supporting JSDI through distinctive budget allocation, although, Jordan has a rich human resource base and educational infrastructures.

REFERENCES

- ANZLIC., 1998. Spatial data infrastructure for Australia and New Zealand. ANZLIC, Barton, Australia.
- Akinyemi, F.O. and E. Uwayezu, 2011. An assessment of the current state of spatial data sharing in Rwanda. *Intl. J. Spatial Data Infrastruct. Res.*, 6: 365-387.
- Chan, T.O., M.E. Feeney, A. Rajabifard and I. Williamson, 2001. The dynamic nature of spatial data infrastructures: A method of descriptive classification. *Geomatica*, 55: 65-73.
- Coleman, D.J. and J.D. McLaughlin, 1998. Defining Global Geospatial Data Infrastructure (GGDI): Components, stakeholders and interfaces. *Geomatica Ottawa*, 52: 129-143.
- Crompvoets, J., A. Bregt, A. Rajabifard and I. Williamson, 2004. Assessing the worldwide developments of national spatial data clearinghouses. *Intl. J. Geog. Inf. Sci.*, 18: 665-689.
- FGDC., 1997. Framework Introduction and Guide. Federal Geographic Data Committee, Reston, Virginia, USA., Pages: 106.
- Groot, R. and J. McLaughlin, 2000. *Geospatial Data Infrastructure: Concepts, Cases and Good Practice*. Oxford University Press (OUP), Oxford, England, UK., ISBN: 9780198233817, Pages: 286.
- Grus, L., 2010. Assessing spatial data infrastructures. Master Thesis, Graduate School of Production Ecology & Resources Conservation, Wageningen University and Research Centre, Wageningen, Netherlands.
- Grus, L., J.W.H.C. Crompvoets and A.K. Bregt, 2007. Multi-view SDI assessment framework. *Intl. J. Spatial Data Infrastruct. Res.*, 2: 33-53.
- Hecimovic, Z., S. Marasovic and J. Crompvoets, 2014. Development of local spatial data infrastructure in Croatia. *J. Spatial Sci.*, 59: 221-234.
- Loenen, B.V. and B.C. Kok, 2017. *Spatial Data Infrastructure and Policy Development in Europe and the United States*. DUP Science Publisher, Delft, Netherlands,.
- Man, W.H.E.D., 2006. Understanding SDI: Complexity and institutionalization. *Intl. J. Geog. Inf. Sci.*, 20: 329-343.
- Masser, I., 2002. A comparative analysis of NSDI's in Australia, Canada and the United States. Master Thesis, GINIE Geographic Information Network in Europe IST-2000-29493, University of Sheffield, Sheffield, England, UK.
- OMB., 2002. To the heads of executive departments and establishments. Office of Management and Budget, Washington, D.C., USA. https://www.whitehouse.gov/omb/circulars_a016_rev/#2
- Okuku, J., A. Bregt and L. Grus, 2014. Assessing the development of Kenya National Spatial Data Infrastructure (KNSDI). *South Afr. J. Geomatics*, 3: 95-112.
- Rajabifard, A. and I.P. Williamson, 2001. Spatial data infrastructures: Concept, SDI hierarchy and future directions. *Proceedings of Conference on GEOMATICS'80, April 29-May 3, 2001, University of Melbourne, Tehran, Iran*, pp: 28-37.
- Rajabifard, A. and I.P. Williamson, 2003. Asia-pacific region and SDI activities. *J. GIS. Dev.*, 7: 30-34.

- Rajabifard, A., I.P. Williamson and M.E.F. Feeney, 2003. Developing Spatial Data Infrastructures: From Concept to Reality. Taylor & Francis, Milton Park, England, UK.,.
- Shariff, A.R.M., A.H. Hamzah, A.R. Mahmud, N.M.Z.N. Yusof and H.M. Ali, 2011. Spatial data infrastructure for Malaysian land administration. Intl. J. Spatial Data Infrastruct. Res., 6: 389-409.
- Tosta, N., 1994. Continuing evolution of the national spatial data infrastructure. Proceedings of the Conference on GIS-LIS'94, October 25-27, 1994, Phoenix Publishing House, Arizona, USA., pp: 769-777.
- Tosta, N., 1995. Data policies and the national spatial data infrastructure: Law and information policy for spatial database. Federal Geographic Data Committee, Reston, Virginia, USA.

Perception of Simulation as an Educational Technique in Health, Education and Applied Sciences by Health Education Students in Universities

Uchenna Cosmas Ugwu, Lawretta Eyuche Ozoemena, Benedeth Anulika Obayi,
Adaustin Ifeoma Okeke and David C.J. Ayalogu
Department of Human Kinetics and Health Education, Faculty of Education,
University of Nigeria, Nsukka, Enugu State, Nigeria

Abstract: Simulation in education has become an increasingly popular tool in teaching and producing highly qualified professionals in health, education and applied sciences. The study aimed to determine the perception of simulation as an educational technique in health, education and applied sciences by Health Education Students (HES) in universities. This study is important as the understanding of the student's perception would help to establish their readiness to engage in simulation. The descriptive research design was utilized. The sample for the study consisted of 320 HES selected from 72 Universities in Nigeria. The 72 universities constitute 72 clusters (cluster sampling techniques). Simple random technique by balloting was used to select 20 universities out of the 72 universities. A total of 16 HES were selected from each of the 20 universities (clusters) using simple random technique. Overall, a total of 320 HES were used for the study. The instrument used for data collection was questionnaire. The descriptive statistics was used for data analysis. The result showed that the perception of simulation as an educational technique in health, education and applied sciences by HES in universities was positive. The study also indicated significant differences between variables of the respondents regarding the phenomenon. Based on the findings, it was concluded that reinforcing and sustaining simulation in health, education and applied sciences during the student's early phases of academic programme and incorporating same in later phases can promote knowledge development, skill acquisition and self-confidence.

Key words: Perception, simulation, educational technique, variables, HES, applied sciences

INTRODUCTION

Globally, the importance of simulation in health, education and applied sciences cannot be overemphasized. This is because, the popularity of simulation as an educational technique in all spheres of teaching and learning has noticeably grown (Tawalbeh and Tubaishat, 2014). Further research indicates that simulation in education has been in application, since, the time of World War 11 (Sanford, 2010). The above research presents facts that simulation is crucial in education, health and applied sciences. As a teaching technique (Lateef, 2010), simulation uses an artificial environment, trying to recreate a real situation in order to allow the students to practice, learn, evaluate, test or develop an understanding of different human actions (Anonymous, 2010). In health education and other science-related disciplines where human anatomy and physiology as well as disease prevention, health promotion and maintenance form the integral part of the

academic program, the use of simulation has been part of teaching, skill acquisition and professional development. Simulation is considered as an attempt to imitate an ideal situation, so that, at a later moment in a real context, there is a better understanding and management of the situation. As a teaching and learning strategy, it consists of a set of activities that seek to replicate real context (Teixeira and Felix, 2011). The successful application of simulation in health, education and applied sciences has a lot of outstanding benefits to the professional development, skill acquisition and knowledge empowerment of the prospective professionals. These benefits are attainable when the student's perception to simulation is considered legitimate, authentic and realistic.

Leigh (2008) indicated that simulation is very important for teamwork, realism and active learning. Other researchers such as Hawkins *et al.* (2008) noted that simulation is more effective in acquiring knowledge and skills through experience and drawing on practical

problem solving and development of technical skills in a safe and controlled environment. As an educational tool and technique, Lateef (2010) supported that simulation can be used to create guided experience needed in acquiring necessary skills and patient's safety. There is also evidence that simulation is mostly effective in the development of psychomotor skills, critical thinking and reflection on practices (Sanford, 2010; Martins *et al.*, 2012), understanding of disease complex processes and the ability to make positive decisions on health issues. The above scenario is suggestive, highlighting the crucial benefits in applying simulation as an educational technique in health, education and applied sciences.

To be specific, health education emphasizes the acquisition of skills and quality education that will improve knowledge, encourage positive attitude and enhance desirable practices in individual. In this context, simulation was considered very essential techniques and was found to have positive influence on various educational outcomes such as self-efficacy, knowledge acquisition and performance (Kaakinen and Arwood, 2009; Norman, 2012). As a potentially powerful teaching approach in health and applied sciences, simulation engages students and requires them to use critical thinking and reasoning and provides an opportunity for reflective learning and integration of knowledge (Peisachovich *et al.*, 2016). This significant role might have prompted the categorization of simulation into three different levels via. high, moderate and low fidelity simulation (World Health Organization, 2011). In education, health and applied sciences, high-fidelity simulation is increasingly popular (Lasater, 2005; Sanford, 2010). This could be attributed to the fact that students have more active role in the learning process and are more motivated to learn (Leigh, 2008), thus, paving ways for gaining more experiences required for their professional development in general and effective service delivery (Dillard *et al.*, 2009) in particular.

Nonetheless, the World Health Organization (2011) stipulated that simulation is more effective in education because students respond and learn best when the learning environment is safe, supportive, challenging and engaging. Studies further emphasized the increased interest in simulation as an approach to teach important psychomotor and critical thinking skills to students (Leigh, 2008; Norman, 2012). These could be the reason why educational curricula adapted to the changing needs of learners and remain updated with the most current educational strategies (Norman, 2012). Till date, a growing body of researches have focused on simulation from different perspectives (Leonard *et al.*, 2010; Wotton *et al.*, 2010; Ogilvie *et al.*, 2011; Sigalet *et al.*, 2012; Tosterud *et al.*, 2013; Baptista *et al.*, 2016; Gharaibeh *et al.*, 2017).

Despite the overwhelming evidence on the positive impacts and effectiveness of simulation on various aspects of teaching and learning (Cant and Cooper, 2010), especially in health, education and applied sciences, simulation can still be perceived as stressful or frustrating by the learners or prospective young professionals such as the health education students in universities. It could also be perceived as not matching the actual practice, knowledge and skill acquisition, stressful environment and uncomfortable way to learn. Such negative indications may impede the process of learning and influence the student's engagement and the fidelity on simulation particularly in applied sciences. Contextually, perception is the view and understanding of idea, concept, action or behaviour. There is evidence that poor planning, organizing and executing simulation can cause frustration to the students and failure to meet the learning objectives (Peisachovich *et al.*, 2016). This makes the present study more unique and crucial for the purpose of empirical generalizations. To our knowledge, no study from Nigeria has determined the perception of simulation as an educational technique in health, education and applied sciences by the students, thus, making the current study the first descriptive survey of this kind. This study aims to find out the perception of simulation as an educational technique in health, education and applied sciences by health education students in universities. The HES was chosen for the present study due to their training in the diverse fields of health, applied sciences, education, disease prevention, patient's safety, health promotion, counseling as well as human anatomy and physiology. It is my expectation that the outcome of this study would help to establish the readiness of the HES and other students to engage in simulation and further helps in preparing effective teaching plans for the prospective health educators in particular and other professional in applied sciences in general. In addition, it would provide a deep insight into the HES readiness to perform simulation. This is because perceptions exert a direct and dynamic influence on student's responses, define how they see situation and how they behave toward simulation.

Objective: This study aimed to determine the perception of simulation as an educational technique in health, education and applied sciences by HES in universities. Specifically, the study provided answers to the following research questions and null hypotheses:

- What is the perception of simulation as an educational technique in health, education and applied sciences by HES in universities?

- What is the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on gender?
- What is the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on student year level?
- What is the perception of simulation as an educational technique in health, education and applied sciences by HES in universities according to prior simulation experience?
- What is the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on the university type?
- What is the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on prior experience in teamwork?
- What is the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on type of teamwork experience?

Hypothesis: The following null hypothesis guided the study and were tested at 0.05 level of significance. There is no statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on gender. There is no statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on student year level. There is no statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities according to prior simulation experience. There is no statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on the university type. There is no statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on prior experience in teamwork. There is no statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on type of teamwork experience.

MATERIALS AND METHODS

The study falls within the paradigm of the descriptive research method. The population of the study comprised

Table 1: Demographic characteristics of the respondents (N = 320)

Variables	Frequency	Percentage
Gender		
Male	150	47
Female	170	53
Student year level		
First	76	24
Second	84	26
Third	70	22
Fourth	90	28
Prior simulation experience		
Yes	185	58
No	135	42
University type		
Public	170	53
Private	150	47
Prior experience in teamwork		
Yes	210	66
No	110	34
Type of teamwork experience		
Workshop	55	17
Seminar	60	18
Lecture	120	38
Conference	85	27

of all the health education students currently on academic program in Universities of Nigeria. The sample for the study consisted of 320 HES selected from 72 universities. The 72 universities constitute 72 clusters (cluster sampling techniques). Simple random technique by balloting was used to select 20 universities out of the 72 universities. A total of 16 HES were selected from each of the 20 universities (clusters) using simple random technique. Overall, a total of 320 HES were used for the study (Table 1). The eligibility criteria to participate in this study included students actively enrolled in the regular undergraduate health education program in universities of Nigeria. The students who did not meet the above criteria such as the sandwich students were excluded from the study.

Instrument for data collection: The instrument for the data collection was self-administered questionnaire designed for the study. The questionnaire “Perception of Simulation Questionnaire (PSQ)” which can be completed in approximately 15 min was composed of two parts-A and B. Part A elicited information on demographic characteristics of the respondents while part B generated data on perception of simulation as an educational technique in health, education and applied sciences. These two parts (A and B) were bundled into one study package for the convenience of the respondents. The PSQ was a Likert-type scale with response options that range from 5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree and 1 = strongly disagree. Reliability testing indicated that the questionnaire had a strong internal consistency (Cronbach’s $\alpha = .87$). The questions were designed allowing responses from the respondents without any bias.

Data collection procedure: In addition to the health education background of the five Research Assistants (RAs) who helped in the distribution of the questionnaire, they were also trained in data collection procedure. The RAs approached the students towards the end of their lectures and from different year levels (i.e., 1st-4th). The RAs provided a brief description of the study and invited eligible students to participate. Informed consent of the HES was duly obtained before the questionnaire was administered. Only those who consented to participate in the study were included. The HES who agreed to participate received a questionnaire package which included a cover letter containing a summary of the study, the participant's right and the researcher's contact information. In the cover letter, potential participants were requested to complete the questionnaire and return immediately to the RAS. The ethical approval for the study was obtained from the Faculty of Education Research Grants Committee, University of Nigeria, Nsukka. The study was free from all forms of physical, psychological, social and economic harm or risk because the data collection process primarily relied on a descriptive non-invasive questionnaire.

Statistical techniques: Statistical Package for Social Science (SPSS) Version 21 was used to analyze the data. Descriptive statistics involving Means (M), Standard Deviation (SD), Frequency (F) and Percentages (%) were used to describe the sample characteristics and answer the research questions. The cut-off point for the weighted mean was 3.00 accrued from the 5-point response options, hence, any item that weighed 3.00 and above signifies positive perception while any item <3.00 implies negative perception of simulation as an educational technique in health, education and applied sciences by HES in universities. One-way Analysis of Variance (ANOVA) and t-test statistics were used to test the postulated null hypotheses of no significant differences at 0.05 level of significance. The hypothesis ought to be accepted when the p-value is below 0.05 while when the p-value is 0.05 and above, the hypothesis was rejected.

RESULTS AND DISCUSSION

A total of 320 HES were studied. Table 1 presents the demographic characteristics of the respondents. Of all, 47% were males while 53% were females. Based on student year level, 24% were in their first year level, 26% second year level, 22% third year level and 28% fourth year level. Approximately 58% had prior simulation experience while only 42% never had simulation experience. About 53% were in public (i.e., government-owned) universities

Table 2: Presenting the perception of simulation as an educational technique in health, education and applied sciences by HES in universities (N = 320)

Perception	Mean	SD	Remark
Average mean value	3.16	202	Positive

while only 47% indicated studying in private universities. Over 66% had prior experience in teamwork while only 34% never had prior experience in teamwork. Only 17 and 18% of the respondents indicated workshop and seminar as type of teamwork experience while 38 and 27% indicated lecture and conference as the type of teamwork experiences, respectively (Table 1).

Available data in Table 2, indicated that the average mean value (M = 3.16; SD = 202) is above the cut-off point of 3.0, implying that the perception of simulation as an educational technique in health, education and applied sciences by HES in universities is positive. This finding is quite encouraging and signifies that HES consider simulation as an encouraging tool in health, education and applied sciences (Table 2).

The study indicated that the average mean score of female HES (M = 3.56>3.00) is above the cut-off point implying positive perception while their male counterparts (M = 2.77<3.00) is below the cut-off point indicating negative perception of simulation as an educational technique in health, education and applied sciences in universities (Table 3). Table 3 further showed that H₀₁ is rejected since the p-value of 0.14 is >0.05. That is to say that, there is significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in university based on gender.

Available data in Table 3 showed the average mean values of HES in first year (M = 2.41<3.00), second year (M = 3.37>3.00), 3rd year (M = 1.29<3.00) and fourth year (M = 3.07>3.00), respectively. As contained in Table 3, the average mean values of HES in second and fourth year were above the cut-off point, indicating positive perception while the average mean values of HES in first and third year levels were below the cut-off point implying negative perception of simulation as an educational technique in health, education and applied sciences in universities. Table 3 further showed that H₀₂ is rejected since the p-value of 0.61 is above 0.05. That is to say that, there is statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on student year level.

The result indicated that HES who had prior simulation experience had average mean value (M = 3.52>3.00) indicating positive while those who never had prior simulation experience had average mean value

Table 3: Presenting demographic differences on perception of simulation as an educational technique in health, education and applied sciences by HES in universities and significant differences between variables (N = 320)

Variables	F-values	Mean	SD	Status	t-Cal	p-values	Remark	Decision
Gender					0.24	0.14	*	Rejected
Male	150	2.77	0.111	Negative				
Female	170	3.56	0.201	Positive				
Student year level					0.32	0.61	*	Rejected
First	76	2.41	0.001	Negative				
Second	84	3.37	0.210	Positive				
Third	70	1.29	0.101	Negative				
Fourth	90	3.07	0.011	Positive				
Prior simulation experience					0.05	0.03	**	Accepted
Yes	185	3.52	0.051	Positive				
No	135	2.46	0.021	Negative				
University type					0.17	0.04	**	Accepted
Public	170	3.52	0.007	Positive				
Private	150	2.49	0.311	Negative				
Prior experience in teamwork					0.20	0.81	*	Rejected
Yes	210	3.61	0.091	Positive				
No	110	2.32	0.181	Negative				
Type of teamwork experience					0.13	0.62	*	Rejected
Workshop	55	3.18	0.010	Positive				
Seminar	60	2.14	0.061	Negative				
Lecture	120	3.31	0.141	Positive				
Conference	85	3.24	0.081	Positive				

*Significant at 0.05 level, **Not Significant at 0.05 level; F = Frequency; SD = Standard Deviation

(2.46<3.00) signifying negative perception of simulation as an educational technique in health, education and applied sciences in universities (Table 3). Table 3 also showed that H_{03} is accepted since the p-value of $0.03 < 0.05$. That is to say that, statistically, there is no significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on prior simulation experience.

Available data indicated that the average mean value of HES in public universities ($M = 3.52 > 3.00$) is above the cut-off point implying positive while those in private universities ($M = 2.49 < 3.00$) is below signifying negative perception of simulation as an educational technique in health, education and applied sciences in universities based on university type (Table 3). Table 3 further indicated that H_{04} is accepted since the P value of $0.04 < 0.05$. That is to say that, statistically, there is no significant difference on the perception of simulation as an educational technique in health, education and applied sciences in universities based on university type.

Available data indicated that the HES who have prior experience in teamwork had average mean value ($M = 3.61 > 3.00$) indicating positive perception while those who never had prior experience in teamwork had average mean value ($M = 2.32 < 3.00$) signifying negative perception of simulation as an educational technique in health, education and applied sciences in universities (Table 3). Table 3 also showed that H_{05} is rejected since the p-value of $0.81 > 0.05$. That is to say that, statistically, there is significant difference on the perception of

simulation as an educational technique in health, education and applied sciences by HES in universities based on prior experience in teamwork.

The study indicated that the average mean values of HES who had experience in workshop ($M = 3.18 > 3.00$), lecture ($M = 3.31 > 3.00$) and conference ($M = 3.24 > 3.00$) are above the cut-off point implying positive perception while those who had experience in seminar ($M = 2.14 < 3.00$) is below the cut-off point indicating negative perception of simulation as an educational technique in health, education and applied sciences in universities (Table 3). Table 3 further showed that H_{06} is rejected, since, the p-value of 0.62 is above 0.05. That is to say that, there is statistically significant difference on the perception of simulation as an educational technique in health, education and applied sciences by HES in universities based on type of teamwork experience.

This quantitative study in its descriptive nature has established that the perception of simulation as an educational technique in health, education and applied sciences by HES in universities is positive. This encouraging result reflects the fact that HES hold strong passion to their educational potentials and and skill acquisitions. The finding is revealing because it shows the readiness of HES to engage in simulation. This finding may be attributed to the fact that the students have a better understanding of the tedious nature of the expected service delivery in the areas of patient's care safety, disease prevention, health education, promotion, counselling, maintenance as well as anatomy and physiology of human system. Aligning with other studies, the present finding can be linked with Tawalbeh and

Tubaishat (2014) who indicated that simulation has gained increased popularity in educational system. This assertion was not surprising since other scholar had indicated that appropriate application of simulation encourages teamwork, realism and active learning (Leigh, 2008). A convincing similarity existed between this finding and others (Gharaibeh *et al.*, 2017). The implication of the finding is that with positive perception of simulation as a tool, remarkable gains are inevitable such as high-fidelity simulation (Lasater, 2007; Sanford, 2010), quality knowledge (Peisachovich *et al.*, 2016), high self-efficacy, knowledge acquisition and clinical performance (Kaakinen and Arwood, 2009; Norman, 2012). The present study considers the above attributes as indices of positive perception to simulation in health, education and applied sciences that are very essential for the student's professional development, knowledge empowerment, skill acquisition and effective service delivery.

This study showed that differences exist within variables of the respondents in relation to perception of simulation as an educational technique in health, education and applied sciences. The result indicated that female HES, those in second year level, fourth year level, HES who had prior simulation experience, HES in public universities, those who had prior experience in teamwork and HES who had teamwork experience in workshop, lecture and conference had positive perception on simulation as an educational technique in health, education and applied sciences in universities. Available data also showed that the male HES, those in first and third year level, HES who had no prior simulation experience, those in private universities, HES who had no prior experience in teamwork and those who had teamwork experience in seminar had negative perception of simulation as an educational technique. These expected findings aligned with the findings of other researchers that perception of individual on a given phenomenon can be affected by variables (Leonard *et al.*, 2010; Sigalet *et al.*, 2012; Tosterud *et al.*, 2013; Baptista *et al.*, 2016; Gharaibeh *et al.*, 2017). Statistically, the study indicated that significant differences exist on the variables of gender, student year level, prior experience in teamwork and type of teamwork experience while none exists on prior simulation experience and university type.

The strength and weakness of the present study lies on the chosen methodology. This study was restricted to selected variables such as gender, student year level, prior simulation experience, university type, prior experience in teamwork and type of teamwork experience, of the respondents. Studies using other variables such as age, marital status, academic qualifications and lifestyle should be conducted for broader research generalization.

The present study was delimited to health education students in universities. It is important to expand the scope by conducting similar study in other tertiary institutions such as colleges of education where health education and applied sciences are offered as specialized courses. This study was descriptive research design adopting descriptive statistics. A more detailed qualitative in-depth survey should be conducted to provide a deeper understanding and finding regarding perception of simulation in health, education and applied sciences by students.

CONCLUSION

From the literature reviewed and data analysis, this study has convincingly established that the perception of simulation as an educational technique in health, education and applied sciences by HES in universities is positive. Nonetheless, this is a reliable indication that the students are eager to engage in simulation. Secondly, it is also justifiable evidence that students learn more and understand facts and practice when the lesson plans are adjusted for simulation activities. Efforts should be geared towards advancing measures that would encourage and sustain simulation in teaching and learning. This constructive initiative would be attainable by encouraging social settings such as workshops, seminars, lectures and conferences that would bring the students together for effective academic interactions and exchange of valuable knowledge and ideas. In addition, reinforcing and sustaining simulation in health, education and applied sciences during the student's early phases of academic programme and incorporating same in later phases can promote knowledge development, skill acquisition and self-confidence.

ACKNOWLEDGEMENT

The researcher acknowledges the unconditional assistance from the five research assistants who helped in the distribution of the questionnaire.

REFERENCES

- Anonymous, 2010. An interactive global simulation community. National League for Nursing, Washington, D.C., USA.
- Baptista, R., F. Pereira and J. Martins, 2016. Perception of nursing students on high-fidelity practices: A phenomenological study. *J. Nurs. Educ. Pract.*, 6: 10-21.

- Cant, R.P. and S.J. Cooper, 2010. Simulation-based learning in nurse education: Systematic review. *J. Adv. Nurs.*, 66: 3-15.
- Dillard, N., S. Sideras, M. Ryan, K.H. Carlton and K. Lasater *et al.*, 2009. A collaborative project to apply and evaluate the clinical judgment model through simulation. *Nurs. Educ. Perspect.*, 30: 99-104.
- Gharaibeh, B., I. Hweidi and A. Al-Smadi, 2017. Attitudes and perception of baccalaureate nursing students toward educational simulation. *Cogent Educ.*, 4: 1-14.
- Hawkins, K., M. Todd and J. Manz, 2008. A unique simulation teaching method. *J. Nurs. Educ.*, 47: 524-527.
- Kaakinen, J. and E. Arwood, 2009. Systematic review of nursing simulation literature for use of learning theory. *Intl. J. Nurs. Educ. Scholarship*, Vol. 6,
- Lasater, K., 2005. High-fidelity simulation and the development of clinical judgment: Student's experiences. *J. Nursing Educ.*, 46: 269-276.
- Lateef, F., 2010. Simulation-based learning: Just like the real thing. *J. Emergencies, Trauma Shock*, 3: 348-352.
- Leigh, G.T., 2008. High-fidelity patient simulation and nursing students self-efficacy: A review of the literature. *Intl. J. Nurs. Educ. Scholarship*, 5: 1-17.
- Leonard, B., E.L. Shuhaibar and R. Chen, 2010. Nursing student perceptions of intraprofessional team education using high-fidelity simulation. *J. Nurs. Educ.*, 49: 628-631.
- Martins, J.C.A., A. Mazzo, R.C.N. Baptista, V.R.D. Coutinho and S.D. Godoy *et al.*, 2012. The simulated clinical experience in nursing education: A historical review. *Acta Paulista Nurs.*, 25: 619-625.
- Norman, J., 2012. Systematic review of the literature on simulation in nursing education. *J. Assoc. Black Nurs. Faculty*, 23: 24-28.
- Ogilvie, S., B. Cragg and B. Foulds, 2011. Perceptions of nursing students on the process and outcomes of a simulation experience. *Nurse Educ.*, 36: 56-58.
- Peisachovich, E.H., R. Gal and S. Johnson, 2016. Experiences of undergraduate nursing students of standardized patient methodology in their transition to nursing practice in Ontario Canada. *J. Nurs. Educ. Pract.*, 7: 1-9.
- Sanford, P.G., 2010. Simulation in nursing education: A review of the research. *Qual. Rep.*, 15: 1006-1011.
- Sigalet, E., T. Donnon and V. Grant, 2012. Undergraduate students' perceptions of and attitudes toward a simulation-based interprofessional curriculum: The KidSIM attitudes questionnaire. *Simul. Healthcare*, 7: 353-358.
- Tawalbeh, L.I. and A. Tubaishat, 2014. Effect of simulation on knowledge of advanced cardiac life support, knowledge retention, and confidence of nursing students in Jordan. *J. Nursing Educ.*, 53: 38-44.
- Teixeira, I.N.D. and J.V.C. Felix, 2011. Simulation as a teaching strategy in nursing education: Literature review. *Interface Commun. Health Educ.*, 15: 1173-1184.
- Tosterud, R., B. Hedelin and M.L. Hall-Lord, 2013. Nursing students perceptions of high-and low-fidelity simulation used as learning methods. *Nurs. Educ. Pract.*, 13: 262-270.
- World Health Organization, 2011. Patient Safety Curriculum Guide: Multi-Professional Edition. World Health Organization, Geneva, Switzerland, ISBN:978-92-4-150195-8, Pages: 270.
- Wotton, K., J. Davis, D. Button and M. Kelton, 2010. Third-year undergraduate nursing students' perceptions of high-fidelity simulation. *J. Nurs. Educ.*, 49: 632-639.

A Seven Tier Architecture of Cloud Database Management System

M. Husain, Yazed. ALsaawy and Ali Tufail
Islamic University, Madinah, Saudi Arabia

Abstract: The cloud computing is a leading driving movement for information technology where the major organizations are using more than one cloud provider, hence, computing will certainly become more distributed than it currently is this poses management, architectural and performance challenges. In this study, we propose a unique seven tier architecture of cloud database management system. The first tier elaborates the physical layer and specifies how to physically store the database files on the secondary storage devices. The second tier, physical middleware layer, provides seamless connectivity regardless of the underlying platform. The third tier, conceptual lower layer, supports logic design, database connectivity, data abstraction and partitioning. The fourth tier, conceptual layer, deals with query processing, searching and security of various databases of the system for the internal processing on data. The fifth tier, conceptual middleware layer, provides seamless connectivity amongst various databases. The sixth tier, external layer, focuses on user's prospective where the same database can be viewed by various user's simultaneously. The seventh tier, external upper layer, provides interface interoperability, interface design and security which are the major issues for cloud databases. We argue that our proposed distinct seven tier architecture would not only help to design a distributed database architecture in an efficient manner but would also optimize the querying/searching in a cloud environment.

Key words: Cloud computing, schema design, hybrid cloud, virtualization, DBaaS, cloud environment

INTRODUCTION

In the traditional database approach, data is usually stored on a single server that serves as the source of all the querying and searching. The data is generally accessed directly or through local area network. However, with the advent of cloud computing the whole new concept of distributed database is now being materialized into reality (Kossmann *et al.*, 2010). On one side, cloud computing gives an option to companies to outsource their database on the other side it brings whole new challenges related to management of database, security, query efficiency, consistency, etc. Nevertheless, the concept of cloud computing is flourishing and major player of industry, like Google, Microsoft, Salesforce.com, Rackspace Amazon EC2, etc. are now competing to provide best cloud platform and related services (Alam *et al.*, 2013; Dinh *et al.*, 2013). Generally, the cloud database management system is deployed in the following three ways (Curino *et al.*, 2011; Singh and Sandhu, 2011).

Virtual machine image: In this approach, the database management system can run only on Virtual Machine (VM) instances and these instances are sold out by the

cloud provider and they are responsible for the infrastructure including uploading or purchasing of the DBMS.

Database as a Service (DBaaS): In this approach, the cloud provider is overall responsible for maintaining the database management system. Mainly it is more attractive due to hardware and software costs.

Managed hosting: This approach provides the facility to install, maintain and manage the overall database accomplishment.

Before deploying cloud database management system the organizations should keep in mind about performance, budget, data governance and staffing which directly affect the organizations.

The cloud computing is a leading driving movement for information technology where the major organizations are using more than one cloud provider, hence, computing will certainly become more distributed than it currently is this poses management, architectural and performance challenges (Curino *et al.*, 2011; Bloor, 2011).

In this study, we propose a seven tier architecture of cloud database management system. These tiers are physical layer, physical middleware layer, conceptual

lower layer, conceptual layer, conceptual middleware layer, external layer and external upper layer. Our architecture, handles the complexity of the distributed database in cloud environment. In order to enhance the efficiency and the optimization of database querying process we have suggested to divide the overall process into seven distinct tiers/layers. Where each layer is responsible to deal with its defined tasks and functionalities while complementing the other layer. We argue that by following our suggested architecture the overall performance of the distributed database will get optimized, efficient and effective.

Literature review: Dimovski (2013) argues that the cloud database has the potential to provide an optimal solution for web and mobile applications where DBaaS and PaaS provides an opportunity to the organization to develop products by not wasting resources. However, he suggests that the cloud is not yet ready and reliable to store highly sensitive enterprise application data but it can be used for testing and development purposes.

Another important aspect of cloud database is the query processing. Stonebraker *et al.* (2005) present query processing technique to enhance the performance of the compressed data. They proposed column oriented algorithm and performed a comparison with the traditional algorithms. Their algorithm created a decision-tree to help the database designers in the decision of compressing a particular column. However, they do not discuss the decompression cost of the various compression algorithms. Also, by Kumar and Bansal (2012) researcher discuss the optimization of database performance in context to lightweight compression schemes. Although, their proposed scheme is well suited for the traditional database, however, it cannot be applied in the cloud database architecture.

Alam *et al.* (2013) have proposed a framework for 5-layered architecture in cloud database management system. Researcher describe functionality of each layer, however, they have overlooked a number of functionalities related to the distributed and the cloud database environment.

MATERIALS AND METHODS

Deployment models: There are various deployment models related to the cloud computing. But generally these models can be divided into four categories. The selection of any of these models takes into account the user requirements (Anonymous, 2017).

Private cloud: This type of cloud infrastructure is deployed for a specific organization where the operation and maintenance is managed by third party at the premises or through in-house.

Community cloud: This type of cloud infrastructure is deployed where same type of organizations exist. It reduces the costs among them and the operation and maintenance is managed by third party at the premises or through in-house.

Public cloud: It is managed by cloud service provider which enables the user to deploy services and develop the system with minimal cost.

Hybrid cloud: This type of cloud infrastructure is deployed for various clouds and they may be different in nature but they should be able to communicate to each other. It is the combination of private and public clouds where the data can be shared and fulfill the requirements (Anonymous, 2017). The two basic concepts are the key features of cloud environment, i.e.

Abstraction: In the cloud environment the user's are unknown about physical data storage locations, application management, administration and the user's identification, etc (Sosinsky, 2010).

Virtualization: The cloud environment provides the facility to share the resources together as and when required from centralized infrastructure and present them as virtual resource.

RESULTS AND DISCUSSION

Architecture of cloud database management system: In this study, we present our seven tier architecture of cloud database system. Following are the seven distinct layers that we propose for our architecture:

- Physical layer
- Physical middleware layer
- Conceptual lower layer
- Conceptual layer
- Conceptual middleware layer
- External layer
- External upper layer (Fig. 1)

Physical layer: This layer deals with the design of the database. It also specifies how to physically store the database files on the secondary storage devices. We can

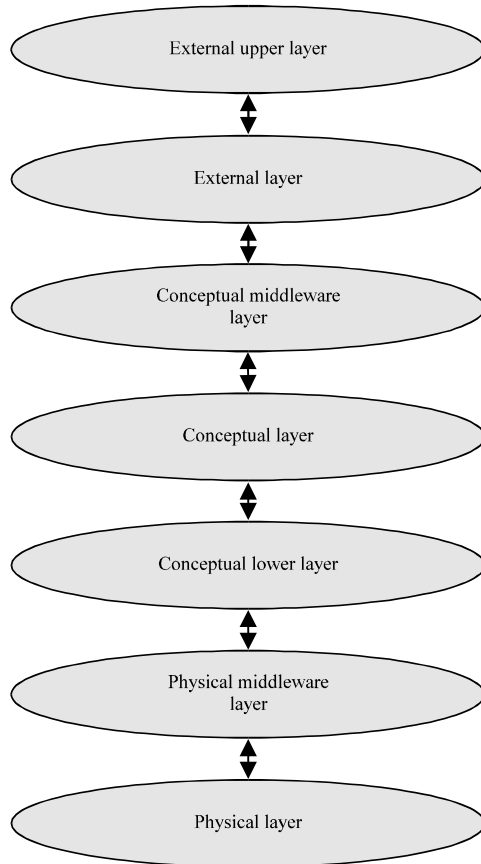


Fig. 1: Architecture of cloud database management system

mention the data characteristics, like data type and data size for a particular attribute. In particular in a cloud environment where the storage can be at multiple locations or servers, this layer can play a pivotal role (Alam *et al.*, 2013). Following are the main functionalities that this layer supports (Fig. 2).

Schema design: It deals with the process of creating a real database. The schema design has to be flexible to research in a cloud environment.

Data storage: It provides the mechanism to store the actual data in the database as per the schema design. The storage technique should cater the storage of data at multiple locations or servers. The technique used here can greatly affect the query optimization.

Data manipulation: Manipulation of data in a cloud should be fast and efficient. If any data has to be replaced or updated it should be done reliably regardless of the data location.

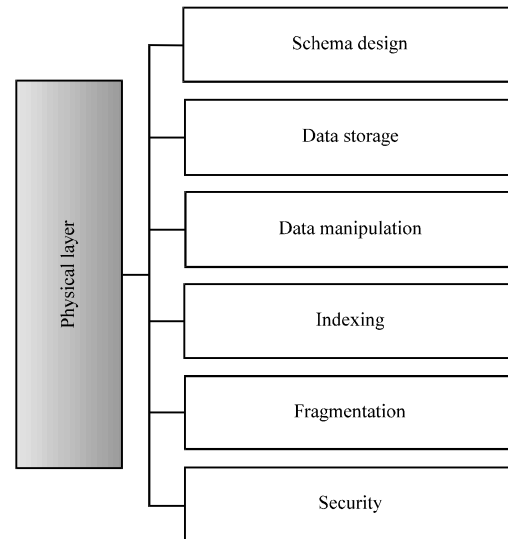


Fig. 2: Physical layer

Indexing: Indexing can be problematic in a cloud environment, since, data can be distributed. An efficient indexing technique should be devised and implemented in order to make an easy access path to the data.

Fragmentation: In this layer, we can divide the data horizontally and vertically for an easy access. This can be useful for cloud environment.

Security: In a cloud environment where data can be distributed across many servers or locations, security can be a very big concern for organizations. In this layer, we can provide security by utilizing encryption techniques. This will make sure that the data is not available in its actual form.

Physical middleware layer: This layer deals with various operating systems environment. It provides seamless connectivity regardless of the underlying platform, i.e., Windows OS, Linux OS, Mac OS (Fig. 3 and 4).

Conceptual lower layer

Logic design: It deals with the logical structure of the entire databases, describes the records and how they are related. The entire database is managed by DBA and it describe the structure of all users.

Database connectivity: The ability of this model is to provide the connectivity among the web server and application server which implement logic queries into the database.

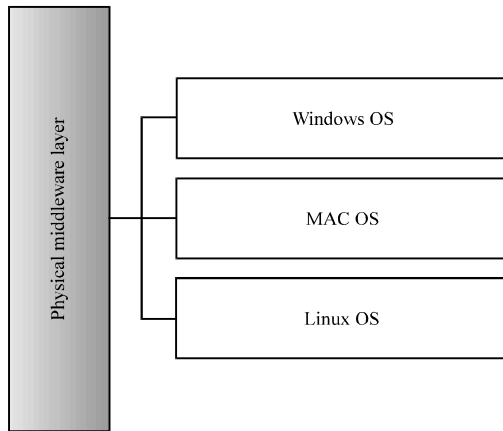


Fig. 3: Physical middleware layer

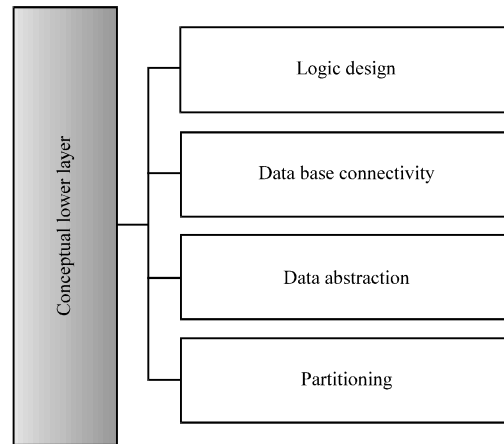


Fig. 4: Conceptual lower layer

Data abstraction: In cloud environment where all the data are not needed to be viewed by the user, it provides the facility to hide the irrelevant data from the users.

Partitioning: In the cloud environment, we tend to have a huge amount of data with large tables. We can use the technique of partitioning by dividing the table into various parts. The idea is to split the large tables to smaller ones, so that, the queries can access only a part of data and information access could be faster.

Conceptual layer: This layer deals with the program, query processing, searching and security of various databases of the system for the internal processing on data. In the cloud various type of data are incorporated to combine the traditional data, so that, the data which is placed on the cloud may be accessed and for this purpose various types of systems are required for cloud databases.

The various languages supported by cloud are such as Microsoft Azure, Google AppEngine, Haskell, Erlang, OpenStack, SQLMR, etc. that gives the results on the basis of their analysis, i.e., we use to design the various forms where the user sends the request and based on that he can get the result after interacting with the corresponding databases which physically resides in some database files on the secondary storage devices. However, it does not matter that how the database has been designed and who is going to use this program irrespective of their qualifications and expertise, language, specialized area, etc (Fig. 5).

Security: In cloud environment, this layer provides security for the interaction with the database without affecting the other layers and the source code of this layer

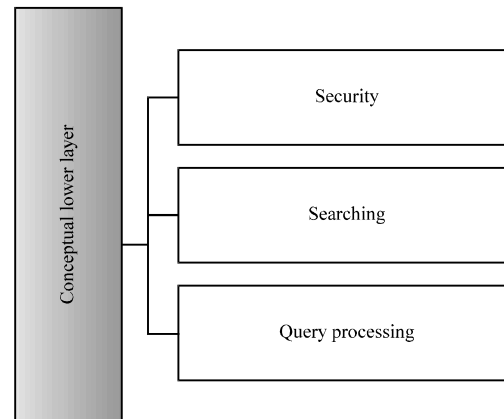


Fig. 5: Conceptual layer

cannot be changed by any other layer or the system. The threats like privacy, modification and fabrication are fully secured and cannot be changed. The information stealing mobile malware, insecure network, insecure marketplaces and proximity-based hacking are the major threats in the cloud environment.

Searching: In this layer, the posed query is interact with the relevant database and search the data in that and after query processing it returns the result in front of the user.

Query processing: In this layer, we provide the input as a query which interact with the database and generate the result to the user. The major issues occurs in cloud environment is the protection of data and query between the cloud, user and possessor.

Conceptual middleware layer: This layer provides seamless connectivity amongst various databases like

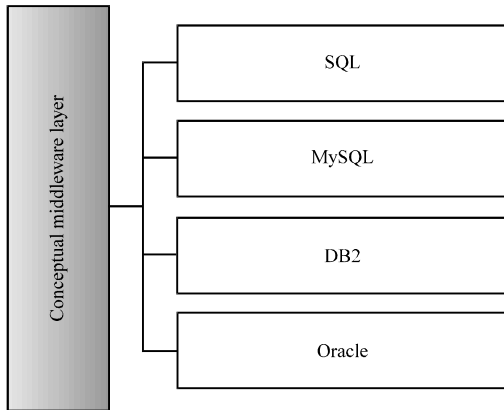


Fig. 6: Conceptual middleware layer

SQL, MySQL, DB2, Oracle. It enables the users to interoperate between any given databases (Fig. 6).

External layer: This layer deals with the user's prospective where the same database can be viewed by various user's simultaneously. It also describes part of the database for particular group of user's. Users can access the by their own customized way according to the need, hence, the same data can be viewed by different users in various ways at the same time.

In this level user interacts with the form/window that appears in front of him and sends the queries to get the result. At this level user does not think that who has designed the form and what was his qualification and when was it designed. The same way at this level the user is not concerned that from where he is getting the answer of his query from how many databases he is getting the result and where they are stored and who was the designer of the database and at what time was it created what was the qualification of the designer, etc (Fig. 7).

DML compiler optimization: In this layer while we wanted to add, delete, change or to perform any other operation onto the database, we need some media to perform it. DML compiler helps to handle these operations in efficient manner. It minimizes resource usage, reduce time complexity and finds near optimum solution.

Multuser interaction: In the cloud environment when we think about a common architecture that are used to implement multi-user database management systems like, client server architecture, file server and teleprocessing.

Seamless connectivity: External layer will help to connect the users seamlessly regardless of the underlying

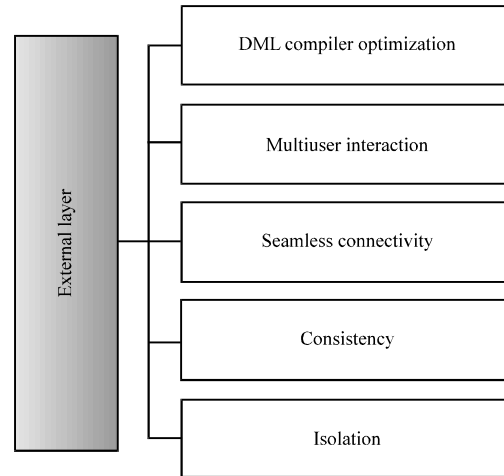


Fig. 7: External layer

database. The users main concern will be to send the query and get the response without worrying about the user device that is sending the query and the underlying architecture and the database responding with the result.

Consistency: In a cloud environment where data can be distributed across many servers or locations and when the transaction process going on unwanted or incorrect data should not insert into the database and should maintain uniformity of the database.

Isolation: In a cloud environment where data can be distributed and if the multiple transactions processing simultaneously then all the transactions should behave like individual transaction and they should not affect to each other. It means each and every transaction executed as if they are independent.

External upper layer

Interface interoperability: Cloud environment provides the facility to interface a system with other systems without any additional features on the part of end user. Several organizations are deploying distributed n-tier client/server applications, most of which require access to data or transactions on existing systems. Standards like TCP/IP, HTP and HTML etc. can serve as a good example of interface interoperability where CORBA and its ORB to provide common interfaces. The interoperability reduces operational cost and complexity, leverages existing investments, rapidly enables deployments (Fig. 8).

Interface design: It provides the interface of the system to varied users. Some examples of interfaces are as follows, menu based interfaces for web clients form Based

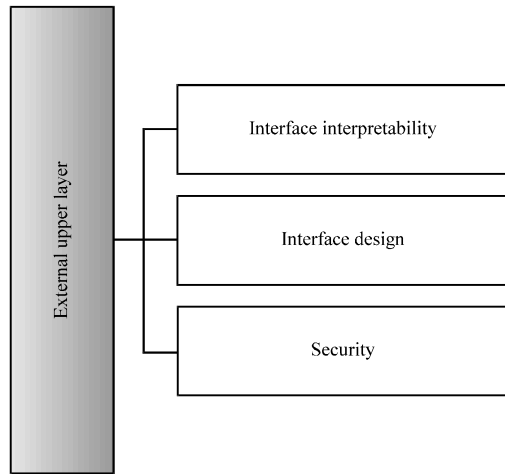


Fig. 8: External upper layer

interfaces, graphical user interfaces, natural language interfaces, interfaces for parametric users and interfaces for the DBA.

Security: Provides a powerful and flexible security mechanism by hiding parts of the databases from certain users. The user is not aware of the existence of any attributes that are missing from the view.

CONCLUSION

Cloud computing is an innovative concept and has been turning in to reality on a fast pace. With cloud computing there comes the challenge of distributed database and the service of distributed querying. In this study, we proposed a distinct seven tier architecture of cloud database management system. The main focus of our proposed architecture is to simplify the complexity of distributed querying of databases spread across different nodes of the cloud. We have defined specific functionality for each layer of the architecture that would help to manage the cloud database efficiently and effectively. We have proposed seven tiers namely physical layer, physical middleware layer, conceptual lower layer, conceptual layer, conceptual middleware layer, external layer and external upper layer. All these tiers cover a vast range of the functionality supporting the cloud architecture. Some of the functionalities are searching, indexing, optimizing, portability, seamless connectivity, interoperability, security, etc. We have also, high lighted various issues and challenges involved with each tier.

REFERENCES

- Alam, B., M.N. Doja, M. Alam and S. Mongia, 2013. 5-layered architecture of cloud database management system. AASRI. Procedia, 5: 194-199.
- Anonymous, 2017. Introduction to cloud computing. Dialogic Corporation, Milpitas, California.
- Bloor, R., 2011. The Suitability of Algebraic Data's Technology to Cloud Computing. The Bloor Group, Austin, Texas,.
- Curino, C., E.P. Jones, R.A. Popa, N. Malviya and E. Wu *et al.*, 2011. Relational cloud: A database-as-a-service for the cloud. Proceedings of the 5th Biennial Conference on Innovative Data Systems Research (CIDR'11), January 9-12, 2011, Asilomar Conference Grounds, Asilomar, California, USA., pp: 235-240.
- Dimovski, D., 2013. Database management as a cloud-based service for small and medium organizations. Master Thesis, Masaryk University, Brno, Czech Republic.
- Dinh, H.T., C. Lee, D. Niyato and P. Wang, 2013. A survey of mobile cloud computing: architecture, applications and approaches. Wirel. Commun. Mob. Comput., 13: 1587-1611.
- Kossmann, D., T. Kraska and S. Loesing, 2010. An evaluation of alternative architectures for transaction processing in the cloud. Proceedings of the ACM International Conference on Management of Data (SIGMOD'10), June 06-10, 2010, ACM, Indianapolis, Indiana, USA., ISBN:978-1-4503-0032-2, pp: 579-590.
- Kumar, N. and K.K. Bansal, 2012. Different compression techniques and their execution in database systems to improve performance. Intl. J. Adv. Res. Comput. Sci. Software Eng., Vol. 2,
- Sakr, S., A. Liu, D.M. Batista and M. Alomari, 2011. A survey of large scale data management approaches in cloud environments. Commun. Surv. Tutorials IEEE., 13: 311-336.
- Singh, T. and P.S. Sandhu, 2011. Cloud computing databases: Latest trends and architectural concepts. Intl. J. Comput. Electr. Autom. Control Inf. Eng., 5: 86-89.
- Sosinsky, B., 2010. Cloud Computing Bible. John Wiley & Sons, Hoboken, New Jersey, ISBN:9781118023990, Pages: 532.
- Stonebraker, M., D.J. Abadi, A. Batkin, X. Chen and M. Cherniack *et al.*, 2005. C-store: A column-oriented DBMS. Proceedings of the 31st International Conference on Very Large Data Bases (VLDB'05), August 30-September 02, 2005, ACM, Trondheim, Norway, pp: 553-564.

Real Option Analysis for Valuating Investments in Information Technology Projects

¹Rabii El Ghorfi, ²Mohamed El Aroussi and ¹Mohamed Ouadou

¹LRIT Associated Unit to the CNRST, Faculty of Sciences, Mohammed V Rabat University,
Rabat, Morocco

²EHTP Engineering School, Casablanca, Morocco

Abstract: This study presents an overview of ROA (Real Option Analysis) for valuating investments in IT (Information Technology) projects. For this purpose, we show how ROA assists the people in charge of the IT to take the most profitable investment decisions and how it encourages managerial flexibility. Then, we detail the mathematical basis of ROA in particular the asset variation modeled by Brownian motions and the computation of the net present value of IT projects. We also classify the ROA literature with the focus of the technique used. Lastly, we discuss the limits of the theory behind ROA. We distinguish between the limits coming from the assumptions related to the mathematical model and the limits coming from the use of intangible criteria in the pricing process. We finally conclude by reporting arguments and counter arguments of an approach based ROA from the analysis of the literature.

Key words: Information technology, real options, information systems, project management, approach, literature

INTRODUCTION

The idea of transferring ROA to real investments traces back to Myers (1974) whereas the application of ROA to the valuation of IT projects started in the early 1990's. The real option theory exploits pricing models derived from finance to model the variation of IT assets. In practice, it includes all intangible assets such as human, technical and informational resources that influence the ongoing IT projects. ROA aims at providing a good estimation of the cash flow of IT projects and thus, helping the decision makers to take the right investment decisions. ROA also presents an attractive alternative because it explicitly accounts for the value of future flexibility in management decision-making (Bardhan *et al.*, 2004).

The real option theory suggests that management should proactively manage strategic investments by creating opportunities for mid-course corrections to investment strategies (Teece, 2007). By using managerial flexibility to deal with risk, managers can avoid risk or bring it down to acceptable proportions, for example by delaying decisions, transferring risk to a third party or by redirecting the course of the project (Fichman *et al.*, 2005). Hence, the real options perspective is a promising complementary theory to existing theories by offering both financial and economic perspective on risk management logic as observed in practice.

ROA in its standard version uses an annual estimation of real options and the possible investment can take place at different points in time which is why the

investment should rather be modelled as an American option (Taudes, 1998). American real options related to IT projects allow investing before the option expires, unlike European options which cannot be exercised before the maturity. In this study, we will consider only American call options, since, it is the most predominant options used in the literature. The valuation of the underlying options thus defined in ROA traditionally relies on the standard BSM (Black Scholes Model), described Black and Scholes (1973). A quantity called NPV (Net Present Value) is computed as a stochastic optimization problem. The objective is to maximize the NPV in most cases. BSM is based on continuous time and assumes that the benefit follows a Wiener process in which options evolve according to continuous stochastic processes, more precisely GBM (Geometric Brownian Motion) processes.

The possibility to stage, abort or defer expected investments in an IT project is represented by real options. The incertitude of the cash flow of an IT project can thus be corrected by applying the right option at the right time. With ROA, The decision makers can determine the optimal time to invest and identify the most profitable option for IT projects. However, the application of ROA is based on several assumptions, which are not always verified (Ullrich, 2013).

Foundation of real options: A real option itself is the right but not the obligation to undertake certain business initiatives such as deferring, abandoning, expanding, staging or contracting a capital investment of a project. IT

Table 1: Types of real options for IT projects

Options	Description
Option to expand	When the progress of an IT project is better than what has been expected, the decision maker could invest in more resources ensuring the proper completion of an IT project
Option to withhold	Depending on the board elements, the decision maker could draw a pessimistic vision and decide to withhold future expected investments
Option to wait and see	The manager could delay an action awaiting more reliable information
Option to abort	If the expected market environment is totally unfavorable, the decision maker could abort investments in IT projects to avoid further losses
Option to defer	The manager could defer investments to some future time when the investment conditions look stable
Option to stage	If there is established inadequacy of the cash flow, the implementation of a project can be carried out in stages
Option to switch	In order to address a resource optimization problem, the input resources can be replaced depending on the course of a project
Option to adjust	The decision maker could decide to reduce or increase expected investments in IT projects
Option to learn	When there is a lack of information, the manager could decide to exercise a learning option

projects may be subject to investment restrictions or otherwise expanding investments. Trigeorgis (1996) presents a good overview of the main real options for IT projects management. Fichman *et al.* (2005) extend the works of Trigeorgis promoting management flexibility in the identified real options. Among these options, there is the option to expand, withhold, abandon and ‘wait and see’ as listed in Table 1. These options enable to invest before their expiration and the investment decision depends on the attractiveness of their IT project.

The option to expand and the option to withhold act on future expected investments, e.g., the reduction of expected investment in technical resources due to cost changes. These two options aim to adapt the investments based on the market conditions. The ‘wait and see’ option is valuable in that it will provide the decision maker with an opportunity to defer investing. Indeed, the right timing for making the investments is particularly crucial to achieve higher returns. The abandon option could occur if the expected market environment is totally unfavourable and thus, expected investments on the overall resources of a project are abandoned. Since, the investments in IT projects cover all IT resources, the presented options are not limited to the financial resources, it may encompass technical, human or informational resources and aims to cost and time savings.

The introduction of a new option in a given model needs a study of compatibility with the existing options of the Model. In the identified real options taken from the literature, we notice some redundancy, for example, the option to ‘wait and see’ consists of the combination of the option to learn and the option to defer. Another example is the option to adjust which combines the option to expand and the option to withhold. Regarding the option to switch, it is responds to the need of a proactive management of the IT while the option to stage responds to the need of cash flow optimization. Finally, the introduction of a new option depends essentially on the context of the organization and its needs.

Literature review: In the literature, many tools have been developed to help people in charge of the IT manage uncertainty and risk in their investments (Brandao and Dyer, 2005; Benaroch and Kauffman, 1999; Wu *et al.*,

2009; Ball and Deshmukh, 2013; Munoz *et al.*, 2011; Bakke *et al.*, 2016; Angelou and Economides, 2008; Ghorfi *et al.*, 2017; Collan *et al.*, 2003; Crasselt and Lohmann, 2016). The standard BSM Model is often extended by applying complementary tools such as decision trees, analytical hierarchical processes, multiple-criteria decision analysis and the construction of learning option.

Classification of the literature: As suggest by Ullrich (2013), we propose to classify the articles that monetarily value IT investments with the help of ROA into three categories: those about investments in standard software, those about investments in individual software and those about investments in new technologies. Table 2 summarizes the identified literature regarding the real options technique they use *J. Eng. Applied Sci.*

Classification of the literature: Hilhorst *et al.* consider individual preferences besides a market valuation and evaluate project-specific risks using a decision tree to compute the expected option value. Angelou and Economides (2008) apply the ‘analytical hierarchy process’ through which the quality of different sources of uncertainty can be considered and calculate the option value considering uncertain cash outflows using a single step binomial tree. Wu *et al.* (2009) acknowledge that traditional option pricing models in general do not have the ability to correctly account for the complexity of IT investments. Therefore, in their study, the researchers formulate the determination of the option value as a stochastic optimisation problem. Thus, the assumptions of ROA become irrelevant to them. Ball and Deshmukh (2013) highlights cooperative real options with incentive pricing scheme into a multi-agent system and compute the expected option value using a binomial decision tree for each task. Ghorfi *et al.* (2017) consider four real options to track the evolution of IT projects. With the help of a Monte Carlo simulation, they calculate optimal time of investing in different scenarios.

In their part, Bardhan *et al.* (2004) account for the uncertain costs of software development projects by applying the Margrabe model and therefore by modelling the cash outflows as a random variable following

Table 2: Classification of the identified literature

Categories	Articles	Description	Focus of technique
Investments in standard software	Hilhorst <i>et al.</i> in 2006	Implementation strategy for the introduction of a capability management information system	Option based risk management+decision tree
	Angelou and Economides (2008)	Prioritising a portfolio of IT projects with interdependencies to follow-up projects of a water supply	Analytical hierarchy process+binomial tree
	Wu <i>et al.</i> (2009)	Implementation of an enterprise resource planning software	Stochastic optimization problem
	Ball and Deshmukh (2013)	Coordinating supply chain and resource allocation	Multi agent system+Binomial tree
	Ghorfi <i>et al.</i> (2017)	Implementation of IT governance best practices	Monte Carlo+Scenario analysis
Investments in individual software	Bardhan <i>et al.</i> (2004)	Valuation of the IT project portfolio of an energy provider	Margrabe model+learning option construction
	Munoz <i>et al.</i> (2011)	Valuation of project investments in wind power generation	Piecewise NPV estimation+Trinomial tree
Investments in new technologies	Bakke <i>et al.</i> (2016)	Evaluating the profitability of investing in electric energy storage	Markov switching+Least squares Monte Carlo
	Benaroch and Kauffman (1999)	Deployment of point-of-sale debit services by an electronic banking network	Black approximation+optimal timing of the investment
Investments in new technologies	Collan <i>et al.</i> (2003)	Enhancing the precision of cash flows of projects in ROA	Fuzzy BSM
	Crasselt and Lohmann (2016)	Tackling short term decisions in ROA	Short term option+Binomial tree

a stochastic process. Munoz *et al.* (2011) do not assume that the NPV abides a GBM process with constant parameters. To estimate these parameters, the researchers introduce a piecewise estimation. Bakke *et al.* (2016) determine the option value and optimal investment time for a battery bank. Their model combine regression analysis with a Monte Carlo simulation in a ROA framework.

Benaroch and Kauffman (2000) assume the cash outflows for the investment to be certain in their article and use a traditional Geometric Brownian Motion (GBM) in order to determine the optimal timing of the investment. Collan *et al.* (2003) introduce a new fuzzy real option valuation method which is built on the use of fuzzy numbers and possibility distributions in order to tackle imprecision and uncertainty in a direct and explicit way. Crasselt and Lohmann (2016) consider the necessity to estimate option values at every decision point as a major drawback and therefore, propose a simplified options-based decision rule using a binomial decision tree.

MATERIALS AND METHODS

Mathematical model: In this study, we highlight the ideas and principles of Brownian motions applied to the valuation of IT asset. In the first part, we recall some basic definitions of probabilities and stochastic calculations in order to explain the valuation of financial assets through Brownian motions. Then, the second part underlines the transition from financial to IT assets.

Brownian motions to model financial assets: Let us consider a random variable X . Now introducing another parameter t referring to time, X_t is called a stochastic or random process. We assume that all X_t are Gaussian

distributions (Normal laws) with a mean μ and a variance σ : $X \sim N(\mu, \sigma)$, this process is now called a Wiener process.

A standard Brownian motion B_t is a wiener process defined with the following expectation and covariance: $E[B_t] = 0$ and $Cov(B_t, B_s) = \min(t, s)$. This process has independent and stationary increments:

- For $k > l > j > i$: $Cov(B_k - B_l, B_j - B_i) = 0$ (independent increments)
- For $t > s$: The variance, $Var(B_t - B_s) = t - s$ (stationary increments)

A Brownian motion x_t with a drift parameter μ and a diffusion coefficient σ is defined as follow:

$$x_t = \mu t + \sigma B_t \tag{1}$$

Where:

- B_t = Standard Brownien motion
- μ = Drift parameter
- σ = Diffusion coefficient

This process abides to a normal distribution $x_t \sim N(\mu t, \sigma^2 t)$ and has independent and stationary increments. The idea behind the standard BSM Model is to model the variation of a financial asset s_t by a Brownian motion. The return on investment r_t of asset s_t is assumed following a Brownian motion as shown in Eq. 2:

$$r_t = \frac{s_{t+dt} - s_t}{s_t} = \mu_t + \sigma B_t \tag{2}$$

This assumption is called the assumption of stochastic process and constant variance (i). Hence, we built a simple Brownian model of asset price movements. Further information on Brownian motions within finance can be

found by Shreve *et al.* (1997). We note that from the previous equation, return r_t abides a Brownian motion whereas asset s_t abides a geometric Brownian motion. We can verify this with a Taylor expansion of the log function assuming a low variation of the asset:

$$r_t \approx \ln(r_t + 1) = \ln\left(\frac{s_{t+\Delta t}}{s_t}\right) \quad (3)$$

Thus, we obtain the value of the asset:

$$s_t = s_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right) \quad (4)$$

Where:

- s_t = Asset value
- r_t = Return on investments
- $\sigma^2/2$ = Risk premium of uncertainty
- B_t = Standard Brownian motion

Equation 4 describes the value of a financial asset s_t which abides a geometric Brownian motion. The solution of this equation relies on ITO formulas (Shreve *et al.*, 1997). The diffusion coefficient represents the volatility of asset s_t in the standard BSM while the drift parameter represents the expectation of return r_t . The volatility can also be defined as the standard deviation of return r_t . The necessary assumptions for the transition from financial assets to real assets of IT projects are reported in the next study.

Net present value of IT assets: The condition of the transition from valuating financial assets in the stock market to valuating real assets of IT projects is the assumption of complete market. This assumption implies that there is a single price of an asset as well as there is a single return on investments for each asset. This means that there is no profitable arbitrage. Even if markets are incomplete for the great majority of the projects, this assumption is often used in the field of continuous time real option valuation and allows the determination of the correct discount rate for the projects (Brandao and Dyer, 2005).

In the standard BSM Model, the value of a call option is defined by its discounted expected terminal value $E[s_t]$ (Benaroch and Kauffman, 1999). The minimum return of the share is assumed to be a fixed-rate asset which has a constant yield. This arises from the assumption of non-fluctuating interest rates. The current value of a call option is thus given by:

$$C_t = e^{-rt} E[s_t] \quad (5)$$

where, r_t is risk free interest rate. In the case of an IT project, C_t represents the cash flow of the project. After exercising a real option, the net present value measures the profitability of investments during the life of the option. The option is profitable for a project when and only when the NPV is positive and greater than the one of other options. In the case of continuous time with the assumption of the certain duration of the option and the assumption of risk-neutrality, the NPV is given by the following Eq. 6:

$$NPV = \int C_t dt = E_Q[\int S_t e^{-rt} dt] \quad (6)$$

Where:

- E_Q = Risk neutral expectation
- NPV = Net Present Value

The net present value of share s_t given by the previous equation depends on the initial price s_0 and the three parameters r_t , μ and σ . However, it may include other parameters such as income, costs, amortizations, loan and taxes. Then, the option price and is often expressed as a function of a share s_t following a GBM and other parameters which vary over time:

$$P_t = u(s_t, t) \quad (7)$$

Where:

- u = Function of s_t and t
- P_t = Option Price

Thanks to the ITO formulas, the new expressions of volatility σ and drift parameter μ are given by Eq. 8 tacking into account all parameters of Eq. 7. Finally, the option price and the net present value are given by Eq. 9 and 10:

$$\mu_0 = \frac{\frac{\partial u}{\partial x} + \mu S_t \frac{\partial u}{\partial s} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 u}{\partial S^2}}{u(t, S_t)}, \quad \sigma_0 = \frac{\sigma S_t \frac{\partial u}{\partial S}}{u(t, S_t)} \quad (8)$$

$$P_t = P_0 \exp\left(\left(\mu_0 - \frac{\sigma_0^2}{2}\right)t + \sigma_0 B_t\right) \quad (9)$$

$$NPV = E_Q[\int P_t e^{-rt} dt] \quad (10)$$

Where:

- P_t = Option Price
- μ_0 = Drift parameter after exercising the option
- σ_0 = Diffusion coefficient after exercising the option

The final net present value depends on μ_0 , σ_0 and r_t and is the basis to determine the most profitable option among the options of the adopted ROA Model. In the

literature, the researchers generally use MIP (Mixed Integer Stochastic) programming combined with decision trees to identify the most profitable option (Brandao and Dyer, 2005; Wu *et al.*, 2009; Ball and Deshmukh, 2013). Other approaches (Munoz *et al.*, 2011; Bakke *et al.*, 2016) relies on regression analysis and Monte Carlo simulations to select the most profitable option.

RESULTS AND DISCUSSION

There is certain limitations in a real option reasoning for the valuation of IT investments. We distinguish two types of limitations:

- The assumptions of the mathematical model
- The intangible criteria in the pricing process

The mathematical assumptions: The major drawbacks due to the mathematical assumptions are: the stochastic process and constant variance assumption implies that volatile changes in the value of cash flow cannot be taken into account. The assumption of the complete market legitimates the application of option pricing models even though the underlying market is not traded. Indeed, the absence of profitable arbitrage makes difficult the price regulation.

The assumption of non-fluctuating interest rates implies that volatile changes in the value of interest rates cannot be taken into account. The assumption of the certain duration of the option supports the idea of the certain cash outflows and the certain duration of the base project. The risk-neutrality assumption considers that the projects remain risk less because of constant rebalancing of discount rates.

The intangible criteria: The integration of non-financial resources creates intangible value drivers. Indeed, the decisions of investing in human, informational and

technical resources cannot be calculated from a pure financial vision which only relies on analyzing the return on the resource investment. Other parameters such as the availability and the state of the resource are hard enough to embed into the pricing process. In the literature, many researchers try to enhance the financial valuation of non-financial resources in order to obtain a more accurate model approaching reality as closely as possible (Wu *et al.*, 2009; Ball and Deshmukh, 2013; Ghorfi *et al.*, 2017). However, if the valuation process ignores the value of intangible assets, this may jeopardize the real option valuation as a whole and therefore the investment decisions.

ROA aims at finding a balance between tolerated risks and benefits in IT projects. In the case of sudden changes, the acceptance of changes is largely depending on non-financial criteria (Fichman *et al.*, 2005). In this case, the level of maturity of processes and resources to respond to changes is an important factor. Among intangible criteria, there is also knowledge sharing, innovation, employee relationships, etc. Since, these criteria are intangible, their valuation is inevitably associated with individual judgments. Then, the integration of these intangible criteria in ROA is one of the most important challenge to reach a precise valuation of investments in IT projects.

Challenge and counter arguments: In this study, we report the main criticisms related to the use of ROA. These criticisms are derived from the analysis of the literature (Benaroch and Kauffman, 2000; Kumar, 2002; Fichman, 2004; Scialdone, 2007; Kumar *et al.*, 2008). The assumptions of ROA may differ from one model to another. However, there is common challenges when applying a ROA based model. Table 3 is built on the works and highlights the main arguments against ROA and the counter arguments. We notice that the presented list is not exhaustive.

Table 3: Arguments against ROA and counter arguments

Challenges	Counter arguments
There is a lack of transparency in options models	ROA is typically complemented by transparent approaches such as decision trees, scoring models, simulations which can give a rough approximation of option value
Absence of a traded market for IT assets, makes difficult to estimate the expected value of future cash flow of a project	NPV of the project can be considered the best estimation for the value of the project as if it were traded on the market The classic NPV approach is no more effective than the real options valuation for the estimation of future project cash flows
Option valuation uses a risk-neutral approach	While it is possible to build a portfolio of assets that tracks the risk characteristics of the investment, option valuation is the same in the risk-neutral world and the real risk-averse world In the event of not being able to build a tracking portfolio, option valuation remains a good approximation that captures flexibilities of the investment decision that are not considered by the classic NPV approach
Absence of a traded market for IT assets, makes difficult to estimate the volatility of a project	In order to determine whether a project can tolerate different sets of assumed levels of volatility, a sensitivity analysis can be performed Meanwhile, analysts can use conservative values of volatility until they gain enough experience

CONCLUSION

In this study, we provided an overview of ROA for valuating investments in IT projects. We described the types of real options and the mathematical model of ROA. Then, we moved on analyzing the literature with a focus on the techniques used. Based on the assumptions of the mathematical model, we highlight the main limits of the theory behind ROA. In addition, we showed how intangible criteria influence the pricing process of ROA. Finally, we reported the criticisms related to the use of ROA based on the analysis of the literature. We also reported the counter arguments.

REFERENCES

- Angelou, G.N. and A.A. Economides, 2008. A decision analysis framework for prioritizing a portfolio of ICT infrastructure projects. *IEEE. Trans. Eng. Manage.*, 55: 479-495.
- Bakke, I., S.E. Fleten, L.I. Hagfors, V. Hagspiel and B. Norheim *et al.*, 2016. Investment in electric energy storage under uncertainty: A real options approach. *Comput. Manage. Sci.*, 13: 483-500.
- Ball, D.R. and A. Deshmukh, 2013. A cooperative options-based strategy for coordinating supply chain and resource allocation decisions. *Intl. J. Manage. Decis. Making*, 12: 259-285.
- Bardhan, I., R. Sougstad and R. Sougstad, 2004. Prioritizing a portfolio of information technology investment projects. *J. Manage. Inf. Syst.*, 21: 33-60.
- Benaroch, M. and R.J. Kauffman, 1999. A case for using real options pricing analysis to evaluate information technology project investments. *Inf Syst. Res.*, 10: 70-86.
- Benaroch, M. and R.J. Kauffman, 2000. Justifying electronic banking network expansion using real options analysis. *MIS. Q.*, 24: 197-225.
- Black, F. and M. Scholes, 1973. The pricing of options and corporate liabilities. *J. Political Econ.*, 81: 637-654.
- Brandao, L.E. and J.S. Dyer, 2005. Decision analysis and real options: A discrete time approach to real option valuation. *Ann. Oper. Res.*, 135: 21-39.
- Collan, M., C. Carlsson and P. Majlender, 2003. Fuzzy black and scholes real options pricing. *J. Decis. Syst.*, 12: 391-416.
- Crasselt, N. and C. Lohmann, 2016. Considering real options in short-term decision making. *J. Manage. Control*, 27: 351-369.
- Fichman, R.G., 2004. Real options and it platform adoption: Implications for theory and practice. *Inf. Syst. Res.*, 15: 132-154.
- Fichman, R.G., M. Keil and A. Tiwana, 2005. Beyond valuation: Options thinking in it project management. *California Manage. Rev.*, 47: 74-96.
- Ghorfi, E.R., M.E. Aroussi, M. Ouadou and D. Aboutajdine, 2017. IT project management under real options theory and IT governance: A case study on organization relocation. *Intl. J. Manage. Decis. Making*, 16: 196-223.
- Kumar, R., A. Haya and N. Yuan, 2008. Information technology portfolio management: Literature review, framework and research issues. *Inf. Resour. Manage. J.*, 21: 1-24.
- Kumar, R.L., 2002. Managing risks in it projects: An options perspective. *Inf. Manage.*, 40: 63-74.
- Munoz, J.I., J. Contreras, J. Caamano and P.F. Correia, 2011. A decision-making tool for project investments based on real options: The case of wind power generation. *Ann. Oper. Res.*, 186: 465-490.
- Myers, S.C., 1974. Interactions of corporate financing and investment decisions implications for capital budgeting. *J. Finance*, 29: 1-25.
- Scialdone, P., 2007. *Valuing Managerial Flexibility*. Cuvillier Verlag, Gottingen, Germany, ISBN:978-3-86727-378-7.
- Shreve, S., P. Chalasani and S. Jha, 1997. *Stochastic Calculus and Finance*. Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.,
- Taudes, A., 1998. Software growth options. *J. Manage. Inf. Syst.*, 15: 165-185.
- Teece, D.J., 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strat. Manage. J.*, 28: 1319-1350.
- Trigeorgis, L., 1996. *Real Options: Management Flexibility and Strategy in Resource Allocation*. MIT Press, Cambridge, Massachusetts, USA., ISBN:9780262201025, Pages: 429.
- Ullrich, C., 2013. Valuation of IT investments using real options theory. *Bus. Inf. Syst. Eng.*, 5: 331-341.
- Wu, F., H.Z. Li, L.K. Chu, D. Sculli and K. Gao, 2009. An approach to the valuation and decision of ERP investment projects based on real options. *Ann. Oper. Res.*, 168: 181-203.

Deep Residual Network for Sound Source Localization in the Time Domain

¹Dmitry Suvorov, ²Ge Dong and ¹Roman Zhukov

¹Center for Space Research, Skolkovo Institute of Science and Technology,
143026 Moscow, Russia

²School of Aerospace Engineering, Tsinghua University, 100084 Beijing, China

Abstract: This study presents a system for sound source localization in time domain using a deep residual neural network. Data from the linear 8 channel microphone array with 3 cm spacing is used by the network for direction estimation. We propose to use the deep residual network for sound source localization considering the localization task as a classification task. This study describes the gathered dataset and developed architecture of the neural network. We will show the training process and its result in this study. The developed system was tested on validation part of the dataset and on new data capture in real time. The accuracy classification of 30 m sec sound frames is 99.2%. The standard deviation of sound source localization is 4°. The proposed method of sound source localization was tested inside of speech recognition pipeline. Its usage decreased word error rate by 1.14% in comparison with similar speech recognition pipeline using GCC-PHAT sound source localization.

Key words: Sound source localization, microphone array, deep neural network, residual network, audio processing, proposed method

INTRODUCTION

The purpose of the research is to develop a sound source localization system based on data obtained from a linear compact microphone array. The system should be resistant to noise and reverberation and also should be able to work in real time on conventional personal computers.

A large amount of noise and reverberation in captured sound is the key problem for distant speech recognition systems (Woelfel and McDonough, 2009). To solve this problem, a sound signal can be captured by microphone array to perform sound source localization and beamforming. In this case, the full process of sound capture and processing will consist of the following steps (Kumatani *et al.*, 2012):

- Sound capture with microphone array
- Sound source localization and tracking
- Beamforming
- Post-filtering

Sound source localization is the key element in this architecture because its accuracy defines quality of algorithms for implementation at further stages. Beamforming and post-filtering use previously defined sound source direction as input parameter.

At the moment there are a large number of methods for sound source localization: weighted GCC-PHAT (Grondin and Michaud, 2015) and its analogs which use sound channels correlation. The baseline Version of GCC-PHAT is presented in Eq. 1 and 2:

$$GCC_{kl}(\tau) = \int \frac{Y_k(\omega) Y_l(\omega) e^{i\omega\tau}}{|Y_k(\omega)| |Y_l(\omega)|} d\omega \quad (1)$$

where, $Y_k(\omega)$ and $Y_l(\omega)$ are discrete fourier transforms of k and l channels of the sound frame from the microphone array. Likelihood of presence of active sound source at direction Θ_i :

$$\log \text{lik}(\theta_i) = \frac{1}{M} \sum_{kl} GCC_{kl}(\tau_{kl}^*(\theta_i)) \quad (2)$$

Where:

M = A number of channels

Θ = A direction (azimuth for a linear microphone array, azimuth and elevation for planar and 3D configurations)

$\tau_{kl}^*(\theta_i)$ = A theoretical delay between k and l channels for Θ_i direction of arrival

IDOA algorithms (Tashev and Acero, 2006) estimating phase delays on different frequencies between channels of captured multichannel sound (Eq. 3-7).

Likelihood of presence of active sound source with frequency ω at direction Θ_i :

$$\log \text{lik}(\theta_i | \omega) = \frac{-\|\text{mod}(\delta(\omega) - \Delta(\omega, \theta_i), 2\pi)\|_2^2}{\left\| \frac{\partial \Delta}{\partial \theta}(\omega, \theta_i) \right\|} \quad (3)$$

Where:

$\Delta(\omega, \Theta_i)$ = A vector of theoretical phase differences between k and zero microphones at frequency ω

ω = The active sound source located at direction Θ_i and δ = A vector of measured phase differences between k and zero microphones at frequency ω

$$\delta_k(\omega) = \angle Y_k(\omega) - \angle Y_0(\omega) \quad (4)$$

$$\delta(\omega) = [\delta_1(\omega), \delta_2(\omega), \dots, \delta_{M-1}(\omega)] \quad (5)$$

Probability of presence of active sound source with frequency ω at direction Θ_i :

$$P(\theta_i | \omega) = \frac{\exp\left(\frac{\log \text{lik}(\theta_i | \omega)}{\sigma}\right)}{\left(\frac{\log \text{lik}(\theta_i | \omega)}{\sigma}\right)} \quad (6)$$

The most probable direction to the wideband sound source:

$$\hat{\theta} = \arg \max_{\theta} (\sum_{\omega} P(\theta_i | \omega)) \quad (7)$$

Scanning of the surrounding area with delay-and-sum beamformer (Valin *et al.*, 2007) or other types of beamformers.

Likelihood of presence of active sound source with frequency ω at direction Θ_i when scanning is performed using delay-and-sum beamformer:

$$\log \text{lik}(\theta_i | \omega) = \frac{1}{M} \sum_{m=0}^{M-1} e^{j\omega \tau_m^*(\theta_i)} Y_m(\omega) \quad (8)$$

The most probable direction to sound source can also be calculated using Eq. 6 and 7. MUSIC algorithms (Ishi *et al.*, 2009) and their modifications. Probability of presence of active sound source with frequency ω at direction Θ_i :

$$\log \text{lik}(\theta_i | \omega) = \frac{1}{\alpha(\omega, \theta_i)^H (I - U_s U_s^H) \alpha(\omega, \theta_i)} \quad (9)$$

Where:

$\alpha(\omega, \Theta_i)$ = The capturing matrix with size M by J (J is a number of sound sources)

U_s = Signal subspace eigenvectors matrix (Tashev, 2009)

The most probable direction to sound source can also be calculated using Eq. 6 and 7. Sound source localization algorithms based on deep neural networks (Yalta *et al.*, 2017) using convolutional and residual layers. Vector of probabilities of presence of sound source at possible directions:

$$P(\theta_0, \dots, \theta_{N-1}) = F(Y_0(\omega), \dots, Y_{M-1}(\omega)) \quad (10)$$

where, N is a number of checking directions. The architecture proposed by Yalta *et al.* (2017) is shown in Fig 1. Human speech localization algorithms based on processing data from microphone array and video camera (Suvorov and Zhukov, 2017).

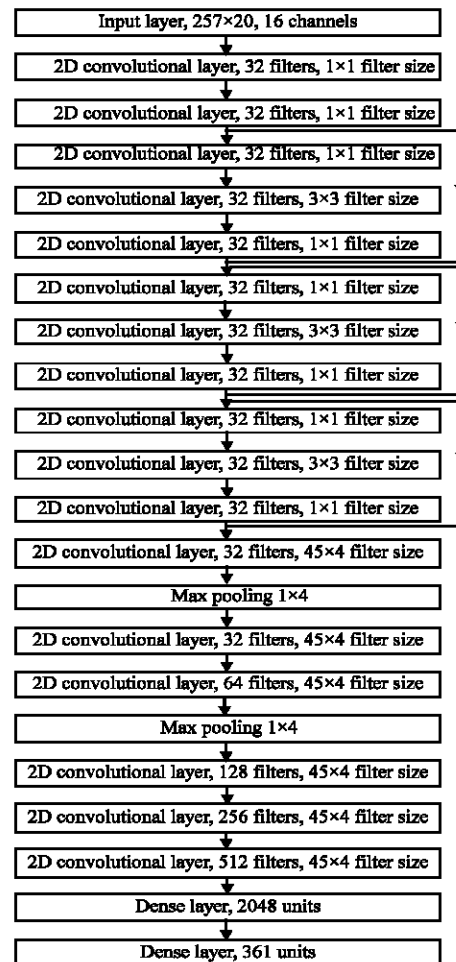


Fig. 1: Residual CNN proposed by Yalta *et al.* (2017). Each convolutional layer is followed by the ReLU non-linearity

All the methods except the one based on neural networks consider the localization problem as a problem of testing the hypothesis about sound source presence in a specific space sector which leads to an increase in required computing power as it is needed to check sound source presence in the surrounding space with a specified step (Tashev, 2009). Also, their implementations use assumptions about the plane front of an acoustic wave (Tashev, 2009) which leads to errors in localization of sound sources located closely to a microphone array (Ronzhin and Karpov, 2008).

The method based on neural networks described by Yalta *et al.* (2017) considers the localization problem as a problem of sound frame classification into sound source direction classes and the classification of active sound source absence. As input data, the algorithm uses the Discrete Fourier Transform (DFT) for every channel with sound duration of about several tens of milliseconds with some previous frames. Necessity in DFT computation for each channel on each iteration and use of two-dimensional convolutional layers, lead to increased computing complexity of the algorithm. Moreover, the algorithm uses only amplitude information from DFT and doesn't use phase information. It can also negatively affect the accuracy of localization.

Further in the study, a sound source localization method based on deep convolutional neural networks using as input, multichannel sound frames with fixed duration from microphone array will be proposed. Unlike in the method introduced in (Yalta *et al.*, 2017) the

network uses only one-dimensional convolutions which significantly reduces its computing complexity. Also the process of training dataset collection, neural network training and system testing will be described.

MATERIALS AND METHODS

Dataset: To perform experiments with deep neural network training, a big dataset of labeled data is required. To solve this problem a python application was developed which plays sound via a speaker whilst simultaneously recording it with an 8-channel microphone array with 3 cm spacing, implemented on the basis of MEMS microphones with PDM interface (Suvorov and Zhukov, 2017) which is shown in Fig. 2. The application randomly chooses and plays a music file for a duration of 30 sec from an array of one-channel sound files from "GTZAN genre collection" collected in the framework by Tzanetakis and Cook (2002). In this way, one-hour multichannel sounds for each direction with a 10° step from $0-180^\circ$ were recorded. One hour of silence was also recorded. Everything was recorded in a 2×3 m room. The sound was recorded with 16 kHz frequency with 16 bit resolution.

As dataset was collected with a linear microphone array, further the task of sound source localization was considered as a task for estimation of azimuth to sound source because the use of linear microphone array makes it impossible to determine an elevation angle for obvious geometric reasons.



Fig. 2: Linear microphone array used for capturing the dataset and real time experiments with proposed sound source localization

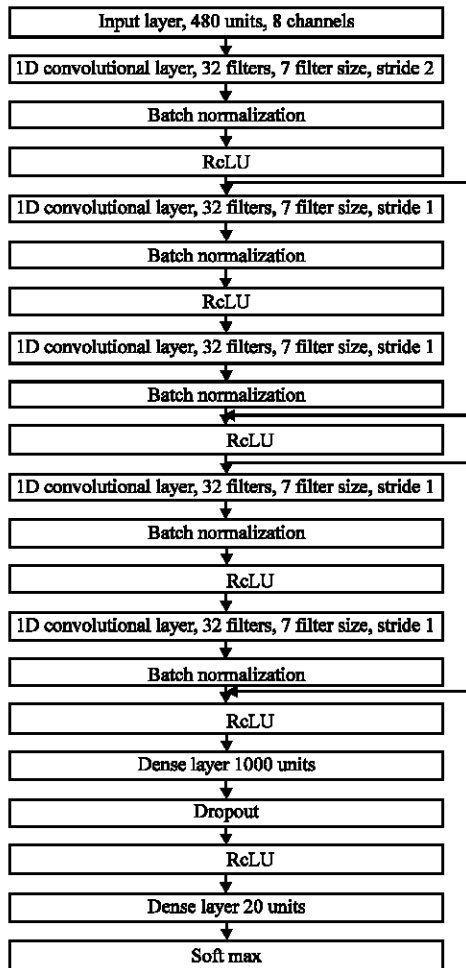


Fig. 3: A deep neural network, used for sound source localization

Neural network architecture: The developed neural network architecturally consists of four big blocks (Fig. 3):

- Input layer, accepting 8-channel sound frames from microphone array with duration of 480 samples (30 m sec) in float format
- First 1D convolutional layer (Eren, 2017) performing primary feature extraction (Eq. 11)
- Block consisting of two residual layers (He *et al.*, 2016). Residual layers allow a delay in overfitting of neural networks and therefore train deeper networks
- Decision-making blocks, consisting of two fully connected layers, create outputting probabilities that a sound frame has a sound from one of the possible azimuths and probability of absence of any active sound sources in the frame:

$$V(x, t) = \sum_{i=x-\frac{L-1}{2}}^{x+\frac{L-1}{2}} \sum_{s=1}^S K\left(i-x+\frac{L-1}{2}, S, t\right) U \quad (11)$$

Where:

$U(x, s)$ = A 1D input signal containing S channels

t = Number of output channel

$K(x, s, t)$ = A matrix of size L by S of the filter for t output channel

After each convolutional layer, a batch norm layer is used to allow to train neural networks with a lesser number of iterations to postpone overfitting (Ioffe and Szegedy, 2015). Batch Normalization Transform is shown in Eq. 12-15.

Mini-batch mean:

$$\mu\beta = \frac{1}{m} \sum_{i=1}^m x_i \quad (12)$$

Mini-batch variance:

$$\sigma_\beta^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu\beta)^2 \quad (13)$$

Normalize:

$$\hat{x}_i = \frac{x_i - \mu\beta}{\sqrt{\sigma_\beta^2 - \epsilon}} \quad (14)$$

Final scale and shift:

$$y_i = \gamma x_i + \beta \quad (15)$$

Where:

m = A batch size

x = A batch of input data

γ and β = Parameters to be learned

After the first fully connected layer, a dropout layer is also used to delay the moment of overfitting to later iterations (Srivastava *et al.*, 2014). Feed-forward operation of the dropout layer:

$$\gamma_j^l = \text{Bernoulli}(p) \quad (16)$$

$$\hat{y}^l = r^l * y^l \quad (17)$$

Where:

r^l = A vector of independent bernoulli random variables each of which has probability p of being 1

$*$ = An element-wise product

ReLU non-linearity (Maas *et al.*, 2013) is used after all the convolutional and fully connecting layers with the exception of the last dense layer:

$$f(x) = \max(0, x) \quad (18)$$

On the last layer SoftMax non-linearity (Maas *et al.*, 2013) is used as it is needed to normalize output of the neural network in such a way that the sum of all probabilities were equal to 1:

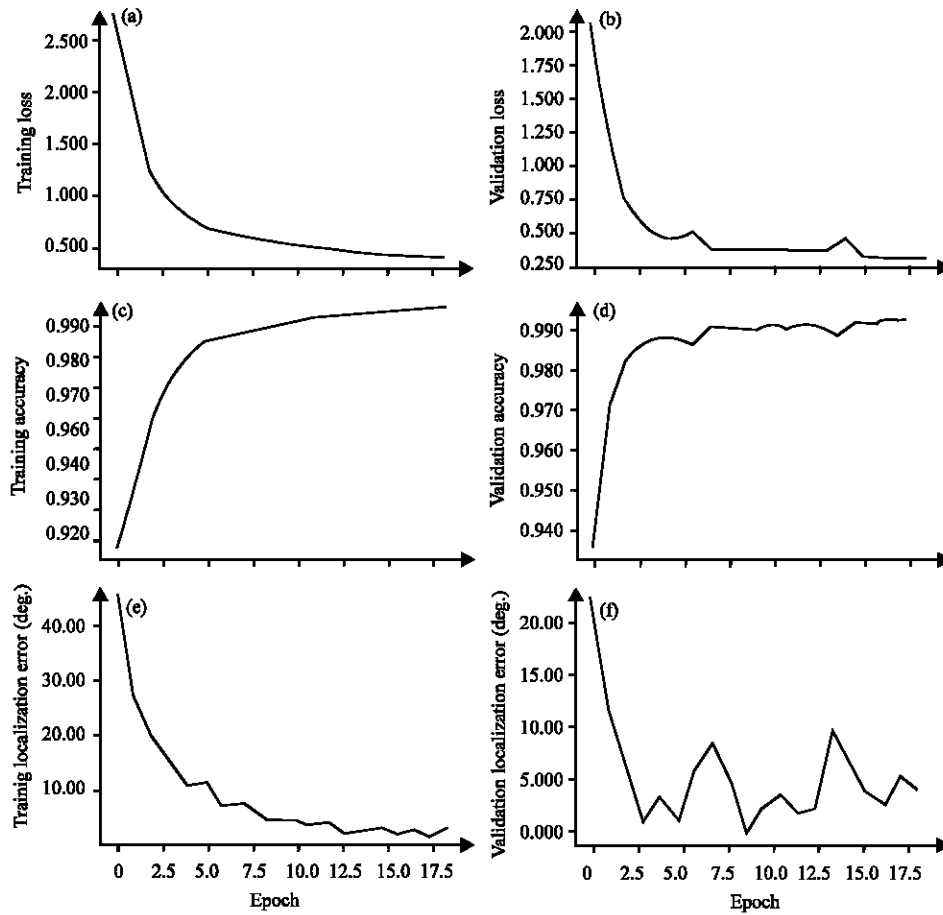


Fig. 4: The learning process of the developed system

$$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (19)$$

Unlike the solution proposed in (Yalta *et al.*, 2017) the network accepts original signals but not its fourier image. It is possible as the fourier transformation is essentially decomposition into narrowband components and therefore one-dimensional convolutional layers are able to learn this decomposition themselves.

Training and testing: A prototype of the proposed system was realized with python based on the Theano and Lasagne libraries. The learning was done with the Adam optimization algorithm (Kingma and Ba, 2015) with a low parameter of training speed (Eq. 21-23). The learning was done in 20 epochs on a NVIDIA GeForce GTX 1070 graphics card using cuda technology. Cross-entropy was used as a loss function (Eq. 24). In the training process, the values of the loss function, the accuracy of the classification and the standard deviation of the azimuth determination error were monitored. In Fig. 4, it can be

seen that overfitting happened only after 18 epochs:

$$w[t+1] = w[t] - \alpha \frac{1}{\sqrt{g[t+1] + \epsilon}} v[t+1] \quad (20)$$

$$g[t+1] = \mu g[t] + (1-\mu) \nabla(L, w[t]) \nabla(L, w[t]) \quad (21)$$

$$v[t+1] = \beta v[t] + (1-\beta) \nabla(L, w[t]) \quad (22)$$

Where:

- t = Iteration number
- L = Loss function
- w = Set of trainable parameters of the network
- ϵ, μ = Scalar parameters of the algorithm
- and β

$$L(p, y) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N y_{ij} \log(p_{ij}) \quad (23)$$

where, matrix p is N×M output of the neural network, matrix y is a one-hot encoded real class identifier, M is set to a number of classes, N is a batch size.

To analyze the training results t-SNE visualization (Maaten and Hinton, 2008) for features generated by the penultimate fully connected layer was implemented (Fig. 5). It clearly shows the cluster structure of features and the mutual arrangement of clusters corresponding to the spatial arrangement of real azimuths which indicates the good quality of the neural network training.

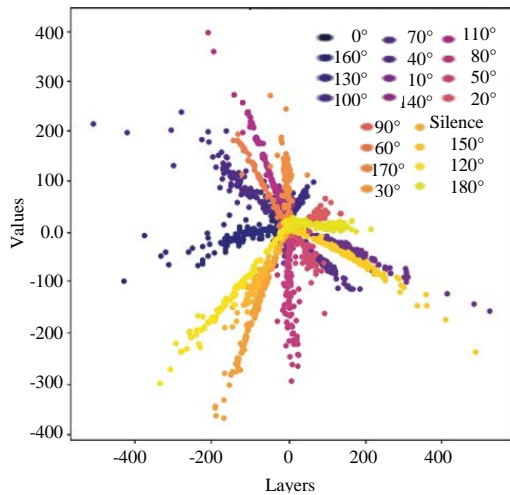


Fig. 5: T-SNE visualization of penultimate layer features

RESULTS AND DISCUSSION

Evaluation in real time: An application was developed which allows sound capture from the microphone array in real time and determines the direction of the sound source azimuth. Using this application, sound source directions were calculated in real time for sound sources in a previously known position. Measurements were done in a room where the training dataset was recorded and in another room that had a significantly different area and filling meaning it had different reverberation parameters. It can be seen in Fig. 6 that average absolute values of sound source direction azimuth determination error did not exceed 12° in both cases which is a good result, considering that the neural network was trained with an azimuth step of 10°. Mostly the same accuracy of localization in the new room and room where the train dataset was recorded, indicates a good generalization property for the trained neural network.

Figure 7 gives an example of results of continuous localization of a stationary source that plays music. It can be seen that localization error is complexity of choice between neighbouring classes of neural networks.

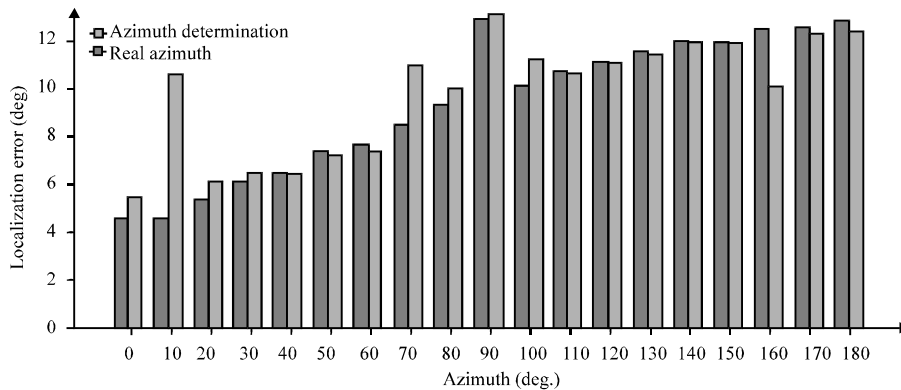


Fig. 6: Dependence of average absolute values of azimuth determination error from real azimuth with active sound source. Light grey bars show measurements conducted in the new room. Dark grey bars show measurements conducted in the room where the training dataset was recorded

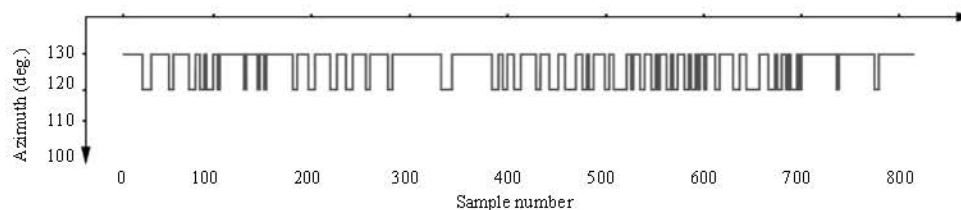


Fig. 7: Values of calculated azimuths for an active sound source as a function of sound frame number with real direction of 110°

Impact on the whole speech recognition pipeline: A comparison of the accuracy of far-field speech recognition was performed with three configurations of speech recognition pipeline to understand the impact of developed sound source localization on the final result of speech recognition. The first speech recognition pipeline didn't use microphone array processing:

- Audio capturing from the first channel of the microphone array
- Voice activity detection using code from the WebRTC project
- Speech recognition using Google speech API

The second speech recognition used sound source localization developed by Grondin and Michaud (2015):

- Audio capturing from the microphone array
- Sound source localization using weighted GCC-PHAT with Kalman filtering
- MVDR beamformer (Tashev, 2009)
- Zelinski post-filter (Aleinik, 2017)
- Speech recognition using Google speech API

Implementations of MVDR beamformer and Zelinski post-filter were used from BTK toolkit. And the last speech recognition pipeline used developed sound source localization based on the residual network:

- Audio capturing from the microphone array
- Proposed sound source localization using residual network with Kalman filtering
- MVDR beamformer
- Zelinski post-filter
- Speech recognition using Google speech API

The 100 phrases were recognized simultaneously through 3 described speech recognition pipelines. Speech recognition pipelines shared the same microphone array during the experiment. Voice sound sources were located on distance 1.5 m from the microphone array at different directions. Word Error Rates (WER) were calculated and compared for results from pipelines (Table 1).

The best result was shown by the solution with proposed sound source localization. High WER is shown by the solution with GCC-PHAT because the width of the beam pattern formed by the MVDR beamformer is lower than the accuracy of the sound source localization achieved GCC-PHAT on used microphone array, so, sometimes beam pattern became orientated not to the sound source. The width of the beam pattern of the MVDR beamformer is about 20° (Fig. 8). Average localization error of the developed sound source

Table 1: WER value for different configurations of speech recognition pipeline

Speech recognition pipeline	WER(%)
Mono audio capturing without any speech enhancement	2.21
Speech enhancement using beamforming and GCC-PHA sound source localization	2.99
Speech enhancement using beamforming and proposed sound source localization	1.85

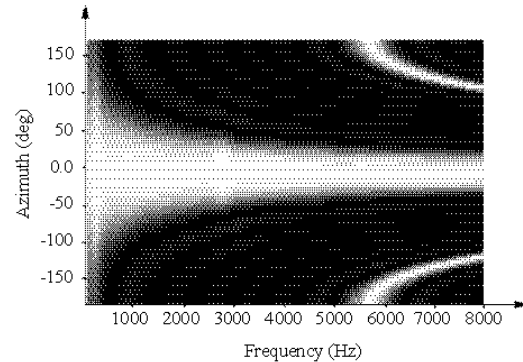


Fig. 8: The beam pattern of the MVDR beamformer in the endfire orientation for used microphone array

localization system is not higher than 12°. So, their combination achieves a good quality of speech enhancement resulting in low WER. Directivity pattern was modeled using following equations (Vary and Martin, 2006):

$$\psi(\omega, \Theta) = |H(\omega)^H u(\omega, \Theta)|^2 \quad (24)$$

MVDR filter coefficient vector (Vary and Martin, 2006):

$$H(\omega) = \frac{\Phi(\omega)_{NN}^{-1} a(\omega)}{a(\omega)^H \Phi(\omega)_{NN}^{-1} a(\omega)} \quad (25)$$

The noise cross-power spectral matrix (Tashev, 2009):

$$\Phi_{NN}(\omega) = \Phi_{N'N'}(\omega) + \Phi_{II}(\omega) \quad (26)$$

The instrumental noise cross-power spectral matrix (Ishi *et al.*, 2009):

$$\Phi_{II}(\omega) = N_i^2(\omega) I \quad (27)$$

where, $N_i(\omega)$ is the magnitude of the instrumental noise in single microphone. The cross-spectral density for an isotropic noise field (Tashev, 2009):

$$\Phi_{ij}(\omega) = N_o(\omega) \sin c\left(\frac{\omega d_{ij}}{v}\right) \quad (28)$$

$$\Phi_{N \times N}(\omega) = \begin{bmatrix} \Phi_{11}(\omega) & \dots & \Phi_{1M}(\omega) \\ \vdots & \ddots & \vdots \\ \Phi_{M1}(\omega) & \dots & \Phi_{MM}(\omega) \end{bmatrix} \quad (29)$$

Where:

- $N_G(\omega)$ = The noise spectrum captured by an omnidirectional microphone
- v = The speed of sound
- d_j = A distance between i and j microphones

The propagation vector for linear microphone array (Vary and Martin, 2006):

$$a(\omega) = \left[a_i \cdot e^{\frac{j\omega p_i}{v}}, i=1, \dots, M \right] \quad (30)$$

where, p_i is a position of i microphone. The unit vector in the required direction of beam pattern (Vary and Martin, 2006):

$$u(\omega, \Theta) = \left[u_i \cdot e^{\frac{j\omega \cos(\Theta) p_i}{v}}, i=0, \dots, M-1 \right] \quad (31)$$

CONCLUSION

A sound source localization method based on deep residual neural networks was developed. It doesn't require a captured signal to be transformed from time domain to frequency domain with fourier transformation which positively affects system performance. The developed method demonstrated good accuracy of the sound source direction azimuth determination with a linear compact microphone array even without the consideration of object dynamics with a Kalman filter or particle filter. As a further improvement to the method, the system can be trained in such a way that will allow us to determine several sound source locations simultaneously.

SUGGESTIONS

Also in future research, the architecture should be complemented by LSTM, BLSTM or GRU layers (Chung *et al.*, 2014) to make the network able to consider object dynamics.

ACKNOWLEDGEMENT

This research was supported by the Russian Innovation Support Fund (project 102GRNTIS5/26071).

REFERENCES

Aleinik, S., 2017. Acceleration of Zelinski post-filtering calculation. *J. Signal Process. Syst.*, 88: 463-468.

Chung, J., C. Gulcehre, K. Cho and Y. Bengio, 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *Proceedings of the NIPS Workshop on Deep Learning*, December 8-13, 2014, NIPS, Montreal, Quebec, Canada, pp: 1-9.

Eren, L., 2017. Bearing fault detection by one-dimensional convolutional neural networks. *Math. Prob. Eng.*, 2017: 1-9.

Grondin, F. and F. Michaud, 2015. Time difference of arrival estimation based on binary frequency mask for sound source localization on mobile robots. *Proceedings of the IEEE-RSJ International Conference on Intelligent Robots and Systems (IROS)*, September 28-October 2, 2015, IEEE, Hamburg, Germany, ISBN:978-1-4799-9994-1, pp: 6149-6154.

He, K., X. Zhang, S. Ren and J. Sun, 2016. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 26-July 1, 2016, IEEE, Las Vegas, Nevada, USA., ISBN:9781509014385, pp: 770-778.

Ioffe, S. and C. Szegedy, 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning*, July 07-09, 2015, Microtome Publishing, Lille, France, pp: 448-456.

Ishi, C.T., O. Chatot, H. Ishiguro and N. Hagita, 2009. Evaluation of a MUSIC-based real-time sound localization of multiple sound sources in real noisy environments. *Proceedings of the IEEE-RSJ International Conference on Intelligent Robots and Systems IROS*, October 10-15, 2009, IEEE, St. Louis, Missouri, USA., ISBN:978-1-4244-3803-7, pp: 2027-2032.

Kingma, D. and J. Ba, 2015. Adam: A method for stochastic optimization. *Proceedings of the International Conference on Learning Representations ICLR*, May 7-9, 2015, San Diego, California, USA., pp: 1-15.

Kumatani, K., J. McDonough and B. Raj, 2012. Microphone array processing for distant speech recognition: From close-talking microphones to far-field sensors. *IEEE. Signal Process. Mag.*, 29: 127-140.

Maas, A., A. Hannun and A. Ng, 2013. Rectifier nonlinearities improve neural network acoustic models. *Proc. ICML.*, 30: 1-6.

Maaten, L.V.D. and G. Hinton, 2008. Visualizing data using T-SNE. *J. Machine Learn. Res.*, 9: 2579-2605.

- Ronzhin, A. and A. Karpov, 2008. [Comparison of methods for localization of multimodal system user by his speech (In Russian)]. *J. Instrum. Eng.*, 51: 41-47.
- Srivastava, N., G.E. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, 2014. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15: 1929-1958.
- Suvorov, D. and R. Zhukov, 2017. Device for synchronous data capturing from the array of MEMS microphones with PDM interface (In Russian). IFI CLAIMS Patent Services Company, Madison, Connecticut.
- Tashev, I. and A. Acero, 2006. Microphone array post-processor using instantaneous direction of arrival. *Proceedings of the International Workshop on Acoustic, Echo and Noise Control*, September 12-14, 2006, IWAENC, Paris, France, pp: 1-4.
- Tashev, I., 2009. *Sound Capture and Processing: Practical Approaches*. John Wiley & Sons, New York, USA., ISBN:9780470319833, Pages: 388.
- Tzanetakis, G. and P. Cook, 2002. Musical genre classification of audio signals. *IEEE Trans. Speech Audio Process.*, 10: 293-302.
- Valin, J.M., F. Michaud and J. Rouat, 2007. Robust localization and tracking of simultaneous moving sound sources using beam forming and particle filtering. *Rob. Auton. Syst.*, 55: 216-228.
- Vary, P. and R. Martin, 2006. *Digital Speech Transmission: Enhancement, Coding and Error Concealment*. John Wiley & Sons, New York, USA., ISBN-13:978-0-471-56018-9, Pages: 607.
- Woelfel, M. and J. McDonough, 2009. *Distant Speech Recognition*. John Wiley & Sons, New York, USA., ISBN:9780470517048, Pages: 594.
- Yalta, N., K. Nakadai and T. Ogata, 2017. Sound source localization using deep learning models. *J. Rob. Mech.*, 29: 37-48.

Reducing Environmental Noise Light Interference in Visible Light Communication Using a Solar Cell

Seong-Ho Lee

Department of Electronic and IT Media Engineering,
Seoul National University of Science and Technology, 232 Gongneung-ro,
Nowon-gu, 01811 Seoul, Korea

Abstract: In this study, we introduce a new method to reduce the interference of environmental noise light in a Visible Light Communication (VLC) system, using a solar cell. In the VLC transmitter a Light Emitting Diode (LED) array was used for the light source and in the VLC receiver a differential detector with a photodiode and a solar cell was used for the light detector. By subtracting the solar cell voltage from the photodiode voltage using a differential amplifier, we removed the noise contained in the photodiode voltage. In the experiment, the Signal-to-Noise Ratio (SNR) was improved by about 15 dB using the solar cell. This configuration is useful for constructing a VLC link without noise light interference in an environment where the noise light interference from a conventional fluorescent lamp or incandescent lamp is severe.

Key words: Visible light communication, LED, environmental noise light, noise reduction, constructing, incandescent

INTRODUCTION

Recently, LED manufacturing technology has been rapidly developing and high-power visible LEDs of several W or more are being mass-produced and widely used as a light source for illumination. Visible LEDs have higher power conversion efficiency than conventional fluorescent lamps or incandescent lamps and the LED illumination is easily controlled by changing the injection current, so the use of visible LEDs is steadily increasing. Since, LED light can be modulated at a speed faster than the human eye can detect, Visible Light Communication (VLC) technology which uses LED light to perform simultaneous illumination and communication is steadily developing (Komine and Nakagawa, 2004; Cheong *et al.*, 2013).

Since, VLC performs both the functions of lighting and communication at the same time, the system must be designed so that the lighting and communication are not affected by each other. In order to keep flicker-free lighting during data transmission, LED modulation using Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK), manchester coding or Pulse Position Modulation (PPM) has been widely used. These are efficient ways of keeping the average LED power constant during data transmission, thus, preventing the LED light from perceptively flickering (Ma *et al.*, 2012; Rajagopal *et al.*, 2012).

Since, the VLC usually transmits data through a free space between the LED and the photodetector, the

transmission condition can be affected by interference from noise light from adjacent lighting facilities such as fluorescent lamps and incandescent lamps. Adopting an optical filter or an electrical filter in the VLC receiver is an efficient way to reduce the influence of noise light (Kahn and Barry, 1997). When the noise frequency is near to that of the signal or when the noise wavelength overlaps the signal wavelength, it may be difficult to prevent noise interference with optical or electrical filters. In that case a differential detection can be an efficient way to reduce the noise light interference (Street *et al.*, 1997). Differential detection is a method of using two detectors with different coupling coefficients to noise and remove noise components by subtracting one detector voltage from the other.

In this study, we newly introduce a differential detection method that removes the interference of low-frequency noise light using a solar cell. In this structure we use two light-detecting components with different bandwidths: one is a high-speed photodiode and the other is a low speed solar cell. The solar cell has the bandwidth of a few tens kHz which is much lower than the photodiode bandwidth, due to its wide receiving area (Lee, 2015). Although, the response speed of a solar cell is very low compared to that of a photodiode, the solar cell detects 120 Hz noise from fluorescent lamps or incandescent lamps that are driven by an AC power source of 60 Hz.

If we use a signal frequency that is lower than the photodiode bandwidth and higher than the solar cell, the

photodiode will detect the signal light mixed with noise light while the solar cell detects only noise light. In this configuration the solar cell is used for two purposes: the AC component of the solar cell voltage is used as a differential input for removing the noise component in the photodiode and the DC component is used as a power supply for biasing the photodiode. This structure is very useful for constructing a noise-free VLC link using LED light in an environment where the interference of noise light due to a conventional fluorescent lamp or incandescent lamp is severe.

MATERIALS AND METHODS

The VLC system is composed of a transmitter and a receiver. The system configuration of the transmitter and a receiver are as follows.

VLC transmitter: The VLC transmitter is illustrated in Fig. 1. It is composed of an Amplitude-Shift-Keying (ASK) modulator, a Field Effect Transistor (FET) and an LED array. Input data is ASCII code generated by a microprocessor and applied to the input of the ASK modulator. The ASK modulated voltage is applied to the gate of the FET. The FET is biased to operate in the linear region using two resistors Ra and Rb.

The drain current of the FET is proportional to the gate voltage and flows to the LED array and as a result the LED output light is proportional to the ASK modulated input data. The LED output light is radiated to the free space and a part of the light reaches the VLC receiver. The optical power of the LED array can be expressed as follows:

$$P(t) = \begin{cases} P_0 + P_s \sin \omega t & \text{when } V_{in} = 1 \\ P_0 & \text{when } V_{in} = 0 \end{cases} \quad (1)$$

Where:

- P_0 = The DC optical power
- P_s = The optical amplitude of the ASK modulated signal
- ω = The subcarrier frequency
- V_{in} = The input voltage that is applied to the input port of the VLC transmitter

The devices used in the VLC transmitter were as follows. Input data was generated by an Atmega 32 microprocessor and the ASK modulator was made using a 200 kHz oscillator and ADG 417 analog switch. The LED array was made of six identical LEDs in the form of a 2H3 array, each LED was a 1 W white LED.

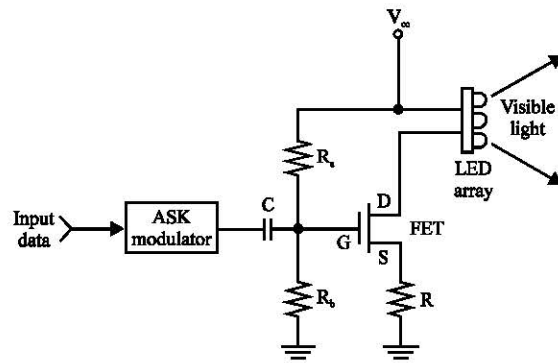


Fig. 1: VLC transmitter

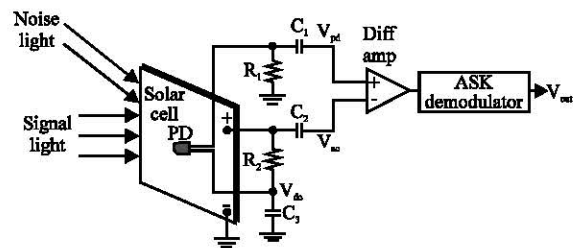


Fig. 2: VLC receiver

VLC receiver: The VLC receiver is schematically shown in Fig. 2. The VLC receiver is composed of a PhotoDiode (PD), a solar cell, a differential amplifier and an ASK demodulator. In Fig. 2, both the photodiode and the solar cell are used for light detection, however, the solar cell has a bandwidth that is much lower than that of the photodiode. R_1 and R_2 are the load resistances of the photodiode and the solar cell, respectively. C_1 and C_2 are the capacitors for AC detection and capacitor C_3 was used to supply DC power to the photodiode using the solar cell voltage. Therefore, in this configuration, the solar cell is used for noise light detection and at the same time supplying DC power to the photodiode.

The photodiode and the solar cell detect the signal light from the VLC transmitter. At this time, noise light from a nearby lighting facility such as a fluorescent lamp or an incandescent lamp may exist at the position of the photodetector. In this environment, the signal light from the VLC transmitter and the noise light from other nearby lighting facilities enter the photodiode and the solar cell at the same time. In general, the receiving bandwidth of a solar cell is significantly lower than that of a photodiode. When the ASK modulation subcarrier frequency in the VLC transmitter is lower than the photodiode receiving bandwidth and higher than that of the solar cell, the AC voltages generated in the photodiode and the solar cell can be expressed as follows:

$$V_{pd}(t) = \begin{cases} \rho_1 R_1 [k_{1,s} P_s \sin \omega t + k_{1,N} P_N(t)] & \text{when } V_m = 1 \\ \rho_1 R_1 k_{1,N} P_N(t) & \text{when } V_m = 0 \end{cases} \quad (2)$$

$$V_{solar}(t) = \rho_2 R_2 k_{2,N} P_N(t) \quad (3)$$

Where:

- $V_{pd}(t)$ and $V_{solar}(t)$ = The output voltages of the photodiode and solar cell, respectively
- $P_N(t)$ = The noise light power
- ρ_1 and ρ_2 = The responsivities of the photodiode and the solar cell, respectively
- R_1 and R_2 = The load resistances of the photodiode and the solar cell, respectively
- $k_{1,s}$ and $k_{1,N}$ = The coupling coefficient from the VLC transmitter and the noise source to the photodiode, respectively
- $k_{2,N}$ = The coupling coefficient from the noise source to the solar cell

In Eq. 2, the first term is the signal voltage from the VLC transmitter and the second is the noise voltage from the environmental noise source such as a fluorescent lamp. In Eq. 3, the solar cell voltage only has the noise term because the solar cell bandwidth is lower than the ASK carrier frequency and it does not detect the ASK carrier frequency. As shown in Fig. 2, the photodiode voltage and the solar cell voltage are applied to the positive and negative input terminals of a differential amplifier, respectively. The output voltage of the differential amplifier is:

$$V_{pd}(t) = G(V_{pd} - V_{solar}) = \begin{cases} G[\rho_1 R_1 k_{1,s} P_s \sin \omega t + \rho_1 R_1 k_{1,N} P_N(t) - \rho_2 R_2 k_{2,N} P_N(t)] & \text{when } V_m = 1 \\ G[\rho_1 R_1 k_{1,N} P_N(t) - \rho_2 R_2 k_{2,N} P_N(t)] & \text{when } V_m = 0 \end{cases} \quad (4)$$

where, G is the voltage gain of the differential amplifier. If we set $\rho_1 R_1 k_{1,N} = \rho_2 R_2 k_{2,N}$ by adjusting the ratio R_1/R_2 , in Eq. 3) the noise term disappears and only the signal term remains. That is:

$$V_{out}(t) = \begin{cases} G \rho_1 R_1 k_{1,s} \sin \omega t & \text{when } V_m = 1 \\ 0 & \text{when } V_m = 0 \end{cases} \quad (5)$$

As a result when the input data is “1” in the transmitter, only the subcarrier signal exists in the receiver and when the input data is “0”, the receiver voltage is 0 V, so that, the ASK signal is normally received without being influenced by the noise light.

The devices used in the VLC receiver were as follows. We used an M11080-12 V silicon solar cell, BPW34 PIN photodiode for light detection and used an opa228 op-amp for the differential amplifier. The ASK demodulator was composed of an envelope detector and

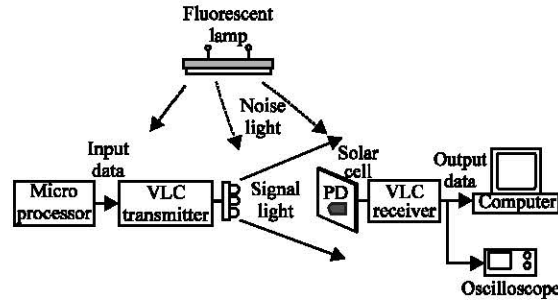


Fig. 3: Experimental setup

a comparator. The envelope detector was made of a diode and RC circuit and an LM7171 op-amp was used for the comparator. The load resistances R_1 and R_2 were 1 k and 2 k Ω , respectively. Capacitors C_1 - C_3 were 0.1, 0.1 and 100 uF, respectively.

Experimental setup: Figure 3 is a schematic diagram of the experimental setup. A VLC transmitter and an LED array were installed on a table. The VLC transmitter circuit was configured as described in Fig. 1. The LED array was made of six identical LEDs in the form of a 2H3 array, each LED was a 1 W white LED. An Atmega 8 microprocessor was connected to the input port of the VLC transmitter and used as a data generator. A 60 W fluorescent lamp was installed on the ceiling of the laboratory for indoor illumination. The fluorescent lamp showed 120 Hz noise and it induced noise voltage in the VLC receiver.

At a distance of about 2 m from the LED array, a VLC receiver with a solar cell and a PIN photodiode was placed. The solar cell was an M11080-12 V and the photodiode was BPW34. The VLC receiver circuit was configured as shown in Fig. 2. The output of the VLC receiver was sent to a computer to see the transmitted character data. We used an HP 52615B oscilloscope to observe the waveforms in the VLC transmitter and the receiver.

Observed waveforms: We programmed the microprocessor to generate ASC2 characters “A\tVLC-CH\r\n” and applied the data to the input port of the ASK modulator whose carrier frequency was 200 kHz. The data rate was 9.6 kbps and the data was in Non-Return-to-Zero (NRZ) code format. The ASK modulated input data was applied to the gate of an FET as shown in Fig. 1. The FET was biased to operate in the linear region and the drain current was proportional to the gate voltage. The drain current flowed to the LED array and it radiated visible light into free space.

Figure 4 shows the waveforms observed with an oscilloscope in the VLC transmitter. Figure 4a shows the

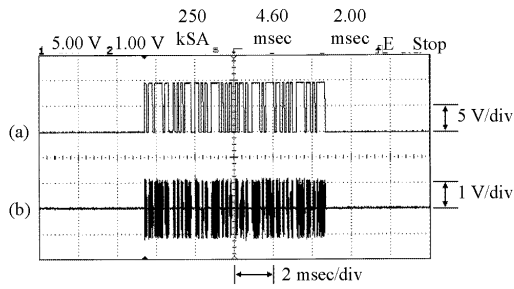


Fig. 4: Waveforms in VLC transmitter: a) Input data waveform and b) ASK modulated waveform

voltage waveform of the data going from the microprocessor to FET gate in the VLC transmitter. The data waveform corresponds to a character string “A\\tVLC-CH\\r\\n” in NRZ code. Figure 4b is the output voltage of the ASK modulator. The black region is a 200 kHz carrier signal and the envelope is the same as the transmitted data. The ASK modulated light from the LED array radiated toward the direction of the VLC receiver. The photodiode and the solar cell in the VLC receiver have different bandwidths. In separate experiments, we measured the bandwidths of the PIN photodiode and the solar cell. In order to measure the bandwidths, we modulated the LED with a sine wave of 1Vpp amplitude and measured the voltages of the photodiode and the solar cell while increasing the modulation frequency.

Figure 5 shows the measured AC responses of the photodiode and the solar cell. Figure 5a is the normalized AC response of the PIN photodiode and Fig. 5b is that of the solar cell. The 3 dB receiving bandwidth is the frequency at which the AC response falls to 1/√2 of its maximum. As shown in Fig. 5, the 3 dB bandwidth of the PIN photodiode and the solar cell were measured to be about 500 kHz and about 8 kHz, respectively. Thus, the carrier frequency 200 kHz in the ASK modulator of the VLC transmitter was lower than the PIN photodiode bandwidth and higher than the solar cell bandwidth. The differential amplifier in the VLC receiver subtracts the solar cell voltage from the PIN photodiode voltage and detects only the signal voltage as shown in the circuit of Fig. 2.

Figure 6 shows the waveforms observed in the VLC receiver. Figure 6a shows a PIN photodiode voltage in which the ASK modulated signal is mixed with the 120 Hz noise from the fluorescent lamp. The peak-to-peak ASK signal amplitude was about 0.13 V, the peak-to-peak noise voltage was about 0.14 V and the Signal-to-Noise Ratio (SNR) was 0.93 (-0.63 dB). Figure 6b is the solar cell voltage in which only 120 Hz noise is detected. Figure 6c denotes the output voltage of the differential amplifier in

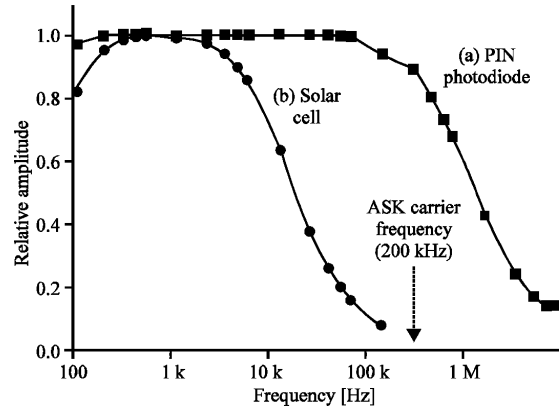


Fig. 5: AC response characteristics: a) PIN photodiode and b) A solar cell

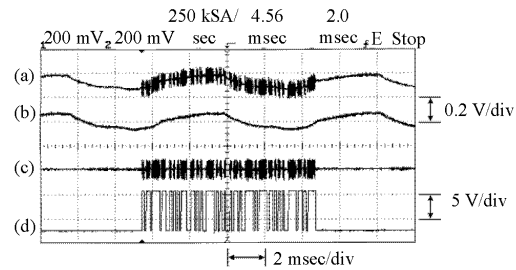


Fig. 6: Waveforms observed in VLC receiver: a) Photodiode voltage; b) Solar cell voltage; c) differential amplifier output and d) ASK demodulated output

the VLC receiver. In this waveform, the 120 Hz noise almost disappeared and only the ASK modulated signal was detected. In this waveform, the peak-to-peak ASK signal voltage was about 4.4 V while the peak-to-peak noise voltage was about 0.8 V and SNR was 5.5 (14.8 dB). Comparing the two waveforms of Fig. 6a and c, the SNR in the photodetector was improved by 15.43dB by using a solar cell. Figure 6d is the recovered NRZ digital waveform observed at the output of the ASK demodulator in the VLC receiver. This waveform was the same as the input data “A\\tVLC-CH\\r\\n” in Fig. 4a.

To check the transmission condition of the characters, we connected this waveform to the serial port of a computer and observed the characters on the monitor. Figure 7 shows the characters displayed on the computer monitor. As shown in Fig. 7, the character string “VLC-CH” was displayed on the monitor. Among the characters sent by the VLC transmitter, “\\t” (horizontal tab), “\\r” (carriage return), “\\n” (line feed) do not appear on the monitor because they are special characters indicating the positions of the characters on the screen.

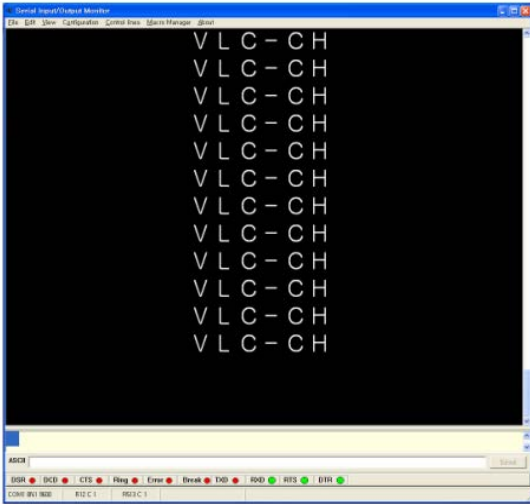


Fig. 7: Characters displayed on a monitor

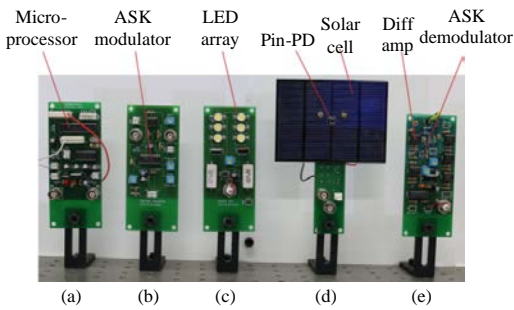


Fig. 8: Circuit boards used in experiments: a) microprocessor; b) ASK modulator; c) LED array; d) PIN photodiode and solar cell circuit and e) Differential amplifier and ASK demodulator

Thus, we experimentally confirmed that the interference of noise light can be prevented by using a solar cell in an environment where noise light could affect the VLC transmission.

Figure 8 shows the circuit boards used in the experiments. Figure 8a is the microprocessor circuit that was used for a data generator in VLC transmitter. Figure 8b is the ASK modulator in the VLC transmitter. Figure 8c shows the 2×3 LED array with FET circuit. Figure 8d shows a PIN photodiode and a solar cell. Figure 8e is the differential amplifier and the ASK demodulator circuit.

RESULTS AND DISCUSSION

We have experimentally demonstrated that we can reduce the interference of environmental noise light using differential detection with a photodiode and a solar cell in

the VLC receiver. We measured their bandwidths and determined the PIN photodiode bandwidth was about 500 kHz and the solar cell bandwidth was about 8 kHz, respectively. The noise frequency from a fluorescent lamp was 120 Hz which is twice the power line frequency of 60 Hz. Therefore, the noise could be detected by the PIN photodiode and the solar cell simultaneously. By setting the ASK carrier frequency higher than the solar cell bandwidth, we could reduce the noise interference by subtracting the solar cell voltage from the photodiode voltage. In experiments we set the ASK modulation carrier frequency at 200 kHz which is between the bandwidths of the PIN photodiode and the solar cell.

In the VLC receiver, we placed the photodiode at the center of the solar cell surface in order to ensure the noise light intensity at the photodiode and at the solar cell would be the same. When the photodiode and the solar cell were exposed to the ASK signal light from the VLC transmitter and the noise light from a fluorescent lamp simultaneously, the photodiode detected the signal mixed with the noise while the solar cell detected the noise only. This is because the photodiode bandwidth was higher than the ASK carrier frequency while the solar cell bandwidth was lower than the ASK carrier frequency. By subtracting the solar cell voltage from the photodiode voltage using a differential amplifier we could detect the signal with the noise almost entirely removed.

CONCLUSION

In experiments we improved the SNR in the VLC receiver by about 15 dB using the solar cell. In this configuration the solar cell was used for two functions: for noise reduction as well as to supply DC power to the photodiode. This configuration is very useful for constructing a noise-free VLC link using LED light in an environment where noise light interference from a conventional fluorescent lamp or incandescent lamp is severe.

REFERENCES

Cheong, Y.K., X.W. Ng and W.Y. Chung, 2013. Hazardless biomedical sensing data transmission using VLC. *IEEE. Sens. J.*, 13: 3347-3348.
 Kahn, J.M. and J.R. Barry, 1997. Wireless infrared communications. *Proc. IEEE*, 85: 265-298.
 Komine, T. and M. Nakagawa, 2004. Fundamental analysis for visible-light communication system using LED lights. *IEEE. Trans. Consum. Electron.*, 50: 100-107.

- Lee, S.H., 2015. A passive transponder for visible light identification using a solar cell. *IEEE. Sens. J.*, 15: 5398-5403.
- Ma, X., K. Lee and K. Lee, 2012. Appropriate modulation scheme for visible light communication systems considering illumination. *Electron. Lett.*, 48: 1137-1139.
- Rajagopal, S., R.D. Roberts and S.K. Lim, 2012. IEEE 802.15.7 visible light communication: Modulation schemes and dimming support. *IEEE. Commun. Mag.*, 50: 72-82.
- Street, A.M., P.N. Stavrinou, D.C. O'brien and D.J. Edwards, 1997. Indoor optical wireless systems: A review. *Opt. Quantum Electron.*, 29: 349-378.

Analysis of the Benefits for Adopting Building Information Modelling (BIM) in Iraq

¹Wadhah A. Hatem, ²Abbas M. Abd and ³Nagham N. Abbas

¹Baquba Technical Institute, College of Engineering, Middle Technical University, Baqubah, Iraq

²Department of Architecture,

³Department of Civil Engineering, College of Engineering, University of Diyala, Baqubah, Iraq

Abstract: Building Information Modeling (BIM) is a unified and comprehensive system for all what is associated with the construction project which includes a set of effective policies, procedures and computer applications that increase the level of performance in construction project during its life cycle. This study investigated the potential benefits of using BIM in Iraq. The quantitative approach was adopted by conducting a questionnaire directed to professionals in the field of construction projects in the public and private sectors. (300 copies) were distributed to the private companies and governmental institutions and departments. The data were subjected to the appropriate statistical analysis and the results showed that the three highest potential benefits of using BIM in Iraq showed saving on the cost of the project, providing high quality and fast data documentation system was minimizing change orders.

Key words: Benefits, BIM, building information modeling, Iraq, providing high quality, documentation

INTRODUCTION

The construction industry is witnessing a paradigm shift that will achieve the highest productivity, efficiency, quality, sustainability and the value of infrastructure and reduce the costs of the life cycle as well reduce time (Arayici *et al.*, 2009). This is consistent with what (Azhar *et al.*, 2008) said about construction industry tends to adopt techniques that reduce the cost of the project, increase the productivity and quality of the project and reduce the project time.

One of these techniques is Building Information Modeling (BIM) which is a technological and procedural shift in the construction industry (Succar, 2009).

Building Information Modeling (BIM) is an advanced tool and process consisting of a combination of virtual features, systems and concepts facilitated within a unified environment (Azhar *et al.*, 2015). It includes the application and keeps integrated digital representation for different information across different project stages (Eastman *et al.*, 2011).

One of main benefits of BIM is the accurate geometrical representation for building parts within an integrated information environment (RCCI., 2007). The BIM as a new phenomenon seeks to renew the practices of the construction industry, so, it is subject to challenges facing its application (Kekana *et al.*, 2014).

For the implementation of BIM, this requires good planning and a coordinated approach with considerations for implementation and innovation management (Smith and Tardif, 2009).

The urgent need for the reconstruction of Iraq after the wars and the conditions that that accompanied with it made the construction industry the most important among the industries, however, it suffers from a series of complex problems that make the progress of projects difficult either in new projects or in the rehabilitation of existing projects. The fragment nature of the research and its negative impact on the project, the weakness in communication and information sharing among different stakeholders, the high cost of projects which is one of the biggest problems, the delays in projects and exceeding the prescribed period and in addition to other problems.

On the contrary, BIM has recently expanded and is able to solve most of the problems of the construction projects and despite the benefits they are not implemented in Iraq as in other countries of the world. Therefore, building a clear understanding of the concept of BIM by investigating the benefits it offers to the projects is important to promote the adoption of BIM in the construction projects of Iraq. Table 1 provide a summary of top BIM benefits according to several previous researchers.

Table 1: Top BIM benefits according to previous researchers

UK (Porwal and Hewage, 2013)	Countries from Europe (Bryde <i>et al.</i> , 2013)	China (Li <i>et al.</i> , 2014)
Limiting incidents and improving site safety Reduce change orders	Saving on the cost of the project Reduce the time of the project Facilitating the construction industry based on prefabricated building components Building maintenance management	Ability to simulate design model Better modeling for complex buildings Reduce design errors and rework Saving on the cost of the project Reduce the time of the project
New Zealand (Stanley and Thurnell, 2014) Calculate estimates directly from the model The ability to visualize the building Minimize rework	Ghana (Armah, 2015) Constructability improvement The ability to visualize the building Improve productivity Detection of clashes Improve the quality of the project	Palestine (Abuhamra, 2015) Provide cooperation between project stakeholders Improve quality of design Improving in sustainable design Study alternatives and support the decision-making process The ability to visualize the building
Italy (Ciribini <i>et al.</i>, 2016) Saving on the cost of the project Reduce the time of the project Ability to simulate design model	China (Doumbouya <i>et al.</i>, 2016) Provide cooperation between project stakeholders Decision-making support Coordination in the project management Reduce risks	Sudan (Osman, 2016) Saving on the cost of the project Detection of clashes Speed up design processes Optimization of construction processes
Brazil (Haraguchi, 2016) Ability to simulate design and construction digitally Project data are available to all stakeholders Provide cooperation between project stakeholders Reduce the time of the project Saving on the cost of the project	Indonesia (Chandra <i>et al.</i>, 2017) Saving on the cost of the project Enhance customer satisfaction Improving the business opportunities Reduce the time of the project Coordination in the management of project	Countries from middle East (Gerges <i>et al.</i>, 2017) Coordination and detection of clashes Collaboration among stakeholders Calculate estimates directly from the model Visualize and manage the sequence of activities through 4D dimension
China and UK (Zhou <i>et al.</i>, 2017) Reduced cost quality improvement Reduced project duration Improved safety	India (NIBT., 2017) Clients will increasingly insist on BIM adoption New collaborative way of working and sharing of information Reduction in costs and errors	UK (NBIML., 2017) Reduction in initial cost Reduction in the overall time Reduction in greenhouse gas emissions

MATERIALS AND METHODS

This study investigates the benefits of BIM application in Iraq. The data were collected by using field survey through design a special questionnaire for this targeted delivery it to the professionals working in the construction project field. Respondents illuminated their views in a set of points in the questionnaire where 300 copies of questionnaires were distributed but the total return was only 273 copies with a response rate 87.7% with 11 incompletes copies, so, the final number of copies is 262. The questionnaire built-in two main parts.

Part 1; Respondent information: This part comprises general information concerning the respondent.

Part 2 ; BIM protentional benefits: This part included a list of the benefits of BIM. This part contains 24 closed questions designed by the 5-Likert scale (Likert, 1932). The scale is (Extremely high benefit, high benefit, moderate benefit, low benefits, extremely low benefit). Each respondent was invited to give a degree of measure to each question according to what he/she believes within the environment of the Iraqi construction sector.

After the collection of the questionnaires, they were arranged, unloaded and analyzed using the Statistical Packaging for Social Science (SPSS) Software Version 24.

Statistical analysis: A set of quantitative statistical techniques have been used to analyze the questionnaire

data to obtain a comprehensive view of the opinions obtained from the specialists in Iraqi Architectural Engineering Construction (AEC) regarding the subject of BIM potential benefits. These statistical techniques will be clarified.

Descriptive statistics: Descriptive statistics can be defined as statistical methods aimed at summarizing data organizing data and simplifying data (Gravetter and Wallnau, 2016). Descriptive statistics that have been used are central tendency measurement and variability measurement.

Relative Importance Index (RII): One of the techniques used in the analysis of data is the Relative Importance Index RII and the purpose of its use is to give a rank for each item in a particular part in the questionnaire. The equations:

$$RII = \frac{\sum W}{(A \times N)} \tag{1}$$

$$RII = \frac{5(n_5)+4(n_4)+3(n_3)+2(n_2)+n_1}{5(n_5 + n_4 + n_3 + n_2 + n_1)} \tag{2}$$

Where:

W = The weight given by respondents for each component (ranging from 1-5)

A = The highest weight (which equals 5)

N = The total number of respondents

RESULTS AND DISCUSSION

General information for respondents: The demographic characteristics of target respondents shown in Table 2.

RII For BIM potential benefits: A value of RII for each item was calculated to obtain the rank of the single item within the rest of the items.

Table 3 shows the mean and standard deviation values for items and the resulting final ranks for each item. The results showed that “Saving on the cost of the project” is the highest potential benefit from BIM application according to the point of view of specialists with (RII = 87.863, mean = 1.61, SD = 0.777). This result is consistent with the Iraqi construction projects suffer from cost overruns due to several factors (Al-Ageeli and Alzobae, 2016) who mentioned these factors.

The potential BIM benefit which was in the second rank is “Provide high-quality and fast data documentation system” with RII = 87.023, mean = 1.67, SD = 0.786. This result can be considered acceptable according to the exposure of many projects after the war and security conditions to the loss of partial or complete documentation of the project in addition to the existence of problems related to the documentation system itself and on the other hand this result agree with (Abdul-Kareem, 2017) who recommended by the need to exist and apply Information Technology IT in the field of documentation in Iraqi construction projects to overcome these problems.

Table 2: The respondent’s profile

Information about categories	Percentage
Work sector	
Public sector	79
Private sector	14
Public and private together	7
Gender	
Male	69
Female	31
Age (years)	
20-30	25
31-40	51
41-50	18
More than 51	6
Academic qualification	
Diploma	3
Bachelor	80
Master	10
PhD	10.6
Other	1
Specialization	
Architect	4
Civil Engineering	52
Electrical Engineering	24
Mechanical Engineering	12
Other	8
Group (Job)	
Designer	14
Consultant	9
Project Manager	10
Site Engineer	50
Contractor	12
Other	5
Practical experience (years)	
<5	30
5-10	24
11-15	17
16-20	17
More than 20	12

Table 3: Rank of BIM potential benefits depending on RII, Mean, SD

No.	BIM potential benefits	Mean	SD	RII	Rank
B1	The ability to visualize the building with different details from different angles and better representation for complex buildings details	1.82	0.889	83.969	9
B2	Increased design efficiency by the ability to simulate the design model	1.85	0.866	83.130	10
B3	Digital virtual construction enables rapid analysis of various performance scenarios related building	1.96	0.888	81.450	14
B4	Link cost, time and other variables to the unified design model	1.74	0.848	85.344	6
B5	Facilitate cooperation between the different stockholder of the project such as the owner, designer, contractor and others to expand the area of interest in several aspects	1.73	0.860	85.725	5
B6	Improving the productivity of estimating the quantities of the project	1.77	0.817	84.733	7
B7	Support decision-making by providing a reliable database for all parts of the project	1.9	0.878	82.137	11
B8	Saving on the cost of the project	1.61	0.777	87.863	1
B9	The high flexibility offered by the BIM design, so that, the change in a particular location of the design followed by automatic change in all variables associated with it	1.97	0.918	80.687	17
B10	Detection of design errors and conflicts between different disciplines at early time and minimizing rework	1.7	0.822	85.954	4
B11	Enhance customer satisfaction	2.16	0.974	76.947	23
B12	Analysis of the building in terms of energy	2.12	1.001	77.634	22
B13	visualize the scheduling of the project this has a role in facilitating the coordination project processes and steps to reach an expected result	1.79	0.783	84.427	8
B14	Provide high-quality and fast data documentation system	1.67	0.786	87.023	2
B15	Minimize construction waste and manage it well	2.00	0.920	80.763	16
B16	Develop risk management and minimize risks as well as site safety planning and reduce accidents	2.10	0.991	77.934	21
B17	Improved construction communication management	2.17	0.947	76.718	24

Table 3: Continue

No.	BIM potential benefits	Mean	SD	RII	Rank
B18	Increase the delivery speed and finish of the project	1.92	0.925	82.061	12
B19	Easy and quick access to any details or sections of the building	2.00	1.041	79.924	18
B20	Increase productivity and quality of the project	1.93	0.914	81.832	13
B21	Minimize change orders	1.69	0.825	86.107	3
B22	Increased efficiency of procurement by processing accurate quantitative data as well as identify the plan and times of purchase	1.93	0.871	81.374	15
B23	Better maintenance management and technical accessories during the life of the building	2.06	0.929	78.779	20
B24	Enhanced location logistics plans and better site equipment management	2.03	0.960	79.313	19

The third potential benefit of BIM is “Minimize change orders” with RII = 86.107, mean = 1.69, SD = 0.825 that change orders have a role in increasing the cost and time of the project. This result is in line with (Nazar, 2014) findings of the benefits of BIM.

CONCLUSION

According to Iraqi professionals views the first benefits of using of BIM in Iraqi construction projects are saving on the cost of the project, providing high quality and fast data documentation system was minimizing change orders.

RECOMMENDATIONS

In order to increase the level of knowledge in BIM, it is recommended that many seminars and conferences be held which in consultation with the experts in the field of construction projects on the other hand to raise the skill level of engineers it is recommended to provide training programs in government centers and institutions. Building and supporting academic projects as well as encouraging researchers in the field of BIM which will facilitate the transfer of expertise and information from the world. The government must play a vital role by providing the main guidelines to institutions in its transition toward BIM.

REFERENCES

Abdul-Kareem, H.I., 2017. Investigation problems of documentation, updating information and following up construction project in Iraq. *J. Eng. Sustainable Dev.*, 21: 193-204.

Abuhamra, L.A., 2015. An investigation into Building Information Modeling (BIM) application in Architecture, Engineering and Construction (AEC) industry in Gaza strip. Master Thesis, Islamic University of Gaza (IUG), Gaza City.

Al-Ageeli, H.K. and A.S.J.A. Alzobae, 2016. The most influential factor on the stumble and failure of the governmental projects. *J. Eng.*, 22: 93-110.

Arayici, Y., F. Khosrowshahi, A.M. Ponting and S.A. Mihindu, 2009. Towards implementation of building information modelling in the construction industry. Proceedings of the 5th International Conference on Construction in the 21st Century (CITC-V) Collaboration and Integration in Engineering, Management and Technology, May 20-22, 2009, University of Salford, Istanbul, Turkey, pp: 29-34.

Armah, N.N.O., 2015. Assessing the benefits and barriers of the use of Building Information Modeling (BIM) in the Ghanaian building construction industry. Master Thesis, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana.

Azhar, S., A. Nadeem, Y.N.K. Mok and H.Y.B. Leung, 2008. Building Information Modeling (BIM): A new paradigm for visual interactive modeling and simulation for construction projects. Proceedings of the 1st International conference on Construction in Developing Countries, August 4-5, 2008, Karachi, Pakistan, pp: 435-446.

Azhar, S., M. Khalfan and T. Maqsood, 2015. Building Information Modelling (BIM): Now and beyond. *Constr. Econ. Build.*, 12: 15-28.

Bryde, D., M. Broquetas and J.M. Volm, 2013. The project benefits of Building Information Modelling (BIM). *Int. J. Project Manage.*, 31: 971-980.

CRCCI, 2007. Adopting BIM for facilities management: Solutions for managing the Sydney opera house. Cooperative Research Center for Construction Innovation, Brisbane, Australia.

Chandra, H.P., P. Nugraha and E.S. Putra, 2017. Building information modeling in the architecture-engineering construction project in Surabaya. *Procedia Eng.*, 171: 348-353.

Ciribini, A.L.C., S.M. Ventura and M. Paneroni, 2016. Implementation of an interoperable process to optimise design and construction phases of a residential building: A BIM pilot project. *Autom. Constr.*, 71: 62-73.

Doumbouya, L., G. Gao and C. Guan, 2016. Adoption of the Building Information Modeling (BIM) for construction project effectiveness: The review of BIM benefits. *Am. J. Civ. Eng. Archit.*, 4: 74-79.

- Eastman, C., P. Teicholz, R. Sacks and K. Liston, 2011. BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors. 2nd Edn., Johan Wiley & Son, Hoboken, New Jersey, ISBN:978-0-470-54137-1, Pages: 304.
- Gerges, M., S. Austin, M. Mayouf, O. Ahiakwo and M. Jaeger *et al.*, 2017. An investigation into the implementation of building information modeling in the middle East. *J. Inf. Technol. Constr.*, 22: 1-15.
- Gravetter, F.J. and L.B. Wallnau, 2016. Statistics for the Behavioral Sciences. Cengage Learning, Boston, Massachusetts, ISBN:978-1-305-50491-2, Pages: 725.
- Haraguchi, L., 2016. BIM implementation: A bibliographic study of the benefits and costs involved. *Civ. Eng. Archit. English Version*, 10: 755-761.
- Kekana, T.G., C.O. Aigbavboa and W.D. Thwala, 2014. Building Information Modelling (BIM): Barriers in adoption and implementation strategies in the South Africa construction industry. Proceedings of the International Conference on Emerging Trends in Computer and Image Processing (ICETCIP'14), December 15-16, 2014, Planetary Scientific Research Center, Pattaya, Thailand, pp: 15-16.
- Li, J., L. Hou, X. Wang, J. Wang and J. Guo *et al.*, 2014. A project-based quantification of BIM benefits. *Intl. J. Adv. Rob. Syst.*, 11: 123-136.
- Likert, R., 1932. A technique for the measurement of attitudes. *Arch. Psychol.*, 140: 1-55.
- NBIML, 2017. NBS international BIM report. The national BIM library, UK.
- NIBT., 2017. Use of BIM in Indian AEC industry. National Institute of Building Technology, Sharanpur, India.
- Nazar, N., 2014. Using the virtual building technology for quantity surveying of buildings. Master Thesis, University of Baghdad, Baghdad, Iraq.
- Osman, M.K., 2016. Implementation and potential benefits of Building Information Modeling (BIM) in Sudan construction industry. Master Thesis, Sudan University of Science and Technology, Khartoum, Sudan.
- Porwal, A. and K.N. Hewage, 2013. Building Information Modeling (BIM) partnering framework for public construction projects. *Autom. Constr.*, 31: 204-214.
- Smith, D.K. and M. Tardif, 2009. Building Information Modeling: A Strategic Implementation Guide for Architects, Engineers, Constructors and Real Estate Asset Managers. John Wiley & Sons, Hoboken, New Jersey, USA., ISBN:978-0-470-25003-7, Pages: 183.
- Stanley, R. and D.P. Thurnell, 2014. The benefits of, and barriers to, implementation of 5D BIM for quantity surveying in New Zealand. *Constr. Econ. Build.*, 14: 105-117.
- Succar, B., 2009. Building information modelling framework: A research and delivery foundation for industry stakeholders. *Autom. Constr.*, 18: 357-375.
- Zhou, Y., L. Ding, Y. Rao, H. Luo and B. Medjdoub *et al.*, 2017. Formulating project-level building information modeling evaluation framework from the perspectives of organizations: A review. *Autom. Constr.*, 81: 44-55.

Wireless Alarming System for Ketosis Proactive Detection in Dairy Cattle

¹Hussain A. Attia and ²Halah Y. Ali

¹Department of Electrical, Electronics and Communications Engineering,

²Department of Biotechnology, American University of Ras Al Khaimah,
Ras Al Khaimah, UAE

Abstract: Ketosis or acetonemia is a common metabolic disease of adult cattle which bred and managed for high production during early lactation period and late gestation. It mainly occurs in dairy cattle that stop using carbohydrates as a source of energy, instead depending on body fats for fuel to meet their energy requirements of lactation. Mobilization of body fat to the liver releases a high rate of ketone bodies. These bodies diffuse across the cell membrane to supply the body with alternative sources of energy to meet their needs of lactation. These ketones are acetone, acetoacetate and β -HydroxyButyrate BHBA, the body gets rid of them in the urine, fresh milk and exhaled breath. One of the best prevention methods is to detect the acetone level in exhaled breath during Subclinical Ketosis SCK before appearance of clinical signs of ketosis. Based on the above, this study proposed a new design of remotely detection system for ketosis proactive in dairy cattle. When the acetone in exhaled breath is exceeding a certain level from an identified cow, SMS message will be sent automatically to the responsible person. This message includes the cow's RFID reader tag number and its high Acetone level. The proposed system includes Global System for Mobile communication (GSM) with Subscriber Identity Module (SIM) card connected to the arduino unit to allow SMS message sending. Fast and accurate response with ability of remotely monitoring and messaging is the main merits of the presented system.

Key words: Ketosis, dairy cattle, acetone sensor, adaptation circuit, arduino UNO, GSM, SIM, SMS, flowchart, investigation

INTRODUCTION

Under normal conditions, there is low concentration of ketone bodies in the blood and exhaled breath but in times of stress (e.g., transition period coupled with starvation) dairy cows are forced to mobilize body fat to meet their glucose needs. An elevated blood β -HydroxyButyrate (BHBA) ≥ 1.2 mmol/L leading to Sub Clinical Ketosis (SCK) that are characterized by absence of clinical signs. Thus, veterinarians and farmers have difficulty finding the affected animal (Elliott-Martin *et al.*, 1997).

Determination of acetone level in exhaled breath was employed as the indicator in the early diagnosis of SCK as it is considered an important endogenous byproduct of the Negative Energy Balance (NEB) and has special odor which can diffuse through the pulmonary alveolar membrane that separates the air in the alveoli from the blood because it is a very thin barrier. This breath detection might offer advantages over the current urine or blood analysis, especially for the farmer and the animal itself (Dobbelaar *et al.*, 1996).

Many research studies are focused on analyzing the samples of breath and offering a desired action based on

analysis result to the parameter under test, A developed device is proposed for nasal breath sampling in dairy cattle (Elliott-Martin *et al.*, 1997), this study showed the possibility of recognizing between the healthy cows and cows with SCK. A multi-functions monitoring and protection system design is proposed by Attia (2016) whereas a new system is proposed by Attia and Ali (2016) for continuous monitoring and protection activities for any gas leakage using a certain type of gas sensor. The study of Dobbelaar *et al.* (1996) focused on analyzing the levels of acetone concentration in cattle through selecting four healthy cows based on non-invasive method. The study by Weng *et al.* (2015) focused on evaluating the level of alcohol in a sample of breath that offers a suitable protection step against the driver who exceeds the acceptable alcohol level at which the driver is still conscious with quick driving reactions. Different system methodologies are proposed by Attia *et al.* (2017), Attia *et al.* (2016a-c), Attia and Getu (2015), Getu and Attia (2015) and Takruri *et al.* (2016) in which the process of the wirelessly monitoring for the related parameters is adopted to have correct decisions.

Based on the mentioned above of the different methodologies for acetone, alcohol and gas monitoring

systems, beside the merits of the non-incentive testing as well as the early high acetone detection with wirelessly messaging ability, this study proposed a detailed electronic design with implementation steps and performance investigation of a new remotely individual detection system for the case of high acetone level cattle.

The proposed system is working on monitoring the SCK and sending individual SMS message to the responsible person about each case of high acetone level. The presented system includes a certain acetone sensor, a suitable adaptation circuit, Global System for Mobile communication (GSM) type SIM800L, RFID reader type RC522 and arduino UNO unit.

MATERIALS AND METHODS

Early detection advantages of acetoneemia in cattle: One third of dairy cows may be affected by some form of metabolic disorders during the transition period, dry period or shortly after calving. Ketosis is considered one of the important and common disorders that can act as a primary disease or as a risk factor for other serious diseases such as mastitis, metritis, retained placenta and displaced abomasum. Ketosis is preceded by SCK before the onset of its symptoms which is characterized by a high concentration of β -HydroxyButyrate (BHBA) ≥ 1.2 mmol/L which equivalent to 2 part per million ppm of acetone in exhaled breath while concentrations that reflect clinical ketosis is ≥ 3.0 mmol/L.

Monitoring strategies of the livestock especially during these critical periods consider as proactive action to provide opportunity for investigation and fast intervention to avoid the consequences and the economic losses due to treatment, reduced performance or death. In addition, the treatment of related infectious diseases or metabolic disorders is seldom successful (Dobbelaar *et al.*, 1996). The proposed system enables the veterinarian and even the herdsman to diagnose the herd (closely and remotely) without need for the difficult and unavailable traditional methods that are either time consuming or requiring expensive equipment and skilled personnel while the proposed system is rapid, easy, low-cost and non-invasive (Weng *et al.*, 2015).

Proactive detection system discription: Figure 1 shows the main block diagram of the proposed system. The detection block is represented by a specialized acetone detector; The selected detector for this function is TGS 822. Detector's location environment condition and the period of taking enough breath samples should be

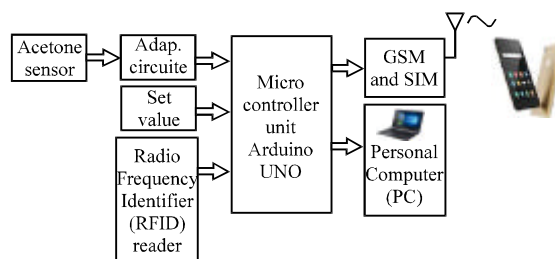


Fig. 1: Block diagram of the proactive detection system

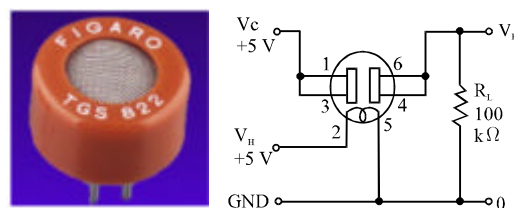


Fig. 2: Adopted acetone sensor; Physical view and the electrical circuit

considered for ideal testing conditions. An adaptation circuit is designed and simulated to have a suitable range of voltage variation with respect to the detected acetone level. The adapted voltage is entered to the microcontroller unit (Arduino Uno) in addition to the information of the Radio Frequency Identification (RFID) reader. RFID reader is designed in the proposed system to recognize each cow individually by fixing the RFID tag at suitable location of cow body based on the physical arrangement of the barn. The acetone level after detection and voltage adaptation will be evaluated and compared with the set value of acetone limit (2 ppm). The cattle have more than this limit of acetone that will be detected as proactive ketosis indication. The flowchart of the designed program will be explained in the next section. Based on the program process when the detected acetone level of an identified cow is more than the set value, SMS message will be sent via GSM card to the responsible person, this message includes the alarm states that the identified cow has high acetone level. This test and SMS message can be repeated for any cow which has another RFID tag. More details and the history of all messages are recorded and can be shown through the connected personal computer.

RESULTS AND DISCUSSION

Acetone sensor and adaptation circuit: The adopted sensor for acetone detection is figaro TGS 822. Figure 2 shows the physical view and the electrical circuit of the selected acetone sensor. Figure 3 shows sensor behavior

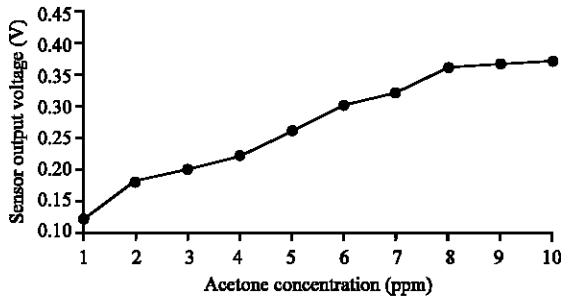


Fig. 3: Acetone sensor behavior of different acetone concentration

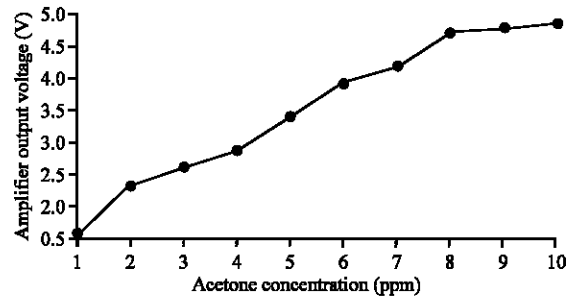


Fig. 5: Amplifier behavior of different acetone concentration

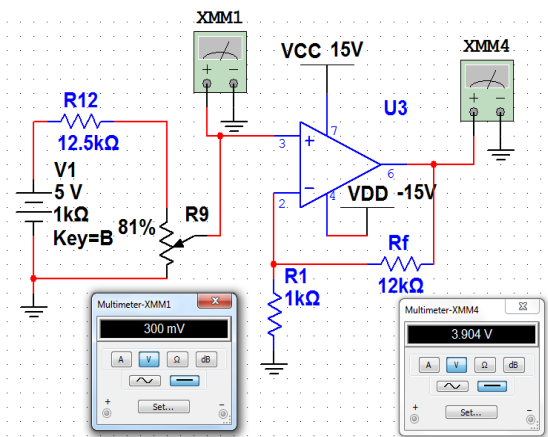


Fig. 4: Simulation of the amplifier with result sample

at normal humidity through the testing records of this sensor which is considered from (Thati *et al.*, 2015). As shown in Fig. 3, the sensor output voltage is varying from 0.12-0.37 V for acetone concentration from 1-10 ppm of a breath sample. Based on the literature studies (Weng *et al.*, 2015; Rydosz, 2015), the acetone concentration more than 2 ppm in exhaled breath sample is evaluated as high acetone level indication. Due to the sensor output voltage is in a low voltage range to have effective monitoring system, this range should be adapted for easy recognizing by microcontroller unit.

The designed adaptation circuit in this study is represented by amplifier circuit with amplification gain of 13; Fig. 4 shows the simulation and a certain result sample of the designed amplifier. Table 1 shows simulation results for different input voltages with respect to related acetone concentration level in ppm whereas, Fig. 5 shows the behavior of the amplifier circuit for range of acetone concentration. Considering the collected results, the user can set the limit of acetone at which the dairy cattle addressed as cow has high acetone level.



Fig. 6: RFID reader type RC522 RFID module

Table 1: Simulation results for different input voltages at different Acetone concentration

Acetone cons. (ppm)	Sensor V_o (V)	Amplifier ($g=13$) V_o (V)
1	0.120	1.560
2	0.180	2.340
3	0.200	2.600
4	0.220	2.860
5	0.260	3.380
6	0.300	3.900
7	0.320	4.160
8	0.360	4.680
9	0.365	4.745
10	0.370	4.810

Radio Frequency Identification RFID reader: The module type of RFID reader is RFID RC522 in which the integrated circuit is working on identifying the ID tags via contact-less communication process. Figure 6 shows the selected RFID reader with two ID tags of different shapes. This module is working by 3.3 V, the operating frequency 13.56 MHz and the operating temperature range is from 40-85°C (Attia and Takturri, 2016).

Arduino Uno: The microcontroller unit which is selected to process and control the proposed system activities is Arduino Uno. Figure 7 shows the terminals arrangement of the unit which includes different types of input and output terminals; digital terminals which can be programmed as input or output. Other terminals are

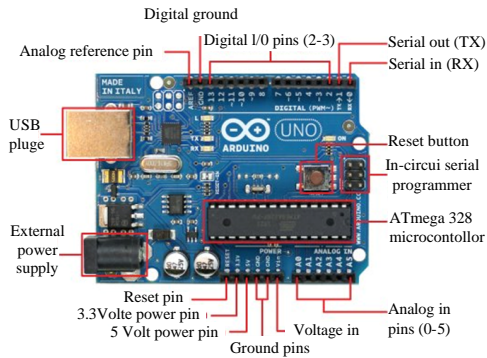


Fig. 7: Micro controlling unit (Arduino Uno)

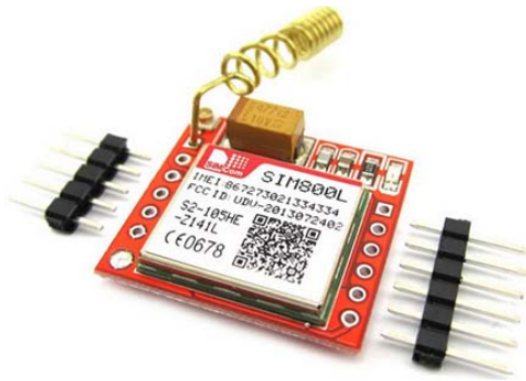


Fig. 8: Global System for Mobile Communications (GSM) module type SIM800L GSM Module

dealing with analog voltages of 0-5V range: these terminals can be programmed to work in input or output mode. About 10 bits is the accuracy of the conversion (Boxall, 2013).

Two analog voltages are entered to the Arduino unit: first analog voltage enters through (A1) and represents the level of monitored acetone. The second analog voltage enters through (A2), it represents the set value of the acetone limit.

Global System for Mobile communications (GSM): Global System for Mobile Communication (GSM) is a standard developed set of 2G and 3G mobile generations which is developed by European Telecommunications Standards Institute (ETSI) (Pavithra and Srinath, 2014). Figure 8 shows the physical shape of the selected GSM card which module is SIM800L. The function of this card is sending and receiving (messaging) the required SMS message between the arduino unit and the responsible person's mobile.

Flowchart of the proactive detection system program: The flowchart of the proposed system program which is

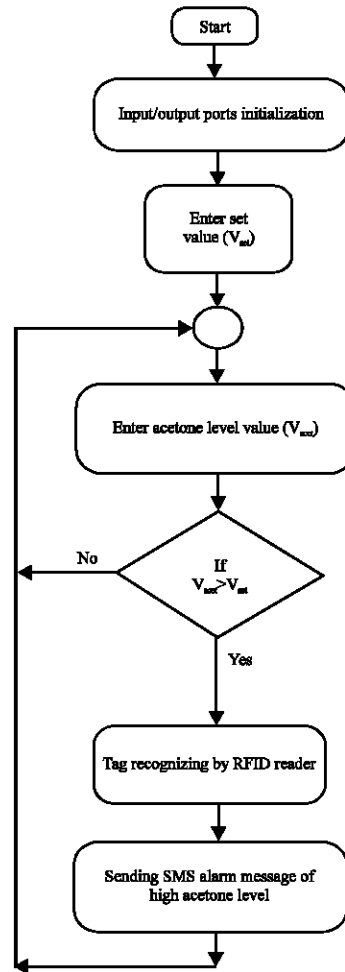


Fig. 9: Flowchart of the proposed diagnostic system

uploaded in Arduino Uno unit is shown in Fig. 9 as shown in the flowchart, the process is started by initializing the ports of the unit based on the required actions. The next step is entering the setting value of acetone limit at which or more the breath sample is considered as it has SCK indication. The setting of this value is affected by the physical arrangement of the test location and this value should be adjusted by expertized person.

The adapted voltage of the sensed acetone level is entered to Arduino unit as a second analog voltage: this voltage will be compared with setting value. The comparison result determines the next step. The result is either the acetone level which is lower than the setting value or greater. In first option, other test will be checked, whereas in second option, RFID reader is activated to identify the tag number of the cow which has high acetone level, then a certain SMS message will be sent automatically via. the SIM800L GSM to the responsible

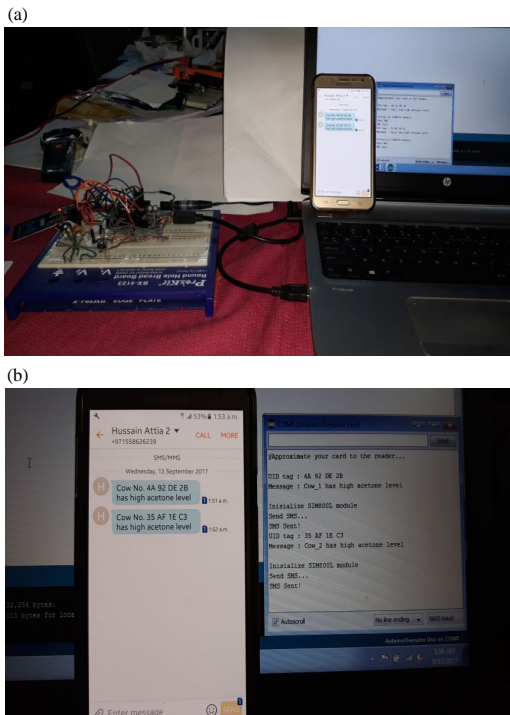


Fig. 10: System implementation: a) Whole system and b) Mobile and personal computer screens zoom in

person's mobile. This message includes the tag number of the cow under test in addition to alarm stating that this cow has high acetone level.

This program process will be repeated for other cases considering that each case has special tag number. For history recording, a personal computer is required. Figure 10 shows the implemented system in addition to the investigation of a two cases of different RFID tag numbers.

CONCLUSION

By considering the importance of the proactive detection of the ketosis in dairy cattle and to take the quick necessary actions, a new remotely monitoring and alarming system design and implementation is done in this study for early detection of SCK in dairy cattle. Detailed design steps of the proposed system are explained and shown including RFID reader, SIM800L GSM with Subscriber Identity Module (SIM) card and Arduino Uno unit as controller. Personal computer is included in the system for data saving ability for all of previous cases. The proposed system is characterized by fast automatically messaging to the responsible side. The presented research offers a promising choice in the field of proactive ketosis detection in dairy cattle.

ACKNOWLEDGEMENTS

The researchers are appreciating the financial support of American university of Ras Al Khaimah (AURAK) and each of School of Engineering and School of Graduate Studies and Research at AURAK.

REFERENCES

- Attia, H., 2016. Multi-functional microcontroller based monitoring alarming and protection system. *Intl. J. Appl. Eng. Res.*, 11: 9491-9496.
- Attia, H., A. Omar, M. Takturi and H.Y. Ali, 2017. Pulse width modulation based decentralized street LED light dimming system. *Intl. J. Power Electron. Drive Syst.*, 8: 1220-1229.
- Attia, H.A. and B.N. Getu, 2015. Design and simulation of remotely power controller. *Intl. J. Appl. Eng. Res.*, 10: 32609-32626.
- Attia, H.A. and H.Y. Ali, 2016. Electronic design of liquefied petroleum gas leakage monitoring, alarm and protection system based on discrete components. *Intl. J. Appl. Eng. Res.*, 11: 9721-9726.
- Attia, H.A. and M. Takturi, 2016. RFID code wireless transceiver system based on amplitude shift keying modulation. *Intl. J. Appl. Eng. Res.*, 11: 7155-7158.
- Attia, H.A., A. Omar and M. Takturi, 2016a. Design of decentralized street LED light dimming system. *Proceedings of the 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, December 6-8, 2016, IEEE, Ras Al-Khaimah, UAE., ISBN:978-1-5090-5306-3, pp: 1-4.
- Attia, H.A., M. Takturi and H.Y. Ali, 2016b. Electronic monitoring and protection system for drunk driver based on breath sample testing. *Proceedings of the 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, December 6-8, 2016, IEEE, Ras Al-Khaimah, UAE., ISBN:978-1-5090-5306-3, pp: 1-4.
- Attia, H.A., S. Ismail and H.Y. Ali, 2016c. Vehicle safety distance alarming system. *Proceedings of the 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, December 6-8, 2016, IEEE, Ras Al-Khaimah, UAE., ISBN:978-1-5090-5306-3, pp: 1-4.
- Boxall, J., 2013. *Arduino Workshop: A Hands on Introduction With 65 Projects*. No Starch Press, San Francisco, California, USA., Pages: 372.

- Dobbelaar, P., T. Mottram, C. Nyabadza, P. Hobbs and R.J. Elliott-Martin *et al.*, 1996. Detection of ketosis in dairy cows by analysis of exhaled breath. *Vet. Q.*, 18: 151-152.
- Elliott-Martin, R.J., T.T. Mottram, J.W. Gardner, P.J. Hobbs *et al.*, 1997. Preliminary investigation of breath sampling as a monitor of health in dairy cattle. *J. Agric. Eng. Res.*, 67: 267-275.
- Getu, B.N. and H.A. Attia, 2015. Remote controlling of light intensity using phone devices. *Res. J. Appl. Sci. Eng. Technol.*, 10: 1206-1215.
- Pavithra, D.S. and M.S. Srinath, 2014. GSM based automatic irrigation control system for efficient use of resources and crop planning by using an Android mobile. *IOSR. J. Mech. Civil Eng.*, 11: 49-55.
- Rydosz, A., 2015. Microsystem in LTCC technology to the detection of acetone in exhaled breath. *Intl. J. Inf. Electron. Eng.*, 5: 98-101.
- Takruri, M., H.A. Attia and M.H. Awad, 2016. Wireless charging in the context of sensor networks. *Intl. J. Appl. Eng. Res.*, 11: 7736-7741.
- Thati, A., A. Biswas, S.R. Chowdhury and T.K. Sau, 2015. Breath acetone-based non-invasive detection of blood glucose levels. *Intl. J. Smart Sens. Intell. Syst.*, 8: 1244-1260.
- Weng, X., W. Zhao, S. Neethirajan and T. Duffield, 2015. Microfluidic biosensor for β -Hydroxybutyrate (BHBA) determination of subclinical ketosis diagnosis. *J. Nanobiotechnol.*, 13: 1-8.

Petrophysical Analysis on Radioactive Sands for Koala Field in Termit Basin, Niger

Muhammad Syafiq Zainal Malek, Lo Shyh-Zung and Abdul Ghani Md Rafek
Department of Petroleum Geosciencet, Universiti Teknologi PETRONAS (UTP),
40150 Seri Iskandar, Perak, Malaysia

Abstract: Investigations in Niger delta adjacent to Termit basin, Niger has widespread occurrence of radioactive sands. About 34-45% of producing wells contains this anomaly which is misinterpreted as shales in petrophysical analysis, specifically from Gamma Ray (GR) log. This causes potential pay zones to be overlooked unless reevaluation of conventional logs is done. A four-step approach is established to define existence of radioactive sands and identifying source of radioactivity based only on conventional wireline logs. Targeted radioactive sands for the research are hydrocarbon bearing. Methods starts with identifying mudcakes in shale lithology layers detected by GR log and cross-referencing with neutron-density (NPHI-RHOB) crossover. By interpretation of NPHI-RHOB crossover and identifying high values in resistivity log, hydrocarbon zones are identified focussing specifically on shale layers with mudcake. Hydrocarbon zones detected as shales contradicts with interpretation as mudcake only sticks on porous and permeable layers specifically on sandstone not shale. To confirm interpreted radioactive sands, NPHI vs. RHOB cross plot is done to identify lithology. This is followed by radioactive source identification by evaluating spectral gamma ray to determine potential source. Lastly, thorium vs. potassium cross plot is used to identify specified minerals present which contributes to high radioactivity. Approximately 25-30 m of accumulated pay zones were found within all five wells. Methods mentioned is to re-evaluate wells in the system without any additional logging tools. The study is to facilitate the development and improve interpretative procedures that might lead to cost-effective re-completion of hidden reservoirs within the same producing system.

Key words: Reservoir characterization, radioactive sands, Termit basin, procedures, reservoirs, hidden

INTRODUCTION

Conventional interpretation based on petrophysics analysis plays an important role especially in reservoir volume calculation. Volume estimation would affect in choosing potential prospect for development. A problem related to volume estimation is misinterpretation of pay zone thickness.

The neighbouring region of Termit basin is the Niger delta which is located on the southern section of Nigeria. Several investigations were carried out from 1971-2012 which confirms that there is wide spread occurrence of radioactive sands in the Niger delta (Chudi and Simon, 2012). Amongst the wells drilled 55-70% are hydrocarbon bearing zones. Out of the producing wells, 35-45% are believed to be radioactive with thickness ranging from 3-10 m (Chudi and Simon, 2012). The mentioned literature, highlights the importance of identifying radioactive sand layers.

Radioactive sands which have potential as reservoirs are overlooked as shale using Gamma Ray (GR) logs.

Presence of radioactive minerals and clays would give high radioactivity leading the GR log to detect these sands as shale. This causes many potential sandstone reservoirs to be overlooked. From the statement, two main objectives can be set. First is to define the presence of radioactive sands in well log and second is to identify the source contributing to radioactivity.

The best way to identify radioactive sands is by analysing thin sections but in the scenario where only well logs are available, a four-step approach can be used. The paper presents examples of reservoir characterization on radioactive sands from Koala field, Termit basin of Niger which includes 5 wells. It discusses the methodology using analysis of only conventional well logs.

Geological setting

Regional geology: The mesozoic plate tectonic has contributed to the development of the West and Central African Rift System (WCARS) during the opening of the Atlantic Ocean via the benue trough and shear zones

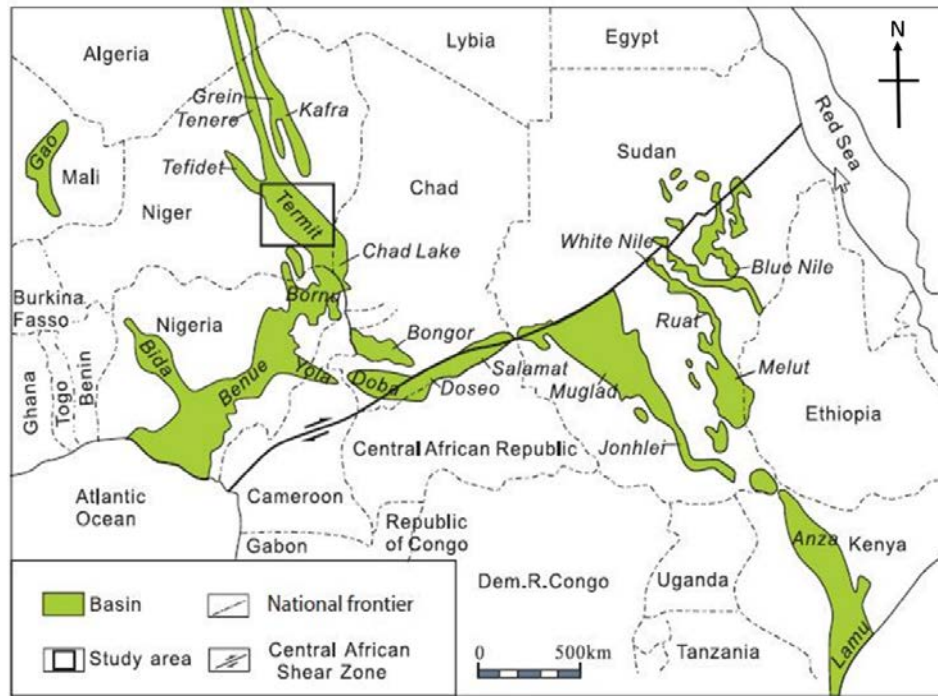


Fig. 1: Location of Termit basin and tectonic settings in the Central-Western African rift system (Liu *et al.*, 2015)

cutting Cameroon (Fairhead *et al.*, 2013). Accordingly, the Cretaceous-Paleogene rifts of Niger, Chad and the Central African rift (C.A.R.) resulted in large part of a geotectonic continuum of the WCARS (Fairhead *et al.*, 2013; Liu *et al.*, 2012). The rift extends to a length of 4000 km which starts from the Gao Trough in Mali to the Anza basin in Kenya (Cratchley *et al.*, 1984). WCARS system is subdivided into two coeval Cretaceous genetically related but physically separated rift subsystems which are the West African rift Subsystem (WAS) and Central African rift Subsystem (CAS) shown in Fig. 1. The main difference between the two subsystems is that WAS is filled by thick sequences of non-marine and marine Cretaceous to Tertiary sediments, however, the CAS is filled by virtually all cretaceous continental sediments (Genik, 1993).

Termit basin: Agadem block is part of the Termit basin and koala field is in the eastern flank of the block. According to the Agadem Field Development Plan (FDP) report, hydrocarbon in Niger is in two large sedimentary basins which are Lullemeden basin in the west and the Eastern (Chad) basin in the east which covers 90% of the national territory.

Termit basin basement is Precambrian while the other components: Donga, Yogou, Madama is Upper Cretaceous, Sokor-1 and 2 of Paleogene and Neogene+Quaternary (N+Q). The lower Cretaceous

thickness is approximately 5000 m and was continental coarse fan delta, underwater fan and lacustrine facies. However, Upper Cretaceous which composes of Donga and Yogou formation was of marine environment while Madama was under the continental environment (Genik, 1992, 1992). This was proven by the lithologies found. Figure 2 shows the geological profile of formations found in Termit basin.

Termit basin is individualized during early Cretaceous and belongs to the northern intracration of the Western African Rift Subsystem (WARS) (Fairhead *et al.*, 2013). According to well report, it has been proven that an abundant oil-generation kitchen by drilling was discovered in seven fields surrounding the Koala field area. The oil and gas discovered have been made in the Eocene reservoirs on tilted fault blocks and uplifted horst blocks in the Agadem block, the central portion of Termit basin. Neponoto Regency capitalized in Bontosunggu has wide are or 837,99 km², and divided into 113 village and 11 subdistricts. Jeneponto Regency has main commodities such as agriculture, plantation, agriculture, fishery, livestock and service sector. Main commodities of agriculture are corn, soybean, potato, banana, pineapple, sweet potato and cassava. Main commodities of fishery are fishery catch, fresh water pond culture, sea culture and brackish cultivation. Main commodities of livestock are cow, sheep, goat, buffalo and horse. Main commodity of service are culture and nature tourism.

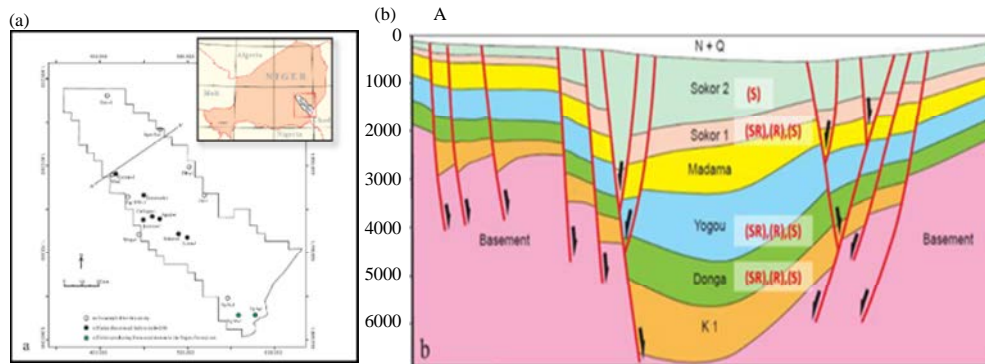


Fig. 2: a) Map of the Agadem permit area in the Termit basin, SE Niger and b) The geological profile of Termit basin (Bottom) (S = Sseal, R = Reservoir, SR = Source Rock) (Liu *et al.*, 2012)

Petroleum system: Termit basin petroleum system is in the in upper Cretaceous and Paleogene (Fairhead *et al.*, 2013; Genik, 1993). The formation ranges involved are Yogou to Sokor-1 and -2 systems. Eocene and Upper Cretaceous sandstones are the main reservoir intervals found in Termit basin. The main oil discoveries occur mainly in the Eocene reservoirs while the source rock and the seal are from the upper Cretaceous and the Eocene marine and lacustrine shales (Genik, 1993; Harouna and Philp, 2012).

There are two main oil groups classified in WCARS The first is derived from marine-paralic source while the second is from lacustrine source. Unique geological features are established to have 5 types which are source rocks, reservoir rocks, regional caps with controlling factors, major play types and favourable accumulation zones in passive rift basin (Fairhead *et al.*, 2013). The petroleum geology of Termit Basin (Wan *et al.*, 2014) are made of:

- Reservoir rock: upper member of Yogou Formation and sandstones from upper Paleogene Sokor-1 Formation (dominant)
- Source rock: principal source rock are in Yogou formation and Lacustrine mudstone in Sokor-1 formation
- Seal: regional top seal from Sokor-2 paleogene formation and intraformational mudstone seal from upper member of Yogou formation

Radioactive sands occurrences: Gamma Ray (GR) log responds based on the concentration of radioactive materials contained in the rock adjacent to the borehole. Enough energy is emitted by the rock for the GR tool to detect. The three-main usage of GR log is well correlation, lithological identification and as shaliness indicator. GR has high affinity for radioactive minerals like Thorium

(Th), Uranium (Ur) and Potassium (K40) which are contained in shale and flag off as high radioactive counts in the log while sands is vice-versa. In the case of sediments originating from nearby granitic highland, sands and gravels found usually consist of high concentration zircon or maybe highly arkosic (high feldspar content) with high potassium content. From the explanation, the question that arises is whether natural GR log can be used as a sole discriminator for sand/shale or reservoir/non-reservoir and how it impacts in choosing reservoir tops or base for net pay calculation. In some cases, clean sandstone can produce high GR response if the sand contains potassium feldspar, glauconite, micas or uranium rich water which is known to be radioactive sands or hard sands. The challenge present in this case is how net sand is defined and what constitutes valid criteria for differentiating net reservoirs from non-reservoirs. Identifying radioactive sands can increase potential pay zones which is important for prospect evaluation.

The aim of research: This study conducted research based only on conventional logs with the objectives of defining existence of radioactive sandstone and identifying the source of radioactivity.

MATERIALS AND METHODS

The methodology conducted in the study area consist of Field Development Plan (FDP) report, 3D seismic data and conventional wireline logs. For this evaluation only the FDP reports and wireline logs were used. Software used was Interactive Petrophysics (IP) by LR synergy for petrophysics analysis. The following steps were carried out for the study:

Phase 1: Interpreting and analysing data of the study area based on the regional geology with petrophysics analysis

and formation identification based on the stratigraphic column. Termit basin basement is Precambrian while the other components: Donga, Yogou, Madama, Sokor-1, Sokor-2 and Neogene+Quaternary (N+Q) is Lower Cretaceous. The focused formation of study is on Sokor-1 formation which has complete petroleum system.

Phase 2: Evaluation of radioactive sands based on conventional petrophysics data from the wireline logs.

M1: Comparison of gamma ray log with calliper log along with density and neutron density logs. This is to identify presence of mudcake and sections with hydrocarbon as well as determining the facies.

M2: Analysing assumed layers through neutron-density (NPHI vs. RHOB) cross plot to confirm defined layers of radioactive sands.

M3: Interpretation of source of radioactivity using spectral gamma ray (uranium, thorium and potassium). Observation on the increase of each individual spectrum indicates different causes of radioactivity.

M4: Identifying contributing minerals using potassium vs. thorium cross plots (Spectral Gamma Ray (SGR) components) to determine the specific mineral causing radioactivity in the sands.

RESULTS AND DISCUSSION

Sandstone interval identification: Interpretation of the sandstone interval was identified from three major traits which are from Gamma Ray (GR) log, resistivity log and neutron-density log. In the GR log, American Petroleum Institute gravity (API) range set for sandstone is 0-75 where 2 types of lithology were set which are clean sandstone (0-40 API) and fine sandstone (40-75 API). Siltstone was set at a range of 75-100 API while shale ranges from 100-150 API. Resistivity of the layers will indicate the potential content of the respective layers whereby high resistivity usually indicates oil or gas (hydrocarbon). In addition, the gap between deep resistivity log and shallow resistivity log will indicate the permeability of identified layers (the bigger the gap, the higher the permeability of the layer). Comparison between the Neutron density Hydrogen Index (NPHI) log as well as the density (RHOB) log indicates low values in both fields for sandstone while shale would have high values. Interpretation of sandstone interval is shown in Fig. 3.

Radioactive sandstone interval identification: M1 comparison of gamma ray log with calliper log along with density and neutron density logs. Initial interpretation of radioactive sands which are hydrocarbon bearing was seen in the three major aspects. An indicator is from gamma ray log respond where it suggests shale lithology but is seen to have presence of mudcake detected by the

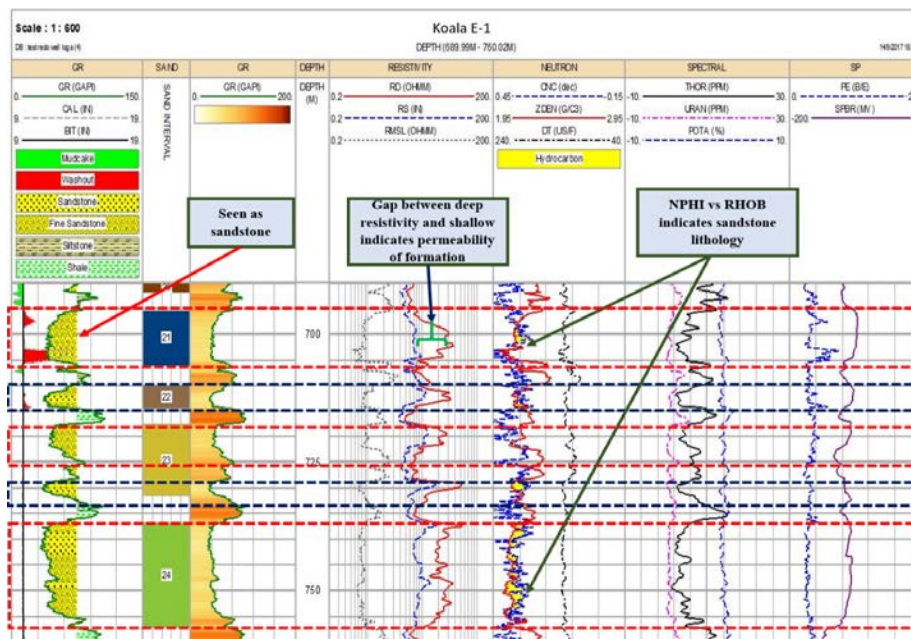


Fig. 3: Example of chosen sandstone interval based on Koala E-1 well

calliper log. Mudcake may indicate formation swell and flow (swelling of shales) into borehole or presence of porous and permeable formations which is a characteristic of sandstone.

To further confirm this analysis, observation from the crossing of NPHI log and RHOB log is used. Cross over may either be hydrocarbon presence or lithological differences of scaling effect. Besides that, NPHI and RHOB log will have low value in sandstone layers. To further confirm presence of hydrocarbon zone, the resistivity log illustrates high value together with obvious separation between shallow and deep resistivity log (high permeability). These three indicators give the first evidence for defining radioactive sands with hydrocarbon. Figure 4 shows examples of radioactive sand interpretation for two zones of Koala CE-1 well. However, zone 7 shows more promising in hydrocarbon bearing potential as the permeability is much higher than zone 6.

M2; Analysing assumed layers through neutron-density (NPHI vs. RHOB) cross plot: The neutron-density cross plot function used for the interpretation is Schlumberger Neutron-Density Raw Rhob = 1.1 mode which includes three lithology reference lines which are sandstone (quartzite) line, limestone line and also the dolomite line.

From Fig. 5, the interpreted radioactive sand layers from the previous section is plotted to further confirm that it is sandstone. Plotted points that fall on the sandstone line verifies that the layers are radioactive sands and not shale. Zones where plots do not fall on the sandstone line are automatically removed as they do not fulfil the characteristics of sandstone. In addition, the referenced lithology lines do include the range of porosity for the zones, in the example it is seen that the interpreted radioactive sand zones range from 25-45% porosity.

M3; Interpretation of source of radioactivity using spectral gamma ray (Uranium, Thorium and Potassium): After conforming the radioactive sands, spectral gamma ray analysis is studied to identify the source of -1 radioactivity. Three common naturally occurring radioactive elements in rocks are uranium, potassium and thorium. Analysis of spectral gamma ray which consist of the three mentioned spectrums each has different indication of the source of radioactivity. Increase in Thorium reading as seen in Fig. 6 indicates the presence of clay minerals and/or heavy minerals, e.g., kaolinite, illite, smectite and chlorite for clays and monazite and/or zircon (found near granitic highlands in unconformities or erosional surfaces) for heavy minerals. Uranium increment is associated with phosphates and organic compounds (shales, plants, shell fragments, euxinic environments). An

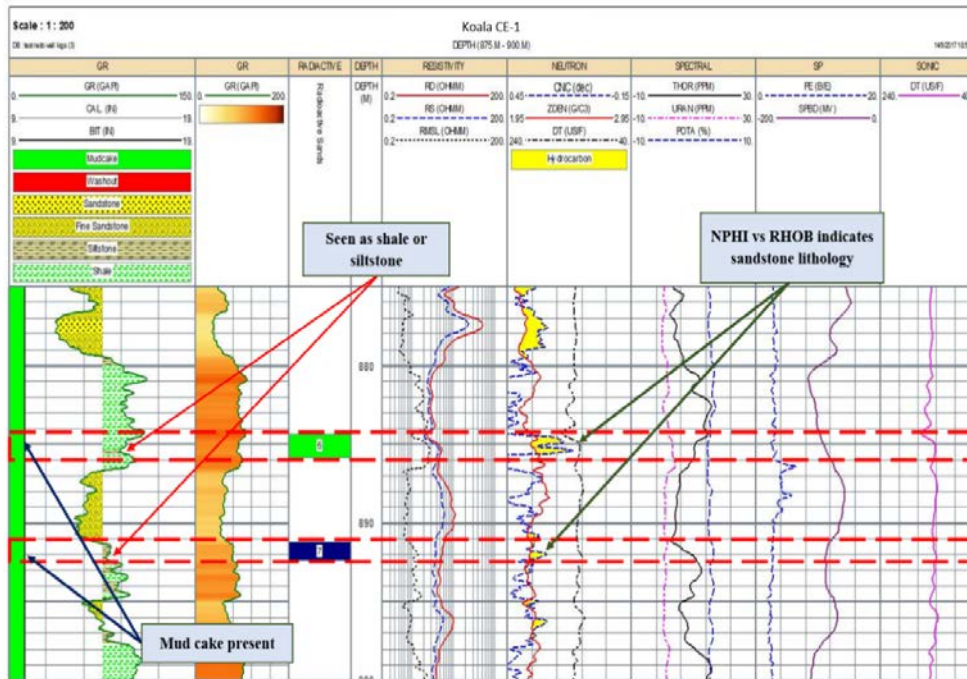


Fig. 4: Radioactive sands interpretation based on present of mudcake with comparison of gamma ray results and NPHI-RHOB crossing in Koala CE

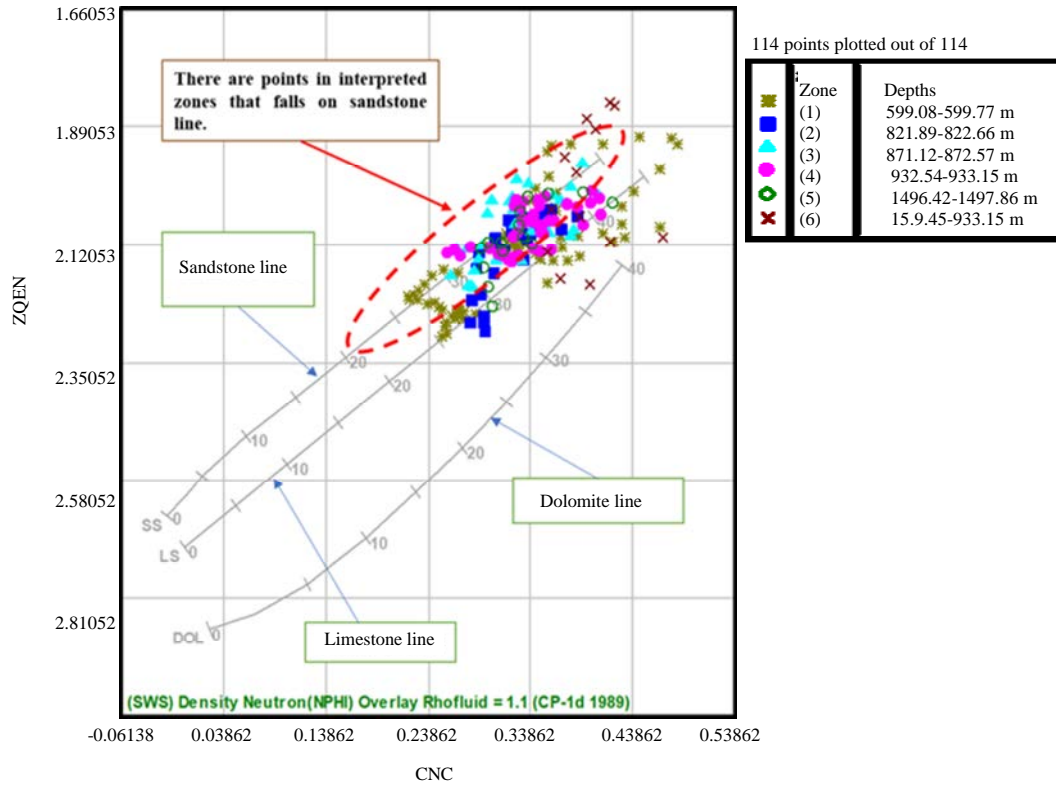


Fig. 5: NPHI vs. RHOB cross plot indicating the interpreted radioactive sands as having characteristics of sandstone

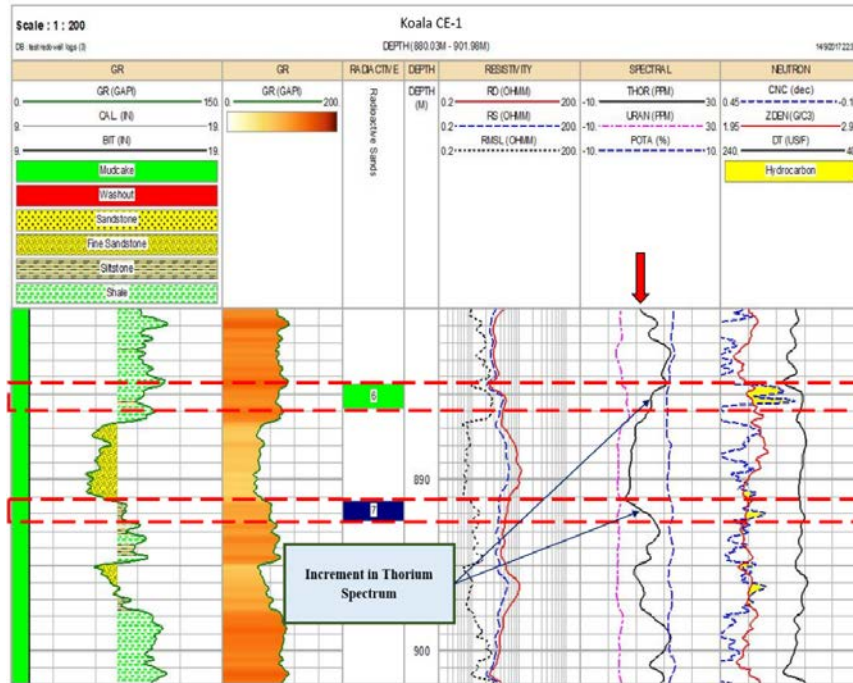


Fig. 6: Spectral gamma ray-thorium spectrum increase in interpreted radioactive sands layer 6 and 7 in Koala CE-1 well

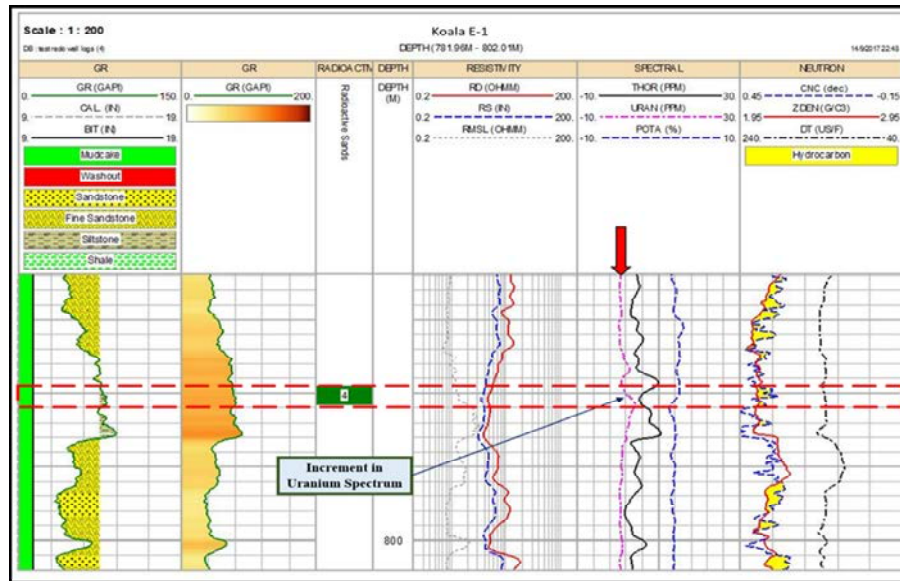


Fig. 7: Spectral gamma ray-uranium spectrum increase in interpreted radioactive sands layer 4 in Koala E-1 well

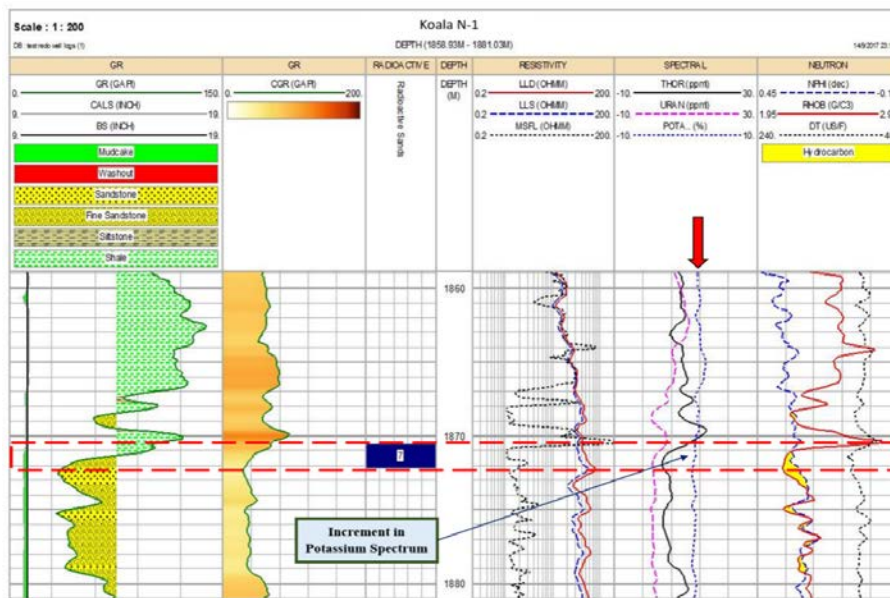


Fig. 8: Spectral gamma ray-potassium spectrum increase in interpreted radioactive sands layer 7 in Koala N-1 well

Table 1: Summarizing gamma ray spectrum and radioactive source	
Spectral gamma ray	Source indicator
Thorium	Heavy minerals Clay minerals
Uranium	Organic compounds Phosphates
Potassium	Presence of mica Presence of K-feldspar

example of increment in uranium reading can be seen in Fig. 7. In addition, sometimes increase in the Uranium

reading in the sandstone layer is due to water saturation where the water is washed out from the shale layer beneath the sandstone thus resulting in the presence of uranium rich water. On the other hand, increase in potassium reading in radioactive sands is due to high mica content and/or K-feldspar mineral as seen in Fig. 8. This relates to sediments originating from nearby granitic highlands. Table 1 summarizes each individual spectrum and its potential source referred from Chudi and Simon (2012).

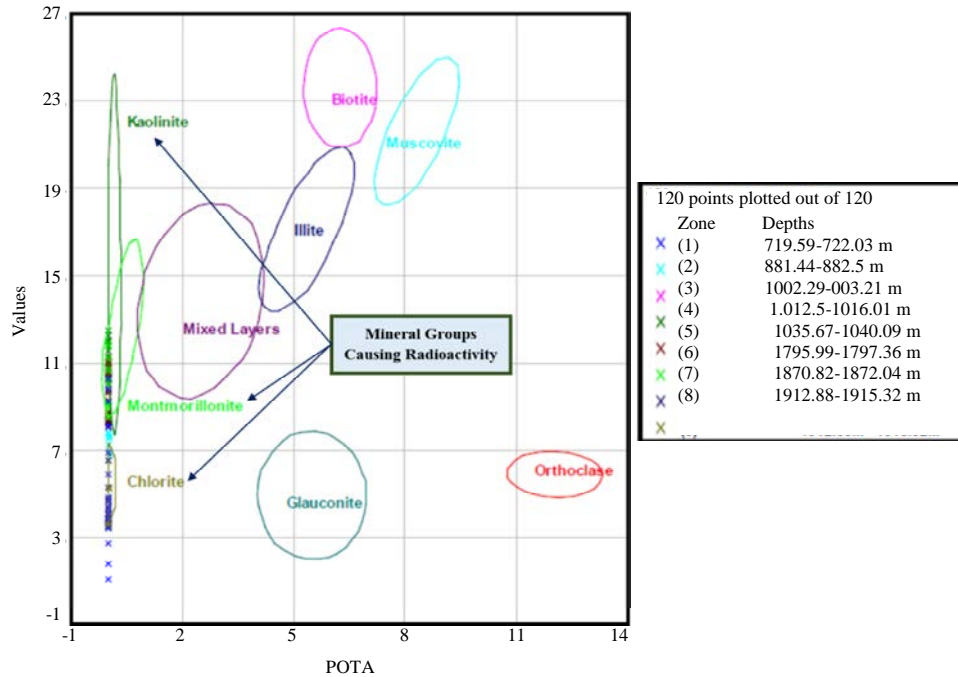


Fig. 9: Potassium vs. thorium cross plot of radioactive sand zones in Koala N-1 well indicating the presence of potential clay minerals of chlorite, montmorillonite and kaolinite

M4; Identifying contributing minerals using potassium vs. Thorium cross plots (spectral gamma ray components):

The potassium vs. thorium cross plot is divided into 9 types of mineral zones. Clay minerals in the range are kaolinite, montmorillonite, chlorite, mixed layers (corrensite and rectorite), illite and muscovite. The balance three are glauconite, biotite and orthoclase. These minerals are said to have significant radioactive characteristics. Glauconite which is also known as green sand is deposited in continental shelf marine environment with slow rates of accumulation. Biotite is considered as a common phyllosilicate mineral within the mica group. Also known as “Black mica” as opposed to “White mica” (muscovite), it is found in igneous and metamorphic rocks. Orthoclase is an important tectosilicate mineral from igneous rocks. Figure 9 shows a good example of the minerals found in the interpreted radioactive sands which consist of clay minerals. From the example, the presence of kaolinite has potential in altering the wettability of the formation to the extend in some cases from oil-wet to water-wet. Mineral identification is important as it helps in choosing suitable composition for drilling mud, since, different clay minerals would affect the porosity and permeability of the reservoirs as well as wettability. Hence, highlighting the importance of identifying the type of radioactive minerals.

CONCLUSION

Radioactive sands are a major feature in petrophysics interpretation and has serious impacts to the exploration evaluation. Identification of the true net pay thickness including radioactive sands is essential for the volumetric calculation. In addition, determining the source of radioactivity gives better understanding on the depositional environment and during well development phase especially in choosing the most suitable mud and cement composition. This is because suitable mud will affect the permeability of the formation. By following the proposed methods, volumetric calculation of the reservoir and the pay zone can be achieved as well as accomplishing positive economic values of the studied area.

From the results, it can be observed that the defined radioactive sands are seen directly adjacent between interbedded sands. Thickness of the anomaly ranges from 0.5-3 m which may not be thick but coverage of area is large enough to give significant effect on volume. Source of radioactivity based on spectral components are dominated by the presence of heavy minerals seen by the subjugation of thorium spectrum. This indicates that radioactive sand interval has clay mineral present. With regards to minerals present, majority of radioactive sands consist of kaolinite, montmorillonite, mixed layers (corrensite and rectorite) and chlorite. Thus, the study

highlights the geological implication on radioactive sand concept and its importance. In addition, it shows that gamma ray sole interpretation is not enough to define sand or shale sequence, especially in areas which are highly interbedded. Overlooking net pay reservoirs due to misinterpretation of radioactive sands could result in the losing of potential barrels of oil.

The mentioned methods can be used to re-evaluate the wells in the system without the need of other additional logging tools. This study illustrates employing improved interpretative procedures that lead to more cost-effective re-completion of hidden reservoirs within the same producing system.

RECOMMENDATIONS

Since, the interpretation of the radioactive sands are based on petrophysics analysis only, the interpretation is an assumption based on statistical analysis from the cross plots as well as standard conventional interpretation from wireline logs. If new wells are explored it is suggested to execute as Many Sidewall Core (SWC), Modular formation Dynamic Tester (MDT) and Formation Multi-Tester (FMT) to increase the understanding of the well. If there is core data and thin section analysis or reports, more accurate interpretation can be established. In addition, from the cross referencing of the core analysis, petrophysics data and seismic data an excellent geological model can be achieved.

REFERENCES

Chudi, O. and R. Simon, 2012. Petrophysical characterization of radioactive sands-integrating well logs and core information: A case study in the Niger Delta. Proceedings of the Nigeria Annual International Conference on Exhibition, August 6-8, 2012, Society of Petroleum Engineers, Richardson, Texas, pp: 1-8.

- Cratchley, C.R., P. Louis and D.E. Ajakaiye, 1984. Geophysical and geological evidence for the Benue-Chad Basin Cretaceous rift valley system and its tectonic implications. *J. Afr. Earth Sci.*, 2: 141-150.
- Fairhead, J.D., C.M. Green, S.M. Masterton and R. Guiraud, 2013. The role that plate tectonics, inferred stress changes and stratigraphic unconformities have on the evolution of the West and Central African Rift system and the Atlantic continental margins. *Tectonophysics*, 594: 118-127.
- Genik, G.J., 1992. Regional framework, structural and petroleum aspects of rift basins in Niger, Chad and the Central African Republic (CAR). *Tectonophysics*, 213: 169-185.
- Genik, G.J., 1993. Petroleum geology of cretaceous-tertiary rift basins in Niger, chad and central African republic. *AAPG Bull.*, 77: 1405-1434.
- Harouna, M. and R.P. Philp, 2012. Potential petroleum source rocks in the Termit Basin, Niger. *J. Pet. Geol.*, 35: 165-185.
- Liu, B., L. Wan, F. Mao, J. Liu and M. Lu *et al.*, 2015. Hydrocarbon potential of upper Cretaceous marine source rocks in the Termit Basin, Niger. *J. Pet. Geol.*, 38: 157-175.
- Liu, B., X. Pan, L.K. Wan, F.J. Mao and J.G. Liu *et al.*, 2012. Polyphase rift evolution of the Termit basin, eastern Niger: Constraints from structural and sedimentary records. *Geosci.*, 26: 319-325.
- Wan, L., J. Liu, F. Mao, M. Lv and B. Liu, 2014. The petroleum geochemistry of the termit Basin, Eastern Niger. *Marine Pet. Geol.*, 51: 167-183.

Performance Analysis for Throughput and BER in High Speed Train Broadband Wireless Mobile Communications System under Different Variable Cases

Raad Farhood Chisab

Technical Institute-Kut, Department of Electrical Technology,
Middle Technical University (MTU), Al-Za'franiya, 10074 Baghdad, Iraq

Abstract: High Speed Trains (HST) is a wonderful way for transporting that raised more attention in wireless communication system within great speed environments. Numerous experiments of HST till 450 km/h are done around the world. The investigation of performance for standard structure design is concentrates on mobile system in the urban with respect of speediness. There are more factors that effect on the activity of this system but the most important effect is the Doppler shift frequency which becomes essential factor that corrupting the performance of the system which increases errors of received signal. The aim of this study is to demonstrate the robust of the system that is serve in HST by investigate the throughput with the Bit Error Rate (BER) for the signal that is transferred from the antenna and which error will be occurred with the same signal when it is received by the user outside or inside the train. The tests are done within different circumstances and variable cases in order to measures the powerful of this system. The results show that the system provides a high throughput with lower BER with variable parameters such as number of antennas, doppler shift frequency and reference channels.

Key words: HST, throughput, BER, reference channel, antenna, frequency

INTRODUCTION

Portable broadband communications have been widely industrialized for global transportation to reply the accumulative crowded of programs downloading. So, if the application still and never changed, these techniques will not appropriate for the requirements of HST travelers as the speed is much high and more stimulating for communication strategies (Luo *et al.*, 2012). Every year, the contribution of automatic structure in numerous productions in the domain increases. The method computerization could decrease energy feeding and growth source productivity. In latest era, many hard works are done on adjusting the conservative GSM structure to High-Speed Train Communication (HSTC). It is not easy to support great-data-rate communication services, like "online video or gaming" within HST up to 450 km/h. So, within the fast growth of HST it is required to cumulative railway controlling data for security observing and conservation that is requests to be transmitted within HST (Zhang *et al.*, 2014). Also, the passenger that use mobile phone request more capacity of the network and more dependable quality without depend on the position or speediness. It can be said that the current scheme "GSM-R" simply offers a supreme rate with fewer than 200 kbps. This is primarily works for railway control and cannot to occupy the necessities of upcoming great data rate transmission (Hiruta *et al.*, 2009).

So, an innovative wireless system is in request for HSR. Main think of design for any wideband mobile scheme is the propagation features for the channel. The dependable and accurate channel model works as the allowing basis for real-world scheme and challenging of the HSR scheme. In the real positioning of a wireless scheme, the genuine network arrangement optimization, ability and everything are applied conferring to the propagation atmosphere (Dong *et al.*, 2010). The small scale fading topographies are critical in strategy which assist designer to made fading measures. This measure is variable like such as "Diversity of transmission/reception, error correction coding and interleaving and equalization procedures". So, the structures of HST channel model meaningfully vary as compare with other communication channel (Cichon *et al.*, 1995). These differences originate from the following specific features of the HSR radio propagation.

Diverse scenarios. HST is regularly located on a viaduct overpass residential or rural zones. The transmitter and receiver meet altered channel surroundings. This variation contingent on topographical position, viaduct tallness, antenna location and other parameters and can be demonstrated in Fig. 1. Channel environments are categorized to more kinds such as: "Viaduct, hilly terrain and tunnel (Liu *et al.*, 2012).

Line of Sight (LOS) ascendancy. The modern types of HST paths are illustrious from conservative outlines by milder curves, lower channels and broader channels. Joint

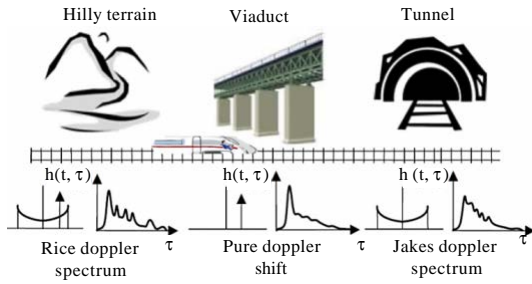


Fig. 1: Variable channel specification with different situation on HST

with tall antenna, these exterior situations produce a pure space to communicate. Dissimilar urban and interior situations with in these exterior situations, the specular LOS factor is robust and additional multipath reverberations happen fewer regularly among the mobile station and base-station.

Great Doppler shift and quick Doppler transition. The extreme speed of HST is currently 480 km/h. this speed lead to get “Doppler shift” of 946 Hz at 2.1 GHz. In case of the base station located remote from the train sideways. The chief aim is that the “Incident angle” among the train and the base station is about 90°. So, the base station is frequently fixed back thirty meter away from train. At once the train crosses the base station, the angle among the received signal and the movement route is altered quickly begun with 0° and end with 180°. This will lead to quick “Doppler transition” (Dimic and Sidiropoulos, 2005).

The meaning of High Speed Train (HST): The high-speed train which is shown in Fig. 2 is a flexible word. In order to know this survey, it involves schemes within the important features (Biguesh and Gershman, 2004):

- Trains aimed for continued works between 200 km/h to about 500 km/h
- Trains that deliver services among populace midpoints or urbanized zones. This service comes with some temporary end
- Trains that normally usage “Semi-permanently” coupled groups of energy engines and coaches of numerous structures
- Trains that ensure dedicate using “Rights-of-way”, mainly external of stations
- Trains that utmost frequently using over-head, continually “Tensioned catenary” to take energy to trains and power engine
- Structures that using certain kind of “Automatic Train Control (ATC)” or “Positive Train Control (PTC)”



Fig. 2: The shape of the high speed train HST

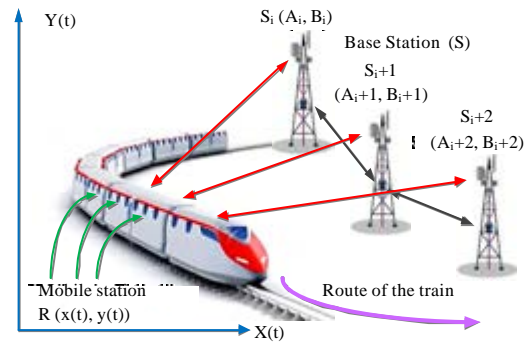


Fig. 3: The communication system between HST and base-stations

The communication system model: Dissimilar community wireless communication scheme, the system of communication that is design to works in railway is devoted to spreading control and transmitting information. The “Base Stations (S_i)” are then positioned alongside the train to assurance high dependability. When making comparison to the weakening produced by “Distance and fading”, the result of elevation difference on received sign is of “No account”. In a actual easy method, the drawing of “Wireless communication system” for HST is shown in Fig. 3. So, the synchronize order of entirely base Stations (S) are fixed as a distinct order $\{(A_1, B_1), (A_2, B_2), \dots, (A_i, B_i), \dots, (A_m, B_m)\}$ in progress (Serbetli, 2008). Within real design, the tracks of the HST are never continuously straightforward. This is because certain difficulties in the route. These difficulties may be hills, rivers and woods. The situation of the “Mobile station (R)” rests on the track within train. This is signified as a “Continuous time function (x(t), y(t))”. The ith base-stations are active within works in the case as (Hamdi and Pap, 2010):

$$(A_i, B_i) \min_{1 < j < m} \sqrt{(A_j - y(t))^2 + (B_j - x(t))^2} \quad (1)$$

In calculation observation, “Discrete sequence” is a singular type of “Multi-partition function”. So, the function for the continuous time of the base station Si (A(t), B(t)) is (Suraweera *et al.*, 2010):

$$(A(t), B(t)) = (A_i, B_i) \min_{|k| < m} \sqrt{(A_i - y(t))^2 + (B_i - x(t))^2} \quad (2)$$

To study the doppler influence as shown within Fig. 2, c(t) and e(t) are definite as” horizontal acceleration and vertical acceleration”, respectively. This are associated with topography and period between neighboring train stations. Within the physical side, the immediate velocity charted to x, y coordinate while the path angle is calculated as:

$$v_x(t) = \int_0^t e(\tau) d\tau \quad (3)$$

$$v_y(t) = \int_0^t c(\tau) d\tau \quad (4)$$

$$\theta_v = \arctan \left(\frac{v_y(t)}{v_x(t)} \right) \quad (5)$$

Therefore, the organization of R is moreover stated as a meaning of acceleration:

$$x(t) = \int_0^t v_x(\tau) d\tau = \int_0^t \int_0^\tau e(\theta) d\theta d\tau \quad (6)$$

$$y(t) = \int_0^t v_y(\tau) d\tau = \int_0^t \int_0^\tau c(\psi) d\psi d\tau \quad (7)$$

Now, the measure of angle among the received signal and the speed is:

$$\theta = \arctan \left(\frac{B(t) - y(t)}{A(t) - x(t)} \right) - \arctan \left(\frac{v_y(t)}{v_x(t)} \right) \quad (8)$$

So, Doppler shift of the straight line of site factor fs will be written as:

$$f_s(t) = \frac{\sqrt{v_y^2(t) + v_x^2(t)} f}{c} * \cos \theta \quad (9)$$

Also, the extreme Doppler spread f_max as:

$$f_{max}(t) = \frac{\sqrt{v_y^2(t) + v_x^2(t)} f}{c} \quad (10)$$

So, we can say that the HST situation for the check of the baseband act is “A non-fading propagation channel” with single tap. Then, Doppler shift is known by Ahmed and Arslan (2008):

$$f_s(t) = f_d \cos \theta(t) \quad (11)$$

f_d is represents the extreme Doppler frequency:

$$\cos \theta(t) = \frac{D_s / 2 - vt}{\sqrt{D_{min}^2 + (D_s / 2 - vt)^2}} \quad 0 \leq t \leq D_s / v \quad (12)$$

$$\cos \theta(t) = \frac{-1.5D_s + vt}{\sqrt{D_{min}^2 + (-1.5D_s + vt)^2}} \quad D_s / v < t \leq 2D_s / v \quad (13)$$

$$\cos \theta(t) = \cos \theta(t \bmod (2D_s / v)) \quad t > 2D_s / v \quad (14)$$

The D_s/2 is the original expanse of the train from base station. It is measures in meters. v is the speed of the train (Patzold and Nguyen, 2004).

MATERIALS AND METHODS

The hardware component of the system: The hardware components that are used for sending and receiving the information via the mobile wireless system in HST can be shown in Fig. 4. In this system, the signal will input as a block of modulation samples in which different types of modulation will be selected according to channel case (QPSK, 16QAM and 64QAM). the three types of modulation is shown in Fig. 5. That means when channel is in healthy condition the 64QAM will be apply with minimum bit error rate BER while when the channel in

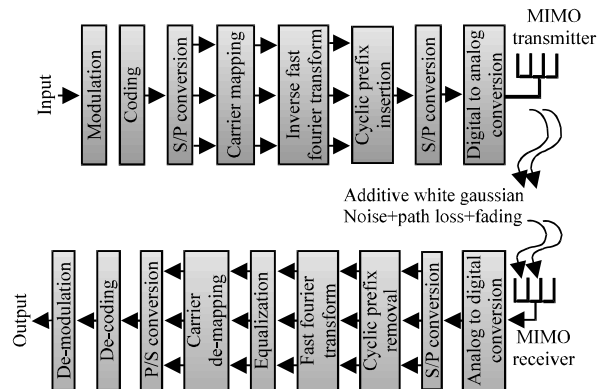


Fig. 4: The block diagram for the communication system in HST

bad condition the QPSK will be chosen in order to reduce the BER. When the channel case is in the middle, the 16 QAM will be taken in order to minimize the Bit Error Rate (BER).

After that the modulated signal is moved to a next block which is the coding then the coded signal enters to serial to parallel block. Inside this block the signal is transferred to parallel form. The amount of parallel numbers corresponds to the amount of carriers in the system. Then the parallel signal pass through the new block which is subcarrier mapping in which the number of carrier will be increased. There are two types of mapping which are Localized Frequency Division Multiple Access (LFDMA) and Distributed Frequency Division Multiple Access (DFDMA). These two types of mapping are shown in Fig. 6.

After that the result signal enters to the important block which is the Inverse Fast Fourier Transform (IFFT). Now the signal passes through the very necessary block which is the cyclic prefix block. To explain the need for this block it can be said that When the OFDM data is serve to broadcast via. an extensive frequency, the inter

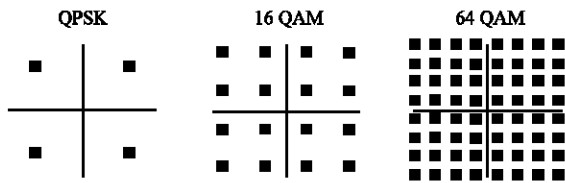


Fig. 5: The three types of modulation

signal interference is so large. This will destroy the orthogonally between subcarriers. This will produce demodulation mistakes and effect in BER. The designers supplement shield which is “Guard Interval (GI)”. Though, in case of GI is unfilled, the orthogonality among sub-carriers is never longer obtainable, since of “Inter Carrier Interference (ICI)” produced by multipath. To remove the “ISI and ICI”, Cyclic Prefix (CP) is put in OFDM transmission. The altered CP spans could be working in changed transmission situations (Fig. 7):

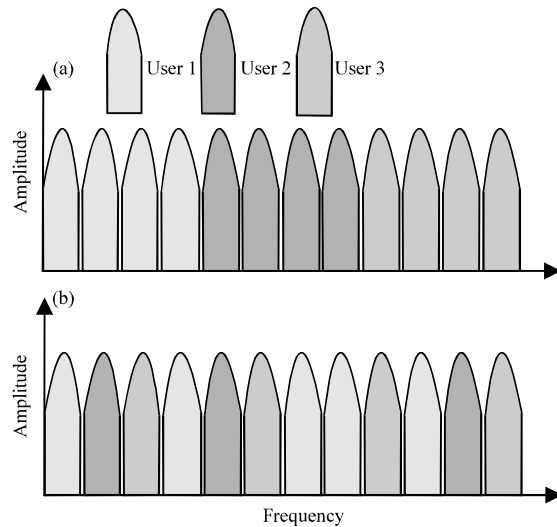


Fig. 6: The two types of subcarrier mapping: a) Localized mode and b) Distributed mode

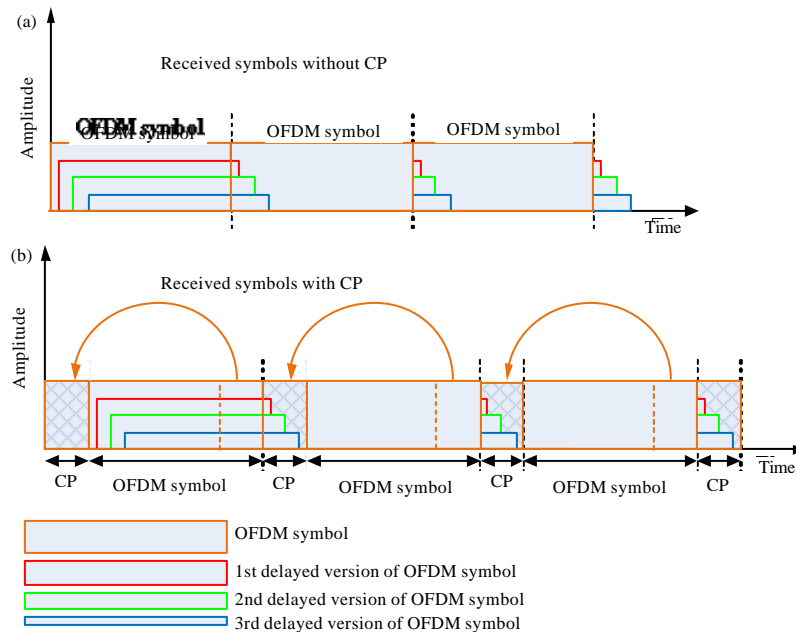


Fig. 7: The effect of adding or remove CP on the received signal

- Shorter cyclic prefix in normal-cell atmospheres to reduce the “Cyclic-prefix overhead”
- Longer cyclic prefix through risky time dispersion. In case of high speed train this type of CP is used

The process of CP is demonstrates in Fig. 7. In Fig. 7, we can understand the effect of adding CP to the symbol and how the signal loss part of the information in case of remove CP. Then the resulting signal after adding CP will be return to serial form using the parallel to serial block.

The signal now returns to analog form in order to transfer through channel to reach the receiving end. Now, the wireless channel suffering from three types of impairments which are frequency, distance and time which can be illustrated in Fig. 8. The distance impairments can be divided into three parts which are multipath fading, shadowing and path loss which is effected on power transmitted and cause the degradation as shown in Fig. 9. From Fig. 10, it can be shown that the data rate of the

signal depends on the quality of the channel because when the channel is bad then error will be increased and retransmitted the signal is occurred and the data received with error will increased.

Doppler spread is additional chief test for understanding wireless broadband HST. Particularly for high speed train communication structure that depends on OFDM. The high channel differences inside one OFDM block terminate the orthogonally among subcarriers that lead to inter-carrier interference relative to the “Doppler frequency”. The orthogonality means the peak point of one carrier equal to zero value of all other carriers. This rule of orthogonality between the carriers can be shown in Fig. 11.

At this time the signal reach to the receiving end and all the blocks in the transmitting will be inverted and one

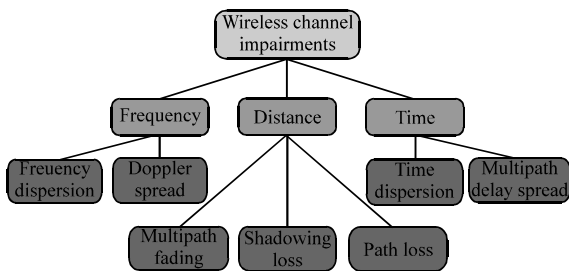


Fig. 8: The parts of wireless channel impairments

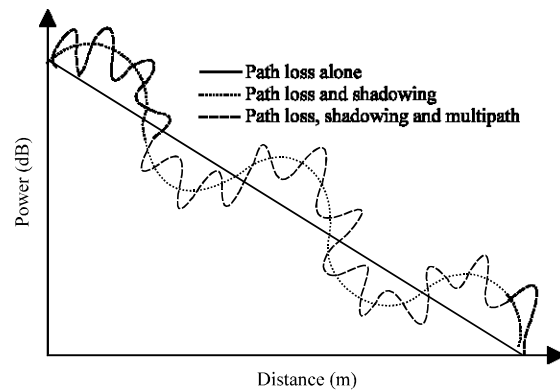


Fig. 9: The effect of distance impairments on power transmitted

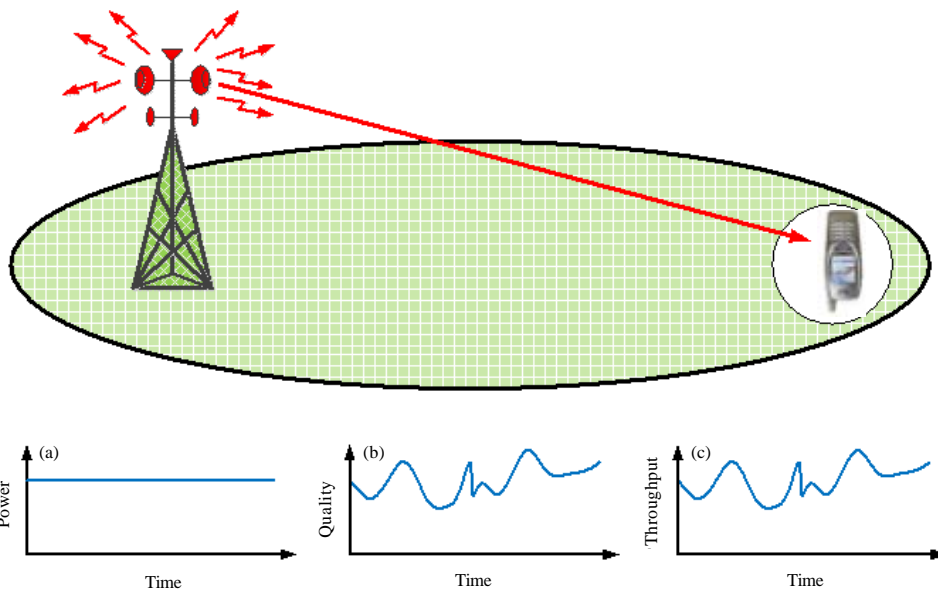


Fig. 10: The effects on channel quality on data rate of signal: a) Transmitted power; b) Channel quality and c) Date rate

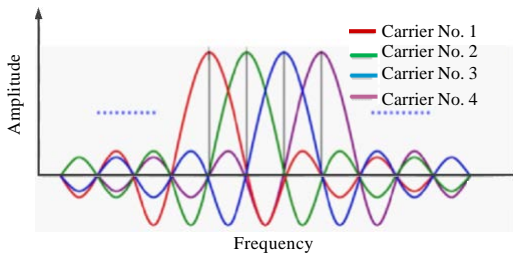


Fig. 11: The orthogonal carriers for communication system

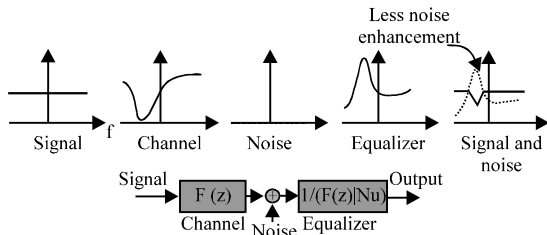


Fig. 12: The simple type of equalization used in communication system

block is adding to reduce the channel effect which is the equalization block. The purpose of equalization is to discover an opposite filter that recompenses the “ISI”. So, the multipath signals turn out to be moved and allied in time. The simple type of equalization can be shown in Fig. 12.

It can be said that the design of the transceiver in the HST is robust against all types of degradation and errors that can be occurred due to the high speed and channel alteration within small period of time. So, the target of designer is to makes the BER be minimized and the throughput is maximized.

RESULTS AND DISCUSSION

In the HST, the route divided into two cases which are: the open space (case 1) which is the most case that is taken in HST while the other case is within tunnel (case 2) that is occurs within some times for small distance. The necessary values for simulation are found in Table 1. After apply the simulation the relations of Doppler shift with time for two cases are shown in Fig. 13 and 14.

In this study, the first situation which is open space (case 1) is taken in consideration. It can be said that a more amount of base station RRHs exist and dispersed in equal space or distance along the train-truck. The distance =Ds with the similar base station number as shown in Fig. 15

In the beginning, when running the simulation the trajectory of power in dB with respect to distance for

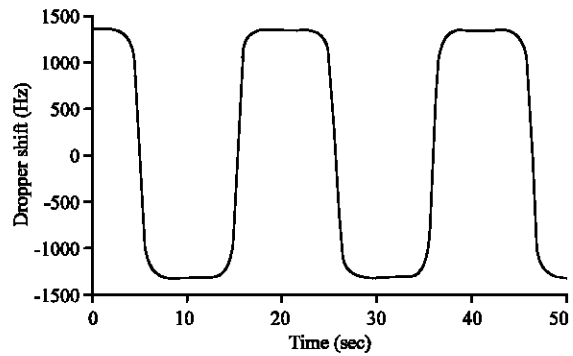


Fig. 13: The doppler shift trajectory for case 1 channel

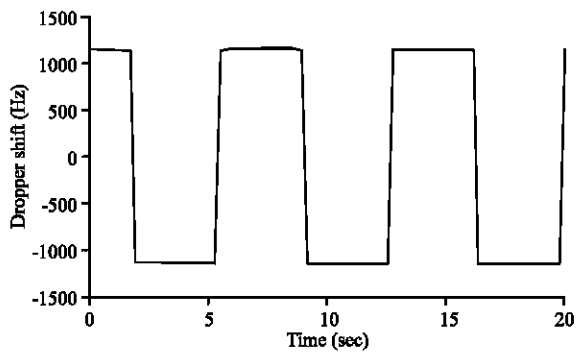


Fig. 14: The doppler shift trajectory for case 2 channel

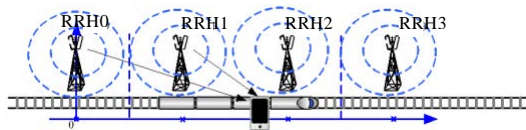


Fig. 15: The open space (case 1) configuration

Table 1: The value of necessary component in HST simulation with two cases of channels

Components	Case 1	Case 2
D_s	1000 m	300 m
D_{min}	50 m	2 m
v	350 km/h	300 km/h
f_c	1340 Hz	1150 Hz

different base stations RRHs which is began from the ID (-1) (RRH₁) that means the train leave this base station and faraway from this base station therefore it can be seen that this power is low as compared with other base station as shown in Fig. 16. The trajectory of Doppler shift in (Hz) with distance in meter for all base station is shown in Fig. 17. Also the trajectory of absolute delay with relative delay for all base stations is shown in Fig. 18.

In this study, the behavior of the high speed train communication system is analyses within different cases with variable factors. The benefit of this study is to know the powerful and drawback points in this system,

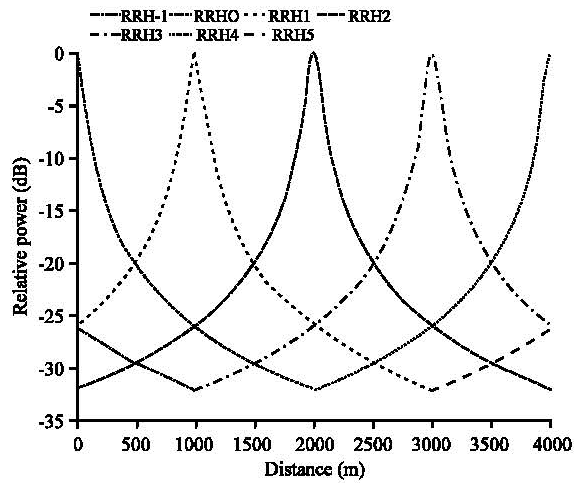


Fig. 16: The relation between relative power and distance

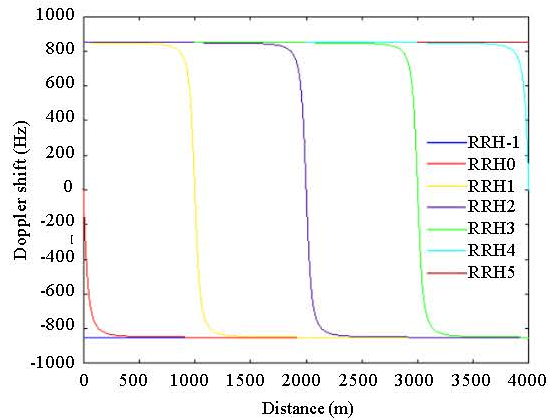


Fig. 17: The relation between doppler shift with distance

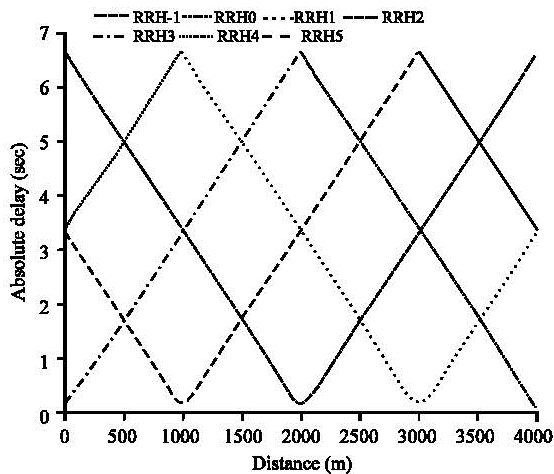


Fig. 18: The relation between absolute delay and distance

especially, when the speed is very high and how can this system stay valid and can obtain the information with

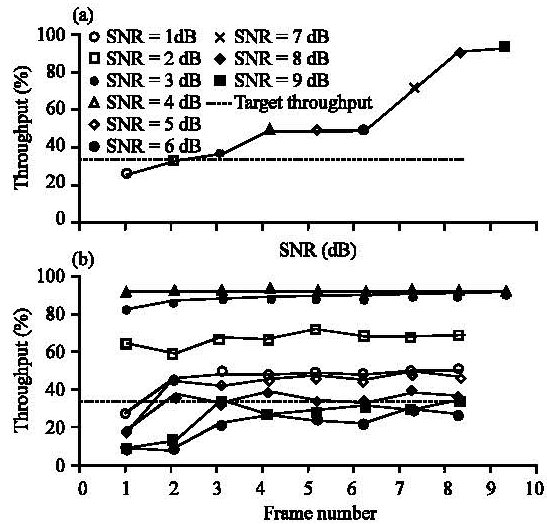


Fig. 19: The performance under the single antenna (SISO):
a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

Table 2: The variables that is used in the simulation

Variables	Quantites			
No. of antenna	1	2	4	8
Doppler frequency (Hz)	5	70	300	750
Channel reference	R0	R1	R3	R5

minimum errors. The high speed train communication system is testes within three main variables which are: the amount of antennas used, Doppler frequency and the channel reference. All details of these factors are shown in Table 2. The main parameter that will take in comparison of performance for the system is the throughput which is the rate of successful message delivery over a communication channel. The data these messages belong to may be delivered over a physical or logical link or it can pass through a certain network node. Throughput is usually measured in bits per second (bit/s or bps) and sometimes in data packets per second (p/s or pps) or data packets per time slot. The system throughput or aggregate throughput is the sum of the data rates that are delivered to all terminals in a network. Throughput is essentially synonymous to digital bandwidth consumption; It can be analyzed mathematically by applying the queuing theory where the load in packets per time unit is denoted as the arrival rate (λ) and the throughput where the drop in packets per time unit is denoted as the departure rate (μ) (Fig. 19-30). In all figures which numbered from 19-30 it can be seen that the “Target throughput” is designed to 30% from the full amount of throughput and it can be noticed that the real throughput is increased from lower values to 100% when the amount of signal to noise ratio increased. The amount

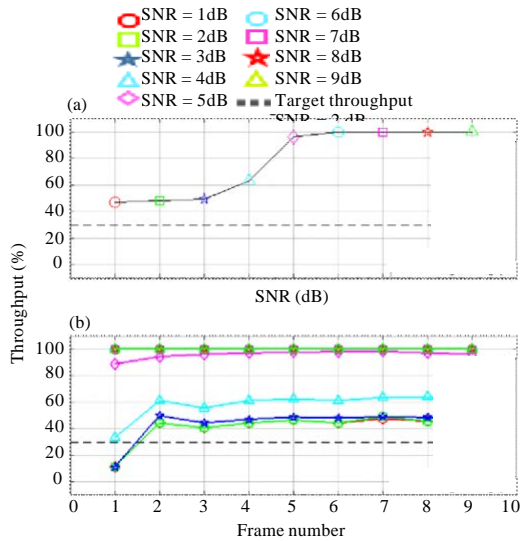


Fig. 20: The performance under the two antennas (MIMO): a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

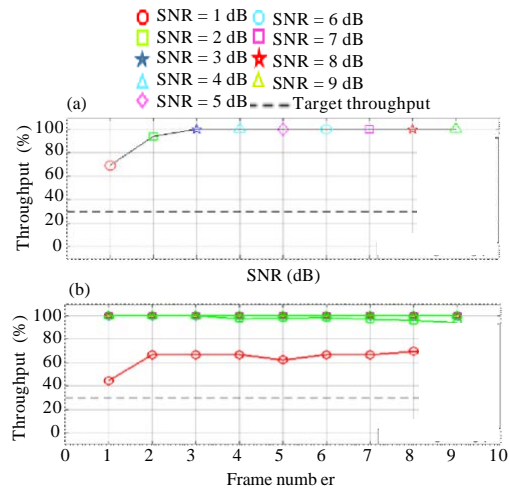


Fig. 21: The performance under the four antennas (MIMO): a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

of throughput is reach to maximum value at different amount of signal to noise ratio which differs from case to others that depends on the variables such that amount of antenna, Doppler frequency and reference channel. Within the first parameter which is the amount of antennas four values of the numbers of antennas are taken which began with single antenna until reach to eight. It can be shown that the throughput is with minimum value within the single antenna and begun with 20% from the maximum value of throughput with

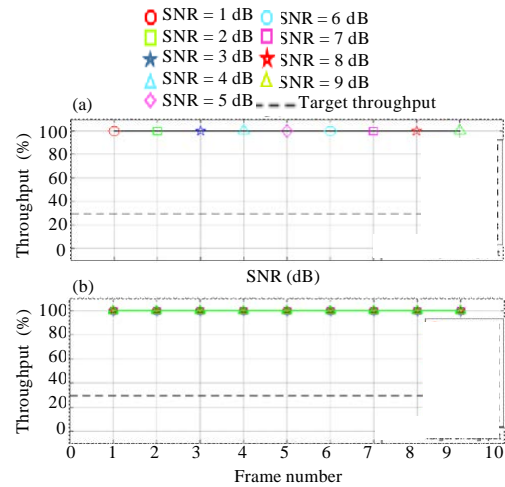


Fig. 22: The performance under the eight antennas (MIMO): a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

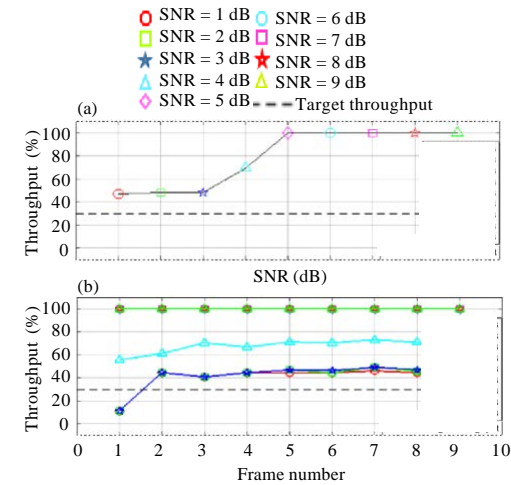


Fig. 23: The performance under doppler frequency equal to 5 (Hz): a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

SNR = 1 dB which is shown in Fig. 19 and raise until reach 100% at SNR = 8 dB. When the antenna become two antennas it can be show that the throughput begun with 47% from optimum at SNR = 1 dB as shown in Fig. 20 and reach to maximum value at SNR = 6 dB. But when the antenna increased and become four antennas the throughput will be increased and begun with 68% from optimum with SNR = 1dB as shown in Fig. 21 and become maximum value at SNR = 3 dB. The final value of changing

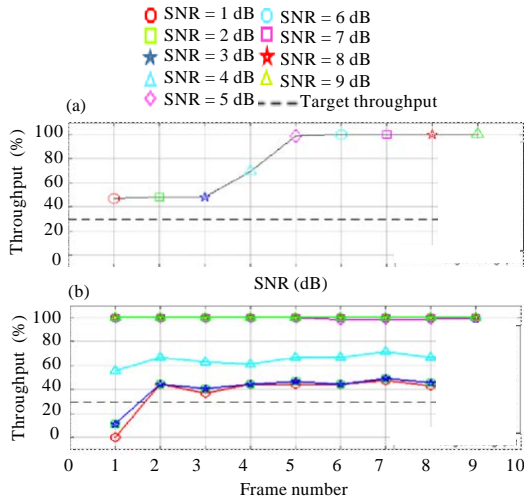


Fig. 24: The performance under doppler frequency equal to 70 (Hz): a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

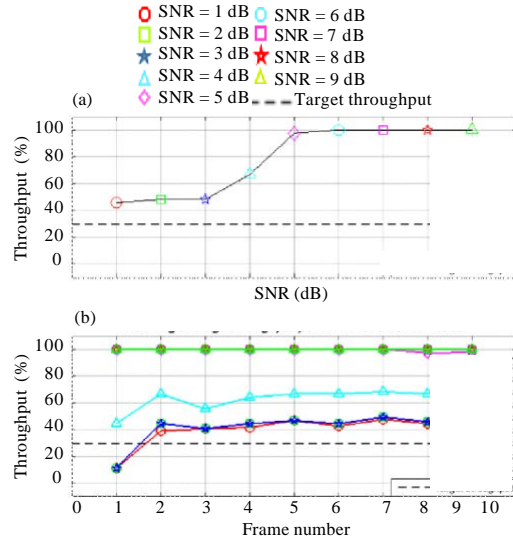


Fig. 26: The performance under doppler frequency equal to 750 (Hz): a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

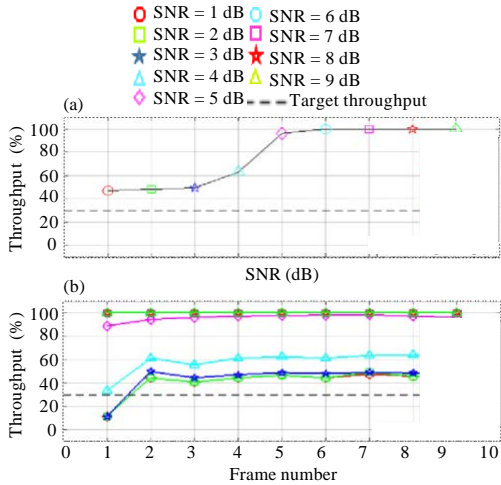


Fig. 25: The performance under doppler frequency equal to 300 (Hz): a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

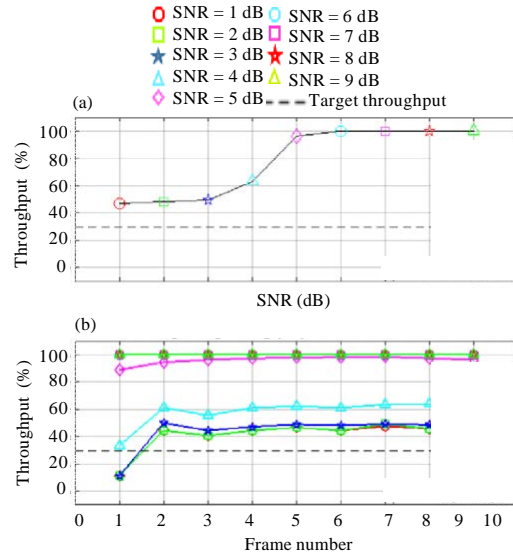


Fig. 27: The performance under the reference channel equal to R_0 : a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

the antenna amount is when the amounts of antenna increased and become equal to eight antennas. In this case which is the best one, it can be shown that the optimum will occurred from the beginning with SNR = 1 dB as shown in Fig. 22 and continued with optimum values for all signal to noise ratio.

The next parameter that will changes in this simulation is the Doppler frequency which relative to speed of the train and carrier frequency that is used in the

process of transmission. This parameter will take in the beginning the value of 5 Hz. In this value, it can be seen that the amount of throughput is begun with about 46% from optimum value at SNR = 1 dB as shown in Fig. 23 and increased until reach the optimum value with SNR = 5 dB. After that the Doppler frequency become

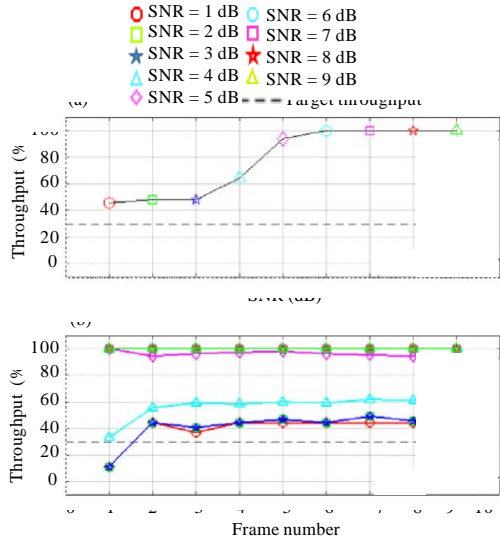


Fig. 28: The performance under the reference channel equal to R_1 : a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

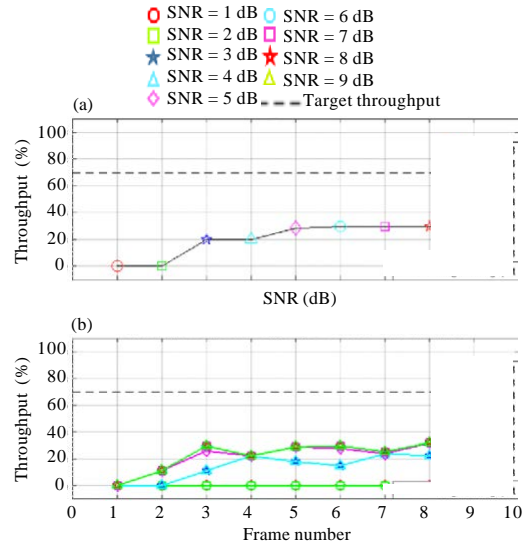


Fig. 30: The performance under the reference channel equal to R_5 : a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

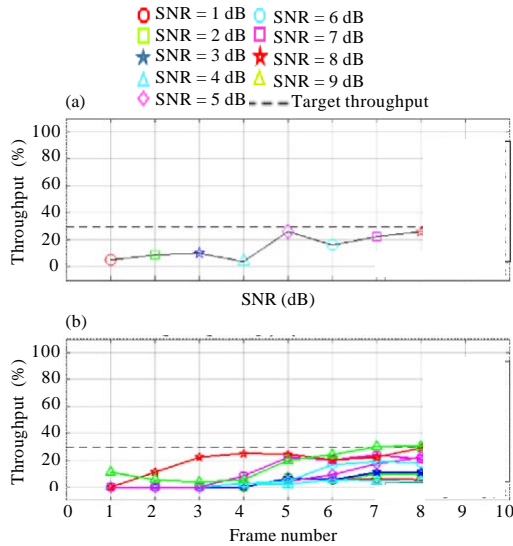


Fig. 29: The performance under the reference channel equal to R_3 : a) Throughput vs. SNR and b) Running average throughput per frame for all SNR value

70 Hz and the throughput begun with 45% from optimum value as shown in Fig. 24 but reach the maximum value in SNR = 5.7 dB. Now, the value of Doppler frequency becomes 300 Hz and the throughput begun with value of 44% from maximum value as shown in Fig. 25 and the optimum value reach at SNR = 5.9 dB. The final value of Doppler frequency is 750 Hz. In this value the

throughput begun with value of 43% from optimum value as shown in Fig. 26 and raised until reach maximum at SNR = 6 dB.

The third parameter that is taken in simulation is the reference channel. In this case four types of reference channel are taken which are R_0 , R_1 , R_3 and R_5 . In the first reference channel which is R_0 the value of throughput is 46% from the maximum value at SNR = 1 dB as shown in Fig. 27 and reach the maximum value at 5.8 dB. In the second case when the reference channel is R_1 the throughput will began from 44% from the optimum value at 1 dB as shown in Fig. 28 and reach the maximum value at 6.1 dB. The third case when change the reference channel to R_3 . In this case, the throughput start with 6% from the optimum value at SNR = 1 dB as shown in Fig. 29 and reach the 36% from the maximum value at 9 dB. The last value of reference channel is R_5 in which the throughput began with 0% from maximum value at 1 dB and reach 32% from maximum value at 9 dB as shown in Fig. 30.

CONCLUSION

The high speed property for this type of train which is called (high speed train “HST”) makes it more complex to deal with the circumstances that will drop the efficiency of this system. So, more than one factor will effects on the working of communication system that exist in this train. In this study, three selected factors that is govern the activity of the communication system are selected to test

and show how can the throughput changes with these factors. In the first factor that is the amount of antennas used in simulation when increase the number of antennas the throughput will improved and the error will be decreased but this need robust control system on the information that will splits between the antennas in order not to mix the information and loss the data so make the tradeoff between the number of antennas and the throughput that will be gained. Within the second parameter which is the Doppler frequency when increase the Doppler frequency the throughput will be decrease and the error is increased in the signal that is transferred though the channel. It can be noticed from third parameter which is the reference channel when increase the reference channel from R_0-R_s , lead to decrease the performance of the system to lowest value within the simulation.

REFERENCES

- Ahmed, S. and H. Arslan, 2008. Evaluation of frequency offset and Doppler effect in terrestrial RF and in underwater acoustic OFDM systems. Proceedings of the IEEE Conference on Military Communications (MILCOM'08), November 16-19, 2008, IEEE, San Diego, California, USA., ISBN:978-1-4244-2676-8, pp: 1-7.
- Biguesh, M. and A.B. Gershman, 2004. Downlink channel estimation in cellular systems with antenna arrays at base stations using channel probing with feedback. EURASIP J. Appl. Signal Process., 2004: 1330-1339.
- Cichon, D.J., T. Zwick and W. Wiesbeck, 1995. Radio link simulations in high-speed railway tunnels. Proceedings of the 9th International Conference on Antennas and Propagation Vol. 2, April 4-7, 1995, IET, Eindhoven, Netherlands, pp: 216-219.
- Dimic, G. and N.D. Sidiropoulos, 2005. On downlink beamforming with greedy user selection: Performance analysis and a simple new algorithm. IEEE. Trans. Signal Process., 53: 3857-3868.
- Dong, W., G. Liu, L. Yu, H. Ding and J. Zhang, 2010. Channel properties of indoor part for high-speed train based on wideband channel measurement. Proceedings of the 5th International ICST Conference on Communications and Networking in China (CHINACOM'15), August 25-27, 2010, IEEE, Beijing, China, ISBN:978-0-9845893-3-3, pp: 1-4.
- Hamdi, K.A. and L. Pap, 2010. A unified framework for interference analysis of noncoherent MFSK wireless communications. IEEE. Trans. Commun., 58: 2333-2344.
- Hiruta, M., M. Nakagawa, S. Haruyama and S. Ishikawa, 2009. A study on optical wireless train communication system using mobile object tracking technique. Proceedings of the IEEE 11th International Conference on Advanced Communication Technology Vol. 1 (ICACT'09), February 15-18, 2009, IEEE, Phoenix Park, South Korea, ISBN:978-89-5519-138-7, pp: 35-40.
- Liu, L., C. Tao, T. Zhou, Y. Zhao and X. Yin *et al.*, 2012. A highly efficient channel sounding method based on cellular communications for high-speed railway scenarios. EURASIP. J. Wirel. Commun. Networking, 2012: 1-16.
- Luo, Q., W. Fang, J. Wu and Q. Chen, 2012. Reliable broadband wireless communication for high speed trains using baseband cloud. EURASIP J. Wireless Commun. Network., 10.1186/1687-1499-2012-285
- Patzold, M. and V.D. Nguyen, 2004. A spatial simulation model for shadow fading processes in mobile radio channels. Proceedings of the 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'04) Vol. 3, September 5-8, 2004, IEEE, Barcelona, Spain, pp: 1832-1838.
- Serbetli, S., 2008. A simple antenna combining framework for Doppler compensation in mobile OFDM systems. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP'08), March 31-April 4, 2008, IEEE, Las Vegas, Nevada, USA., ISBN:978-1-4244-1483-3, pp: 3037-3040.
- Suraweera, H.A., H.K. Garg and A. Nallanathan, 2010. Performance analysis of two hop amplify-and-forward systems with interference at the relay. IEEE. Commun. Lett., 14: 692-694.
- Zhang, J., F. Jin, Z. Tan, H. Wang and Q. Huang *et al.*, 2014. Performance analysis of high-speed railway communication systems subjected to co-channel interference and channel estimation errors. IET. Commun., 8: 1151-1157.

Review Paper on Wireless Network-on-Chip Architecture

¹Ammar A. Ali, ²Bashar S. Al-Dabbagh and ³Laith Alzubaidi

¹College of Islamic Science, University of Baghdad, Baghdad, Iraq

²Autonomous University of Madrid, Madrid, Spain

³University of Information Technology and Communications, Baghdad, Iraq

Abstract: As fabrication technologies continue to provide smaller transistors, the number of processing cores that in multi-core processing systems increased. Large amount of data is transferred between these processing cores using Network on Chip (NoC). Fast NoCs are required for this task. Recently, wireless networks concepts is proposed to be used in NoCs. In the past decade, attention has been paid to Wireless Network on Chip (WiNoC). This research investigates the latest improvements in the field of WiNoC since the year 2014. It is intended for this research to cover the state of the art of the latest proposed research in this era. Several research study have been studies in this research. The study's focus varied on different topics starting from system level down to the circuit levels. All of these levels have been covered in this research. Tens of research papers have been studied in preparation for this research. Only ten study which have vital contributions are presented. The discussed contributions varied on different levels starting form circuit level up to application level. Challenges and suggestions for future research are presented at the end of this study.

Key words: WiNoC, MPS, parallel processing, architecture, latency, power area, CMOS

INTRODUCTION

Adaptive using advanced fabrication technologies con-tributes to reduce size, increase speed and decrease power consumption of transistors. This improvement led to produce more complex integrated circuits where complete computer system has been built on a single IC (system-on-chip). The SoC contains all functional elements of a computer system including a processor on chip memory and all other peripheral components (Martin and Chang, 2001). The multi-core-system-on-chip is a more sophisticated version of SoC. This system is an integration of a number of one or more types of heterogeneous processing units, general purpose processors, DSPs or specific processors (Devigo *et al.*, 2015). These processing units (cores) simultaneously cooperate to implement complex tasks. It consists of a number of hardware processing cores that cooperate to process large scale tasks simultaneously. As a parallel computing system, SoC is demanded on many scientific and commercial applications like weather forecasting, bioinformatics applications, data analysis, CAD or video rendering.

To get high efficiency of this multiple processor system on chip, processing units (cores) have to be connected by a high efficiency network. This network should satisfy special requirements. These requirements are high bandwidth, low latency and low power consumption. Figure 1 shows an example of multi-core

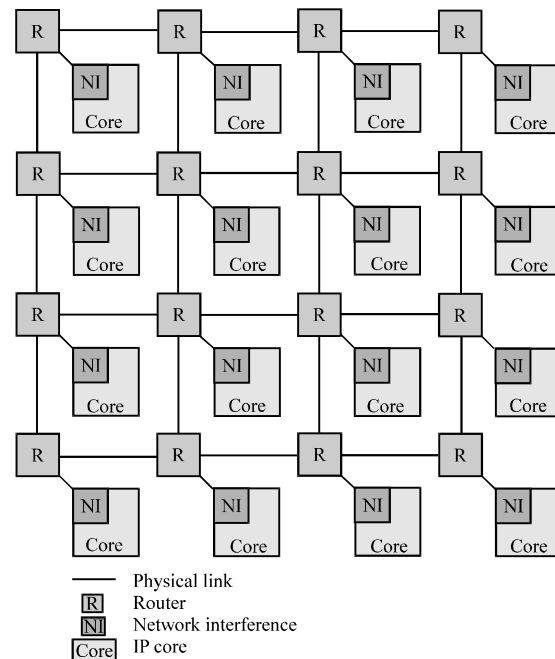


Fig. 1: An example of multi-core system on chip. It consists of a number of processing cores that are connected using connection links. In this particular case 2D mesh network is used (Liu *et al.*, 2012)

system on chip. It consists of a number of processing cores that are connected using connection links. In this particular case 2D mesh network is used. The links are connected using routers that are distributed at each cross. The processing cores are connected to the network through the network interface unit.

Conventional electrical interconnects that use metal wire has several problems such as high latency and limited off-chip bandwidth. They also, dissipate high energy compared to performing a single floating point operation. Wireless and photonic interconnects are strong candidate to solve this problem. This research investigates the latest update in this topic and it focuses on the research study published, since, the year 2014.

Motivation: WiNoC has been an emerging technique in multi-core processing system. The performance of such system can be increased by providing a high speed communication networks. In the past decade, attention has been paid to WiNoC. This research investigates the latest research in the field of WiNoC, since, the year 2014. It is intended for this research to cover the state of the art of the latest proposed research in this era. Several research study have been studies in this research. The study’s focus varied on different topics starting from system level down to the circuit levels. All of these levels have been covered in this research.

Research questions: The goal of this research is to study the WiNoCs and how they can be adopted in the NoC. In addition, it studies the architectures and performances of the recently proposed research. The four main questions of this research is trying to answer are:

- Why do wireless network have advantages over conventional electrical networks?
- How are the concepts of WiNoC can be adopted in NoC and what are the recently proposed architectures?
- What are the latest state of the art contributions?
- What are the challenges facing improvements of WiNoC?

MATERIALS AND METHODS

wireless networks: Several factors stands by suggesting the wireless link over wired in multi-core processing systems. In this study, we will try to analyze some of these factors. Most of the information in this study has been adopted from research by Abadal *et al.* (2015) and Laha *et al.* (2015). The study is tries to answer the first two research questions that this research is trying to

answer: Why do wireless network have advantages over conventional electrical networks? And how are the concepts of WiNoC can be adopted in NoC and what are the recently proposed architectures?

It has been observed that when the distance of electrical meta wire links increases, the capacitance and resistance also increases. In addition, electrical wires require a number of inverters to serve as repeaters while the signal is propagated. These two factors, wire distance and repeaters, represents some of the reasons behind the high energy per bit that the conventional electrical link has. For this reason, long distances electrical wires suffer from high energy dissipation and high latency. In addition to these factors, wired links required number of repeaters and switches that in order to use them in a network. These components consumes substance amount of the overall energy fed into the network. For comparison purposes the energy transferred bits module is used. Figure 2 shows the energy per bit for three main interconnect technologies including electrical, wireless and photonic. This figure has been adopted from a research by Abadal *et al.* (2015). The energy per bit in this figure for the electrical and photonic links is calculated using the DSENT modeling and simulation tool (Sun *et al.*, 2012). The energy data for the wireless link is estimated using trends in wireless interconnect technology. Such data are available at Kaya *et al.* (2013). According to the data in this figure wireless interconnects have constant and low energy per bit compared to the conventional electrical links for more than 10 mm distance. It is also shown that

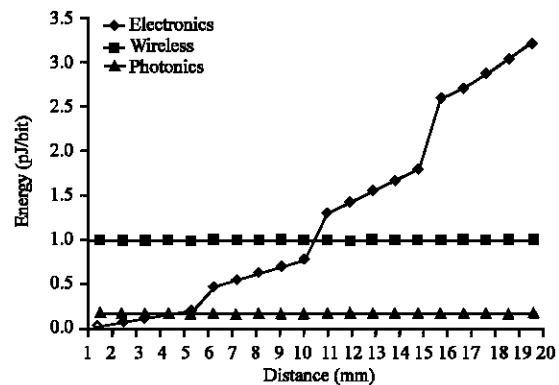


Fig. 2: Energy per bit for three interconnect technologies candidate for building NoC on multi-core systems: electrical (blue color), wireless (red color) and photonics (green color). It is clear how the energy consumed in the electrical networks is dramatically increased as the number of transferred bits is increased (Laha *et al.*, 2015)

the energy consumed by the wireless links is independent of the number of transferred bits. In the other hand, energy consumed in the electrical links is dramatically increased as the number of transferred bits is increased. Photonic interconnects show low energy dissipation per bit. In the same time, however, this does not include the energy consumed by the laser and the energy consumed in other devices such as splitters. Counting these extra energy costs in addition to the transmitting energy of the photonics will push the photonics out of the comparison by high difference. In this case the wireless links are the stronger candidate unless small, low-power and cheap lasers can be fabricated on silicon substrates. For this reason, wireless interconnects have two advantages over the photonic in terms of the technology required for fabricating such networks and power consumed per bit transferred of it.

In addition to energy, several other factors are important in evaluating wireless links. Both bandwidth and latency are important factors for network on chip technologies. Wireless signals propagate very fast compared to their technologies. With near the speed of light, data transferred over wireless links can be received in one clock cycle considering all the time starting from after modulation to receiving the signal by the transducer of the receiver. Meanwhile, electrical links that use metal wires could take about three or four cycles to transmit a signal for 20 mm within the same chip. The number of cycles is increased to order of hundreds at transmitting signals between chips. Bandwidth is an important factor based on what the interconnect technologies are compared. In this factor, the wireless technology is limited by the frequency bands available for such communication tasks. In wireless, channels each with different frequency band is required to increase the bandwidth. In the other hand the bandwidth of the electrical connections can be directly increased by adding extra wires. However, this will increase the area of the systems and power dissipated by these wires. From the above discussion, it is easy to see that wireless networks have limited bandwidth but every efficient power consumption and low latency when they are used for long distance communication that are more than 10 mm.

For this reason, wireless networks have been used in the multi-core processing systems in addition to the conventional wire networks for the long distance communication for their favorable characteristics, low power consumption and low latency. Several other benefits that wireless links have over other type of connections such as:

- Low power consumption for long distance on-chip communication
- Wireless networks are compatible with CMOS fabrication technologies

- Can adopt existed network architectures
- Easy to design and build compared to photonic connections

Wireless network on chip: Over the last five years, many research study suggested techniques to improve the WiNoC. The goal of all of these efforts is to introduce network modules that are efficient to use in the current commercial parallel computing systems. This study is trying to answer the second question that is raised in research motivation part. What are the latest state of the art contributions? Several techniques have been suggested each with different level such as:

- Circuit level
- Transceiver level
- Architecture level
- Application level

This study introduces the research that has been suggested at each of these levels.

Circuit level: Different hardware components are used in the WiNoC including transceivers, routers or antennas. These components form most of the costs of the WiNoC in terms of area and power consumption. Designing these components specifically for the WiNoC could improve these metrics. Different novel and circuit level components designs have been proposed. By Yu *et al.* (2015) proposed circuit level On Off Keying (OOK) demodulator. OOK has low complexity architecture compared to other modulators. It is also considered on the top of the most energy efficient modulation circuits when it is used for communication of distance of order of millimeters. It is already discussed that most of the WiNoC designs require specific performance metrics such as small area overhead, low power consumption of about 1 pJ per bit and high bandwidth. In order to achieve such requirements, Yu and his team proposed an OOK demodulator with a novel differential envelop detector ED topology that has dual gain-boosting techniques. In addition, it is equipped with a single-stage Base Band (BB) peaking amplifier. This amplifier is provided with Actively-Enhanced Tunable Inductor (AETI) load which is used to extend the RX bandwidth. Figure 3a shows the proposed OOK circuit schematic. The proposed OOK is fabricated using a 1P9M 65-nm CMOS technology. The area of the overall system is 0.25 with the external pads while the active area is about 0.043. The measurements of the on-wafer tests are taken using cascade summit-11000 probe station. An Agilent PNAX network analyzer is used for S-parameter measurement up to 67 GHz.

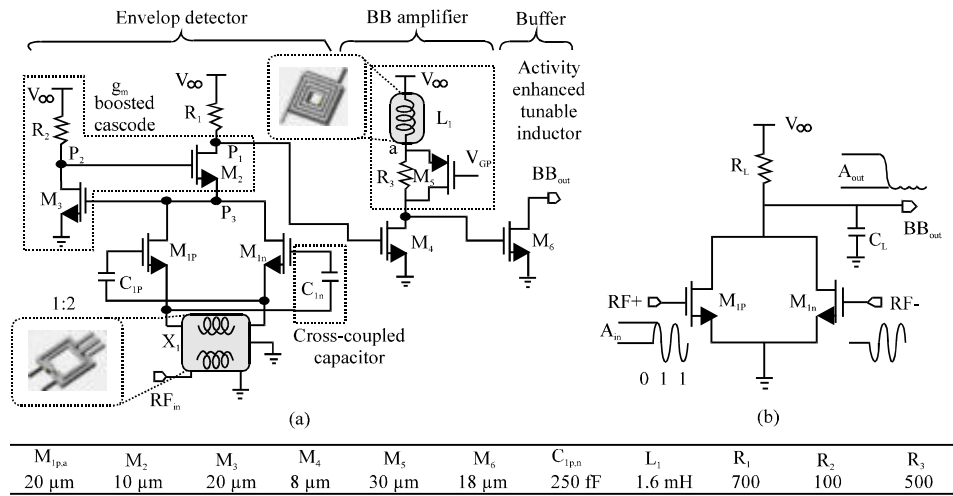


Fig. 3: a) Schematic of the proposed On Off Keying (OOK) demodulator. The sizes of the component of the design are listed in the figure. These sizes are based on the 60 nm technology and b) Schematic of conventional differential envelop detector (Yu *et al.*, 2015)

The PNAX is used to generate the 60 GHz carrier signal. The BB data is analyzed by an Agilent N4960A Bit-Error-Rate Tester (BERT). A Tektronix DSA8300 sampling oscilloscope is used for the evaluation of the demodulated BB output from the OOK demodulator. Results show that the peak data rate of the proposed modulator is 18.7 Gb/sec. This speed comes at low power cost bit-energy efficiency of 0.25 pJ/bit. This means that the overall power consumed was 4.6 mW that is taken from a 1-V supply. The researcher claimed that this OOK demodulator has a high data rate and very low power consumption compared to the state of the art modulators. Research has been conducted to improve the antennas used in WiNoC. Almost all of the WiNoC architectures use millimeter-wave omni antennas that use token passing protocol to control access to the shared wireless medium. This protocol allows only one point to point communication at a time which limit the performance of the network. In this scheme a number of non-overlapping channels are required in the same chip to cover the demand of data transfer in the multi-core chip which requires complex filters in the transceivers. This solution is costly in terms of power and area. Mondal *et al.* (2017) have proposed directional antennas that allows multiple wireless interconnect pares to communicate simultaneously to improve performance and energy-efficiency. To avoid interference, the Wireless Interface (Wis) are placed according to an algorithm that is designed for this specific purpose. This algorithm is also responsible for the routing operations for network. The architecture uses directional Planar Log-Periodic Antennas (PLPAs). The proposed Directional Wireless Network-On-Chip (DwiNoC) architecture is shown in Fig. 4. It is built based on small

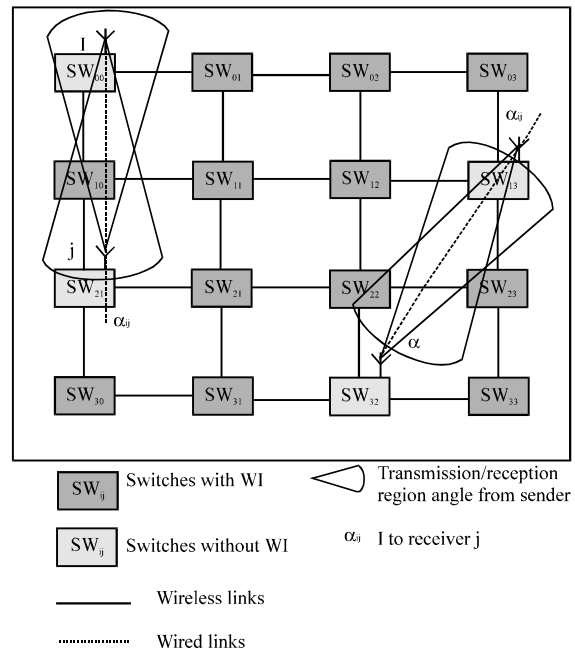


Fig. 4: The proposed directional wireless network-on-chip (DwiNoC) architecture. It is built based on small-world network topology (Mondal *et al.*, 2017)

world network topology. In this topology wire links are used between nodes in sub networks while the directed wireless links are used to connect between the sub networks. The proposed DwiNoC architecture is tested using cycle accurate simulator. It is used to models the progress of the data flits accurately over each clock cycle. HFSS tool is used to produce results of the PLPA

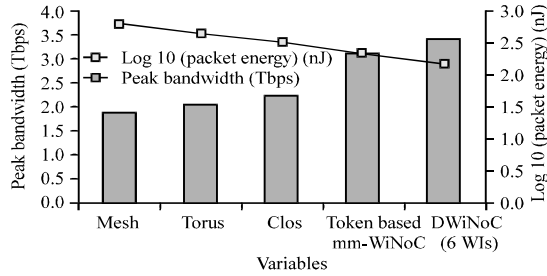


Fig. 5: Maximum obtained bandwidth and power consumed per transmitted packed of the proposed NoC architecture and several other NoC architectures. All of these networks have the same system size of 64-core (Mondal *et al.*, 2017)

antenna. These characteristics are used to produce an accurate model of the energy consumption, bandwidth and reliability for the network.

Two system sizes of 64 and 256 cores are considered in the experiments. Results showed that the proposed DWiNoC overcomes other conventional NOC in terms of peak band-width and power consumption. Figure 5 shows simulation results of energy dissipation and peak bandwidth of a 64 core system. This system is tested using uniform random traffic distribution. The figure also compare different other NoC architectures such as mesh touros, Clos and mm-Wave WiNoC. It is clear from the figure that the proposed research has gain over all other networks.

Transceiver level: In addition to the circuit level, efforts have been paid to improve the transceivers and transducers used in the WiNoC. Transceivers represent important components of any wire-less networks. They come with high cost of area and power consumption. For this reason, several research groups have been working on improving the transceivers in WiNoC. Agyeman *et al.* (2015) proposed a novel transducer design to generate wireless signals with high signal strength. This research is focusing on reducing error rate of the WiNoC to the point where it is competitive with the conventional NoC. Conventional wire network have extremely low Bit Error Rate (BER). It is estimated to be of around 10^{-14} while the error rate of millimeter-wave is of around 10^{-7} as it is estimated by the research presented in this paragraph.

Losing one message in multi-core systems could lead to high effects on the performance of the whole system. For this reason, this research group is trying to improve the BER of the WiNoC by proposing a WiNoC fabric with a novel transceiver. This device is used to launch surface wave signals into a commercially available thin metal conductor. This material is coated with a low cost

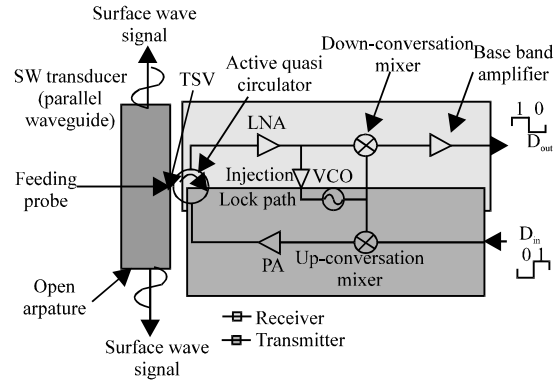


Fig. 6: Block diagram of the proposed transceiver (Agyeman *et al.*, 2015)

dielectric material. For a reliable transmission, the proposed transceiver should have a low power consumption circuit with a wide bandwidth and high data throughput. The proposed transceiver is shown in Fig. 6. It is an update of low-power non-coherent On-Off Keying (OOK) modulator. An up-conversion mixer and a Power Amplifier (PA) are embedded in the transmitter. In the other side, the receiver is equipped with a Low Noise Amplifier (LNA), a baseband amplifier and a down conversion mixer. A single injection-lock Voltage Controlled Oscillator (VCO) is used for both the receiver and transmitter to reduce the area and power consumption of the proposed design. The proposed research is tested using cycle-accurate experiments. An extended version of Noxim simulator which is an open source system simulator is used to perform the experiments. Results showed that the proposed transducer has a very high band width range of 45-60 GHz. The tested device is also found to have high efficiency in terms of average packet delay and power consumption compared to other mm-wave WiNoCs.

Architecture level: A number of research groups have suggested architecture level improvements to increase the efficiency of the WiNOC. Most of these improvements focus on reducing latency and power consumption while keeping an eye on the area of the architecture. The wireless transceivers and the associated antenna imply power and area overhead. Therefore, the number wireless links that can be used within the multi-core system on chip is limited. Rezaei *et al.* (2014) proposed a hierarchical WiNoC. They called it hierarchical wireless-based architecture and for short they gave it the acronym HiWA. In the HiWA, the network is divided into groups of smaller networks, subnets. The communications within the subnets are carried over wire links. A single-hop wireless connection

Table 1: Latency results of the proposed algorithm. The numbers in this table are normalized to the dama results to ease comparison (Rezaei *et al.*, 2015)

Variables	NN (Mondal <i>et al.</i> , 2017)	BN (Agyeman <i>et al.</i> , 2015)	INC (Rezaei <i>et al.</i> , 2014)	DAMA
AWD	1.25	1.30	1.42	1.00
AWMD	1.29	1.32	1.36	1.00
NMRD	1.24	1.62	1.21	1.00
ICR	1.52	1.21	2.01	1.00

channel is used for communication between the nodes. This means that both wired and wireless links are used in this hybrid network. The proposed HiWA is based on a 2D mesh that is divided into square subnets. Each subnet is connected to a center baseline router. The router in each subnet is replaced with wireless routers that communicate with the neighbour subnets. Two different size HiWA: 225 and 256 are built. Figure 7 shows the 256 nodes HiWA. The nodes without color are baseline connection nodes while the green nodes are wireless and red presents the borders of each block. The wireless routers placement is designed in such a way to decrease the number of the wireless routers that that are required to perform the inter-subnets communication. The goal of this step is to improve the performance of the network and energy efficiency. An algorithm is custom designed for this purpose. In addition to the router placement, several other aspects of the architecture have been considered. An addressing method is used for routing packets between the nodes. The system is managed using a routing algorithm that also take in account avoiding deadlock in the routing the packet around the network.

Application level: The contributions in these levels try to design the code in such a way that is efficient to be implemented on systems that utilizes WiNOC. In multi-core systems, applications will enter and leave as sets of communicating tasks at run time might. In some of these cases the overall system might face a highly dynamic research load. The tasks load is controlled by the map-ping of the system manager. It is the manager’s responsibility to allocate application’s tasks onto the system in a way that efficiently utilize the system resources. Rezaei *et al.* (2015) proposed a Dynamic application Mapping Algorithm (DMA) for WiNOC systems. The goal of this research is to reduce both internal and external congestion. Congestion has a negative impact on the network performance which dramatically affects on the delay of the system. Without an efficient mapping mechanism, the wireless routers will be heavily used very. The proposed algorithm has been implemented on the HiWA that is already presented in this study (Rezaei *et al.*, 2014). The tests included different scenarios with several applications. The number of tasks within the applications varied from 4-35 tasks. Task graph generator

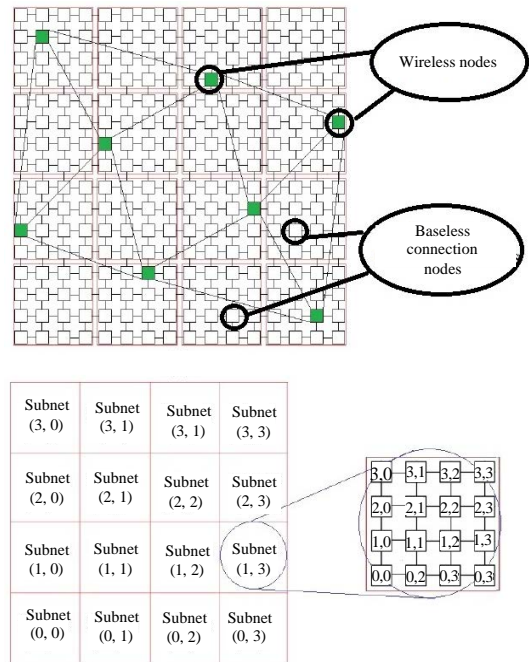


Fig. 7: Hierarchical Wireless Network Architecture HiWA; a) 256 nodes HiWA (The white nodes are baseline connection nodes while the green nodes are wireless and red presents the borders of each block) and b) Physical structure of the 256 nodes HiWA along with its addressing method (Rezaei *et al.*, 2014)

tool is used to generate the tasks. The tests are performed on an open source simulator (XMulator). Several metrics have been used in evaluating the algorithm such as Average Manhattan Distance (AMD), Average Weighted Manhattan Distance, Mapped Region Dispersion (MRD), Normalized Mapped Region Dispersion (NMRD) and Internal Congestion Ratio (ICR). Table 1 shows comparison results of the latency of the proposed algorithm against other algorithms. It can be seen that the DAMA is more efficient than the NN algorithm by 25% in average of all evaluation tests. It has also shown that DAMA has 40 and 20% improvements compared to the BN algorithm at reducing congestion. In the same time, DAMA has 50 and 15% gain over the INC in internal and external congestion, respectively.

RESULTS AND DISCUSSION

This study discusses the results obtained in from all the research presented in the previous study. Several challenges are facing using the NiWoC in commercial application. It is important keep in mind that the technologies utilized for electrical wire NoC are much more optimized compared to the technologies used with the nanophotonic or WiNoC devices (Abadal *et al.*, 2015). The technologies used in WiNoC are still in their infancy. They could be improved in the following years. The fourth question raised at the beginning of this paper answered here. The challenges and improvement aspects of the WiNoC can be grouped into levels.

Circuit level: Most of the circuits presented for the WiNoC are merely updated versions of the circuits used in traditional networks. Novel and optimized circuit topologies that are designed specifically for WiNoC could lead to a important improvements in reducing the area and power consumption. The underlying technology limits the performance of the wireless transceiver that is proposed so far in the literature. The raises the question about the compatibility between the proposed circuits and the CMOS technologies used for the multi-core systems.

Transceiver level: Transceivers and transducers are important components of the WiNoC. They consume much of the power of the network and require large area of the chip. For this reason, a transceiver design that tries to trade-offs between transceiver energy consumption, radiated power in one side and received power in the over side is highly recommended to improve energy efficiency.

Architecture level: Several architectures have been adopted for WiNoC. Most of them are based on the traditional network typologies such as 2D mesh torus and tree. These typologies might look good fits for the WiNoC at the schematic level. However, high schematic level does not necessarily reflect the actual substrates fabrication level. It is vital to keep in mind this fact while designing such architectures.

Application level: To the best of my knowledge almost all the research that has been presented, so, far lacks testing mechanisms at the application levels. The designs proposed in the field are tested on simulation software under hypothetical scenarios and with randomly generated data. A standard application level bench mark is required for evaluation and comparison with other research.

CONCLUSION

Parallel processing has been heavily used for divers of application including weather forecast, data mining and machine indulgent. The multi-core system on chip is one of the highly growing era of parallel processing. Such systems contains number of processing units that collaborates to process large amount of data concurrently which acquires data transfer between these processing units. Wireless network on chip has been presented as a solution for the demand high traffic communication in the multi-core system on chip. It is used for its promising characteristics such as high bandwidth, small area and low power consumption. However, WiNoCs are still evolving to reduce their complicity and overcome the performance of the conventional wire NoC. This research is trying to study the improvements that have been made to the designs of the WiNoC in the past half decade. A number of research study that have important contribution in this field has been presented. The focus of these contribution varied on different levels between circuit and technology level up the application level. At the end of this research the challenges that are facing the improvement of WiNoC are discussed along with some points for future improvements.

REFERENCES

- Abadal, S., M. Iannazzo, M. Nemirovsky, A. Cabellos-Aparicio and H. Lee *et al.*, 2015. On the area and energy scalability of wireless network-on-chip: A model-based benchmarked design space exploration. *IEEE. ACM. Trans. Netw.*, 23: 1501-1513.
- Agyeman, M.O., K.F. Tong and T. Mak, 2015. Towards reliability and performance-aware wireless network-on-chip design. *Proceedings of the 2015 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFTS'15)*, October 12-14, 2015, IEEE, Amherst, Massachusetts, ISBN:978-1-4799-8631-6, pp: 205-210.
- Devigo, R., L. Duenha, R. Azevedo and R. Santos, 2015. Multiexplorer: A tool set for multicore system-on-chip design exploration. *Proceedings of the 2015 IEEE 26th International Conference on Application-specific Systems, Architectures and Processors (ASAP'15)*, July 27-29, 2015, IEEE, Toronto, Canada, ISBN:978-1-4799-1925-3, pp: 160-161.
- Duraisamy, K., Y. Xue, P. Bogdan and P.P. Pande, 2017. Multicast-aware high-performance wireless network-on-chip architectures. *IEEE. Trans. Very Large Scale Integr. (VLSI) Syst.*, 25: 1126-1139.

- Kaya, S., S. Laha, A. Kodi, D. Ditomaso and D. Matolak *et al.*, 2013. On ultra-short wireless interconnects for NoCs and SoCs: Bridging the 'THz Gap'. Proceedings of the 2013 IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS'13), August 4-7, 2013, IEEE, Columbus, Ohio, USA., ISBN:978-1-4799-0066-4, pp: 804-808.
- Laha, S., S. Kaya, D.W. Matolak, W. Rayess and D. DiTomaso *et al.*, 2015. A new frontier in ultralow power wireless links: Network-on-chip and chip-to-chip interconnects. IEEE. Trans. Comput. Aided Des. Integr. Circuits Syst., 34 186 198-10.1109/TCAD.2014.2379640.
- Liu, F., H. Gu and Y. Yang, 2012. DTBR: A dynamic thermal-balance routing algorithm for Network-on-Chip. Comput. Electr. Eng., 38: 270-281.
- Martin, G. and H. Chang, 2001. System-on-chip design. Proceedings of the 4th International Conference on ASIC, October 23-25, 2001, IEEE, Shanghai, China, pp: 12-17.
- Mondal, H.K., S.H. Gade, M.S. Shamim, S. Deb and A. Ganguly, 2017. Interference-aware wireless network-on-chip architecture using directional antennas. IEEE. Trans. MultiScale Comput. Syst., 3: 193-205.
- Rezaei, A., F. Safaei, M. Daneshtalab and H. Tenhunen, 2014. HiWA: A hierarchical wireless network-on-chip architecture. Proceedings of the 2014 International Conference on High Performance Computing and Simulation (HPCS'14), July 21-25, 2014, IEEE, Bologna, Italy, ISBN:978-1-4799-5312-7, pp: 499-505.
- Rezaei, A., M. Daneshtalab, D. Zhao, F. Safaei and X. Wang *et al.*, 2015. Dynamic application mapping algorithm for wireless network-on-chip. Proceedings of the 23rd Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP'15), March 4-6, 2015, IEEE, Turku, Finland, ISBN:978-1-4799-8491-6, pp: 421-424.
- Sun, C., C.H.O. Chen, G. Kurian, L. Wei and J. Miller *et al.*, 2012. DSENT-a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling. Proceedings of the 6th IEEE/ACM International Symposium on Networks on Chip (NoCS'12), May 9-11, 2012, IEEE, Copenhagen, Denmark, ISBN:978-1-4673-0973-8, pp: 201-210.
- Yu, X., H. Rashtian, S. Mirabbasi, P.P. Pande and D. Heo, 2015. An 18.7-Gb/s 60-GHz OOK demodulator in 65-nm CMOS for wireless network-on-chip. IEEE. Trans. Circuits Syst. I Regul. Pap., 62: 799-806.

Investigation of Effective Hybrid FRP and Steel Reinforcement Ratio for Concrete Members

Hayder H. Alkhudery, Oday M. Albuthbahak and Haider A.A. Al-Katib
Faculty of Engineering, Kufa University, Najaf, Iraq

Abstract: Fiber Reinforced Material (FRP) has become an effective alternative reinforcement for concrete constructions owing to their higher ultimate strength in tension and strongly anticorrosive properties. However, the low ductility behaviour of FRP Reinforcing Concrete structure (FRP-RC) is appeared comparing to high ductility of conventional Steel Reinforcing Concrete structures (S-RC). Successfully hybridization of reinforcement by combining FRP and steel rebar presents effective solution to migrate corrosion of steel rebar in steel reinforcing concrete constructions and advanced ductility of FRP-RC structures. Also, a good improvement is achieved in structural performances of concrete beams in terms of ultimate strength, central deflection, number cracks and crack width. In this study, a theoretical prediction analysis model of Hybrid Reinforcing Concrete (H-RC) members reinforced by FRP and steel rebar presents. The moment-curvature relationship is predicted to achieve load-deflection curve and bending moment capacity of beams. A good agreement has been indicated from verification between results of predicted theoretical model and 19 experimental results of other researcher. The hybrid reinforcement in HRC section can be replaced by equivalent steel ratio to conduct limitations of reinforcement ratio. The ductility index of H-RC members is measured to reference S-RC member while the ultimate strength index is measured to reference FPRC member. A critical limitations of (A_f/A_s) ratio is found for H-RC members for specified mechanical properties of concrete, steel and FRP rebar also, this limitations are affected by the equivalent steel ratio of hybrid concrete section. A parametric study of different parameters specially that influencing the amount of hybridization reinforcing ratio (A_f/A_s) is involved to explore the critical margin of (A_f/A_s) ratio utilizing the ductility and ultimate strength requirements. Also, various modes of failure were conducted to the corresponding area of reinforcing FRP and steel rebar.

Key words: Theoretical prediction model, FRP and steel reinforcements, hybrid reinforced concrete members, flexural investigation, ductility index, ultimate strength

INTRODUCTION

In the last decades Fiber Reinforcing Polymers (FRPs) are using as reinforcing materials for the concrete constructions because of their higher corrosion resistance and a high ratio of strengthen to weight. Moreover, they are nonconductive to electrical power and nonmagnetic, making them attractive as alternatives to traditional steel reinforcing rebar in concrete structures with severe environments or special situations such as transparency is required. Besides all these advantages, FRP rebar have mainly two disadvantages which are brittleness and low elasticity modulus. For pure FRP-Reinforcing Concrete (FRP-RC) beams, elastic behaviour of FRP rebar is linearly upward to failure (i.e., without yield point) which is caused brittle failure mode without warning rather than ductile. Codes of design adopted over reinforcing FRP principle for designing criterion which is further gradual and reduced probability to sudden collapse with more deformability grade. Furthermore, ductility lack for FRP rebar due to low of elasticity constant comparing with steel rebar, consequently, FRP-RC members suffer

large deflection and widening crack width than Steel Reinforcing Concrete (S-RC) members having the same cross section and reinforcing ratio. However, designing criterion of FRP-RC structures based on service ability limit states which ensure structural behaviour instead of the strength to verify safety and functional of the structures within expect life. Structural performances of concrete structures are enhanced by merging FRP and steel reinforcements (hybrid reinforcements) in order to migrate corrosion problem caused by Steel Rebar in SRC structures and advanced ductility of FRP-RC structures. Hybrid reinforcements (FRP and steel rebar) present a realistic and appropriate solution for structural concrete design. Tan (1997) investigate hybrid reinforced concrete members were reinforced by Aramid-Fiber Reinforcing Polymer (AFRP) and steel rebar. He indicated that sufficient serviceability parameters have been utilized when the amount of AFRP reinforcement area less than one half of the whole reinforcement (Tan, 1997). Aiello and Ombres (2002) concluded by placing steel rebar on interior level and FRP rebar close to exterior face of tension region, the stiffness of hybrid reinforcing

members is increased. Thus, the steel reinforcement within hybridization section have less involvement to the ultimate strength, whilst it provides an effective contribution regarding ductile and structural stiffness performance also, the crack widening and spacing is diminished (Aiello and Ombres, 2000). Leung and Balendran (2003) were explored the flexure behaviour of H-RC members reinforced by GFRP and steel rebar which located at different effective depth. They indicated that at stage of loading after yielding of steel rebar is occurred, the ultimate strength of H-RC member is increased as the FRP reinforcement strength is increased (Leung and Balendran, 2003). Qu *et al.* (2009), carried out experimental program for H-RC members with hybrid reinforcing steel and (GFRP) rebar, they indicated by presenting a sufficient amount of steel to GFRP reinforcement, the beam's ductile and serviceability can be enhanced. Lau and Pam (2010) were reported the hybridization of reinforcement strongly improved the ductility of H-RC members in comparison with FRP-RC members throughout experimental program as well. Hawileh (2015) studied H-RC beams using numerical approach represented by FEA model with AFRP and steel reinforcements, he found that the compressive strength of concrete has moderate effected on flexural behaviour features (Hawileh, 2015). Ilker *et al.* presented a numerical model for H-RC section. They found that the FRP reinforcement a core of loading resistance after steel is yielding (Kara *et al.*, 2015).

Previous research programs were concentrated mainly on the strength evaluation, failure models, load-deflection response and cracking patterns in H-RC beams. These beams were reinforcing by various FRP rebar types and containing a large range of reinforcing ratio with dissimilar in geometry and mechanical parameters. The obtaining tested results ensured the efficiency of steel rebar to improve the structural behaviour of H-RC beams in terms of ductility index and serviceability limitations comparing with FRP-RC members. However, the ductility issue remains focal of structural capability aspects of H-RC members, yet, needed more investigation. It stills unrecognized the conditions to be introduced H-RC members having ductility features similar to that of S-RC members. Extra investigation of experimental and theoretical programs should be provided to clarify reinforcement hybridization features.

In this study, a theoretical model has been introduced to study the flexure behaviour of H-RC concrete members. This model based on forces equilibrium and compatibility of strain to establish the moment-curvature relations (M- θ), the load-deflection and flexural capability of H-RC beams. A good agreement has been indicated from

verification between results of predicted theoretical model and 19 experimental results of other researcher. The hybrid reinforcement in H-RC section can be replaced by equivalent steel ratio to conduct limitations of reinforcement ratio. The ductility index of H-RC members is measured to reference S-RC member while the ultimate strength index is measured to reference FPR-RC member. A critical limitations of A_f/A_s ratio is found for H-RC members with specified mechanical properties of concrete, steel and FRP rebar and this limitations are affected by the equivalent steel ratio of hybrid concrete section. Also, a parametric study of different parameters specially that influencing the amount of hybrid reinforcing ratio (A_f/A_s) is involved to explore the critical margin of A_f/A_s ratio utilizing the ductility and ultimate strength requirements. Also, various modes of failure were conducted to the corresponding area of reinforcing FRP and steel rebar.

MATERIALS AND METHODS

Theoretical model: A theoretical model is introduced to formulate a flexure response of FRP/steel reinforced concrete (i.e., H-RC) beam. Using equilibrium of forces and compatibility of strains, the moment curvature relation is established of H-RC member. Then mid-span deflection of the beam determined using relative curvature of mid-span. The adopted approach illustrated in more details as follows.

Stress-strain relationships of materials: In this investigation, the stress strain relations of implemented materials concrete and reinforcing steel and FRP rebar are presented in Fig. 1. Other material models can be adopted in proposed model. The adopted model of stress strain relation for concrete under compressive loading that proposed by Hognestad *et al.* as described in Eq. 1 and Fig. 1a (Qu *et al.*, 2009):

$$f_c(\epsilon_c) = f'_c \left(\frac{2\epsilon_c}{\epsilon_{co}} - \left(\frac{\epsilon_c}{\epsilon_{co}} \right)^2 \right) \text{ if } \epsilon_c \leq \epsilon_{co} \quad (1)$$

$$f_c(\epsilon_c) = f'_c \left(1 - 0.15 \frac{(\epsilon_c - \epsilon_{co})}{(\epsilon_{cu} - \epsilon_{co})} \right) \text{ if } \epsilon_{co} \leq \epsilon_c \leq \epsilon_{cu} \quad (2)$$

Where:

- f'_c = Cylinder concrete compressive strength
- ϵ_c = Concrete compressive strain
- ϵ_{co} (= 1.8 f'_c/E_c) = Concrete strain with peak stress
- ϵ_{cu} = 0.0038 = Ultimate concrete compressive strain
- E_c (= 4700 f'_c) = Concrete elasticity constant (ACI, 318-05, 2005)

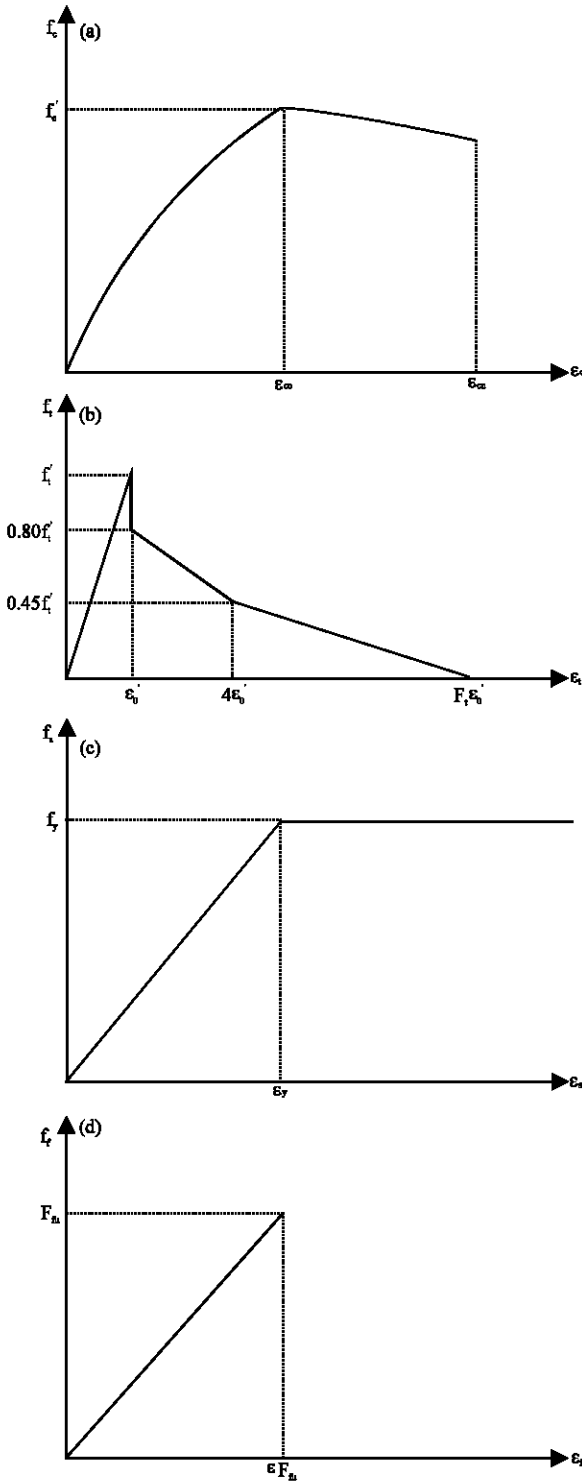


Fig. 1: Concrete, steel and FRP reinforcement stress strain relations (Qu *et al.*, 2009; Bischoff and Paixao, 2004; Nayal and Rasheed, 2006): a) Concrete in compression; b) Concrete in tension; c) Steel-reinforcement and d) FRP-reinforcement

Bischoff and Paixao (2004) determined the tensile strength for H-RC members, they concluded that the less concrete tensile strength is exhibited by S-RC beam in comparison with GFRP-RC members. Nayal and Rasheed (2006) investigated the behavior of S-and FRP-RC members in terms of tensile strength effects, they used a reverse integration technique of non-linear numeric evaluation and experiment results to provide a concrete tensile strength model for concrete member with hybrid reinforcements as described in Eq. 2 and shown in Fig. 1b:

$$f_t = E_c \epsilon_t \quad \text{if } 0 \leq \epsilon_t \leq \epsilon'_0 \quad (3)$$

$$f_t = 0.45 f'_t + 0.35 f'_t \left(\frac{4\epsilon'_0 - \epsilon_t}{3\epsilon'_0} \right) \quad \text{if } \epsilon'_0 \leq \epsilon_t \leq 4\epsilon'_0 \quad (4)$$

$$f_t = 0.45 f'_t + 0.35 f'_t \left(\frac{F_t \epsilon'_0 - \epsilon_t}{F_t \epsilon'_0 - 4\epsilon'_0} \right) \quad \text{if } 4\epsilon'_0 \leq \epsilon_t \leq F_t 4\epsilon'_0 \quad (5)$$

$$F_t = 124 - 0.05 \lambda \quad (6)$$

$$\lambda = \frac{E_s A_s + E_f A_f}{n_s d_s + n_f d_f} \quad (7)$$

Where:

- f_t = Concrete tension stress
- ϵ_t = Concrete tension strain
- $\epsilon'_0 = f'_t / E_c$, f'_t (= $0.263 \sqrt{f'_c}$) = Concrete tension strength
- F_t = Proportional factor of the maximum strain in tension to the strain of concrete cracking
- E_s = Elasticity constant of steel rebar
- A_s = Steel area
- E_f = Elasticity constant of FRP rebar
- A_f = FRP area
- n_s = Steel rebar number
- d_s = Steel rebar diameter
- n_f = FRP rebar number
- d_f = Rebar FRP diameter
- λ = kN/mm
- F_t = kN/mm

The multilinear descending mode of concrete in tension reflects initially and secondary cracking of concrete. Researchers were suggested special values, $F_t = 100$ to the FRP rebar and $F_t = 10$ to the steel rebar. Since, the parameter F_t designate the variance of S-and FRP-RC members in tensile strength. Also, factor λ , collected the sectional geometry parameters which showing the tensile strength behaviour and that is

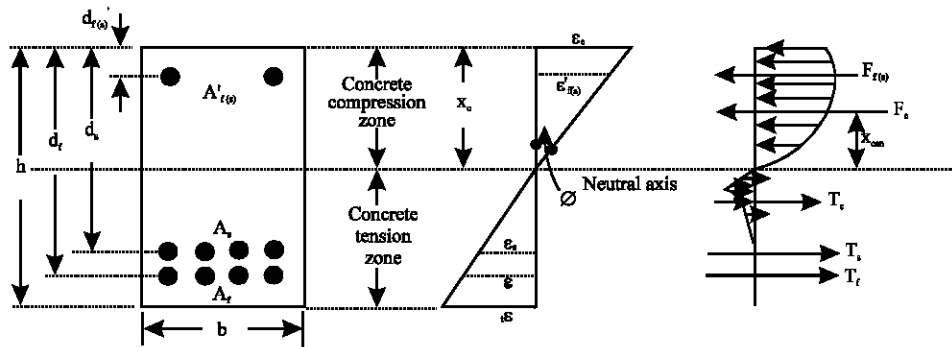


Fig. 2: Hybrid FRP and steel reinforcing concrete section stress, strain distribution and forces: a) Hybrid reinforcing concrete section; b) Strain division and c) Stress division and forces

explained the variance of F_s magnitudes for steel in comparison with FRP reinforcing rebar (Bischoff and Paixao, 2004). A bi-linear elastic perfect plastic is the adopted material model of reinforcing steel with yield stress f_y as shown in Fig. 1c and given in Eq. 8 and 9:

$$f_s = E_s \epsilon_s \quad \text{if } 0 \leq \epsilon_s \leq \epsilon_y \quad (8)$$

$$f_s = f_y \quad \text{if } \epsilon_s > \epsilon_y \quad (9)$$

Where:

- f_s = Steel tension stress
- E_s = Steel elasticity constant
- ϵ_s = Steel tension strain
- ϵ_y = Steel yielding tension strain

And linear elastic up to rupture, the stress strain relationship of FRP rebar as presented in Fig. 1d and described by Eq. 10 and 11:

$$f_f = E_f \epsilon_f \quad \text{if } \epsilon_f \leq \epsilon_{fu} \quad (10)$$

$$f_f = 0 \quad \text{if } \epsilon_f > \epsilon_{fu} \quad (11)$$

Where:

- f_f = Tensile stress in FRP rebar
- E_f = FRP modulus of elasticity
- ϵ_f = Tensile strain in FRP rebar
- ϵ_{fu} = Failure tensile strain for FRP rebar
- F_{fu} = Failure tensile stress for FRP rebar

Moment-curvature relationship: Figure 2 shows H-RC rectangular beam section having steel and FRP reinforcing rebar at top and bottom, the process of analysis begins by supposing a little initial strain value for outer concrete level of compression zone (i.e., $\epsilon_c = 0.00001$) with incremental strain 0.00001 upward till concrete crushing strain ($\epsilon_{cu} = 0.0038$). For each value of assumed strain ϵ_c one loop carried out through repeated steps. Initially, the

depth of neutral axis is assumed (x_c). Consequently, strains for the extreme fiber of concrete in tension zone and FRP-/steel-reinforcement rebar at top and bottom level is determined. The consistent assuming criterions comprised, the study plane before bending is continues plane after bending and perfectly bonding through surfaces of reinforcements and surrounding concrete. Next these strains values used to calculate the corresponding stresses for each material using their coincided stress-strain relationship presented in Fig. 1. Then, the relative forces are determined and the equilibrium condition of forces is satisfied. These steps are repeated using iterative method called bi-section method in which two values of the neutral axis depth (x_c) is assumed one for each sides of the correct value (i.e., $x_{c, top} = 0$ at the top level and $x_{c, bottom} = h$ where h = overall section depth). Finally, the corrected value is utilized with accuracy of $\xi = \pm 0.0001$ for equilibrium forces condition.

The strain at each level in hybrid reinforced concrete section is linear proportion to its dimension from the neutral axis as shown in Fig. 2, the tensile strain in concrete at extreme fiber in tension zone described as follows:

$$\epsilon_t = \frac{h-x_c}{x_c} \epsilon_c \quad (12)$$

Since, a fully bonding is assumed between reinforcement rebar and surrounding concrete, thus, the strain initiated in tension and compression steel and FRP rebar stated in Eq. 13-16 as follows:

$$\epsilon'_f = \frac{d'_f - x_c}{x_c} \epsilon_c \quad (13)$$

$$\epsilon_f = \frac{d_f - x_c}{x_c} \epsilon_c \quad (14)$$

$$\epsilon'_s = \frac{d'_s - x_c}{x_c} \epsilon_c \quad (15)$$

$$\epsilon_s = \frac{d_s - x_c}{x_c} \epsilon_c \quad (16)$$

Where:

ϵ'_f and ϵ_f = The strain at top and bottom FRP reinforcing rebar

d'_s and d_s = The top and bottom FRP-reinforcing depth, respectively

ϵ'_s and ϵ_s = The strain at top and bottom steel-reinforcing rebar

d'_s and d_s = The top and bottom steel-reinforcing depth, respectively

To determine the resultant for the internal concrete compression forces at any loading stage, rational integration approach is used along the neutral axis depth with transformation the compression concrete stress (f_c) in term of x_c instead of ϵ_c as described in Eq. 17-19:

$$F_c = b \int_0^{\epsilon_{ca}} f_c(\epsilon_c) d\epsilon_c = b \int_0^{x_{co}} f_c(x_c) \frac{\epsilon_{co}}{x_{co}} dx \quad (17)$$

$$f_c(x_c) = f'_c \left(\frac{2 x_c}{x_{co}} - \left(\frac{x_c}{x_{co}} \right)^2 \right) \text{ if } x_c \leq x_{co} \quad (18)$$

$$f_c(x_c) = f'_c \left(1 - 0.15 \frac{(x_c - x_{co})}{(x_{cu} - x_{co})} \right) \text{ if } x_{co} < x_c \leq x_{cu} \quad (19)$$

Where:

F_c = Compression forces resultant of concrete

b = Beam width

ϵ_{ca} = Assumed concrete strain

x_{co} = Neutral axis depth at ϵ_{co}

x_c = Neutral axis depth at ϵ_c

$f_c(x_c)$ = Compressive stress function in terms of neutral axis depth

x_{cu} = Neutral axis depth at ϵ_{cu}

The tension forces resultant of concrete (T_c) is determined based on the internal tensile stress of concrete in tension zone with corresponding value of tensile strain (ϵ_t) and (ϵ_c) (Eq. 12). Simple triangle and trapezoidal relationships are used to evaluate internal forces as shown in Fig. 3, then concrete tension resultant is obtained at each stage of loading as described in Eq. 20-23:

$$T_c = \frac{b}{2} \times x_t \times f_t(\epsilon_t) \text{ if } 0 \leq \epsilon_t \leq \epsilon'_0 \quad (20)$$

where $x_t = (h - x_c)$.

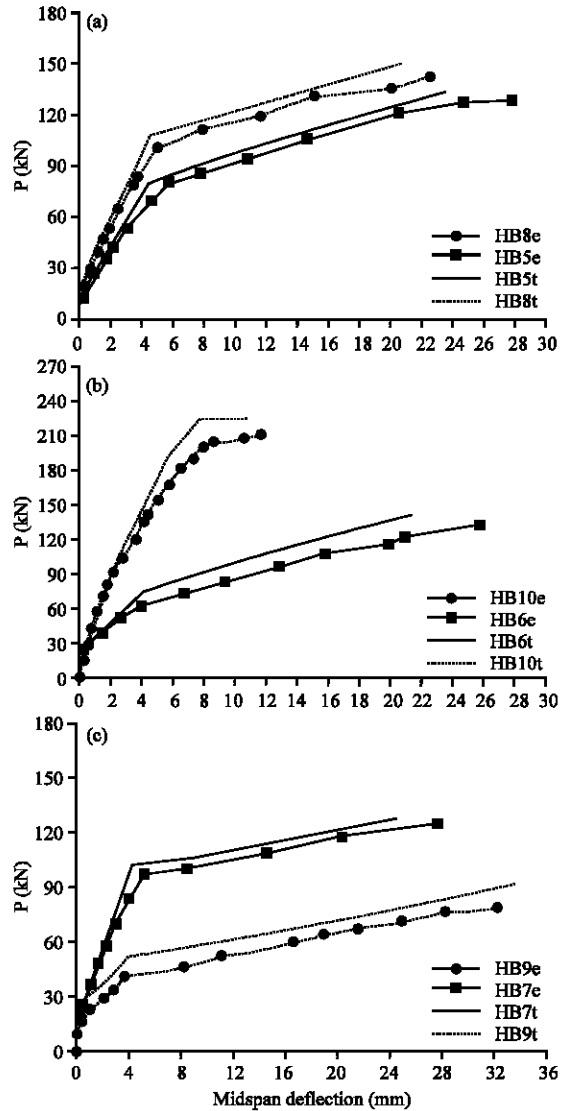


Fig. 3: a-c) Comparisons of load deflection curves of experimental tests (Qu *et al.*, 2009) and prediction theoretical model (Qu *et al.*, 2009)

$$T_c = T_{co} + \frac{b}{2} \times (x_t - x_{t0}) \times (0.8 f'_t + f_t(\epsilon_t)) \text{ if } \epsilon'_0 < \epsilon_t \leq 4\epsilon'_0 \quad (21)$$

where, $T_{co} = b/2 \times x_{t0} \times f'_t$, $x_{t0} = x_t \times \epsilon'_0 / \epsilon_t$.

$$T_c = T_{co} + T_{co4} + \frac{b}{2} \times (x_t - x_{t04}) \times (0.45 f'_t + f_t(\epsilon_t)) \text{ if } 4\epsilon'_0 < \epsilon_t \leq F_t \epsilon'_0 \quad (22)$$

where, $T_{co4} = b/2 \times (x_{t04} - x_{t0}) \times 1.25 f'_t$, $x_{t04} = x_t \times 4\epsilon'_0 / \epsilon_t$.

$$T_c = T_{co} + T_{co4} + T_{coFt} \text{ if } \epsilon_t > F_t \epsilon'_0 \quad (23)$$

where, $T_{coFt} = b/2 \times (x_{coFt} - x_{t04}) \times 0.45 f'_t$, $x_{coFt} = x_t \times F_t \epsilon'_0 / \epsilon_t$.

The components of force for the FRP and steel rebar for the top and bottom of cross section are calculated according to the reinforcing as illustrated in Eq. 24-31:

$$T_s = A_s E_s \epsilon_s \text{ if } 0 \leq \epsilon_s \leq \epsilon_y \quad (24)$$

$$T_s = A_s f_y \text{ if } \epsilon_s > \epsilon_y \quad (25)$$

$$C_s = A'_s E_s \epsilon'_s \text{ if } 0 \leq \epsilon'_s \leq \epsilon_y \quad (26)$$

$$C_s = A'_s f_y \text{ if } \epsilon'_s > \epsilon_y \quad (27)$$

$$T_f = A_f E_f \epsilon_f \text{ if } 0 \leq \epsilon_f \leq \epsilon_{fu} \quad (28)$$

$$T_f = 0 \text{ if } \epsilon_f > \epsilon_{fu} \quad (29)$$

$$C_f = A'_f E_f \epsilon'_f \text{ if } 0 \leq \epsilon'_f \leq \epsilon_{fu} \quad (30)$$

$$C_f = 0 \text{ if } \epsilon'_f > \epsilon_{fu} \quad (31)$$

Where:

- T_s = Tensile force of bottom steel rebar
- A_s = Reinforcing area of bottom steel rebar
- C_s = Compressive force of top steel rebar
- A'_s = Reinforcing area of top steel rebar
- T_f = Tensile force of bottom FRP rebar
- A_f = Reinforcing area of bottom FRP rebar
- C_f = Compressive force of top FRP rebar
- A'_f = Reinforcement area of top FRP rebar
- E_f = Elasticity modulus of FRP rebar (assuming same behaviour of FRP rebar in tension and compression)

Taking into account the equilibrium condition of forces, the forces resultant of hybrid reinforced cross section shown in Fig. 2 must be equated to required accuracy (i.e., $\xi = \pm 0.0001$) in the prediction theoretical model as given in Eq. 32:

$$F_c + C_s + C_f - T_s - T_f = \xi \quad (32)$$

Equation 32 represents the final repeating steps in iterative process (i.e., bi-section method) to find the corrected value of neutral axis depth x_c where the sufficient equilibrium accuracy is reached. From the strain distribution shown in Fig. 2b, the beam curvature ϕ can be evaluated as described in Eq. 33:

$$\phi = \frac{\epsilon_c}{x_c} \quad (33)$$

Since, all internal forces of hybrid cross section are established for assumed value of compression concrete strain at extreme top fiber (ϵ_c), then the corresponding stage of loading can be identified by calculating the applied moment M_{app} by taking a moment about any horizontally axis such as neutral axis as specified in Eq. 34:

$$M_{app} = F_c (x_{cen}) + C_s (x_c - d'_s) + C_f (x_c - d'_f) + T_s (d_s - x_c) + T_f (d_f - x_c) \quad (34)$$

Where:

M_{app} = Applied moment of H-RC section

x_{cen} = Distance between the axis of concrete compressive resultant and neutral axis

Mid-span deflections: Since, the moment-curvature relationship of H-RC beam has been provided, the flexural rigidity of the beam EI_{eff} can be developed at each stage of loading as in Eq. 35:

$$EI_{eff} = \frac{M}{\phi} \quad (35)$$

The central span deflection (Δ) of H-RC members can be calculated at each loading stage using the formula of elastic deflection with specific loading system, instance, the mid-span deflection of simple supporting members having two-point loading ($P/2$) can be obtained using Eq. 36:

$$\Delta = \frac{23}{1.296} \frac{PL^3}{EI_{eff}} \quad (36)$$

Where:

Δ = Central span deflection

P = Applied load

L = Member span

For simply supported members with central point load (P), the mid-span deflection can be determined as illustrated in Eq. 37:

$$\Delta = \frac{1}{48} \frac{PL^3}{EI_{eff}} \quad (37)$$

RESULTS AND DISCUSSION

Verification of theoretical predicted model with test results: Verification of the theoretical model of hybridization FRP/steel reinforcing concrete member is carried out through a comparison of the prediction model results and the tested results for the same cross sectional

shape (i.e., rectangular cross section) with various conditions of geometrical, mechanical properties and loading pattern that is to examine the suitability of adopted model to analyze H-RC members. Table 1 shows the geometry and mechanical details of (19) H-RC members. Mainly, comparisons are introduced in terms of the bending moment capacity ratio and load deflection responses that is to investigate the coincidence degree of the prediction technique in applications.

Flexural capacity ratio: Since, all tested beams were designed to fail in flexural failure as described in literature, the flexural capacity of each beam is compared between the prediction theoretical model and tested beams as illustrated in Table 2 where $M_{u, th}$ is the ultimate moment of theoretical model and $M_{u, exp}$ is the ultimate moment of experiment. The statistic coefficients mean, standard

deviation and variation factor of the flexural capacities ($M_{u, th}/M_{u, exp}$) are 1.01, 11.88 and 1.41%, respectively and the variation of ratio ($M_{u, th}/M_{u, exp}$) in range of 0.79-1.12 reasonably well except HB10. As seen, a good agreement is achieved through comparison results, so, the theoretical model can be used in analysis of hybrid reinforcement concrete members.

Load-deflection response: The load-deflection relationship of H-RC member reflects the structural behaviour of material combination, concrete, steel and FRP reinforcement. Figure 3 presents the load deflection curve obtained using adopted technique in comparison with the results of HB5-HB10 which are tested by Qu *et al.* (2009). These tested results represent a good sample to proof the validity of prediction theoretical model because it contains a wide range values of the

Table 1: Geometric and mechanical details of tested HRC-beams in literature

Beam No.	b×h×L (mm)	n _s b _{as} (mm)	A _s (mm ²)	b _{af} (mm ²)	A _r (mm ²)	A _s ' (mm)	E _s (GPa)	F _y (MPa)	E _r (GPa)	f _{tu} (MPa)	f _c ' (MPa)	f _c ' (MPa)	Load system
HB1 (Aiello and Ombres 2000)	b=150	2ø8	100.5	27.5	88.3	100.50	200	465	49.0	1674	36.56	4.03	4-point
HB2 (Aiello and Ombres 2000)	h=20	2ø8	100.5	210	157.0	100.50	200	465	50.1	1366	36.56	4.03	4-point
HB3 (Aiello and Ombres 2000)	L=2700	2ø8	100.5	310	235.5	100.50	200	465	50.1	1366	36.56	4.03	4-point
HB4 (Aiello and Ombres 2000)		2ø8	100.5	27.5	88.3	100.50	200	465	49.0	1674	36.56	4.03	4-point
HB5 (Qu <i>et al.</i> , 2009)		2ø12	226.1	212.7	253.2	-	200	363	45.0	782	26.48	10.31	4-point
HB6 (Qu <i>et al.</i> , 2009)	b=180	1ø16	201.0	215.9	396.9	-	200	336	41.0	755	26.48	10.31	4-point
HB7 (Qu <i>et al.</i> , 2009)	h=250	2ø16	401.9	29.5	141.7	-	200	336	37.7	778	27.52	10.58	4-point
HB8 (Qu <i>et al.</i> , 2009)	L=2100	2ø16	401.9	212.7	253.2	-	200	336	45.0	782	27.52	10.58	4-point
HB9 (Qu <i>et al.</i> , 2009)		1ø12	113.0	29.5	141.7	-	200	363	37.7	778	32.52	11.82	4-point
HB10 (Qu <i>et al.</i> , 2009)		6ø16	1205.8	215.9	396.9	-	200	336	41.0	755	32.52	11.82	4-point
HB11 (Lau and Pam, 2010)	b=280	2ø25	981.7	119	283.5	56.55	181	336	39.5	588	41.30	13.21	3-point
HB12 (Lau and Pam, 2010)	h=380	2ø20	628.3	225	981.7	56.55	190	597	38.0	582	39.80	12.84	3-point
HB13 (Lau and Pam, 2010)	L=4200	2ø25	981.7	219	567.1	56.55	196	550	39.5	588	44.60	12.99	3-point
HB14		1ø10	78.5	212	226.2	100.50	200	520	50.0	1000	40.00	11.70	4-point
HB15	b=230	2ø10	157.1	212	226.2	100.50	200	520	50.0	1000	40.00	11.70	4-point
HB16	h=300	2ø12	226.2	212	226.2	100.50	200	520	50.0	1000	40.00	11.70	4-point
HB17	L=3700	2ø10	157.1	216	402.1	100.50	200	520	50.0	1000	40.00	11.70	4-point
HB18		2ø12	226.2	216	402.1	100.50	200	520	50.0	1000	40.00	11.70	4-point
HB19		2ø16	402.1	216	402.1	100.50	200	520	50.0	1000	40.00	11.70	4-point

Table 2: Comparison ultimate bending strength of hybridization reinforcing concrete members

Beam No.	A _r /A _s	M _{u, exp} (kN.m)	M _{u, th} (kNm)	(M _{u, th} /M _{u, exp})
HB1 (Aiello and Ombres 2000)	0.88	25.14	19.79	0.79
HB2 (Aiello and Ombres 2000)	1.56	28.41	25.21	0.89
HB3 (Aiello and Ombres 2000)	2.34	35.55	29.62	0.83
HB4 (Aiello and Ombres 2000)	0.88	25.14	20.94	0.83
HB5 (Qu <i>et al.</i> , 2009)	1.12	38.28	39.68	1.04
HB6 (Qu <i>et al.</i> , 2009)	1.98	39.66	42.76	1.08
HB7 (Qu <i>et al.</i> , 2009)	0.35	36.36	38.95	1.07
HB8 (Qu <i>et al.</i> , 2009)	0.63	42.57	44.09	1.04
HB9 (Qu <i>et al.</i> , 2009)	1.25	23.55	29.86	1.27
HB10 (Qu <i>et al.</i> , 2009)	0.33	63.30	70.83	1.12
HB11 (Lau and Pam, 2010)	0.29	147.00	157.49	1.07
HB12 (Lau and Pam, 2010)	1.56	261.00	226.6	0.87
HB13 (Lau and Pam, 2010)	0.58	229.00	238.75	1.04
HB14	2.88	47.62	51.86	1.09
HB15	1.44	53.55	57.06	1.07
HB16	1.00	58.94	61.73	1.05
HB17	2.56	68.30	71.03	1.04
HB18	1.78	64.71	68.23	1.05
HB19	1.00	83.53	86.21	1.03
The average	1.01			

hybrid reinforcement ratio between reinforcing area of FRP (A_f) and reinforcing area of steel (A_s) under same circumstance geometrical and mechanical conditions. Although, the samples chosen in Fig. 3a-c have considerable difference in hybrid reinforcement ratio (A_f/A_s), the comparison of corresponding curves of experimental results and theoretical model are convergent with acceptable degree. Thus, the prediction theoretical model can be used in applications and parametric studies to distinguish the effectiveness of mechanical and geometrical properties of hybrid reinforcements members.

In general view, the load deflection curve of H-RC section contains three stages, namely uncrack section stage, cracked section stage and yielding section stage as shown in Fig. 3. Each phase affected by geometrical and mechanical properties of collective materials and relative relation among them. Initially, first stage is a linear part terminated when the members starting cracked and secondly stage part had a reduction in grade than the first line owing to decrease in beam stiffness after cracking. Finally, yield section segment started when steel reinforcement is yielded which caused higher reduction in beam stiffness and FRP reinforcement plays the role in resisting the load in this stage.

Stiffness tension effects: Including the tension stiffness in theoretical predictions increased the accuracy of theoretical results especially up to the stage of serves loading (i.e., about 65% of ultimate load), since, the effects of tension stiffness is decreased with increased loading (Qu *et al.*, 2009). The predicted theoretical model is employed to study the effects of tension stiffness in structural performance H-RC members on the load deflection response as presented in Fig. 4. The geometrical and mechanical properties of case study used in Fig. 4 are simply supported beam with four-point loading, the dimensions $L = 4500$ mm, $b = 250$, $h = 400$ mm, $d_s = d_t = 357$ mm and the mechanical properties of concrete $f'_c = 28$ MPa, $f'_t = 3.28$ MPa steel rebar $f_y = 420$ MPa, $E_s = 200$ GPa and GFRP rebar $f_{ts} = 1000$ MPa, $E_f = 40$ GPa. In design calculation, the tension stiffness in concrete is neglected for simplicity in this case the first linear segment in load-deflection curve is omitted and the load-deflection curve become two linear parts divided by yielding point as shown in Fig. 4. Also, a small reduction in stiffness beam, yield strength is appeared while approximately the same ultimate strength value as seen. In addition, the tension stiffness is affected by the area of steel/FRP reinforcement with same (A_f/A_s), so that as the area of reinforcement increased the effects of tension

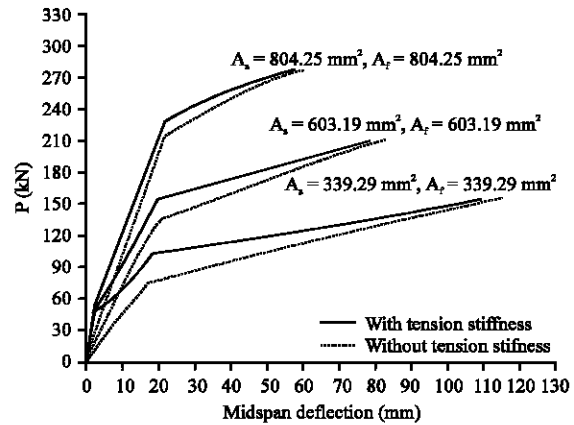


Fig. 4: Tension stiffness effects on load deflection curve of H-RC member

stiffness is decreased due to increasing in flexural capacity which is diminished the effects of tension stiffness.

Frp reinforcement type effects: The type of FRP reinforcement in H-RC members effects shown in Fig. 5. The same dimensions and material properties of case study are used with the same area of reinforcements $A_f = 603.19$ mm², $A_s = 603.19$ mm² and AFRP rebar $f_{ts} = 1300$ MPa, $E_f = 60$ GPa and CFRP rebar $f_{ts} = 1500$ MPa, $E_f = 120$ GPa. The effects of FRP reinforcement is null in the first stage of load-deflection curve because of low loading phase and uncrack concrete section while the effects are slightly improvements in the second stage owing to higher value of elastic modulus of steel reinforcement compared with those elastic moduli of FRP reinforcement and clearly effects of FRP reinforcements appeared on the third stage of load-deflection curve due to yielding of steel rebar and the FRP rebar become the core of reinforcements resistance at this stage.

Reinforcements ratio of H-RC section: The reinforcement ratio of H-RC section effectively affects the structural behaviour of members, since, the value of reinforcement ratio indicate the type of expected failure modes. Also, the structural behaviour of H-RC section with any (A_f/A_s) ratio is similar of S-RC members with the same equivalent steel reinforcement ratio ($\rho_{eq,s}$), since, they coincided up to the yield point stage of loading which is an important loading stage for design purpose as shown in Fig. 6. The difference between them occurred in the third stage of load-deflection curve due to present FRP-reinforcement in H-RC section which can be considered as the reserve strength. Thus, it is useful to specify the hybrid

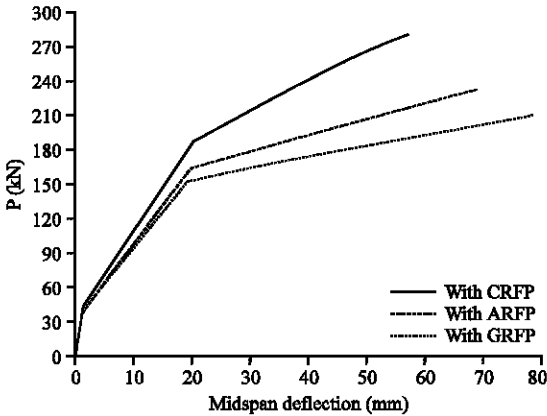


Fig. 5: Effects of FRP types on load deflection curve of H-RC member

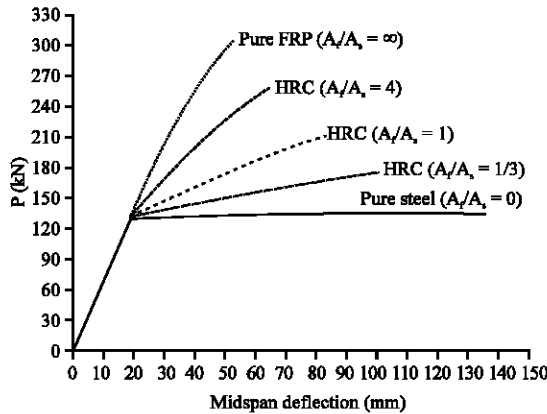


Fig. 6: Load-deflection curve of H-RC section with same equivalent reinforcement ratio ($\rho_{eq,s} = 0.811\%$) and different (A_f/A_s) ratio

reinforcement ratio in H-RC section in terms of the equivalent steel reinforcement ratio ($\rho_{eq,s}$) rather than the equivalent FRP reinforcement ratio ($\rho_{eq,s}$), since, the latter linear up to failure (i.e., without yield point). The equivalent steel or FRP reinforcement ratio of H-RC section are defined as following:

$$\rho_{eq,s} = \rho_s + \rho_f \frac{E_f}{E_s} \quad (38)$$

$$\rho_s = \frac{A_s}{b d_s} \quad (39)$$

$$\rho_f = \frac{A_f}{b d_f} \quad (40)$$

Where:

ρ_s = steel reinforcement ratio

ρ_f = FRP reinforcement ratio

$$\rho_{eq,f} = \rho_s \frac{E_s}{E_f} + \rho_f \quad (41)$$

For clarification the effects reinforcements ratio on structural behaviour of H-RC members, the tension stiffness is neglected. Figure 6 presents the load deflection curves of case study of H-RC section with different values of (A_f/A_s) ratio, pure FRP-and pure steel- reinforcement sections are included with the same equivalent steel ratio (i.e., $\rho_{eq,s} = 0.811\%$) for all cases as details shown in Table 3.

It can be observe that the difference in behaviour among all cases under consideration in Fig. 6 commenced at the point of steel yielding. Although, all cases of H-RC beam have the same $\rho_{eq,s}$ ($= 0.811\%$) but the ultimate strength is increased as the (A_f/A_s) ratio increased owing to increase in the amount area of FRP-reinforcement.

Failure modes of hybrid-RC section: The prediction theoretical model is employed to indicate the failure mode of H-RC members. The failure can be happened in H-RC section when one or more of material components fail individual or simultaneously. The reinforcement ratio of H-RC section effectively affected the governing type of failure mode which can be distinguished according to limitations of reinforcement ratio specified in Eq. 45. When concrete crushing with steel reinforcement yielding simultaneously at which FRP reinforcement not ruptured yet in this case the balanced steel reinforcement ratio ($\rho_{s,b}$) can be calculated using consistent assumptions for equilibrium and compatibility relationships as follows:

$$\rho_{s,b} = 0.85 \beta_1 \frac{f'_c}{f_y} \frac{E_s \epsilon_{cu}}{f_y + E_s \epsilon_{cu}} \quad (45)$$

Where, $\beta_1 = 0.85 - 0.05 f'_c - 28/7$ when steel reinforcement yielding, then concrete crushing with FRP reinforcement ruptured simultaneously in such case the balanced FRP reinforcement ratio ($\rho_{f,b}$) can be obtained as:

$$\rho_{s,b} = 0.85 \beta_1 \frac{f'_c}{Ff_{fb}} \frac{E_s \epsilon_{cu}}{Ff_{fb} + E_s \epsilon_{cu}} \quad (46)$$

In addition, a critical condition (i.e., “Not balanced” and it is occurred in tension zone) can be utilized when steel reinforcement yielding with FRP reinforcement ruptured simultaneously at which concrete crushing not yet occurred the critical steel-FRP reinforcement ratio (ρ_{stc}) can be obtained as:

Table 3: Reinforcements details of HRC-section with same value of equivalent reinforcement ratio ($\rho_{eq,s}$)

Case No.	A_s (mm ²)	ρ_s (%)	A_f (mm ²)	ρ_f (%)	$\rho_{eq,s}$ (%)	$A_{eq,s}$ (mm ²)	A_f/A_s
Pure-steel	723.83	0.811	0	0	0.811	723.83	0
HRC	678.58	0.760	226.250	0.254	0.811	723.83	1/3
HRC	603.19	0.676	603.190	0.676	0.811	723.83	1
HRC	402.12	0.451	1608.550	1.8020	0.811	723.83	4
Pure-FRP	0.00	0.000	3619.150	4.0550	0.811	723.83	8

$$\rho_{sf,c} = \frac{f_y A_s + F_{fu} A_f}{f_{fu} b d_f} = \rho_s \frac{f_y}{f_{fu}} + \rho_f \quad (47)$$

According to the first material component failure and limitations of reinforcements ratio specified in Eq. (45-47) the failure modes of H-RC members can be illustrated as following.

Concrete crushing failure mode: This undesirable failure occurred if the equivalent steel reinforcement ratio ($\rho_{eq,s}$) is greater than the balanced steel reinforcement ratio ($\rho_{s,b}$) (i.e., over reinforced section) in which concrete crushing while steel-and FRP-reinforcements not yielding not ruptured yet, respectively.

FRP rupturing failure mode: This inadmissible failure mode occurred when ($\rho_{eq,s}$) ratio is less than (ρ_{fb}) or ($\rho_{sf,c}$) ratio is less than (ρ_{fb}).

Steel yielding failure mode: This is ductile and desirable failure mode, since, yielding of steel causes extensive previous warning. It is occurred when ($\rho_{eq,s}$) ratio is less than ($\rho_{s,b}$) (i.e., under reinforced section) and to control steel reinforcing yield followed by concrete crushing rather than FRP rebar ruptured, the value of ($\rho_{sf,c}$) ratio must be greater than (ρ_{fb}).

Flexural design guidance of H-RC section singly reinforced: To mitigate the variation of practical strength of S-RC section and to ensure ductile failure mode ACI 318M-14 adopted approach of tension controlled section, so that, the nominal flexural strength determined when the tensile strain of outer steel layer is equal or more than ($\epsilon_s = 0.005$) at which the concrete crushing is occurred simultaneously (ACI., 2006). Using tension controlled section in H-RC section, the upper limit of ($\rho_{eq,s}$) ratio is the maximum equivalent steel reinforcement ratio ($\rho_{s,max}$) which is determined in terms of ($\rho_{s,b}$) as:

$$\rho_{eq,s} = \rho_s + \rho_f \frac{E_f}{E_s} \leq \rho_{s,max} \quad (48)$$

$$\rho_{s,max} = \rho_{s,b} \frac{E_s \epsilon_{cu} + f_y}{E_s \epsilon_{cu} + 0.005 E_s} \quad (49)$$

In other hand, the same adopted approach of ACI 440.1R-06 for FRP-RC section (ACI., 2006) is used to ensure concrete crushing prior FRP-rupturing failure mode. Thus, the maximum equivalent FRP reinforcement ratio (ρ_{fb}) which the upper limit of ($\rho_{eq,f}$) ratio is obtained regarding (ρ_{fb}) as follows:

$$\rho_{sf,f} \geq \rho_{fb,max} \quad (50)$$

$$\rho_{fb,max} = 1.4 \rho_{fb} \quad (51)$$

Also to ensure steel yielding occurred before FRP-rupturing failure mode, the condition of steel yielding failure mode illustrated as follows:

$$\rho_{sf,c} \geq \rho_{fb,max} \quad (52)$$

Basis on the concept of ACI 440-06 (i.e., $\phi Mn = Mcr$ where Mcr is the moment of cracking section) to specify the formula of minimum equivalent steel reinforcement ratio ($\rho_{s,min}$) for H-RC members as follows [12]:

$$\rho_{s,min} = \frac{0.25 \sqrt{f'_c}}{f_y} \geq \frac{1.4}{f_y} \quad (53)$$

Ductility index of H-RC member: The definition of ductility can be summarized as the ability of structure to withstand the applied loading up to the ultimate load with large deformation before collapse. In ordinary S-RC structure, ductility index is recognized by a ratio of ultimate to the first steel-reinforcement yield deformations. However, Rashid *et al.* (2005) concluded that the traditional definition of ductility index in terms of deformability found to be inadequate for structure without yield point such as AFRP-RC beams. In energy basis, the ductility is identified as the ability of structure to absorbing the energy up to collapse. Tan (1997) suggested the ductility index as ratio between the total energy determined as area below the load-deflection curve and total elastic energy of corresponding case. Aiello and Ombres (2002) predicted (df) parameter which is represented as a ratio of area below the moment-curvature curve to that area determined up to

detecting curvature value. Qu *et al.* (2009) concluded that the ductility HRC members varied with (A_f/A_s) ratio and modulus of elasticity for FRP and steel reinforcements. Pang *et al.* (2015) proposed a definition of ductility index (μ_h) with energy and deformability basis for H-RC members as follows (Kara *et al.*, 2015):

$$\mu_h = \psi \frac{D_{uh}}{D_{uy}} \geq (\mu_D) \quad (54)$$

$$\psi = \frac{U_H}{U_s} \leq 1.0 \quad (55)$$

Where:

- D_{uh} = Curvature, rotations and displacement at ultimate stage of loading
- D_{uy} = Curvature, rotations and displacement at first steel yielding
- H_h = Area below moment-curvature curve of H-RC member
- U_s = Area below moment-curvature curve of equivalent S-RC member
- ψ = Factor of ductility reduction
- μ_D = Traditional definition of SRC beams (Kara *et al.*, 2015)

By reviewing the verification test results of last proposed ductility index, it is found that the factor of ductility (ψ) values gives a significant meaning for beam's ductility rather than the proposed ductility index (μ_h), since, the factor ψ is scaled the ductility index of H-RC member with respect to reference point (i.e., $\psi = 1.0$) represented by S-RC member reinforced with equivalent steel reinforcement ratio ($\rho_{eq,s}$). Figure 7 shows the modified ductility index (ψ) versus the (A_f/A_s) ratio for the case study under consideration with range values of ($\rho_{eq,s}$) ratio from ($\rho_{eq,s} = \rho_{s,min} = 0.33\%$) to ($\rho_{eq,s} = \rho_{s,b} = 3.01\%$). It seems that as the (A_f/A_s) increased the ductility of H-RC members decreased as expected, since, the amount of steel reinforcement decreased. In addition, the range of variation of ductility index (ψ) is reduced as the equivalent steel reinforcement ratio ($\rho_{eq,s}$) is increased because of increasing the amount of reinforcing steel rebar which eliminated the effect of FRP reinforcement in other words as the ($\rho_{eq,s}$) ratio increased the ratio between load causing yielding steel to ultimate load (P_f/P_s) is increased.

Effective range of (A_f/A_s) ratio of hrc member: In HRC members, the steel reinforcement is introduced to improve ductility and stiffness while the FRP reinforcements essentially to increase the ultimate strength and to migrate corrosion of steel reinforcement. These two criteria are

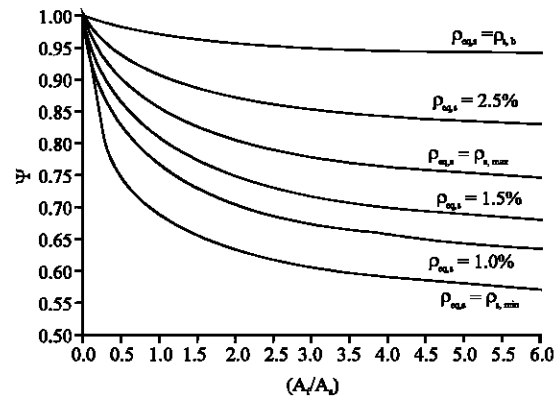


Fig. 7: Variation of ductility index (ψ) of H-RC case study member against (A_f/A_s) ratio

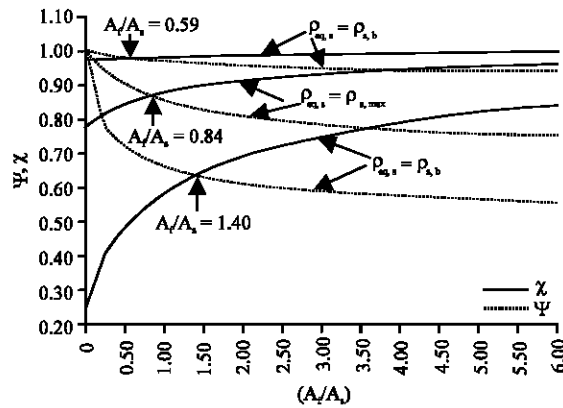


Fig. 8: Effective limits range of (A_f/A_s) ratio for HRC members

acted in opposite direction, thus by intersecting these two criteria with various amount of equivalent steel reinforcement ratio, the critical behaviour of H-RC beams can be achieved. The theoretical prediction model is used to investigate the effective range of (A_f/A_s) ratio in H-RC beams for the case study. Figure 8 represents the variation of ductility index (ψ) and normalized ultimate strength index ($X = P_{u,h}/P_{u,f}$) versus the (A_f/A_s) ratio of case study considered in this study. where, $P_{u,h}$ = Ultimate strength of H-RC members and $P_{u,f}$ = Ultimate strength of FRP-RC members reinforced by FRP reinforcement with the same equivalent steel reinforcement ratio ($\rho_{eq,s}$). Three values of ($\rho_{eq,s}$) ratio were used, namely ($\rho_{eq,s} = \rho_{s,min} = 0.33\%$, $\rho_{eq,s} = \rho_{s,max} = 1.98\%$ and $\rho_{eq,s} = \rho_{s,b} = 3.01\%$) to indicate limits of the critical range of ($\rho_{eq,s}$) ratio that meets the ductility index (ψ) and ultimate strength index (ψ) at same level as shown in Fig. 8. It shows that the critical range of the (A_f/A_s) ratio for the case study is limited between ($A_f/A_s = 1.40$ for $\rho_{eq,s} = \rho_{s,min} = 0.33\%$) and ($A_f/A_s = 0.59$ for $\rho_{eq,s} = \rho_{s,b} = 3.01\%$).

CONCLUSION

Theoretical model of analysis hybrid reinforced FRP and steel concrete member is presented using principle of strain compatibility and forces equilibrium, the moment-curvature relation, deflection and ultimate moment strength are established. A good agreements have been shown between theoretical prediction results and experimental results of other researcher. The adopted stress-strain relationships for concrete in compressive and tensile zones in the theoretical model are suitable with acceptable degree of coinciding in load-deflection response.

Tension stiffness of concrete in H-RC section has significant effects in the first part of load-deflection curve and this effect is lessened in second stage of curve while it is diminished in the third stage and ultimate strength of H-RC members.

The effectiveness type of FRP reinforcements in H-RC member is augmented as it's modulus of elasticity increased, since, the FRP reinforcements plays a role in the third stage of load deflection curve of HRC member after yielding of steel is occurred.

The principle of converting the FRP reinforcement in HRC section into equivalent steel reinforcement can be adopted to apply provisions of ACI-318M-14 in designing approach with modifying formula, since, the recommended failure in HRC is ductile failure mode initiated by steel yield then concrete crushing while FRP rebar not yet ruptured. The equivalent steel reinforcement ratio is affected by elasticity modulus of FRP rebar.

The ductility index of H-RC beam represented by the ratio of total absorption energy of H-RC beam to total absorption energy of S-RC beam reinforced with equivalent steel ratio. While the ultimate strength index of H-RC beam identified as the ratio between the ultimate bending strength of H-RC member to that for FRP-RC beam reinforced by FRP reinforcement with the same equivalent steel reinforcement ratio. The variation of these indices are opposed as the equivalent reinforced ratio (A_f/A_s) is increased due to increase of FRP reinforcing area with respect to decrease of steel reinforcement.

Critical range of hybrid reinforcement (A_f/A_s) ratio are specified that meets the ductility index (ψ) and normalized ultimate strength index (X) at the same level.

REFERENCES

- ACI, 318-05, 2005. Building Code Requirements for Structural Concrete and Commentary. ACI Committee 318, American Concrete Institute, Farming Hills, MI, USA., pp: 430.
- ACI., 2006. 440.1R-06: Guide for the Design and Construction of Structural Concrete Reinforced with FRP Bars. American Concrete Institute, Michigan, USA., ISBN:9780870312106, Pages: 44.
- Aiello, M.A. and L. Ombres, 2000. Load-deflection analysis of FRP reinforced concrete flexural members. *J. Compos. Constr.*, 4: 164-171.
- Bischoff, P.H. and R. Paixao, 2004. Tension stiffening and cracking of concrete reinforced with Glass Fiber Reinforced Polymer (GFRP) bars. *Can. J. Civil Eng.*, 31: 579-588.
- Hawileh, R.A., 2015. Finite element modeling of reinforced concrete beams with a hybrid combination of steel and ARAMID reinforcement. *Mater. Des.*, 65: 831-839.
- Kara, I.F., A.F. Ashour and M.A. Koroglu, 2015. Flexural behavior of hybrid FRP-steel reinforced concrete beams. *Compos. Struct.*, 129: 111-121.
- Lau, D. and H.J. Pam, 2010. Experimental study of hybrid FRP reinforced concrete beams. *Eng. Struct.*, 32: 3857-3865.
- Leung, H.Y. and R.V. Balendran, 2003. Flexural behaviour of concrete beams internally reinforced with GFRP rods and steel rebars. *Structur. Surv.*, 21: 146-157.
- Nayal, R. and H.A. Rasheed, 2006. Tension stiffening model for concrete beams reinforced with steel and FRP bars. *J. Mater. Civil Eng.*, 18: 831-841.
- Pang, L., W. Qu, P. Zhu and J. Xu, 2015. Design propositions for hybrid FRP-steel reinforced concrete beams. *J. Compos. Constr.*, 20: 04015086-04015086.
- Qu, W., X. Zhang and H. Huang, 2009. Flexural behavior of concrete beams reinforced with hybrid (GFRP and steel) bars. *J. Compos. Constr.*, 13: 350-359.
- Rashid, M.A., M.A. Mansor and P. Paramasivam, 2005. Behaviour of aramid fiber reinforced polymer reinforced high strength concrete beams under bending. *J. Compos. Construct.*, 9: 117-127.
- Tan, K.H., 1997. Behaviour of hybrid FRP-steel reinforced concrete beams. Proceedings of the 3rd International Symposium on Non-Metallic (FRP) Reinforcement for Concrete Structures (FRPRCS'97) Vol. 2, October 14-16, 1997, Japan Concrete Institute (JCI), Tokyo, Hokkaido, Japan, ISBN:4931451004, pp: 487-494.

An Expansion in Classification of Eye Gaze Detection Methods

¹Jumana Waleed, ²Ahmed Brisam and ¹Taha Mohammed Hasan

¹Department of Computer Science, College of Science, University of Diyala, Baqubah, Iraq

²College of Agriculture, University of Al-Qadisiyah, Diwaniyah, Iraq

Abstract: Eye gaze, often regarded as a language of non-verbal communication can be studied and estimated using scientific detection methods and contains a tremendous scope of applications in daily lives. Investigations in eye gaze detection can give ideas about what a person is thinking about based on how and where he is looking. Nowadays, the eye gaze detection methods utilize an incorporation of convenient hardware and affined software to detect the gaze direction for providing new applications ranges. In this study, the objective is to provide an inclusive expansion in classifying the recently existing methods based on the eye gaze detection by considering various paradigms such as wearable and unwearable devices, active and passive systems, feature and appearance approaches, the estimation of head movements and the used algorithm. This study includes the major preliminaries, usefulness and abuses as well as the necessities of future expansion in the domain.

Key words: Communication, inclusive, appearance, necessities, passive, preliminaries, scientific

INTRODUCTION

The eye gaze movements supply a fundamental interfacing with a person's purposes and thinking. A noticeable view of the psychology field is the capability of investigating the internal working of the brains via. computing how different muscles of the eye are contracted (Fig. 1).

The researches of eye gaze detection methods may expose what is the thinking of person about on the basis of where he/she is looking. Thus, the eye gaze detection allows various applications, like working as an alternate mechanism of input to make the users capable of doing

different jobs such as utilizing a smartphone to determine the visual attention of the users providing a modern field of human interaction environment for a perfect life to the disabled and old people reducing the road accidents by alarming the drivers under various driving conditions; controlling the video games representing a helpful means for medical diagnostics like mental disease diagnostic as well as improving medical education and therapy controlling the humanoid robots removing the awkwardness related to industrialization of sports allowing an extra reliability with a strong performance for security and authentication purposes.

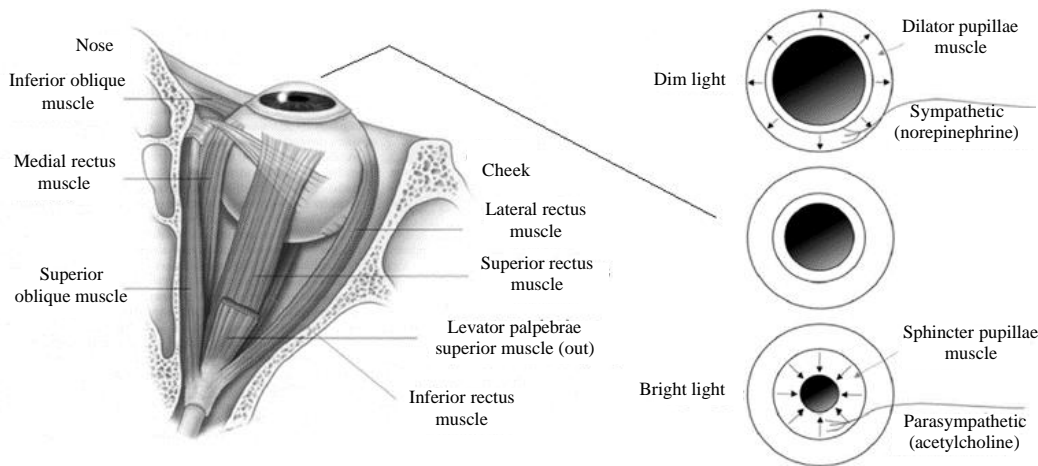


Fig. 1: Demonstrates the eye-muscles which are supply the eye gaze movement

MATERIALS AND METHODS

Classification of eye gaze detection methods: The methods adopted for the detection of eye gaze can be classified into several types based on the perimeter chosen for classification. For instance, the classification differs when the physical proximity between the subject and the system is studied than when the algorithm used for gaze detection is studied. Further, new perimeters of classification of eye gaze detection systems are being introduced as new developments are being done rapidly in this field. In this study, different methods of eye gaze detection are classified under some broad criteria of differentiation.

Methods based on wearable/non-wearable devices: One of the most popular classification of eye gaze detection methods is based on the wearability and non-wearability concepts. Based on these concepts, the methods of eye gaze detection are classified in to attached and remote methods (Al-Rahayfeh and Faezipour, 2013). The attached (wearable) gaze detection methods generally consist of a single/multiple lenses attached to a device, generally glasses in order to detect the gaze coordinates of the eyes. In this methods, the person requires to lay glasses frames on the helmets or cameras. Helmets provide good solutions for detecting the eye gaze, none the less, there is a limitation in expanding the use of such methods because of their heavy weight. As an example for using helmets in detecting of eye gaze (Pires *et al.*, 2013). present a method can be successfully implemented in sport applications with a fast computing and handling of the eye location changes. While the spectacles or glasses can achieve a high publicity, since, they are not heavy and accurate as in the work in which McMurrrough *et al.* (2013) use a wearable eye gaze glasses (Fig. 2) for monitoring the right eye equipped with an infrared-monocular camera in addition to the used of reflective markers to detect the head movements and this research provides a high degree of accuracy.

Another arrangement of wearable eye gaze glasses with a low-cost webcam was introduced by Jen *et al.* (2016) for detecting the region of interest and this work also provides high degrees of accuracy and robustness. Depending on the recently existing eye gaze detection methods, the wearable based classification methods are implemented in various applications. Kim *et al.* (2012) present a monitoring system for both eyes equipped with three mini cameras and two mirrors for extracting the regions of interest depending on the detection of binocular gaze. Yip *et al.* (2016) present a controlling interface for surgical manipulators based on wearable



Fig. 2: Represents the wearable eye gaze device (McMurrrough *et al.*, 2013)

eye gaze detection method. And with application of virtual reality, Tripathi and Guenter (2017) propose a method based a wearable device to detect an eye gaze with a relatively good degree of accuracy. The remote (non-wearable) methods of eye gaze detection do not require any attachment to the subject's body. Rather, the hardware in this system can be assembled in a close proximity to the subject in order to feed the necessary data to the software involved. This type of methods can be particularly beneficial to operate machineries and equipment's remotely, allowing the controller to be at a safe distance while performing tasks that are risky. Also such method can be applicable for people with motor disorders to control appliances that input eye signals to perform tasks. An application of remote based methods is proposed by Ebisawa and Fukumoto (2013). Another application is the concept of a tele operative mobile robot has been proposed by Gego *et al.* (2017).

There are different fields of both attached and remote eye gaze detection methods because of both types of detection methods have unrivaled usable benefits. But comparing with other methods based on wearable devices, the methods based on non-wearable devices are widely applied, gratitude for their convenience and easy utilization.

Methods based on active/passive systems: The eye gaze detection methods can be categorized into active and passive systems. Active detectors use infrared illuminants, a light emitting diode and cameras for adding lighting to obtain a suitable recognition for the eye. they are a less effectiveness at more distances and in the daybreak. The passive systems are operating in visual light. These systems are so, natural yet, they are very sensitive to lighting condition. The most present eye gaze

detection methods necessitate somewhat a high cost and employment of infrared illuminators and a light emitting diode techniques that affect the user eyes (Palinko *et al.*, 2016). For this reason, it is quite necessary to design a low cost eye gaze detection method based on a software.

Recently, a lot of researches have been worked with the active and passive eye gaze detection methods. Hughes *et al.* (2017) proposed an accurate active detection method under the perfect script of lighting environment (32 light emitting diodes) and a fixed head movement but when moving the head, the detection of eye gaze will be very difficult. Findings by Smith *et al.* (2013) show that, while the active systems of gaze detection were generally more accurate, the passive system is better in terms of distance sensitivity and pose variation. Further, the efficiency of the passive system can be increased by feeding the system with more memory (the number and variations of pictures of different faces in different poses and gazes) that are used as reference to detect the gaze of the subject. As a passive based method in the application of driver assistance, Maralappanavar *et al.* (2016) decide if the drivers are distracted or not via estimating their gaze directions.

One complication for both the active and passive systems of eye gaze can arise when the subject is either far-sighted or near-sighted and wears glasses. It is because glasses often reflect environmental lights which makes detecting eye gaze more difficult. This problem can even combine with a difficulty to detect the pupil when using an active gaze detection method (Chinsatitf and Saitoh, 2016). However, one major advantage of active eye gaze system over the passive gaze system is former's usefulness to control robotics that require the input of a controller's facial and body gestures in order to perform.

Methods based on appearance/feature approaches: The eye gaze detection methods are divided into two categories appearance based methods and feature based methods. The appearance based eye gaze detection methods are the ones that treat the cropped eye images to be detected as points in a high dimensional space instead of treating them as a combination of certain geometrical figures like contours and corners (Zang and Liu, 2012). The core usefulness of appearance based methods is that they require low hardware and these methods can be implemented without using high resolution cameras or extra sources of light. While the core disadvantage of this classification is that it has low accuracy between 2-3° with a fixed head.

Feature based methods occupy into consideration different merits of the human eyes to detect a distinguishing set of features such as pupil contour, the corners of the eye and corneal reflections of near infrared light sources (Wang *et al.*, 2016). These methods can be partition into regression based and model based methods. In the regression based methods, the vector between the pupil center and corneal reflections is detected and geometrically mapped via a polynomial regression-function to gaze coordinates on the screen. While the model based methods utilize a geometrical pattern of the eye of human straight with near infrared illuminator and various cameras and mirrors for developing an accurate detecting of eye gaze directions. Comparing with the appearance based methods, this classification has a higher accuracy but it requires extra hardware like various sources of light or cameras. Also, there is a difficulty may cause method fail when the near infrared glints in the eye cornea are disappearing owing to the head movement or when more glints are appearing in the eye via other near infrared sources of light (Bazrafkan *et al.*, 2015).

There are many researchers are tried to fix the problems appeared in appearance and feature based methods. Wang *et al.* (2016, 2018) improve that appearance based methods can be very effective when training big amounts of data. Zhang *et al.* (2017) propose a full face appearance based trained architecture method to detect accurate and robust eye gaze directions under various challenges of light conditions and head poses. Without any extra hardware (only a webcam on the screen of the computer), Jariwala *et al.* (2016) presented a geometric feature based method to detect the eye gaze directions with a high degree of accuracy.

The methods based on head pose estimation: The estimation of the eye gaze of a human can also be made by analyzing the head pose instead of detecting for minute features like pupil or iris. Therefore, the head pose estimation system is more effective when the subject of study is not in a close proximity of the detection system to effectively detect the configuration of optical and facial muscles. The head pose method of eye gaze detection can be applied to study multiple subjects at the same time, thus, making it suitable for applications like studying the behavioral patterns of crowds for crowd management, searching for a particular person in a crowd and so on (Baxter *et al.*, 2015). The head pose estimation systems are further studied under two different domains namely visual surveillance and human computer interactions (Mukherjee and Robertson, 2015). Both domains of head pose detection differ for one another in aspects of the quality of input and range of detection. The visual

surveillance system is the type of estimation system than can work from a fairly long range between the system taking the input and the subject. It is mostly applied to detect patterns and superficial details such as the speed and direction of a subject in motion, common gestures and patterns of crowd movement etc. (Chen and Odobez, 2012). Since, this domain only concerns with the superficial subjects and can work from a long range, it does not demand very high quality of images as input data. In contrast, Human Computer Interactions (HCI) is a detail-based domain of detection. This type of detection system applies the approach where facial landmark detection is performed (Fanelli *et al.*, 2011). Therefore, this approach requires the input of high resolution images and can only work in a short range, i.e., generally, 1-2 m between the hardware and the subject. Further, unlike the visual surveillance methods, the human computer interaction methods are suitable for indoor applications where the subject in interest are static and facing directly towards the input mechanism. However, it is possible to combine the usefulness of both the domains and to avoid their limitations by combining the concepts of both the domains together in an applicable system as proposed by Mukherjee and Robertson (2015). Their approach involved the combination of two different networks, one trained to capture depth and detail and the other to capture RGB in order to unify the problems of both the domains. Their results further show that the combination of these two domains into one system opens up scope for faster and more reliable detection systems that are highly applicable. Valenti *et al.* (2012) provide a preferable visual eye gaze estimation via. joining the eye location with the information of the head movement. For smart interacting between the user and the smart TV, Nguyen *et al.* (2013) present an eye gaze detection method based on head pose estimation to control the smart TV. Another application of controlling various assistive technologies is proposed by Al-Rahayfeh and Faezipour (2014) in

which several controlling signals are promoted via. the classification of eye gaze direction and the detection of head movement. In the application of the drowsiness in the driving assistance systems, Choi and Kim (2014) propose a relatively fast calculation method for detecting the head pose and the eye gaze to reveal the drowsy driver.

Methods based on used algorithm: The used algorithm based eye gaze detection methods, rather than detecting for certain facial reference points or analyzing images for certain facial features and poses, uses a set of calculable rules in order to determine the eye gaze of an individual. Already, this classification of eye gaze detection methods can be utilized in the field of driver fatigue monitoring, since, the past few years. Lopar and Ribariæ (2013) suggest two different bases of eye gaze detection, one for eye detection and the other for face detection further, both the types of detection methods contained further divisions of detection types in practice. George and Aurobinda (2016) have provided fast and reliable detection algorithm for eye localization using low resolution images which searches for the center of the iris using the geometrical features. Besides, several other algorithmic techniques for eye gaze detection are being introduced in researches these days, expanding the field and introducing new approaches all the time.

RESULTS AND DISCUSSION

The performance analysis: The previously presented classifications are briefly given in Table 1. This table introduces a comparison that is based on the following criteria: the application domain, wearable and unwearable devices, active and passive systems, feature and appearance based systems and the head movements detection.

Table 1: A comparison between different classifications of eye gaze detection methods

Researchers (years)	Field of application	Attached methods	Remote methods	Active system	Passive system	Feature based system	Appearance based system	Head poses
Pires (2013)	Motor sports application	✓		✓		✓		
McMurrough <i>et al.</i> (2013)	Educational and research purposes	✓		✓		✓		✓
Jen (2016)	Driver assistance systems and assistive techniques	✓			✓		✓	
Kim <i>et al.</i> (2012)	Psychology and robot utilization	✓		✓			✓	
Yip <i>et al.</i> (2016)	Surgical manipulator	✓			✓		✓	
Tripathi and Guenter (2017)	Virtual reality	✓		✓		✓		✓
Ebisawa and Fukumoto (2013)	-		✓	✓		✓		✓
Gego <i>et al.</i> (2017)	Human robot interaction		✓		✓		✓	
Hughes <i>et al.</i> (2017)	Input device for computer interaction		✓	✓		✓		
Smith <i>et al.</i> (2013)	Human object interaction							

Table 1: Continue

Researchers years	Field of application	Attached methods	Remote methods	Active system	Passive system	Feature based system	Appearance based system	Head poses
Shweta <i>et al.</i> (2016)	Driver assistance system		✓		✓		✓	
Wang <i>et al.</i> (2016)	Human attention understanding		✓		✓		✓	
Wang <i>et al.</i> (2018)	-		✓		✓		✓	✓
Zhang <i>et al.</i> (2017)	Computer vision tasks		✓		✓		✓	✓
Jariwala <i>et al.</i> (2016)	Commercial gaze estimation systems with simple webcams		✓		✓		✓	✓
Mukherjee and Robertson (2015)	Scene interaction detection		✓		✓		✓	✓
Valenti <i>et al.</i> (2012)	-		✓		✓		✓	✓
Nguyen <i>et al.</i> (2013)	Smart TV controlling		✓		✓		✓	✓
Al-Rahayfeh and Faezipour (2014)	Various assistive techniques		✓		✓		✓	✓
Choi and Kim (2014)	Driver assistance systems		✓		✓		✓	✓
George and Routray (2016)	-		✓		✓		✓	

CONCLUSION

To conclude, the classifications and applications of eye gaze detection methods are diverse and expanding. Each classification of the methods has its own usefulness and abuses. Unfortunately, these methods based classification cannot provide entirely exhaustive objectives. In the previous sections, a representative insight for classifying the eye gaze detection methods in various applications is presented. The need to find a beneficial eye gaze detection method includes high degrees of reliability and accuracy, robust head movements detection and robustness to spacious varied lighting conditions. In addition to the necessity for increasing the performance of the method and decreasing the cost.

REFERENCES

Al-Rahayfeh, A. and M. Faezipour, 2013. Eye tracking and head movement detection: A state-of-art survey. *IEEE. J. Trans. Eng. Health Med.*, 1: 2100212-2100223.

Al-Rahayfeh, A. and M. Faezipour, 2014. Application of head flexion detection for enhancing eye gaze direction classification. *Proceedings of the 2014 IEEE 36th Annual International Conference Engineering in Medicine and Biology Society (EMBC'14)*, August 26-30, 2014, IEEE, Chicago, Illinois, USA., ISBN:978-1-4244-7929-0, pp: 966-969.

Baxter, R.H., M.J. Leach, S.S. Mukherjee and N.M. Robertson, 2015. An adaptive motion model for person tracking with instantaneous head-pose features. *IEEE. Sig. Process. Lett.*, 22: 578-582.

Bazrafkan, S., A. Kar and C. Costache, 2015. Eye gaze for consumer electronics: Controlling and commanding intelligent systems. *IEEE. Consum. Electron. Mag.*, 4: 65-71.

Chen, C. and J.M. Odobez, 2012. We are not contortionists: Coupled adaptive learning for head and body orientation estimation in surveillance video. *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'12)*, June 16-21, 2012, IEEE, Providence, Rhode Island, ISBN:978-1-4673-1226-4, pp: 1544-1551.

Chinsatitf, W. and T. Saitoh, 2016. Improvement of eye detection performance for inside-out camera. *Proceedings of the 2016 IEEE/ACIS 15th International Conference on Computer and Information Science ICIS'16*, June 26-29, 2016, IEEE, Okayama, Japan, ISBN:978-1-5090-0807-0, pp: 1-6.

Choi, I.H. and Y.G. Kim, 2014. Head pose and gaze direction tracking for detecting a drowsy driver. *Proceedings of the 2014 International Conference on Big Data and Smart Computing (BIGCOMP'14)*, January 15-17, 2014, IEEE, Bangkok, Thailand, ISBN:978-1-4799-3920-6, pp: 241-244.

Ebisawa, Y. and K. Fukumoto, 2013. Head-free, remote eye-gaze detection system based on pupil-corneal reflection method with easy calibration using two stereo-calibrated video cameras. *IEEE. Trans. Biomed. Eng.*, 60: 2952-2960.

Fanelli, G., J. Gall and L.V. Gool, 2011. Real time head pose estimation with random regression forests. *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'11)*, June 20-25, 2011, IEEE, Providence, Rhode Island, ISBN:978-1-4577-0394-2, pp: 617-624.

Gego, D., C. Carreto and L. Figueiredo, 2017. Teleoperation of a mobile robot based on eye-gaze tracking. *Proceedings of the 12th Iberian Conference on Information Systems and Technologies (CISTI'17)*, June 21-24, 2017, IEEE, Lisbon, Portugal, ISBN:978-1-5090-5047-5, pp: 1-6.

George, A. and A. Routray, 2016. Fast and accurate algorithm for eye localisation for gaze tracking in low-resolution images. *IET. Comput. Vis.*, 10: 660-669.

- Hughes, J., S. Rhodes and B.E. Dunne, 2017. Eye gaze detection system for impaired user GUI control. Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS'17), August 6-9, 2017, IEEE, Boston, Massachusetts, USA., ISBN:978-1-5090-6390-1, pp: 1348-1351.
- Jariwala, K., U. Dalal and A. Vincent, 2016. A robust eye gaze estimation using geometric eye features. Proceedings of the 3rd International Conference on Digital Information Processing, Data Mining and Wireless Communications (DIPDMWC'16), July 6-8, 2016, IEEE, Moscow, Russia, ISBN:978-1-4673-9380-5, pp: 142-147.
- Jen, C.L., Y.L. Chen, Y.J. Lin, C.H. Lee and A. Tsai *et al.*, 2016. Vision based wearable eye-gaze tracking system. Proceedings of the 2016 IEEE International Conference on Consumer Electronics (ICCE'16), January 7-11, 2016, IEEE, Las Vegas, Nevada, ISBN:978-1-4673-8363-9, pp: 202-203.
- Kim, M.Y., S. Yang and D. Kim, 2012. Head-mounted binocular gaze detection for selective visual recognition systems. *Sensors Actuators A: Phys.*, 187: 29-36.
- Lopar, M. and S. Ribaric, 2013. An overview and evaluation of various face and eyes detection algorithms for driver fatigue monitoring systems. Proceedings of the 2nd Workshop on Croatian Computer Vision (CCVW'13), September 19, 2013, University of Zagreb, Zagreb, Croatia, pp: 15-18.
- Maralappanavar, S., R. Behera and U. Mudenagudi, 2016. Driver's distraction detection based on gaze estimation. Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI'16), September 21-24, 2016, IEEE, Jaipur, India, ISBN:978-1-5090-2030-0, pp: 2489-2494.
- McMurrough, C.D., V. Metsis, D. Kosmopoulos, I. Maglogiannis and F. Makedon, 2013. A dataset for point of gaze detection using head poses and eye images. *J. Multimodal User Interfaces*, 7: 207-215.
- Mukherjee, S.S. and N.M. Robertson, 2015. Deep head pose: Gaze-direction estimation in multimodal video. *IEEE. Trans. Multimedia*, 17: 2094-2107.
- Nguyen, D.T., K.Y. Shin, W.O. Lee, Y.G. Kim and K.W. Kim *et al.*, 2013. Gaze detection based on head pose estimation in smart TV. Proceedings of the 2013 International Conference on ICT Convergence (ICTC'13), October 14-16, 2013, IEEE, Jeju, South Korea, ISBN:978-1-4799-0696-3, pp: 283-288.
- Palinko, O., F. Rea, G. Sandini and A. Sciutti, 2016. Robot reading human gaze: Why eye tracking is better than head tracking for human-robot collaboration. Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'16), October 9-14 2016, IEEE, Daejeon, South Korea, ISBN:978-1-5090-3763-6, pp: 5048-5054.
- Pires, B., M. Hwangbo, M. Devyver and T. Kanade, 2013. Visible-spectrum gaze tracking for sports. Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, IEEE, Portland, Oregon, ISBN:978-0-7695-4990-3, pp: 1005-1010.
- Smith, B.A., Q. Yin, S.K. Feiner and S.K. Nayar, 2013. Gaze locking: Passive eye contact detection for human-object interaction. Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, October 08-11, 2013, ACM, St. Andrews, Scotland, UK., ISBN:978-1-4503-2268-3, pp: 271-280.
- Tripathi, S. and B. Guenter, 2017. A statistical approach to continuous self-calibrating eye gaze tracking for head-mounted virtual reality systems. Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV'17), March 24-31, 2017, IEEE, Santa Rosa, California, ISBN:978-1-5090-4823-6, pp: 862-870.
- Valenti, R., N. Sebe and T. Gevers, 2012. Combining head pose and eye location information for gaze estimation. *IEEE. Trans. Image Process.*, 21: 802-815.
- Wang, Y., T. Shen, G. Yuan, J. Bian and X. Fu, 2016. Appearance-based gaze estimation using deep features and random forest regression. *Knowl. Based Syst.*, 110: 293-301.
- Wang, Y., T. Zhao, X. Ding, J. Peng and J. Bian *et al.*, 2018. Learning a gaze estimator with neighbor selection from large-scale synthetic eye images. *Knowl. Based Syst.*, 139: 41-49.
- Yip, H.M., D. Navarro-Alarcon and Y.H. Liu, 2016. Development of an eye-gaze controlled interface for surgical manipulators using eye-tracking glasses. Proceedings of the 2016 IEEE International Conference on Robotics and Biomimetics (ROBIO'16), December 3-7, 2016, IEEE, Qingdao, China, ISBN:978-1-5090-4365-1, pp: 1900-1905.
- Zang, Y. and H. Liu, 2012. A real-time video-based eye tracking approach for driver attention study. *Comput. Inf.*, 31: 805-825.
- Zhang, X., Y. Sugano, M. Fritz and A. Bulling, 2017. It's written all over your face: Full-face appearance-based gaze estimation. Proceedings of the 2017 IEEE International Workshops on Computer Vision and Pattern Recognition (CVPRW'17), July 21-26, 2017, IEEE, Honolulu, Hawaii, ISBN:978-1-5386-0734-3, pp: 2299-2308.

Multi-Criteria Analysis and Advanced Comparative Study of Self-Organization Protocols in Wireless Sensor Network

Adil Hilmani, Abderrahim Maizate and Larbi Hassouni
RITM-ESTC/CED-ENSEM, University Hassan II km 7,
El Jadida Street, B.P. 8012 Oasis, Casablanca, Morocco

Abstract: Nowadays, Wireless Sensor Networks (WSN) play a crucial role in various fields such as the army, health, the environment and so on. Emerging technologies such as the internet of things, smart applications and smart grids stimulate the deployment of autonomous, self-configuring, large-scale Wireless Sensor Networks (WSNs). These sensors which are small nodes, collect data in a wellfield and transfer it to a new sink where all data waiting to be exploited by applications (web, mobiles, ...) are stored. These new sensors have limitations which must be taken into consideration when deploying a WSN network, especially their strictly limited energy. Thus, in order to efficiently use the energy of these nodes to improve and increase the lifetime of the entire network and reduce the energy consumed during data transmission and processing, several self-organization algorithms have been proposed to create different network architectures based on required applications. In this research, we will study and examine the various WSN-based self-organization algorithms, highlighting their principles in order to compare them according to different metrics and on multiple requirements such as load balancing, energy efficiency, complexity of the algorithm, ..., etc. To perform this multi-criteria comparison, we will use the ROC (Rank Order Centroid) multicriteria analysis method to make the decision when designing and implementing a new efficient and effective self-organization algorithm that will meet all metrics and to all criteria in order to create a performant and sustainable WSN network.

Key words: Wireless sensor networks, network architecture, self-organization, energy efficiency, load balancing, multi-criteria analysis method (ROC)

INTRODUCTION

Micro Electromechanical Systems technology (MEMS) and wireless communication have recently seen tremendous progress and growth which has allowed the development of small-scale wireless sensors with limited energy resources, low storage capacity and low power communication (Kalantary and Taghipour, 2014; Lan and Mehmet, 2010; Akyildiz *et al.*, 2002). These wireless sensor nodes are deployed in an area of interest by forming WSN wireless sensor arrays which detect and collect the information to be transferred to the Base Station (BS or sink) as illustrated in Fig. 1. This information will then be sent to the user via external networks (ethernet, WiFi, 3G/4G, satellite, ...). The wireless sensor nodes are composed of different units as illustrated in Fig. 2.

The detection or sensing unit which consists of two modules: a sensor which makes it possible to detect the information in the field according to the phenomenon to be monitored and the characteristic of the sensor

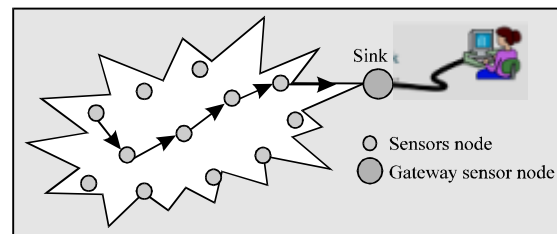


Fig. 1: General architecture of a WSN

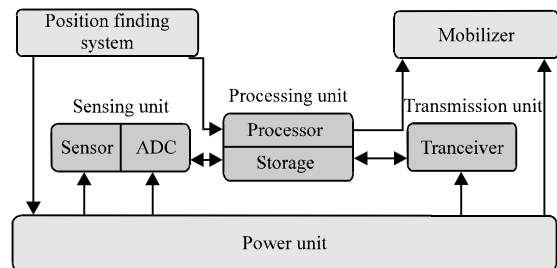


Fig. 2: Internal architecture of a sensor

(Ex: the presence of an object, temperature, humidity, pressure, ...) and an Analog-Digital Converter (ADC) which communicates with the processing unit to transfer the detected information:

- C The processing unit for saving and processing the data collected from the field
- C The communication unit which transmits and receives data from the outside
- C The energy unit which is responsible for feeding the node

Sensor nodes are individual nodes that know nothing about the network and do not have an existing fixed infrastructure, they are often totally decentralized. Then, these nodes have to organize themselves, unlike traditional wired networks in an autonomous way to form a network topology, so that, they can communicate and transfer the detected data to the sink (Sirsikar *et al.*, 2014). That is to say, before starting to transfer the detected data, all the sensor nodes that are in the capture zone will self-organize themselves in an autonomous and efficient way to build a solid topology which considers their limits in energy, storage capacity, processing and transmission power (a short radio range) and ensures the reception of data by the sink.

For this reason, several algorithms have been developed in recent times which allow the sensor nodes to self-organize and adopt a topology that will be the basis on which the data will be transferred to the base station and also an optimal structure to increase the life of these nodes and at the same time of the network.

These topologies and network architectures are divided into three types of categories (Al-Karaki and Kamal, 2004) flat architecture, location architecture and hierarchical architecture (Fig. 2). In a flat architecture and location, all nodes have the same role in the network and they have the same resources whereas in a hierarchical architecture, the network is divided into several levels of responsibility and the nodes do therefore not play the same role. The main techniques used in the hierarchical architecture are the construction of clusters or clustering, chains and trees (tree structure).

The algorithms based on the hierarchical architecture are more efficient than those based on the flat architecture on the location architecture in terms of energy consumption since they make, it possible to minimize the total aggregate transmit power on the nodes in the selected path and to balance the load between the nodes to extend the service life of the network (Walters *et al.*, 2007).

The main objective of this study is to examine the different hierarchical algorithms used to form and construct network architectures and topologies in the form of clusters or chains or trees for WSN based on different criteria and metrics while considering the constraints of the sensor nodes to provide a long service life for the network and to balance the energy consumption and the load between these nodes.

MATERIALS AND METHODS

Chain-based algorithms: The chain-based algorithms are based on the construction of a string from all the nodes scattered in the zone of interest in order to transmit all the detected information to the base station. Among these algorithms the most common are PEGASIS (Lindsey and Raghavendra, 2002; Tan and Korpeoglu, 2003) is a protocol based on the construction of the chain. In each round, a leader is elected to lead the construction of this chain (Fig. 3). The construction of the chain begins with the node furthest from the base station.

The global knowledge of the network by the nodes begins with the node furthest from the base station and it is done with the choice of the nearest neighbors by transferring the distance and information of the neighboring node until covering all the nodes. This procedure is also carried out in the opposite direction so that all the nodes will have a global vision on the whole network (Fig. 4).

The data transmission begins with the leader which sends a token to both ends to send it their detected data (Fig. 5). In each round, a new leader is chosen to rebuild a chain and collect the detected data.

EECB (Yu and Song, 2010) is a protocol based on the construction of a chain of nodes which uses the same algorithm as PEGASIS except that they avoid the formation of the long links between the nodes (Fig. 6) and it chooses the leader on the basis of a residual energy formula and the distance between the base station and each sensor node in the network (Eq. 1):

$$d_i = \sqrt{(x_i - x_{BS})^2 + (y_i - y_{BS})^2} \tag{1}$$

$$Q_i = \frac{E_{\text{residual-}i}}{d_i}$$

PDCH (Linping *et al.*, 2010), PEGASIS double cluster head is a new protocol that uses the same notions as PEGASIS. It proceeds in two phases. In the first phase, it creates several levels in the form of concentric circles whose center is the base station. In the second phase, it elects at each level two leader (or two cluster heads) to build the chain at the level. The red circles show the

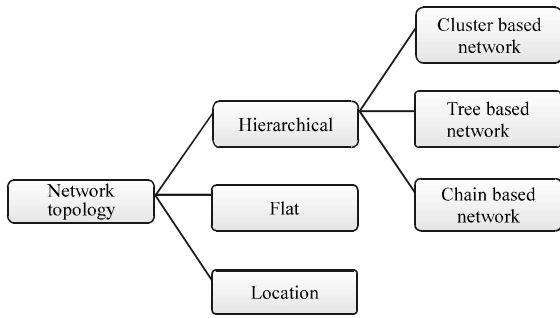


Fig. 3: Classification of network topologies

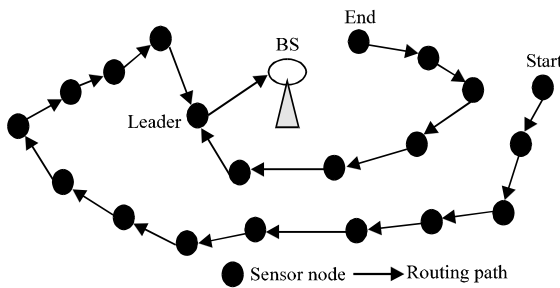


Fig. 4: Illustration of the PEGASIS protocol

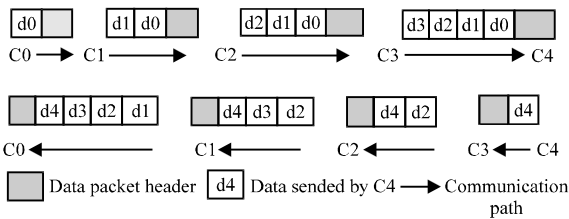


Fig. 5: Global network knowledge procedure

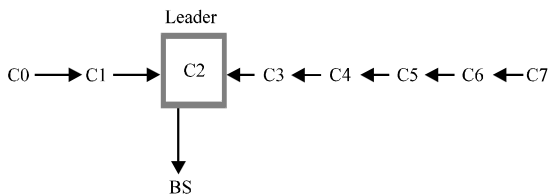


Fig. 6: Transmission of data to the leading node

nodes whose number of branches is >2 . These notes are therefore more likely to be selected to be a group leader than the other notes of each level.

The selection of cluster heads is based on the same Eq. 1 to choose two cluster heads (main cluster head and secondary cluster head) in each level among the possible cluster heads available (red circles). One of the two CHs collects data of the chain in its level and sends it to the other CH which merges the chain of its level with the one it receives from the level to which it is

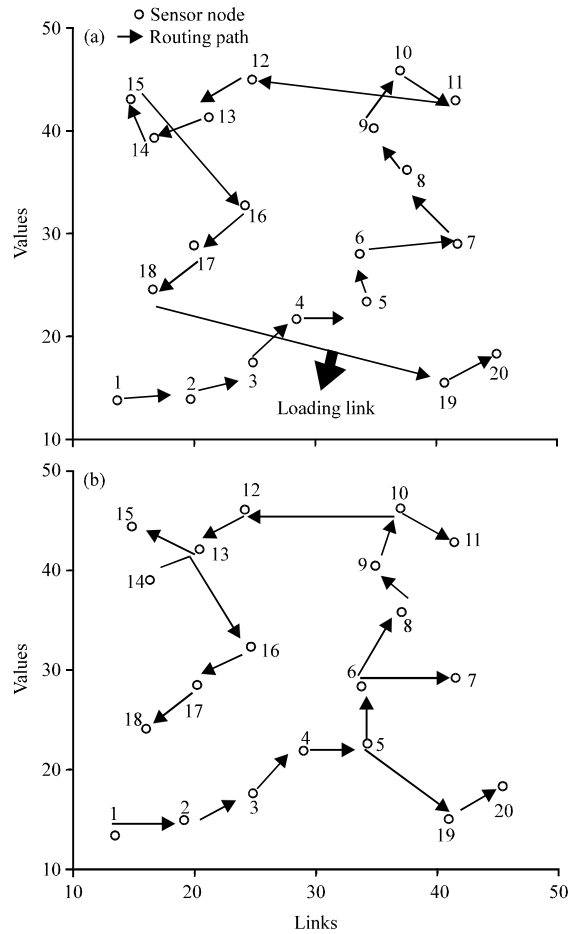


Fig. 7: a) PEGASIS with Long Link (LL) and b) EECB without Long Link (LI)

adjacent if it exists. The latter then sends the result chain to the other adjacent level but closer to the base station. This process is initiated at the furthest level and is repeated until the global chain arrives at the base station (Fig. 7).

CCM is a new algorithm applied in an organized and well distributed network where each node has uniform coordinates (i, j) (i = the number of the string and j is the number of the node in this chain) (Tang *et al.*, 2012).

This protocol is based on the construction of the chains whose leaders form a cluster. In each round, chain leaders are chosen and among them, a cluster head is elected to transmit to the base station his chain and all other chains received from other leaders.

ECCP is a hybrid hierarchical protocol based on the formation of chains in clusters. The selection of CH is based on the calculation of the weight of each node in a radius r and the node with the highest weight is elected a CH. Each CH forms its own cluster and the farthest node begins constructing the chain in the latter using the same processes as PEGASIS (Sheikhpour *et al.*, 2012).

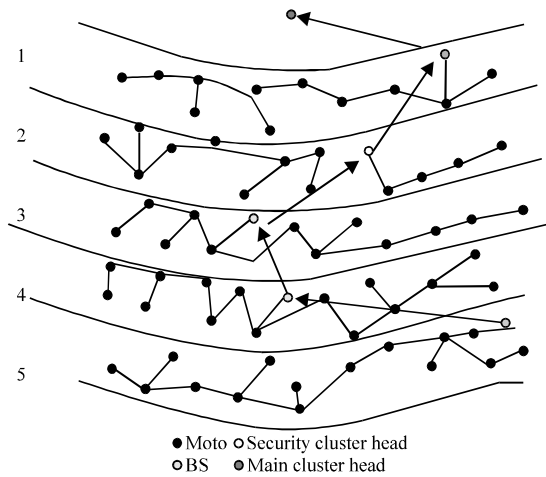


Fig. 8: Process of chain construction in PDCH

The construction of clusters is composed of two stages, the first is the selection of CHs and the second is the formation of clusters. In the step of selecting the CHs, each node broadcasts a message in a radio that contains its location and its residual energy. Each node that receives this message, it realizes the update of its table of neighborhood which contains the distance to its neighbors and it calculates its weight. After calculating the weight, each node sends its weight broadcast to all the nodes that are found in its radio r . The node with the bigger weight is elected a CH in his radio:

$$\text{Weight}_i = RE_i * \sum_{j=1}^{\text{number of neighbours}} \frac{1}{\text{dist}^2(v_i, v_j)} \quad (2)$$

Where:

- RE_i = The residual energy of the node i
- $\text{dist}(v_i, v_j)$ = The distance between the node i and the node j (for all the neighbors of the node i in his radio r)

SCBC is a protocol that is based on the creation of clusters in the surveillance zone and in each cluster a chain is constructed that contains a leader which is a CH or SCH (Secondary CH) (Tan and Viet, 2015). The division of the cluster surveillance zone is carried out by applying an algorithm which calculates the angular value of each cluster and also the angular angles of each node and after it decides the nodes that belong to each cluster. The number of nodes of each cluster is fair whose clusters have the same number of sensor nodes (Fig. 8).

Once the clusters are formed, the CHs selection step in each cluster begins. This step is based on the residual energy of each node in a cluster and a value of the cost that has a relation to the distance of the base station. If a node has a residual energy higher than E_{avg} and has a maximum cost then this node is elected as a CH. The

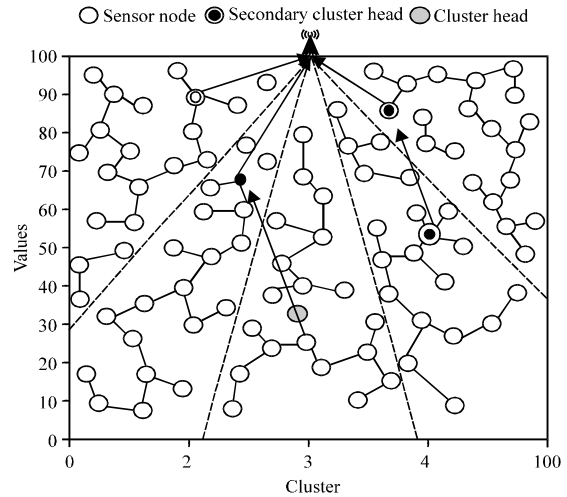


Fig. 9: Network topology with SCBC

selection of the secondary CH in each cluster is carried out in each round and is based on the average distance between the CHs selected in the previous step and the base station with the base station ($R/2$) range and Also with the number of SCHs which must not exceed $k/2$ (k is the cluster number). With these conditions, the CHs will be far from the base station and the SCHs will be closer to the base station and will work as much routers for these remote CHs (in a cluster there must not be a CH and SCH in the same round).

Pegasis multi-chain is a protocol based on the construction of chains in 4 regions of the surveillance zone (Patel and Munjani, 2016). In a area 100×100 with 100 nodes, each zone will contain 25 nodes in a dimension 50×50 and a base station. In each region, a node farthest from the BS of its region begins to create a chain whose leader is chosen in each round in a random manner (Fig. 9).

Tree-based algorithms: In this part, we will study some of the most used algorithms which are based on the self-organization of all the nodes that exists in the capture area to form a more suitable tree topology to transfer the collected data to the base station.

PEDAP is a hierarchical protocol based on tree construction using the prim spanning tree minimal algorithm which performs the branch cost calculation to know the optimal parent node of each node (Tan and Korpeoglu, 2003) (Fig. 10). The calculation is carried out by the base station.

The base station sends to each node in the network information such as: the parent node, the child nodes and the time slot number to send the data to the root node.

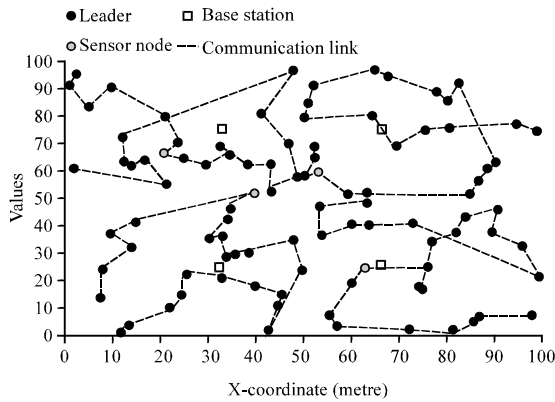


Fig. 10: PEGASIS multi-chain architecture

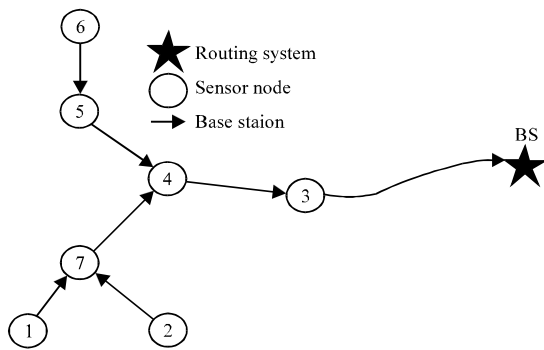


Fig. 11: The tree of a minimum spanning tree

TREEPSI is a protocol that works with rounds (Satapathy and Sarma, 2006). The construction of the tree begins with a node i which is the root and from this root the procedure of building the tree begins (Fig. 11). All the nodes in the monitoring zone have uniform initial energy, they have the location information and also have the ability to adapt the transmission power according to the distance at which the destination is located. The selection of node i as a root node is random (such as PEGASIS in turn) but this node must have a residual energy sufficient to transmit the data to the base station (the nodes know the location of the base station).

TBC is a protocol based on the construction of trees in clusters (Kim *et al.*, 2010). In each round, nodes are elected as CHs (root nodes) and trees will be built in each cluster from these nodes.

Once the CHs form the clusters using the same procedure as LEACH, these CHs have become roots in each cluster. Each root detects the member node farthest from its cluster and its distance is named d_{max} . Once the value d_{max} is determined, then divide the cluster to levels

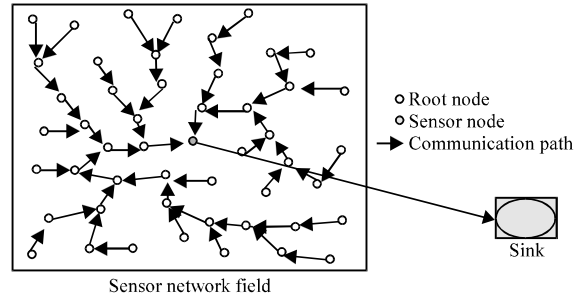


Fig. 12: Construction of the tree in TREEPSI

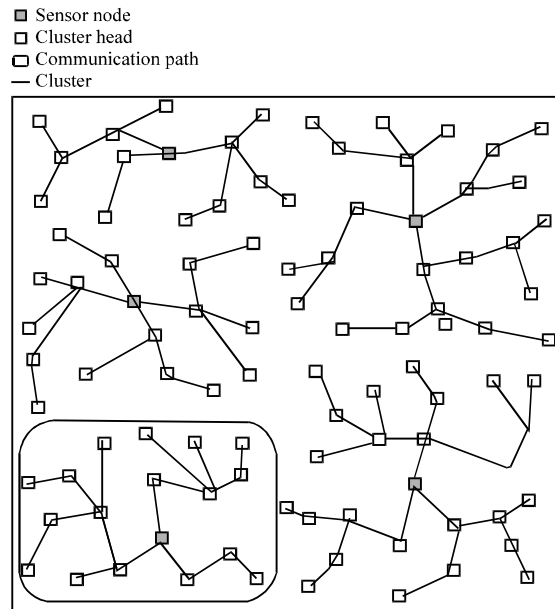


Fig. 13: The network topology with TBC

according to the alpha value, the number of levels in a cluster is d_{max}/α . These levels are circles around the root node. Each node in a level i detects the nearest neighbors and which are in level $i-1$ (Fig. 12).

Each node sends this detected data to its parent node, this node receives data from the child nodes, fusions them and passes them to the parent node in the cluster. When the data of all nodes that are members of the cluster are received by the CH (root node), it sends them to the BS.

TRP is a protocol based on the construction of a tree where the base station is the root node (Gong and Jiang, 2011). In each round tree formation begins with each node that searches for a parent node closer to the base station. This procedure is performed by all nodes in the network to build the desired structure before to begin transferring data from the most remote nodes to the root node (Fig. 13).

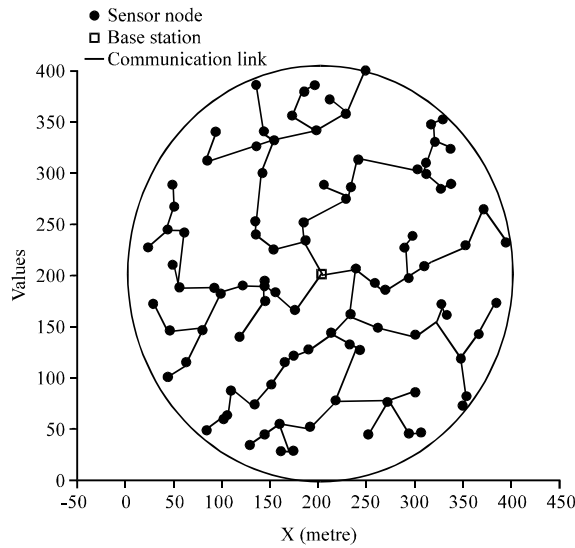


Fig. 14: Tree construction with TRP

GSTEB is a protocol based on the same procedures as TRP except that the base station does not play the role of the root node (Han *et al.*, 2014). In this protocol, the root node is chosen in each round by the base station based on the highest residual energy of the nodes and this information is broadcast to all nodes of the network so that each node selects its parent node according to the distance with the root node (Fig. 14).

There is another case in the GSTEB protocol where the base station plays the role of the root node. In this case, all nodes in the network know that the root node is the base station. Then, each node looks for a relay node that is a neighbor node closer to the root than itself and that has a higher residual energy of all these neighbors (which will cause minimal consumption between all neighbors, one of these consumption is the sum of the consumption of the sensor node towards a relay node and that of the relay node to BS), this node will be selected as the parent node. If the sensor node can not find an appropriate parent node, it will transmit its data directly to BS (Fig. 15).

SSTBC is a protocol based on the creation of clusters in the zone of interest and in each cluster, it will have a node CH which will play the role of the root to form a minimum spanning tree using the prim algorithm (Tan and Viet, 2015). This protocol also uses a process to minimize the energy for nodes that relies on disabling the radio from a number of unnecessary nodes that can detect the same information to suppress the redundant data. In this protocol, the base station is responsible for the formation of clusters, sleep planning and the construction of minimum spanning trees.

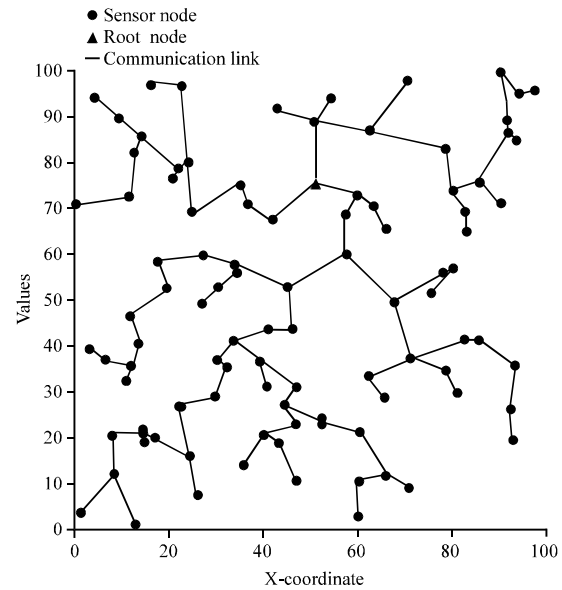


Fig. 15: Network topology in GSTEB with a different root node than the base station (bs located at (50, 175 m))

In each round, all nodes send a message to the base station to inform them of their locations and residual energies. The base station divides the zone of interest into 5 clusters, only in the first round. After each round, the base station partitions the nodes detection area to virtual square grids whose size is less than a distance threshold value. The candidate node with the highest residual energy will be in active mode and other nodes will be in standby mode in the same grid for the current round. In each round, the base station selects a CH in each cluster (Fig. 16). The selection is based on the calculation of the average energy and also of a cost, the node in a cluster that has a residual energy higher than that of the average energy and also has the highest cost, this node is elected like a CH in this round. Once the zone of interest is divided into grids and 5 clusters and in each cluster a CH is selected. So, the node CH will play the role of the root to build minimum spanning tree.

Cluster-based algorithms:In this family of algorithms, sensor nodes self-organize to form clusters. Each cluster has a cluster head which is responsible to transmit all data from its cluster to the base station. The most used algorithms of this family are:

LEACH (Heinzelman *et al.*, 2000) is the first algorithm based on hierarchical clustering whose CHs Cluster Heads) is are selected randomly and in turn according to the threshold $T(n)$ given by:

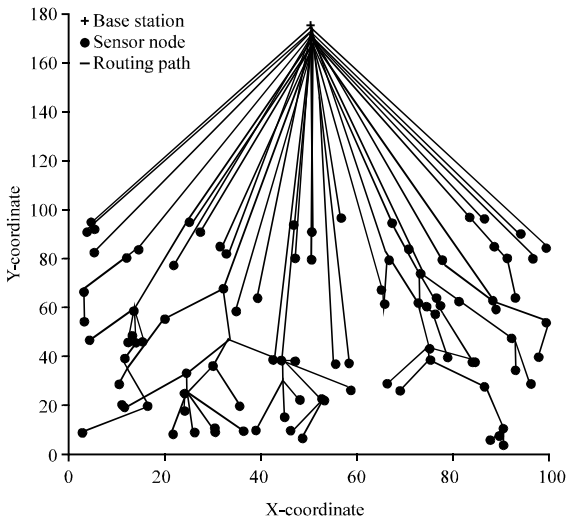


Fig. 16: Network topology in GSTEB with base station as root

$$T(n) = \frac{P}{1-P \times (r \bmod \frac{1}{P})} \forall n \in G$$

$$T(n) = 0 \forall n \notin G$$

Where:

- P = Desired percentage of cluster heads (p = 0.05)
- r = Current round
- G = Set of nodes that have not been cluster-heads in the last 1/P rounds

Once the CHs are chosen, they start sending advertising messages to build the clusters. Each non-cluster-head node decides the cluster to which it will belong for this round. This decision is based on the received signal strength of the advertisement (Fig. 17).

EECS is a protocol based on the formation of clusters using almost the same procedures as LEACH, except that the selection of CHs is different (Ye *et al.*, 2005). The choice of the CH is based on the residual energy of all the candidate nodes in the network. This energy is calculated by the base station when it receives the message “Hello” from the different nodes.

DAIC is a protocol that uses two zones to build clusters in the network (Gautam and Pyun, 2010). The zones are created based on the vertical distances between each sensor node in the network and the base station. These distances are computed by the base station. The zones are divided into primary zone which contains CHs and CH gateway and the secondary zone contains only

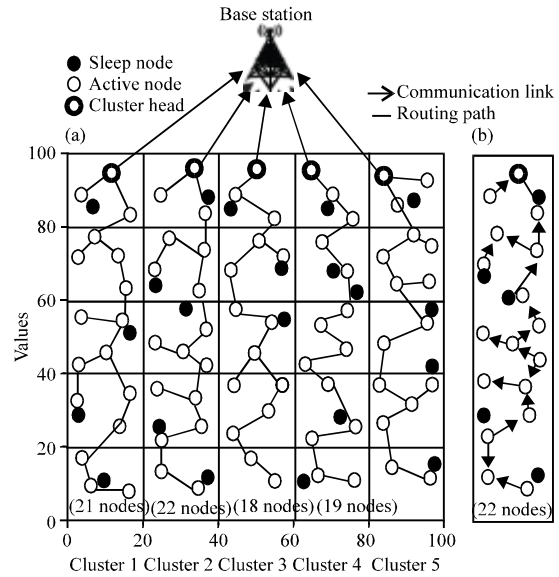


Fig. 17: a) The 100-node in 100×100 m network with virtual grid 5×5 m and b) An example for minimum spanning tree

Chs. The CHs of the primary and secondary zones form clusters in their zones and the gateway CHs make connections between them to the base station (Fig. 18).

CBRP is a protocol based on the construction of clusters using the CHs selected with the same procedure as LEACH (Zarei *et al.*, 2010). To transmit the data collected by the CHs to the base station, a tree is created by them using the minimum spanning tree.

EADUC is a hierarchical protocol that uses the same algorithm as LEACH to select CHs and clustering and also uses the concept of tree mapping between CHs to transmit data to the base station (Yu *et al.*, 2011). The transmission of data from the CH to the base station is based on calculating the distance between the CH and the BS and the distance DIST_TH (this algorithm defines an appropriate TH DIST to ensure that all CH can find their next hop). In the case where the distance from CH to BS is smaller than DIST TH, then the CH communicates directly with the BS and defines the BS as its next jump. Otherwise, it communicates with the BS using another CH.

EADC is a protocol based on the construction of clusters with the same algorithms and procedures as the EADUC protocol (Yu *et al.*, 2011). The only difference is that the CH selects the next CH with a higher relay value and closer to the BS than its next hop. If several neighboring CHs have the same “Relay” value, CH

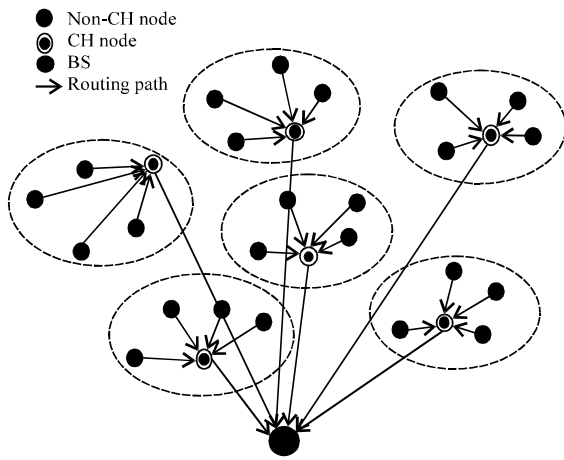


Fig. 18: Architecture LEACH

chooses the one with the larger $d(s_j, BS)$ value to avoid premature death of CHs close to the BS due to the excessive transfer of data.

MH-LEACH (Neto et al., 2014; Wang et al., 2009):

The principle of this protocol is based on the selection of the CHs to form clusters according to the following threshold $T(n)$:

$$T(n) = \frac{P}{1 - P \times (r \bmod \frac{1}{p})} Q$$

$$Q = \left\{ \frac{E_{residual}}{E_{initial}} + \left(1 - \frac{E_{residual}}{E_{initial}}\right) \frac{N_num/P}{H_times+1} \right\}$$

Where:

- $E_{residual}$ = The node current energy
- $E_{initial}$ = The maximal energy of the node
- N_num = The number of adjacent nodes
- H_times = Represents the times of the node has been selected as cluster head

Once the CHs are selected and the clusters are formed, the transmission of the data to the base station uses the creation of the initial and final routing tables according to the RSSI power received by the closest neighbors on the one hand and on the other hand according to the distance of the base station to ensure the sense of data transmission (Fig. 20a, b).

NEAHC (Ke et al., 2016) is a protocol based on the formation of clusters by the selection of CHs in each round using the threshold T_h which is based on the residual energy:

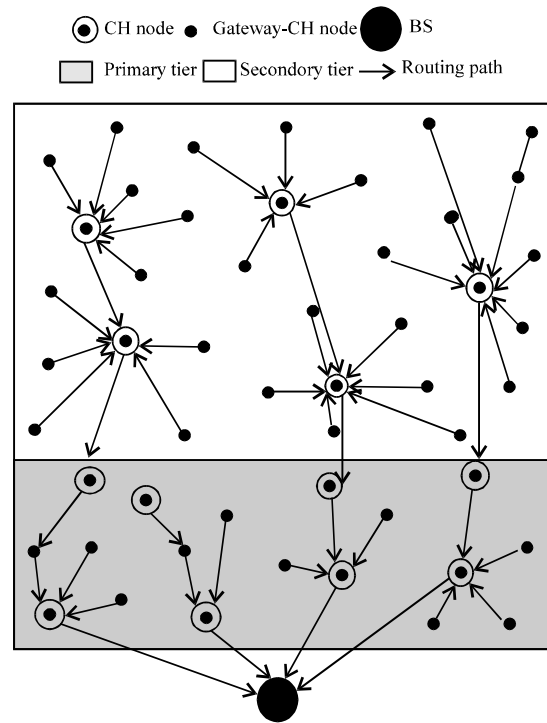


Fig. 19: The network topology in DAIC

$$T_h = \begin{cases} \frac{P}{1 - P \left[r \bmod \left(\frac{1}{p} \right) \right]} \frac{1}{1 + e^{-\alpha E_i}}; n \in G \end{cases}$$

Where:

- p = The desired percentage of cluster heads
- r = The current round
- G = The set of nodes that have not become the cluster head in last $1/p$ rounds

" $\alpha > 0$ is a constant, $1/(1+e^{-\alpha E_i})$ is a monotone increasing function about E_i ; If the chosen random number less than T_h , then n becomes the cluster head for the current round r . This function of threshold makes the nodes with more residual energy have a better chance to become the cluster heads, thus, balancing the energy consumption of the network.

After the cluster formation, each CH selects the data transmission path to the base station using the residual energy of all neighboring CHs and also the transmission cost. The chosen CH is either the direct communication to the BS or the communication to an optimal neighboring CH according to the residual energy and the cost of communication (Fig. 19).

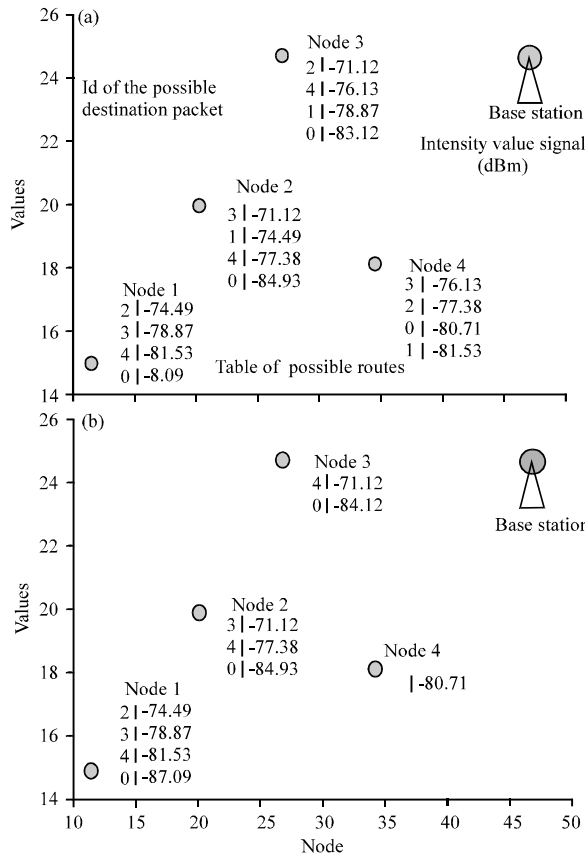


Fig. 20: Initial and final routing tables

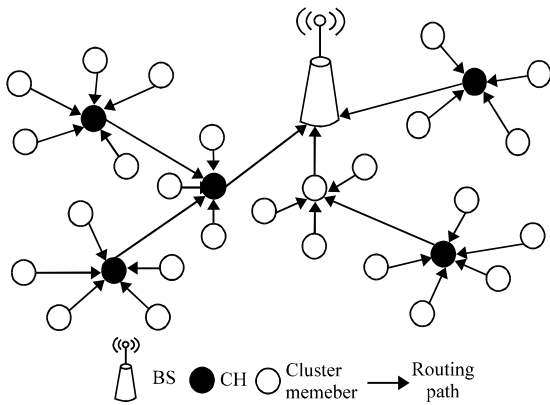


Fig. 21: Multi-hop network topology in NEAHC

Figure 20 and 21 the NEAHC protocol consumes a lot of energy resources in the calculations made to look for the optimal neighbor for each CH knowing that these CHs remain the most of the times turn on to collect data from their cluster which will cause their depletion to abruptly and also cause a decrease in the lifetime of the network. On the other hand, this protocol uses an

adequate and efficient formula for the selection of the CHs in each round which improved the balance of the load between the different nodes of the network and also increases the stability of the system.

RESULTS AND DISCUSSION

Comparative study: In this study, we will carry out a comparison study based on the ROC multi criteria analysis method between the different algorithms studied in this study according to different metrics and performance criteria.

Comparison between the different algorithms studied according to their families on the basis of different metrics: In order to compare the different types of self-organization algorithms studied in the study, we propose some criteria and metrics that describe the behavior and architecture of wireless sensor networks and also affect their limitations and deployments as load balancing, energy efficiency, scalability, stability, delivery time and complexity of the algorithm (Liu, 2012; Abbasi and Younis, 2007; Boyinbode *et al.*, 2011; Haneef and Zhongliang, 2012).

Table 1 shows the comparison between the different algorithms studied according to their families on the basis of different criteria using a conventional notation system.

This system is used for the purpose of a better and an easy understanding of the study and does not provide a standardized notation for the comparison.

Compariso between the different self-organization algorithms studied using the ROC multicriteria analysis method: In this comparative study, we will compare the different self-organization algorithms studied according to the most used criteria and metrics (load balancing, scalability, delivery time, complexity of the algorithm energy efficiency, stability) using the ROC multicriteria analysis method. This method is based on the calculation of the score of each self-organization algorithm studied according to the criteria and the chosen metrics. To perform this calculation, we first have to determine the weight associated with each metric for this reason we used a multicriteria analysis method (Roy, 2005). Multi-Criteria Decision Analysis (MCDA) is a powerful and critical tool and method that can be applied to many complex decisions. It allows in delicate and complex cases to properly analyze the phenomena or the systems in order to take a definitive decision to implement an action or a change by obtaining solutions that meet the criteria chosen in principle.

Table 1: Comparative table of different algorithms studied according to their families on the basis of different metrics

Algorithms by family	The performance metrics					The complexity of the algorithm
	Load balancing	Energy efficiency	Scalability	Stability	The delivery time	
Chain-based algorithm						
PEGASIS	Moderate	Low	Very Low	Moderate	Very long	High
EECB	Moderate	Moderate	Very Low	Moderate	Very long	High
PDCH	Moderate	Moderate	Long	Low	Long	High
CCM	Low	Moderate	Moderate	Moderate	Moderate	Low
ECCP	Low	Low	Moderate	Low	Long	High
SCBC	High	High	Moderate	Low	Long	High
PEGASIS multi chain	Low	Low	Moderate	Low	Moderate	Moderate
Tree-based algorithms						
PEDAP	Moderate	High	High	High	Very long	Low
TREEPSI	Low	Moderate	Moderate	Low	Very long	Moderate
TBC	Moderate	Moderate	Moderate	Moderate	Short	Moderate
TRP	High	Moderate	Low	Low	Long	High
GSTEB	Low	Moderate	Moderate	Low	Long	Moderate
SSTBC	Low	Low	Moderate	Moderate	Moderate	Very high
Cluster-based algorithms						
LEACH	Moderate	Moderate	Low	Moderate	Very short	Low
EECS	Moderate	Low	Low	Moderate	Short	Moderate
DAIC	Low	Moderate	Moderate	Moderate	Moderate	High
CBRP	Moderate	Moderate	Low	Moderate	Moderate	High
EADUC	Low	Low	Low	Low	Short	High
EADC	Moderate	Low	Low	Moderate	Moderate	High
MH-LEACH	High	Moderate	Moderate	Moderate	Moderate	High
NEAHC	High	Moderate	Moderate	Moderate	Moderate	Very high

Table 2: Calculation of criteria weights

Variables	Load balancing	Energy efficiency	Scalability	Stability	The delivery time	The complexity of the algorithm	Control
R1	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.00
R2	0.7500	0.2500	0.0000	0.0000	0.0000	0.0000	1.00
R3	0.6111	0.2778	0.1111	0.0000	0.0000	0.0000	1.00
R4	0.5208	0.2708	0.1458	0.0625	0.0000	0.0000	1.00
R5	0.4567	0.2567	0.1567	0.0900	0.0400	0.0000	1.00
R6	0.4083	0.2417	0.1583	0.1028	0.0611	0.0278	1.00
Average	0.6245	0.2162	0.0953	0.0426	0.0169	0.0046	1.00

The score is calculated according to a number of criteria according to which the list is not exhaustive. The overall score is obtained by adding the partial scores (criteria) affected by relative weights. In the analysis of decisions, this operation is called additive synthesis or aggregation. As regards the evaluation of the relative weight of the criteria, there are several methods of multicriteria decision analysis. We have selected (ROC) (Barron, 1992; Danielson and Ekenberg, 2016; Hajoui *et al.*, 2015) for its simplicity and proven effectiveness. Several methods of weight selection including Equal Weights (EW) and Rank-Order Centroid (ROC) weights have been proposed and evaluated (Butler and Olson, 1999). A common finding of these studies is that ROC weights appear to be better than other strategic patterns in terms of choice accuracy.

This method is a simple way to give weight to a number of items classified according to their importance. Policy makers can generally classify items much more easily than giving them weight. The centroid method

affects the following weights where W_1 is the largest weight of the objective, W_2 is the weight of the second most important objective and so on:

$$D = \left\{ W_1 \geq W_2 \geq \dots, W_m \geq 0 \text{ and } \sum_{j=1}^m W_j = 1 \right\}$$

This method takes these rows as inputs and converts them into weights for each of the elements. The conversion is based on the following equation:

$$W_j = \frac{1}{m} \left(\frac{1}{j} + \frac{1}{j+1} + \dots + \frac{1}{m} \right)$$

To begin the comparative study. First, the criteria must be ordered in descending order according to the importance of the latter: load balancing <energy efficiency> <scalability> <stability> <the delivery time> the complexity of the algorithm (Table 2). Table 3 shows the result of the weight calculation of each criterion using the following equation:

Table 3: Criteria weights

Variables	Load balancing	Energy efficiency	Scalability	Stability	The delivery time	The complexity of the algorithm	Control
Weights	0.6245	0.2162	0.0953	0.0426	0.0169	0.0046	1.00

Table 4: The score of each self-organization algorithm according to the associated criteria using the indicated notation system

Algorithms by family	The performance metrics							Global score	Average per family
	Load balancing	Energy efficiency	Scalability	Stability	The delivery time	The complexity of the algorithm			
Chain-based algorithms									
PEGASIS	5	3	1	5	1	3	4.1101	4.2962	
EECB	5	5	1	5	1	3	4.5425		
PDCH	5	5	3	3	3	3	4.6817		
CCM	3	5	5	5	5	7	3.3283		
ECCP	3	3	5	3	3	3	2.9153		
SCBC	7	7	5	3	3	3	6.1551		
PEGASIS multi chain	3	3	5	3	5	5	3.0433		
Tree-based algorithms									
PEDAP	5	7	7	7	1	7	4.5967	4.5245	
TREEPSI	3	5	5	3	1	5	3.4081		
TBC	5	5	5	5	7	5	4.8775		
TRP	7	5	3	3	3	3	5.9307		
GSTEB	3	5	5	3	3	5	3.6663		
SSTBC	3	3	5	5	5	1	3.4913		
Cluster-based algorithms									
LEACH	5	5	3	5	9	7	4.6961	4.7185	
EECS	5	3	3	5	7	5	3.7883		
DAIC	3	5	5	5	5	3	3.7423		
CBRP	5	5	3	5	5	3	4.8007		
EADUC	3	3	3	3	7	3	3.5003		
EADC	5	3	3	5	5	3	4.3683		
MH-LEACH	7	5	5	5	5	3	6.0497		
NEAHC	7	5	5	5	5	1	6.2311		

$$W_i^{ROC} = \frac{1}{N} \sum_{j=1}^N \frac{1}{j}$$

The control column ensures that all weights are normalized (sum of weights = 1). Consequently, we have the weight of each criterion as shown in Table 3.

Table 4 shows the notation attributed to each self-organization algorithm according to the criteria and metrics studied. The scoring system adopted is as follows:

- C Less important
- C Equally important
- C Moderately important
- C Much more important
- C Considerably more important

Table 4 also shows the overall score calculated using the following formula: global score = (0.6245*load balancing)+(0.2162*energy efficiency)+(0.0953*scalability)+(0.0426*stability)+(0.0169* the delivery time y)+(0.0046* the complexity of the algorithm). Figure 22 illustrates the results of the overall score in the form of a graph.

According to the comparative study carried out in this study is based on the study of the different types

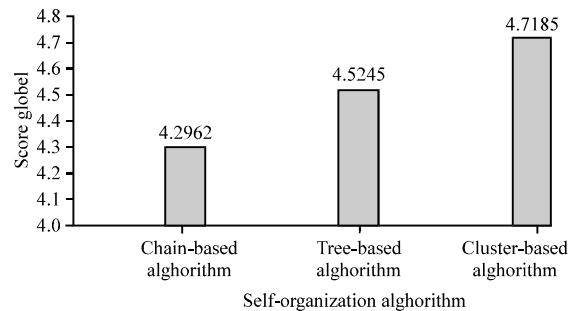


Fig. 22: The overall score as a graph

of self-organization algorithms in the WSNs using some criteria and metrics based on the multicriteria analysis method ROC, we can notice that the family which best meets performance and metric criteria is the family of clustering algorithms that allow better management of energy efficiency and also for the improvement of network life.

CONCLUSION

In recent years, Wireless Sensor Networks (WSNs) have achieved tremendous feat in various fields for better collaboration, deployment and efficiency in data detection by sensor nodes according to the service requested by

applications. There is not a predefined infrastructure for network deployment, they use an efficient self-organization algorithm that help the sensor nodes to form an appropriate topology and transmit the sensed data to a base station.

In this study, we have carried out an in-depth study on the different types of self-organization algorithms most common in WSNs. During the analysis, we have found that several algorithms have been developed for improving energy efficiency and also increasing network lifetime by using clustering and the other algorithms have been focused in load balancing and the improvement of energy consumption in multi-hop communication using chains or trees. But the choice between protocols depends on the desired application, i.e., there can be an algorithm of a specific family (cluster, tree or chain) works very widely in a given application on the other hand another algorithm of another family can give poor results despite these better performances and these better criteria. For this reason, we will develop an approach based on self-organization algorithm that adapts to the most applications and which will give better performances.

SUGGESTIONS

This approach will be applied in the management of an intelligent parking system in order to improve the system and also facilitate the task for users to find a vacant parking space remotely in an efficient and practical way.

REFERENCES

- Abbasi, A.A. and M. Younis, 2007. A survey on clustering algorithms for wireless sensor networks. *Comput. Commun.*, 30: 2826-2841.
- Akyildiz, I.F., W. Su, Y. Sankarasubramanian and E. Cayirci, 2002. A survey on sensor networks. *IEEE Commun. Mag.*, 40: 102-114.
- Al-Karaki, J.N. and A.E. Kamal, 2004. Routing techniques in wireless sensor networks: A survey. *IEEE Wireless Commun.*, 11: 6-28.
- Barron, F.H., 1992. Selecting a best multiattribute alternative with partial information about attribute weights. *Acta. Psychol.*, 80: 91-103.
- Boyinbode, O., H. Le, A. Mbogho, M. Takizawa and R. Poliah, 2011. A survey on clustering algorithms for wireless sensor networks. *Proceedings of 2010 13th International Conference on Network-Based Information Systems*, September 14-16, 2010, IEEE, Takayama, Japan, ISBN:978-1-4244-8053-1, pp: 358-364.
- Butler, J. and D.L. Olson, 1999. Comparison of centroid and simulation approaches for selection sensitivity analysis. *J. Multicriteria Decis. Anal.*, 8: 146-161.
- Danielson, M. and L. Ekenberg, 2016. Trade-Offs for Ordinal Ranking Methods in Multi-criteria Decisions. In: *Group Decision and Negotiation Theory, Empirical Evidence and Application*, Bajwa, D., S. Koeszegi and R. Vetschera (Eds.). Springer, Cham, Switzerland, ISBN:978-3-319-52623-2, pp: 16-27.
- Gautam, N. and J.Y. Pyun, 2010. Distance aware intelligent clustering protocol for wireless sensor networks. *J. Commun. Networks*, 12: 122-129.
- Gong, B. and T. Jiang, 2011. A tree-based routing protocol in wireless sensor networks. *Proceedings of the International Conference on Electrical and Control Engineering*, September 16-18, 2011, Yichang, China, pp: 5729-5732.
- Hajoui, O., R. Dehbi, M. Talea and Z.I. Batouta, 2015. An advanced comparative study of the most promising NoSQL and NewSQL databases with a multi-criteria analysis method. *J. Theor. Appl. Inf. Technol.*, 81: 579-588.
- Han, Z., J. Wu, J. Zhang, L. Liu and K. Tian, 2014. A general self-organized tree-based energy-balance routing protocol for wireless sensor network. *IEEE. Transac. Nucl. Sci.*, 61: 732-740.
- Haneef, M. and D. Zhongliang, 2012. Design challenges and comparative analysis of cluster based routing protocols used in wireless sensor networks for improving network life time. *Adv. Inform. Sci. Service Sci.*, 4: 450-459.
- Heinzelman, W.R., A. Chandrakasan and H. Balakrishnan, 2000. Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, January 7, 2000, IEEE, New York, USA., ISBN:0-7695-0493-0, pp: 1-10.
- Kalantary, S. and S. Taghipour, 2014. A survey on architectures, protocols, applications and management in wireless sensor networks. *J. Adv. Comput. Sci. Technol.*, 3: 1-11.
- Ke, W., O. Yangrui, J. Hong, Z. Heli and L. Xi, 2016. Energy aware hierarchical cluster-based routing protocol for WSNs. *J. China Univ. Posts Telecommun.*, 23: 46-52.
- Kim, K.T., C.H. Lyu, S.S. Moon and H.Y. Youn, 2010. Tree-based clustering (TBC) for energy efficient wireless sensor networks. *Proceedings of the 2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, April 20-23, 2010, IEEE, Perth, Australia, ISBN: 978-1-4244-6701-3, pp: 680-685.
- Lan, F.A. and C.V. Mehmet, 2010. *Wireless Sensor Networks*. John Wiley & Sons, Hoboken, New Jersey, USA., Pages: 493.

- Lindsey, S. and C.S. Raghavendra, 2002. PEGASIS: Power-efficient gathering in sensor information systems. Proceedings of the Conference on Aerospace Vol. 3, March 9-16, 2002, IEEE, Big Sky, Montana, USA., ISBN:0-7803-7231-X, pp: 3-1125-3-1130.
- Linping, W., B. Wu, C. Zhen and W. Zufeng, 2010. Improved algorithm of PEGASIS protocol introducing double cluster heads in wireless sensor network. Proceedings of the 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE) Vol. 1, August 24-26, 2010, IEEE, Changchun, China, ISBN:978-1-4244-7957-3, pp: 148-151.
- Liu, X.X., 2012. A survey on clustering routing protocols in wireless sensor networks. *Sensors*, 12: 11113-11153.
- Neto, J.H.B., A. Rego, A. Cardoso and J. Celestino, 2014. MH-LEACH: A distributed algorithm for multi-hop communication in wireless sensor networks. Proceedings of the 2014 13th International Conference on Networks ICN, February 23, 2014, Nice, France, ISBN:978-1-61208-318-6, pp: 55-61.
- Patel, B. and J. Munjani, 2016. Chain based routing protocol for wireless sensor network. *Intl. J. Comput. Appl.*, 134: 21-24.
- Roy, B., 2005. Paradigms and Challenges. In: *Multiple Criteria Decision Analysis: State of the Art Surveys*, Figueira, J., S. Greco and M. Ehrgott (Eds.). 1st Edn. Springer, USA., pp: 3-24.
- Satapathy, S.S. and N. Sarma, 2006. TREEPSI: Tree based energy efficient protocol for sensor information. Proceedings of International Conference on Wireless and Optical Communications Networks, April 2006, Bangalore, pp: 4-4.
- Sheikhpour, R., S. Jabbehdari and A. Khademzadeh, 2012. A cluster-chain based routing protocol for balancing energy consumption in wireless sensor networks. *Intl. J. Multimedia Ubiquitous Eng.*, 7: 1-16.
- Sirsikar, S., A. Chunawale and M. Chandak, 2014. Self-organization architecture and model for wireless sensor networks. Proceedings of the 2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies (ICESC), January 9-11, 2014, IEEE, Nagpur, India, ISBN:978-1-4799-2102-7, pp: 204-208.
- Tan, H.O. and I. Korpeolu, 2003. Power efficient data gathering and aggregation in wireless sensor networks. *ACM. Sigmod Rec.*, 32: 66-71.
- Tan, N.D. and N.D. Viet, 2015. SCBC: Sector-chain based clustering routing protocol for energy efficiency in heterogeneous wireless sensor network. Proceedings of the 2015 International Conference on Advanced Technologies for Communications (ATC), October 14-16, 2015, IEEE, Ho Chi Minh City, Vietnam, ISBN:978-1-4673-8372-1, pp: 314-319.
- Tan, N.D. and N.D. Viet, 2015. SSTBC: Sleep scheduled and tree-based clustering routing protocol for energy-efficient in wireless sensor networks. Proceedings of the 2015 IEEE RIVF International Conference on Computing and Communication Technologies-Research, Innovation and Vision for the Future (RIVF), January 25-28, 2015, IEEE, Can Tho, Vietnam, ISBN:978-1-4799-8043-7, pp: 180-185.
- Tang, F., I.Y. You, S. Guo, M. Guo and Y. Ma, 2012. A Chain-cluster based routing algorithm for wireless sensor networks. *J. Intell. Manuf.*, 23: 1305-1313.
- Walters, J.P., Z. Liang, W. Shi and V. Chaudhary, 2007. *Wireless Sensor Network Security: A Survey*. In: *Security in Distributed, Grid, Mobile and Pervasive Computing*, Xiao, Y. (Ed.). CRC Press, Boca Raton, Florida, ISBN: 9780849379253, pp: 367-417.
- Wang, G., H. Zhu, H. Dai, L. Wu and B. Xiong, 2009. The clustering algorithm of wireless sensor networks based on multi-hop between clusters. Proceedings of the 2009 WRI World Congress on Computer Science and Information Engineering Vol. 3, March 31-April 2, 2009, IEEE, Los Angeles, California, USA., ISBN:978-0-7695-3507-4, pp: 177-181.
- Ye, M., C. Li, G. Chen and J. Wu, 2005. EECS: An energy efficient clustering scheme in wireless sensor networks. Proceedings of the 24th International Performance Computing and Communication Conferences, April 7-9, 2005, IEEE Press, New York, pp: 535-540.
- Yu, J., Y. Qi, G. Wang, Q. Guo and X. Gu, 2011. An energy-aware distributed unequal clustering protocol for wireless sensor networks. *Intl. J. Distrib. Sens. Networks*, Vol. 2011,
- Yu, Y. and Y. Song, 2010. An energy-efficient chain-based routing protocol in wireless sensor network. Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCSM) Vol. 11, October 22-24, 2010, IEEE, Taiyuan, China, ISBN:978-1-4244-7235-2, pp: V411-V486.
- Zarei, B., M. Zeynali and V.M. Nezhad, 2010. Novel cluster based routing protocol in wireless sensor networks. *IJCSI. Intl. J. Comput. Sci. Issues*, 7: 32-36.

Performance Evaluation of Transport Protocols for Mobile Ad Hoc Networks

¹Ali H. Wheeb, ²Ameer H. Morad and ³Maad Issa AL Tameemi

¹Al-Khwarizmi College of Engineering, University of Baghdad, Baghdad Iraq

²Department of Information and Communication, Al-Khwarizmi College of Engineering,
University of Baghdad, Baghdad, Iraq

³Department of Computer Engineering, College of Engineer, University of Baghdad, Baghdad, Iraq

Abstract: Mobile Ad hoc Networks (MANETs) is a wireless technology that plays an important role in several modern applications which include military, civil, health and real-time applications. Providing Quality of Service (QoS) for this application with network characterized by node mobility, infrastructure-less, limitation resource is a critical issue and takes greater attention. However, transport protocols effected influential on the performance of MANET application. This study provides an analysis and evaluation of the performance for TFRC, UDP and TCP transport protocols in MANET environment. In order to achieve high accuracy results, the three transport protocols are implemented and simulated with four different network topology which are 5, 10, 30 and 50 nodes, respectively using well known Network Simulator (NS-2.35). Moreover, Constant Bit Rate (CBR) considered as a traffic source and On-demand Distance Vector (AODV) as the routing protocol. For evaluation performance, QoS metrics such as end-to-end delay, packet delivery ratio, throughput and jitter are measured. The results show that delay and jitter of TFRC are slightly less than UDP and TCP whereas UDP has the significantly better performance wise throughput.

Key words: MANET, TFRC, UDP, TCP, network topology, QoS metric

INTRODUCTION

The Mobile Ad Hoc Network (MANETs) a paradigm rising wireless communication technology (Conti and Giordano, 2014). Its minimum configuration and flexibility have made it suitable for different application such as vehicular system (Dharmaraja *et al.*, 2016) disaster recovery (Liu and Kato, 2016) and military operations (Meena and Vasanthi, 2016). Basically, MANET is a set of mobile nodes are connected via wireless links with the infrastructure-less networks. These mobile nodes can move dynamically which lead to changing topology constantly (Khan *et al.*, 2017). A hop is a link between two nodes established while both nodes are in the transmission range of each other. Whilst, a multi-hop path can be existed by connecting two nodes among other nodes which act as router nodes. Moreover, multi-hop communication is widely used in the MANET in order to increase network capacity (Patnaik *et al.*, 2015).

The mobility of nodes, congestion and wireless link nature are the main reason for packet loss and end-to-end delay in mobile ad hoc networks. Mobility leads to change the network topology dynamically. So, it may cause packet drop in various ways. A packet dropped in

intermediate nodes if the next hop link has broken or the queue that stores incoming packets is full. Also, it may be dropped if the route from source to destination is not available. Congestion occurs in a network when the data traffic exceeds the maximum bandwidth of a communication link (Zhang *et al.*, 2015).

In fact, the performance of a MANET strongly depends on the efficiency of the transporting protocol that is used. The TCP-Friendly Rate Control (TFRC) (Handley *et al.*, 2008) and User Datagram Protocol (UDP) (Postel, 1981) and Transmission Control Protocol (TCP) (Postel, 1981) are transport layer protocols provide communication between application. Here, TFRC is a connection-oriented protocol provides a congestion control mechanism and fairness, making it suitable for real-time application where smooth bitrate is important. Because connectionless and message-oriented are the features in UDP, it has attracted multimedia application. Whereas TCP is a highly reliable connection protocol provides a congestion control mechanism and guaranteed delivery of data. Table 1 explained a comparison of the features and service provided by three transport protocols.

Table 1: Features and service provided by TFRC, UDP, AND TCP

Features	TFRC	UDP	TCP
Connection oriented	Yes	No	Yes
Message oriented	No	Yes	No
Reliable	Yes	No	Yes
Congestion control	Yes	No	Yes
Sequence number	Yes	No	Yes

However, the transport protocols TFRC, UDP and TCP are working fine in wired networks and can support different application. But these protocols face challenges when research on wireless networks or mobile ad hoc network environment. Studies must be conducted to know the best transport protocol for the application. Hence, this study presents a performance evaluation of TFRC, UDP and TCP transport protocols in mobile ad hoc networks. Furthermore, four different network topology used in the simulation in order to get accurate results. In addition, QoS metrics namely end-to-end delay, throughput, packet delivery ration and jitter measured for the performance evaluation. The simulation result obtained may give a good idea of selecting the best protocol for applications in MANET environment.

Lierature review: Many studies are mainly focused on the evaluation and comparison of the performance of transport protocols in mobile ad hoc network. By Nor and Dakkak (2016) researcher compared the performance of TFRC and SCTP transport protocols in MANET. Two scenarios implemented in the simulation with respect to background traffic. The result shows that SCTP has higher throughput whilst TFRC has low delay value. Gharge and Valanjoo (2014) presented a performance evaluation of TCP variants over protocol in mobile ad hoc networks. TCP variants evaluated over four different routing protocols in two scenarios which are link failure and signal noise scenario.

By Rajaboina *et al.* (2016) researcher evaluated and compared the performance of TCP, UDP and TFRC protocols in static wireless ad hoc networks. The simulation of the three transport protocols is divided into two modes which are independently mode and interoperation mode. Based on the simulation result, UDP outperforms in term of throughput as compared to other protocols. TCP is fairness than TFRC. Sharma and Patidar (2016) evaluated the performance of proposed TCP in a mobile ad hoc network. The proposed-TCP simulated and compared with different TCP variants to achieve better results.

By Wheeb (2015, 2017) researcher evaluated and compared the performance of UDP, SCTP, TFRC and DCCP protocols for different application traffic in a wired

network environment. For the simulation, NS2 is used. Three different scenarios with different parameters are implemented in order to get high accuracy result. The result of this study shows that throughput of SCTP is higher than other protocol whereas DCCP performance is good in term of delay. Xiang and Yang (2018) evaluated reliability performance for mobile ad hoc networks based on transmission reliability.

MATERIALS AND METHODS

Simulation enviroment and parameter setting: The performance of transport protocols is studied under different condition of network size on mobile multi-hop ad-hoc networks. Specifically, four network topology used in the simulation experiments for evaluation of TFRC, UDP, TCP protocols which are 5, 10, 30 and 50 nodes respectively. In addition, NS-2.35 (Issariyakul and Hossain, 2011) used as the simulation tool, since, it is preferred by researchers interested in the field of networking. Figure 1a-d depicts the topologies of 5, 10, 30 and 50 nodes in the sumlation.

In order to simulate a network environment, the setting of simulation network parameter is required. Table 2 displays simulation parameters used in the study.

Finally, regarding all the simulation experiments, Constant Bit Rate (CBR) used as a traffic source, since, it generates a data rate similar to that generated by real-time applications. Moreover, On-demand Distance Vector (AODV) implemented as the routing protocol in all mobile nodes of the network. The mobile nodes distributed randomly and move using a random waypoint algorithm with varied pause time. Queue length setup to 100 packet max.

RESULTS AND DISCUSSION

In the simulation experiment, four performance metrics measured to evaluate the performance of the three different transport protocols. This metrics are end-to-end delay, packet loss rate, throughput and jitter (Floyd, 2008; Nor *et al.*, 2017) (Fig. 1a-d and Table 2).

End-to- end delay: The time taken by a packet to transmit from sender to receiver is called end-to-end delay. For the real-time application, end-to-end delay represents an important factor. However, end-to-end delay contains the sum of processing delay, queuing delay and propagation delay, etc. it is measured in seconds. The following equation is used to measure end-to-end delay value.

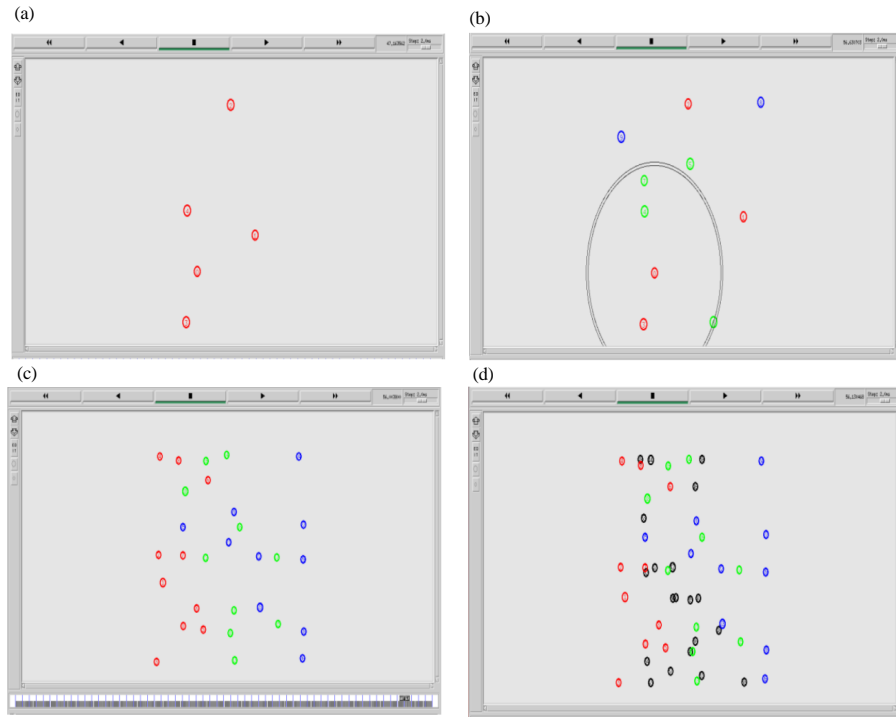


Fig. 1: a) Simulation topology of 5 nodes; b) Simulation topology of 10 nodes; c) Simulation topology of 30 nodes and d) Simulation topology of 50 nodes

Table 2: parameter setting of simulation

Parameter	Values
Simulator	NS 2.35
Channel	Wireless
Propagation	Two ray ground
Area (x)	600 m
Area (y)	600 m
MAC	802-11
Queue	Drop Tail-PriQueue
Routing protocol	AODV
Number of nodes	5, 10, 30, 50
Transport protocols	TFRC, UDP, TCP
Packet size	512 bytes
Traffic type	CBR
Simulation time	60 sec
Mobility model	Random way point

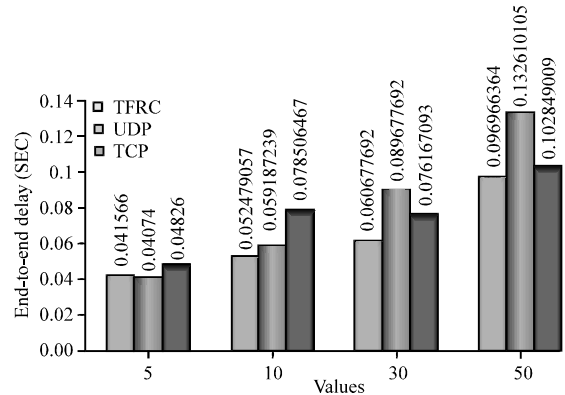


Fig. 2: End-to-End delay of TFRC, UDP and TCP for various node density

$$\text{End-to-end delay} = T_r - T_s \quad (1)$$

Where:

T_s = Sending time of the measured packet

T_r = Receiving time for the same packet

As illustrated in Fig. 2, considerable variance in end-to-end delay is observed between the TFRC, UDP and TCP protocol. End-to-end delay of TFRC is low and less than that of UDP and TCP. UDP gives high end-to-end delay at a higher number of nodes, hence, the performance of UDP decrease as network density increases. At 5 and 10 nodes, end-to-end delay of TCP is higher than TFRC and UDP while at 30 and 50 nodes

end-to-end delay of TCP higher than TFRC but less than UDP. To conclude TFRC provide better performance delay at MANET environment.

Packet delivery ratio: Due to the wireless link nature, packet loss occurs in MANET more than in wired networks. Packet Delivery Ratio (PDR) is the rate between the number of packets received and the number of packets sent across the network. However, QoS of application reduced when the packet loss is increased. Equation 2 explains the method used to measure the PDR.

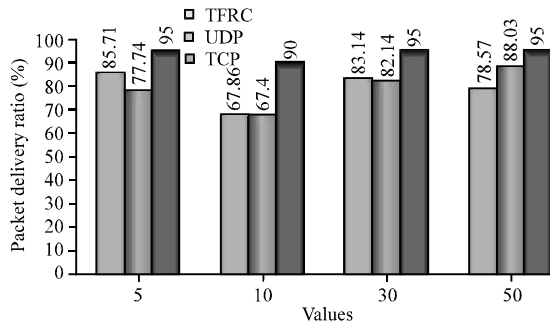


Fig. 3: Packet delivery ratio of TFRC, UDP and TCP for various node density

$$PDR = \frac{\sum \text{Total number of packet received}}{\sum \text{Total number of packet send}} \quad (2)$$

Figure 3 shows that the packet delivery ratio of TCP is relatively higher than UDP and TFRC. Both in TFRC and UDP loss many packets at 10 nodes. Notice clearly that UDP drop packet more than two protocols in all network nodes because it does not have a congestion control mechanism. Whilst, the packet delivery ratio of TCP is high and independent with consideration of increase network nodes. The result emphasizes the reliability of TCP performance in an environment of the MANET.

Throughput: Throughput is the amount of packet transmitted from source nodes to destination nodes, through the network at a specified time. Usually, it is rated in terms of packets per second or bits per seconds. To achieve good performance, the throughput must be relatively high. The value of throughput is measured by the Eq. 3.

$$\text{Throughput} = \frac{\text{Received packets}}{\text{Last packet } S_T - \text{first packet } S_T} \quad (3)$$

where, S_T is sending time of a packet. According to Fig. 4 at 5 nodes throughput of TFRC is relatively less as compared to UDP and TCP. Also, Throughput of UDP is little better than the throughput of TCP. As the number of nodes increases to 10, 30 and 50 nodes the traffic load increase in the network. At the same time, UDP achieves the highest throughput superior on TFRC and TCP. The reason is UDP send packets at a constant transmission rate regardless the network state. Furthermore, the overall throughput of TCP is higher than that of TFRC because TCP investing the bandwidth efficiently. Based on the results, UDP performs best among other two protocols at all network density.

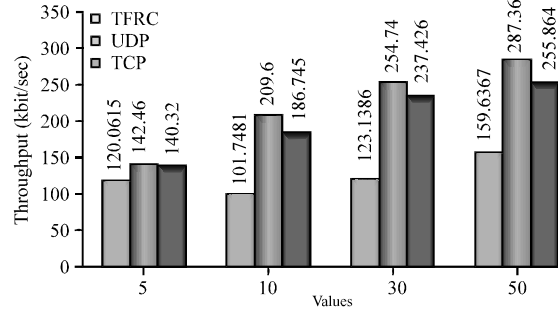


Fig. 4: Throughput of TFRC, UDP and TCP for various node density

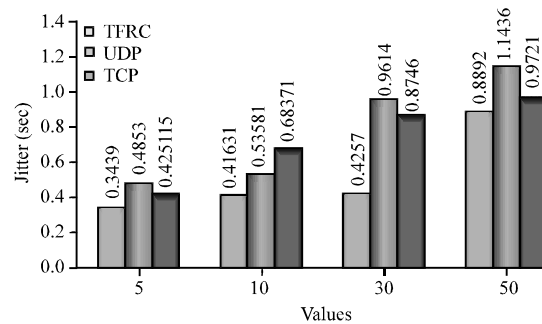


Fig. 5: Jitter of TFRC, UDP and TCP for various node density

Jitter: Jitter is the variation of packet arrival delay. It can be measured as the difference of delay of the current packet and previous packet. The streaming application such as a video stream or audio application is sensitive to delay and prefer a low value of jitter. Jitter is measured using Eq. 4:

$$\text{Jitter} = \text{Delay}_n - \text{Delay}_{n-1} \quad (4)$$

Where:

Delay_n = Delay of the current packet

Delay_{n-1} = Delay of the previous packet

Jitter performance analysis of TFRC, UDP and TCP protocols at 5, 10, 30 and 50 nodes is shown in Fig. 5. The result indicates that the TFRC is the best protocol concerning the jitter value. Also, it can be observed that UDP gives a high value of jitter and increases when the number of nodes increases. The reason behind this is jitter related to delay time. Jitter in TCP is better compared to UDP but less compared to TFRC. The reason is TCP monitor the network state. If congestion occurs in the network, then it reduces the data rate and stays monitor until next event which leads to the high value of jitter.

CONCLUSION

The performance of TFRC, UDP and TCP in MANET was analyzed and evaluated. Four different network size used in the simulation. AODV consider as the routing protocol in mobile nodes and CBR employ as data traffic. Moreover, the performance of the three transport protocols evaluated according to QoS metrics. The experiment results show that TFRC performs better other protocol in term of end-to-end delay and jitter. Further, UDP gives a higher throughput subsequently it suitable for video applications. On the other hand, TFRC is appropriate for real-time applications like VoIP in MANET environment.

SUGGESTIONS

In future research, non-standard transporting protocols and a high number of mobile nodes can be included.

REFERENCES

- Conti, M. and S. Giordano, 2014. Mobile ad hoc networking: Milestones, challenges and new research directions. *IEEE Commun. Mag.*, 52: 85-96.
- Dharmaraja, S., R. Vinayak and K.S. Trivedi, 2016. Reliability and survivability of vehicular ad hoc networks: An analytical approach. *Reliab. Eng. Syst. Safety*, 153: 28-38.
- Floyd, S., 2008. Metrics for the evaluation of congestion control mechanisms. *J. Inf.*, 1: 1-23.
- Gharge, S. and A. Valanjoo, 2014. Simulation based performance evaluation of TCP variants and routing protocols in Mobile Ad-hoc Networks. *Proceedings of the International Conference on Advances in Engineering and Technology Research (ICAETR'14)*, August 1-2, 2014, IEEE, Unnao, India, ISBN:978-1-4799-6394-2, pp: 1-8.
- Handley, M., J. Padhye, S. Floyd and J. Widmer, 2008. TCP Friendly Rate Control (TFRC): Protocol specification. *Internet Eng. Task Force*, 1: 1-58.
- Issariyakul, T. and E. Hossain, 2011. *Introduction to Network Simulator (NS2)*. Springer, Boston, Massachusetts, USA., ISBN:978-1-4614-1405-6, Pages: 510.
- Khan, M.S., D. Midi, M.I. Khan and E. Bertino, 2017. Fine-grained analysis of packet loss in MANETs. *IEEE Access*, 5: 7798-7807.
- Liu, J. and N. Kato, 2016. A Markovian analysis for explicit probabilistic stopping-based information propagation in postdisaster ad hoc mobile networks. *IEEE Trans. Wirel. Commun.*, 15: 81-90.
- Meena, K.S. and T. Vasanthi, 2016. Reliability analysis of mobile ad hoc networks using universal generating function. *Qual. Reliab. Eng. Intl.*, 32: 111-122.
- Nor, S.A. and O. Dakkak, 2016. Comparative Study on the Performance of TFRC and SCTP Over AODV in MANET. In: *Wireless Communications, Networking and Applications*, Zeng, Q.A. (Ed.). Springer, New Delhi, India, ISBN:978-81-322-2579-9, pp: 363-370.
- Nor, S.A., R. Alubady and W.A. Kamil, 2017. Simulated performance of TCP, SCTP, DCCP and UDP protocols over 4G network. *Procedia Comput. Sci.*, 111: 2-7.
- Patnaik, S., X. Li and Y.M. Yang, 2015. *Recent Development in Wireless Sensor and Ad-Hoc Networks*. Springer, Berlin, Germany,.
- Postel, J., 1981. *Transmission control protocol*. Master Thesis, Defense Advanced Research Projects Agency, Arlington County, Virginia, USA.
- Rajaboina, R., P.C. Reddy, R.A. Kumar and N. Venkatramana, 2016. Performance comparison of TCP, UDP and TFRC in static wireless environment. *Intl. J. Inf. Comput. Secur.*, 8: 158-180.
- Sharma, N. and G. Patidar, 2016. Improved congestion control mechanism using modified hybrid-TCP in mobile ad-hoc networks. *Proceedings of the 2nd International Conference on Computational Intelligence and Communication Technology (CICT)*, February 12-13, 2016, IEEE, Ghaziabad, India, ISBN:978-1-5090-0211-5, pp: 312-315.
- Wheeb, A.H., 2015. Performance comparison of transport layer protocols. *Intl. J. Adv. Res. Comput. Sci. Software Eng.*, 5: 121-125.
- Wheeb, A.H., 2017. Performance evaluation of UDP, DCCP, SCTP and TFRC for different traffic flow in wired networks. *Intl. J. Electr. Comput. Eng.*, 7: 3552-3557.
- Xiang, S. and J. Yang, 2018. Performance reliability evaluation for mobile ad hoc networks. *Reliab. Eng. Syst. Safety*, 169: 32-39.
- Zhang, X.M., Y. Zhang, F. Yan and A.V. Vasilakos, 2015. Interference-based topology control algorithm for delay-constrained mobile ad hoc networks. *Mobile Comput. IEEE Trans.*, 14: 742-754.

Evaluation of Strategic Leadership Practices and Critical Success Factors for Competitiveness in Iraqi Manufacturing Sector a Descriptive Analysis

¹Hamad Karem Hadrawi, ²Araden Hatim Khudair Al-Obaidi and ³Zayad Ali Abbas
¹University of Kufa, Faculty of administration and economics, Kufa, P.O. Box 21 Kufa, Iraq
²University of Al-Mustansiriya, Faculty of administration and economics, Baghdad, Iraq
³University of Baghdad, Faculty of administration and economics

Abstract: The current state of Iraq economy and the manufacturing sector makes it is necessary to search for the successful practices that are appropriate to the success of the companies. This research aims to evaluate strategic leadership practices in Iraqi's cement manufacturing sector with a view of proffering managerial implications to improve competitiveness. A sample of 134 engineers in the two most active cement factories in the middle of Iraq was determined for this study. Descriptive analysis was used to show respondent's views of strategic leadership practices and critical success factors for competitiveness. Findings revealed that employee's responsiveness towards competitiveness low as opposed to strategic leadership. More interestingly and as expected is the fact that employee's strong inclinations, evidenced by high mean scores that refer to strategic leadership practices are exist in Iraqi manufacturing sector. But it need some steps to make it success and it is affected by the Iraqi circumstances.

Key words: Iraq, strategic leadership, critical success factor for competitiveness, manufacturing, cement, competitiveness

INTRODUCTION

The Iraqi cement manufacturing is suffering from deterioration under difficult conditions, therefore, the Iraqi Ministry of Industry has launched a campaign to save it and protect this national product and develop strategies with strategic leadership capable of saving the industry and its factories.

Iraqi national cement companies as primarily responsible for providing a competitive environment for their plants, especially in light of changing economic conditions afflicting Iraqi manufacturing companies, especially, Kufa and Najaf cement companies which belong to the Southern cement company which became rival of many imported products, especially, Iranian products that found a good opportunity in Iraqi markets. This reflected negatively on local products in addition, Iraqi cement companies have been affected during previous years by challenges and global variables including the global economic crisis. The demand for the cement fell in world markets, moreover the circumstances of the Iraqi war against terrorism for long time also affected demand in the domestic market.

Iraq has been an exporting state of this material and it is now an importer due to insufficient production in the local market plus the poor conditions for nearly 30 years. Iraq has gone through difficult circumstances starting

from the Iran-Iraq war in the 1980's (Marr, 2018). The first Gulf war and economic blockades in the 1990s and the second Gulf war in 2003 (Chatelard, 2009) and what followed it from the war against terrorism that significantly affected the manufacturing sector in Iraq including cement plants.

Despite all these challenges, Iraqi cement companies primarily Southern cement company continued to succeed in fulfilling its role towards shareholders as a profitable company as it got the quality certificate for most of its products and managed to get partially into global competition market.

All these circumstances call for the need to find successful means contribute to strengthen the competitive advantage and responsiveness and keeping pace with global changes and developments. Because of intense competition in this age, it became necessary for the Southern cement company to adopt competitive strategies and survive in an environment of complexity and the constant and rapid change which makes the attention towards building and developing competitive strategies guarantee sustainability and edging over its competitors. Perhaps among the tools that have proved their success in many countries is adopting strategic leadership (Ireland and Hitt, 1999; Haque *et al.*, 2017). That its importance stems from what awaits business organizations from future competition, building and

developing competitive advantages that achieve competitive superiority enabling the company to withstand the acute and severe competition. Consequently, it became necessary for the company to adopt strategic leadership practice which able to meet the different requirements and competitive strategies.

The performance and success of the organizations closely linked to the performance and success of the strategic leader (Lennick *et al.*, 2011). Despite the fact that leadership is not the only element that distinguishes it but it is an important element in it (Al-Alawi *et al.*, 2007) and with increasing pressures and the biggest challenges facing organizations at the moment, we need strategic leadership that have a major role in transforming the cases of degradation and descent into strong organization able to compete and succeed (Nixon *et al.*, 2012).

For that there must be a strategic leader able to transfer the vision to people in middle management and executive levels regarding the process of excellence (Jeston and Nelis, 2014) and this vision can create a climate of participation and put new and creative ideas and help to create conditions conducive to success (Hitt *et al.*, 2001). Many of the ancient and modern literature confirm that many organizations with modest and simple beginnings in terms of possibilities and limited financial resources but although it had achieved marked successes (Mauro, 2011) as other studies confirm that successes in various fields such as economic, political, security and military fields, all due to excel in strategic leadership (Hitt *et al.*, 2001).

Today, the need to a strategic leadership in working in has appeared where it seems more difficult to research (House *et al.*, 2013). Many studies have proven in this area that the effective strategic leadership leads to innovation, creativity, excellence, leadership and promoting competitiveness and it recorded many success stories (Bolman and Deal, 2017; Jeston and Nelis, 2014). Studies have shown that strategic leadership is one of the main and essential pillars to the success of organizations of all types and styles (Schoemaker *et al.*, 2013). This is largely dependent on attracting qualified efficiencies that is capable of dealing with accelerating changes in the research environment and fierce competition at every level. This requires strategic leadership to cope with changes and developments in such environment (House *et al.*, 2013).

In sum, the purpose of this study is to evaluate the level of strategic leadership and critical success factors for competitiveness in Iraqi manufacturing sector.

Literature review

Strategic leadership: Leadership is an important hub on which the activities of various organizations rely, especially, in the modern world where there is a big change in different areas of business world, especially in

the last decade of the last century when there is an increase in the need to a wise leadership with the skills and abilities that embrace the successful visions in order to keep pace with rapid change, age requirements, survival and growth (Bolman and Deal, 2017). In general, strategic leadership is defined as the ability and wisdom to make decisions about goals, strategies and tactics, through the combination of leadership and management on the one hand and strategic intent with tactical events on the other (Pisapia, 2009), this strengthens the ability to shape the future of organizations with a vision of a mechanism to achieve goals and aspirations in the fastest way (Oetinger, 2004). Accordingly, strategic leadership has become a process of organizations transition from being to the place where their leader wants to be (Eisner *et al.*, 2014). As many literatures mention strategic leadership involves creating visions, cognitive activity ,strategic thinking, information technology and change to the future (Kapferer, 2012; Pugh, 2016).

Creating vision: Having a vision and being able to communicate is clearly important aspects of leadership, it is a clear picture of what the organization will be like in the future (Leithwood and Riehl, 2003). Develop a long-term strategic vision reflects the personal views of the inspiring leadership (Hitt *et al.*, 2001). If the strategic leader can clarify his own point of view and involve his subordinates, they will support his strategic vision, this will representing privacy of the leader and at the same time representing a concept for everyone in the organization (Macmillan and Tampoe, 2000).

Cognitive activity: Cognitive thinking demands cognitive skills elucidates this as applying thorough thoughts and experiences in order to attain knowledge and understand a particular event. At any instant, the act of cognition should be very competent, allowing one to take advantage of personal experiences, accommodate the thoughts and develop a guided behavior (Sun and Hui, 2012).

Strategic thinking: Strategic thinking is closely related to both the formulation and implementation of strategies by the leader as well as to the strategic performance of the organizations. It includes strategic planning, strategic analysis, controlling and strategic leadership. Accordingly, it includes all the features that can be classified as "strategic" (Liedtka, 1998). The importance of strategic thinking is that it is a competitive tool suited to the global competitive environment and that the organization's strategies and limits are to achieve the highest profits (Baloch and Inam, 2009).

Information technology: Information technology has redefined the world economic concepts through the

introduction of digital economy (Ukwandu and Nnamocha, 2013), information technology defined as computers and related digital communication technology has the border power to reduce the cost of coordination, communication and information processing (Buckley and Casson, 2010), information technology has redefined productivity, education, social and political landscape of the world.

Change to the future: Predicting the future and change are difficult. Therefore, the organizations resort to adopting mechanisms and various means to explore the unknown future in order to cope with the changes that are expected to occur (Argote, 2011). Change to the future is new ways for organizations to change reality for the better. It is a tool of changing the situation to increase effectiveness and achieve goals through positive future change. It is done by developing a future scenario for the change process it specializes in creating a series of possible events that are likely to happen in future (Oetinger, 2004).

Critical success factors for competitiveness: A broad range of factors that can influence the success of a project has been mentioned in the literature. For example, much has been stated about project success. However, no systematic research exist in characterizing a collective set of critical success factors to implementing competitiveness in Iraqi manufacturing sector. Critical success factors can be define as areas in which results, if they are satisfactory will ensure successful competitive performance for organization (Rockart, 1979; Wong, 2005) viewed them as those critical areas of planning and actions that must be practiced in order to achieve effectiveness. In terms of competitiveness, they can be viewed as those activities and practices that should be addressed in order to ensure its successful implementation (Turner and Muller, 2005). That focuses on critical success factors a competition that adds value to the customer and all stakeholders (Trkman, 2010).

Based on insights gleaned from the study of practices and experience of leading companies in the leadership field, Iraqi Ministry of Industry report highlighted four common key success factors for competitiveness in Iraqi companies. These include manager concern for employees which is essential element for success (Alexandrov *et al.*, 2007) working environment that affect the process of decision making within the organization (Daft, 2010), appreciation and rewards that motivates employees and achieves values of loyalty and dedication the organization's objectives (Yaday, 2016) and social relationships which is positively reflected on the performance of the employees (Bain *et al.*, 2001).

MATERIAL AND METHODS

Sampling: To evaluate the level of strategic leadership practices and critical success factors for competitiveness, we select a group of engineers in two of the big Iraqi cement factories (Najaf and Kufa) as the population of this study. It included 134 respondents as a sample and it is randomly.

Instrument: For obtaining information on satisfaction, we collected primary data via. a 5-point labeled Likert type scale questionnaires which anchored from “strongly agree” to “strongly disagree”. The questions about strategic leadership were divided into five subgroups, Creating a Vision (CV), Cognitive Activity (CA), Strategic Thinking (ST), Information Technology (IT) and Creating Future (CF). Moreover, the question about critical success factors for competitiveness comprised four different groups, Manager Concern for his Employees (MCE), Working Environment (WE), Appreciation and Rewards (AR) and Social Relationship (Rel). Each subgroup contained five questions. All analysis were performed in R environment and (SPSS V.23).

In respect to measure the reliability of the instrument, the reliability coefficients cronbach's alpha shows that all variables presented values bigger that 0.7, thus, indicating a reliable questionnaire.

Demographic profile: We investigate the influence of three main demographic factors in the responses, namely, gender (men and women), age, divided into four categories (<30, 30-39, 40-49, >50 years old) and years of experience (<10, 10-19, >20 years).

A total of 134 people answered the questionnaire (90.3% male and (9.7% female and most of the employees from 40-49 years old (56.7%), followed by the group aging between 30-39 (19.4%) those younger than 30 years old (10.4%) and lastly by a group older than 50 years old (13.4%). Regarding experience, the majority of responders (55.2%) with 10-19 years of experience after that the group <10 years of experience with 29.1%, followed by the group had more than 20 years of experience with 15.7%. Table 1 describes the demographic profile of respondents.

Table 1: Demographic profile of respondents

Characteristics	Frequency (N=134)	Percentage
Gender		
Male	121	90.3
Female	13	9.7
Age		
<30	14	10.4
30-39	26	19.4
40-49	76	56.7
50-above	18	13.6
Years of experience		
<10	21	29.1
10-19 20-above	55.2	15.7

RESULTS AND DISCUSSION

Descriptive analysis for strategic leadership and critical success factors of competitiveness described in Table 2 and 3.

The first set of question corresponded to the capacity of strategic leadership create a vision in the employees (Fig. 1a). The question L1 asks how the vision reflects the expectation for a distinct system. Overall, the most likely answers are agreeing (42.4%) or strongly agreeing (34.6%). We also observed the same for the question L2 which evaluate how the vision and goals are essential to the personal goals of an employee (47.6% agree and 31.6% strongly agree). The question L3 ask if the vision articulates what the respondent aspire to be a working system, here, again there a high probability of they agree (43.1%) or strongly agree (26.4%) with this

shared view. Question L4 review that the majority of respondents are likely to agree (40.5%) or strongly agree (32.7%). Finally, the last question on the vision (L5) asked whether or not the employees participate in crafting the vision of their organization and the results showed that the majority actively participate on that process, since, the majority strongly agree (40.5%) or strongly agree (38.9%).

The next set of question evaluates the level of cognitive activity to question whether or not the respondents agree with how the knowledge is transmitted, shared or required within the organization (Fig. 1b). Question L6 ask if the organization provide a high level of knowledge and the results showed that the employees are highly likely to agree with that provisioning (48.7% probability of agreeing and 31.0 of strongly agreeing). Question L7 also showed a similar pattern when asking for the participation on a team that allows knowledge exchange for our

Table 2: Descriptive analysis for strategic leadership practices

Items	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13
Mean	4.07	4.06	3.87	3.96	4.14	4.02	3.88	3.81	3.62	4.05	3.97	3.73	3.89
SE	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.05
SD	0.86	0.82	0.91	0.95	0.84	0.83	0.77	0.89	0.92	0.76	0.83	0.85	0.89
1%	0.70	-	0.40	-	-	0.40	0.40	0.70	0.40	0.40	0.40	0.40	0.40
2%	3.30	5.20	8.20	10.40	3.70	4.80	4.80	7.80	13.80	3.00	4.80	7.10	6.70
3%	19.00	15.60	21.90	16.40	18.20	16.70	19.30	23.40	24.20	15.20	19.30	29.40	23.00
4%	42.40	47.60	43.10	40.50	40.50	48.70	58.50	46.10	46.50	54.30	48.70	45.40	43.50
5%	34.57	31.60	26.39	32.71	38.90	31.00	17.84	21.93	15.24	27.10	26.77	17.84	26.39
SKEW	0.74	0.67	0.52	0.64	0.65	0.70	0.64	0.53	0.39	0.70	0.61	0.31	0.49
KUR	0.35	0.02	0.34	0.49	0.38	0.34	0.72	0.07	0.52	0.84	0.18	0.27	0.28
Item	L14	L15	L16	L17	L18	L19	L20	L21	L22	L23	L24	L25	-
Mean	3.51	3.95	3.56	3.30	3.39	3.56	3.60	3.34	3.30	3.21	3.01	2.86	-
SE	0.06	0.05	0.07	0.06	0.06	0.07	0.07	0.07	0.05	0.07	0.07	0.07	-
SD	0.92	0.81	1.08	1.02	1.05	1.08	1.07	1.07	0.87	1.20	1.18	1.23	-
1%	0.40	-	7.10	5.90	6.70	6.70	5.90	1.10	1.50	8.60	10.40	15.20	-
2%	16.00	2.60	7.80	14.10	9.70	9.30	10.00	27.10	14.10	24.20	26.00	25.30	-
3%	29.40	27.50	23.40	33.50	34.90	21.90	19.00	23.80	45.40	19.00	27.90	30.10	-
4%	41.30	42.00	45.40	36.80	35.30	45.70	48.30	32.30	30.50	34.20	23.80	17.10	-
5%	13.00	27.90	16.40	9.70	13.40	16.40	16.70	15.60	8.60	14.10	11.90	12.30	-
SKEW	0.20	0.21	0.81	0.42	0.49	0.79	0.84	0.00	0.02	0.22	0.04	0.18	-
KUR	0.70	0.81	0.22	0.26	0.09	0.11	0.18	1.13	0.15	1.01	0.89	0.86	-

Table 3: Descriptive analysis for critical success factors for competitiveness

Items	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Mean	4.07	4.19	3.96	4.00	4.03	3.93	3.74	3.68	4.01	3.93
SE	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05
SD	0.73	0.60	0.73	0.74	0.79	0.78	0.81	0.82	0.82	0.85
1%	0.40	0.40	0.40	0.00	1.10	0.40	1.10	0.00	0.70	2.20
2%	1.50	-	2.60	4.50	2.20	4.80	5.20	8.60	4.50	3.00
3	16.70	7.80	18.20	14.10	16.40	17.10	26.40	28.30	15.20	17.50
4%	53.53	63.90	57.99	58.36	53.16	56.88	53.16	49.40	51.67	54.30
5%	27.88	27.90	20.82	23.05	27.14	20.82	14.13	13.80	27.88	23.00
SKEW	0.57	0.52	0.59	0.66	0.87	0.70	0.62	0.32	0.83	1.04
KUR	0.65	2.34	0.90	0.62	1.53	0.76	0.73	0.32	0.92	1.86
Item	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Mean	3.97	4.10	3.90	4.06	4.10	3.99	3.91	3.80	3.67	4.01
SE	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
SD	0.89	0.90	0.88	0.89	0.88	0.85	0.84	0.87	0.86	0.81
1%	0.70	0.70	1.10	-	-	0.70	0.70	1.10	0.40	0.70
2%	6.30	6.30	5.90	5.90	4.80	6.70	6.70	6.30	10.00	4.10
3%	17.50	11.90	19.70	18.20	19.30	11.50	15.60	23.80	27.10	15.60
4%	46.10	43.90	48.00	39.40	36.80	54.60	54.60	48.70	47.60	52.40
5%	29.40	37.20	25.30	36.40	39.00	26.40	22.30	20.10	14.90	27.10
SKEW	0.74	0.99	0.72	0.65	0.63	0.95	0.81	0.60	0.38	0.82
KUR	0.28	0.70	0.44	0.37	0.47	1.04	0.74	0.32	0.28	0.98

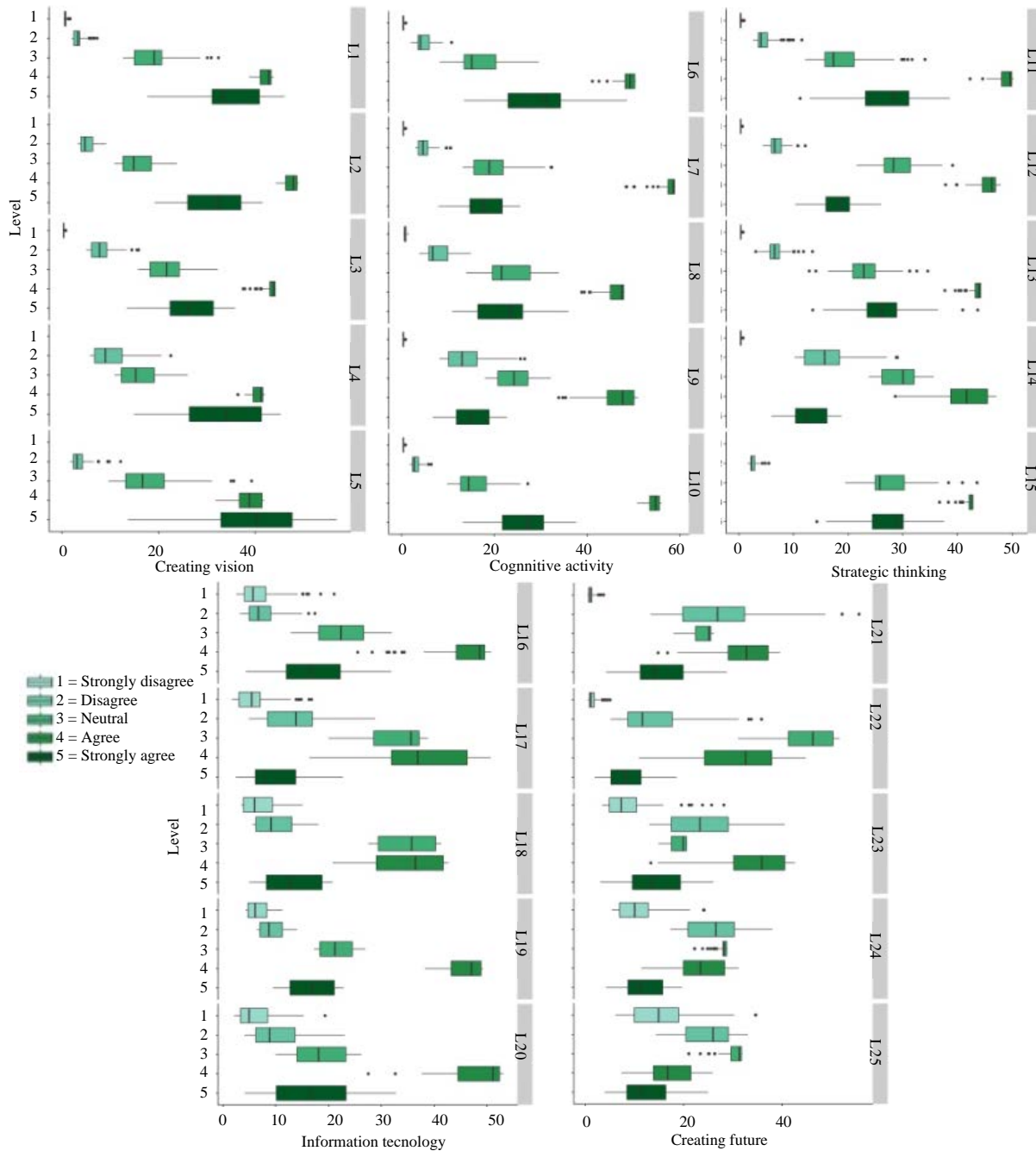


Fig. 1: Probability, according to a logistic regression model of respondents choosing one of the five possible answers for each one of the statements regarding the strategic leadership practices

sampling group it is highly likely that they participate in such a group (58.5% agree and 17.8% strongly agree). Question L8 reveal how the employees see the leadership as an encourager of sharing knowledge, according to the interviewed answers this is highly likely to occur with a total probability of 68% of respondents agreeing or

strongly agreeing with that. This is important, since, the question L9 ask whether or not the respondents possess enough knowledge to perform their research our results show that although, it is highly likely the employee agree with possessing enough knowledge (46.5%), the second most common result is a neutral answer (24.2%).

Then, the last question for this subgroup evaluate the working environment regarding the group interest in self-development here again we observe a high probability of a positive answer (54.3% agree and 27.1% strongly agree), thus, demonstrating the common interest in developing cognitive skills.

The questionnaire also evaluated the perceptions on strategic thinking with five questions (Fig. 1c). Question L11 evaluated how likely a respondent is in not spending time rethinking things that are done, according to our results, the likelihood of the employee wasting time is smaller than 4.8% (disagreeing and strongly disagreeing). However when questioned about re-evaluating the past activities (L12) we found an increasing likelihood of our respondents doing that with 46.2% chance of agreeing and 18.6% chance of strongly agreeing. Interestingly, it is also high the likelihood of the employees to get rid of unwanted thoughts as suggests by the answers given in the question L13 (43.5% agreeing, 26.4% strongly agreeing). In turn, question L14 investigated how much time usually spent thinking back over embarrassing or disappointing moments and our results showed that it is highly likely that the employees spend time on that (41.3% probability of agreeing). However, the second most probable answer is a neutral option (29.4%) and disagree presented only 16% chance of occurring. Our results also, showed that if the employee rethinks in a situation it usually will do that from another point of view (L15) with a 42.0% of agreeing with this approach and 27.9% chance of strongly agreeing.

The next set of questions covered questions regarding information technology (Fig. 1d). The question L16 showed that the majority of respondents agree (45.4%) or strongly agree (16.4%) that in their organization technology is widely used in developing new systems but we also observed a high chance of neutral responses (23.4%). Next question, L17 investigated how the organization is concerned about the quality of technology and the results showed that both agree and neutral responses presented nearly the same probability (36.8% and 33.5%, respectively). The same phenomena occurred when we asked whether or not the organization possess appropriate information technology (L18) with 35.3% chance of agreeing and 34.9% chance of a neutral response. However, there is a slight change when in L19 we inquire the respondents about the experience of employees in information technology, the results reveal that the majority of respondents agree (45.7%) or strongly agree (16.4%) that they have enough experience on information technology. We also observed a major consensus that the organization possesses the information technology that allows strategic changes (L20). In total, there is a 67.3% chance of employees agreeing or strongly agreeing with that statement.

The last set of question regarding leadership evaluates the organization capacity showing a long-term perspective (creating future, Fig. 1e). According to the responses to the question L21, there is no consensus of employees in a long-term future. The chance of agreeing is 32.3%, followed by disagreeing with 27.1 and a 23.8% chance of a neutral answer. Next question L22 investigates how the perception regarding the achievement of future goals here the respondents possess a more neutral opinion (45.4%) followed by a chance of agreeing (30.5%) that all the employees are confident about the achievement of future goals. When we investigate whether or not the respondent agrees that their working environment helps to create strategic opportunities (L23), the results showed divided opinions with a 34.2% chance of agreeing followed by 24.2% chance of disagreeing with this statement. Opinions also remained divided when we evaluate how the respondents agree that the employees offer new ideas (L24), the changes of a neutral answer (27.9%) slightly surpass the chances of disagreeing (26%) and agreeing (23.8). Lastly when we checked the degree of confidence, the respondents possess in the strategic objectives of the organization we observe a high level of discordance. Although, the most was a neutral opinion (31.1%) the disagree and strongly disagree options together presented higher values (25.3 and 15.2%, respectively) than the counterparts (agreeing with 17.1% and strongly agreeing with 12.3%).

The second part of the questionnaire allows us to obtain the perception regarding the critical success factors. In that, the first set of five questions evaluated the manager concern for employees (Fig. 2a). The first question (S1), the respondents, inform whether or not they believe to be necessary for the research team. Our results showed that they are highly likely to perceive themselves as an essential part of the research team (53.5 agreed and 27.9% strongly agreed). The question S2 evaluated how the respondents consider that their jobs allow them to use a variety of skills and abilities here again, positive answer occurred more often (63.9% chance of agreeing and 27.9% of strongly agreeing). Moreover when the respondents faced a question regarding the recognition of their accomplishments (S3) our results showed that this is highly likely to occur (78.8% chance of agreeing or strongly agreeing in receiving some recognition). Interestingly, the results also showed a high chance of the respondents research with someone who encourages their development (S4) with an overall 58.4% probability of agreeing and 23.0% chance of strongly agreeing with that statement. Then, question S5 evaluated how the leadership followed its commitments. According to the respondent's opinion this is a common phenomenon, since, it is highly likely that they either agree (53.2%) or strongly agree (27.1%).

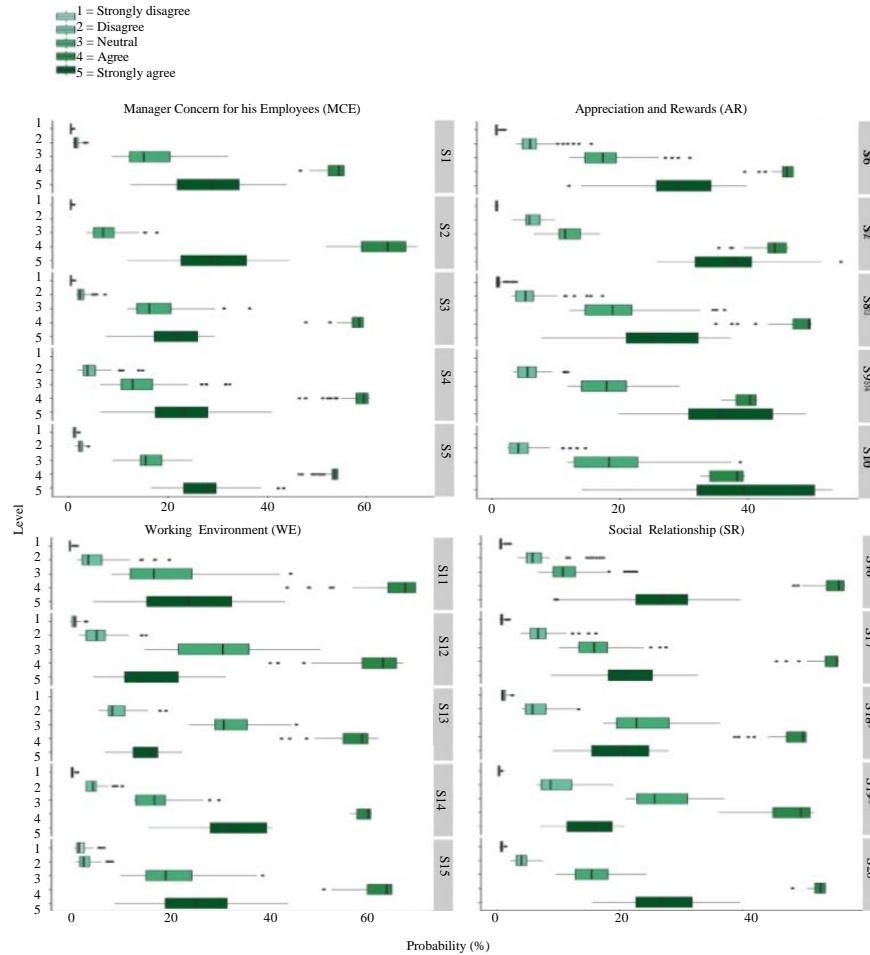


Fig. 2: Probability, according to a logistic regression model of respondents choosing one of the 5 possible answers for each one of the statements regarding the critical success factors of competitiveness

The next set of question allows the respondents to evaluate their working environment (Fig. 2b). First, the question S6 checks whether or not the working environment is open and accepts individual differences. findings reveal that the majority of respondents are highly likely to agree with that statement (56.9% agreeing and 20.1% for strongly agreeing). However when the question S7 inquires about the tools and resources available for a good performance, results showed that despite the high chance of the respondents agree they have the proper tools and resources (53.2%) the chance of a neutral response (26.4%) is 2.1 times bigger than the probability of strongly agreeing with that (14.1%). The results also reveal the same trend for S8 that deals with the balance work and personal life, so, although, the results showed a high chance of the respondents find good ways to mention the research/life balance (49.4% of agreeing and 13.8% of strongly agreeing) we observe a 28.3% chance

of a neutral answer. We also evaluated the respondent's perception regarding the decision-making. Specifically, we assessed whether or not the decisions depend on environmental analysis (S9). In general, this seems to occur highly often, since, both positive answers presented a high likelihood (52.7% chance of agreeing and 27.9% of strongly agreeing). Lastly, S10 evaluated if the employee has the opportunities to learn and grow in the organization which seems to occur highly often, since, there is a high likelihood of a respondent agree (54.3%) or strongly agreeing (23%).

We also evaluated the organization success by investigating the employee's perceptions of appreciation and rewards (Fig. 2c). The first question (S11) evaluates the respondent's perception in whether or not they receive a deserved recognition for their research. We observed a high chance of occurring a recognition, since, a total probability of 75.5% of a respondent agrees or

strongly agree with that happening. We also noted a high likelihood of they received satisfying benefits for good research (S12) with a total 81.1% (agreeing or strongly agreeing).

Interestingly, the vast majority of respondents seem to agree (48%) or strongly agree (25.3%) that their promotion depends on their effectiveness (S13). This combined with statement S14 that evaluates whether or not the respondents received adequate promotion and opportunities (39.4% chance of agreeing and 36.4% of strongly agreeing) suggests that rewards occurred following employee's expectations. This becomes particularly important when we observe that according to the answers for question S15 there is a high probability (75.8%) of a respondent receive a verbal or written recognition for their research.

Finally, our last set of questions evaluated the aspects of social relationships (Fig. 2d). Statement S16 evaluated whether or not the respondents know someone to ask for any suggestion. results showed that the chances of that occurring are high with a total of 81%. Another relevant point of our investigation is if the respondent perceives himself as an equal among others (S17). According to our responses, they indeed perceive equality among each other (54.6% agreeing and 22.3% strongly agreeing). Moreover, in case of need for advice, the responded revealed a high chance of knowing someone (48.7% agreeing and 20.1% strongly agreeing). However, when evaluated their perception regarding someone take the pride from their accomplishments, we observed that a neutral response occurred more often (27.1%) than a strongly agree (14.9%). However, the principal outcome remains the agree (47.6%). The respondents also revealed that they are highly likely to quickly become close to their colleagues with a total of 79.5% chance of agreeing or strongly agreeing with the statement S20.

The greatest GDP producer in Iraq, since, its independence has been oil-derived products from the oil sources, Iraq can generate sulphur, natural gas as well as phosphates (OPEC, 2018). This is so devastating to the economy as well as governance of the state. In essence, there is a cognate link between a country depending on one export (especially, natural resources) to both internal and external conflict, abuse of democracy and corruption.

For a better future, the economic stability should focus on manufacturing (Sebastian and Muhammed, 2017). In the entire governance, the leaders have a degree of influence to be drivers of change of such magnitude (Goleman *et al.*, 2013). This is possible when the government itself creates a link between strategic leadership and competitiveness. The nature of the

product and consumer satisfaction in order to cause diversity in the Iraqi manufacturing sector (Chang *et al.*, 2016).

This study has evaluated the level of strategic leadership practices as (creating vision, cognitive activity, strategic thinking, information technology, change to the future) and critical success factor for competitiveness as (manager concern for employees, working environment, appreciation and rewards, social relationships).

As it was discovered by the findings that strategic leadership had the capacity to create a vision and the vision reflects employees expectations and it is essential to the goals.

The results illustrate that employees needs more knowledge to improve their ability and skills to perform research, so, the leadership has to encourage individuals to share their knowledge with each other. Moreover, the employees are wasting time and not exploit it in strategic thinking. The results also show that although employees are highly efficient in dealing with technology, there is a gap in the use of modern and high quality technology in the implementation of business. That emphasizes the need to implement strategic change in technology.

The company has no consensus of employees in a long-term future and confident about the achievement of future goals. The formulation of new ideas is limited in light of the difficult conditions experienced by various manufacturing sectors in the country. According by Boga and Ensari (2009), leadership strongly defines the concept of success in organizations, playing a major and pivotal role in determining the key factors and requirements for future success. Our results showed that in general that strategic leadership practices are exist in Iraqi manufacturing companies. But it need some steps to make it success and it is affected by the Iraqi circumstances.

The results shows that the strategy requires attention to the sharing of knowledge for workers as this contributes to the improvement of their competence in carrying out the tasks entrusted to them. Several studies have confirmed that knowledge sharing contributes to improving the success and innovation of the organization (Ritala *et al.*, 2015), it is also reflected in the social relationship, especially with regard to shared skills and knowledge within the working group (Widen-Wulff and Gimman, 2004). On the social relationship, especially with regard to shared skills and knowledge within the working group.

The results showed that there is a wasted time in this company and the employees does not care much about the time, especially for strategic thinking that not widely adopted within the company. The importance of strategic thinking is great for improving leadership abilities and

achieving competitiveness (Kapferer, 2012). Furthermore, there is a need for the other factors, first the adoption of modern technology in the company, especially, production technology, secondly, leaders need to stimulate organizational components through an effort-based reward system and working environment, thirdly, leadership must ensure effective resource allocation (information technology and skills exchange) to enhance competitiveness.

CONCLUSION

The objective of this study is to study strategic leadership practices and critical success factors for competitiveness in the Iraqi manufacturing sector. The present study supports previous studies that linked transformational leadership and perceived organizational success (Lowe *et al.*, 1996). We expand previous research by demonstrating the relationship between different subjects in strategic leadership and competitiveness (Ireland and Hitt, 1999; Haque *et al.*, 2017).

LIMITATIONS

Despite its advantages our current study faces some limitations. First, the size and location of the sample reduces the generalization of the results. Secondly, the sample should include a larger number of staff from different organizations.

RECOMMENDATIONS

Moreover, since, concepts create facts (Jussim, 1991), we emphasize that in order to ensure a clear understanding of the vision and purpose of the company, leadership must establish a reasonable and equitable system of benefits and rewards. Once an employee understands that looking for the company's vision and purpose will bring benefits, the company intensifies the path towards competitiveness. In addition, the evaluation of performance perceptions must become a focal point for ensuring that everyone has an accurate assessment of their efforts.

REFERENCES

- Al-Alawi, A.I., N.Y. Al-Marzooqi and Y.F. Mohammed, 2007. Organizational culture and knowledge sharing: Critical success factors. *J. Knowledge Manage.*, 11: 22-42.
- Alexandrov, A., E. Babakus and U. Yavas, 2007. The effects of perceived management concern for frontline employees and customers on turnover intentions. *J. Serv. Res.*, 9: 356-371.
- Argote, L., 2011. Organizational learning research: Past, present and future. *Manage. Learn.*, 42: 439-446.
- Bain, P.G., L. Mann and A. Pirola-Merlo, 2001. The innovation imperative: The relationships between team climate, innovation and performance in research and development teams. *Small Group Res.*, 32: 55-73.
- Baloch, Q.B. and M. Inam, 2007. Strategic thinking: Catalyst to competitive advantage. *J. Managerial Sci.*, 1: 1-26.
- Boga, I. And N. Ensari, 2009. The role of transformational leadership and organizational change on perceived organizational success. *Psychologist Manager J.*, 12: 235-251.
- Bolman, L.G. and T.E. Deal, 2017. *Reframing Organizations: Artistry, Choice and Leadership*. 6th Edn., Jossey-Bass, San Francisco, California, USA., ISBN:9781119281818, Pages: 494.
- Buckley, P.J. and M. Casson, 2010. Models of the Multinational Enterprise. In: *The Multinational Enterprise Revisited*, Buckley, P.J. and M. Casson (Eds.). Palgrave Macmillan, Basingstoke, England, UK., ISBN:978-1-349-35424-5, pp: 147-176.
- Chang, H.J., M. O'Boyle, R.C. Anderson and C. Suttikun, 2016. An fMRI study of advertising appeals and their relationship to product attractiveness and buying intentions. *J. Consum. Behav.*, 15: 538-548.
- Chatelard, G., 2009. Migration from Iraq between the Gulf and the Iraq wars (1990-2003): Historical and sociospatial dimensions. *Compas*, Oxford, England, UK. <https://halshs.archives-ouvertes.fr/halshs-00364961/document>
- Daft, R.L., 2010. *Organization Theory and Design*. 10th Edn., West Publishing, Co., San Francisco, CA., USA.
- Eisner, A., G. McNamara, G. T. Lumpkin and G. Dess, 2013. *Strategic Management, Creating Competitive Advantages*. 7th Edn., McGraw-Hill Education, New York, USA., ISBN:9780077636081, Pages: 544.
- Goleman, D., R. Boyatzis and A. McKee, 2013. *Primal Leadership: Unleashing the Power of Emotional Intelligence*. Harvard Business Publishing, Boston, Massachusetts, USA., ISBN:978-1-4221-6803-5, Pages: 306.
- Haque, A.U., R. Faizan and A.Cockrill, 2017. The relationship between female representation at strategic level and firm's competitiveness: Evidences from cargo logistic firms of Pakistan and Canada. *Pol. J. Manage. Stud.*, 15: 69-81.
- Hitt, M.A., R.D. Ireland and R.E. Hoskinsson, 2001. *Strategic Management: Competitiveness and Globalization*. South-Western College Publishing, USA., Pages: 549.

- House, R.J., P.W. Dorfman, M. Javidan, P.J. Hanges and M.F.S. De Luque, 2013. *Strategic Leadership Across Cultures: The GLOBE Study of CEO Leadership Behavior and Effectiveness in 24 Countries*. SAGE Publications, Thousand Oaks, California, USA., ISBN:978-1-4129-9594-8, Pages: 430.
- Ireland, R.D. and M.A. Hitt, 1999. Achieving and maintaining strategic competitiveness in the 21st Century: The role of strategic leadership. *Acad. Manage. Perspect.*, 13: 43-57.
- Jeston, J. and J. Nelis, 2014. *Business Process Management: Practical Guidelines to Successful Implementations*. 3rd Edn., Routledge, Abingdon, England, UK., ISBN:978-0-415-64176-0, Pages: 652.
- Jussim, L., 1991. Social perception and social reality: A reflection-construction model. *Psychol. Rev.*, 98: 54-73.
- Kapferer, J.N., 2012. *The New Strategic Brand Management: Advanced Insights and Strategic Thinking*. 5th Edn., Kogan Page, London, England, UK., ISBN:978-0-7494-6515-5, Pages: 493.
- Leithwood, K. and C. Riehl, 2003. *What We Know About School Leadership*. National College for Teaching and Leadership, Nottingham, England, UK.,
- Lennick, D., F. Kiel and K. Jordan, 2011. *Moral Intelligence 2.0: Enhancing Business Performance and Leadership Success in Turbulent Times*. Pearson Education, USA.,
- Liedtka, J.M., 1998. Strategic thinking: Can it be taught?. *Long Range Plan.*, 31: 120-129.
- Lowe, K.B., K.G. Kroeck and N. Sivasubramaniam, 1996. Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *Leadersh. Quart.*, 7: 385-425.
- Macmillan, H. and M. Tampoe, 2000. *Strategic Management: Process, Content and Implementation*. Oxford University Press, Oxford, England, UK., ISBN:9780198782292, Pages: 378.
- Marr, P., 2018. *The Modern History of Iraq*. 4th Edn., Taylor & Francis, Abingdon, England, UK., ISBN:9780429974069, Pages: 480.
- Mauro, P., 2011. *Chipping Away at Public Debt: Sources of Failure and Keys to Success in Fiscal Adjustment*. John Wiley & Sons, Hoboken, New Jersey, USA., ISBN:978-1-118-04338-7, Pages: 279.
- Nixon, P., M. Harrington and D. Parker, 2012. Leadership performance is significant to project success or failure: A critical analysis. *Intl. J. Productivity Perform. Manage.*, 61: 204-216.
- Oettinger, B.V., 2004. A plea for uncertainty: Everybody complains about uncertainty, but it might be a good thing to have. *J. Bus. Strategy*, 25: 57-59.
- Pisapia, J.R., 2009. *The Strategic Leader: New Tactics for a Globalizing World*. Information Age Publishing Inc. (IAP), Charlotte, North Carolina, USA., ISBN:978-1-60752-152-5, Pages: 225.
- Pugh, L., 2016. *Change Management in Information Services*. 2nd Edn., Routledge, Abingdon, England, UK., ISBN:978-0-7546-4665-5, Pages: 229.
- Ritala, P., H. Olander, S. Michailova and K. Husted, 2015. Knowledge sharing, knowledge leaking and relative innovation performance: An empirical study. *Technovation*, 35: 22-31.
- Rockart, J.F., 1979. Chief executives define their own data needs. *Harv. Bus. Rev.*, 57: 81-93.
- Schoemaker, P.J., S. Krupp and S. Howland, 2013. *Strategic leadership: The essential skills*. *Harv. Bus. Rev.*, 91: 131-134.
- Sebastian, B. and M.A. Yildirim, 2017. *Iraq's manufacturing sector: A challenging path to more products*. Master Thesis, The Lebanese Center for Policy Studies, Beirut, Lebanon.
- Sun, R.C.F. and E.K.P. Hui, 2012. Cognitive competence as a positive youth development construct: A conceptual review. *Scient. World J.* 10.1100/2012/210953
- Trkman, P., 2010. The critical success factors of business process management. *Int. J. Inform. Manage.*, 30: 125-134.
- Turner, J.R. and R. Muller, 2005. The project manager's leadership style as a success factor on projects: A literature review. *Project Manage. J.*, 36: 49-61.
- Ukwandu, E. and P. Nnamocha, 2013. The effects of information technology on global economy. *Soc. Sci.*, 8: 606-609.
- Widen-Wulff, G. and M. Ginman, 2004. Explaining knowledge sharing in organizations through the dimensions of social capital. *J. Inf. Sci.*, 30: 448-458.
- Wong, K.Y., 2005. Critical success factors for implementing knowledge management in small and medium enterprises. *Ind. Manage. Data Syst.*, 105: 261-279.
- Yadav, A., 2016. *A study of factors affecting employee satisfaction and measures to attract retain and motivate employees in organized retail sector with special reference to Rajasthan*. Ph.D Thesis, Banasthali Vidyapith, Vanasthali, India.