

3rd GCARSET 2013 Global Conference for Academic Research on Scientific and Emerging Technologies

# GCARSET PROCEEDING 09-10 MARCH 2013

BEST WESTERN PREMIER HOTEL KUALA LUMPUR, MALAYSIA

# International Journal of Scientific & Engineering Research (ISSN 2229-5518)

Conf. Chair: Prof. Raouf F. Conf Organizer: Dr. Hosam A.S. Welcome Message from Prof. Raouf GCARSET Chair

The GCARSET conference serves as an interdisciplinary venue for inspiring new ideas, presenting cutting-edge studies and encouraging collaborations between scholars in the area of emerging technologies. It brings together researchers from various disciplines and areas like: computer studies, educational studies, electronic engineering, and more.

We are proud that we have been given the opportunity to host GCARSET 2013 in Malaysia. Malaysia is a home to many business and management companies, and is a hub for communities and academic institutions that study business and economics. As such, it can serve as a great location for a conference like this.

On a personal note, organizing a conference is not exactly what we are trained for. It requires the help of many volunteers that contributed many hours to make it successful. A special thanks to Dr. Hosam A.S., who helped with the organizing of the conference.

So, it is time to enjoy our conference. My hope is that the conference will serve as a locus for interdisciplinary, a space for discourse and collaboration.

I look forward to seeing you in our next events.



## LIST OF PAPERS

- Paper ID 501: -nucleus Interactions at 200 GeV/c via Neural Network Technique
- Paper ID 502: A Genetic programming for modeling Hadron-nucleus Interactions at 200 GeV/c
- Paper ID **510**: Analysis optimum transmit power of 10 Gbits optical CDMA system in fiberto-the-home access network
- Paper ID **513**: Commodity Price Effect on Stock Market: A Markov Switching Vector Autoregressive Approach
- Paper ID **514:** Fitting Finite Mixture Model to Exchange Rate Using Maximum Likelihood Estimation
- Paper ID 515: Variable Twist Angle of Flexible Electromagnetic Hyper Redundant Robot
- Paper ID **520**: Experimental Study in the Process Parameters in Laser Percussion Drilling
- Paper ID **524**: Digital Prototyping: A Case Study on its Viability in Enhancing Small and Medium Ceramic Industries
- Paper ID **527**: Variable Twist Angle of Flexible Electromagnetic Hyper Redundant Robot
- Paper ID **528**: Application of Binary Particle Swarm Optimization in Automatic Classification of Wood Species using Gray Level Co-Occurence Matrix and K-Nearest Neighbor
- Paper ID **529**: A Comparative Study of the Application of Swarm Intelligence in Kruppa-Based Camera Auto-Calibration
- Paper ID **531**: A SmartInk: Mediator Solution of Technology Learning Dilemma for Digital Note application
- Paper ID **532**: Transverse product effect on CSEM with double Hydrocarbon reservoir in seabed logging
- Paper ID **533**: Evaluation of Electric Field Components Response for Offshore Hydrocarbon Detection
- Paper ID **535**: Preliminary Overview of Potential Recommendations and Future Directions in Pursuit towards the Progression of Nanotechnology in Malaysia

- Paper ID **536**: Investigating Deep Target Reservoirs Using CSEM Methods
- Paper ID **537**: High efficient Fibre Plant Utilization by Multiple PON infrastructure based on Frequency Re-use Approach for Scalable FTTH Networks
- Paper ID **538**: Animation as a Problem Solving Technique in Mechanical Engineering Education
- Paper ID **541**: GaD-eM: An Adaptive Game Design Model for Malaysian Higher Education (HE)
- Paper ID **542**: Dynamic Self-Adaptive Music-Inspired Optimization Algorithm for the Hippocampus Localization in Histological Images: A Preliminary Study
- Paper ID 543: Monitoring, Tracking and Quantification of Quality of Service in Cloud Computing
- Paper ID **546:** A Comprehensive Survey on Quality of Service Implementations in Cloud Computing
- Paper ID **548**: Shared Services in Collaborative System Configurations
- Paper ID **549:** Case Study: Using Spreadsheet to Deploy Chained Business Rules
- Paper ID **553**: Determination of Mathematical Model and Torque Estimation of sEMG Signals based on Genetic Algorithm
- Paper ID **556**: Adoption of Iris Recognition System in Financial Sectors: A Case of Saudi Arabia Point of View
- Paper ID **557**: Mobile learning in the sphere of physical problems solving using the semantic structures method
- Paper ID **559**: The effect of the seeding power on RSOA-based colorless ONU performance in FTTH-PON
- Paper ID 560: The Transformation of Saudi Arabia through Tablet-Based Technology
- Paper ID 561: Towards The Use of Advance Technology in E-Learning: A Review
- Paper ID **563**: Synthesis and Characterization of Silver Oxide Nanoparticles by a Novel Method
- Paper ID 566: CMOS based Thermal Energy Generator For Low Power Devices

- Paper ID 567: Before and After: User's Knowledge Maturity within Personal Information Management
- Paper ID **568**: Overhead Analysis as One Factor Scalibility of Private Cloud Computing for IAAS Service
- Paper ID **571**: The Impact of Integrating HCI in Software Processing Models
- Paper ID **572:** Comparative Analysis of Ecological Footprint of Two Different Neighbourhoods in Minna
- Paper ID **573**: Impact of public space utilization frequencies on tacit knowledge sharing
- Paper ID **575**: E-Learning Acceptance among Tertiary Education Students



# P(P<sup>-</sup>)-nucleus Interactions at 200 GeV/c via Neural Network Technique

\*Mahmoud Y. El-Bakry Tabuk University, Faculty of Science, Department of Physics, Tabuk, KSA. \*\*A. Radi, \*\*\*El-Sayed A. El-Dahshan Ain Shams University, Faculty of Science, Department of Physics, Abbassia, Cairo, Egypt. \*\*The British University in Egypt (BUE). \*\*\*Egyptian E-Learning University- 33 El-mesah St., El-Dokki- Giza- Postal code 12611. M. Tantawy \*Ain Shams University, Faculty of Education, Department of Physics, Roxi, Cairo, Egypt. \*Moaaz A. Moussa Buraydah Colleges, Al-Qassim, Buraydah, King Abdulazziz Road, East Qassim University, P.O.Box 31717, KSA. moaaz2030@Yahoo.com

**Abstract**— $P(P^-)$ -nucleus interactions at 200 GeV/c have been studied. Two cases are considered; the charged pions multiplicity distribution and the negative ones. The neural net work (NN) technique has been adopted to study the same two cases, the trained NN shows a better fitting with experimental data than the PTFM calculations do. The NN simulation results are satisfactory and prove a vital and strong presence in modeling  $P(P^-)$ -nucleus interactions at 200 GeV/c. From paper to paper; we prove that the NN technique is better than the old conventional ones.

\_\_\_\_\_ 🔶 \_\_\_\_\_

Index Terms— Artificial intelligence (AI), machine learning (ML), NN technique, multiplicity distribution, (h-A) interactions.

### **1** INTRODUCTION

Experimental data on hadron-nucleus (h-A) interactions at high energies are required for understanding high energy interactions. They provide a useful link between hadron

interactions. They provide a useful link between hadronhadron (h-h) interactions and the complex phenomena of nucleus-nucleus (A-A) interactions. These types of interactions investigate space time picture and highlight on phenomenon which doesn't exist in (h-h) such as cascade, multi-collisions, gray particles, etc. There are various models for (h-A) interaction like collective tube model [1], diffractive excitation model[2], energy flux cascade model [3], quark model [4], interanuclear cascade model [5] multiple scattering model [6], hydrodynamical model [7], and many others.

From view point of parton two fireball model (PTFM), nucleons are treated as composite objects of loosely bound states of the spatially separated constituents (quarks) which in turn are composed of point-like particles (partons) [8]. This may allow one to consider the nucleons as consisting of a fixed number of partons. This nucleon structure has been used in different models [8-10] along with other assumptions to describe h-A interactions. PTFM, which is proposed by Hagedorn [11] and developed by Tantawy [12], has been used to explain the high energy interactions of hadrons and nuclei [12-18]. All these studies showed qualitative predictions of the measured parameters [19-23].

Analogous to the theoretical approach based on different views, development in the artificial intelligence (AI) field has given the neural networks a strong presence in high-energy physics [24-27]. Neural networks are composed of simple interconnected computational elements operating in parallel. These artificial neural networks (ANNs) are trained, so that a particular input leads to a specific target output. The objective of this paper is to extract the multiplicity distribution of charged pions for h-A collisions at 200 GeV/c using NNM compared to PTFM. Section 2 presents parton two fireball model PTFM at high energies for the multiparticle production in hadron-nucleus h-A collisions. The NN model is described in Sections 3, 4. The results and discussion of both models are compared in Section 5.

### 2 PARTON TWO FIREBALL MODEL (PTFM)

Multiplicity of the created charged particles and other parameters in h-A interactions can be determined only by the overlapping volume participating in the interaction at a given impact parameter [12, 14, 17, 18].

Let us assume that a proton with mass m and radius  $r_0$  is incident on a nucleus of radius R. The overlapping volume at any impact parameter, V (b), is given by,

$$V(b) = \pi (r_0 + R - b) \left[ \frac{2}{3} r_0^2 - \frac{1}{3} r_0 (b - R) - \frac{1}{3} r_0 (b - R)^2 \right]$$
(1)

If we define a dimensionless impact parameter  $x = \frac{b}{(r_0 + R)}$ ,

then the fraction of partons from the projectile that participate in the interaction at a given impact parameter, Z(x), can be given by,

$$Z(\mathbf{x}) = \left(\frac{1}{2} + \frac{3}{4}\mathbf{A}^{1/3} - \frac{1}{4}\mathbf{A}\right) + \frac{3}{4}\left(\mathbf{A}^{2/3} + \mathbf{A} - \mathbf{A}^{1/3} - 1\right)\mathbf{x}$$
$$-\frac{3}{4}\left(\mathbf{A}^{1/3} + \mathbf{A} + 2\mathbf{A}^{2/3}\right)\mathbf{x}^{2} + \frac{3}{4}\left(\frac{1}{3} + \frac{1}{3}\mathbf{A} + \mathbf{A}^{2/3} + \mathbf{A}^{1/3}\right)\mathbf{x}^{3}$$
(2)

and the statistical impact parameter distribution is given by

$$\mathbf{P}(\mathbf{x}) \, \mathbf{d}\mathbf{x} = 2 \, \mathbf{x} \, \mathbf{d}\mathbf{x} \tag{3}$$

The total probability of peripheral collisions  $P_{\rm Pr.}$  will be given by,

$$P_{Pr.}(x) = \int_{a}^{1} 2 x \, dx$$
 (4)

Where,  $a = \frac{(A^{1/3} - 1)}{(A^{1/3} + 1)}$ 

The total probability of central collisions will be given by

$$P_{Cent.} = \frac{(A^{\frac{1}{3}} - r_0)^2}{(A^{\frac{1}{3}} + r_0)^2}$$
(5)

From equations (2) and (3) using least square fitting technique Z function distribution can be written in the form [14, 17],

$$P(z) dz = 2 \sum_{k=0}^{k=9} C_k Z^k \sum_{k=0}^{k=8} k C_k Z^{k-1} dZ \qquad (6)$$

Where,  $C_k$  values are represented in table (1),

### 2.1 CHARGED PION PRODUCTION FOR HADRON- NUCLEUS

TABLE 1  $C_{\iota}$  Values for Considered Interactions

Collision Type	$p^{\pm}-Ar^{40}$	$p^{\pm} - Xe^{131}$	$p^{\pm} - Au^{197}$
<i>C</i> <sub>9</sub> =	-318.4882903	0	-206.4492348
$C_8 =$	1433.1975229	-13.252443	929.0215018
$C_{7} =$	-2718.5422389	33.4122421	-1762.202021
$C_6 =$	2826.6421281	-18.4012711	1832.2735619
$C_5 =$	-1755.406211	-21.3332567	-1137.881365
$C_{4} =$	666.0383358	33.3692337	431.7361847
$C_3 =$	-152.2887535	-18.035544	-98.7158647
$C_{2} =$	20.2506544	4.8202546	13.1267791
$C_1 =$	-1.8432316	-0.8927745	-1.1948108
$C_{0} =$	0.9937953	0.9904029	0.995978

#### COLLISIONS

The number of created pions from each fireball (  $n_o$  ) will be given by,

$$n_{o}(Z) = \frac{Z(x)T_{0}}{\varepsilon} = \frac{Z(x)Q}{2\varepsilon}$$
(7)

Where, Q is the free energy which is carried by the two fireballs,  $\mathcal{E}$  is the energy required for the creation of one pion. From Eqs. (6, 7), the probability of the emission of any number of pions ( $n_0$ ) from one fireball in the peripheral collision can be obtained in the form:

$$P(n_o)dn_0 = \int_{n_0}^{n_0+1} 2 \sum_{k=0}^{k=9} C_k \left(\frac{2\varepsilon n_0}{Q}\right)^k$$
$$\sum_{k=0}^{k=8} k C_k \left(\frac{2\varepsilon n_0}{Q}\right)^{k-1} \frac{2\varepsilon}{Q} dn_0 \qquad (8)$$

It is clear that at a given impact parameter, Eq. (7) gives the total number of created particles (i.e. charged and neutral particles).

According to the above scheme, the charged multiplicity distribution will be given by,

$$\mathbf{P}(\mathbf{n}_{ch}) = \sum_{n=1}^{nch} \Phi(n) \Phi(n_{ch} - n)$$
(9)

; 
$$n_{ch} = 2, 4, 6, \dots, Q/\varepsilon$$

Where, 
$$\Phi(\mathbf{n}) = \sum_{n_0} \Psi(n_2) P(n_0)$$

,  $\Psi(n_2)$  is the Poisson distribution of the form,

$$\Psi(\mathbf{n}_{2}) = \frac{\mathbf{N}^{n_{2}}}{\mathbf{n}_{2}!} p^{n_{2}} e^{-\mathbf{N}\mathbf{P}}$$
(10)

Where, N: is the number of pairs of created particles from one fireball ( $N = \frac{n_0}{2}$ ),  $n_2$  the number of pairs of charged pions,  $n_2 = \frac{n-1}{2}$ , P the probability that the pair of pions is charged. The number of negative particles from one fireball equals the

half of new created charged pions  $n_{-} = \frac{n_{ch}}{2}$ 

The probability distribution of negative particles  $P(n_{_{_{_{_{_{_{_{_{_{_}}}}}}}}})$  is the same as the probability distribution of charged particles

$$P(n) = P(n_{ch} = 2n)$$
 (11)

; 
$$n = 0, 1, 2, 3, \dots, Q/2\varepsilon$$

### **3** ARTIFICIAL NEURAL NETWORKS (ANNS)

ANNs are composed of interconnecting artificial neurons (programming constructs that mimic the properties of biological neurons). Artificial neural networks may either be used to gain an understanding of biological neural networks, or for solving artificial intelligence problems without necessarily creating a model of a real biological system. The real, biological nervous system is highly complex: artificial neural network algorithms attempt to abstract this complexity and focus on what may hypothetically matter most from an information processing point of view.

The neuron transfer function, f, is typically sigmoid or step function that produces a scalar output (n) as in Eq. (12):



 $n = f \sum_{i} w_i I_i + b \tag{12}$ 

Where,  $I_i$ ,  $w_i$ , b are the  $i^{th}$  input, the  $i^{th}$  weight and b the bias respectively.

"Figure 1. Neuron model"

A network consists of one or more layers of neurons. A layer of neurons is a number of parallel neurons. These layers are configured in a highly interconnected topology as shown in figure (1).

### 4 TRAINING OF THE H-A-ANN

Neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between elements. Training in simple is to make a particular input leads to a specific target output. The weights are adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically many such input/target pairs are used, in this supervised learning, to train a network.

The proposed ANNs in this paper was trained using Levenberg–Marquardt optimization technique. This optimization technique is more powerful than the conventional gradient descent techniques [27-31].

The Levenberg-Marquardt updates the network weights using the following rule,

Where, J is the Jacobean matrix of derivatives of each error with respect to each weight.  $J^T$  is the transposed matrix of J; I is the identity matrix that has the same dimensions of  $J^T J$ ,  $\mu$  is a scalar; changed adaptively by the algorithm and e is an error vector.

The only requirement for this method is a considerably large memory for large problems. The initial training weights were also chosen using the Nguyen–Widrow random generator in order to speed up the training process [27-31].



$$\Delta W = (J^T J + \mu I)^{-1} J^T e$$



### 5 RESULTS AND DISCUSSION

Charged and negative pions multiplicity distributions, Eq. (9, 11), are calculated for  $p^{\pm} - Ar^{40}$ ,  $p^{\pm} - Xe^{131}$ ,  $p - Au^{197}$  and  $p - He^4$  assuming  $\varepsilon$  in Eq. (11) is given by :  $\varepsilon = a n_o + b$ 

Where, a = 0.04, b = 0.35 as in references [14, 17, 18]. The results of these calculations are represented in figure 2 (a, b, c, d, e) and figure 3 (a, b, c, d, e, f) along with experimental data [32, 34] which show fair agreement with the corresponding experimental data. It can be seen from figs. (2, 3) that charged and negative pions multiplicity distributions are not in accordance with the experimental data for heavy nuclei although the situation becomes better for the light ones. The emission of secondary particles is assumed to follow a Poisson distribution. As mass number increases the multiplicity distribution is not broaden but its peak is shifted to high numbers.

We have also calculated the same collisions by using ANN model and these calculations are represented in figs (2, 3) along with the same experimental data [32-34]. We have also found great variations compared to PTFM.

Different configurations of network structure were investigated to achieve good mean squared error (MSE) and good performance for the network using the input-output arrangement. The input and target vectors are randomly divided into three sets (validation set, training set, testing set), 80% of the vectors are used to train the network and 20% of the vectors are used to validate how well the network generalized.



The proposed neural network model of charged and negative pions multiplicity distributions of  $p^{\pm} - Ar^{40}$ ,  $p^{\pm} - Xe^{131}$ ,  $p - Au^{197}$  and  $p - He^4$  collisions at 200 GeV/c have three inputs  $(n_{ch}, P_{Lab}, A)$ , one output  $P(n_{ch})$  and two hidden layers of 23, 22 neurons, for charged pions, two hidden layers of 24, 21 neurons for negative ones. The transfer functions of the first and second hidden layers were chosen to be a tan sigmoid, while the output layer was chosen to be a pure line. The trained method which used to train the ANN model is Levenberg-Marquardt optimization technique, with number of epochs=19, performance of order  $10^{-5}$  for charged pion production and epochs=10, performance of order  $10^{-4}$  for negative ones.



NN model, (.....) PTFM, ( $\Delta$ ) experimental data"

It is should be emphasized that when the mass increases the multiplicity distributions of charged and negative pion production using ANN are consistent with all regions of the experimental data (low, medium, high multiplicity). In contrary with PTFM, the theoretical calculations are inconsistent with the experimental data especially at high multiplicity. That is why; we use the ANN technique because it is able to exactly model the multiplicity distribution for different beams in hadron nucleus interactions.

### APPENDIX

Our obtained function for charged and negative pions in h-A interactions is generated using the obtained control NN parameters as follows:

The structure of the network is 3-23 -22-1 for charged pions and 3-24 -21-1 for negative ones. The obtained equation which describes the multiplicity dustribution of charged and negative pions in h-A interactions for different beams (projectiles), different mass numbers (nucleii) at the same energy is given by:

$$P(n_{ch}) = pureline[\{net.LW(3,2). \tan sigmoid.\{net. LW(2,1). \tan sigmoid \{net.IW(1,1).P + net.b(1)\} + net.b(2)\} + net.b(3)\}]$$

Where, pure line is linear transfer function, tan sigmoid is hyperbolic tangent sigmoid transfer function.

*P* is the input which is  $(n_{ch}, P_{Lab}, A)$ .

net.LW(3,2) linked weight between the second hidden layer and the output.

*net.LW*(2,1) linked weights between the first and the second hidden layer.

*net.IW*(1,1) linked weights between the input layer and the first hidden layer.

*net.b*(1) is the bias of the first hidden layer.

*net*.*b*(2) is the bias of the second hidden layer.

*net.b*(3) is the bias of the output layer.

### REFERENCES

- [1] S. Fredriksson, Preprint Ref. TH. 2720-CERN, August (1979).
- [2] A. Dar and J. Vary Phys Rev D 6, 2412 (1972).
- [3] K. Gottfried Phys. Rev. Lett 32, 957 (1974).
- [4] Yoichiro Nambbu, The Conference of Quarks Scientific American (1976).
- [5] V. S. Barashenkov et al., Sov. Phys. Usp 16, 31 (1973); Yu. P. Nikitin et al, Sov. Phys. Usp 20, 1 (1977).
- [6] A. Capella and A. Krzywicki, Phys. Rev. D 18, 3357 (1978).
- [7] D. Amati et al., Nuovo Cimento. 26, 896 (1962); F. Zachariasen, Phys. Rep., C 2, 1 (1971).
- [8] R. P. Feynman, photon-Hadron Interactions (Benjamin, Reading, Massachusetts, 1972).
- [9] E. Fermi, Prog. Theor. Phys. 5, 570 (1950).
- [10] J. Ranft, Phys. Lett, B 31, 529 (1970).
- [11] R. Hagedorn, Nuovo Cimento Suppl. 3, 147 (1965); R. Hagedorn and J. Ranft, Nuovo Cimento Suppl.6, 169 (1968).
- [12] M. Tantawy Ph. D. Dissertation, Rajasthan University (Jaipur-India) (1980).
- [13] D.M.Habashy, Ph. D. Dissertation, Ain Shams University (Cairo-Egypt) (2011).
- [14] Ehab G. Abbas, M. Sc. Thesis Ain Shams University (Cairo-Egypt) (2010)
- [15] Moaaz A. Moussa, M. Sc. Thesis Ain Shams University (Cairo-Egypt) (2009)
- [16] S. Gamiel, M. Sc. Thesis Ain Shams University (Cairo-Egypt) (1997)
- [17] M. El-Mashad, Ph. D. Dissertation, Cairo University (Cairo-Egypt) (1994).
- [18] M.Y. El-Bakry, M. Sc. Thesis, Ain Shams University (Cairo-Egypt) (1987).

- [19] T. I. Haweel, M. Y. El-Bakry, K. A. El-Metwally, Chaos, Solitons and Fractals 18, 159 (2003).
- [20] M. Y. El-Bakry, Chaos, Solitons and Fractals 18, 995(2003).
- [21] M. Tantawy, M. El-Mashad, S. Gamiel and M. S. El-Nagdy Chaos, Solitons and Fractals 13, 919 (2002).
- [22] M. Tantawy, M. El-Mashad, M. Y. El-Bakry. Indian J Phys.72A, 110 (1998).
- [23] M. Tantawy, M. El Mashad, M. Y. Elbakry, The 3rd International Conference on Engineering Mathematics and Physics (ICEMP) Faculty of Eng., Cairo university, 23rd–25th Dec. (1997).
- [24] AK Hamid, Can. J. Phys. 7, 76 (1998)
- [25] P. Bhat, Using neural networks to identify jets in hadron-hadron collisions. Proceedings of The 1990 Summer Study on HEP, (1990), Research Directions--the Decade, Snowmass, Colorado
- [26] RP. Lippman, IEEE Acoust Speech Signal Process Mag. (1987)
- [27] M.T. Hagan, Menhaj M.B, IEEE Trans Neural Networks 6:861-7 (1994)
- [28] El-Sayed El-dahshan, A. Radi. Cent. Eur. J. Phys. 9(3), 874-883 (2011)
- [29] S. Y. El-Bakry, El-Sayed El-dahshan, M. Y. El-Bakry. Ind. J. Phys., 85 (9), 1405-1413 (2011)
- [30] S. Y. El-Bakry, El-Sayed El-dahshan, S. Al-Awfi and M. Y. El-Bakry. ILNouvo Cimento, 125B (10), (2010)
- [31] El-Sayed El-dahshan, A. Radi and M. Y. El-Bakry. Int. J. Mod. Phys. C, 20(11), 1817-825 (2009)
- [32] C. De Marzo et al, Phys Rev, D 26, 1019 (1982).
- [33] D. H. Brick et al., Phys. Rev. D 39, 2484 (1989).
- [34] W. Bell et al., Phys. Lett. B128, 349 (1983).

# Analysis optimum transmit power of 10 Gbits optical CDMA system in fiber-to-the-home access network

N. Ahmed, S. A. Aljunid and R. B. Ahmad

**Abstract**— In this paper, the optimum transmits power for an optical code division multiple access (OCDMA) is analyzed at different data rates and transmission distance. We used Enhance Double Weight (EDW) codes as a signature address in designing the system because this code can accommodate more number of simultaneous users under considerable standard Bit-Error-Rate (e.g.  $\leq 10^{-9}$ ). The induced EDW codes for OCDMA system can suppress multiusers interference and increase the bit-error-rate performance with optimum transmit power. The numerical simulations have been taken into the account to carry out analysis. We ascertained by simulation results that the optimum power is decreases with the distance and high bit rate and maintain error floor transmission rate  $(10e^{-09})$ . Therefore this system can be considered as a promising solution for optical access network such as Fiber-to-the-Home access network.

Index Terms— Optical Code Divission Multiple Access (OCDMA), Enhance Double Weight (EDW), Fiber Brag Grating (FBG), NAND subtraction, AND Subtraction, Fiber –to-the-Home (FTTH), Multiple Access Interference (MAI).

\_\_\_\_\_

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

### 1. INTRODUCTION

ORE than one decade, the rapid advance has been real-**V** Lized in the optical fiber communication technology with significant reduction of loss in single mode fiber. The main progress has been obtained in high sensitivity and high-speed optical detectors, development of high-speed semiconductor laser diodes, and the advance of optical amplifiers to the fiber communication transmission capacity. These advances have had an important impact on the field of optical telecommunications system and widely used in many more applications of communication engineering [1]. Optical Code Division Multiple Access (OCDMA) has been recognized as one of the most important technologies for supporting many users in shared media simultaneously, and in some cases can increase the transmission capacity of an optical fiber [2]. OCDMA has been exciting developments in short haul optical networking because this system can supports both wide and narrow bandwidth applications on the same network. In addition, it connects large number of asynchronous users with low latency. The jitter, permits quality of service guarantees to be managed at the physical layer which offers robust signal security and has simplified network topologies. In order to suppress the MAI effect sufficiently, many subtraction techniques with codes fixed in-phase cross correlation have been proposed [3-4].

The detection is one of the important processes to design the system transmitter and receivers. In general, there are two well known basic detection techniques, namely coherent and incoherent [5]. The knowledge of the phase information of the carriers keeps big impact when coherent detection send detection signal. On the other hand, the incoherent detection has no such kinds of information. Alternatively, the incoherent OCDMA is performed in a unipolar approach and coherent is performed in a biopolar behavior with the coding operation. The less hardware complexity of incoherent detection makes a popular candidate compared to coherent detection. Moreover, the incoherent detection does not need phase synchronization. The application of the coherent technique will be more difficult than incoherent technique. Therefore, we have chosen the incoherent detection technique based on spectral-amplitudecoding (SAC) for this research. However, the cross correlation function is always generated in the incoherent code words. As results, the multiple access interference (MAI) is generated in the system due to this cross correlation, which can be reduced by using a suitable detection technique in OCDMA systems.

Many kinds of detection techniques are available and already proposed by many researchers [3, 6, 7 and 8]. The well known detection techniques are the complimentary subtraction technique [6 and 7], the AND subtraction technique [7], the spectral direct detection technique (SDD) [8] and the XOR subtraction technique [4]. However, all these detection techniques have various limitations. Although some of these detection techniques has successfully reduced the MAI effect but still suffering from the poor signal quality, which is considered as a big limitation of the existing detection techniques. The Enhance Double Weight (EDW) [8] code was successfully applied in the complimentary and AND subtraction techniques but the problem of poor signal

Dr. N. Ahmed is currently workingas senior lecture in school of computer and communication engineering in University, Malaysia Perlis (UniMAP)PH++6049798686. E-mail:nasim751@yahoo.coml.com

Pro.Dr.S.A.Aljunid is currentlyworking as dean of R&D in University, Malaysia Perlis (UniMAP). E-mail:syed@ unimap.edu.my

Prof. Dr. R. B. Ahmad is currently working as a dean of school of computer and communication engineering, University Malaysia Perlis(UniMAP. Email:badly@ unimap.edu.my

quality remain the same. In order to solve this problem, we introduce a new detection technique named the NAND subtraction detection technique [9], which can reduce the MAI in a significant amount. OCDMA is a multiplexing technique taking the advantage of the large bandwidth of fiber and flexibility of code division multiple accesses. Now it is getting more and more attention, because of its prominent advantages, such as unique confidentiality, scalability, asynchronous access, convenient network management and etc [10]. Therefore, this proposed OCDMA system would then be a prospective candidate for the future FTTH access network.

The remainder of this paper is organized as follows. In Sec II, we review MDW code construction. We discuss about detection technique in Sec III. The Sec IV shows the system architecture. Network simulation setup is shown in Sec V. Results and discussions are shown in Sec VI, and finally some conclusions are drawn in Sec VII

### 2. EDW Code Construction

The basic matrix for the EDW code consists of  $K \times N$  matrix depending on the value of the code weight. The general form of the basic matrix of an EDW code with weight *W* is shown in Fig. 1, where all the component matrices [A<sub>1</sub>], [A<sub>2</sub>], ..., [A<sub>w</sub>] depend on *W*. The basic matrix consists of the minimum number of *K* and *N* for the specific number of code weight. From the basic matrix, a larger number of *K* can be achieved using the mapping technique as bellow:

		÷		÷	÷	7
		÷		÷	÷	
[H] =	$A_1$	÷	$A_2$	÷	 ÷	$A_w$
		÷		÷	÷	
		÷		÷	÷	

Fig.1. General form of the EDW code matrix [12].

The size of each matrix consist of  $K_a \times N_a$ , Where:

$$K_a = W \tag{1}$$

and

$$N_a = \frac{\sum_{j=l}^{w} j}{W}$$
(2)

The basic matrix for EDW consists of a  $3 \times 6$  matrix. The component matrices are [A<sub>1</sub>], [A<sub>2</sub>], and [A<sub>3</sub>]. The size of matrix [A] is  $3 \times 2$ , after using Eqs. (1) and (2). The combination sequence for each matrix is 2 and 1. The basic EDW code denoted by (6, 3, 1) is shown [9]. The basic matrix consists of a chip-

combination sequence of 1,2,1,2... (alternating 1's and 2's) for the columns. A chip combination is defined as the summation of the spectral chips (1's and 0's) for all users (or rows) in the same column with each code sequence allowed to overlap at most, once with every other sequence in the columns of the matrix.

#### 2.1. NAND Subtraction Technique

The mobility of the digital electrons in NAND gate is three times higher than AND gate as well as NOR gate [11]. This statement refers to the digital logic gates (AND, OR, NAND). However, in our proposed system the idea of NAND is used as an operation, not as a digital gate. Considering this point of view, the authors brought the concept of the NAND subtraction technique in our study. In the NAND subtraction detection technique, the cross-correlation  $\theta_{\bar{X}\bar{Y}}(K)$  is substituted by  $\theta_{(\bar{X}\bar{Y})Y}$ , where  $\theta_{(\bar{X}\bar{Y})}$  represents the NAND operation between X and Y sequences. For example, let X = 1100 and Y = 0110 therefore the NAND is  $(\bar{X}\bar{Y})Y = 0010$ . Fig. 2 shows the implementation of NAND subtraction detection technique and Table 1 shows the comparisons between complementary and NAND subtraction detection technique using EDW codes.

Table.1Comparison of complementary and NAND subtraction detection technique

	Complementary Subtraction			NAND Subtraction				
	λ	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$
Х	1	1	0	0	1	1	0	0
Y	0	1	1	0	0	1	1	0
			$\theta_{XY}$	, = 1	$\theta_{XY} = 1$			
		$\frac{1}{X = 0011}$			$\theta_{\widetilde{X}\widetilde{Y}} = 1011$			
	$\theta_{\overline{XY}} = 1$			$\theta_{(\widetilde{X}\widetilde{Y})Y} = 1$				
Ζ		<i>Z</i> =	$\theta_{XY}$	$-\theta_{\overline{XY}}=0$	Z =	$\theta_{XY}$ -	$-\theta_{(\widetilde{X}\widetilde{Y})}$	$y_{Y} = 0$

Note that  $\lambda_i$  (where *i* is 1, 2....*N*) is the column number of the codes which also represents the spectral position of the chips. Therefore, MAI can be cancelled using both techniques. However, NAND subtraction detection technique can generate extra weight as shown in Table 1. This is due to the fact that when the code weight is increased, the signal power increases as well; hence, increases the signal-to-noise ratio. Therefore, The OCDMA performance is improved significantly using the NAND subtraction detection technique.



Fig.2. Implementation of NAND subtraction technique

### 3. Network Simulation Setup

A simple schematic block diagram consisting of three users are illustrated in Fig. 3. This system each chip has a spectral width of 0.8 nm. The tests were carried out at the rate of (10 Gbits), for 30 km. The fiber used had the values of parameters taken from the data which are based on the G.652 Non Dispersion Shifted Fiber (NDSF) standard. This included the attenuation, group delay, group velocity dispersion, dispersion slope and effective index of refraction, which were all wavelength dependent. The non-linear effects such as the Four Wave Mixing and Self Phase Modulation (SPM) were also activated and specified according to the typical industrial values to simulate the real environment as close as possible. The system specifications listed in Table 1 were used throughout the simulation. At 1550 nm wavelength, the attenuation coefficient was 0.25 dB/km, and the chromatic dispersion coefficient was 18ps/nm-km and the polarization mode dispersion (PMD) co-efficient was 0.5 ps/sqrt (km). The transmit power used was between -5 dBm to 5 dBm out of the broadband source. The noises generated at the receivers were set to be random and totally uncorrelated. The dark current value was 5 nA and the thermal noise co-efficient was  $1.8 \times 10^{-23}$ W/Hz for each of the photodetectors. The performance of the system was evaluated by referring to the bit error rate, received power, transmit power and the eye pattern. Fig.3 shows the system simulation setup which has transmitter and receiver. At the transmitter, the coherent light (Lesar) source is used. The pseudo-random bit sequence generator and non-returnto-zero (NRZ) pulse generator to generate the input signals, an external modulator to modulate the input signal into the optical signal. The modulated signal is transmitted through single

mode optical fibe. The function of the encoder was amplitude spectrally encode the source according to the specific used code which is Enhance Double Weight (EDW) code with a weight three. At the receiver side consist of FBG which act as a decoder, PIN photo-detectors and low pass filters. The decoding scheme used was NAND subtraction. A subtractor is used for subtract the overlapping data from the desired one.

TT 1 1 1 C	6		• • • • •
label 1. Su	mmary of	system s	pecifications
10001.1.00	many or	o y bicini b	pectitention

	J 1
Data rate	10 Gbits, 2.5 Gbits and 662 Mb/s
Channel spacing	0.8-nm
Insertion loss	0.25 dBm
Mux/Dmux insertion loss	2 dBm
Fiber attenuation	0.25dB/km



Fig.3. Simulation setup for the OCDMA system with NAND subtraction detection technique

### 4. Results and Discussion

Fig. 4 depicts the relationship between received power and the measured BER for NAND subtraction technique at (10 Gbits, 2.5 Gbits and 622 Mbps) data rates. The optical fiber length was fixed (30 km) for the analysis. It clearly shows that, the system using NAND subtraction technique at various bit rate 10 Gbits, 2.5 Gbits and 622 Mbps the BER is 4.05e-10, 2.02e-<sup>24</sup> and 1.63e<sup>-39</sup> respectively when the minimum transmit power is -2 dBm. (10-13) and received power -19.189 dBm, -20.586 dBm and -21.700 dBm respectively. The Fig. 4 demonstrates that at low bit rate (622 Mbps) BER and received power is better as compare to high bit rate. However, at high bit rate (10 Gbits) the system can have standard error free transmission value BER 4.05e<sup>-10</sup> with optimum received power (19.189 dBm). The main objective of this analysis is to find out the optimum received power with standard error free (10e-09) transmission value for high bit rate system. On the other hand Fig.5 presents the effect of transmit power against BER at various bit rate 10 Gbits, 2.5 Gbits and 622 Mbps for fixed fiber length (30

km). As it is seen from Fig. 5 that at 10 Gbits the system require -2 dBm transmit power to maintain the error floor transmission rate where as 2.5 Gbits and 622 Mbits require transmit power -4 dBm and -5 dBm. Although, is this study we have considered maximum input power 5 dBm. At transmit power 5 dBm the system BER is 2.83e<sup>-33</sup>.



Fig. 4. Received power versus BER for different bit rates and a fixed transmission distance (30 km)



Fig. 5. Transmit power versus BER for different bit rates and a fixed transmission distance (30 km)

As shown in Fig.6, that the system using new detection technique shows BER in fixed bit rate in respect to the different fiber length (15 to 30 km). It is found that the system can transmit excellent signal up to 30 km at 10 Gbits. The transmit power 0 dBm is taken for analyzing the performance. It is also seen from fig.6 that as the distance increase the error floor become worst and received power is high. Though, the received power is slightly high but the signal quality is exactable range which satisfies the error free transmission rate (10e-<sup>09</sup>). On the other hand, Fig. 7 shows the effect of transmit power on the BER. It is also found that, as the input power is increased the

error floor is decries. To conclude all the results, we found that the system pwer distribution and BER is very much depended on bit rate. Moreover, the performance of the NAND subtraction technique is evident at all rates with supportable distance to support by the conventional technique. In Fig. 7 (a, b and c) shows the measured eye patterns at (10 Gbits, 2.5 Gbits and 622 Mbps). It clearly illustrates that using new detection technique the system had a better performance with a larger eye opening at 10 Gbits data rate.



Fig.6. Received power versus BER for different transmission distance with a fixed bit rate of 10 Gbits



Fig. 7. Transmit power versus BER for different transmission distance with a fixed bit rate of 10 Gbits



(a). 10 Gbits bit rate



(b). 2.5 Gbits



(c). 622 Mbps

### 6. CONCLUSSION

In this paper, optical CDMA system has been analyzed using numerical simulation to find out the optimum transmit power at different bit rates and transmission distance. The NAND subtraction technique has been applied in the receiver to improve the system performance. The Enhance Double Weight (EDW) code is used as a signature address of the system. The analysis has revealed that the system at 10 Gbits, the received power is acceptable when transmitting power is -2 dBm. It has been also shown that the system can maintain error floor transmission rate (10e<sup>-09</sup>) with low transmit power (-4 dBm) for 30 km. To consider these advantages, it is concluded that the proposed system can be suitable for Fiber-to-the-Home (FTTH) access network to fulfill the consumer's demands.

### REFERENCES

- A.F. Hillal, S. A. Aljunid and R. B. Ahmaad, "Effect of random diagonal code link of an OCDMA scheme for high-speed access networks," Journal of Opt. Fiber Technol, vol. 15, no. 3, pp. 237-240, June 2009, doi.org/10.1016/j.yofte.2008.11.001. (Optical Fiber Technology)
- Z. H. Wei and Ghafouri-Shiraz, "Codes for spectral-amplitudecoding optical CDMA systems," Journal of Lightwave Technol, vol. 50, no. 8, pp. 1284–1291, Aug 2002, doi:10.1109/JLT.2002.800301, 2002.12.10. (Lightwave Technology)
- [3] E.D.Smith and R. J. Blaikie, "Performance enhancement of spectralamplitude-coding optical CDMA using pulse-position modulation," IEEE Trans. Commun., vol. 46, no. 9, pp. 1176-1185, Sep 1998, doi:10.1109/26.718559, 6-08-2002 (IEEE Transactions)
- [4] H.Y.Ahmed, I.Faye, N.B.M.Saad and S.A.Aljunid," OCDMA system: new detection scheme and encoder- decoder structure based on fiber bragg gratings (FBGS) for Vcc code," Int. Journal of Computer and Application vol.32, no.4, pp. 2878-2881, April 2010, doi: 10.2316/Journal.202.2010.4.202-2881 (Int. J. of Comp. Appl.)
- [5] W.Huang, M.H.M. Nizam, I. Andonovic and M. Tur," Coherent Optical CDMA (OCDMA) Systems used for High Capacity Optical Fiber Networks System Description, OTDMA Comparison, and OCDMA/WDMA Networking," Journal. of Lightwave Technol, vol.18, no.6, pp. 765-778, June 2000, doi:10.1109/50.848384 (Lightwave Technol)
- [6] L. Nguyen, B. Aazhang, J. F. Young, "All-Optical CDMA With Bipolar Codes," Elect. Letters, vol.31, no.6, pp. 469-470, March 1995, doi: 10.1049/el: 19950296 (Electronic letters).
- [7] S.A. Aljunid, M.Ismail, A.R. Ramli, M. A. Borhanuddin, and Mohamad Khazani Abdullah, "A New Family of Optical Code Sequences for Spectral-Amplitude-Coding Optical CDMA Systems," IEEE Photon Technol. Letters, vol. 16, no.10, pp. 2383-2385, October 2004, doi: 10.1109/LPT.2004.833859, (IEEE photonic letters).
- [8] M. K. Abdullah, N. Feras. Hasoon, S.A. Aljunid, Sahbudin Shaari, "Performance of OCDMA Systems with new spectral direct detection (SDD) technique using enhanced double weight (EDW) code," Journal of Optical Communication, vol. 281, no. 18, pp. 4658-4662, Sept 2008, doi: http://dx.doi.org/10.1016/j.optcom.2008.06.029, (Optical communication)
- [9] N.Aahmed, S.A.aljunid, R.B. Ahmad, H. A. Fadhil, and M.A. Rashid," Performance Improvement of Spectral Amplitude Coding-Optical Code Division Multiple Access Systems Using NAND Detection with Enhanced Double Weight Code" Optical Review, vol. 19,

no.1, pp. 1-6, Oct 2011, doi: 10.1007/s10043-012-0001-2, (Optical review)

- [10] J. Huang, Y. Chang and C. Hsu, "Hybrid WDM and Optical CDMA Implemented over Waveguide-Grating-based FTTH Networks," Optical Fiber Technolo, vol 13, no. 3, pp. 215-225, July 2007, doi: 10.1016/j.yofte.2007.02.002, (Optical Fiber).
- Dr. N. Ahmed is currently workingas senior lecture in school of computer and communication engineering in University, Malaysia Perlis (UniMAP)PH-+6049798686. E-mail:nasim751@yahoo.coml.com
- Pro.Dr.S.A.Aljunid is currentlyworking as dean of R&D in University, Malaysia Perlis (UniMAP). E-mail:syed@ unimap.edu.my
- Prof. Dr. R. B. Ahmad is currently working as a dean of school of computer and communication engineering, University Malaysia Perlis(UniMAP. E-mail:badly@ unimap.edu.my

# Commodity Price Effect on Stock Market: A Markov Switching Vector Autoregressive Approach

<sup>1</sup>Phoong Seuk Wai, <sup>2</sup>Mohd Tahir Ismail, and <sup>3</sup>Sek Siok Kun

<sup>1,2,3</sup>Universiti Sains Malaysia School of Mathematical Sciences 11800 Minden, Penang, Malaysia 1wai\_phoong@yahoo.com, 2mtahir@cs.usm.my, <sup>3</sup>sksek@usm.my

**Abstract**— Real economic data always present nonlinear properties such as asymmetry and radically change in the series through time. Missing data and jumps as well as breaks also common reported in economic time series model. Thus, linear models are no longer suitable used in estimate the economic data and markov switching vector autoregressive model (MS-VAR) is applied in study the economic model. This paper will present the analysis of commodity price effect on stock market by using markov switching vector autoregressive model. Oil price and gold price are selected to represent the commodity price and the influence of those prices on Malaysia, Singapore, Thailand and Indonesia stock market are investigated. Furthermore, the mean adjusted markov switching vector autoregressive model (MSM-VAR) and mean adjusted heteroskedasticity markov switching vector autoregressive model (MSMH-VAR) are investigated to determine the suitable specification of the model in providing a more significance and reliable result when analysis the economic relationship model between oil price, gold price and stock market returns.

Index Terms— Markov Switching Vector Autoregressive Model, Mean Adjusted, Heteroskedasticity, Stock Market and Commodity Price

### I. INTRODUCTION

Nonlinear time series models became popular through the past two decades in economic and financial field. Fan and Yao (2003) [1] define that nonlinear time series is not a linear stochastic process but is generated by nonlinear dynamic properties such as time varying, higher-moment structures, asymmetric cycles, and jumps or breaks exist in the time series sequence. Markov switching autoregressive model is one of the popular nonlinear time series model and proposed by Hamilton (1989) [2]. The Hamilton model is then extended to Markov switching vector autoregressive model (MS-VAR) by Krolzig.

According to Krolzig (1997) [3], MS-VAR model have become increasingly prominent in applications since it is able to detect the classical business cycle phases and detect the difference in terms of average growth rates of the economy. Moreover, MS-VAR model also can increase the reliability of analysis of the business cycle, and provide useful information in the study of economic relationship because of its properties on detection the high level regime. Therefore, the perception of the current state of economy can be improving.

In addition, MS-VAR model are a nonlinear model which commonly applied in economic and financial time series to analysis and predict the economy and financial relationship, and give implication on the economic theories since MS-VAR model able to estimate the dramatically changes or breaks that happened in the data through time such as at the last decade oil price shocks, structural change or break in the stock market prices are happen because of the events like Asian Financial Crisis and '9-11'. Commodity prices such as gold and oil always relate to the movement of the stock market prices. Historical evidence was shown that when the global economy was in the recession period, the demand of oil and gold will be increased. Although oil and gold are limited reservation but oil belongs to a basic of energy sources which affect the country's economy as states by Sauter and Awerbuch (2003) [4] and gold is a popular investment. Both of these two commodity prices implies on the country's GDP especially in stock market index.

In this study, MS-VAR model is choose to use in analysis the implication of the world oil price (OP) and gold price(GP) in Malaysia stock market return (KLCI), Singapore stock market return (STI), Thailand stock market return (SETI) and Indonesia stock market return (JCI) since commodities are the primary inputs to many manufactured things and services.

Oil price is an important commodity that able to directly influence the world market especially in stock market are included in this study since Bhar and Hammoudeh (2011) [5] describe that oil price is positively leads the stock market in a normal state. Besides that, gold which is a popular investment of the investors also will be take account in the study. This is because there are investors believe that gold is a liquid assets which can be use to buy a low-value stock or pay other assets and gold also believe that has less investment risk compare to the others commodity. Thus commodity prices pay role in affecting the general price level in a country. While "Trading on the stock market is set to be brisk as bulls attempt to break out from bear territory" which reported in New Straits Times on 21st March 2011 show that stock market pay an important role in the country economy and as an indicator to reflect any changes in the country economy. So, the relationship between OP, GP and stock market index are examined in this study.

The structural of this paper is as follows. The section 2 discusses the methodology of the study. Section 3 presents the results and discussion. Lastly, section 4 presents the conclusion of this study.

## II. METHODOLOGY

Stationary test is used to detect whether the time series data have a constant mean, variance or auto-covariance since in real world data, extreme changes always exist in economic model. Many economic data exhibit nonstationary behavior such as regime switching and jumps therefore Augmented Dickey-Fuller (ADF) test and Kwiatkowski-Philips-Schmidt-Shin (KPSS) test are important to apply in the study to check the occurrence of the random walk or unit root problem in the time series data.

Next is cointegration test where Johansen test is used to examine the properties of the series (check the existence of the stationary linear combination). Furthermore, the presence of cointegration in the series must take account the method to test the occurrence of the unit roots in the variables.

Although Ordinary least square regressions can given an accurate result in large samples, but this test is not used in estimate the non-stationary variables since cointegration test is a powerful tool in estimate the existence of the unit root in nonlinear time series data and has strength when the time series data have a limited length. Therefore, if the variables in the set are not cointegrated, then MS-VAR model will applied in study the economic relationship model since MS-VAR class of models provides a convenient framework to analysis multivariate changes in regimes.

The general idea of the MS-VAR model is the ndimensional vector time series vectors,  $y_t = (y_{1t}, ..., y_{nt})$ 

depend upon an unobserved regime variable,  $s_r \in \{1,...,n\}$ which represents the probability of being in a particular state to the switching mechanism in various states. The conditional probability density of the observed yt is given by

$$P(y_{t}|Y_{t-1}, s_{t}) = \begin{cases} h(y_{t}|Y_{t-1}, \beta_{1}) \text{ if } s_{t} = 1\\ \cdot\\ \cdot\\ h(y_{t}|Y_{t-1}, \beta_{n}) \text{ if } s_{t} = n \end{cases}$$

Where  $\beta_n$  is the VAR parameter vector with regime n and Yt-1 is the past information. Thus, the parameter vector,  $\beta$  depend on the regime at time t.

If the mean vector  $\mu(s_t)$ , vector of autoregressive process  $A(s_t)$  and matrix of variance-covariance  $\sum(s_t)$  are subject to shift in regime; thus, the general form of the equation MS-VAR will be

$$Y_{t} = \mu(s_{t}) + \sum_{m=1}^{r} A_{m} (Y_{t-m} - \mu_{t-m}) + \varepsilon_{t}$$

$$\varepsilon_{t} \Box \text{ i.i.d } N(0, \sum_{s})$$
(1)

While the mean adjusted MS(2)-VAR(p) model can be generalize as

$$y_{t} - \mu(s_{t}) = A_{1}(s_{t})(y_{t-1} - \mu(s_{t-1})) + A_{2}(s_{t})(y_{t-2} - \mu(s_{t-2})) + u_{t}$$
(2)

Where  $u_t \square \text{NID}(0, \sum(s_t))$  and  $\mu(s_t), A_1(s_t), ..., A_2(s_t), \sum(s_t)$  are parameter shift functions.

# **III. RESULTS AND DISCUSSION**

ADF test and KPSS are as the first step in the testing procedure in this study because want to check the stationarity of the series. The following graph presented all the behavior of the variables before undergoing the tests.



Fig. 1. Time series properties of the variables

Results from ADF and KPSS show that all the variables in the model including OP, GP, KLCI, STI, SETI and JCI are non-stationary and able to transform into stationary series after taking first differencing process.

Next, Johansen Cointegration Test is used to test the cointegrated relations between the series since all the series has the same integration order. There are two types Johansen Cointegration Test that are Johansen Trace Test and Johansen Maximum Eigenvalue Test. According to Johansen (1991) [6], these two tests have different hypothesis where the hypothesis of the trace statistic test in this study are

 $H_0$ : There is at most *r* cointegrating relations

 $H_1$ : There are *m* cointegrating relations

While the hypothesis of maximum eigenvalue statistic test in this study are

 $H_0$ : There is at most *r* cointegrating relations

 $H_1$ : There are r + 1 cointegrating relations

Where cointegration order, r = 0, 1, 2, 3, 4 and m is the total of cointegrating relations that may exist in the test. Before start analysis the cointegration relations between the variables, number of lags need to be determine first because it may affect the accuracy of the results. Therefore, VAR model of estimating the endogenous variables in the model are carry out to determine the lag order of the cointegration test. Lag order is important since higher lag will results lost many data thus VAR Lag Order

Selection Criterion is test before investigate the existence of the cointegrated relations between the variables and the outputs are summarizes and grouped in the tables below:

TABLE I Johansen Cointegration Test of OP, GP, KLCI, STI, SETI and JCI

	Johansen Cointegration Test					
Hypothesis	Trace	e Test	Maximum	n-Eigen test		
on no. of	0.05	Trace	0.05	Maximum-		
CE	Critical	statistic	Critical	Eigen		
	value		value	statistic		
None	83.93712	89.7910*	36.63019	30.0542		
		(0.0176)		(0.2389)		
At most 1	60.06141	59.7368	30.43961	23.7657		
		(0.0532)		(0.2688)		
At most 2	40.17493	35.9711	24.15921	18.7642		
		(0.1244)		(0.2274)		
At most 3	24.27596	17.2070	17.79730	12.8326		
		(0.2982)		(0.2388)		
At most 4	12.32090	4.3743	11.22480	3.0957		
		(0.6568)		(0.7729)		
At most 5	4.129906	1.2786	4.129906	1.2786		
		(0.3015)		(0.3015)		

The value in () is represent to the p-value \* denotes rejection of the hypothesis at the 0.05 level

Lag order 2 is choose for estimate the time series model because of the suggestion result from Final Prediction Error and Akaike Information Criterion (AIC). At most five cointegrating equations may exist in the model since six variables are tested that are OP, GP, KLCI, STI, SETI and JCI. Results show that, there are no cointegrating relations exist between the parameters. So, MS-VAR is used in estimate the economic relationship of the model.

MSM(M)-VAR(p) model and MSMH(M)-VAR(p) model are build up to estimate the economic relationship between the variables but MSMA(M)-VAR(p) model and MSMAH(M)-VAR(p) model are not implement in analysis the economic relationship model since the autoregressive parameter in the economic model is not vary. The purpose to choose mean as a varying factor when build up the model is because it is probable to assume that mean able be adjusted to a new level after a translation from one state to another state or also means that a permanent regime switching in mean  $\mu(st)$  can cause an immediate jump of the observed time series vector to a new level.

TABLE II MS(2)-VAR(1) models for OP, GP, KLCI, STI, SETI and JCI

		MSM(2)-VAR(1)				
	OP	GP	KLCI	STI	SETI	JCI
$\mu_1$	-0.014	-0.030	-0.196	-0.152	-0.203	-0.336
$\mu_2$	0.003	0.006	0.010	0.010	0.007	0.013
σ	0.062	0.043	0.076	0.072	0.097	0.096
Matrix Of Transition Probabilities, p <sub>ii</sub>						

	<i>S</i> <sub><i>t</i>-1</sub> = 1		5	$s_{t-1} = 2$
$s_t = 1$	0.6167		0.3833	
$s_t = 2$	0.0121		(	).9879
Regime Properties				
	No. of	Probability		Duration
	Observations			
$s_t = 1$	9.1	0.0306		2.61
$s_t = 2$	258.9	0.96	694	82.74

LR = Likelihood Ratio linearity test

The values in ( ) is represent to the t-values indicates that the p-value is significant at 5% level

		MSMH(2)-VAR(1)									
	С	P	GP	KLCI	STI	SETI	JCI				
$\mu_1$	0.0	800	0.002	-0.007	-0.005	-0.008	-0.017				
$\mu_2$	0.0	)03	0.006	0.010	0.011	0.007	0.013				
<b>σ</b> 1	0.0	)97	0.058	0.132	0.122	0.150	0.174				
<b>0</b> 2	0.0	)39	0.035	0.047	0.045	0.075	0.067				
		N	latrix Of Tr	ansition Pr	obabilities,	$p_{ij}$					
			$S_{t-1} =$	1		$s_{t-1} = 2$					
$s_t = 1$			0.723	30		0.2770					
$s_t = 2$			0.118	35		0.8815					
			Re	gime Prope	erties						
			No. of	Probability		No. of Probabilit		Dura	ation		
		Ob	servations								
$s_t = 1$			81.400	0.29	0.299700		0.299700		0.299700 3.61000		000
$S_t = 2$			186.600		00300	8.44	000				

LR = Likelihood Ratio linearity test

The values in () is represent to the t-values \*\* indicates that the p-value is significant at 5% level

The role of OP, GP, KLCI, STI, SETI and JCI are assessed with the MS(2)-VAR(1) model to estimate the relationship of the economic model. First regime is represent to the recession state and second regime is the growth state since regime 1 coefficients is smaller than regime 2. In economy, recession state also known as "Bear" market and growth state also known as "Bull" market.

MSM(2)-VAR(1) model is a mean adjusted model after a switch in regime with variant mean. All coefficients of the lag component in first regime in MSM-VAR model are negative. This shows that a decreasing of monthly return is happened on oil price with average 1.4% and in gold price with average 3%. While the second regime which represent the "Bull" market reported an increasing of average 0.3% in oil price and 0.6% in gold price are positively increasing in average of the KLCI, STI, SETI and JCI monthly index. Among these four selected South East countries; JCI market reported the largest increasing and decreasing country in the monthly return either in "Bull" market or in "Bear" market.

The transition probability p11 = 0.61167 and p22 = 0.7879suggest that the second regime is the persistence state since the transition probability of the growth regime is higher than the recession regime. Furthermore, the computed transition probability,  $Prob(s_t = 1|s_{t-1} = 2) = 0.3833$ reported that an increasing transition probability occurs

from a crisis regime (regime 1) to the growth regime (regime 2).

Number of observations that including in the growth state is 259 which is 96.94% of the total sample size and the average duration of each regime are 2.61% on regime 1 and 82.74% on regime 2. This denotes that regime 2 is the dominant phases in the economic model and asymmetric property is present. Crisis economy 1997 may act as a significant issue that indicates a shift of the series of OP, GP, KLCI, STI, SETI and JCI from first regime to second reaime.

Besides that, MSMH(2)-VAR(1) model which estimate vary mean and variance of the variables has the same representation state that the first regime (st = 1) is the "Bear" market and second regime (st = 2) is the "Bull" market. When comparing the MSM-VAR model and MSMH-VAR model, OP and GP have different sign in first regime (recession state). In MSMH-VAR model, OP recorded a high volatility of 0.94% and a high positive average monthly probability of 0.8% in regime 2 compare to 0.16% of volatility and 0.003 or 0.3% of average monthly return in MSM-VAR model.

While gold price recorded 0.2% average of monthly price and  $\sigma 2(st = 1) = (0.058)2 = 0.00336 = 0.34\%$  of monthly volatility in regime 1 and a high positive average of monthly price with small volatility value in regime 2.

The	transition	probability	of	state	1	is
Prob	$s_t = 1   s_{t-1} = 1 = 0$	).7230 and	Prob	$s_t = 1 \left  s_{t-1} \right  = 1$	2) = 0.8	8815

describes that regime 2 has the higher probability than regime 1. A higher probability of transition from recession state to growth state also reported in the MSMH-VAR models and the duration of these two regimes also show same conclusion with MSM-VAR model that a significant asymmetries is presence in the business cycle. While the transition probabilities can be write in a matrix form,

D _	0.7230	0.2770	
r =	0.1185	0.8815	





Fig. 2. MS-VAR Probabilities

Graph 2 show two type of specifications regime properties in mean adjusted after a changes in regime and each MS-VAR model including three panels that are: first panel present how the inferred regime probabilities are translated into switching in the mean growth rate of the variables, second panel show the filtered and smoothed probabilities of regime 1 and third panel reported the filtered and smoothed probabilities of regime 2.

Smoothed probability is the optimal inference of the turning points on the regime at time t which estimate the business cycle by using the whole sample information,  $\Pr(s_t = j | Y_T)$ . While filtered probability is the optimal inference on the state variables at time t by using the information at time t only,  $\Pr(s_t = j | Y_t)$  and this two probability are sketch to explain the economic model of the variables OP, GO, KLCI, STI, SETI and JCI.

Regime 1 of each model depicts very precisely the

recession of the year 1998 and 2009 in MSM-VAR model. However, smoothed probabilities of MSMH-VAR model indicate a long recession from end of the year 1997 until the year 2000, and a few short recession period such as at 1990, 1994, middle year of 2001 and 2008. While regime 2 is represent to the growth period in the business cycle.

Furthermore, MSM-VAR and MSMH-VAR model display that the smoothed probabilities of all models show that the downswings are sudden and much shorter but the upswings are more gradual and highly persistent.

When comparing the log-likelihood, AIC, HQ and SC among the MSM-VAR and MSMH-VAR model, results are recorded in the following to make a comparison on which model able to given a more reliable and significance result.

TABLE III Comparison of the MSM(2)-VAR(1) and MSMH(2)-VAR(1) model

Critorion	$MSM(2) \setminus AP(1)$	
Cillenon		
	IIIOUEI	model
Log Likelihood	2163.1114	2319.4936
AIC	-15.6128	-16.6231
HQ	-15.2307	-16.1280
SC	-14.6614	-15.3904

According to the results that summarizes in the table above, AIC, HQ and SC of the MSMH-VAR model able to provide approximate value than MSM-VAR model when estimate the relationship model between the variables of OP, GP, KLCI, STI, SETI and JCI. Generally, it is a common misconception that the log-likelihood value must be negative and the AIC value is positive; but if the likelihood is derived from the probability density then it is reasonable for the value is exceed 1 or a positive loglikelihood, and hence has a negative AIC. In algebraically, lower AIC is represent the better performance of the model, thus the more negative value of AIC which recorded by MSMH(2)-VAR(1) model able to provide a more significance results compare to MSM(2)-VAR(1) model. This is proven by HQ and SC outputs. Moreover, MSMH(2)-VAR(1) model has a higher log-likelihood than MSMH(2)-VAR(1), therefore MSMH-VAR model can be concluded as more suitable than MSM-VAR model in estimate the commodity price effect on KLCI, STI, SETI and JCI market index.

### **IV. CONCLUSION**

In this paper, MS-VAR model that proposed by Krolzig is used in analysis the commodity price that are oil price and gold price effect on the four selected South-East Asia countries stock market index from December 1989 until May 2012. The two regime markov switching model with first autoregressive order is used in estimate the economic relationship model. First regime in this paper is the recession state while the second regime represent to the growth state. Markov switching vector autoregressive models with mean adjusted after a shift are applied in study the relationship between commodity price effect and stock market returns on Malaysia, Singapore, Thailand and Indonesia. While all the data including World oil price, world gold price, Malaysia stock market index, Singapore stock market index, Thailand stock market index and Indonesia stock market index are taken from DATASTREAM and transform into a natural logarithm form before start the estimation. Oil price and gold price are proved that able to affect the stock market index on Malaysia, Singapore, Thailand and Indonesia.

### ACKNOWLEDGMENT

The authors wish to thanks to their respective university for the financial support and computing facilities that make the works possible to be completed.

### REFERENCES

- [1] J. Fan, and Q. Yao, *Nonlinear Time Series: Nonparametric and Parametric Methods.* New York: Springer, 2003.
- [2] J. D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*,

vol.

- 57, pp. 357-384, 1989.
- [3] H. M. Krolzig, Markov-switching Vector Autoregression. Berlin: Springer, 1997.
- [4] R. Sauter and S. Awerbuch, "Oil Price Volatility and Economic Activity: A Survey And Literature Review," *IEA Research Paper*, International Energy Agency, Paris, August 2003
- [5] R. Bhar, and S. Hammoudeh, "Commodities and financial variables :

Analyzing relationship in a changing regime environment," International

Review of Economics and Finance, vol. 20, pp. 469-484, 2011.

# Fitting Finite Mixture Model to Exchange Rate Using Maximum Likelihood Estimation

<sup>1</sup>Phoong Seuk Yen and <sup>2</sup>Mohd Tahir Ismail <sup>1,2</sup>Universiti Sains Malaysia School of Mathematical Sciences, 11800 Minden, Penang, Malaysia

<sup>1</sup>yen\_phoong@hotmail.com, <sup>2</sup>mtahir@cs.usm.my

Abstract- Exchange rate has great influence to the inflation and economic growth for a country. The importance of currency is that the great influence on import and export prices with the changes of exchange rate. Thus, maximum likelihood estimation (MLE) is used to fit finite mixture model. In this paper, a twocomponent mixture of normal distribution is used to analysis the return value of nominal monthly exchange rate for Malaysia, Thailand and Philippines by using maximum likelihood estimation. The data collected for this paper is taken from July 2005 until September 2012.

*Key words*: Exchange rate, maximum likelihood estimation, finite mixture model, mixture of normal distribution

### I. INTRODUCTION

Finite mixture model is increasingly getting attention over the years because it is a flexible method of modelling. Fields in which finite mixture model has widened applied include economics, engineering, meteorology, financial, biology, medicine, genetics, marketing, physical, social sciences and neural networks. In statistics, the usefulness of applied finite mixture model include modelling of over-dispersed data, fitting the zero-inflated or hurdle models, estimating heavytailed densities, modelling heterogeneity in a cluster analysis, calculating of switching regressions, allows measuring a two component distribution and in estimating a mixing probability in the data. In addition, finite mixture model also able to measure the observable variables, describe the complex system on the data analysis and measure the sensitivity and specificity of the diagnostic and screening procedures.

Finite mixture model made its first recorded appearance in a statistical literature paper by Newcomb [1] in modelling outliers of the data. A few years later, Pearson [2] applied method of moments to fit finite mixture model in order to estimate the parameters in the model. In Pearson study, a mixture of two univariate normal components is used to measure the ratios of forehead to body lengths for 1,000 crabs. Then, method of moments has been widely used and this can be proved by refer to the studies of Charlier and Wicksell [3] and Doetsch [4]. According to Charlier and Wicksell [3], method of moments is extended to the case of bivariate normal components. Then, Doetsch [4] further the study of Charlier and Wicksell [3] to the case of more than two univariate normal components.

With the advent of high speed computer, maximum likelihood estimation is introduced to fit finite mixture model by Fisher [5]. The main reason of applied maximum likelihood estimation instead of method of moment is that maximum likelihood estimation enable large sample sizes being analysis and the time consume is faster than method of moment. Then, Rao [6] applied maximum likelihood estimation to estimate two components of biological problem with equal standard deviations. However, Hasselblad [7] has different idea in his study. In Hasselblad [7] study, maximum likelihood estimation is used to estimate unequal variances with more than two components. Then, maximum likelihood estimation has received more attention and many statisticians start applied the maximum likelihood estimation in their study such as Peters and Coberly [8], Duda and Hart [9] and Hosmer [10] were applied maximum likelihood estimation in estimated the parameters in a mixture distribution by expressed it into an iterative form.

According Horton and Laird [11], maximum likelihood estimation is used to estimate the parameters of joint distribution and solving the missing covariates in the model. Furthermore, Yang [12] applied maximum likelihood estimation to examine biology data. A data set of HIV-1 gp120 *env* gene is analysed in order to identify a number of sites in which under the positive selection.

Moreover, maximum likelihood estimation also popular in used in meteorology field and this can refer to the studies of Castiglioni et al. [13] and Owolawi [14]. Castiglioni et al. [13] described that maximum likelihood estimation is an effective statistical method in examined the rain drop data. While in Owolawi [14] study has similar idea with Castiglioni et al. [13] that applied maximum likelihood estimation in estimated the raindrop size distribution and rain attenuation.

The main reason in which maximum likelihood estimation is vital in statistics is because it has higher probability of being close to the quantities to be estimated and has lower variance as the sample size increases in compared to other methods. In addition, the method also performs statistically well understood and provides a consistent approach when developed for a large variety of estimation situations.

In this study, a two-component mixture of normal distribution is analysis by using maximum likelihood estimation in order to model the financial time series data. The data that mention is the returns value of nominal monthly exchange rate for Malaysia, Thailand and Philippines.

The structural of this paper is as follows. The section 2 presents the literature review. Section 3 presents the sample and data that collected. Section 4 presents the methodology meanwhile section 5 is the results and discussion. Lastly, section 6 presents the conclusion of this study.

### II. SAMPLE AND DATA

The variable that applied in this paper is the return of nominal monthly exchange rate for Malaysia, Thailand and Philippines from July 2005 until September 2012. This paper contains 86 observations for each country. The unit of currency for Malaysia is in Ringgit Malaysia (RM) meanwhile the unit of currency for Thailand and Philippines are in Thai Baht and Peso. The data is collected from yahoo finance. While the statistical software package that use in the study is SAS version 9.3.

#### III. METHODOLOGY

In order to estimate the data, the formula for finite mixture model and maximum likelihood estimation is provided in this section. The general formula for mixture of normal distributions is

$$f(x_t) = \pi \varphi_1(\mu_1, \sigma_1^2) + (1 - \pi) \varphi_2(\mu_2, \sigma_2^2)$$
(1)

where  $\varphi_i(\mu_i, \sigma_i^2)$  denotes the probability density function of a normal distribution with mean  $\mu_i$  and variance  $\sigma_i^2$ . Meanwhile,  $\pi$  represents the weight of the normal distribution which is the first regime. Since the total weight for a component is equal to one, therefore, the weight of the second regime is  $1 - \pi$ .

Then, maximum likelihood estimation is applied in this paper to fit the mixture of normal distribution. The general formula for maximum likelihood estimation is as follows.

By assuming that  $X_1, ..., X_n$  are independent data, then the joint density function for all observations have to be specifies and the formula for joint density function is

$$f(x_1,...,x_n|\theta) = f(x_1|\theta) \times f(x_2|\theta) \times ... \times f(x_n|\theta)$$
<sup>(2)</sup>

However, if fixed parameters are provided for this function, the function will be known as likelihood function and the formula is

$$L(\theta|x_1,...,x_n) = f(x_1,...,x_n|\theta) = \prod_{i=1}^n f(x_i|\theta)$$
(3)

In general, the likelihood function is more convenient to work with the logarithm, called the log-likelihood.

$$\operatorname{In}L(\theta|x_1,...,x_n) = \sum_{i=1}^n \operatorname{In} f(x_i|\theta)$$
(4)

or the average log-likelihood

$$\hat{l} = \frac{1}{n} \ln L \tag{5}$$

where the hat over l denotes the expected log-likelihood of observation in the model.

### IV. RESULTS AND DISCUSSIONS

Before the analysis being started, a descriptive statistics table is displays to describe the basic information for the return value of nominal monthly exchange rate for each country. Information such as mean, standard deviation and skewness for the return of exchange rate for Malaysia, Thailand and Philippines are listed in table 1 meanwhile figure 1 display the histogram with normal curve of return value for exchange rates for Thailand, Malaysia and Philippines.

TABLE I. Summary Statistics

	Country					
Statistics	Malaysia	Thailand	Philippines			
Mean	-0.002361	-0.003426	-0.003446			
Standard Deviation	0.01434	0.01569	0.01540			
Skewness	0.233	0.453	0.324			



Fig. 1. Empirical Distributions of exchange return rates

By referring to the table 1 and figure 1, it can be concluded that the mean value for Malaysia is the largest in compared to Thailand and Philippines. The reason that Malaysia has highest mean value is because the value of currency in Malaysia is more valuable than others countries. However, the value of currency for Philippines is the lowest among these three countries; therefore, it has the smallest mean value.

Furthermore, table 1 is listed out all the standard deviation for the selected South Asia countries. Thailand has the largest standard deviation than others countries which indicates that it has a largest dispersion or more volatile. While Malaysia has the smallest standard deviation which illustrate that Malaysia has the smallest dispersion on the exchange return rate or less volatile than other countries.

Table 1 also listed the skewness of the distribution. A positive value is found on the skewness for all three countries which indicates that the exchange rate in these three countries can be altering by the changing of the United States currency. Moreover, all three countries are skewed to right which means that the distribution for Malaysia, Thailand and Philippines are positively skewed and this can be refer to the figure 2.

In this paper, a two-component mixture normal distribution is estimated by using maximum likelihood estimation and the normal distribution that mention in table 2 refers to recession state and growth state. The recession state in this study also known as first normal meanwhile the growth state represents the second normal. Two tables are provided which include a parameter estimates for normal model table and a mixing probabilities table with five effective parameters and the result in these tables are summarize in the table 2. Table 2 described the summary result that analysis by using maximum likelihood estimation. Meanwhile figure 2 denotes the mixture distribution graph for return of exchange rate to Malaysia, Thailand and Philippines.

	IADLE II.						
Maximum Like	lihood Estimation for	r normal mixture					
	Malaysia						
Country	Normal 1	Normal 2					
Weight	0.9437	0.0563					
Mean	-0.0041	0.0266					
Variance	0.0002	0.00003					
-							
Thailand							
Country	Normal 1	Normal 2					
Weight	0.5783	0.4217					
Mean	-0.00114	-0.00657					
Variance	0.0001	0.0004					
C I	Philip	opines					
Country	Normal 1	Normal 2					
Weight	0.0737	0.9263					
Mean	-0.0197	-0.0022					
Variance	7.30E-7	0.0002					

TADIEII

Both the table 2 and figure 2 described that the weight for Thailand and Malaysia in first normal shows that heavier than second normal and this indicates that the recession state is more dominant than growth state for Thailand and Malaysia. However, the economy in Philippines not affected by those event because the growth state in Philippines is more dominant than recession state.

Moreover, the mean for both Thailand and Philippines shows negative value for both recession state and growth state. This can be explained that the currency in Thailand is declining because the economy is not recovered from the impact of September 11 attacks. For Philippines, the negative mean value in both normal densities denotes that the currency in Philippines is declining because of the global economics crisis which beginning at 2008. However, Malaysia shows different mean value in compared to Thailand and Philippines. Malaysia has negative value in recession state and positive value in growth state. This indicates that the economy in Malaysia is recovered from the crisis.

Although variance is a not necessarily tool in measure the up and down swings of the bankroll goes through but a large value of variance represents a wider swings of bankroll. Result in table 2 show that all variance for normal 1 and 2 for Thailand, Malaysia and Philippines are small thus it can be concluded that the bankroll swing in these three countries are less. This is reliable since all these countries are label as developing countries so there might have a limited or less bankroll for the country. In addition, the result is valid, reliable and significant since the variance is small.









#### V. CONCLUSION

This paper applied maximum likelihood estimation to fit finite mixture model in order to estimate returns of nominal monthly exchange rate for Malaysia, Thailand and Philippines. A two-component mixture of normal distribution is analysing in this paper and the first normal is categorized as recession state while second normal denotes as growth state for the nominal exchange rate data. Result shows that the economy in Malaysia has recovered, but not for the economy Thailand and Philippines.

#### ACKNOWLEDGMENT

This paper is supported by the financial and facilities from the respective university.

#### REFERENCES

- S. Newcomb, "A generalized theory of the combination of observations so as to obtain the best result," *American Journal of Mathematics*, vol. 8, pp.343-366, 1886.
- [2] K. Pearson, "Contributions to the theory of mathematics evolution," *Philosophical Transactions of the Royal Society of London*, vol. A185, pp.71-110, 1984.
- [3] C. V. L. Charlier and S. D. Wicksell, "On the dissection of frequency functions," *Arkiv for Mathematics Astronomi och Fysik*, vol. 18(6), pp. 85-98, 1924.
- [4] G. Doetsch, "Die elimination des dopplereffekts auf spektroskopische feinstrukturen und exakte bestimmung der komponenten," Zeitschrift f`r Physik, vol. 49, pp. 705-730, 1928.
- [5] R. A. Fisher, "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society of London*, vol. 222, pp. 309-368, 1992.
- [6] C. R. Rao, "The utilization of multiple measurements in problems of biological classification," *Journal of the Royal Statistical Society*, vol. B10, pp. 159-203, 1948.
- [7] V. Hasselblad, "Estimation of parameters for a mixture of normal distributions," *Technometrics*, vol. 8, pp. 431-444, 1966.
- [8] B. C. Peters and W. A. Coberly, "The numerical evaluation of the maximum likelihood estimate of mixture proportions," *Communications in Statistics-Theory and Methods*, vol. 5, pp. 1127-1135, 1976.
- [9] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
- [10] D. W. Hosmer, "A comparison of iterative maximum likelihood estimates of the parameters of a mixture of two normal distributions under three different types of sample," *Biometrics*, vol. 29, pp. 761-770, 1973.
- [11] N. J. Horton and N. M. Laird, "Maximum Likelihood Analysis of Generalized Linear Models with Missing Covariates," *Statistical Methods in Medical Research*, vol. 8, pp. 37-50, 1998.
- [12] Z. H. Yang, "Maximum Likelihood Analysis Of Adaptive Evolution In Hiv-1 Gp120 Env GenE," Pacific Symposium on Biocomputing, pp. 226-237, 2001.
- [13] S. Castiglioni, L. Lombardi, E. Toth, A. Castellarin, and A. Montanari, "Calibration of rainfall-runoff models in ungauged basins: A regional maximum likelihood approach," *Advances in Water Resources*, vol. 33, pp. 1235–1242, 2010.
- [14] P. Owolawi, "Raindrop Size Distribution Model for the Prediction of Rain Attenuation in Durban," *Piers Online*, vol. 7(6), pp. 516-523, 2011.

# Variable Twist Angle of Flexible Electromagnetic Hyper Redundant Robot

Amir Sharizam Ismail, Samsi Md Said, Ishkandar Baharin

**Abstract**— Denavit Hartenberg (D-H) Kinematic structure representation of robot having revolute joint normally have fix arm length a, offset distance d, twist angle  $\alpha$  and only varied joint angle  $\theta$ . This paper introduce two (2) inputs variable of twist angle  $\alpha$  and joint variable  $\theta$  besides fix value of arm length a and offset distance d. Variable value of  $\alpha$  and  $\theta$  are dependent on polar electromagnetic actuator actuation sequence. The actuator extension and their combination will determine the joint variable value of joint angle  $\theta$  and twist angle  $\alpha$ . This electromagnetic muscle is arranged in four (4) polar arrays around disk body. Even though the electromagnetic muscles are classified as prismatic joint, their sequence and combination of operation producing two joint variables of joint angle  $\theta$  and twist angle $\alpha$ . Output from this forward kinematic model is the end effector *NOAP* matrix and its yaw, pitch and roll orientation representation. This novel module is intended for building a high assemblage of Hyper Redundant Robot. A low cost polycarbonate plastic and spring are used for structure of flexible body module. Coil of more than 1000 *turns* is employ in order to generate distance extension. The Programmable Logic Controller (PLC) is used to manage the sequence of energizing the coils for having value of  $\alpha$  and  $\theta$ .

\_\_\_\_\_

Index Terms— Snake Robot, Low Cost, Flexible Body, Kinematic Model

### **1** INTRODUCTION

DENAVIT Hartenberg [1] representation of single degree of freedom robot requires four parameters such as arm length *a*, offset distance *d*, twist angle  $\alpha$  and joint angle  $\theta$ . Arm length *a* is the arm extension distance between adjacent joint, offset distance d is the distance between the origin of two neighbouring coordinate frame along Z reference axis, twist angle  $\propto$  is the angular different between Z axis of both joints and joint angle is the angular displacement between X axes measured about reference Z axis of the specified degree of freedom. Normally joint angle  $\theta$  become joint variable for revolute joint and offset distance d is a joint variable for prismatic joint [2].

This kind of representation is used to model industrial robot successfully from 1950's. Since the current trend of robotic research is diversifying into service, humanoid and biomimetic robot which having large numbers of degree of freedom, a new approach is required to serve this computational intensive requirement. Using Denavit Harternberg requires four parameters to describe a single degree of freedom or joint. In hyper redundant robot design normally requires more than 30 degree of freedom [3]. Refer to Chirikjian and Burdick [4], the term hyper-redundant is to describe robots which have a very large number of independent degrees of freedom. A simple approach is needed in order to model just one module or segment of snake robot.

Since D-H algorithm is an established and standard method to model any robot, a modification of D-H system is required for high degree of freedom robot. Maintaining D-H approach with novel simplification will retain the robustness of robot modeling while reducing the computational burden in hyper redundant robot design. This simplified approach will accommodate four degree of freedom into a single representation of Denavit Hartenberg algorithm.

In this work a robot module or segment consist of four prismatic joints is being developed. The combination of this module will become an assemblage for multiple degree of freedom of hyper redundant robot or snake robot. Normal approach is to treat a single prismatic joint as a single degree of freedom with four Arm matrices. A unique approach in this research is to model this four degree of freedom into single arm matrix representing a segment or a module. The actuator extension and their combination will determine the joint variable value of joint angle  $\theta$  and twist angle  $\alpha$ . This electromagnetic muscle is arranged in four (4) polar arrays around disk body. Even though the electromagnetic muscles are classified as prismatic joint, their sequence and combination of operation producing two joint variables of joint angle  $\theta$  and twist angle  $\alpha$ . Output from this forward kinematic model is the end effector NOAP matrix and its yaw, pitch and roll orientation representation. This novel module is intended for building a high assemblage of Hyper Redundant Robot. Employing this actuation sequence representation technique 40 DOF system requires only ten arm matrices in contrast to normal modeling of using 40 arm matrices.

Amir Sharizam Ismail is currently pursuing master's degree program in control & robotic engineering technology in Universiti Kuala Lumpur, ID.51311212118. E-mail: amirsharizam@mfi.unikl.edu.my

<sup>•</sup> Samsi Md Said is co-supervisor and a senior lecturer at Universiti Kuala Lumpur Malaysia France Institute, Email: samsi@mfi.unikl.edu.my

Ishkandar Baharin is a member of IFR Industry Associations and Professor at Universiti Kuala Lumpur Malaysia Type equation here.France Institute, E-mail: ishkandar@mfi.unikl.edu.my

### **2** PROTOTYPE OF ROTATING SEGMENT

### 2.1 Actuator Description

The actuator is cylindrically symmetrical configuration shown in Figure 1. It consists of three polycarbonate plastic layers, a spring and four armatures in polar array.



It is a special design electromagnet that consists of a coil and a movable iron core called the *armature*. When current flows through a wire, a magnetic field is set up around the wire. When the coil of the solenoid is energized with current, the core moves to increase the flux linkage by closing the air gap between the cores. The movable core is usually springloaded to allow the core to retract when the current is switched off. The force generated is approximately proportional to the square of the current and inversely proportional to the square of the length of the air gap. This high-speed electromagnetic actuator is highly consistent of high pressure application. Therefore, the new actuator can provide high response over short strokes [5] and is suitable for variable twist angle of flexible electromagnetic hyper redundant module.

### 2.2 Programmable Logic Controller

A Programmable Logic Controller (PLC) is a digital computer used for automation of electromechanical processes [6]. A PLC is an example of a hard real time system since output results must be produced in response to input conditions within a limited time, otherwise unintended operation will result.

For this work, we used PLC Omron CP1A-CPU20 with 8 bit input and 8 bit output relay. Table 1 is the specification of the PLC that we used:

PLC Type: Compact Voltage Supply: 100VAC-240VAC Model number: CQM1H-CPU51 I/O capacity: 512 points Program capacity: 7.2 Kwords DM Area size: 6 Kwords CPU Unit built in I/O points: 16 DC Inputs Communication ports: Peripheral/ RS-232C Inner board: Supported Communication Units: Supported

We controlled the sequence of the four bits to make variation

of twist angle  $\alpha$ . Figure 2 and Table 2 shown the inputs and outputs wiring and list of input and output PLC.

To operate the sequence, we programmed the PLC using ladder language to active these four actuators of a module.



Figure 2: Wiring I/O PLC for One Segment

1 able 2. List of input / Output I LC	Table 2:	List	of I	nput /	Output	PL	C
---------------------------------------	----------	------	------	--------	--------	----	---

Item	Description	PLC i/o Address
S1	Switch 1 Push Button (N/O)	000.00
S2	Switch 2 Push Button (N/O)	000.01
S3	Switch 3 Push Button (N/O)	000.02
S4	Switch 4 Push Button (N/O)	000.03
EMA1	Electromagnetic Actuator 1	100.00
EMA2	Electromagnetic Actuator 2	100.01
EMA3	Electromagnetic Actuator 3	100.02
EMA4	Electromagnetic Actuator 4	100.03

### **3** OPERATION OF ROTATING TWIST ANGLE MODULE

### 3.1 Denavit Hartenberg Modeling

The module consists of reference disk and orientation disk. The XYZ axis is attached on the reference disk while the NOAP orientation axis is at the outer disk (Refer Figure 2). The difference between origins of these two frames along X axis is arm length *a*. The offset distance *d* is zero since theoretically there is no displacement among the frame's origin along Z axis. Twist angle  $\alpha$  is a product of prismatic actuator combination. The similar value of angular variable  $\theta$  is achieved with varied value of twist angle  $\alpha$ . The bang-bang control scheme is employed in electromagnetic actuator since the objective of this work to show the effect of twist angle variation on *NOAP* homogeneous orientation and position matrix. In future, servo control for electromagnetic extension will be explored in order to give varied value of joint angle  $\theta$  and twist angle  $\alpha$ .

### 3.2 Variable Twist Angle, $\alpha$

The combination of these four electromagnetic actuator will determine the twist angle  $\alpha$ . Twist angle  $\alpha$  is the angle between Z reference axis and Z transition axis about X transition axis. By energizing actuator number one, the frame angular displacement is about 90°. If both actuator 1 and 2 is being energized the transition frame will rotate further to become 135°. Refer Figure 3 and Table 3 for complete transition axis rotation and their derived twist angle  $\alpha$ .



Figure 3: Rotating Frame of Z-Axis

Denavit Harternberg modeling consist of four sequence from (1) to (4) of transformation axes.

Where  $Trans[z_{i-1}, d_i]Rot[z_{i-1}, \theta_i]$  is the  $z_{i-1}$  screw motion and  $Trans[n_i, a_i]Rot[n_i, \alpha_i]$  is the  $x_i$  screw motion.

$$(1)T_{z,\theta} \to (2) T_{z,d} \to (3) T_{n,a} \to (4) T_{n,\alpha} \tag{1}$$

Only the last transformation related to twist angle is being modified to simplified overall segment representation. This four transformation being multiplied to become arm matrix A.

$$A_i = T_{z,a} T_{z,\theta} T_{n,a} T_{n,\alpha} \tag{2}$$

Value of  $\theta$  can is related to actuator extension (in mm) as equation (3)

$$\theta = \tan^{-1} \left[ \frac{2\Delta_X}{D} \right] \tag{3}$$

Where  $\Delta_x$  is the positive or negative incremental extension of linear electromagnetic actuator and D is diameter disk segment. The combination four bit control will determine the value of twist angle  $\alpha$  as matrix (4).

$$T_{w}(bit, bit, bit, bit)_{ni,\alpha i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha_{i} & -\sin \alpha_{i} & 0 \\ 0 & \sin \alpha_{i} & \cos \alpha_{i} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(4)

The state of bit 1 and bit 2 as in table 1 will determine the value of twist angle  $\alpha$  and the related matrix (5)

$$T_{W}(1,1,0,0)_{135}^{\circ} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -.707 & 0.707 & 0 \\ 0 & 0.707 & -0.707 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(5)

Table 3 show the total twist angle generation and their bit's

$$T_{\mathcal{W}}(1,0,0,0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

### sequence

$$T_{W}(1,1,0,0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -.707 & 0.707 & 0 \\ 0 & 0.707 & -0.707 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(7)  
$$T_{W}(0,1,0,0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(8)  
$$T_{W}(0,1,1,0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.707 & 0.707 & 0 \\ 0 & -0.707 & -0.707 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(9)

$$T_{w}(0,0,1,0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(10)  
$$T_{w}(0,0,1,1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.707 & 0.707 & 0 \\ 0 & -0.707 & 0.707 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(11)  
$$T_{w}(0,0,0,1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(12)

Table 3: D-H Frame Assignment with Variable Twist Angle

Sequence of Operation (ON=1/OFF=0)							
1	1	1	0	0	0	0	0
2	0	1	1	1	0	0	0
3	0	0	0	1	1	1	0
4	0	0	0	0	0	1	1
α i	90°	135°	180°	225°	270°	315°	360°
d <sub>i</sub>	0	0	0	0	0	0	0
$\theta_{i}$	15°	15°	15°	15°	15°	15°	15°
a <sub>i</sub>	25	25	25	25	25	25	25
$cos(\alpha_i)$	0	-0.707	-1	-0.707	0	0.707	1
$sin(\alpha_i)$	1	0.707	0	-0.707	-1	-0.707	0

The arm matrix Ai is as equation (13).

$$A_{i} = Rot_{z, \ell i} Trans_{z, d i} Trans_{n, a i} Tw(bit, bit, bit, bit)_{n, c i}$$
(13)

These arm matrixes normally represent single degree of freedom but in this work it represents the whole single module of robot,  $A_i = M_i$ .  $M_i = Rot_{z,\theta_i} Trans_{z,d_i} Trans_{n,a_i} Tw(bit,bit,bit,bit)_{n,\alpha_i}$ 

$$T_i^n = \prod_{i=1}^n M_{i-1}^i M_i^{i+1} \dots M_{i+1}^n$$
(15)

(14)

$$T_{i}^{n} = \begin{bmatrix} n_{x} & o_{x} & a_{x} & P_{x} \\ n_{y} & o_{y} & a_{y} & P_{y} \\ n_{z} & o_{z} & a_{z} & P_{z} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(16)

The position and orientation representation of the *Yaw*  $\phi$ , *Pitch* $\theta$ , and *Roll*  $\psi$  for this twist module can be derived from matrix (16) as in (17) to (19).

 $\phi = atan2(n_y, n_x) \tag{17}$ 

$$\theta = atan2(-n_z, \cos\phi, n_z + \sin\phi, n_y)$$
(18)  
$$\psi = atan2(\sin\phi, a_x - \cos\phi, a_y, -\sin\phi, o_x + \cos\phi, o_y)$$
(19)

$$\psi = \operatorname{atan2}(\operatorname{sin\phi}.a_x - \cos\phi.a_y, -\sin\phi.o_x + \cos\phi.o_y) \tag{19}$$

### **4** RESULT ANALYSIS

Consider the combination of individual module into a hyper redundant robot system represented symbolically in Figure 4. The configuration can be scaled into any number of N degree of freedom. Tables 4 summarize the Denavit Harternberg table for the total segment or module. The result for every individual segment and their incremental value is displayed in Table 5 and Table 6.



Figure 4: Three-Link Segment with Twist Angle,  $\alpha = 90^{\circ}$ 

Table 3: D-H Parameters for Six Segments

Tw(1,0,0,0)							
Link	a <sub>i</sub>	$\alpha_i$	$d_i$	$\theta_{i}$			
1	25	90°	0	15°			
2	25	90°	0	15°			
3	25	90°	0	15°			
4	25	90°	0	15°			
5	25	90°	0	15°			
6	25	90°	0	15°			

Table 5: Result NOAP for Six DOF

	Variable Twist Angle, $\alpha = 90^{\circ}$ for each joint module							
	Position (mm)				Orientation (°)			
Module	P <sub>xi</sub>	P <sub>yi</sub>	$ \begin{array}{c c} P_{yi} & P_{zi} & Yaw, \emptyset & Pitch, \theta \\ atan2(n_2 & atan2(-n_{zi}, cos \emptyset) \\ + sin \emptyset. n_{yi}) \end{array} $		Pitch, $\theta$ $atan2(-n_{zi}, cos \phi)$ $+ sin \phi . n_{yi})$	$ \begin{array}{c} \text{Roll, } \psi \\ atan2(sin\emptyset. a_{xi} \\ -cos\emptyset. a_{yi}, -sin\emptyset. o_{xi} \\ +cos\emptyset. o_{yi}) \end{array} $		
J 1	24.1	6.5	0.0	15.0	0.0	90.0		
J 2	47.5	12.7	6.5	15.0	-39.4	180.0		
J 3	71.7	12.5	12.7	-0.5	-45.0	-86.0		
J 4	96.7	12.7	12.5	0.5	134.7	3.8		
J 5	120.8	19.4	12.7	15.5	-6.5	93.8		
J 6	144.1	25.4	19.4	15.5	-6.5	-176.3		

### Table 6: Result NOAP for Six DOF

	Variable Twist Angle, $\alpha$ = 360° for each joint module								
	Position (mm)			Orientation (°)					
Module P <sub>xi</sub>		P <sub>yi</sub>	P <sub>zi</sub>	Yaw, Ø atan2(n <sub>y</sub>	Pitch, $\theta$ $atan2(-n_{zi}, cos \emptyset$ . $+ sin \emptyset. n_{yi})$	$ \begin{array}{l} \text{Roll, } \psi \\ atan2(sin\emptyset. a_{xi} \\ -cos\emptyset. a_{yi}, -sin\emptyset. o_{xi} \\ +cos\emptyset. o_{yi}) \end{array} $			
J 1	24.1	6.5	0.0	15.0	0.0	0.0			
J 2	45.8	19.0	0.0	30.0	0.0	0.0			
J 3	63.5	36.6	0.0	45.0	0.0	0.0			
J 4	76.0	58.3	0.0	60.0	0.0	0.0			
J 5	82.4	82.4	0.0	75.0	0.0	0.0			
J 6	82.4	107.4	0.0	75.0	0.0	0.0			



Figure 5: Gait Motion for Variable Twist Angle  $\alpha$  at XY Axis



Figure 6: Gait Motion for Variable Twist Angle  $\alpha$  at XZ Axis

Gait Motion with Variable Twist Angle



Figure 7: XYZ View

**5** CONCLUSION

It is shown that spiral configuration of hyper redundant or snake robot can be built by adding rotation of 45° into variable twist angle in every module such as in Figure 8. This spiral shape is very important for gait motion [7] intended to climb tree or any round object. Using the twist angle value of 360° v will give planar curve of snake robot configuration useful for

### Denavit-Hartenberg (D-H) is an accepted robot modeling technique since 1950's. It is being used successfully to model Six degree of Freedom modern industrial robot in industries. Since the current trend of robotic is to move toward the service and military application which requires robot in great number of degree of freedom deviate from a standard six de-

serpentine motion useful surface robot locomotion.







Figure 9: Four Type of Angle of Flexible Hyper Redundant Robot with Variable Twist Angle α

gree of freedom, a new mathematical modeling is required.

This work shows the improvement of Denavit - Hartenberg usage in robot modeling which reduce the computational burden. Normally one degree of freedom robot requires one arm matrices as proposed by D-H method. This research proposes the compression of four degree of freedom robotic system into single module M. The robustness of Denavit Hartenberg is retained while improving the computational efficiency. The outcome of this work is lay foundation for the new robotic era that mimic biological creature such as snake robot which posses a high number of degree of freedom. The spiral gait motion for climbing or planar surface locomotion can be achieved easily by employing the incremental twist angle or constant value of 360° of twist angle.

### REFERENCES

- J Denavit, R S Hartenberg," A Kinematic Notation for Lower-Pair Mechanisms based on Matrices", Journal of Apllied Mechanics, pp. 215-221, 1955.
- [2] Norton, Robert, "Design of Machinery (4<sup>th</sup> Edition)".Boston, MA: McGraw Hill Higher Education, pp. 33. ISBN 978-0-07-312158-1.
- [3] Chirikjian, G S, Burdick, J W. "Design and Experiments with 30 DOF Robot". IEEE, Robotics and Automation, pp. 113-119, Vol. 3, 1993.
- [4] Chirikjian, G S, Burdick, J W. "A Hyper-Redundant Manipulator". IEEE, Robotics and Automation Magazine 1(4), pp. 338-341.
- [5] Jun Man et al "Novel High-Speed Electromagnetic Actuator With Permanent-Magnet Shielding for High-Pressure Applications" IEEE Transaction on Magnetics, Vol. 46, No. 12, December 2010.
- [6] E.A. Parr." Industrial Control Handbook, Industrial Press Inc. 1999, ISBN 0-8311-3085-7.
- [7] Gray, J. "The Mechanism of Locomotion in Snakes". Journal of Experimental Biology, 23(2), pp. 101-120.
- [8] S. Hirose."Biologically Inspired Robots: Snake-Like Locomotors and Manipulators". Oxford University Press, 1993.
- [9] Burdick, J.W.; Radford, J.; Chirikjian, G.S.; "A 'sidewinding' locomotion gait for hyper-redundant robots," Robotics and Automation, 1993.
- [10] Kevin J. Dowling, "Limbless Locomotion: Learning to Crawl with a Snake Robot", PhD Thesis, The Robotics Institute, Carnegie Mellon University, December 1997.
- [11] Kevin J Dowling, William L. Whittaker." *Limbless Locomotion: Learning to Crawl with a Snake Robot*". Robotic and Automation, 1999.
- [12] Hartenberg, Richard Scheunemann; Denavit, Jacques, "Kinematic Synthesis of Linkages", McGraw-Hill series in Mechanical Engineering. New York. McGraw-Hill, pp 435.
- [13] T. Kajima and Y. Kawamura, "Development of a high-speed solenoid valve: Investigation of solenoids," IEEE Trans. Ind. Electron., vol. 42, no. 1, pp. 1–8, Feb. 1995.
- [14] J. Kim and J. Chang, "A new electromagnetic linear actuator for quick latching," *IEEE Trans.Magn.*, vol. 43, no. 4, pp. 1849–1852, Apr. 2007.

# Experimental Study in the Process Parameters in Laser Percussion Drilling

B. B. Satapathy, J. Rana, K. P. Maity, S.Biswal

Abstract – In order to service and grow in the present day of global competition, every manufacturing concern focuses mostly on two factors i.e. productivity and quality. Laser drilling is a popular non-traditional machining technique for producing large numbers of cooling holes of various sizes (less than mm diameter) and angles in modern aerotech components such as turbine blades, injector nozzles, and combustion chambers. Although the productivity of microhole in very high, the quality of hole (straightness, circularity, aspect ratio, heat affected zone, micro-crack spatter rate etc.) is poor due to unique nature of the process. In the present paper experimental investigation has been carried out using Taguchi's L9 orthogonal array technique to optimize the main laser and process parameters to get a quality hole in case of in plane carbon steel specimen. Pulsed Nd: YAG laser beam has great ability for micro-machining of plane carbon steel materials because of high laser beam intensity at low mean beam power, good focusing characteristics due to very small pulse duration and less heat effected zones. The quality characteristics such as aspect ratio (depth / diameter), heat affected zone, spatter deposition, hole circularity are studied though scanning electron micro-scope (SEM). The specific advantages of this Taguchi's technique of optimization are that with a very less number of experiments, optimization is possible.

Index Terms: Aspect ratio, Drilling rate Hole circularity, HAZ, Laser drilling ,Optimization, , Taguchi's technique,

### 1 INTRODUCTION

The automobile, modern aerospace manufacturing units has been using Laser drilling to produce cooling holes of various sizes in turbine blades, nozzles guide vanes, combustion chamber etc. The high intensity laser beam falls on a very small area, the target materials gets heated, melted and vaporized due to the pressure of assist gas, the mechanism of material removal involves both vaporization and melt ejection resulting in conical hole with lot of spatter deposition on the hole periphery at the top surface. The word laser is an acronym for 'Light Amplification by Stimulated Emission of Radiation'. It has the capability to deliver the energy in order of 106 W/cm<sup>2</sup>, and can be varied over a wide range of temporal and spatial distribution. Moreover, the laser light is highly coherent and monochromatic in nature. Materials processing makes the use of the thermal and photonic effect associated with the interaction of laser beam with various engineering materials. Being a noncontact, inertia less tool, laser has tremendous potential in improving the processing speed.

- Asst. Professor, Mech. Engg. Department, P.K.A.College of Engg.India.cell no-09437201161
   e.mail:- <u>bbs\_pkace@gmail.com</u>
- Professor and Head, Mech. Engg. Department, V.S.S.U.T India.
- Professor and Head, Mech. Engg. Department, N.I.T, India..
- Manager (Mechanical), M &MHIC, L & T ECC Div. India.



Fig.1. Laser drilling set-up.

The laser machining is based on the interaction of laser light with the outer most atoms of solid matter (opaque). Because of complex process, a small amount of material is removed from the surface of the solid. In both the cases, short to ultra short laser pulses when applied to the solid body, a small amount of material is removed in a controlled way. Keeping in view of the potentiality of the laser drilling process, the present work has been planned in this area. Low et al [1] studied the Characteristics of spatter formation under the effect of different laser parameters during laser drilling. Guo et al [2] developed spatter free laser drilling of alumina ceramics based on gel casting technology and new laser micromachining technique using a mixed mode ablation. Wang et al [3] experimentally studied the effect of assist gases such as oxygen, argon and nitrogen on the speceman. Schoonderbeek et al [4] micro structures of recast layer and reported that these are independent studied the influence of pulse width on the quality of hole using a excimer laser. Nedialkov et al [6] investigated "Laser drilling of AlN ceramics using nanosecond pulses, also similar type of work are found in [5] N.N Nedialkov et al.

Campbell et al [7] analyzed the ultra short pulse laser micromachining parameters for optimization of shallow hole drilling. Naeem, et al [8] developed "Laser percussion drilling of aerospace materials using high peck power fiber delivered 1 Amp – Pumped pulsed Nd – YAG Laser. Ghoreishi et al [9] analyzed "Optimization of effective factors in geometrical specifications of laser percussion drilled hole. The present paper deals with experimental investigation of laser and process parameters using SEM analysis for quality optimizations such as hole circularities, heat affected zone and aspect ratio, spatter deposition on a medium carbon steel specimen. G. Taguchi, [10] "Introduced Quality Engineering the appropriate S/N ratio for these quality characteristics (circularity) is "smaller the better" and" higher the better.

### 2 EXPERIMENTAL DETAILS:

A steel plates of 50 mm long, 15 mm wide and 5 mm thick were used for the drilling operation. The main factors (air flow rate, pulse frequency, pulse width) are considered to characterize laser-drilled hole. The specimen was mounted on the platform with the arrangement to regulate the flow of assist gas and the laser beam was focused with a lens of focal plane as shown in Fig.1. A set of holes were drilled on one longitudinal surface of the specimen following Taguchi's L<sub>9</sub> orthogonal array. Three controllable parameters such as pulse width, pulse frequency and air flow rate were varied as presented in Table 1.

SI. No	Parameters	Unit	Low level	Medi- um level	High level
1	Pulse width	μs	500	700	900
2	Pulse frequency	S <sup>-1</sup>	1	2	3
3	Air flow rate	lit/min	5	15	25

When a laser beam is focused in to the work piece surface, a portion of the beam is absorbed on the surface and some part is reflected back which depends upon the nature of surface finish and wave length of laser beam. The photons of absorbed laser beam interact with the outer most atom of the work piece surface and heat is developed. Because of the focused and high energy density of the incident laser beam, a portion of the work piece material gets melted and vaporized. In principle, laser drilling is governed by an The same optimum result can also be obtained by only 9 set of experiments following Taguchi's L<sub>9</sub> orthogonal array. The Taguchi's method requires both analysis of mean response (output value) and analysis of variance using appropriately chosen "signal to noise (S/N) ratio" derived from a quadratic loss function. In the present experimental work, three output parameters i.e. hole circularity; aspect ratio and drilling rate were studied. The error of hole circularity is decided by the difference of radial distance between the minimum and maximum inscribing circles. That means smaller the difference, better will be the circularity. Therefore, the appropriate S/N ratio for these quality characteristics (circularity) is "smaller the better" and is given by:

energy balance between the irradiating energy from the laser beam and the conduc tion of heat into the work piece, the energy losses to the environment and the energy required for phase change in the work piece. The mechanism of the laser drill- ing process is explained with the help of the Fig.1(a).



Fig.2. Mechanism of Laser Drilling.

$$\eta_{LB} = -10 \log_{10} \left( \frac{1}{r} \sum_{i=1}^{r} y_i^2 \right)$$

For a better quality hole, the aspect ratio (depth/diameter) should be as high as possible. Similarly for higher productivity, the drilling rate should be higher. Therefore, the appropriate S/N ratios for these two quality characteristics are "higher the better" and is given by:

$$\eta_{HB} = -10 \log_{10} \left( \frac{1}{r} \sum_{i=1}^{r} \frac{1}{y_i^2} \right)$$
The various steps in this technique are:

- The output parameters to be optimized.
- The levels of controllable and uncontrollable parameters are idetified.
- The orthogonal array matrix is prepared.
- Experiments are performed as per orthogonal array matrix.
- Using the experimental observation, average output parameters for each set of process parameters and appropriate (S/N) signal to noise ratios are determined.
- The data are analyzed through a set of graphs to obtain the optimum level combination of parameters.

#### 3. RESULTS AND DISCUSSION.

The present study involved investigation of various characteristic associated with pulsed Nd-YAG laser drilling using processing of the metal. After the drilling operation, micrographs of the top surface of the holes were taken using scanning electron microscope (SEM) to show the circularity, heat affected zone .A transverse section is made near the diametral plane of the micro-hole after polishing the cut surface of the micro hole, the different parameters such as hole diameter, thickness of Heat affected zone, approximate depth etc. were measured under a scanning electron microscope. The side view of few micro holes is also presented. The measured parameters are presented in Table.2 and average values of output parameters such as aspect ratio (depth/diameter) thickness of heat affected zone and corresponding S/N ratio are calculated and presented.



Fig. 3: SEM photograph at process parameters

Pulse width =  $500 \ \mu$ s, Number of pulse per second =3, average power =  $1.7 \ W$ , Air flow rate =5 lit/min, time =  $20 \ sec$ .

Pulse width = 700 μs, Number of pulse per second = 3,average power = 2 W, Airflow rate = 15 lit/min, time = 20 sec

Pulse width = 900  $\mu$ s, Number of pulse per second = 3, average power = 3 W, Airflow rate = 25 lit/min, time = 20 sec.



Fig 4 .Super imposition of 9 different predicted hole contours for comparision .

Fig.4 shows the different isotherm by changing parameters and all the 9 isotherm in a single graph respectively.

Table - 2: Process parameters and correspondingoutput with S/N values.

#### Journal of xxx & xxx, Volume 3, Issue 3, March-2012 ISSN XXXX-XXXX

12.41	4.173	225	33.43	46.95	939	-29.82	30.986	20	3.0	15	з	006	9
12.94	4.435	170	31.53	37.7	754	-35.78	61.49	20	25	U1	2	006	60
14.99	5.618	110	29.80	50.5	618	-35.42	59.043	20	15	25	1	006	7
13,47	4,713	150	30.97	3535	707	-9.81	3.093	20	2.0	5	3	700	6
14.63	5.387	150	32.13	40.4	808	-32.84	43.878	20	1.8	25	2	700	5
17.16	7.215	130	33.42	46.9	938	-37.45	74,528	20	1.6	15	1	700	4
11.26	3.658	155	29.05	28.350	567	-14.69	5.426	20	1.7	25	u	500	ω
6.47	2.107	300	29.99	31.6	632	-35.20	57.565	20	11	15	2	500	N
11.0	3.547	265	33,44	47.0	940	-20.9	11.18	20	8.0	5	1	500	1
(db)	)	Ē	rate (db)							(min)	Pulse/s)	-	
ratio	(depth/dia	diameter	drilling	(µm/s)	Ē	circularity (db)	(%)	(5)	Ì	Qit	(No.of	(LLS)	No.)
for aspec	ratio	hole	ratio for	rate	hole	error of hole	circularity	time	power	rate	Freq.	width	(Test
S/N ratio	Aspect	Average	SN	Drilling	Depth of	S/N ratio for	Error in hole	Drilling	Average	Air flow	Pulse	Pulse	SL No.

It clearly shows from the table-2 that hole of high aspect ratio is possible for a pulse width of  $500\mu$ s(lowest), pulse frequency of 3(highest) and with average power of 1.7 W. A good quality hole is simply

represented by higest aspect ratio. That means micro hole having less diameter with higher depth. It is observed from the Taguchi's parametric optimization technic that a deper and narrow hole can be possible by lesser pulse width and higher pulse frequency. The reason is obvious because, lesser pulse width does not allow much time for the heat to diffuse into the neighbouring metal matrix, and higher pulse frequency frequency favours higher vapoization/ melting rate compared to lower pulse frequency.

#### **4** CONCLUSIONS

The following conclusions are obtained from this experimental work:

- To obtain a best circular shaped hole, the optimum combination of process parameters is: pulse width, 500 μs; pulse frequency, 3; and assist gas flow rate, 5 lit/min.
- High productivity and high quality cannot be obtained simultaneously. One has to make a compromise between these two.
- The optimum combination of process parameters for high drilling rate and high aspect ratio is: pulse width; 700 μs, pulse frequency; 1 and assist gas flow rate; 15 lit/min.
- The hole taper varies widely which is obvious from the wide variations of aspect ratios. (i.e. from 2.107 to 7.215) Taguchi's parametric optimization technique is a simple and economic tool to decide the optimum combination of process parameters for obtaining best results from any experiment or process.

#### **REFERENCES.**

- D.K.Y. Low, L. Li and A. G. Corfe, "Characteristics of patter formation under the effect of different laser parameters during laser drilling", Journal of materials processing Technology, Vol. 118 (2001), 179-186.
- [2] D. Guo, K. Cai, J. Yang and Y. Huang, "Spatter free laser drilling of alumina ceramics based on gel casting technology", Journal of European Ceramic Society, 23 (2003) 1263-1267
- [3] X.Y. Wang, G.K.L. Ng, Z Liu, L. Li. And L. Bradley, " EPMA micro analysis of recast layers produced during laser drilling of type 305 stainless steel", Thin Solid film, 453 (2004) 84-88.
- [4] B.C.A. Schoonderbuk, R.M. Hofstra, K.J. Boller and J.Meijer, "The influence of the pulse length on the drilling of metals with an excimer laser", Journal of laser Application, 16, (2004) 85-91
- [5] N.N Nedialkov, P.A. Atanasov, M. Sawczak and Q. Slinski, "Laser drilling of AIN ceramics using nanosecond pulses", Proceeding of SPIE-The international society for optical engineering, (2005) 5777.
- [6] N.N. Nedialkov, M. Sawczak, M. Jadraque, P.A. Atanasov, M.Maictin and G. Slikinski, "Effect of laser drilling on surface and material properties of AIN ceramics", Proceeding of SPIE-The international society for optical engineering, (2005) 5958.
- [7] B.R. Compbell, V.V. Semak and J.G. Thomas "An analysis of ultra short pulse laser micromachining parameters for optimization of shallow hole drilling", Proceeding of SPIE-The international society for optical engineering, (2006) 6261.
- [8] M. Naeem, " Laser percussion drilling of aerospace materials using high peck power fibre delivered 1 Amp – Pumped pulsed Nd – YAG Laser", Laser Institute of America, 99(2006) 148-156
- [9] M. Ghoreishi and O.B. Nakhjavani, "Optimization of effective factors in geometrical specifications of laser percussion drilled hole", Journal of materials processing Technology, (2007)
- [10] G. Taguchi, "Introduction to Quality Engineering", UNI Publication, Newyork. (1986)

## Digital Prototyping: A Case Study on its Viability in Enhancing Small and Medium Ceramic Industries

#### Ab. Aziz Shuaib, Olalere Folasayo Enoch

(Only author names, for other information use the space provided at the bottom (left side) of first page or last page. Don't superscript numbers for authors ) **Abstract** Small and Medium Industries are important to almost all economies in the world, particularly to developing countries and, within that broad category, especially to those with major employment and income distribution challenges. They contribute to the output and creation of jobs; also they are a nursery for the large firms of the future. Hence, this paper seeks to explore further on findings from previous research work which experimentally investigates the viability of digital and rapid prototyping in small and medium scale ceramic industries. The findings from the research revealed that digital and rapid prototyping are viable in reducing development time and improving prototype accuracy; however, the research shows that rapid prototyping is an expensive approach and thus, proposed conventional method of prototyping (hand turning and carving) to be still valid in small and medium ceramic industries. Therefore, based on these findings, this further research seeks to explore how digital prototyping can be used to enhance conventional method of prototyping. Rather than replacing conventional method with digital and rapid prototyping can be enhanced by integrating it with digital prototyping (3D CAD model) into development process as this will help designers/carvers to visualize products digitally throughout product development process, thereby communicating the whole concept and idea being proposed. Also, this will build their creative skills and make them think "outside the box" thereby making them capable of producing innovative and exclusive prototypes/products.

Index Terms - Ceramics, Computer Aided Design, Digital and Rapid Prototyping, Hand Turning and Carving, Pottery, Small and Medium Industries.

\_\_\_\_\_

\_\_\_\_ **♦** 

#### **1** INTRODUCTION

mall and Medium Industries (SMIs) are important to almost all Deconomies in the world, particularly to developing countries and, within that broad category, especially to those with major employment and income distribution challenges. They contribute to output and creation of jobs; also they are a nursery for the large firms of the future. SMIs are, on average a good deal less complicated structurally than are corporations and other large firms. Small and medium ceramic industries are firms that manufactures structured clay products, e.g. pottery, clay roof and floor tiles, clay bricks, e t c [1]; however, this research will focus on pottery wares (such as table wares, kitchen wares, and decorative wares). Pottery products play a very important role in our daily life. This is because, apart from their decorative look, they are primarily hygiene products. This is one of the chief reasons for their wide usage in homes etc. Therefore, pottery is seen as a sector which is able to generate local community economy as well as nation's economy.

The pottery industries has a long history, with the first instance of functional pottery vessels being used for storing water and food, being thought to be around since 9,000 or 10,000BC. However, the industry has been modernizing continuously, by newer innovations in product design, quality etc. Hence, the sector has a multibusiness potential that are yet to be fully explored. Thus, this paper seeks to investigate how small and medium scale pottery industries can be enhanced with the application of digital prototyping. The research is a further study on findings from previous research work which experimentally investigates the viability of digital and rapid prototyping in small and medium scale ceramic industries. The findings from the research revealed that digital and rapid prototyping are viable in reducing development time and improving prototype accuracy; however, the research shows that rapid prototyping is an expensive approach and thus, proposed conventional method of prototyping (hand turning and carving) to be still valid in small and medium ceramic industries. Therefore, based on these findings, this further research seeks to explore how digital prototyping can be used to enhance conventional method of prototyping. Instead of replacing conventional method with digital and rapid prototyping, this research tried to enhance the conventional method by integrating it with digital prototyping. The findings from this further research revealed that the conventional method of prototyping can be enhanced by integrating digital prototyping (3D CAD model) into development process as this will help designers/carvers to visualize products digitally throughout product development process, thereby communicating the whole concept and idea being proposed and also help them to identify possible faults at the early stage of product development. Application of digital prototyping will also improve the prototype accuracy, build designers/carvers creative skills and make them think "outside the box" thereby enhancing them to be capable of producing innovative and exclusive prototypes/products.

## 2 Computer Aided Design

Computer Aided design (CAD) also known as Computer Aided Design and Drafting (CADD) is the use of computer technology for the process of design and design documentation [2]. The use of computer aided tools has impacted significantly in the execution of technological innovation in almost all spheres of industrial design products development. Computer aided designs (CAD) in ceramics production has been expressed through conceptualization of ideas to realization of processes in the production of ceramics prototypes that utilize drawing seed, quality production, quick modifications, production innovation, cost and time effectiveness [3]. The scope of CAD tools has been extended to include the whole spectrum of design initiation and decision making through to technical design, with the subsequent link to the production plant and machinery [4]. The application of CAD as tools in ceramic prototyping includes digital and rapid prototyping.

## 2.1 Digital Prototyping

Prototyping is a fundamental design initiation which involves the construction of working models of conceived products for mass production [5]. A prototype is the first or original example of product that has been or will be copied or developed; it is a model or preliminary version [6]. Digital prototyping helps product developers to design, iterate, optimize, validate and visualize their products digitally throughtout the product development process [7]. Companies often adopt digital prototyping with the goal of improving communication between product development stakeholders, getting products to market faster and facilitating product innovation. Some of the CAD softwares used by designers for digital prototyping includes; AutoCAD, Siemens NX CAD, Auto desk Algo, Solidworks, Auto desk Inventor, Iron CAD, EPLAN Platform, Pro ENGINEER, Solid Edge, Catia.

### 2.2 Rapid Prototyping

Rapid prototyping is the automatic construction/fabrication of physical objects directly from CAD data sources, using additive maufacturing technology. These systems add and bond materials in layers to form objects, and they are known by the name additive manufacturing, additive fabrication, three dimensional printing, solid freeform fabrication (SFF) and layer manufacturing [8].

With this additive technologies, object can be form with any geometric complexity or intricacy without the need for elaborate machine setup of final assembly [9]. Also, rapid prototyping systems reduce the construction of complex objects to a manageable straight forward and relatively fast process. This has result in their used by engineers as a way to reduce time to market in manufacturing, to better understand and communicate product designs, and to make rapid tooling to manufacture those products. Surgeons, architects, artist and individuals from many other disciplines also routinely use the technology.

## **3 PREVIOUS RESEARCH**

The previous research used Kelantan traditional pottery production as a case study, it investigated and illuminate on the systematic process used by Kelantan Potters for production. This was achieved by several visit to three pottery centre in Kelantan and the outcome of this revealed how Kelantan potters create models by hand turning or carving the concepts in plaster (POP) using turning wheel [10]. However, despite the prodigious skills of the carvers, the method never produced accurate models [11]. Therefore, the researcher proposed that replacing the conventional method with digital and rapid prototyping will help to reduce the development time and also improve the prototype accuracy. In order to test the viability of this proposed method, digital method of product development was developed which has four stages has illustrated in Fig. 1 below, after which the method was used to develop and produce a ceramic product (see Fig. 2).



Figure 1: Digital method of product development



method

The findings from this previous research revealed that digital and rapid technology is a viable tool for reducing development time and enhancing prototype quality (accuracy), producing exclusive designs and durable prototypes. However, the research shows that rapid prototyping is an expensive approach, and thus, proposed conventional method of prototyping (hand turning and carving) to be still valid in small and medium ceramic industries. Therefore, based on these findings, this further research seeks to explore how digital prototyping can be used to enhance conventional method of prototyping.

## 4 FURTHER STUDY

On the basis of these previous findings, further research was performed to explore how digital prototyping can be used to enhance the conventional method by improving the accuracy and also help designers in producing exclusive designs. This was achieved by first developing exclusive "handle free mug" design.

## 4.1 Design Stage (Using Solidworks software)

During the design process, there were several iteration before finally arriving at the final design (see Fig. 3). The first product idea produced was Design A, after which the design was improved on and the result of the improvement is design B. However, after visualizing the design, few adjustments were made which result to design C. Design C was first taken as the final design, but during testing and evaluation, some fault were identified which will prevent it's manufacturability. Therefore, these faults were addressed and this brought about the final design (D). This is one of the advantages of digital prototyping because, it utilizes drawing seed, quality production, quick modifications, production innovation, cost and time effectiveness [3]. Therefore with digital prototyping, designers can visualize design, explore design alternatives, verify suitability of design, test the manufacturability of the proposed design and virtually explore a complete product before it's produced.



Figure 3: Pictures illustrating the development of product idea

#### 4.2 Test and Evaluation

After creating three dimensional design of the product, they were tested and evaluated at this stage to determind their manufacturability. The test includes draft analysis which was done digitally using Solidworks software; the important of the draft analysis is to identify the possible fault that may prevent the model from demoulding during the casting process. Colours are used in draft analysis; green represents positive draft, red represents negative draft while yellow represents parts that needs draft. When the design draft was first analyzed, some parts were yellow which means they need adjustment for easy de-moulding (see Fig. 4). Therefore, the faulted parts were adjusted and re-analyzed (see Fig. 5).



Figure 4: Draft analysis of design



Figure 5: Re-Analyzed design

After the draft analysis, the design was evaluated by showing the different projections and sectional views of the design (see Fig. 6). After which the evaluated design was printed and given to the carvers to produce the prototype using conventional method.



Figure 6: Evaluated design

### 4.3 Creation of the Model (using conventional method)

Due to the extensive drawing given to the carvers, they got the clear view of the design and were motivated to produce the prototype using conventional method. After producing the prototype, it was evaluated. The evaluation revealed some imperfection at the edges of the conventional model. Therefore, in order to address this imperfection, template of the product were designed with Solidworks, printed to scale on a cardboard paper and the paper was cut to the exact shape of the template (See Fig 7). Then, the paper template was folded round the model (see Fig. 8) so as to guide in correcting the imperfection. After the corrections were made, the model was found to be almost 100% accurate (see Fig 9). The model was then used to create the plaster mould (Fig. 10) for mass production.



Figure 7: Picture of the paper template



Figure 8: Template folded round the model.



Figure 9: Picture of the final model produced with conventional method



Figure 10: Picture of the Mould

## CONCLUSION

Since building of rapid prototype from CAD model has been proven to be expensive in previous research; this research revealed that the conventional method of prototyping can thus be enhanced by integrating digital prototyping (3D CAD model) into development process. This will help designers to explore design alternatives, verify suitability of design, test and evaluate design to identify possible faults. It will also help in visualizing products digitally throughout product development process thereby communicating the whole concept and idea being proposed. This will build designer's creative skills and make them think innovatively, thereby making them capable of producing good and exclusive products.

## REFERENCES

 R. Paul. An introduction to the technology of pottery (2nd ed.). Institute of Ceramics & Perbamon Press, 1988

- [2] Wikipedia. Computer-aided design. en.wikipedia.org/wiki/ computeraided\_design. October, 2011
- [3] C. Woodward. Computer aided industrial design for ceramics and glass industries. Finland Dest Artes Oy. 1996
- [4] Design, C.-A. In The Columbia Encyclopedia (6th ed.). Retrieved February 26, 2009, from http://www.encyclopedia.com/1E1-ComputerAD.html. 2008
- [5] O.S. Adelabu. & I.B. Kashim. Developing computer aided tools for ceramic prototyping in nigeria. In Proceedings of the 1<sup>st</sup> International Conference on; Man, Technological Advancement and Sustainable Environment, FUTA SET 2010. 25-27 Oct., Pg 458-465. 2010
- [6] C.K. Chua. & K.F. Leong. (2<sup>nd</sup> Edition). Rapid prototyping: Principles and applications. World Scientific Publishing Co. Plc. Ltd. 2003
- [7] Wikipedia. Digital prototyping. en.wikipedia.org/wiki/Digital\_ prototyping, July 24, 2012
- [8] Wikipedia. Rapid prototyping. en.wikipedia. org/wiki/rapid prototyping. October 26, 2011
- P.K. Venuvinod & W. Ma. Rapid prototyping: Laser-based & other technologies. Boston, MA: Kluwer Academic. 2004
- [10] F.E. Olalere. The act of mass production in Kelantan traditional pottery, Malaysia. IISTE Journal (Arts & Design Studies), vol 3, 2012. Pg 18-25, ISSN 2224-6061 (Paper) ISSN 2225-059X (Online). 2012
- [11] F.E. Olalere. The use of computer aided design for prototyping: An approach to enhance competitiveness in ceramic production.Dept of Product design, Universiti Malaysia Kelantan. Malaysia. 2013. (Thesis)
  - Ab. Aziz Shuaib is an Assoc. Professor and a lecturer in Faculty of Creative Technology and Heritage, Universiti Malaysia Kelantan. Malaysia. His research interests are Architecture, design theory, traditional art and design and heritage.PH-0199838319. E-mail: aziz@umk.edu.my.
  - Olalere Folasayo Enoch is currently pursuing masters degree program in Product Design in Universiti Malaysia Kelantan, Malaysia, PH-0109812721. E-mail: folasayoidd@yahoo.com

## Performance Evaluation of Custom Power Devices in Power Distribution Networks to Power Quality Improvement: A Review

Mahmoud zadehbagheri, Naziha Ahmad Azli, Askar bagherinasab, Shahrin bin Md Ayob

Abstract—Power distribution networks are considered the main link between power industry and consumers and they are exposed to public judgment and evaluation more than any other section. Thus, it is essential to study power quality in distribution section. On the other hand, power distribution networks have always been exposed to traditional factors such as voltage sag, voltage swell, harmonics and capacitor switching which destruct sinusoidal waveforms and decrease power quality as well as network reliability. These are the abnormalities which are imposed to the networks by the consumers resulting in bad effects on other consumers and the network equipment. Thus, power suppliers are committed to provide the consumers with a reliable power of acceptable quality. These goals are achieved by applying power electronic devices, called Custom Power devices, in power distribution networks to solve the problem of abnormalities which disturb power quality. The present study tries to introduce and compare different types of Custom Power Devices and explain their application and the simulation results show the role of these devices in power quality enhancement.

Index Terms— Custom Power Devices, Distribution Networks, Power Quality, Voltage Abnormalities.

\_\_\_\_ 

#### **1** INTRODUCTION

Ower quality started to gain high importance for power supply companies and low voltage consumers since the late 1980s. In this regard, the power distribution companies tried to improve power quality following the consumers' requests. The reasons behind the increasing attention on this issue can be as follows: If a component fails, severe consequence will emerge due to the complex interconnection of the systems- significant increase of harmonics in power systems power supplier companies pay more attention to power quality due to increasing knowledge of the consumers about power quality issues- higher sensitivity of the existing electrical equipment against the different kinds of disturbances manifested in power distribution networks[1,2,7].Today,electrical equipment producers present their products based on the power quality level. Based on the available standards in this regard, development of power electronic devices and the offered models to compensate for the variations in power distribution networks aim to provide the highest level of power quality for the consumers. These power quality equipment's are the power electronics devices which are connected together either in series or shunt and their performance is monitored by an intelligent digital control system [3,4,9].

#### **2 POWER QUALITY**

\_\_\_\_\_

Power quality is a term used to describe the rate of voltage quality in A.C power systems [7,8]. An ideal status for a power network is achieved when a balanced/undistorted threephase voltage with a sinusoidal waveform and constant magnitude and frequency is delivered and any deviation from the ideal status is regarded as weakness in power quality. Electricity consumers evaluate power quality based on its influence on the operation of their equipment and devices [2]. Since sensitivity to power quality disturbances varies from one device to another, different consumers show different perspectives towards the delivered power quality. For example, variation in voltage magnitude may not disrupt operation of a heater. However, emergence of this phenomenon in any lighting device may lead to variation in the amount of produced light resulting in eyestrain. From the perspective of one consumer, voltage variation is not considered a noticeable disruption whereas another consumer may consider this phenomenon as an important disturbance in power quality. Therefore, it is difficult to identify a unanimous criterion for power quality. Recently, extensive efforts have been made by other organizations such as International Electro technical commission (IEC) and Institute of Electrical and Electronics Engineers (IEEE) to identify and classify various power quality disturbances and set a limit for each of them. In this study, it is tried to introduce some methods applied to mitigate various types of power quality disturbances. Therefore, this paper presents a comprehensive survey of custom power devices for mitigation of Power quality problems. [5,6].

#### 2.1An overview of Power quality Improvement Methods

There are some long-established methods to mitigate power quality disturbances. Since voltage sags and interruptions have been regarded as the most widespread type of power quality disturbance in the past, various methods have been proposed to confront these two phenomena, Motor-generator

Mahmoud Zadehbagheri is with the Universiti Teknologi Malaysia, Faculty of Electrical Engineering, 81310 Skudai, Johor, Malaysia .(zbmahmoud2@live.utm.my)

Naziha Ahmad Azli is with the Universiti Teknologi Malaysia, Faculty of Electrical Engineering, 81310 Skudai, Johor, Malaysia. (naziha@fke.utm.my).

Askar Bagherinasab is with the Universiti Teknologi Malaysia, Faculty of Electrical Engineering, 81310 Skudai, Johor, Malaysia. (Askar.ba9heri@uahoo.com).

Shahrin bin Md Ayob is with the Universiti Teknologi Malaysia, Faculty of Electrical Engineering, 81310 Skudai, Johor, Malaysia. (shahrin @fke.utm.my).

sets (to confront interruptions) and Ferro resonance Transformers (to confront voltage sags of large loads) are among these methods [10,11]. Devices such as Load Tap Changing Transformers, line voltage drop compensators and shunt capacitors are common to solve the problem of long-term voltage changes. Surge arrestors are applied to protect against transient overvoltage's . Passive filters are used to reduce the harmonic distortion level of current flow and network voltage. Technology advancement in power electronics has played a significant role in the emergence and evolution of power quality improvement devices .The following parts of the present study introduce some of these devices and explain how they operate[5,6,7].

### **3 CUSTOM POWER DEVICES**

Applying power electronic devices to improve energy transmission systems under the concept of Flexible AC transmission systems (FACTS) is almost a widespread method. However, applying these devices in distribution systems has not been considered until recent years due to the expansion of distribution systems and absence of need to these devices. In the last two decades, high power quality has become essential due to significant expansion in the use of electronic devices at all levels of consumers .(industrial, commercial and domestic) and high sensitivity of these devices to various types of power abnormalities. One of the power quality improvement solutions is using FACTS devices in distribution systems with different application and controlling strategy which is proposed under the term Custom Power. In fact, reliability and power quality are the goals considered in distribution networks and these goals are achieved through Custom Power devices [12,13]. Applying custom power devices is one of the two ways for power improvement in shops, offices and houses .The other way is for the companies that develop the products being consumed in the distribution networks to improve their products. There are many reasons to use custom power devices among which the following reasons can be pointed out [14]: most factors that damage power and influence sensitive loads are derived from the system itself. If the consumers' power quality issues are solved by improving performance of power production sources (distribution networks), the former can be given better services. Finally improving power quality through using custom power devices costs less than solving the problems of one or more specific customers. Fig.1 illustrates the application of FACTS devices as well as custom power devices [15,30].



Fig. 1 Application location of FACTS and Custom Power devices

#### 4 GENERAL CLASSIFICATION OF CUSTOM POWER DEVICES

Generally, custom power devices fall into two main categories as shown in Fig. 2 . The first category contains those devices that do not need DC energy storage sources and are mainly applied as static keys in the system and play the role of ordinary circuit breakers and advantage of these switches than conventional mechanical switches is their high switching speed. Generally speaking, these devices are subcategories of solid state breakers(SSBs). Among these devices, solid state transfer switches(SSTS) are the most widelyused and are usually located to back up sensitive loads and generally according to type and sensitivity of load , in addition to primary feeder, secondary feeder is also required. [15]. Unlike SSBs, the devices that need DC energy storage sources have complicated power and control circuits. Moreover, these devices may be employed in applications such as compensating active and reactive power, eliminating harmonics and compensating for unbalanced voltage etc. These devices can undertake different roles based on their applications. For example, these devices can be applied as an active power filter for removing harmonics. This type of custom power devices falls into three groups: shunt( Distribution Static Compensator), series (Dynamic Voltage Restorer) and series-shunt(Unified Power Quality Conditioner). The following subsections elaborate on each of the three groups of custom power devices [16,31].



Fig. 2 Types and configuration of Custom Power Device

#### 4.1 Distribution Static Compensator (D-STATCOM)

Compensating for reactive power in a distribution network plays a crucial role in the power quality improvement, power factor correction and maintaining constant voltage distribution. Among the various existing conditioners, voltage controllers for voltage source converters quickly respond to reactive power demands. Thus, they are used for power factor correction and voltage regulation. Distribution static compensator (D-Statcom) is a type of VSC-based controllers regarded as an appropriate alternative for Static VAR Compensator( SVC) and they can take the capacitive reactive power current from the power system to produce reactive power [7,11]. D-STATCOM is used to compensate for power quality disturbances such as unbalanced voltage, voltage drop and voltage variations that occur in a millisecond. In such a short time, D-STATCOM can inject active and reactive power into the system for sensitive loads compensation. Active power injection must be conducted through an energy storage system. D-STATCOM can operate like a synchronous voltage source

which has variable phase and magnitude. In this case, the terminal voltage is controlled leading to power factor correction. In addition, this device can be used as an active filter to eliminate current harmonic [17,18]. When D-STATCOM is used along with a SSB and an energy storage source, it operates like an uninterruptable power supply (UPS) for short-term interruptions. Moreover, D-STATCOM can be used as an efficient device for voltage flicker compensation [10,11]. Fig. 3 illustrates the main function of a D-STATCOM. These functions are harmonic compensation for load, negative sequence current and reactive current. D-STATCOM operates like a synchronous condenser without mechanical inertia and controls fluctuations easily and quickly using its power electronics properties. As the D-STATCOM operates, line parameters are supervised permanently and compared with the optimal signal strength. Then, the required current and voltage are created to eliminate disturbances. Thus, this device does not allow the load-created disturbances into the system.



Fig. 3 Main Functions of D-STATCOM

It is worth mentioning that STATCOM was used for the first time in sawmills in England, Colombia and Canada to control flickers caused by large induction motors. D-STATCOM has the rating of 2 MVA and a load of 2.6 MVA with the power factor of 0.85. Voltage variations in these factories were caused by device current difference while cutting the logs. These current flow changes cause voltage variation at the point of common coupling( PCC). Using D-STATCOM eliminates these voltage variations [19,20,21,35].

#### 4.2 Dynamic Voltage Restorer (DVR)

Dynamic voltage restorer is a powerful electronic device that can compensate for voltage sags and voltage swells caused by disturbances and is structurally similar to D-STATCOM and consists of three single-phase inverters and a DC bus. However, a DVR's transformers are connected to a distribution line in series and controls magnitude and voltage phase of the two sides of load by injecting voltage into the network in a millisecond. By using a D-STATCOM, it is aimed to protect sensitive loads against negative effects of disturbing loads. However, using a DVR aims to protect sensitive loads against disturbances of the system. The main application of the DVR is to improve voltage sag and voltage sinking through three-phase voltage injection into the distribution network voltage in series or synchronously. In different applications and capacities, DVR is used with nominal capacity of 2-4 MVA to inject voltage of 0.3-0.5 per unit in 300-500 milliseconds. The nominal capacity of a DVR is calculated in the following way:



Fig. 4 Actual schemes of a DVR

Designing the control system in a DVR is of high importance since it determines its response speed and how it can compensate for voltage sag and voltage swell. Control circuits are used to analyze control signal timing parameters such as magnitude, frequency and phase shift that are injected by the DVR. The injected voltage is supplied based on the control signal and by means of the existing inverters on the power circuit. A DVR can determine the amount of voltage sag or voltage swell by measuring the network's three-phase voltage and comparing it with the reference voltage and then exerts the appropriate command signal into the power electronics converter. An AC/DC converter is a voltage-source-type converter that produces the required voltage and injects it into the network through a series transformer. In the case that only short-term voltage variations are considered, using a DC capacitor as an energy restoration component suffices. However, a power supply must be used on the side of the DC converter if permanent voltage regulation and voltage distortion compensation. Fig.5 is a simple illustration of a DVR on the distribution network[23,24,25]



Fig. 5 an overview of a DVR and its Operation

$$S_{DVR} = S_l \times V_{inject} \tag{1}$$

#### 4.3 Unified Power Quality Conditioner (UPQC)

A UPQC is the most comprehensive system for power quality improvement of distribution networks suggested so far. Structurally, UPQC is a combination of a D-STATCOM and a DVR that are common in a short circuit. As illustrated in Fig. 6 .In fact, a UPQC comprises of two PWM-controlled converters that use one DC bus jointly.Two parameters of current and voltage are used as reference on the control circuit. In this structure, the parallel branch is responsible for load current harmonic compensation, reactive power compensation, power factor correction and no-load current correction. In contrast, the series branch is responsible for compensating for harmonic components of network voltage, correcting three-phase voltage unbalance, compensating for voltage sag and voltage swell, voltage flicker compensation and damping harmonic fluctuations between load and network [26,27,28,32].



Fig.6 Schematic of UPQC

Fig.7 shows a combination of series and parallel compensation in which network voltage harmonic is represented by a voltage source *Vc* and network current harmonic is represented by a current source *Ic*. In the consumer bus, a diode converter has been considered at the sample of voltage and current harmonic production and sensitive loads have been connected to the harmonic in the same bus. The produced harmonics in the network influence the feeder bus and power supply source leading to distortion in the produced voltage.



Fig.7 Combination of series and parallel compensation

Adding a parallel active filter on the load side the UPQC results in universal power quality conditioning system (UPQS). Consequently, UPQS is derived from UPQC. The main goal of a UPQS is compensating for source voltage and load current faults such as line voltage sag and voltage swell, temporary interruptions, unbalance, flicker, reactive distortion and unbalanced current. Fig.8 illustrates an overview of a UPQS used for power quality improvement and it contains a harmonic generating load with parallel connection(diode and thyristor) and sensitive load. The feeder also has harmonic distortions. The p-q theory is used in controlling both the series and parallel components [28,29,33].



Fig.8 Overview of UPQS

## **5 SIMULATION RESULTS**

#### 5.1 D-STATCOM

In this section, a D-STATCOM's performance in a sample distribution network is investigated using computer simulation through PSCAD/EMTDC software. The sample distribution network includes linear, non-linear and unbalanced loads .The D-STATCOM's performance is investigated against one of the disturbances such as harmonic distortions.



Fig. 9 Waveforms of current (KA) with the presence of D-STATCOM in the network A: Sum of low voltage load current B: Current waveforms feeder loads C: Injected current by D-STATCOM

As illustrated in Fig. 9, the D\_STATCOM has compensated for all lower order harmonic components of the load.

#### 5.2 DVR

In this section, a DVR's performance in a sample distribution network is investigated. Fig.10 ,and Fig.11 shows the voltage of one of the phases' load with and without DVR. It is illustrated that a DVR is successful in compensation leading to power quality improvement in the distribution network.



Fig. 10 Voltage of one of the phases' load (A) with and without DVR



Fig. 11 Voltage (Pu) of one of the phases' load (A) with and without  $\mathsf{DVR}$ 

#### 5.3 UPQC

Fig.12 shows the waveform of PCC bus's three-phase voltages when subjected to a three-phase short circuit to the ground. According to Fig. 12, the occurrence of a short circuit in one of the buses leads to voltage sag. Fig.13 shows the waveform of PCC bus's three-phase voltages after compensation by a UPQC. Fig.13 shows that the UPQC has delivered balanced sinusoidal voltage to the load by the compensation.



Fig. 12 Voltage sag resulted from occurrence of three-phase short circuit to the ground



Fig. 13 Compensation of voltage sag resulted from occurrence of threephase short circuit to the ground

The ability of these devices in power quality improvement against various disturbances has been compared and it is available in Table 1 in the appendix.

### **6 CONCLUSTION**

In this study, some quantities in power quality have been pointed out according to the IEEE and IEC standards. Moreover, application of custom power devices for improving these quantities are also mentioned. According to this study, D-STATCOM is used in distribution networks to compensate for power quality problems such as unbalanced load, voltage drop, voltage fluctuation, unbalanced voltage and harmonic distortions. A DVR is responsible for voltage sag and swell protection , voltage balancing and compensating for voltage harmonic distortions while UPQC is responsible for compensating for load current harmonic, reactive power compensation, power factor correction, correcting non-load current and regulating DC circuit voltage.

#### REFERENCES

[1] Arrillaga, J., Y.H. Liu, and N.R. Watson, Semiconductor Power Devices, in Flexible Power Transmission. 2007, John Wiley & Sons, Ltd. p. 21-56

[2] Pal, Y., et al., A Review of Compensating Type Custom Power Devices for Power Quality Improvement. 2008 Joint International Conference on PowerSystemTechnology.2008.1004-1011.

[3] Recommended Practice for Monitoring Electric Power Quality, 2009. IEEEStd.1159,IEEE

[4] Coteli, R., et al., Phase Angle Control of Three Level Inverter Based D-STATCOM Using Neuro-Fuzzy Controller. Advances in Electrical and Computer Engineering, 2012.12 (1): p.77-84
[5] M.Brent Hughes and John S. Chan, "Canadian National Power Quality Survey", Power tech Labs Inc , Canada and B.C Hydro, Canada, 1992.

[6] J. Bumett, "Survey of Power Quality in High-Rise Air Conditioned Buildings", Power Electronics and Variable-Speed Drives Conference, 26-28October1994.

[7] Bilgin, H.F. and M. Ermis, Design and Implementation of a Current-Source Converter for Use in Industry Applications of D-STATCOM. Ieee TransactionsonPowerElectronics,2010.25(8):p.1943-1957.

[8] Najafi, E., A.H.M. Yatim, and Ieee, A D-STATCOM Based On Goertzel Algorithm for Sag Detection and a Novel Current Mode Controller. Iciea 2010: Proceedings of the 5th Ieee Conference on Industrial Electronics and Applications, Vol 2. 2010. 441-446 [9] Xi, Z., et al., Improving Distribution System Performance with Integrated STATCOM and Supercapacitor Energy Storage System, in 2008 Ieee Power Electronics Specialists Conference, Vols 1-10. 2008. p. 1390-1395

[10] Cetin, A., et al., Reactive power compensation of coal convetor drives by using D-STATCOMs, in Conference Record of the 2007 Ieee Industry Applications Conference Forty-Second Ias Annual Meeting, Vols. 1-5. 2007 p.1731-1740.

[11] Xu, S., et al., Development of a D-STATCOM prototype based on cascade inverter with isolation transformer for unbalanced load compensation.2005

[12] J. A. Oliver, R. Lawrence, B. B. Banerjee, .How to Specify Power-Quality-Tolerant Process Equipment. IEEE Industry Applications Magazine, Sept/Oct 2002.

[13] E. Acha , V.G. Agelidis , O. Anaya-Lara, T.J.E. Miller, .Power Electronic Control in Electrical Systems. MPG Books Ltd., BODMIN, Corn-wall,FirstPublished2002.

[14] Shiva Swaminatham, Rajat K. Sen, .Review of Power Quality Application of Energy Storage Systems .Contractor Report, SAND98-1513
[15] CUSTOMPOWER - State of the art., Report by CIGRE WG14.31,2000.
[16] T.Devaraju and V.C. Veera Reddy, Role of custom power devices in Power Quality Enhancement: A Review .International Journal of EngineeringScienceandTechnologyVol.2(8),2010,3628-3634

[17] Cetin, A. and M. Ermis, VSC-Based D-STATCOM With Selective Harmonic Elimination. Ieee Transactions on Industry Applications, 2009. 45(3): p. 1000-1015.

[18] Yao, X.-l., et al., LCL output filter design and the influence to the D-STATCOM compensation characteristic. 2011 International

Conference on Electronics, Communications and Control. 2011.

[19] Noroozian, R., A Performance Comparison of D-STATCOM and DC Distribution System for Unbalanced Load Compensation. International Review of Electrical Engineering-Iree, 2012. **7**(2): p. 4194-4207

[20] Meng, X.-p., et al., Controlling Study of D-STATCOM Based on Reinforcement Learning Adaptive PID. 2009 Ieee International Conference on Automation and Logistics. 2009. 1208-1211

[21] Masdi, H., et al., Design of a prototype D-statcom for voltage sag mitigation. National Power & Energy Conference: PECon 2004, Proceedings. 2004. 61-66.

[22] Antonio Moreno-Munoz, Daniel Oterino, Miguel," Study of sag compensation with DVR," IEEE MELECON 2006, May 16-19, Benalmadena (Malaga), Spain

[23] M.Osbome , R.H. Kitchin And H .M. Ryan," Custom Power Technology In Distribution System An Overview, " IEE Symposium, Reliability And Security Of Distribution Systems1995.

[24]: P. Daehler , R. Affolter ," Requirements And Solutions For Dynamic Voltage Restorer , A Case Study," ABB Review 1/2001.

[25] Ravilla Madhusudanl, Ramamohan Rao," Modeling and Simulation of a Dynamic Voltage Restorer (DVR) for Power Quality Problems Voltage Sags and Swells," IEEE- International Conference On Advances In Engineering, Science And Management (ICAESM -2012) March 30, 31,2012
[26] M. Aredes, "Active Power Line Conditioner" Dr.-Ing. Thesis,

Technische University Berlin, Berlin, Germany, March 1996. [27] M. Aredes , J. Hafner , K. Heuman, "A Combined Series and Shunt Active Power Filter", IEEE/KTH Stockholm Power Tech Conf , Power

Electronic, Sweden, June 1995, PP.237-242.
[28] D.Graovac, V. Katic And A. Rufer, "Power Quality Compensation Using Universal Power Quality Conditionering Systems (UPQS)," IEEE Power Engineering Review, December 2000. [29] Vinod K hadkikar," Enhancing Electric Power Quality Using UPQC: A Comprehensive Overview ," IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 27, NO. 5, MAY 2012.

[30] Olimpo Anaya-Lara and E. Acha," Modeling and Analysis of Custom Power Systems by PSCAD/EMTDC" IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 17, NO. 1, JANUARY 2002.

[31] Yash Pal, A. Swarup, and Bhim Singh," A Review of Compensating Type Custom Power Devices for Power Quality Improvement," Power System Technology and IEEE Power India Conference, 2008.

[32] Nikita HariK ,Vijayakumar, Subhranshu Sekhar Dash," A VersatileControl Scheme for UPQCfor Power Quality Improvement," IEEEInternationalConference(ICETECT),2011.[33]

Jovica V. Milanovic', and Yan Zhang," Modeling of FACTS Devices for Voltage Sag Mitigation Studies in Large Power Systems, "IEEE TRANS-ACTIONS ON POWER DELIVERY, VOL. 25, NO. 4, OCTOBER 2010.

[34] R.J. Nelson et al., "Requirements for Dynamic Voltage Restoration to Relieve Distribution System Voltage Sags", American Power Conference, April 1995.

[35] Stewal1 M. Ramsay Patrick E. Cronin, "Using Distribution Static Compensator (D-STATCOMs) To Extend The Capability Of Voltage-Limited Distribution Feeders" UMSGroup, 2001.

#### APPENDIX

#### TABLE I ABILITIES OF CUSTOM POWER DEVICES ON POWER QUALITY IMPROVEMENT

Type of device	Voltage Sag	Voltage Swell	Transient	Voltage fluctuations	Harmonic	Interruption
D- STATCOM	-	*	*	*	-	-
DVR	*	*	*	*	*	-
UPQC	*	*	*	*	*	-
UPQS	*	*	*	*	*	-
SA	-	-	*	-	-	-
SSTS	*	*	-	-	-	*
SSCB	-	-	-	-	-	*
APF	-	-	*	-	*	-

SA= Surge Arrester, APF= Active power filter

# Application of Binary Particle Swarm Optimization in Automatic Classification of Wood Species using Gray Level Co-Occurence Matrix and K-Nearest Neighbor

Ahmad Fariz Hasan, Mohd Fairus Ahmad, Mohd Nasir Ayob, Shamsul Amir Abdul Rais, Nor Hidayah Saad, Amar Faiz Zainal Abidin, Ali Abussal, Nur Anis Nordin, Syahrul Hisham Mohamad, Safirin Mohd Karis, Hazriq Izzuan Jaafar

Abstract— This paper proposed an application of Binary Particle Swarm Optimization in automatic classification of wood species. The images of wood species are taken from Universiti Teknologi Malaysia's CAIRO Wood Database which consists of 25 species. The features of the images are extracted using Gray Level Co-Occurrence Matrix. Then, Binary Particle Swarm Optimization is use to optimize feature selection and parameters related to it. The result indicates that the proposed approach obtained a better result compared to previous literatures with fewer features used as input for the classifier.

Index Terms — binary particle swarm optimization; computational intelligence; gray level co-occurrence matrix; k-nereast neighbour; optimization; pattern recognition; wood recognition

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

## 1. INTRODUCTION

UTOMATIC classification of wood species is not something new, there are numerous researches has been done on the area [1-14]. Based on our literature survey, researchers in Centre for Artificial Intelligence and Robotics (CAIRO), Universiti Teknologi Malaysia (UTM) and Computer Vision and Intelligent Systems (CVIS), Universiti Tuanku Abdul Rahman (UTAR) have done extensive researches on the performance of the automatic classification of tropical wood species. In this paper, we try to experiment with the application of Binary Particle Swarm Optimization (BPSO) [15] in features selections of Gray Level Co-Occurrence Matrix (GLCM) [16]. The objective is to investigate whether there is any improvement in classification rate if the features and parameters of GLCM used for classification were chosen by using BPSO. Optimized features not only improve the accuracy rate of the classification but also reduce classification time because only features that

- Shamsul Amir Abdul Rais is currently a lecturer at Universiti Malaysia Perlis, Malaysia. Email: shamsulamir@unimap.edu.my.
- Nor Hidayah Saad is currently a lecturer at Universiti Malaysia Perlis, Malaysia. Email: hidayahsaad@unimap.edu.my.
- Amar Faiz Zainal Abidin is currently a tutor at Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Malaysia. Email: amarfaiz@fke.utm.my.
- Ali Ábuassal is an alumni of Universiti Teknologi Malaysia.
- Nur Anis Nordin is currently a student of Universiti Teknologi MARA.
- Syahrul Hisham Mohamad is currently a student of Universiti Teknologi Malaysia. Email: syahrulhisham@utem.edu.my.
- Safirin Mohd Karis is currently a lecturer in Universiti Teknikal Malaysia Melaka. Email: safirin@utem.edu.my.
- Hazriq Izzuan Jaafar is a lecturer in Universiti Teknikal Melaka, Malaysia

relevant are selected as input. The model proposed is taken from a literature [17]. The result obtained indicates that there is a slight improvement in the performance of the k-Nearest Neighbor classifier.

## 2. STATE OF ART

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

As mentioned earlier that CAIRO, UTM, itself had done extensive researches on this area. A literature by M. Khalid et al. [2] back in year 2008 is one of the earliest literatures written in this area. The authors proposed a system that covered the entire process of wood species recognition from hardware implementation for images acquisition to software implementation for automatic classification. For images acquisition, the author used an monochrome CCD industrial camera with а magnification tube. The acquired images are transferred to computer using interface card. The raw image is then preprocessed using several methods: high pass spatial filtering, contrast enhancement and histogram equalization. Then, GLCM is used to extract features from the preprocessed images. These features are used as inputs for the Artificial Neural Network (ANN) classifier to perform classification. The wood images obtained (better known as CAIRO wood database) by the literature has becomes a benchmark of other literatures that come after it. The CAIRO wood database also increase in term of the number of wood species covered as the researches progresses.

R. Yusof *et al.* [6] proposed the use of Gabor Filter (GF) to multiply a single wood image into two images. This will create additional features for the GLCM to work with. The authors then extended the model in [10] by having a higher order GF to produce more additional images. The result obtained indicates improvement compared to [6].

Prasetiyo *et al.* [8] had done an extensive comparison of the performance of several features extraction

Ahmad Fariz Hasan is an alumni of Universiti Teknologi Malaysia and currently a lecturer at University Malaysia Perlis, Malaysia. Email: farizhasan@unimap.edu.my.

Mohd Fairus Ahmad is currently a lecturer at Universiti Malaysia Perlis, Malaysia. Email: fairusahmad@unimap.edu.my.

Mohd Nasir Ayob is currently a lecturer at Universiti Malaysia Perlis, Malaysia. Email: nasirayob@unimap.edu.my

methods and several classification methods. The features extraction methods used are GLCM, Linear Binary Pattern (LBP), Wavelet, Ranklet, Granulometry, and Law's Masks. The classification methods used are Artificial Neural Network (ANN), k-Nearest Neighbour (k-NN), Support Vector Machine (SVM), Linear Discrete Classifier (LCD), and Quadratic Discrminant Classifier (QDC). The authors discussed the performance from two views: computation time and classification rate.

In [13], M. Khalid *et al.* proposed an improved system based on multi-features extractor and classifier. The improved system proposed features extraction using both, Binary Gray Level Aura Matrix (BGLAM) and Statistical Properties of Pores Distribution (SSPD). The features extracted then are clustered into several clusters using K-Means and Kernel Discriminant Analysis (KDA). Two classifiers used in the literature are LCD and k-NN. In the same year, M. Khalid *et al.* [14] proposed the use of Genetic Algorithm (GA) in selecting the features as the input for the classifier. The paper test the performance of GA by using three feature extraction methods: GLCM, SSPD and BGLAM, and two classification methods: k-NN and LDA. It is proven that GA improves the performance of the system.

Other than CAIRO UTM, Computer Vision and Intelligent Systems (CVIS) Universiti Tuanku Abdul Rahman (UTAR) also actively performing researches on the wood species recognition. In year 2007, Y. T. Jing *et al.* [1] proposed a hardware model of an embedded system of a computer vision based wood recognition system. As for initial finding, the authors decided to use Matlab for executing the proposed approach. The proposed approach consists of GLCM as features extraction method and ANN as classifier. The proposed approach is test using five species of tropical wood from CAIRO wood database.

Then, Y. T. Jing *et al.* [3] proposed a one-dimensional GLCM algorithm in year 2008. Again, five species of tropical wood from CAIRO wood database are a part of the test database use by the authors. In a different literature [5], the same authors make a concise comparative study of implementation of two features extraction methods (GLCM and Gabor filters) and three classifiers (Verification-Based, k-NN and Covariance Matrix). In the same year, the authors proposed a rotational invariant GLCM algorithm for wood species recognition. The algorithm proposed that the features values are calculated based on the minimum value of the eight rotational variations of GLCM.

There are also several literatures done by other than CAIRO UTM and CVIS UTAR. In year 2009, R. Bremananth *et al.* proposed a wood species recognition system that can identify 10 species of Indian wood species. The image acquisition is done using a consumer digital camera. Then, the acquired images are resized and conversions from RGB to grayscale are done. GLCM is use to extract statistical-based features from the images. A simple correlation method is use to classify the wood species. V. Piuri and F. Scotti [9] suggested a more complex system in performing wood type classification. The authors suggested the use of spectrometer, 473nm DPSS laser and long pass filter to acquired wood information at fluorescence spectra. The

proposed approach able to recognize 21 type of woods use in the Woodtechnology Gmbh.

A. Malik et al. [11] proposed wood species recognition by image segmentation method of the wood micrographs. Seven types of woods consist of five hardwoods and two softwoods are use as case study. The images acquired at a magnification of 1500 times from the original images. Then, several methods in preprocessing are used to enhance and prepare the images for features extraction. The methods are median filtering, histogram equalization, thresholding, edge detection and dilation. The features extracted are parameters related to the tracheids of the wood: number of tracheids, average circularity of tracheids, average rectangularity of tracheids, average area per tracheid, and average distance between tracheids. The authors used several classification methods which are LDC, Logistic Regression, Naïve Bayes, k-NN, SVM, and ANN.

L. Sun *et al.* in [12] proposed a new wood recognition method based on texture analysis. The images captured using Olympus SZ61TRC. The features extraction method proposed by the authors using Gabor wavelet while the k-NN is chosen as the classifier.

## 3. MOTIVATION

The motivation of having automated wood species recognition had been described in great details in [2]. In [14], GA is use to optimize the features selection process but ignore the parameters selection for features extraction and classifiers. This project attempted to use BPSO not only to optimize features selection process but also optimize parameters use in GLCM and k-NN. BPSO will find the best combination of GLCM parameters (d,bin) features and k-NN parameters (k). There are a few advantages of having optimized features and parameters of feature extraction method and classifiers. First and the most significant advantage is the performance of the proposed system is expected to improve. Second, the computation time taken should be lesser as the number of features selected is reduced to the minimal amount.

## 4. METHODOLOGY

The proposed model for automatic classification of wood texture is similar to [14, 17] where the model is as shown in Figure 1. The model consists of four main components: database of wood texture images, GLCM as features extraction method, k-NN as classifier, and BPSO as optimization method.



Fig 1: The proposed approach

#### 4.1 Wood Texture Image

The wood texture images are taken from the CAIRO wood database as mentioned in [2, 6, 8, 10, 13, 14]. There are 25 types of woods consist of the images of the woods textures. Figure 1 listed the scientific name and images of the wood species. The total number of images used in the experiment is 1250 images. 50 images per wood species.

en permen		0 '		, F
Campnos	Mangifera	Dyera	Durio	Canarium
perma	Foetida	Costulata	Lowianus	Apertum
Auriculat				
um				
	计上于		The second	
Kokoona	Lophopet	Dillenia	Anisoptera	Neobalanocar
Littoralis	alum	Reticulat	Costata	pus Heimii
	Javanicu	а		
	m	and the second second second second		
	<u></u>	D: 11		
Parashore	Shorea	Dialium	Intsia	Koompassia
a	Macropter	lndum	Palembanic	Excelsa
Densiflora	а	1	а	
Densmora		AND DECOMPOSITION OF	CC BERNELSER	
	4	and the second		
			u	
Koompass	Pithecello	Sindora	Artocarpus	Myristica
Koompass	Pithecello bium	Sindora Coriacea	Artocarpus Kemando	Myristica Iners
Koompass ia Malaccens	Pithecello bium Splendens	Sindora Coriacea	Artocarpus Kemando	Myristica Iners
Koompass ia Malaccens is	Pithecello bium Splendens	Sindora Coriacea	Artocarpus Kemando	Myristica Iners
Koompass ia Malaccens is	Pithecello bium Splendens	Sindora Coriacea	Artocarpus Kemando	Myristica Iners
Koompass ia Malaccens is	Pithecello bium Splendens	Sindora Coriacea	Artocarpus Kemando	Myristica Iners
Koompass ia Malaccens is Scorodoca	Pithecello bium Splendens Palaquiu	Sindora Coriacea Tetramer	Artocarpus Kemando	Myristica Iners Pentace
Koompass ia Malaccens is Scorodoca rpus	Pithecello bium Splendens Palaquiu m	Sindora Coriacea Tetramer ista	Artocarpus Kemando Gonystylus Bancanus	Myristica Iners Pentace Triptera
Koompass ia Malaccens is Scorodoca rpus Borneensi	Pithecello bium Splendens Palaquiu m Impressin	Sindora Coriacea Tetramer ista Glabra	Artocarpus Kemando Gonystylus Bancanus	Myristica Iners Pentace Triptera
Koompass ia Malaccens is Scorodoca rpus Borneensi s	Pithecello bium Splendens Palaquiu m Impressin ervium	Sindora Coriacea Tetramer ista Glabra	Artocarpus Kemando Gonystylus Bancanus	Myristica Iners Pentace Triptera
Koompass ia Malaccens is Scorodoca rpus Borneensi s	Pithecello bium Splendens Palaquiu m Impressin ervium	Sindora Coriacea Tetramer ista Glabra	Artocarpus Kemando Gonystylus Bancanus	Myristica Iners Pentace Triptera
Koompass ia Malaccens is Scorodoca rpus Borneensi S	Pithecello bium Splendens Palaquiu m Impressin ervium	Sindora Coriacea Tetramer ista Glabra	Artocarpus Kemando Gonystylus Bancanus	Myristica Iners Pentace Triptera

Fig 2: Example image of each wood species

## 4.2 Gray Level Co-Occurrence Matrix

GLCM [16, 18-19] is use as feature extraction method for the proposed approach. GLCM table is tabulated based on the transition of the gray–level between the pixel of interest and its neighbors. GLCM consists of two parameters:  $\theta$ , d, and bin.  $\theta$  is the angle of between the pixel of interest and its neighbor. d is the distance between the pixel of interest and its neighbor. bin is the number of gray-level use. Instead writing the Matlab source code by our own, we choose to use source code provided by [20].

Feature	Feature	Feature	Feature
Number	(mean, range)	Number	(mean,
	_		range)
1, 2	Autocorrelation	23, 24	Sum of
			squares
3, 4	Contrast	25, 26	Sum of
			average
5, 6	Correlation	27, 28	Sum of
	(Matlab)		variance
7,8	Correlation [18,	29, 30	Sum of
	19]		entropy
9, 10	Cluster	31, 32	Difference
	prominence		variance
11, 12	Sum variance	33, 34	Difference
			entropy
13, 14	Sum entropy	35, 36	Information
			measure of
			correlation
			(Info A)
15, 16	Entropy	37, 38	Information
			measure of
			correlation
			(Info B)
17, 18	Homogeneity	39, 40	Inverse
	(Matlab)		difference
19, 20	Homogeneity	41, 42	Inverse
	[19]		difference
			normalized
21, 22	Maximum	43, 44	Inverse
	probability [19]		difference
			moment
			normalized

Table 1: Features extracted from texture image

As stated earlier, BPSO is use to select optimized values of the parameters in GLCM. In this project the value of GLCM's parameters can vary based on Equation (1) and Equation (2). While all possible  $\theta$  values as in Equation (3), are used in the proposed approach. Therefore, for each value of d, and bin, GLCM algorithm can extract up to 44 features (as shown in Table 1). This is based on the approach suggested by [17, 20].

$$d = \{1, 2, 3, 4, 5, 6, 7, 8\}$$
(1)

 $bin = \{2, 8, 64, 256\}$ (2)

$$\theta = \{0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}\}$$
(3)

#### 4.3 k-Nearest Neighbour

k-NN is used as the classifier due to its simplicity. Classifier k-NN classify a test data based on the majority class of the k-th nearest neighbor of the test data. If there is more than one majority class, the average Euclidean distance is use as a tie-break where majority class with smallest value of the average Euclidean distance will be chooses as the winner. In k-NN, there is only one parameter that need to be set by the user: k. Although, each possible value of k can be attempt to identify the optimal classification performance, the process is tedious and time consuming. Instead, BPSO can be used to find optimized value of k for a suitable range of time period. The proposed approach limits the value of k from integer value of 1 to 16 (as stated in Equation (4)).

$$1 \le k \le 16 \text{ and } k \in \mathbb{N}^+ \tag{4}$$

Each features selected by BPSO as input for the k-NN classifier will be normalize to the range [0, 1]. This to ensure all input features have the same weightage in the training and testing process.

#### 4.4 Binary Particle Swarm Optimization

J. Kennedy and R. Eberhert introduce BPSO [15] in 1997 as a discrete version of Particle Swarm Optimization (PSO). BPSO had been successfully employed in numerous engineering optimization problems such as routing problem in PCB holes drilling process [21], VLSI problem [22-24], and DNA sequence design problem [30]. BPSO differs slightly from PSO in two ways: each dimension of search space in BPSO can be either in the state 0 or 1, and the particle position is updated based on the sigmoid value of the velocity. In this project, each particle in BPSO represents a candidate solution of the optimization problem as shown in Equation (5).

$$\boldsymbol{s} = [bin, d, features, k]^T$$
 (5)

where:

- 1<sup>st</sup> 2 bits for *bin* (00 = 2, 01 = 8, 10 = 64, and 11 = 256).
- Next, 3 bits for *d* (000 = 1, 001 = 2, ... 1111 = 8).
- Next, 46 bits for features selections (1 = select, 0 = off). Bit 6 is for the first feature number, Bit 7 is for the second feature number, and so on.
- Last, 4 bits for k in k-NN (0000 = 1, 0001 = 2, ...
   1111 = 16).

The fitness formulation of each agent or particle is as shown in Equation (6). This equation is better known as classification rate.

$$f(\mathbf{s}) = \frac{100\% \times T}{N}$$

where *T* is number of test data correctly classified, and *N* is the total number of test data.

(6)

The first step in executing BPSO algorithm is to set the BPSO parameters, as listed in Table 2. These values are initialized based on user's desired values, and problem-dependent. Particle's velocity,  $\boldsymbol{v}$  and particle's position, **s**, are randomly assigned based on the boundary of the search space. In this case, **s** can be either 0 or 1 while v is usually clamped at the range of [-2, 2]. Each particle will propose a solution as model in Equation (5). Then, based on the proposed solution by the particles, relevant GLCM parameters value are used, and selected GLCM features are extracted from the wood texture images. These features then are normalized before use as inputs for k-NN classifier. The k-NN classifier will return the classification rate of the proposed solution by the particle. The process is repeated for all particles. After that, the particle's personal best, p, and the swarm's global best, g, are updated according to the condition in Equation (7) and Equation (8).

$$\boldsymbol{p}_{i}^{k+1} = \begin{cases} \boldsymbol{p}_{i}^{k}, \ f(\boldsymbol{s}_{i}^{k}) \leq f(\boldsymbol{p}_{i}^{k}) \\ \boldsymbol{s}_{i}^{k}, \ f(\boldsymbol{s}_{i}^{k}) > f(\boldsymbol{p}_{i}^{k}) \end{cases}$$
(7)

$$\boldsymbol{g}^{k+1} = \begin{cases} \boldsymbol{g}^{k}, \ f(\boldsymbol{s}^{k}_{i}) \leq f(\boldsymbol{g}^{k}) \\ \boldsymbol{s}^{k}_{i}, \ f(\boldsymbol{s}^{k}_{i}) > f(\boldsymbol{g}^{k}) \end{cases}$$
(8)

Then, v is updated for next iteration, **k** for each bit, **d**, using Equation (9).

$$\boldsymbol{v}_{id}^{k+1} = \omega \boldsymbol{v}_{id}^{k} + \boldsymbol{r}_{1} c_{1} (\boldsymbol{s}_{id}^{k} - \boldsymbol{p}_{id}^{k}) + \boldsymbol{r}_{2} c_{2} (\boldsymbol{s}_{id}^{k} - \boldsymbol{g}_{d}^{k})$$
(9)

Algorithm 1: BPSO Algorithm for the proposed approach 01: Initialize all particles with a random position and velocity in the search space 02: while stopping condition not met 03: **for** each particle **do** Calculate the fitness of the particles using Equation 04:(6) 05: if particle fitness better than previous *p* then 06: Set particle fitness value as new **p** 07: end if if particle fitness value better than the current 08: g then Set fitness value as the new *g* 09: 10: end if end for 11: for each particle do 12: 13: Update particle velocity according to Equation (9) 14:Update the particle position according to Equation (10)15: end for 16: end while 17: Present g solution

where *i* is the particle's number.  $r_1$ ,  $r_2$ , and  $r_3$ , are random number of [0,1].  $c_1$  and  $c_2$  are cognitive parameter and social parameter, respectively. Next, **s** is updated using Equation (10) based on the probability of the normal distribution. The process of repeated of evaluation proposed solution of the particles offered to particles position updated are repeated until maximum iteration is reached, and the global best solution is presented. This can be summarized by Algorithm 1.

$$s_{id}^{k+1} = \begin{cases} 1, & r_3 < \frac{1}{1+e^{-\nu_{id}^{k+1}}} \\ 0, & r_3 \ge \frac{1}{1+e^{-\nu_{id}^{k+1}}} \end{cases}$$
(10)

## 5. EXPERIMENTAL RESULT AND DISCUSSION

Table 2 shows the BPSO parameters and value chosen for the proposed approach. 5-fold cross validation method for training and testing is used to measure the performance of the proposed approach. At an instance, the ratio of testing and training data is 1:4.

 Table 2: BPSO parameters and their values for the proposed approach

Parameters	Value
Inertia weight, $\omega$	0.7
Cognitive component, $c_1$	1.42
Social component, $c_2$	1.42
Random: $r_1, r_2$ , and $r_3$	[0, 1]
Velocity clamping, $ v $	2
Number of iteration, k	500
Number of particle, <i>i</i>	50

Table 3 indicates the result obtained by [8] and ours. Although, the performance cannot be compare directly due to variance of experimental methodology by [8], it can be clearly seen that the application of BPSO reduces the complexity of experimental methodology required as all the parameters of the features extraction method and classifier are chosen autonomously and efficiently by BPSO. It can be also seen that BPSO able to reduce the number of features use as input for classifier by 91.58% (assuming all possible combinations of k, d, bin, and number of features in Equation (5) are used as inputs for classification).

 Table 3: Result obtained by [8] and ours

Parameter			Praset	iyo [8]			Our
							S
k	2	3	5	7	9	11	5
θ			0°,45	<sup>°°</sup> , 90°, 8	۵ 135°		
d		1,	2, 3,	., 9, & 1	10		2
bin		256					
Features		520 (	$4 \times 10$	× 1 ×	(13)		124
Accuracy	63.5	64.3	63.5	63.5	56.3	63.5	68.4

## 6. CONCLUSION

This paper introduces reader to the application of BPSO to optimize the process of parameters and features selection of GLCM and k-NN. The methodology of the proposed approach is explained in great details. The result indicates that there is potential for further study due to the good result obtained. Further study can be extended using different optimization strategies and different classifiers.

### 7. REFERENCES

[1] J. Y. Tou, P. Y. Lau, Y. H. Tay. "Computer Vision-based Wood Recognition System." Proceeding of International Workshop on Advanced Image Technology, pp. 197-202, 2007.

[2] M. Khalid, E. L. Y. Lee, R. Yusof, M. Nadaraj. Design of An Intelligent Wood Species Recognition System." International Journal of Simulation Systems, Science & Technology, vol. 9, no. 3, pp. 9-19, 2008.

[3] J. Y. Tou, Y. H. Tay, P. Y. Lau, "One-dimensional Greylevel Co-occurrence Matrices for Texture Classification." International Symposium on Information Technology, pp. 1-6, 2008.

[4] R. Bremananth. "Wood Species Recognition System." World Academy of Science, Engineering and Technology, pp. 873-879, 2009.

[5] J. Y. Tou, Y. H. Tay, P. Y. Lau. "A Comparative Study for Texture Classification Techniques on Wood Species Recognition Problem." Fifth International Conference on Natural Computation, vol. 5, pp. 8-12, 2009.

[6] R. Yusof, N. R. Rosli, M. Khalid. "Tropical Wood Species Recognition Based on Gabor Filter." International Congress on Image and Signal Processing, pp. 1-5, 2009.

[7] J. Y. Tou, Y. H. Tay, P. Y. Lau. "Rotational Invariant Wood Species Recognition through Wood Species Verification." First Asian Conference on Intelligent Information and Database Systems, pp. 115-120, 2009.

[8] Prasetiyo, M. Khalid, R. Yusof, F. Meriaudeau. "A Comparatiev Study of Feature Extraction Methods for Wood Texture Classification." Sixth International Conference on Signal-Image Technology and Internet-Based System, pp. 23-29, 2010.

[9] V. Piuri, F. Scotti. "Design of an Automatic Wood Types Classification System by Using Fluoresence Spectra." IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 40, no. 3, pp. 358-366, 2010.

[10] R. Yusof, N. R. Rosli, M. Khalid. "Using Gabor Filters as Image Multiplier for Tropical Wood Species Recognition System."
12th International Conference on Computer Modeling and Simulation, pp. 289-294, 2010.

[11] A. Mallik, J. T. Saavedra, M. F. Fernandez, S. Naya. "Classification of Wood Micrographs by Image Segmentation." Chemometrics and Intelligent Laboratory Systems, vol. 107, no. 2, pp. 351-362, 20011.

[12] L. J. Sun, Z. W. Ji, H. J. Wang. "A New Wood Recognition Method Based on Texture Analysis." Applied Mechanics and Materials, vol. 58-60, pp. 613-617, 2011.

[13] M. Khalid, R. Yusof, A. S. M. Khairuddin. "Improved Tropical Wood Species Recognition System based on Multi-feature Extractor and Classifier." World Academy of Science, Engineering and Technology, pp. 1702-1708, 2011.

[14] M. Khalid , U. Khairuddin, R. Yusof. "Optimized Feature Selection for Improved Tropical Wood Species Recognition System." ICIC Express Letters, Part B: Applications, vol. 2, no. 2, pp. 441-446, 2011.

[15] J. Kennedy, R. Eberhart. "A Discrete Binary Version of the Particle Swarm Optimization." IEEE International Conference on Systems, Man, and Cybernatics, vol. 5, pp. 4104-4108, 1997.

[16] R. M. Harlick, K. Shanmugam, I. Dinstein. "Textual Features of Image Classification." IEEE Transactions on Systems, Man and Cybernatics , vol. 3, no. 6, 1973.

[17] K. Ahmed, G. Karim, B. M. Mihammed, B. Nacera, A. Mohamed. "A Hybrid Approach for Automatic Classification of Brain MRI Using Genetic Algorithm and Support Vector Machine." Leonardo Journal of Science, vol. 17, p. 71-82, 2010.

[18] L. Soh, C. Tsatsoulis. "Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices." IEEE Transactions on Geoscience and Remote Sensing, vol. 37, no. 2, 1999.
[19] D. A. Clausi. "An Analysis of Co-Occurrence Texture Statistics as a Function Grey Level Quantization." Canada Journal of Remote Sensing. vol. 28, no. 1, pp. 45-62, 2002.

[20] A. Uppuluri. "Matlab Codes for GLCM Texture Features." URL: http://goo.gl/LOCjK [Date Accessed: 15th December 2011].

[21] M. H. Othman, A. F. Z. Abidin, A. Adam, Z. M. Yusof, Z. Ibrahim, S. M. Mustaza, L. Y. Yang. "A Binary Particle Swarm Optimization Approach for Routing in PCB Holes Drilling Process." First International Conference on Robotic Automation System, pp. 201-206, 2011.

[22] Z. M. Yusof, T. Z. Hong, A. F. Z. Abidin, M. N. Salam, A. Adam, K. Khalil, J. A. A. Mukred, N. S. Husin, Z. Ibrahim. "A Two-Step Binary Particle Swarm Optimization Approach for Routing in VLSI with Iterative RLC Delay Model." Third International Conference on Computational Intelligence, Modelling & Simulation, vol. 2, no. 3, pp. 43-47, 2011.

[23] Z. M. Yusof, A. F. Z. Abidin, M. N. A. Salam, K. Khalil, J. A. A. Mukred, M. K. Hani, Z. Ibrahim. "A Binary Particle Swarm Optimization Approach for Buffer insertion in VLSI Routing." International Journal of Innovative Management, Information & Production, vol. 2, no. 3, pp. 34-39, 2011.

[24] Z. M. Yusof, A. F. Z. Abidin, A. Adam, K. Khalil, J. A. A. Mukred, M. S. Mohamad, M. K. Hani, Z. Ibrahim. "A Two-Step Binary Particle Swarm Optimization Approach for Routing in VLSI." ICIC Express Letters, vol. 6, no. 3, pp. 771-776, 2012.

## A Comparative Study of the Application of Swarm Intelligence in Kruppa-Based Camera Auto-Calibration

Ahmad Fariz Hasan, Ali Abuassal, Mutaz Khairalla, Amar Faiz Zainal Abidin, Mohd Fairus Ahmad, Shamsul Amir Abdul Rais, Nor Hidayah Saad, Nur Anis Nordin, Ismail Mohd Khairuddin, Asrani Hj Lit, Hazriq Izzuan Jaafar

**Abstract** – This paper presented a comparative study of the application of two Swarm Intelligence algorithms: Particle Swarm Optimization and Firefly Algorithm in automatic camera calibration problem. The fitness function used in the camera calibration problem is based on the Kruppa's equation. A case study from a dataset provided by Le2i Universite de Bourgoune is taken for benchmarking the performance of both algorithms. The result is compared with previous literatures. Result obtained from these algorithms indicates there is potential for further improvement.

Index Terms— auto camera calibration; computational intelligence; firefly algorithm; particle swarm optimization; swarm intelligence; kruppa equation; fundamental matrix.

#### **1** INTRODUCTION

In camera auto calibration problem, provided essential matrix and fundamental matrix, intrinsic parameters can be found by minimizing the cost function. The intrinsic parameters are the aspect ratio and skew represents the principal point. This paper employed two Swarm Intelligence algorithms: Particle Swarm Optimization and Firefly Algorithm in optimizing the intrinsic parameters by minimizing the cost function. The cost function chosen is the Kruppa's equation.

#### 2 STATE OF THE ART

Auto camera calibration problem required the algorithm to find optimal combination of camera calibration's parameters which keep the cost function at a minimal value. This problem is not a new problem where there are numerous literatures had attempted to solve the problem.

Y. Zhang and Q. Ji proposed the application of GA for camera calibration in year 2001 [6]. Genetic Algorithm is use to optimize interior and exterior camera parameters. The experimental result use synthetic and real images.

- Ali Abuassal is an alumni of Universiti Teknologi Malaysia.
- Mutaz Khairalla is an alumni of Universiti Teknologi Malaysia.
- Amar Faiz Zainal Abidin is currently a tutor at Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Malaysia. Email: amarfaiz@fke.utm.my.
- Mohd Fairus Ahmad is currently a lecturer at Universiti Malaysia Perlis, Malaysia. Email: fairusahmad@unimap.edu.my.
- Shamsul Amir Abdul Rais is currently a lecturer at Universiti Malaysia Perlis, Malaysia. Email: shamsulamir@unimap.edu.my.
- Nor Hidayah Saad is currently a lecturer at Universiti Malaysia Perlis, Malaysia. Email: hidayahsaad@unimap.edu.my..
- Nur Anis Nordin is currently a student of Universiti Teknologi MARA.
- Ismail Mohd Khairuddin is currently a lecturer of Universiti Malaysia Pahang.
- Asrani Hj Lit is currently a lecturer of Universiti Malaysia Sarawak.
- Hazriq Izzuan Jaafar is a lecturer in Universiti Teknikal Melaka, Malaysia.

The author claimed that GA produced an excellent performance in term of convergence, accuracy and robustness.

In year 2008, J. Z. Tao *et al.* proposed the application of Particle Swarm Optimization (PSO) in camera autocalibration process [1]. In the paper, the author decided to find only theaspect ratio of the intrinsic parameters which consist of  $f_u$  and  $f_v$  (focal length in pixels along the axes of the image). The skew,  $\gamma$  is let to 0. While the other two intrinsic parameters:  $u_0$  and  $v_0$  are ignored based on recommendation by [2]. The fitness or cost function use in optimizing the parameters are as recommended by [3].

In the same year, K. Bilal and J. Qureshi investigate the application of several nature inspired optimization algorithms in tuning the parameters in camera autocalibration. The algorithms use are Simulated Anneling (SA), Genetic Algorithm (GA), and PSO [4]. The main objective is to benchmark the performance of the algorithms based on several criteria: algorithm efficiency, algorithm accuracy, algorithm reliability, and calibration error (at different noise level). The authors concluded that if the application requires reliability, GA would be more suitable. While if the application requires precision, SA or PSO can be use.

On the following year, X. Song *et al.* proposed another implementation of PSO for single camera calibration [5]. A detailed experimental setup are explained in the paper. The result indicates that PSO provides satisfying calibration accuracy.

J. Li *et al.* proposed a hybrid of GA & PSO in order to improve the accuracy of the camera auto-calibration [7]. The simplified Kruppa's equation is use as cost function. The result indicates that the proposed approach produced 100% success rate compared to PSO (97%) and GA (98%).

Ahmad Fariz Hasan is an alumni of Universiti Teknologi Malaysia and currently a lecturer at University Malaysia Perlis, Malaysia. Email: farizhasan@unimap.edu.my

## **3** MOTIVATION

Motivation to use Swarm Intelligence (SI) algorithms to optimize parameters in camera auto-calibration problem had been mentioned concisely in [1]. J. Z. Tao *et al.* mentioned that SI better than the traditional optimization algorithms:

Traditional optimization algorithms required initial value and sensitive to it. Most SI algorithms which based on Swarm Intelligence (SI) does not require initial value as the agent are randomly assigned in the search space at the initialization phase. Traditional optimization algorithms have higher chance to trap in local minima. Swarm Intelligence algorithms have less chance to trap in local minima due to their stochastic nature. Traditional optimization algorithms required extensive knowledge on the mathematical foundation behind the optimization strategies. All SI algorithms are nature inspired which make them easier to relate to. The algorithms also do not contain many mathematical equations.

Traditional optimization algorithms are more suitable to solve linear problem while SI algorithms can tackle both (linear and non-linear) problems quite well.

In this paper, the implementation of SI algorithms has been extended to five parameters compared to three in [1]. Implementation of Firefly Algorithm (FA) on camera auto-calibration also never been done before.

#### 4 CAMERA AUTO-CALIBRATION

The main objective of camera calibration is to obtain the intrinsic camera parameters for a given number of images. In auto-calibration, as the word suggested, the camera calibration does not require supervision by the user. The intrinsic matrix is as shown in Equation (1).

$$K = \begin{bmatrix} f_u & \gamma & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$
(1)

where  $[u_0, v_0]$  is the skew ratio,  $f_u$  is the product of focal length and magnification factor,  $\varepsilon$ . The magnification factor,  $\varepsilon$  is defined by Equation (2).

$$f_v = \varepsilon \, \times \, f_u \tag{2}$$

As suggested by previous authors [1, 3], a point, *p* is on the absolute conic case, vector  $\mathbf{x} = (x, y, z)^T$  satisfy Equation (3).

$$\boldsymbol{x}^T \boldsymbol{x} = \boldsymbol{0} \tag{3}$$

Based on (3) we can then extend the work of [3], where now each point, p must satisfy Equation (4) or Equation (5).

$$p^T K^{-T} K^{-1} p = 0 (4)$$

$$p^T \omega \, p = 0 \tag{5}$$

While the dual absolute conic of  $\omega$ ,  $\omega^*$  is as Equation (6).

$$\omega^* = KK^T \tag{6}$$

R. I. Hartley in [3] simplified the Kruppa's equation to Equation (7) where values of r and s come from diagonal matrix, D which is described as Equation (8). Column vector, U are  $u_1$ ,  $u_2$ , and  $u_3$ . Column vector, V are  $v_1$ ,  $v_2$ , and  $v_3$ .

$$\frac{r^2 v_1^T \omega^* v_1}{u_2^T \omega^* u_2} = \frac{r s v_1^T \omega^* v_2}{-u_2^T \omega^* u_1} = \frac{s^2 v_2^T \omega^* v_2}{u_1^T \omega^* u_1}$$
(7)

$$D = \begin{bmatrix} r & & \\ & s & \\ & & 1 \end{bmatrix}$$
(8)

By assigning each part of Equation (7) as  $J_1$ ,  $J_2$ , and  $J_3$ , we can obtained Equation (9) to Equation (11).

$$J_{i1} = \frac{r^2 v_1 {}^T \omega^* v_1}{u_2 {}^T \omega^* u_2} \tag{9}$$

$$J_{i2} = \frac{rsv_1^T \omega^* v_2}{T} \tag{10}$$

$$J_{i3} = \frac{-u_2^T \omega^* u_1}{s^2 v_2^T \omega^* v_2} \frac{1}{u_1^T \omega^* u_1}$$
(11)

The optimized value of the parameters of intrinsic matrix can be obtained by finding value of parameters that minimize Equation (12).

$$Error = \sum_{z=1}^{im-1} \sqrt{J_{i1}^{2} + J_{i2}^{2} + J_{i3}^{2}}$$
(12)

where *im* is number of images. Note that Equation (12) is use as fitness function for SI algorithms.

#### 5 SWARM INTELLIGENCE

Swarm Intelligence is an emerging field in Computational Intelligence (CI) where all the algorithms are inspired by the cooperative knowledge of nature. All SI algorithms consists of three main components: initial random position in the search space, fitness comparison between the agents and agent trying to improve their fitness by learning from other agents. Most of the algorithms differ on the third component where each algorithm has different learning methods based on the nature it inspired from.

One of the earliest SI introduced is PSO by J. Kennedy and R. Russell [8]. The movement of the flocking birds inspires the algorithm. The main gist of PSO is that the entire population will try to replicate their historical success and in the same try to follow the success of the best agent in the population. The movement of the agents in the search space dimension shows this attempt.

Firefly Algorithm was introduced by Xin-She Yang in 2007 which fundamentally based on the mating behavior of fireflies [9]. Instead the entire population trying to replicate the best agent success, agents in FA tries to compare its fitness with its neighbors. It will try to improve its fitness by learning from all agents that has better fitness than it.

For this experiment, both algorithms can be modeled using the same model. The proposed model suggests that the relationship of agent's position,  $s_I$  with intrinsic matrix parameters can be generalized as Equation (13).

$$\boldsymbol{s}_m = [f_u \ \beta \ \boldsymbol{u}_0 \ f_v \boldsymbol{v}_0]^T \tag{13}$$

Example,  $s_2 = [800 \ 0 \ 256 \ 900 \ 256]^T$  means that the 2nd agent suggesting that the parameters of the intrinsic matrix should be as follows:  $f_u = 800, \beta = 0, u_0 = 256, f_v = 800, v_0 = 256$ . Another point worth mentioning here is the selection of the fitness function. Fitness function is the function that the agents use to benchmark their proposed solution. This function must have a numerical value. As mentioned earlier, the fitness function of the SI algorithms is as shown in Equation (12) which can be rewritten as Equation (14).

$$f(\boldsymbol{s}_{l}) = \sum_{z=1}^{im-1} \sqrt{J_{i1}^{2} + J_{i2}^{2} + J_{i3}^{2}}$$
(14)

### 6 MODELING CAMERA AUTO-CALIBRATION PROBLEM USING FIREFLY ALGORITHM

The algorithm starts by generating initial population of agent, randomly. Here the agent is the firefly. The fireflies' positions are evaluated using the fitness function in Equation (14). Light intensity,  $I_m$  is formulated to be equal to the inverse value of the firefly's fitness function as shown in Equation (15).

$$I_m = \frac{1}{f(\boldsymbol{s}_m)} \tag{15}$$

From here on the algorithm will start looping until stopping criteria are fulfilled. For this study, maximum iteration, z is chosen as stopping criteria where the algorithm will stop when the iteration, t reached maximum iteration, z.

For each iteration, each agent will move toward to other agent with greater light intensity. The movement of this agent is bounded by Equation (16).

$$\boldsymbol{s}_m = \,\boldsymbol{s}_m + \,\beta_0 e^{-\gamma r_{mu}^2} (\boldsymbol{s}_m - \boldsymbol{s}_u) + \alpha \boldsymbol{e}_m \tag{16}$$

where r is the distance between two agents in Euclidean distance. Given agent m and agent n, the Euclidean distance can be calculated using Equation (17).

$$r_{mu} = \|\boldsymbol{s}_m - \boldsymbol{s}_u\| \tag{17}$$

 $\beta_0$  is the agent's attractiveness at r = 0.  $\gamma$  is absorption coefficient.  $\alpha$  is randomization parameter which in range [0,1].  $\boldsymbol{e}_m$  is a vector random number taken from uniform distribution.

Algorithm 1: Firefly Algorithm for camera auto-calibration 01 Set fitness function,  $f(s_m)$  according to Equation (13) where  $s_m = [s_{m1}, s_{m2}, ..., s_{mn}]^T$  02 Generate randomly initial population of agent,  $s_m$  where m = 1, 2, ..., q

03 Find agent's light intensity,  $I_m$  at  $S_m$  using Equation (15) 04 Define light absorption coefficient,  $\gamma$ 

05 while z < t

- $06 \quad \text{for } m = 1 \text{ to } q$
- $101 \text{ m}^2 + 10 \text{ q}^2$ 07 for u = 1 to q
- 08 if  $I_m < I_u$
- 09 Move agent *m* towards *u* using Equation (16)

10 Evaluate new solution using Equation (14), update

 $I_m$  using Equation (15) and global best if necessary

11 end if

12 end for u

13 end for *m* 

14 end while

15 Post process results and visualization

The fitness of the new agent's position is evaluated and the light intensity is updated. If the fitness obtained smaller than the global best record, the new fitness will become the new global best and the agent's position is kept as the best solution found so far.

#### 7 MODELING CAMERA AUTO-CALIBRATION PROBLEM USING PARTICLE SWARM OPTIMIZATION

Similar to FA where PSO starts by randomly assigned the particle position based on Equation (13). Then the particle fitness is calculated using Equation (14).

Algorithm 2: Particle Swarm Optimization for camera autocalibration 01 Initialize all particle by randomizing position based on

Equation (13)

02 while z < t

- 03 for m = 1 to q
- 04 Calculate fitness for particle using Equation (14)
- 05 if the particle fitness is better than previous *pbest* then
- 06 Set the particle fitness value as new *pbest*
- 07 if the *pbest* is better than previous *gbest*
- 08 Set **pbest** as new**gbest**
- 09 end if
- 10 end if
- 11 end for *m*
- 12 for m = 1 to q do
- 13 Calculate particle velocity according to Equation (18)
- 14 Update the particle position according to Equation (19)

15 end for

16 end while

17 Post process results and visualization

Then the **pbest** and **gbest** will be updated if the particle has a better fitness value compared to the current **pbest** and **gbest** values. Then, the particle velocity,  $v_m^{z+1}$  is updated using Equation (18).

$$\boldsymbol{v}_m^{z+1} = \omega \boldsymbol{v}_m^z + r_1 c_1 (\boldsymbol{v}_m^z - \boldsymbol{pbest}_m) + r_2 c_2 (\boldsymbol{v}_m^z - \boldsymbol{gbest})$$
(18)

where  $r_1$  and  $r_2$  are random values [0,1],  $c_1$  is cognitive component and  $c_2$  is social component. After that, the

particle position is updated using Equation (19).

$$s_m^{z+1} = s_m^z + v_m^{z+1} \tag{19}$$

The process is repeated until reaching the maximum iteration. The *gbest* is taken as the best found solution.

#### 8 IMPLEMENTATION AND SIMULATION RESULT

To compare the performance between the algorithms, the algorithms are tested using a dataset provided by Le2i Universite de Bourgoune [10]. The algorithm is written in MATLAB environment and the simulation is performed 10 times on a laptop equipped with 1.80GHz Intel Pentium Core 2 Duo processor with 2GB RAM.

Table 1 listed out the parameters values of SI algorithms used throughout this simulation.Both SI algorithms used same values for common parameters for benchmarking purposed. Table 2 listed out the fitness value obtained from the simulation done.

	PSO	FA						
Co	ommon Parameters							
Number of agents, i	100	100						
Number of iterations, t	1000	1000						
Number of computa-	10	10						
tions								
	PSO Parameters							
Inertia weight, $\omega$	0.9	Not applicable						
Cognitive component,	1.42	Not applicable						
$c_1$								
Social component, $c_2$	1.42	Not applicable						
$r_1$ and $r_2$	Random [0,1]	Not applicable						
FA Parameters								
α	Not applicable	0.01						
β	Not applicable	0.1						
γ		0.001						

Table 1: Parameters of PSO and FA

	: Fitness value obtained of PSO, FA and [11]
--	----------------------------------------------

	PSO	FA	Levenberg-
			Marquart
Best	1.8481	2.4394	1.6051
	$\times 10^{-7}$	$\times 10^{-4}$	$\times 10^{-13}$
Worst	1.3582	2.7532	Not available
	$\times 10^{-4}$	$\times 10^{-4}$	
Mean	$1.404 \times 10^{-5}$	2.669	Not available
		$\times 10^{-4}$	
Standard	4.2787	9.500	Not available
Devia-	$\times 10^{-5}$	$\times 10^{-6}$	
tion			

In [11], the author implements self-calibration method proposed by [12]. The optimization strategy use is Levenberg-Marguart Algorithm (LMA) which is suitable for non linear system. Result indicates that LMA produces the best result, PSO at the second place, and FA in last place. The disadvantage of LMA is that it requires initial/rough estimation value of the intrinsic parameters while SI algorithms does not. From Table 2, it can be also seen that PSO is more accurate while FA is more precise. Table 3 listed the best and worst found parameters value of PSO, FA and LMA. The PSO and FA values are round up to one decimal point.

From Table 3, one can notice that if the proposed value of the best found PSO being round up to nearest integer, it will produces the same result like LMA. In general, SI algorithms are really high precision-value algorithm. To solve this problem, the user usually defines the number of precision of the parameters values.

10001	00.101	unicee	o reareat	- p - o p		<i>j</i> ± <i>c ci</i>			1
		PSO			FA		Lev M	venbe arqua	erg- art
Best	[799.9 0 0	0.0 799.9 0	255.9 255.9 1	[799.4 0 0	19.4 751.7 0	277.5 261.5 1	800 0 0	0 800 0	256 256 1
Worst	989.9 0 0	0.0 1000.0 0	304.8 203.0 1	[716.2 0 0	58.6 670.0 0	270.8 254.9 1	ap	Not plical	ole

Table 3: Parameters values proposed by PSO, FA and [11]

#### 9 CONCLUSION

This chapter introduces reader to the application of SI to find optimized values of the intrinsic matrix's parameters for pinhole camera. The methodology of the proposed approach is explained in great details. The result indicates that there is potential for further study due to the good result obtained. Further study can be extended using different optimization strategies and fitness functions.

#### REFERENCES

- J. Z. Tao, W. Wenhuan, W. Min, "Camera Autocalibration from Kruppa's Equations Using Particle Swarm Optimization," Proceeding of International Conference on Computer Science and Software Engineering, pp. 1032-1034, 2008.
- [2] S. Bougnoux, "From Projective to Euclidean Space Under Any Practical Situation, A Criticism of Self-Calibration," Proceeding of 6th International Conference on Computer Vision, pp. 790-796.
- [3] R. I. Hartley, "Kruppa's Equations Derived from the Fundamental Matrix," IEEE Transactions on Pattern Analysis and Machine Learning, vol. 19, no. 2, pp. 133-135, 2007.
- [4] K. Bilal, J. Qureshi, "Nature Inspired Optimization Techniques for Camera Calibration," Proceeding of International Conference on Emerging Technologies, pp. 27-31, 2008.
- [5] X. Song, B. Yang, Z. Feng, T. Xu, D. Zhu, Y. Jiang, "Camera Calibration Based on Particle Swarm Optimization," Proceeding of International Congress on Image and Signal Processing, pp. 1-5, 2009.
- [6] Q. Ji, Y. Zhang, "Camera Calibration with Genetic Algorithms," IEEE Transactions on Systems, Man, and Cybernatics, Part A: Systems and Humans, vol. 31, no. 2, pp. 120-130, 2001.
- [7] J. Li, Y. Yang, G. Fu, "Camera Self-Calibration Method Based on GA-PSO Algorithm," Proceeding of IEEE International Conference on Cloud Computing and Intelligence Systems, pp. 149-152, 2011.
- [8] J. Kennedy, R. Eberhart, "Particle Swarm Optimization," Proceeding of IEEE International Conference on Neural Networks, vol. 4, pp. 1942-1948, 1995.

- [9] X. S. Yang, "Firefly Algorithms for Multimodal Optimization," Lecture Notes in Computer Science, vol. 5792, pp. 169-178, 2009.
- [10] University of Burgundy, Camera Autocalibration Dataset, URL: http://goo.gl/qxerh [Date Accessed: 19th January 2012].
- [11] G. Lemaitre, "3D Digitization: Camera Autocalibration," URL: http://goo.gl/s8CTT [Date Accessed: 19th January 2012].

## A SmartInk: Mediator Solution of Technology Learning Dilemma for Digital Note Application

#### Mogeeb A. A. Mosleh, Mohd Sapiyan Baba

**Abstract**— This paper proposed a novel solution for technology learning dilemma found on current note taking application. SmartInk system prototypes implemented with specific mediation tools to improve technology research progress for transferring note taking into digital age. Technology learning dilemma was addressed here as one of critical issues exists in most developed tools of note taking application. Initial results of conducted experiments for evaluation SmartInk prototype proof the strong ability of systems in solving current problem in terms of system efficiency, effectively, and usability. Our conclusion drawing here that mediation tools can be developed by using the powerful of technology to bridge the gap between traditional tasks of note taking and digital environments without losing learning consistency.

Index Terms— Digital Note, Educational Technology, e-Learning, Intelligent E-Learning Systems, Learning Development, Note Mediator, SmartInk, Simulations in Education, Technology Learning Dilemma.

### **1** INTRODUCTION

Note-taking is an important activity for supporting learning behaviours made by most people to record events and capture information for enhancing their memorization. Note taking is a shorthand process which allow large amount of information to be summarize and written on paper quickly. It is a complex human behaviour related to personal information management with a variety of underlying mental processes, and cognitive interactions [1]. Note taking research had begun early on 1920s, when Crawford performed experimental study to examine the impacts of taking notes during lecture on student's performance [2]. Education and learning theories research reported that note taking process has two essential functions for supporting learner activities, encoding and external storage. Encoding is a function of note taking that improves learning by affecting learner cognitive process and working memory. While, External storage is the produced notes which used to record information for reviewing purposes [3], [4], [5][6]. Moreover, research found that taking notes improves the ability of learn, integrate, and capture knowledge [7],[8],[9]. Furthermore, research reported that note taking improve students learning achievement and their academic performance because it is effect their recalling of external memory [10][11]. In addition, research reported that about 99% of students are writing notes, and 96% of them consider note taking as essential activity of their academic tasks [12][13].

Recently, technology was involved to change learning environment by replacing the traditional learning tools with digital devices such as projector, lecture slides, etc. Technology had achieved well in serving learning and education area, however a slow progress of technology was made for transferring note taking into digital age. Unfortunately, note taking as education tool is still struggling to be existed traditionally, even we are in the middle of digital revolution [14],[15]. Necessity of transferring note taking into digital era is become more importance due the widely increasing of information resource while manual note taking become insufficient to process these huge amounts of information. Moreover, digital note is providing us with several advantages such as auditability, legibility, portability, searchability, and extra functionality such as ability for indexing, linking, and information extraction.

#### **2 RELATED WORKS REVIEW**

\_\_\_\_\_

#### 2.1 Current Note Taking Tools

Several note taking applications have been developed to support note taking process in digital environments with various criteria such as active learning, active reading, information assimilations, and collaboration tools, where we classified them based on user group targets and system functionality components. For example, active learning tools were designed to support user activities of taking notes during classroom such as StuPad, NoteTaker, Classroom Presenter, E-Notes, and DyKnow [16],[17],[18],[19],[20]. While other system is concerned to support users for taking notes during reading information resources where annotation, highlighting, and adding comments functions were got high consideration for elaborating resource materials such as DigitalDesk, XLibris, PapierCraft, Paper-Top -Interface, and InkSeine [21],[22],[23],[24],[25]. Furthermore, some developed tools were focused more about the learning gain of collaboration, and sharing notes between users such as Tivoli, Livenotes, CoScribe, and NotePals [26], [27], [28], [29]. Additionally, several tools were developed to facilities note taking in handheld, and Tablet PC devices for performing specific tasks such as creating map concepts, enabling offline learning, annotation, sharing course structure, and semantic indexing [30],[31],[32],[33].

Many research groups were participated to move note taking towards electronic formats, and advised the suitability of us-

Author Mogeeb is currently pursuing PhD degree program in Artificial Intelligent field in University of Malaya, Malaysia, PH-+60132773344. Email: MogeebMosleh@Yahoo.com

<sup>•</sup> Co-Author **Prof.Mohd Sapiyan** is a staff member in Artificial Intelligent Department, University of Malaya, Malaysia, PH-+60379676300. E-mail: pian@um.edu.my

Faculty of Computer Science & Information Technology, University of Malaya, 50603, Kualalumpur, Malaysia.

ing hand-held devices, Taplet PCs, and personal computers for note taking activities [16],[29],[34]. Hence, current note taking tools were widely diverse in their user targets, function components, interface layout, and their ways in achieving note taking tasks.

Based on our review of current tools, We found that current note taking tools not only fails in supporting the activities of taking notes, however they are surviving from several issues such as usability, mentality, and knowledge serving. Furthermore, developed tools were impacted learning features negatively, and included some learning deficiency. In this study, we investigated most of developed note taking tools to discover the issues and problems in behind the failure of current technology in supporting note taker tasks. Investigated study was performed to examine current note taking tools with its effects on current user practice, and its effects on learning functions. Although, there has been an increasing of note taking application over last decade, however current tools were miscarried, and failed to signify the traditional advantages and tasks of note taking.

#### 2.2 Research Problem

Based on our review of current tools, We found that current note taking tools not only fails in supporting the activities of taking notes, however they are surviving from several issues such as usability, mentality, and knowledge serving. Furthermore, developed tools were impacted learning negatively, and included some learning deficiency. In this study, we investigated most of developed note taking tools to discover the issues and problems in behind the failure of current technology in supporting note taker tasks. Investigated study was performed to examine current note taking tools with its effects on current user practice, and its effects on learning functions. Although, there has been an increasing of note taking application over last decade, however current tools were miscarried, and failed to signify the traditional advantages and tasks of note taking. Since, focusing was given mainly in this study to address the technology impacts on learning practice of note taker, we investigated current note taking tools based on learning theories of note taker constrains and limitation. Our finding was drawn here that most of existing tools of note taking are suffering from various issues of learning insufficiency which summarized here under a term of technology learning dilemma. Actually, we found that this problem exists in preceding tools of note taking because of three main issues as described below.

Linearity issues vs. Free Form: most if not all of developed tools were used a text editors for creating and editing note documents which leads to break freedom roles of note taker activities. Using keyboard for writing note in linear way was recommended by specific studies because of faster typing speed, and gain advantages of electronic notes [17],[35]. However, the free form tools are recommended by other research for supporting users to write their notes in flexible matters, and efficient way [24],[36]. This issue made developers to confuse about the appropriate tool which is most convenience for taking notes either linear or free form tool.

**Confliction of traditional and digital advantages:** As discussed above, linearity and free form options were lead a conflictions issue to be occurred between the gain advantages of

traditional and digital note taking. This confliction occurs because the linear system is gain advantages of digital notes but linearity is impacted users current practice, and increased system inefficiency in terms of disturb users, reduce their attention, and increase their cognitive load. Vice versa, users are habitually taking notes in a free form way which provided users with learning features by reducing user time and cognitive load, and also supported user graphical familiarity features where free form tools is caused to loss the advantages of representing note digitally. In contrast, this confliction in selecting the main tool for creating notes leads to delay the process of transferring notes into digital forms, and cause most developed application to fail in representing the note taking activities digitally.

Learning deficiency: Accordingly, current technology is contained some major learning deficiency which impact negatively learning process. For example, a copy-past function had been reported its impact learning negatively because it allows note taker to copy the text without even reading it. Our finding is agreed with similar research that using copy-past function on note taking application decreased the learner ability for memories knowledge well [14],[32]. Another example of learning deficiency is occurred in sharing feature, where studies reported that including such function in note taking system can be affected note takers to change their note taking behaviours, styles, and sometime lead other students to depend on active user notes instead of writing their own notes. Moreover, education research showed that sharing function is affecting the encoding function of note taking process which impacted learning outcomes [37],[38].

Specifically, we found that technology learning dilemma is exist in developed note taking tools because of many issues including the design decision mostly mad by system developers whom doesn't related to education area, the design tools developed without special consideration about learning roles and theories, and most developed tools had not evaluated well.

## **3 DESIGN DECISION**

The main objective of this work is to offer a limited solution for technology learning dilemma of note taking application, and to contribute with others research for allowing note taking process to take place in digital age. Thus, for accomplishing these goals we identified a set of key objectives for mediating note taking application as listed below:

- Solving the technology learning dilemma by designing specific tools to acquire maximum gain for both advantages of traditional and digital notes.
- **2-** Introducing the powerful of technology to mimic the context of traditional notes, and to improve user gain learning functionality.
- **3-** Making the process of taking note with technology more realistic for simplify the transferring process of note taking tasks into digital age.

#### 3.1 SmartInk Mediator

Consequently, a prototype SmartInk is designed to be integrated with several tasks, where technology was employed here to mediate note taking tasks that had conflictions or learning dilemma issues. In this research we avoided to design any tools that caused learning deficiency such as copy past, auto-summarization, and sharing functions because they required extensive studies for getting the right design decision.

Mediator note is used for providing changeover between traditional and electronic notes for utilizing the best advantages rather than trying to emulate each traditional task within digital note applications. Furthermore, SmartInk mediation tools was designed to provide user with the ability to create notes using handwriting with pen-based technology, where a background processes run to convert inked notes into digital form. Hence, using this technique is allowed both linearity and free form benefits to be gain. SmartInk was designed to allow users for entering their notes using free forms while transferring processes run in silent mode to generate electric copy of user notes. The SmartInk system is designed to generate the electric version of user notes automatically as illustrated in figure.

#### 3.2 SmartInk Design

SmartInk system functions is delivered from user activities during note taking where we tried to implement these function to be similar as much possible with traditional activities. Table 1 is shown the list of developed functions of SmartInk system.

Functional require-	Specification Context Description
ments	
Write notes	Using handwriting for writing notes
Handwrite drawing	Drawing diagrams and shapes
Annotation notes	Using handwriting to annotate material
Select note elements	Select specific note elements.
Erase note elements	Delete word, sentences, and diagrams.
Move note elements	Change the location of note elements
Highlight note ele- ment	Highlight specific note elements
Search notes	Search about specific note elements.
Index, Tag, and link-	Ability to index, tag, or linking notes with
ing notes	other resources.
Query	Initiate query operation
Import lecture slide	Include lecture materials for annotating and write notes.
Access Notes	Each user has account and data storage space.
User Login	It's a function of user authentication
Access, Brows, and	Functions that support user to open spe-
Navigate Notes	cific folder, subject, and page notes to
	brows note contents.
	Also, Its functions that facilitate note
	browsing through multi page views, and
	display the suitable view for user re-
	quests.

Table 1: SmartInk Functions

Server-client and data repository models were used for implementation SmartInk system architecture as shown in Figure 1. The client server model is used here to specify the higher level of SmartInk abstraction for addressing the interaction process between user interface clients with data repository in sever side. We identified the client side here as note taker devices, where server side represent the data service machine which contains the learning material, user notes, and lecture slides. Furthermore, client-server model of SmartInk was designed to support parallel interaction with data repository for multiple users simultaneously.



Figure 1. SmartInk Architectures models.

Although, Data repository model is used to store a large amount of data such as user information scheme, user notes, lecturer material, and electronic resource material in a central repository of SmartInk server side for allowing users to access, control, and maintains their own data schema. In addition, data repository model included database structure, and metadata objects where database structure store user information, user authentication roles, and user activity events, and the metadata represent the properties and attributes of notes, form layout, and user note documents in XLM formats.

#### 3.3 SmartInk Implementation

SmartInk prototype was implemented to work on Tablet PC devices with pen-based technology to achieve our key objectives. Moreover, SmartInk prototype coded by using Microsoft C#.Net and Microsoft SQL server 2008. We selected C#. Net because it's a platform independent language which can be integrated within various operating system, and Tablet devices. Also, Microsoft SQL server selected here because it is suitability of implementing SmartInk models for creation data repository schemes, manages user storages, and storing the various types of note files.

Furthermore, several classes developed to implement SmartInk functions and interaction process with note taker users including **NoteDocument**, **NotePage**, **NoteElement**, **NoteTransformer**, **NoteViewer**, **NoteUser**, **XMLNote**, and **NoteAgent**.

The **NoteDocument** class is the superclass of SmartInk system which responsible for adding, removing, and organizing note topics, subjects, pages.

The **NotePage** class is a child of **NoteDocument** class which used to define the stored information in each page. **NotePage** is designed to be responsible for adding and removing note elements, keep information about page properties such as dimension, margins, and default page view.

The **NoteElement** class is a child of **NotePage** which used to identify the note elements in each pages, it is used for storing note element such as handwrite text and diagrammatic with element properties such as type, location, and dimension of each element.

The **NoteTransformer** class is a subclass of **NotePage** which responsible to track the user ink and to categorize user stroke components into shape, lines, paragraph, and chunks. Also, it is responsible to store ink note in data repository, where the ink note is converted into associated text note, and the diagrammatic shape is transferred into associated image.

The **NoteAgent** class is a child of **NoteDocument** which introduced here to perform special intelligent tasks that assist in improving system functionality and performance. It's responsible for introducing powerful tools of technology for developing SmartInk mediation tools. **NoteAgent** class is responsible for improving converting process of ink notes by providing users with built-in system dictionary and user dictionary for performing word converting process automatically such as autocorrecting, detecting misspelling words, and identifying the abbreviation words, keywords, and indexed words.

The **XMLNote** is a subclass of **NoteTransformer** which responsible for creating XML schema. It included several functions to converting resulted note text into .XSD file, and to convert XSD file into appropriate SQL script for creating, and modifying the database repository schema.

The **NoteViewer** is a subclass of **NoteDocument** which responsible to display the proper interface layout of user views such as colour, page grid, and highlight tools. The **NoteUser** class is also a child of **NoteDocument** class which responsible for creating and managing users and groups, and identifying their roles, permissions, and storage space with ability to record various user activities on system repository, and ability to manage the parallel user sessions.

In addition, ADO.Net was used for integrating LINQ tools in our system to provide a high level of data abstraction for XML data and query operation. The primary purpose for using LINQ in SmartInk prototype was to unify and simplify the delegation between XML tools such as XQuery, XPath, and XSLT objects.

In contrast, the transferring process of ink notes into XML files is considered the kernel of mediation strategy applied in this study. Besides, transferring process of ink notes into XML schema involved with several steps to generate the electric version of in XML format as shown in Figure 2.



Figure 2. Transferring Ink note into Digital Note.

#### 3.4 SmartInk Interface

Since, university, institute, and learning organization is responsible to support learner with all necessary facilities, equipment's, and materials for learning process, we expected that organizations are responsible to afford a tablet device on learner place connected via networks to facilitate note taking process digitally during classroom.

One of key objective solutions were to keep the current practice of traditional process unchanged as much possible for designing transparent interface of SmartInk system. Thus, we derived SmartInk prototype interface from traditional tools to offer probably the familiarity of traditional approach to support users with more confidently and comfortably tools of note taking. Similarly, Interface is designed based on recommended of many research to use pen as a reliable input device with a tablet PC as paper surface without crowded elements of menus or toolbars to allow users to perform actions as quickly as possible[14],[20],[38]. SmartInk prototype initial interface is shown in Figure 3, where panel tools in left area of interface is designed to be autohide during writing notes.





In addition, SmartInk interface was designed to reduce the overload functionality and to avoid the inefficiency issues reported which existed on most of previous system. On other hand, using pen with Tablet device give users a sense of consistency, and reduce their activities if compared with other systems where they limited pen functionality for drawing, while mouse used for selecting and positioning, and keyboard used for entering text such as OneNote, DyKnow, EverNote, and NoteTaker.

#### 3.5 SmartInk Context Functions

Several system functions are developed to mediate note taking activities within SmartInk such as creating, entering, erasing, moving, tagging, linking, highlighting, and searching tasks as described below. Microsoft Tablet PC API functions were used to implement SmartInk system functions.

**Creating, Adding ink notes:** users allowed entering notes simply by handwriting on Tablet device screen using pen tips as shown in Figure 4. Ink Collector object in NotePage class was used to manage and collect ink objects from digitizer. Moreover, InkCollector object implemented to capture ink input from system interface with efficient event sink to render this input in real time, and save strokes in associated Ink object which used as input for recognition module.



#### Figure 4. SmartInk note sample.

**Converting ink notes into text**: SmartInk prototype implemented to allow the ink note to be converted into text file and XML file automatically. Divider object is used in SmartInk system to analyse the ink elements, classify them into a group of data strokes, and save the results of layout in DivisionResults objects. The Divider object is used mainly to improve the recognition process of note elements via categorizing ink elements into related components for separating the text and drawing objects.



#### Figure 5. Divider Object Process Representation.

Next, the Ink strokes of text objects are sent to the recognition engine module, where recognition module is design here with limitation to recognize English language only. RecognizerContext object is run asynchronously to recognize a given collection of data strokes with two dictionary types Microsoft embedded dictionary and user dictionaries. In case of recognized text is not identified in both dictionaries, the unidentified word is stored in text files normally, and user notification is set to remind user for identifying it properly. The user dictionary is a file in data repository which identified mainly user abbreviation, shorthand word, and special user glossary words. Using this approaches allowed SmartInk to guess the best matches word with alternative possibilities about words, with capabilities to break the alternate segments into separate words and perform autocomplete, and autocorrecting for ink notes automatically. Finally, the results from recognition algorithm are stored in text file where each chunk is stored as text file, where XML file is generated to specify word properties such as coordinates, highlighted, indexed, and other XML attributes. In addition, ink note, text note, and XML note files associated together with index value for storing, and manipulating process, where each associated text file consider as an electric copy of original ink note file as shown in Figure 6.



Figure 6. Demo for Ink note with generated text note. Saving ink notes in data repository: SmartInk is designed to save ink notes in a database to facilitate digital tasks document such as editing, searching, and querying. In SmartInk server side the database is designed to store both versions of ink and electric notes in two individual tables, they are ink\_note and txt\_note tables. The ink notes table contains a unique identifier, the image data in fortified GIF persistence format, and length of data array. While the associated text note table contains unique integer index, the ink identifier, ink words in text, and chunk attributes such as the left, top, right, and bottom values of bounding box. A Useful metadata of ink notes is stored in database tables such as bounding box values, length of ink strokes, and other ink attributes such highlighted, bolded, indexed, tagged and underlined.

**Selecting Note element:** Selection process is considered as the pre tasks for achieving specific tasks of note taking such as deleting, transferring, highlighting, tagging, and linking functions. SmartInk prototype is allowed the selection of ink elements via pen tips, where InkOverlay object used to trace the selected area and return the strokes collection of user selection. The select process is designed to detect pen tips movements by determine the beginning and ending of stroke locations. A bounded box appears on selected area with three options for user desired operation such as deleting, moving, and highlighting tasks as shown in Figure 7.



Figure 7. Example for Selection process in SmartInk. Erasing, Highlighting, and Transferring note elements: The note element can be selected, sized, and handled appearance of four corners of bounding box to perform the note taking activities on note elements. After selecting elements, users can just simply strokes the pen tip on desired options for performing their tasks as represented in Figure 8.



Figure 8. Examples for Erasing, Highlighting, and Moving note. Searching, and querying: Searching and querying about ink notes are considered as one of fundamental purpose of digital notes which supports users for accessing specific information through one or more documents quickly. SmartInk prototype was designed to store ink notes in a database table as discussed previously. SmartInk prototype is designed to support users for searching about specific topics, paragraph, and words either by note contents or creation dates. Searching process is allowed users to initiate search or a batch search for purpose of creating query index of ink notes. In addition, searching procedure is executed to retrieve the index, ink identifier, and matched words from text\_note table. Then the index, and ink identifier is used to retrieve the related ink document files where the stored attributes of chunks is used to identify the position of retrieved words. These functions is implemented with internal procedure in database to retrieve data fast, and display the results of found words in ink note files as shown in Figure 9.

	Title InkNote1	Date:	:
Searching Resultes			
General Notes	Note taking is an inform purpare. Researcher hause Carfilen increage student architen main trus function of ,	ut which be tearing that is taking the and and tearing process of Encoding IT <i>Encoding</i>	ér
	This Example show select Deleting process	ion, Transfer, Highighting, an of Emart Note protatype.	4

Figure 9. Example for Searching Process in SmartInk.

#### 3.6 SmartInk Evaluation

Two combination evaluation approaches of survey questionnaires, and observation experiments were conducted for getting user feedbacks about system usability, and for validating efficiency of SmartInk functions. Students from different majors with less diverse are volunteered to use SmartInk system

during their classroom for taking notes for seven weeks. Total of 35 students were volunteered for purpose of evaluation SmartInk system, where volunteers were 23 male and 12 female, from Twintech University at Yemen [39]. The hardware devices used for evaluating SmartInk system inside classroom were various Tablet PC devices; we used two Acer Iconia Tablet Pc, two Compag TC1100 Tablet Pc, and one iPad Tablet. a total of five Tablet devices used for evaluation SmartInk prototype, where the SmartInk system was customized by using MonoDevelop software ver. 3.0.6 to work on apple OS [40]. Normal PC with powerful performance is used as a server machine and connected with device via intranet network. Five students are selected to use SmartInk system with the provided Tablet devices for taking their notes during classroom for one week only. At the beginning of each week, short introduction for ten minutes was given to the volunteers about using SmartInk in Tablet device for taking notes where volunteers had informed to take their notes by using English language only. Moreover, web based questionnaires is used to evaluate SmartInk usability based on System Usability Scale approaches [41]. In addition, server log files which contents the volunteer's notes and their activities are used to observe the efficiency of SmartInk functions.

#### **4 DISCUSSION**

Evaluation experiments conducted here were designed to evaluate SmartInk system within two dimension, system usability and efficiency.

**System Usability:** Student feedbacks about SmartInk system were collected at the end of experiment period. SUS approached was used to calculate the usability scale of SmartInk system which considered as a simple and common approaches for giving usability scale as numerous value varied between 0 - 100. Analysing data of student feedbacks, and calculated the usability score based on SUS approaches was showed an excellent scored resulted approximately 93.2%. This higher rate is giving to the system due the specific design which constrained to be similar with traditional tools in appearance and functionalities. Students were very satisfied about using SmartInk to take their notes in Taplet PC devices.

System Efficiency: Accordingly, evaluation SmartInk system functionality and performance is conducted by observing students note contents and log events of user activities which stored in data repository. The server logs provide us with a very accurate and detailed summary about user activities and their usage including inked notes, electric notes, associated image diagrams, and user activities logs about usage functions such as create notes, highlight, erasing, tagging, and searching. We compared the user note contents with generated system note to evaluate SmartInk mediator efficiency. 25 ink note samples from different users with their electric text notes were selected randomly. We compared the ink contents with generated notes and found that around 76% of ink notes samples had transferred correctly, while around 11% of ink note had been identified in user dictionaries, and the remaining of 13% of ink note contents converted with some errors. These errors were mostly occurred because user writing notes in different axis, user font is not clear even to us, and due the space distance leaving between letters or words.

In addition, server loges was contented 454 entries of user activities created during experiment period, 193 entries was about creating, and deleting new notes, while the other enters were divided between other user tasks as shown in Table 2 below.

System loge entries	No. of	% user
	Entries	Usage
Creating notes	127	27.9%
Deleting notes	66	14.5
Deleting note elements	43	9.4
Moving note elements	26	5.7
Highlighting note elements	86	18.9
Searching & Querying tasks	57	12.6
Tagging, indexing, linking	49	10.8

Table 2. Summery for Server Log entries and Usag	Summery for Server Log entri	ies and Usag
--------------------------------------------------	------------------------------	--------------

We used the observation methods to verify that students had used all available SmartInk functions, and to know which functions were used frequencies, and which is rarely used. Table 2 showed that users had used some SmartInk system heavily including creating notes, highlighting note elements, deleting notes, and searching functionality. However some SmartInk functions were used regularly such as Tagging, Indexing, Linking, Erasing element, and moving element features. On other hand, we noticed that each user had used all system functions; however we need to perform other experiments to evaluate the intended of users with system functionality.

#### **5** CONCLUSION

In this study, we found that mediation tools can offers a new vision for developing appropriate tools to mimic the actions of performing traditional note taking tasks. Furthermore, evaluation SmartInk was showed that using this technique makes user able to perform activities in more efficient time, and reduce action overhead. In addition, mediation tools developed in SmartInk system were approved its functionality, efficiency, and usability for solving technology learning dilemma which considered as one of essential problems that prevent note taking to be exists in digital era. In addition, current powerful of technology tools should be employed well to mediate the complicated tasks of learning which resisting to be existed in digital media. Even there are some constrains and limitation in this study such as handwriting accuracy, input language, unexpected user behaviour for drawing diagrams. However, our technique is still promising to provide note taker with the appropriate solution of several issues in previous note taking systems. The future of note taking is drawn based on learning advantages, and demand mainly to mediate the note taking activity, style, and behaviour into digital environments.

#### REFERENCES

- A. Piolat, T. Olive, and R.T. Kellogg, "Cognitive Effort During Note Taking," Applied Cognitive Psychology, vol. 19, no. 3, pp. 291-312, 2004. (Journal citation)
- [2] C.C. Crawford, "Some Experimental Studies of the Results of College Note-Taking," Journal of Educational Research, vol. 12, no. 5, pp. 379-86, 1925. (Journal citation)

- [3] F.J. Di Vesta and G.S. Gray, "Listening and Note Taking," Journal of Educational Psychology, vol. 63, no. 1, pp. 8-14, 1972. (Journal citation)
- [4] K.A. Kiewra and B.M. Frank, "Encoding and External-Storage Effects of Personal Lecture Notes, Skeletal Notes, and Detailed Notes for Field-Independent and Field-Dependent Learners," The Journal of Educational Research, pp. 143-48, 1988. (Journal citation)
- [5] R.F. Flippo and D.C. Caverly, Handbook of College Reading and Study Strategy Research. Taylor & Francis Group, 2000. (Book style)
- [6] R.J. Peper and R.E. Mayer, "Generative Effects of Note-Taking During Science Lectures," Journal of Educational Psychology, vol. 78, no. 1, pp. 34, 1986. (Journal citation)
- [7] B.A. Donald, What's the Use of Lectures?, edited by Jossey-Bass San Francisco2000. (Book style)
- [8] J.J. McConnell, "Active Learning and Its Use in Computer Science," Paper presented at the ACM SIGCSE Bulletin, 1996. (Conference proceedings)
- [9] R.J. Peper and R.E. Mayer, "Note Taking as a Generative Activity," Journal of Educational Psychology, vol. 70, no. 4, pp. 514, 1978. (Journal citation)
- [10] J.L. Fisher and M.B. Harris, "Effect of Note Taking and Review on Recall," Journal of Educational Psychology, vol. 65, no. 3, pp. 321, 1973. (Journal citation)
- [11] G.O. Einstein, J. Morris, and S. Smith, "Note-Taking, Individual Differences, and Memory for Lecture Information," Journal of Educational Psychology, vol. 77, no. 5, pp. 522, 1985. (Journal citation)
- [12] R.A. Palmatier and J.M. Bennett, "Notetaking Habits of College Students," Journal of Reading, pp. 215-218, 1974. (Journal citation)
- [13] P. Van Meter, L. Yokoi, and M. Pressley, "College Students' Theory of Note-Taking Derived from Their Perceptions of Note-Taking," Journal of Educational Psychology, vol. 86, no. 3, pp. 567-76, 1994. (Journal citation)
- [14] A. Bauer and K.R. Koedinger, "Note-Taking, Selecting, and Choice: Designing Interfaces That Encourage Smaller Selections," Paper presented at the Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries, 2008. (Conference proceedings)
- [15] K. Kim, S.A. Turner, and M.A. Pérez-Quiñones, "Requirements for Electronic Note Taking Systems: A Field Study of Note Taking in University Classrooms," Education and Information Technologies, vol. 14, no. 3, pp. 255-83, 2009. (Journal citation)
- [16] G.D. Abowd, C.G. Atkeson, A. Feinstein, C. Hmelo, R. Kooper, S. Long, and N. Sawhney, "Teaching and learning as multimedia authoring: the classroom 2000 project," In Proceedings of the fourth ACM international conference on Multimedia, pp. 187–198, 1997. (Conference proceedings)
- [17] N. Ward, and H. Tatsukawa, "A Tool for Taking Class Notes," International Journal of Human-Computer Studies, vol. 59, no. 6, pp. 959-981, 2003.
- [18] R. Anderson, L. McDowell, and B. Simon, "Use of Classroom Presenter in Engineering Courses," Paper presented at the Frontiers in Education, 2005. FIE'05. Proceedings 35th Annual Conference, 2005. (Conference proceedings)
- [19] M.A. Wirth, "E-notes: Using electronic lecture notes to support active learning in computer science," ACM. SIGCSE Bulletin, vol. 35, no. 2, pp. 57–60, June 2003. (Conference proceedings)
- [20] D. Berque," An evaluation of a broad deployment of dyknow software to support note taking and interaction using pen-based computers," J. Comput. Small Coll., vol. 21, no. 6, pp. 204–216, 2006. (Journal citation)
- [21] P. Wellner, "Interacting with Paper on the Digitaldesk," Communications of the ACM, vol. 36, no. 7, pp. 87-96, 1993. (Conference proceedings)
- [22] B.N. chilit, G. Golovchinsky, and M.N. Price, "Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations," Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, 1998. (Conference proceedings)
- [23] F. Guimbretière, "Paper Augmented Digital Documents," Paper presented at the Proceedings of the 16th annual ACM symposium on User interface software and technology, 2003. (Conference proceedings)

- [24] H. Mitsuhara, Y. Yano, and T. Moriyama, "Paper-Top Interface for Supporting Note-Taking and Its Preliminary Experiment," Paper presented at the Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, 2010. (Conference proceedings)
- [25] K. Hinckley, S. Zhao, R. Sarin, P. Baudisch, E. Cutrell, M. Shilman, and D. Tan, "Inkseine: In Situ Search for Active Note Taking," Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, 2007. (Conference proceedings)
- [26] E.R. Pedersen, K. McCall, T.P. Moran, and F.G. Halasz, "Tivoli: An Electronic Whiteboard for Informal Workgroup Meetings," Paper presented at the Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems, 1993. (Conference proceedings)
- [27] M. Kam, J. Wang, A. Iles, E. Tse, J. Chiu, D. Glaser, O. Tarshish, and J. Canny, "Livenotes: A System for Cooperative and Augmented Note-Taking in Lectures," Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, 2005. (Conference proceedings)
- [28] J. Steimle, O. Brdiczka, and M. Muhlhauser, "Coscribe: Using Paper for Collaborative Annotations in Lectures," Paper presented at the Advanced Learning Technologies, 2008. ICALT'08. Eighth IEEE International Conference on, 2008. (Conference proceedings)
- [29] R.C. Davis, J.A. Landay, V. Chen, J. Huang, R.B. Lee, F.C. Li, and J. Lin, "Notepals: Lightweight Note Sharing by the Group, for the Group," Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, 1999. (Conference proceedings)
- [30] K. Luchini, C. Quintana, and E. Soloway, "Pocket Picomap: A Case Study in Designing and Assessing a Handheld Concept Mapping Tool for Learners," Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, 2003. (Conference proceedings)
- [31] K. Nakabayashi, T. Hoshide, M. Hosokawa, T. Kawakami, and K. Sato, "Design and Implementation of a Mobile Learning Environment as an Extension of Scorm Specifications," Paper presented at the Advanced Learning Technologies, 2007, ICALT 2007. Seventh IEEE International Conference on, 2007. (Conference proceedings)
- [32] A. Varadarajan, N. Patel, B. Maxim, and W.I. Grosky, "Analyzing the Efficacy of Using Digital Ink Devices in a Learning Environment," Multimedia Tools and Applications, vol. 40, no. 2, pp. 211-239, 2008. (Journal citation)
- [33] P. Chiu, A. Kapuskar, S. Reitmeier, and L. Wilcox, "Notelook: Taking Notes in Meetings with Digital Video and Ink," Paper presented at the Proceedings of the seventh ACM international conference on Multimedia, vol. 1, 1999. (Conference proceedings)
- [34] M. Lin, W.G. Lutters, and T.S. Kim, "Understanding the Micronote Lifecycle: Improving Mobile Support for Informal Note Taking," Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, 2004. (Conference proceedings)
- [35] D. Pilon, J. Raymond, and P. Raymond, "Software Tools to Improve Note-Taking in the Classroom," Jorn. of Education, IEEE Transactions, vol. 40, no. 4, pp. 3, 1997. (IEEE Transactions)
- [36] Reimer, Y. Jacobs, E. Brimhall, C. Cao, and K. O'Reilly, "Empirical User Studies Inform the Design of an E-Notetaking and Information Assimilation System for Students in Higher Education," Computers & Education, vol. 52, no. 4, pp. 893-913, 2009. (Journal citation)
- [37] M. Grabe and C. Kimberly, "Evaluating the Advantages and Disadvantages of Providing Lecture Notes: The Role of Internet Technology as a Delivery System and Research Tool," Journal of Internet and Higher Education, vol. 8, no. 4, pp. 291-298, 2005, ISSN-1096-7516. (Journal citation)
- [38] R.B Larson, "Enhancing the Recall of Presented Material," Comput. Educ., vol. 53, no. 4, pp. 1278-84, 2009. (Journal citation)
- [39] http://www.iutt.edu.ye, (URL link 2013).
- [40] http://xamarin.com, (URL link 2013).
- [41] J. Brooke, "SUS: A quick and dirty usability scale," 1996. (Technical report)

## Transverse product effect on CSEM with double Hydrocarbon reservoir in seabed logging

Adeel Ansari<sup>a</sup>, Afza Bt. Shafie<sup>b</sup>, Abas B Md Said<sup>a</sup>, Seema Ansari<sup>c</sup>, Muhammad Rauf<sup>d</sup>

<sup>a</sup> Computer Information Systems Department, <sup>b</sup> Fundamentals and Applied Sciences Department, <sup>d</sup>Electrical and Electronic Engineering Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, Tronoh, Perak, Malaysia

<sup>c</sup>Electrical Engineering Department, Institute of Business Management, Pakistan

**Abstract**— In this paper, simulations are performed to reaffirm the relationship between the resistivity and the thickness of the Hydrocarbon reservoir by considering double layers of Hydrocarbon in seabed logging application. In order to establish this correlation, various simulation models are carried out using Computer Simulation Technology (CST) tool and the results obtained from each simulation is plotted as graphs using MATLAB. The simulations are performed, by varying the resistivity and thickness of the second Hydrocarbon layer at various target depths. The results obtained from the simulations, illustrate that the resistivity and the thickness of the Hydrocarbon have a direct relationship by a factor of 1 and that the electric field strength is affected by considering multiple layers of Hydrocarbons, so as to know that the electromagnetic waves transmitted reach the bottom layer of Hydrocarbon and not just only the first upper layer of the Hydrocarbon reservoir under the seabed floor.

Index Terms— Sea Bed Logging, Hydrocarbon, Controlled Source Electromagnetic, resistivity, thickness, Electric Field.

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

#### **1** INTRODUCTION

In 2000 the first SBL survey was performed offshore Angola [1]. Since then, the interest in electromagnetic methods for subsurface exploration has increased. Today, six years after, electromagnetic methods are attractive for the petroleum industry as complementary tools to seismic methods, or even standalone tools, for remote sensing of the subsurface. In a controlled-source EM (CSEM) survey [2], it is necessary to interpret the measurements in such a way that a prediction of the presence of hydrocarbons in the sedimentary layers can be made. The mechanism in seabed logging is thoroughly elaborated in the following sections along with the simulation models using Computer Simulation Technology (CST) and MATLAB for graph plotting.

In this study, we focus on reaffirming the relationship between the resistivity and thickness of the hydrocarbon reservoir in seabed logging application by considering the simulation of having double layers of Hydrocarbon. Also, this simulation will also let us determine whether the electric field strength is affected by the multiplicity of Hydrocarbon layers, which will assist us in understanding the nature of the hydrocarbon layer. By the presence of variations within the electric field intensity, we can easily predict whether the CSEM sounding is reaching the multiple layers of Hydrocarbon within the oceanic lithosphere. The seawater depth is kept constant at 2000m whilst the target depth of the second layer hydrocarbon is varied from 500m to 900m and for each target depth, the resistivity and thickness is varied from 100 units to 10 units by a unit interval of 20 units. The electric field is measured over different offsets using various simulation models and graphs.

## 2 METHODOLOGY

#### 2.1 The Sea Bed Logging Method

Sea bed logging uses active source electromagnetic (EM) sounding technique in detecting subsurface hydrocarbon. The CSEM method uses a horizontal electric dipole (HED) source to transmit low frequency (typically 0.01 – 10Hz) signals to an array of receivers that measure the electromagnetic field at the seafloor [5]. The method relies on the large resistivity contrast between hydrocarbon-saturated reservoirs, and the surrounding sedimentary layers saturated with aqueous saline fluids. Hydrocarbon reservoirs typically have a resistivity of 30-500  $\Omega$ m), whereas the resistivity of the over and underlying sediments is typically less than a few  $\Omega$ m. Both the amplitude and the phase of the received signal depend on the resistivity structure beneath the seabed [3].

By studying the variation in the resistivity contrast of the Hydrocarbon layer and thickness, as the transmitting source is towed through the receiver array, the effects of the Electric field at different offsets can be determined at scales of a few tens of meters depths of several kilometers. According to [11], as depicted by Figure 1, the receivers record the EM responses as a combination of energy pathways including signal transmitted directly through seawater, reflection and refraction via the sea-water interface, refraction and reflection along the sea bed and reflection and refraction via possible high resistivity subsurface layers.

In the following sections it will be demonstrated that this resistivity contrast and thickness has a detectable influence on SBL data collected at the sea bed above the reservoir. The effect of the reservoir is detectable in SBL data at an appropriate frequency of 0.1 to 10 Hz [9]. For this simulation, the frequency is set to 0.125 Hz and the current is 1250

A at the transmitter end.



Figure 1. A Schematic diagram of Seabed logging application showing the direct waves, air waves, reflected waves and refracted waves.

#### 2.2 Simulation Model 1

The simulation model proposed here contains no Hydrocarbon reservoir, so as to determine the Electric Field at varying offset during the absence of hydrocarbon reservoir for sea water depth at 2000m, 1000m and 100m. The result obtained from this model will then be used in comparison with the graphs obtained from simulation model 2 and 3, which contains the presence of Hydrocarbon reservoirs.



Figure 2. Simulation Model 1 showing no presence of Hydrocarbon reservoir.

#### 2.3 Simulation Model 2

The simulation model proposed here contains a 100m thickness of Hydrocarbon Layer with a resistivity contrast of 100  $\Omega$ m. In Simulation Model 2, the resistivity contrast and thickness is kept constant and the Electric Field strength against the varying Offset for seawater depth of 2000m is determined.



Figure 3. Simulation Model 2 showing presence of Hydrocarbon reservoir of 100m thickness, at target depth of 1000m.

#### 2.4 Simulation Model 3

The simulation model proposed here contains another 100m thickness of Hydrocarbon Layer with a resistivity contrast of 100  $\Omega$ m and is 20km wide than compared with the bottom layered Hydrocarbon reservoir which is 50km wide. In Simulation Model 3, the thickness and resistivity is varied to determine the Electric Field strength against the varying Offset for seawater depth of 2000m. The second layer of Hydrocarbon is 500m below the seabed floor and this target depth is varied from 500m to 900m with a 100m interval.



Figure 4. Simulation Model 3 showing two layers of Hydrocarbon reservoir present.

## 2.5 Assumptions

- The representation of the layer of Hydrocarbon reservoir is considered as a rectangular cuboid.
- For simulation 3, the resistivity of the second Hydrocarbon is varied with a constant thickness of 100m at each target depth.
- For simulation 3, the thickness of the second Hydrocarbon is varied with a constant resistivity of 100  $\Omega$ m at each target depth.

#### 3 RESULTS AND DISCUSSIONS

In this paper, the simulations are performed using Computer Simulation Technology (CST) tool and MATLAB R2009b. The simulation results are obtained using CST whereas, the plotting of the graphs and the result estimation were developed using MATLAB programming. Assumptions are being made while conducting this simulation. The environment is assumed to be free from internal and external disturbances, no bathymetry effect, no various shapes of hydrocarbon reservoirs as well as other aspects which we may find in real world survey. This work will be improved later by taking into considerations of real sea bed environment that has many challenges and obstructions in it.

The developed simulation as in Figure 2,3 and 4 is used to model a plane layer of the sea bed environment, by setting the sea water (of 2000m), sediments and size and location of the hydrocarbon trap. This model shall be used to understand the electric field variations with varying resistivity of the second Hydrocarbon at constant thickness of 100m, as well as, the other model in

Figure 4 where the thickness of the Hydrocarbon layer is varied at constant resistivity of  $100\Omega$ m. The parameters of each medium are set as follows:

Parameters	Air	Sea water	Oil	Soil
Electric Permittivity E	1.006	80	4	30
Electrical Conductivity (S/m)	1E-11	4	0.01*	1.5
Thermal Conductivity (W/K m)	0.024	0.593	0.492	2
Density (kg/m <sup>3</sup> )	1.293	1025	800	2600

\* Electrical conductivity will vary for simulation model 3.

This research aims to determine whether the electric field measured can easily determine the presence of two Hydrocarbon reservoirs by varying the resistivity (from  $100\Omega$ m to  $10\Omega$ m) and thickness (from 100m to 10m) of the second Hydrocarbon at different target depths (500m, 600m, 700m, 800m and 900m). All the parameters are maintained; only in simulation model 3, the resistivity of the Hydrocarbon layer is decreased gradually by  $20\Omega$ -m and starting from  $100\Omega$ -m for each seawater depth. Also, the same parameters are maintained but keeping the resistivity constant at  $100\Omega$ -m and varying the thickness of the Hydrocarbon layer by gradually decreasing it from 100m to 10m by 20 m decrement, for each seawater depth.

Results from Simulation Model 1 showing no Hydrocarbon present for seawater depth of 2000m.



Results from Simulation Model 2 where Hydrocarbon reservoir is present at target depth of 1000m with 100m thickness and resistivity of  $100\Omega m$ 



Figure 5: Electric field is plotted against the offset for sea water depth of 2000m, showing no presence of Hydrocarbon.

Figure 6: Electric field is plotted against the offset for sea water depth of 2000m with a Hydrocarbon layer at target depth of 1000m.

Results from Simulation Model 3 where second layer of Hydrocarbon reservoir is present at target depth of 500m with varying resistivity and thickness.



Figure 7: Electric field is plotted against the offset for sea water depth of 2000m, with varying resistivity of second layer Hydrocarbon reservoir (left) and with varying thickness (right) at 500m target depth compared with no presence of Hydrocarbon.

Results from Simulation Model 3 where second layer of Hydrocarbon reservoir is present at target depth of 700m with varying resistivity and thickness.



Figure 8: Electric field is plotted against the offset for sea water depth of 2000m, with varying resistivity of second layer Hydrocarbon reservoir (left) and with varying thickness (right) at 700m target depth compared with no presence of Hydrocarbon.

Results from Simulation Model 3 where second layer of Hydrocarbon reservoir is present at target depth of 900m with varying resistivity and thickness.





Figure 9: Electric field is plotted against the offset for sea water depth of 2000m, with varying resistivity of second layer Hydrocarbon reservoir (left) and with varying thickness (right) at 900m target depth compared with no presence of Hydrocarbon.

Graph showing second layer of Hydrocarbon reservoir is present at target depth of 900m with the bottom layer Hydrocarbon reservoir at target depth of 1000m.



Figure 10: Electric field is plotted against the offset for bottom layer Hydrocarbon at target depth 1000m compared with second layer of Hydrocarbon at target depth 900m. There is a sudden raise in the electric field at 1.5km and a decline at 3.5km due to the presence of the second layer of Hydrocarbon reservoir.

Graph showing second layer of Hydrocarbon reservoir (20km by 20km) is present at target depth of 500m with the bottom layer Hydrocarbon reservoir at target depth of 1000m compared with curve showing only Hydrocarbon reservoir (50km by 50km) at target depth of 500m.



Figure 11: Electric field is plotted against the offset for no Hydrocarbon layer with bottom layer Hydrocarbon at target depth 1000m compared with second layer of Hydrocarbon at target depth 900m. There is a sudden raise in the electric field at 1.5km and a decline at 3.5km due to the presence of the second layer of Hydrocarbon reservoir.

#### Tables

TABLEI.SHOWSPERCENTAGEDIFFERENCE OF FIGURE 6

Offset (m)	HC layer of target depth of 1000m (50km by 50km)
0-2,500	74%
2,501-5,000	77%
5,001-10,000	86%
10,001-25,000	73%

TABLE II. SHOWS PERCENTAGE DIFFERENCE

110011110		
Offset (m)	HC of target depth of 1000m (50km by 50km)	HC of tar- get depth 900m (20km by 20km)
0-2,500	74%	77%
2,501-5,000	77%	81%
5,001-10,000	86%	91%
10,001-25,000	73%	75%

#### TABLE III. SHOWS PERCENTAGE DIFFERENCE OF FIGURE 11

	HC layer at t	arget	HC layer at target			
Offset (m)	depth !	500m	depth	500m		
	(50km by 50	km)	(20km by 20km)			
0-2,500		89%		78%		
2,501-5,000		91%		83%		
5,001-10,000		96%		93%		
10,001-25,000		87%		76%		

Tabular results of Figure 7, 8 and 9 showing percentage differences for the two Hydrocarbon layers(bottom layer at target depth of 1000m and second layer of Hydrocarbon target depth is varied from 500m to 900m).

TABLE IV. SHO	OWS TARGET	DEPTH O	f 500m of	SECONE	HYDROC.	ARBON L	AYER.	

Offset (m)	100Ωm	100m	80Ωm	80m	60Ωm	60m	40Ωm	40m	20Ωm	20m	10Ωm	10m
0-2,500	78%	78%	78%	78%	77%	77%	77%	77%	76%	75%	76%	75%
2,501-5,000	83%	83%	82%	82%	82%	82%	81%	81%	80%	78%	79%	78%
5,001-10,000	93%	93%	92%	92%	92%	92%	91%	91%	90%	88%	89%	88%
10,001-25,000	76%	76%	76%	76%	76%	76%	76%	76%	76%	76%	76%	76%

TABLE V. SHOWS TARGET DEPTH OF 700M OF SECOND HYDROCARBON LAYER.

Offset (m)	100Ωm	100m	80Ωm	80m	60Ωm	60m	40Ωm	40m	20Ωm	20m	10Ωm	10m
0-2,500	78%	78%	78%	77%	77%	77%	77%	76%	76%	75%	76%	75%
2,501-5,000	82%	82%	82%	82%	81%	81%	80%	80%	79%	78%	79%	78%
5,001-10,000	92%	92%	92%	92%	91%	91%	91%	90%	90%	88%	89%	88%
10,001-25,000	75%	75%	75%	75%	75%	75%	75%	75%	75%	75%	75%	75%

TABLE VI. SHOWS TARGET DEPTH OF 900M OF SECOND HYDROCARBON LAYER.

Offset (m)	100Ωm	100m	80Ωm	80m	60Ωm	60m	40Ωm	40m	20Ωm	20m	10Ωm	10m
0-2,500	77%	77%	77%	77%	77%	77%	76%	76%	76%	75%	75%	75%
2,501-5,000	81%	81%	81%	81%	80%	80%	80%	79%	79%	77%	78%	77%
5,001-10,000	91%	91%	91%	91%	90%	90%	90%	89%	89%	87%	88%	87%
10,001-25,000	75%	75%	75%	75%	75%	75%	75%	74%	75%	74%	75%	74%

Table I and Figure 6, shows the percentage difference of the bottomed layer Hydrocarbon (dimensions 50km by 50km) at target depth of 1000m (green curve) with no Hydrocarbon present (blue curve). As apparent from the graphs and tabular results there is a significant difference between the two curves, the green curve is greater in strength due to the presence of the Hydrocarbon.

Tables II and Figure 10, shows the percentage difference of the bottomed layer Hydrocarbon with the no Hydrocarbon presence. From figure 6, we can easily decipher the presence of a single Hydrocarbon layer from that which has a double layer of Hydrocarbon (one at target depth of 1000m of dimensions 50km by 50km and other at target depth of 900m with dimensions 20km by 20km). From the results it is clear that there is a sudden raise at 1.5km and at 3.5km due to the presence of the other Hydrocarbon layer at 900m target depth. This sudden raise would not be present as visible in figure 6, the curve with single layer Hydrocarbon. Hence from the results, we can easily analyze the nature of the Hydrocarbon layer, whether it is singly layered or with a double layer of varying dimensions. Table III and Figure 11, shows the percentage differences are taken with no Hydrocarbon presence, firstly with a single layer Hydrocarbon at target depth 500m of dimensions 50km by 50km (green curve) and with double layer Hydrocarbon, with one layer at target depth of 1000m with dimensions 50km by 50km and the second is at target depth of 500m of dimensions 20km by 20km (red). From the graphs and from the tabular results, it is clearly illustrated that the green and red curves are both dissimilar and do not have any similarity, hence the electric field strength recorded for both the given scenarios are easily distinguishable and there can be no possibilities of any similarities between the two.

Table IV and Figure 7, shows the percentage differences when the second layer hydrocarbon (of dimensions 20km by 20km) is at target depth of 500m with the no hydrocarbon presence. The resistivity is varied from  $100\Omega m$  to  $10\Omega m$ , keeping the thickness constant at 100m and the same is performed by varying the thickness from 100m to 10m.

The same is performed with Table V and Figure 8, but the second layer hydrocarbon (of dimensions 20km by 20km) is at a target depth of 700m and in Table VI and Figure 9, the second
layer hydrocarbon is at target depth 900m, touching the bottom hydrocarbon layer, which is of target depth 1000m (refer to simulation model 3).

From Table IV to VI and Figure 7 to 9, it is clearly illustrated that the resistivity parameter is directly proportional to the thickness parameter of the Hydrocarbon layer. At higher values of resistivity and thickness, they are both equivalent to each other but only at 20m (of thickness) and/ or  $20\Omega m$  (of resistivity), there is a slight difference by 1%, which is considered negligible and has no significance. Therefore, the following relation can be considered:

#### R=kT

Where **R** is the resistivity of the Hydrocarbon, **T** is the thickness depth of the Hydrocarbon and **k** is a constant, which is 1, as from simulation experiments, these two properties are both equivalent to each other.

From figure 10 and 11, it is clearly observed that the electric field strength obtained with a single Hydrocarbon layer is different from the electric field strength obtained with double Hydrocarbon layers. Due to the presence of double layer of Hydrocarbon, the electric field strength obtained is greater than the electric field observed with single layered Hydrocarbon. At offset 1.5km and 3.5km, the presence of the second Hydrocarbon layer, causes a sudden increase in the intensity of the electric field strength and a decrease, clearly signifying, under the oceanic lithosphere, that there are multiple Hydrocarbon layers rather than just one. So therefore, from the conclusion of our simulation results and experiments, we have deduced and reaffirmed that the resistivity and thickness parameters of the Hydrocarbon are directly proportional to each other when we have single Hydrocarbon layer and/ or multiple Hydrocarbon layers. Also, another conclusion is made that electric field intensity is different for both single layered Hydrocarbon and for double layered Hydrocarbon, so via electric field strength, we may also determine the multiplicity of Hydrocarbon layers under the seabed floor.

#### **4** CONCLUSION

In this study, simulations were performed to attain a better understanding of the relationship between resistivity and thickness of multiple Hydrocarbon layers within the seabed logging model. In simulation model 2, the electric field intensity of a single Hydrocarbon layer at target depth 1000m with resistivity and thickness constant at  $100\Omega m$  and 100m is observed. In simulation model 3, a second Hydrocarbon layer is introduced with dimensions 20km by 20km at target depth of 500m. The target depth of this second Hydrocarbon is varied from 500m till lastly 900m and for each target depth, the resistivity and thickness parameters are varied from 100 units till lastly 10 units. From the simulations results obtained, it is clearly evident that both the resistivity and thickness of the Hydrocarbon maintain a direct relationship with each other and this relationship is very much evident at greater values than compared with negligible values. Also, results obtained with single layered Hydrocarbon were compared with double layered Hydrocarbon and it was noted that the results obtained from both the simulations, showed different electric field strengths, hence it becomes easy to decipher the multiplicity of the Hydrocarbon layers under the seabed floor, as the double layer Hydrocarbon had a greater electric field strength intensity than with the single layer Hydrocarbon. Therefore, as conclusion it is re-affirmed that the direct relationship between resistivity and thickness is retained for even multiple layers of Hydrocarbon under the seabed and that the electric field strength is easily affected by the multiplicity of Hydrocarbon layers.

#### ACKNOWLEDGMENTS

The authors would like to express their gratitude and appreciation to those who have contributed and facilitated towards the success of this paper and more particularly to Universiti Teknologi PETRONAS for the facilities provided to carry out the simulation work.

#### REFERENCES

- S. E. Johansen, H.E.F. Amundsen, T.Rosten, S. Ellingsrud, T.Eidesmo, A.H. Bhuiyan, "Subsurface Hydrocarbon Detected by Electromagnetic Sounding", Technical Article, First Break Volume 23, March 2005.
- [2] L.M.MacGregor, M.Tompkins, "Imaging Hyrocarbon Reservoirs using Marine Controlled-Source Electromagnetic Sounding," in Offshore Technology Conference, 2-5May 2005, paper OTC 17163.
- [3] Peter Weidelt," Guided Waves in Marine CSEM and the Adjustment Distance in MT: A Synopsis", Electromagnetic Colloquium, Czech Republic, Oct 1-5, 2007.
- [4] Andreis, D., L. MacGregor," Controlled-Source Electromagnetic Sounding in Shallow Water: Principles and applications", Geophysics Journal, vol. 73, pp 21-32, 2008.
- [5] Constable, S. C., and C. Weiss, 2006, Mapping thin resistors and hydrocarbons with marine EM methods: Insights from 1D modeling: Geophysics, 71, no. 2, G43–G51.
- [6] EN Kong H. Westerdhal "Seabed Logging: A possible direct hydrocarbon for deepsea prospects using EM energy". Oslo : Oil & Gas Journal,2002. -May 13, 2002 edition.
- [7] Constable, S., "Marine electromagnetic methods—A new tool for offshore exploration: The Leading Edge", 25, 438–444, 2006.
- [8] Eidesmo, T., Ellingsrud, S., MacGregor, L.M., Constable, S.C., Sinha, M.C., Johansen, S., Kong, F.N., Westerdahl, H., "SeaBed Logging (SBL), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas using controlled source electromagnetic sounding", First Break 20, 144–152, 2002.
- [9] Ellingsrud, S., Eidesmo, T., Sinha, M.C., MacGregor, L.M., Constable, S.C., "Remote sensing of hydrocarbon layers by SeaBed Logging (SBL): results from a cruise offshore Angola", Leading Edge 20 (10), 972–982, 2002.
- [10] Cox, C.S. Constable, S.C., Chave, A.D, Webb S.C., "Controlled source Electromagnetic Sounding of the oceanic Lithosphere," Nature Magazine, 1986, 320, pp 52-54.
- [11] P. M. van den Berg, A. Abubakar, T. M. Habashy,"Removal of Sea Surface Related Wave<sup>-</sup>elds from CSEM Data", Progress In Electromagnetics Research Symposium, Cambridge, USA, July 2-6, 2008.
- [12] L. J. Gelius, "Multi-component Processing of Sea Bed Logging Data", PIERS Online, Vol. 2, NO. 6, 2006.
- [13] Tomas Zeman, "BSS Preprocessing Steps for Separation Improvement", CTU FEE, Dept. of Circuit Theory, May 2000.

## Evaluation of Electric Field Components Response for Offshore Hydrocarbon Detection

Muhammad Rauf, Noorhana Yahya, Vijanth Asirvadam, Elisha Tadiwa, A'fza Shafie, Adeel Ansari, Hasnah Mohd Zaid, Hassan Soleimani

**Abstract**—This study is aimed to provide new insights on determining E field components for the sea bed logging application. In the scenario of increasing interest in CSEM since last two decades, it is desirable to evaluate the individual field component for the better hydrocarbon presence response. 1D forward modelling is carried out that have the capability to simulate offshore hydrocarbon detection using resistivity contrast analysis. Initially the study supports the proved Ex component of the in-line antenna with orientation along x direction as powerful response as compare to Ey and Ez field by providing a maximum of 93% difference with and without hydrocarbon. However the analysis of further results identifies that even with a weaker response, Ey component with same antenna orientation carries better information for hydrocarbon presence, a maximum of about 100% difference with and without hydrocarbon. Thus outcomes from this research have a clear potential for selecting the informative component of E-field to further the experiments for enhanced hydrocarbon detection.

Index Terms— Sea bed logging; Electromagnetics; 1D simulation; CST, Antenna geometry

#### 1. INTRODUCTION

Seismic method has been the most useful approach to detect offshore hydrocarbon reservoir since past four decades. The technique typically uses sound waves released through the water. The reflected wave by rock underneath the sea floor is observed to determine the presence of potential hydrocarbon [1]. However a step from its abilities, the method is found incapable of distinguishing the water and oil reservoirs which may increase the risk factor for well drilling. The marine Controlled Source Electromagnetic (CSEM) method, which is introduced in the beginning of this century [2, 3], given the name seabed logging have gained intensive interest for the offshore hydrocarbon detection. Since then CSEM has evolved from a past significant technology to become a promising emerging tool for de-risking hydrocarbon exploration as witnessed by a number of the consequent successful surveys and result [4][5][6][7].

CSEM for sea bed logging method uses the powerful mobile horizontal electric dipole (HED) towed at a certain height of about 20-40m above the sea floor. HED emits the high alternating current, low voltage waveform typically in the low frequency range of about 0.1 to 10Hz [8]. The underlying principle of marine CSEM is that EM waves attenuate more in a conductive medium like water saturated rocks while less in a

 Muhammad Rauf is currently pursuing Phd program in Electrical and Electronic Engineering Department in Universiti Teknologi PETRONAS, Malaysia.30750, E-mail: mrauf57@gmail.com.

- Elisha Tadiwa and Adeel Ansari are currently pursuing PhD program in Computer Information System Department in Universiti Teknologi PETRONAS, Malaysia.30750, E-mail: tadiwalee@gmail.com, adeel.ansari@hotmail.com
- A'fza Shafie, Hasnah Mohd Zaid and Hassan Soleimani are currently Sr. Lecturers at Applied Sciences Department in Universiti Teknologi PETRONAS, Malaysia.30750. E-mail: afza@petronas.com.my, hasnamz@petronas.com.my, hassan.soleimani@petronas.com.my

resistive medium like hydrocarbon saturated reservoirs due to skin effect. Moreover, some saturating fluids have varying conductivity levels too. Thus the generated EM waves while travelled down through several mediums are mainly refracted then received by the EM detectors [9], which measure the amplitude and phase relationship of the signal based on the resistivity contrast of the mediums. The received waves consist of mainly four kinds; direct waves, reflected and refracted waves (from subsurface), guided waves (through the hydrocarbon layer) and reflected and refracted waves (from sea air interface-Airwaves) as shown in the Figure 1 below. The most considerable is the guided waves, which carry the required data of hydrocarbon location. Therefore the guided wave must have enough energy to be received with reliable information of sub-surface.



Figure 1: Types of waves received at EM detector. Usually hydrocarbon beneath sea bed is uneven in surface. Therefore effects of several types of waves can easily be received at EM detectors depends on transmitter towing orientation, water depth and target depth.

It is now well know that the main aspect of CSEM is the diffusive behavior of electromagnetic fields in the conductive medium. Therefore at certain offsets usually near or shallow water with deep target, the secondary fields like direct and airwaves dominates as compare to the primary field that carries the relevant information of the subsurface. In the presented condition, it is desirable to evaluate the field response in the paradigm having less effect of the secondary waves.

Noorhana Yahya is currently the Dean of Faculty of Science & Information Technology and Professor at Fundamental and Applied Sciences Department in Universiti Teknologi PETRONAS, Malaysia.30750. E-mail: noorhana\_yahya@petronas.com.my.

Vijanth Asiroadam is currently Assoc. Prof. at Electrical and Electronics Engg. Department in Universiti Teknologi PETRONAS, Malaysia.30750. Email: vijanth@gmail.com

The response of the resistivity contrast effect depends on the HED excitation mood either galvanic or inductive [2]. Both modes are specifically the function of the HED geometry. The HED orientation can be defined in terms of the source receiver azimuth, the angle between the dipole axis and the line joining the source and receiver. At an azimuth of 90° (broadside geometry) inductive effects dominates due to skin depth whereas at azimuth of 0° (in-line geometry) galvanic effects are much stronger. Thus a study carried out to evaluate the response of a 1-dimensional model comprising of 100 m thick layer of the hydrocarbon bearing reservoir, overburden of 1000 m and overlain by an 800-m thick seawater layer. The outcomes mentioned that source at 0.5 Hz HED at the seafloor; the radial amplitudes are 20 times more sensitive to the presence of the thin resistive layer as compared to the azimuthal component. The obvious reason is that the radial dipole geometry carries a vertical component of electric current, which is guided back by the thin resistor, whereas the azimuthal dipole fields are largely horizontal and perturbed little [7].



Figure 2: Transmitter orientation; Azimuthal and Radial field for the transmitter geometry evaluation as in-line or broadside.

Increasing scope of 3D forward modelling is well documented for the sea bed logging data interpretation [10, 11, 12] because of extended target oil reservoir bearing is needed. However the experimental design considering merely frequency range, offsets and oil detection can be easily addressed by 1D modelling [7, 13]. Thus the sufficient works on evaluating field response with respect to the offsets and water depths encouraged this research to examine the components of the merely E field for the suitable selection among them for further the experiments. Simulation based modelling using CST uses Maxwell equations to solve the EM propagation with respect to the provided medium like air, hydrocarbon and sea water in the x, y and z plane as shown in Figure 3. In this study, the fix inline geometry of the dipole antenna along x direction is chosen. Therefore, the Maxwell equation below can be used to predict the propagation of EM waves (Electric Field Strength) in seawater (lossy dielectric) [14].

$$\begin{aligned} \mathbf{E}_{\mathbf{x}} &= \mathbf{E}_{0} \exp^{(j\omega t - \gamma \mathbf{y})} & (1) \\ \mathbf{E}_{\mathbf{y}} &= \mathbf{E}_{0} \exp^{(j\omega t - \gamma \mathbf{x})} & (2) \end{aligned}$$

Where j is the current density (A/m2) and  $\gamma$  is the propagation constant (m-1) in the time domain which can further be expressed in terms of  $\alpha$ , the attenuation constant in Np/m and  $\beta$ , the phase constant in rad/m.

$$\gamma = \alpha + j\beta \tag{3}$$

Even though the transmitter in-line orientation along x direction proved Ex and Hz components as the better choice [15, 16], however this simulation study identifies that the Ex component of E-field is not adequate enough for the better resistivity contrast information. Simulations based on 1D forward modelling are carried out using a conventional straight antenna. The results mention that although Ex field shows strong response however Ey component have the better capabilities as compared to Ex related to the hydrocarbon presence at far offset.

#### 2. MODELLING METHODOLOGY

The Marine CSEM method requires modelling tools for the characterizing, mapping and detection of hydrocarbon reservoir. 1D forward modelling have potential to investigate the effect of electromagnetic fields (E, D, B and H) with respect to several transmitting frequencies and current, distance between the source and the seabed, sea water depth, thickness of overburden, hydrocarbon and under-burden layer etc. In this research, Computer simulation technology (CST) is selected for modelling of the real seabed environment by using a conventional straight EM antenna for deep target hydrocarbon detection. CST uses a pattern to discritize Maxwell's equations at low frequency to investigate the resistivity contrast.

Certain steps are involved in order to generate the CST simulated mode. This investigation uses the parameters for 1D modelling inspired by recent studies. Arranging background parameters is considered as the first step that consist of setting model area 50 x 50 km to replicate the real seabed environment with fixed target position, air thickness at 500 m, sea water depth 2000 m, overburden thickness 1000 m, hydrocarbon thickness 100 m and under burden 900 m. All the layers are allotted with their specific conductivities and permeability values. Second step is to set parameters for aluminum antenna. In the case, length of 270 m, frequency of 0.125 Hz and a current of 1250 A, along with x direction is used as shown in Figure 2. Subsequently electric boundary conditions are applied then run low frequency full wave solver to initiate simulation for sea bed model [14].

It is evident from a number of the successful surveys that hydrocarbon usually exists as the thin-bed layer, approximately parallel to the sea floor [3, 4, 7]. In this simulation study, towing the transmitter over the conduction and resistive mediums is represented by the absences and the presence of hydrocarbon respectively. Since the amplitude and phase response of the resistive medium (presence of hydrocarbon) increases in the measured electric and magnetic field, thus it is taken as reference for the comparison in order to have a clear hydrocarbon presence response. Figure 3 shows a 3-dimensional sea bed model to perform the simulations.



Figure 3: Sea bed logging model showing different resistive mediums in 3D. Transmitter with current path (white arrow direction) is placed 30m above the sea bed while receivers are placed all along the axis, shown in blue line to measure field response at every offset.

#### 3. SIMULATION RESULTS

Since 1D modelling has the ability to evaluate the behavior of the CSEM method as the function of its individual electric field components. This research simulates each of them as per described in the modelling methodology. Figure 4 shows the plot for Ex, Ey and Ez component including the presence of hydrocarbon. All three components of electric field response are measured with conventional HED antenna within the proposed area (50 km x 50 km) having deep water (2000 m) where no airwaves effect takes place. The comparison of amplitude response clearly mentions that using the in-line geometry of the transmitter, Ex proves to be the strongest among other two. The same behavior is observed while eliminating the hydrocarbon layer. From the Figure 3, it verifies that strength of E field in x direction due to well guided by some critical angle and less attenuation.



Figure 4: E field component response with hydrocarbon. Ex and Ey shows a customary behaviour of response while Ez shows abnormal due to broadside transmission against the receiver across the water level surface.



Figure 5: E field component's response without hydrocarbon. Ex, Ey and Ez with apparently same response behavior in contrast with presence of hydrocarbon.

From above Figures 4 and 5, E field in x direction qualified as the most dominant with strong response for further the experiments related to seabed logging environment. Now this study will analyze the individual component of E-field with and without hydrocarbon with the percentage difference obtained between them. The differences in the response have the clear potential to finalize the informative signal in terms of hydrocarbon presence instead of strong amplitude response elements.



Figure 6: Ex field strength with and without hydrocarbon received at the detectors along the surface of subsurface



Figure 7: Ey Ex field strength with and without hydrocarbon received at the detectors along the surface of subsurface

Figures 6 and 7 show the evaluation of Ex and Ey field component in the presence and absence of hydrocarbon. Specifically Figure 7 demonstrates the effect of down going vertical component of E Field. The steep attenuation of Ey field without high resistive layer led to a broader difference when the hydrocarbon is placed as resistive medium. Table 1 and 2 shows the maximum difference with and without hydrocarbon cast by Ex and Ey component of E field.

Table 1: Percentage difference of Ex field with and without hydrocarbon vs source receiver offset

11	jaroearoon vo	source receiver	onset
Offset	Ex-Field	Ex-Field	Percentage
	(With HC)	(Without HC	Difference
0	2.01E-09	2.98E-10	85%
5009	2.49E-09	3.32E-10	87%
10003	4.59E-09	4.70E-10	90%
15012	1.16E-08	8.51E-10	93%
20006	6.44E-08	3.16E-08	51%
25000	1.97E-04	1.97E-04	0%
30009	6.35E-08	3.08E-08	52%
35003	1.15E-08	8.49E-10	93%
40012	4.58E-09	4.70E-10	90%
45006	2.49E-09	3.32E-10	87%
50000	2.01E-09	2.98E-10	85%

Table 2: Percentage difference of Ey field with and without hydrocarbon vs source receiver offset

,			
Longth	Ey-Field,	Ey-Field,	Percentage
Lengui	(With HC)	(Without HC)	Difference
5009	1.71E-10	4.25E-15	100%
10003	4.32E-10	5.32E-13	100%
15012	1.03E-09	8.40E-11	92%
20006	1.13E-08	1.92E-08	-70%
25000	1.54E-18	1.55E-18	-1%
30009	1.12E-08	1.89E-08	-69%
35003	1.03E-09	8.24E-11	92%
40012	4.31E-10	5.27E-13	100%
45006	1.70E-10	4.18E-15	100%

#### 4. RESULTS VALIDATION

The validation of the Ey component selection is supported by further simulations carried out with varying water depths from 2000m until 500m. The results strengthen the claim of Ey selection followed by deep until potential shallow water depth. The outcomes in terms of the percentage difference with and without hydrocarbon with respect to E field, Ey proved to be the most stable choice, as Ex showed a nearly undetectable difference in the presence of hydrocarbon while lowering water depth. The validation of the results as shown in Figure 8 until 11 which describes that the Ex field while lowering down the water depth, have severely affected by air waves. In the contrast, the vertical component of E field, Ey while having low magnitude response provided better delineation in terms of percentage difference for hydrocarbon presence.









Figure 10: Ex and Ey comparison with and without hydrocarbon at water depth 1000m



hydrocarbon at water depth 500m

#### 5. DISCUSSION

In shallow water depth, the possibility of the airwaves to produces an unwanted response is high. Thus for the investigation, initially deep water is focused to avoid air wave effects. The results are generated using 1D forward modelling that exhibits a realistic distribution of individual E field response. Figure 4 and 5 describe the E field components comparison as the function of offset, distance between transmitter and receiver. This 1D simulation supports the proved strength of Ex field over Ey and Ez component with and without hydrocarbon. The clear reason for x field domination is its diffusion pattern in the subsurface. Ex component enters in the subsurface with critical angle. This entering pattern helps the component to be guided well back from higher resistive layer and show up with strongest response. The x components of downward moving EM waves have less possibility to decay or attenuate thus can show a considerable value of the percentage difference between with and without hydrocarbon as shown in Figure 6 and Table 1.

In the contrast, the distinguishing results are shown in Figure 7 and Table 2. Although the y components of downward moving EM waves showed weak response, however the enhanced percentage difference between with and without hydrocarbon is observed as compared with Ex. The reason is the rate of attenuation of Ey field due to Eddy current in the conductive medium. Since Ey is purely a vertical component thus when the EM waves cross deeper into the medium without hydrocarbon, the eddy current turned incrementally weaker which lead the field response weaker as compare to the x field in a conductive medium. While in the presence of hydrocarbon, the direct waves dominate vigorously at near offsets. Therefore while neglecting the response at near offsets, a potentially promising wider percentage difference is observed at far offset. Eventually the enhanced percentage difference with and without the presence of hydrocarbon with y component of E-field showed a clear potential to lead the future research considering it as more better replacement of evaluating E field in x direction.

#### 6. CONCLUSION

In CSEM, it is desirable to evaluate the individual electric field component for enhanced hydrocarbon detection in seabed

logging application. This study focused on evaluating x, y and z component of specifically E field. The 1D simulation results identify that with the in-line geometry of the transmitter, Ex proved to be the strongest among other field component in the presence and absence of hydrocarbon. However in the comparative analysis between Ex and Ey component for better percentage difference with and without hydrocarbon, Ey proved to be the better choice with approximately 100% difference as far offset. The observed reason is due to its distinguishing trait as the vertical component of the E field while transmitter geometry along x direction, skin effect and steep attenuation in conductive medium which lead to bring up with a wider percentage difference in the presence of high resistive medium like hydrocarbon. While the stronger Ex field is found capable of distinguishing hydrocarbon up to maximum 93%.

#### ACKNOWLEDGMENT

The authors would like to express their gratitude and appreciation to those who have contributed and facilitated towards the success of this paper and more particularly to PRGS fund, Universiti Teknologi PETRONAS for the facilities provided to carry out the simulation work.

#### REFERENCES

- [1] Dyke Kate Van, "Fundamentals of petroleum" Texas: Petroleum Extension Services, 1997. Fourth Edition
- [2] Eidesmo, T., Ellingsrud, S., MacGregor, L.M., Constable, S., Sinha, M.C., Johansen, S., Kong, F.N., Westerdahl, H., 2002. Sea Bed Logging (SBL), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas. First Break 20, 144-152.
- [3] Ellingsrud, S., Sinha, M., Constable, S., Eidesmo, T., MacGregor, L., Johansen, S., 2002. Remote sensing of hydrocarbon layers by seabed logging (SBL): results from a cruise offshore West Africa. The Leading Edge 21, 972-982.
- [4] Constable, S., and L. J. Smka, 2007, An introduction to marine controlledsource electromagnetic methods for hydrocarbon exploration: Geophysics, 72, WA3-WA12.
- [5] Srnka, L. J., J. J. Carazzone, M. S. Ephron, and E. A. Eriksen, 2006, Remote reservoir resistivity mapping: The Leading Edge, 25, 972–975.
- [6] Smit, D., Saleh, S., Voon, J., Costello, M., Moser, J., 2006. Recent controlled source EM results show positive impact on exploration at shell. In: SEG New Orleans 2006 Annual Meeting.
- [7] Steven Constable and Chester J. Weiss. Mapping thin resistors and hydrocarbons with marine EM methods: Insights from 1D modelling. GEOPHYSICS, VOL. 71, NO. 2 (MARCH-APRIL 2006); P. G43–G51.
- [8] Janniche Iren Nordskag and Lasse Amundsen. 2007. Asymptotic airwave modeling for marine controlled-source electromagnetic surveying. GEOPHYSICS, Vol. 72, No. 6.
- [9] Ellingsrud, S., Eidesmo, T., Johansen, S., Sinha, M. C., Mac-Gregor, L. M., and Constable, S., 2002, Remote sensing of hydrocarbon layers by seabed logging SBL: Results from a cruise offshore Angola: First Break, 21, no. 10, 972–982.
- [10] Hesthammer, J., Stefatos, A., Boulaenko, M., Vereshagin, A., Gelting, P., Wedberg, T., Maxwell, G., 2010, CSEM technology as a value driver for hydrocarbon exploration. Marine and Petroleum Geology: 27 1872-1884.
- [11] K.C. Chiadikobi \*1, O.I. Chiaghanam \*2, A.O. Omoboriowo #3, M.V Etukudoh and N.A. Okafor. Detection of Hydrocarbon Reservoirs Using the Controlled-Source Electromagnetic (CSEM) Method in the "Beta" Field, Deep-Water Offshore Niger Delta, Nigeria. Int. J Sci. Emerging Tech. Vol-3 No 1 January, 2012.

- [12] G. Michael Hoversten, Gregory A. Newman, Nathan Geier, and Guy Flanagan. 3D modeling of a deepwater EM exploration survey. GEOPHYSICS, VOL. 71, NO. 5 \_SEPTEMBER-OCTOBER 2006, P. G239-G248.
- [13] Hesthammer, J., A. Stefatos, M.Boulaenko, S. Fanavoll, J. Danielsen, (2010), CSEM performance in the light of well results: The Leading Edge, 29, 29-34.
- [14] Noorhana Yahya, Nadeem Nasir, Majid Niaz Akhtar, Muhammad Kashif, Hasnah Mohd Zaid, Afza Shafie. Modeling of Antenna for Deep Target Hydrocarbon Exploration. Journal of Electromagnetic Analysis and Applications, 2012, 4, 30-41.
- [15] Deng Ming, Wei Wenbo, Zhang Wenbo, Sheng Yan, Li Yanjun, Wang Meng. Electric field responses of different gas hydrate models excited by a horizontal electric dipole source with changing arrangements. PETROL. EXPLOR. DEVELOP., 2010, 37(4): 438–442.

## Preliminary Overview of Potential Recommendations and Future Directions in Pursuit towards the Progression of Nanotechnology in Malaysia

#### Kasthoory Rajalingam

**Abstract** — The emergence of a new technology is characterized by the authenticity of its atypical features that makes it distinguishable and distinctive from other existing technologies. Not a single emerging technology is characteristically identical as the other. This form of an unusual complexity that embeds extremely promising potential is what makes it stand out as a new and maturing technology. Even if it is a stand alone or a hybrid formation or a composition of multiple disciplines put together, this technology becomes a resourceful contributor and a source of magnificent derivative innovations and services. One that does not need any introduction to scientists and researchers worldwide is none other than nanotechnology. Nevertheless in terms of awareness and level of knowledge absorption, the term "nano' lingers only in the minds of the common public as just a buzz word. The scientific activity that takes place at the microscopic nanoscale has spurred the gigantic interest of governments all over the world because of its soon to be seen progressive and massive visionary benefits. Rest assured, comes along with it is a degree of foreseeable destruction if the right kind of planning fall short in taking place. Additionally, the translation of lab nano-prototypes into fully fledged innovations has been slow paced resulting in sluggish infiltration into the commercial arena due to many identified barriers. Since the 6 year establishment of Malaysia's National Nanotechnology Initiative (NNI) in 2006, planning has been taking place; yet there has not been a policy plan for nanotechnology that has been crafted out to address these pitfalls. Therefore, this paper suggests a preliminary overview of potential recommendations and policy directions in pursuit towards the future progression and sustainability of nanotechnology in Malaysia.

Index Terms - Nanotechnology, policy, NNI, recommendations, multidisciplinary, innovation

#### 1. INTRODUCTION

set of recommendations have been meticulously outlined for the main aim of progressing and augmenting the development of nanotechnology to a higher level in terms of R&D policy, incentives, initiatives, endowments, safety, awareness, linkages, infrastructure, education, entrepreneurship and availability of statistics. A thesis study diagnosing the barriers of R&D and commercialization of nanotechnology was conducted from 2011 - 2012. These preliminary recommendations were generated based on those findings.

#### 2. RECOMMENDATIONS FOR FUTURE DIRECTIONS

The following are recommendations to enhance the development of nanotechnology in terms of research and development (R&D) and commercialization:

i. R&D policy concerning nanotechnology must take into consideration the multiplicity and diversity of economic sectors so that standards and procedures can be developed and molded to suit each sector appropriately. The R&D policy will also be dissimilar from county to country's environmental setting. Therefore, it is unlikely that one country's R&D policy will fit the needs of our country's nanotechnology R&D agenda. However, there are the several countries' R&D policies that our country can indoctrinate in terms of improving its effectiveness and efficiencies with the foremost mission of positively bringing out not just inventions but innovations that will

be successful in the marketplace. [It must be said that even developing countries have suffered pitfalls in terms of R&D and commercialization of nanotechnology]. Therefore, R&D policy need not just only spell out the needs of research and development in a particular area

but also pronounce clearly and explicitly that its ultimate agenda is to steer its way towards the commercialization of a R&D prototype or invention. With that in mind, the nanotechnology R&D policies should be meticulously planned with various parties involved (better still if every economy sector is represented) without allowing any room for ambiguities that can possibly cause mediocrities concerning nanotechnology prototypes and products.

- The nanotechnology R&D policy should take into ii. careful contemplation the time factor involved between research and commercialization with special deliberation on the requirement for clinical testing, robustness of the technology, compatibility and other post - development technical procedures that make nanotechnology a cut above the other technologies in terms of complexity. It should put in place the necessity to adhere towards a certain time limit to produce a product from the time of its basic research until applied research right up to off shooting a fully fledged product in order to prevent resources provided by the government from becoming unworthy of its cause. This action will also impact on the positive utilization of government resources.
- The nanotechnology R&D policy needs to manifest iii. a clause that if ever universities are in any way paralyzed by way of not being able to provide a prototype that suit the needs of the market, then they will be indoctrinated to amalgamate or bring their scientific ideas to the attention of a firm (prior development) who will be able to furnish them with the business perspectives and current market trend so as to prevent these prototypes or inventions from sitting in the shelf later. In other words, these firms will guide these universities (not in terms of scientific expertise) but with the current market needs prior development of a prototype. This way, the cost of making the product can be advocated earlier on and a straight forward financial rundown can be envisioned prior development. This will impact on the positive utilization of government resources endowed on universities.
- iv. In order to bring out an invention or a product on time for market release, the government should be able to provide incentives or assistances from the perspective of reducing the time factor. These incentives and assistances should come in the form of assigning product engineers and design specialist to work closer with scientists in the labs in the course of their research and development. These product engineers or design specialist can be

either the direct product of a university podium itself or the product of years of firm experience. However, the selection of these personnel should be based on their sharp expertise on nanotechnology. Scientists can request them for their projects with the condition that these research projects will be completed within a stipulated time frame.

- Since the primary reason why governments v. provide endowments to universities to perform research is for them to be innovatively productive meaning to say, bringing out innovations that will be successful in the market place. This is the era of commercialization. Except for the medical research of a particular drug to cure cancer and AIDS which takes numerous years, these government endowments are not for the purpose of allowing scientists to remain stationary at the phase of nanotechnology basic research but to progress to the stages of applied research. If ever this being the the government research case, then and development council should hold an investigative committee to enquire the solid reasons for delaying applied research considering that a great amount endowments have been supplied to research universities to bring out obvious applications. This would help construct a precipice between research in the university and commercialization in our country.
- An independent agency (separate entity) needs to vi. be set up to look into the safety of nanotechnology inventions prior to mass production or market release. Malaysia's recently initiated National Nanotechnology Directorate (NND) cannot absorb all these responsibilities. Exploring and probing into the safety of nanotechnology should be a sole and designated role or portfolio authorized to an independent agency. Since it will take a while for the public to become aware of the issues surrounding nanotechnology, the government should take full responsibility in determining whether or not these products are safe and healthy for public usage. At the outset or the beginning of its initiation, it would be advisable to seek the advice of other countries' experts who have experiential knowledge of its disastrous characteristics and behaviors so as to prevent the anomalies circulating nanotechnology from out bursting within our country's environment. These experts who will form the agency need to come from a multidisciplinary array of know how's

biology, entailing chemistry, physics and technology management to better understand the incongruities and discrepancies of this technology such as exploring how the technical and precisional intricacies that influence the nano chemical or nano - biological activity can be better managed. The agency will be given the sole authority for certifying nano products. Some products are labeled nano but not necessarily contain any nano materials or nano particles. There is no restriction or any provisions for this at the moment. Therefore, products that claims to be embedding nano material and which is created to boost the impact of the solution for the purpose intended for need to be thoroughly examined by the agency. If found that there is no nano material or particle contained in the product, the agency should issue a marker (sticker tag) on the product indicating this for the awareness of the people or even stating its possible side effects in the case of prolonged use. Or else, how will people come to know? However, this agency should be positioned under the technological wing of a higher ministry who will be able to monitor its activities.

vii. Awareness needs to be spread to the society as a whole on what nanotechnology is all about. The aspect of awareness can only be thoroughly said to have reached a saturation point if only there is at least a certain amount of understanding absorbed by the many different types of people who form the society at large. They have to be rationally warned of the side effects and dangers of unapproved and uncertified nanotechnology products or told to realize its gains in order to fully embrace its potential. Therefore, the government needs to start organizing two way interactive talks on the subject of nanotechnology and its role in society so that it is not entirely misconstrued. The talks should be structured in a way to be able to provide a surface view of the subject in lay man's terms. Schools, organizations, non research universities and shopping complexes are the best platforms to conduct this initiative but it has to be done in a continuous manner because people tend to forget in the long term. The next best platform would be the media - channels that are widely viewed by the people. Another appropriate method would be to construct unsophisticated yet informative booklet or CDs which will be easy to comprehend and understand - on the subject of nanotechnology and provide them to schools and

non research universities. This successful turnout of this agenda could possibly become a catalyst towards increasing the demand and supply of nanotechnology products in the future.

- viii. There should be a fixed criterion and requirement that a product needs to comply with in order to be categorized as a nano product. There should be an explicitly stated specification of what percentage (%) of nano component needs to be embedded in order to be declared as a nano product.
- Government initiatives and its missions with ix. regards to nanotechnology need to be in coordination with one another. Except for grants being issued to universities by the ministries, currently there are no obvious linkages between the two. Many research activities are being conducted by research institutes (based in universities) but they are not being scrutinized or monitored to find out how productive are these research activities. Therefore, there should be a comprehensive plan crafted by the endowment agency/ministry to track the progress of these research funded activities that includes making physical visits to scientists' labs/workshop where the research is being conducted; and grants to be segmented conditionally based on various phases of research outputs (and not be made in lump sum). This will augment the standards of nanotechnology research productivity in our country.
- Since many universities are said to be lacking very x. crucial infrastructure required to conduct nano these pertinent infrastructure research, for nanotechnology should be purchased and given to universities directly by the government instead of assigning the universities the responsibility to make the purchase themselves; and also the high allocation given restructure to universities. Meaning to say, the allocation can be restructured in a way that it will consider only the cost of materials, human capital (excluding equipment) and cost of maintenance. Being very exorbitantly costly, the universities have complained of not having enough from their allocation to set aside for paraphernalia. Therefore, this dilemma can be resolved if the endowing party provides the university in the form of paraphernalia instead of monetary. Once this matter is dealt with, then the government can proceed to examine to what extent has this

initiative made a difference to the standard of nanotechnology research productivity.

- xi. University PhD and MSc students coming from science backgrounds should be instructed to study the maintenance manual of the necessary equipment so that in lieu of suppliers, graduates will able to conduct the maintenance on the equipment as part of their practical training or hands on training experience. This will ensure that these postgraduates will understand first hand of the ins and outs of the functionalities of microscopy equipment used for nanotechnology.
- xii. Nanotechnology equipment that far exceeds the minimum cost threshold of government estimated expenditure can be placed in a centralized unit of each university. This can prevent the hassle and time depletion for one university from visiting another university to use specific equipment. Instead of purchasing equipment for each science faculty/department, one unit can be utilized by all science faculties according to time allocations. This aspect is taking in consideration the verity that these equipment are not necessarily used 24/7.
- xiii. Above paraphernalia, it is the scientists, researchers, technopreneurs and entrepreneurs who are greatly required in field of the nanotechnology. In order for many transformations from prototype to product to flourish, the number of hours and number of specific expertise need to be amplified.
- xiv. Encourage large local companies (e.g oil and gas) to prioritize nanotechnology research as part of their policy and provide opportunities and grants to PhD research students to work with them in nanotechnology. In this sense, professors from our country's premier universities can become affiliated with these companies on a contract basis. On top of this, these large companies can endow universities with research grants to conduct further research in nanotechnology.
- xv. The Registrar of Companies need to annually assess the number of companies involved in nanotechnology in this country and find out how many are legitimately registered nanotechnology companies. Physical annual visits need to take place to see for themselves whether these companies really exist.
- xvi. The Department of Statistics (DOS) and the Ministry of Science, Technology and Innovation need to measure the number of scientists/engineers involved in nanotechnology locally and globally,

number of student enrollments and number of degrees/majors conferred in the area of nanotechnology in universities (if any). Many universities claim that they have many graduates specializing in the field of nanotechnology; however there is no evidence to support this claim. Therefore, it would benefit the science community if these two (2) organizations were to carry out a census to measure these statistics even if the number is small; so that the science community is aware. If the cost of carrying out census of this sort is excessive, then the government needs to also consider this cost in their annual budget allocated for nanotechnology.

xvii. In addition to this, apart from the Department of Statistics (DOS) and the Ministry of Science, Technology and Innovation providing information on spending by each country on nanotechnology, it will be more beneficial if governmental statistical surveys and census reports begin tracking figures on 'how' each country spends the total amount of governmental spending on nanotechnology.

- xviii. Government should provide tax exemptions for SMEs that conduct R&D in nanotechnology. These tax exemptions should be offered for at least 10 years (not 5 years which is the minimum number given to SMEs conducting any type of R&D) since it takes a lot of high investment to venture into nanotechnology and returns could only be seen in a long term basis. However, in this case, the Registrar of Companies needs to collaborate with organizations like MIDA to monitor the progress and existence of these companies. It is of no use giving out tax exemptions to firms that claim to be conducting R&D in paper but do not conduct any type of R&D in reality.
  - From the perspective of funding, financial aid by xix. public and private to startups to be classified into different stratums; one of them being nanotechnology startups. Owing the fact that financial institutions have provided many forms of fiscal aid to SMEs during the past decade; it's time for these institutions to further prioritize their SME aid into different stratums. Meaning to say, focus should be directed towards the SMEs/startups involved in nanotechnology.
  - xx. The main reason for SMEs for not venturing into the field of nanotechnology is because high technology can be risky business. Therefore, the Ministry of Finance should be able to provide some

kind of incentive such as *"guarantor-ship"* or a helping hand if in case these companies fail.

- xxi. The Ministry of Education together with the higher tertiary universities should begin crafting ways to develop a creative curriculum for kindergarten students to study not the the basic but the "pre basic" aspects of nanotechnology. At present, it will not be seen as useful but in the near future, however, it will serve worthy in the long run. When pre- school students are able to play video games and computer games - that which was unimaginable 50 years ago, but which is comprehensible today. The integration of nanotechnology in the pre - school curriculum will be able to help boost their mind's eye of the movement of tiny particles (referring to atoms and molecules) picturesquely and serve as "pre foundation" or preparation to boldly take on science subjects when offered and taught in schools.
- xxii. The subject of nanotechnology should be incorporated into the undergraduate curriculum particularly in Management, Information Technology, Social Sciences degrees and also in the MBA curriculum. It should be provided as a core subject or offered to students as a minor/major option.
- xxiii. The missions of each individual member of an industry - academia partnership should be made clear right from the beginning. The different directions headed by both industry and academia should be able to ultimately reach a common goal. Even though research publications still measure up as a standard benchmark used in university rankings and there are very few scholars cum patentees in our country especially in the field of nanotechnology; over the years, academia has also recognized the importance of patenting. But it is the industry that is not willing to have a positive outlook towards the importance of academic publications. Industries in Malaysia should be made aware that many inventions or successful innovations have been the result of conversions from paper to prototype. Industries pride themselves with their own 'publications' but it is the university academic research publications that are certified as qualified. If ever industry is hesitant in disclosing data for the purpose of university publications, then this is where the relevant parties should identify their needs and together craft a constructive and productive work plan to address

this issue. As it is, there are so many partnerships, but no evident innovations as proof of output.

#### 3. CONCLUSION

These preliminary recommendations and potential policy directions stated above are the creative epitome of what needs to be done and holds lots of promise if correctly implemented by the rightful authorities.

 Kasthoory Rajalingam is currently pursuing a PhD degree at the Department of Science and Technology Studies, Faculty of Science, University of Malaya. Her research study focuses on nanotechnology R&D and commercialization. She is also a research fellow of the university. Email: kasthooryrajalingam@gmail.com

## Investigating Deep Target Reservoirs Using CSEM Methods

Tadiwa Elisha Nyamasvisva <sup>1</sup>, Noorhana Yahya<sup>2</sup>, Halabi Hasbullah <sup>1</sup> <sup>1</sup>Department of Computer & Information Sciences, <sup>2</sup>Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS,

Bandar Seri Iskandar, 31750 Tronoh Perak, Malaysia.

tadiwalee@gmail.com 1, noorhana\_yahya@petronas.com.my2, halabi@petronas.com.my3

Abstract—In this paper we describe a technique called Sea Bed Logging (SBL) as a tool to investigate deep sea target reservoirs. SBL an application of marine CSEM sounding, which can be applied to detect and characterize hydrocarbon bearing reservoirs in deep water areas. The basis of the approach is the use of a mobile horizontal electric dipole (HED) source and an array of seafloor electric field receivers. The transmitting dipole emits a low frequency electromagnetic signal both into the overlying water column and downwards into the seabed. The array of sea floor receivers measures both the amplitude and the phase of the received signal that depend on the resistivity structure beneath the seabed. A survey consisting of many transmitter and receiver locations can be used to determine a multidimensional model of subsea floor resistivity. A hydrocarbon reservoir can have resistivity perhaps 10–to 100 times greater. With an in-line antenna configuration the transmitted electric field enters the high resistive hydrocarbon layer under a critical angle and is guided along the layer. Electromagnetic signals constantly leak from the layer and back to the seafloor. The guiding of the electric fields significantly alters the overall pattern of current flow in the overburden layer. These capabilities are harvested in this paper in an effort to determine the depth at which hydrocarbon can be detected. A survey is done for two fields one with an offset of 25Km and another with an offset of 50Km. the depth of the Hydrocarbon reserves are varied from 1000m to 3000m at intervals of 100meters. The data collected is analyzed and summarized in the preceding sections.

\_\_\_\_\_

#### **1.0** INTRODUCTION

Measurements of electrical resistivity beneath the seafloor have traditionally played a crucial role in hydrocarbon exploration and reservoir assessment and development. In the oil and gas industry, sub-seafloor resistivity data has, in the past, been obtained almost exclusively by wire-line logging of wells. However, there clear advantages to developing noninvasive are geophysical methods capable of providing such information. Although inevitably such methods would be unable to provide comparable vertical resolution to wire line logging, the vast saving in terms of avoiding the costs of drilling test wells into structures that do not contain economically recoverable amounts of hydrocarbon would represent a major economic advantage. Several electromagnetic methods for mapping sub-seafloor resistivity variations have been developed [1,2]. Here we concentrate on marine controlled source electromagnetic (CSEM) sounding in the frequency domain. This technique has been successfully applied to the study of oceanic lithosphere and active spreading centers [1,3,4,5,8].In this paper we describe a technique called Sea Bed Logging (SBL), developed by Statoil [6], an application of marine CSEM sounding which can be applied to the problem of characterizing hydrocarbon bearing detecting and reservoirs in deep water areas.

The method relies on the large resistivity contrast between hydrocarbon saturated reservoirs, and the surrounding sedimentary layers [7]. Hydrocarbon reservoirs typically have a resistivity of a few tens of  $\Omega$ m or higher, whereas the resistivity of the over and underlying sediments is typically less than a few  $\Omega$ m [8]. It will be demonstrated that this resistivity contrast has a detectable influence on SBL data collected at the sea bed above the reservoir. The effect of the reservoir is detectable in SBL data at an appropriate frequency, and if the horizontal range from source to receiver is of the order of 2-5 times the depth of burial of the reservoir in typical situations [6].

The most crucial factors for the success of the SBL technique in practical applications related to hydrocarbon reservoirs is related to survey geometry. Subsea bed structure can be represented by a strata of horizontal layers as in figure 1. The upper layer represents sediments above a reservoir (the overburden). The middle layer, corresponding to a hydrocarbon reservoir, has resistivity perhaps 10 to–100 times greater, due to a high saturation of non-conducting hydrocarbon occupying much of the pore spaces. The deepest layer, below the reservoir, again has low resistivity due to its similarity to the overburden layer. Applying this to our model of a sub-seafloor structure containing a resistive hydrocarbon reservoir of 100m, we can deduce that the effect of the reservoir on the survey

results will depend strongly on the direction of flow of the currents generated by the transmitter or the direction of the E-fields.

The sections below discus the simulated model and the results of the data obtained herein the simulations.

#### 2.0 METHODOLOGY

The simulations were carried out to determine the depths at which deep target hydrocarbon reserves can be detected. The simulations are done using CST Studio suite 2009 covering an area of 50km by 50Km and the second model coves 100Km by 100Km. The models are of a fixed depth of 8600m, composed of layers of air, sea water, sediments (upper and lower burden) and hydrocarbon. The control is a model without hydrocarbon as shown in Figure 1. Compositions of the layers in the models are given in the sections below. Table 1 shows the parameter values for the components as used in the simulation.

MATERIAL	Epsil On	Mue	El. Cond	Rно (кс/м <sup>3</sup> )	Therm . Cond (W/K/ M)
Air	1.0 06	1	1E- 011	1.1	0.02 5
Hydroca rbon	4	1	0.00 1	800	0.49 2
Sea Water	80	1	4	1025	0.59 3
Sediment	30	1	1.5	2600	2

All the simulations were conducted in CST simulator with the parameter values as tabled in table 1. Simulations were run for 22 models of 25 000 meters offset and model height of 8 600meters. Table 2 below shows the stratification of the models from Model 1 without Hydrocarbon (No. H.C.), Model 2 with Sediment over burden of 1000m to Model 22 with sediment over burden of 3000m. Stratification of the Model 1, model 2 and model 22 are shown in Figures 1 below.



Figure1: Model1 without hydrocarbon, models 2 and 3 with hydrocarbon at different levels (1000m and 3000m)

Table 2 shows the stratification values in meters for the components making the model layers. The first model id the no hydrocarbon model followed by 22 models with hydrocarbon layers of 100m at varying depths of 100m from a depth of 1000m to 3000m.

	Mod el 1. No H.C.	MOD EL 2. 1000 METE RS	Mod el 3. 1100 mete rs	Mod el 4. 1200 mete rs			Mod el 22. 3000 Mete RS
Air	500	500	500	500			500
Sea Wat er	3000	3000	3000	3000	  -	  -	3000
Sedi ment (U.B.)	5100	1000	1100	1200			3000
H.C.		100	100	100			100
Sedi ment (L.B.)		4000	3900	3800			2000
Mod el Thickn	8600	8600	8600	8600			8600

Table 2. Stratification values for the models layers.

ESS				

The simulations are done to verify and determine the extent at which hydrocarbon can be detected at greater depth. It is known that at certain frequencies the detection of hydro carbon reservoirs is hindered due to the depth of the reservoir from the sea surface. These models have been designed to eliminate the chances of having air wave related anomalies. The sea water depth has been pegged at 3000 meters to ensure that the propagated waves are not distorted by the air wave. It is well known that at these depths 3000m and more, the air wave does not have any visible effect.

The results obtained from the set ups 1 through 22 will help to determine at a larger extent the cut off zone from which hydro carbon cannot be detected. The following sections shows the simulation results for frequencies of 0.25Hz, 0.125Hz and 0.0625Hz.

Figure 2 shows the far offset result of simulating H.C. at depths of 1000m to 3000m. At varying depths of 100m intervals. For a frequency of 0.0625Hz, a magnitude versus offset (MVO) analysis gives a return of more than 20% in resistivity values for the no hydrocarbon graph and the with hydrocarbon graphs from depths of 1000m only until 2500m. Below 2500m the confidence levels fall below 20%. This means that at these levels CSEM methods are not able to determine the presence of H.C. reserves. Tables 3, 4, 5 and 6 shows the percentages of H.C. variations in this set up.



Figure 2: 20000m to 25000m offset values for 0.0625Hz frequency with H.C. depths of 1000m. to 3000m.

Table 3.

HC.	1000	1100	1200	1300	1400	1500
Depth/	m	m	m	m	m	m
Offset						
0	315.57%	478.77%	207.49%	203.17%	186.21%	168.64%
400	242.25%	391.92%	155.07%	155.26%	144.05%	131.66%
800	188.45%	324.80%	115.15%	116.48%	109.08%	100.44%
1200	193.25%	323.40%	116.00%	110.00%	101.36%	92.51%
1600	237.03%	372.45%	147.73%	131.98%	119.46%	107.73%
2000	288.92%	427.49%	185.41%	157.27%	139.26%	123.67%
2400	294.55%	434.72%	189.55%	152.47%	133.59%	117.91%
2800	342.37%	487.52%	225.61%	175.08%	151.35%	132.27%
3200	453.51%	607.81%	309.63%	240.87%	207.19%	180.02%
3600	590.10%	754.96%	412.72%	325.09%	280.31%	243.56%
4000	642.04%	814.27%	454.36%	353.47%	304.67%	264.93%

Table 4.

HC. Depth/	1600m	1700m	1800m	1900m	2000m				
Offset									
0	78.27%	74.49%	177.18%	161.08%	66.94%				
400	61.00%	58.77%	146.37%	133.68%	53.93%				
800	46.11%	45.20%	120.05%	110.11%	43.87%				
1200	41.76%	41.51%	113.41%	104.04%	43.67%				
1600	47.32%	47.04%	126.42%	115.76%	50.91%				
2000	50.93%	50.64%	137.57%	125.54%	57.77%				
2400	45.03%	45.57%	132.17%	120.83%	55.56%				
2800	45.62%	46.31%	140.33%	127.87%	59.06%				
3200	59.29%	58.31%	170.77%	154.21%	71.41%				
3600	82.35%	78.01%	212.22%	190.06%	87.34%				
4000	85.93%	80.88%	223.45%	199.84%	89.02%				
	Table 5.								

HC. Depth/ Offset	2100m	2200m	2300m	2400m	2500m
0	52.35%	40.82%	82.94%	29.23%	22.36%
400	42.21%	32.94%	68.95%	23.18%	17.73%
800	34.43%	27.00%	57.00%	18.92%	14.60%
1200	34.45%	27.27%	54.23%	19.63%	15.40%
1600	40.35%	32.15%	60.18%	23.59%	18.69%
2000	45.96%	36.91%	64.81%	27.80%	22.35%
2400	44.17%	35.52%	61.93%	26.94%	21.75%
2800	46.90%	37.81%	64.24%	29.16%	23.74%

32	200	56.51%	45.58%	74.89%	35.89%	29.50%
36	500	68.61%	55.08%	89.51%	43.90%	36.24%
40	000	69.48%	55.55%	92.18%	44.36%	36.70%

Table 6	5.
---------	----

HC.	2600	2700	2800	2900	3000
Depth/	m	m	m	m	m
Offset					
0	16.94%	12.26%	-1.82%	-0.35%	-0.31%
400	13.44%	9.75%	-1.62%	-0.42%	-0.35%
800	11.21%	8.25%	-1.47%	-0.49%	-0.38%
1200	12.08%	9.09%	-1.38%	-0.49%	-0.36%
1600	14.86%	11.35%	-1.10%	-0.41%	-0.28%
2000	18.12%	14.13%	-0.60%	-0.21%	-0.09%
2400	17.76%	13.94%	-1.04%	-0.22%	-0.07%
2800	19.63%	15.61%	-0.74%	-0.06%	0.10%
3200	24.75%	20.03%	0.86%	0.39%	0.52%
3600	30.68%	25.16%	3.16%	0.98%	1.05%
4000	31.28%	25.84%	3.81%	1.13%	1.21%

Tables 3-5 shows that at depths of 1000m to 3500m, the percentage difference in the model without H.C. and the one with H.C. have a difference with a significance level of more than 20%. Table 6 show the percentages are less than 20% meaning that H.C. reserves are not identifiable at these depths at a frequency of 0.0625Hz.

Figure 3 shows the graphs resultant from the simulation of the same model at a higher frequency of 0.125Hz. Resultant MVO graphs Cleary shows that at 1000m depth, the significance level to determine the presence of HC is conspicuously high. The other levels have their graphs close to the no hydrocarbon graph. This means that at a frequency of 0.125Hz with the set up parameters described in table 1, hydrocarbon reserves cannot be detected.

Figure 4 shows the simulation with a frequency of 0.25Hz. From this graphs it can be seen none of the 21 levels of hydrocarbon depths can be seen at the far off set.



Figure 3: 20000m to 25000m offset values for 0.125Hz frequency with H.C. depths of 1000m. to 3000m.





Figure 4: 20000m to 25000m offset values for 0.125Hz frequency with H.C. depths of 1000m. to 3000m.

Figure 5: 40000m to 50000m offset values for 0.0625Hz frequency with H.C. depths of 1000m. to 3000m.

Table 7.					
HC.	1000	1500	2000	2500	3000
Depth					
Offset					
0	19.27	0.19	-	-	-
	%	%	0.31%	0.55%	0.18%
400	19.27	0.19	-	-	-
	%	%	0.31%	0.55%	0.18%
800	16.10	-	-	-	-
	%	0.02%	0.75%	0.83%	0.10%
120	13.85	-	-	-	-
0	%	0.23%	1.06%	1.05%	0.05%
160	11.63	-	-	-	0.00
0	%	0.48%	1.35%	1.28%	%
200	11.75	-	-	-	-
0	%	0.57%	1.52%	1.38%	0.01%
240	12.42	-	-	-	-
0	%	0.62%	1.67%	1.45%	0.03%
280	15.76	-	-	-	-
0	%	0.69%	1.84%	1.56%	0.03%
320	19.21	-	-	-	-
0	%	0.72%	1.92%	1.59%	0.03%

360	23.95	-	-	-	0.04
0	%	0.59%	1.56%	1.34%	%
400	26.53	-	-	-	0.04
0	%	0.46%	1.46%	1.22%	%

Figure 5 is resultant graphs of the simulation of simulation of the same model with an offset of 50 km. H.C. depths were varied in the same order as described before. The far offset values viewed from 40000m to 50000m are tabulated in table 7. The values are all less than 20% meaning to say that at a frequency of 0.0625Hz with a model of 50Km offset, H.C. reserves cannot be detected.

#### 3.0 DISCUSSION

At depths between 1000m and 3000m three frequencies were simulated for models 50km by 50km and 100km by 100km. the stratification of the models were the same for the two models. Hydrocarbon depths were varied between 1000m and 3000m at intervals of 100m. all together 21 simulations were done for each model. The frequencies were 0.25Hz, 0.125Hz, and 0.0625Hz. Several observations were made from the resultant MVO analysis.

It was observed that first, frequency levels play a significant role in CSEM based methods of scavenging for hydrocarbon reserves. At a depth of 1000m already a frequency of 0.25Hz could not give enough resistivity differences to ascertain the presence of Hydro carbon. In both the 100km and 50km models this frequency could not register the required 20% minimum resistivity difference between the with and without hydrocarbon simulations.

A frequency of 0,125Hz could manage a percentage of more than 20% at the depth of 1000m only for the 50km model and none at the 100km model. The two frequencies are therefore not small enough to penetrate the resistive overburden to see the H.C. layers at 25km and 50km offsets.

For the 0.0625Hz frequency a significant difference was noted only with the 50km model. Resistivity differences of more than 20% were registered for the depths more 1000m to 2500m. the significant difference shows that the smaller the frequency the deeper the skin depth. Beyond 2500m below the sea bed, resistivity values recorded were less than the 20% significance level. For the 100km model the result were otherwise. No hydrocarbon depth level managed to register more than 20% resistivity differences.

#### 4.0 CONCLUSION AND FUTURE WORK

The result gathered from this work shows that the detection of Hydrocarbon reserves using CSEM methods

depends on the depth of the reserves from the sea bed. The deeper the reserves the more difficult it is to find the reserves. The size of the field also plays a significant role in the detection of H.C. reserves. The larger the reserves the lesser the differences in resistivity levels at the far offset for models with H.C. and without H.C. This work has also shown that for deep target reserves, smaller frequencies are have a better skin depth which makes them more favorable for deep target reserves. At the resistivity level prescribed in Table 1, a frequency of 0.0625Hz will be best to detect H.C. reserves at depths less 2500m. beyond 2500m the resistivity of the over burden does not allow enough penetration to give a significant difference in the MVO analysis.

This work has maintained the resistivity values of air, water, sediment and H.C. the same. Future studies would vary these values and determine the extend at which the H.C reserves can be detected below the sea floor.

#### **5.0 ACKNOWLEDGEMENTS**

The authors would like to express their gratitude and appreciation to those who have contributed and facilitated towards the success of this research and most particularly to PRGS Universiti Teknologi PETRONAS for the facilities provided to carry out the simulation works.

#### REFERENCES

- Sinha, M.C., Patel, P.D., Unsworth, M.J., Owen, T.R.E. and MacCormack, M.G.R. [1990] An active source electromagnetic sounding system for marine use. Marine Geophysical Research 12, 29.68.
- [2]. Chave, A.D., Constable, S. and Edwards, R.N. [1991] Electrical exploration methods for the seafloor. Electromagnetic Methods in Applied Geophysics (ed. Nambighian, M.N.), Vol. 2, pp. 931.966, Society of Exploration Geophysicists.
- [3]. Young, P.D. and Cox, C.S. [1981] Electromagnetic active source sounding near the East Pacific Rise. Geophysical Research Letters 8, 1043.1046.
- [4]. Constable, S. and Cox, C. [1996] Marine controlled source electromagnetic sounding. II: The PEGASUS experiment. Journal of Geophysical Research 97, 5519.5530.
- [5]. MacGregor, L.M., Constable, S.C. and Sinha, M.C. [1998] The RAMESSES experiment . III: Controlled source electromagnetic sounding of the Reykjanes Ridge at 57°45¢N. Geophysical Journal International 135, 773.789.

- [6]. Eidesmo, T., Ellingsrud, S., Kong, F.N., Westerdahl, H. and Johansen, S. [2000] Method and apparatus for determining the nature of subterranean reservoirs. Patent application number WO 00/13046, filed August 1998.
- [7]. Yahya, N., Nasir, N., Akthar, M., Kashif, M., Daud, H., & Zaid, H. (2011). Magnitude versus offset study with Em transmitter in different resistive medium. *Journal of applied sciences*.
- [8]. Ellingsrud, S., Sinha, M., Constable, S., Eidesmo, T., MacGregor, L. and Johansen, S. [in press] Remote sensing of hydrocarbon layers by Sea Bed Logging (SBL): results from a cruise offshore. West Africa, in press.
- [9]. Farrelly, B., et al., "Remote characterization of hydrocarbon filled reservoirs at the Troll field by Sea Bed Logging," EAGE Fall Research Workshop, Rhodes, Greece, September 19 - 23, 2004.

## High efficient Fibre Plant Utilization by Multiple PON infrastructure based on Frequency Re-use Approach for Scalable FTTH Networks

Mohammad Syuhaimi Bin Ab-Rahman, Ibrahim M M Mohamed

**Abstract**— We have proposed a hybrid time-division multiplexing (TDM)/dense wavelength-division multiplexing (DWDM) scheme in order to build a scalable and flexible passive optical network (PON) that satisfies current and future bandwidth demands. Increasing efficiency and capacity in PON networks were our main objectives. To this end, we proposed to involve an arrayed waveguide grating (AWG) in the proposed configuration and exploit its two properties of wavelength cyclic (WC) and free spectral range (FSR). Wavelength cyclic property has been exploited in order to allow different bit rate optical line terminals (OLTs) to use the same frequency band and handle their traffic over a common fibre. Free spectral range has been targeted in order to allow each optical network unit (ONU) to handle its traffic via the same AWG input/output ports as well as to multiply system capacity. The performance of the proposed PON architecture has been examined by using OptiSystem and Matlab software packages under BER constraints. It has been verified that the architecture was able to transmit different bit rate services simultaneously (622Mbps, 1Gbps, 2.5Gbps and 10Gbps) over a 24Km shared feeder. This variety of service is provided to 16 passive remote terminals (PRTs) with 16 ONU group for each. Each group can accommodate up to 16 ONU, total of 256 ONU/PRT, resulting in over all system capacity 4096 ONU.

Index Terms— Arrayed waveguide grating, Fiber-to-the-home FTTH, Free spectral range, Optical hybrid schemes, TDM-PON, WDM-PON, Wavelength cyclic property.

### 1 INTRODUCTION

ODERN telecommunication network is customary divided into three main segments; access, metro and backbone segments. Access segment is the basic level in which service is provided to end users through central office while metro and backbone segments are the main two higher levels that are respectively responsible for multiplexing functions and long-haul transmission. Twisted pairs and coaxial cables have been traditionally exploited in access segment while optical fibers have been exploited in metro and backbone segments for long time due to their very huge capacity and low losses. Figure 1 shows schematically the aforementioned levels of a modern telecommunication network.

Along with the continuous development of new services and applications such as digital TV and VoD, the demand for broader bandwidth per user increases rapidly. Therefore, copper-based access networks become not able to satisfy such increasing demand. Thus evolution to fiber-based access networks becomes inevitable. The history of passive optical network standardization began in 1990 when there was an expected growth of bandwidth demand. In 1995, FSAN held its forum including some popular operators such as British Telecom, NTT and Bell South in order to develop a common

 Mohammad Syuhaimi (phone:+60192100019, syuhaimi@olsi.eng.ukm.my), Ibrahim Mohamed (phone:+60123602663, engibrahim\_2007@yahoo.com) are with Spectrum Technology Research Group (SPECTECH) Department of Electrical, Electronics and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia (phone: +603-89216448; fax: +603-89216146).

standard of optical access systems. Later, FSAN recommendations have been adopted by the International Telecommunication Union ITU [1] [2] [3] in the form of ITU-T G.983 standard which is known as broad band passive optical network (BPON). BPON provides 622Mbps and 155Mbps for downstream and upstream respectively over 20 km between 32 Optical Network Unit (ONU) and one Optical Line Terminal (OLT).



networks

In 2001, IEEE802.3 standard group began to invest their efforts in order to develop a P2MP PON standard (IEEE802.3ah) to be agreed up on later in 2004 [4]. IEEE802.3ah provides 1Gbps for both Downstream and Upstream respectively over 10-20 km between 16 ONU and one OLT. ITU-T study group 15 (SG15) was also working in parallel with IEEE802.3 standard group to develop new PON standard [5] [6] [7] called gigabit PON (G.984). GPON provides 2.5Gbps for Downstream and 1.25Gbps for Upstream between 64 ONU and one OLT.

Along with the ever-increasing demand for broader bandwidth as a result of the continuous development of new bandwidth-hungry services such as HDTV and 3D-video, evolution to next-generation PON becomes more and more insistent. To this end, IEEE [8] and ITU-T [9] have ratified their new standards, IEEE802.3av (10G-EPON) and ITU-T G.987 (XG-PON) in 2009 and 2010 respectively. IEEE802.3av specifies symmetric 10Gbps for downstream and upstream, and asymmetric 10Gbps downstream and 1Gbps upstream. ITU-T G.987 also specifies symmetric 10Gbps for both downstream and upstream, while asymmetric 10Gbps downstream and 2.5Gbps upstream. Although all of the aforementioned PON standards specify TDMA as an access technique, each one uses a different data framing technique. For example, G.983 specifies ATM for data framing whereas IEEE802.3ah specifies Ethernet as data framing technique while G.984 specifies another data framing technique called G-PON encapsulation mode (GEM). Although TDMA PON offers a cost-effective approach as service is provided over a common wavelength, its main drawback is the sharing-nature which posses a real challenge towards future upgrade. Figure 2 shows a schematic diagram of TDM-PON.



Another technology competing towards next-generation optical access adoption is WDM-PON. WDM-PON was initially proposed based on exploiting the large counts of wavelength available in optical fibers. It is deemed as more securable than TDM-PON as it provides a virtual P2P connection (dedicated pair of wavelength for each ONU). Additionally, it is more scalable than TDM as it ensures self upgradeability for each ONU. Moreover, it allows coexistence over an open access environment where optical distribution network can be shared by different operators. However, the main challenges in adopting WDM-PON in the next-generation PON was based on cost consideration. Since WDM-PON relies on narrowband filtering, each ONU will be enforced to handle traffic over a stable pair of wavelength, which imposes additional requirements on the transmitter. Figure 3 show a schematic diagram of WDM-PON. To exploit high data rate provided by TDM-PON and large number of wavelengths offered by WDM-PON, a hybrid TDM/WDM-PON was proposed. In accordance with the wavelength grid used, TDM/WDM-PON can be categorized either as TDM/CWDM-PON or TDM/DWDM-PON. The first commercial colorless gigabit TDM/WDM-PON was proposed by Lee et al [10] for Korea Telecom. The contribution of this paper is to develop a new optical access scheme that is able to satisfy current and future demands. The design was based on study the characteristics of different optical components including arrayed waveguide

gratings (AWGs) and the possibility to integrate them together in order to build a scalable and flexible TDM/DWDM-based passive optical network. Focus on AWG properties of wavelength cyclic (WC) and free spectral range (FSR) was made in order to: realize frequency re-using by several OLTs, enable each ONU to handle its traffic over a same input output port and multiply system capacity.



#### **2 AWG ARCHITECTURE AND PRINCIPLES OF OPERATION**

AWG is an optical device, mainly consists of an array of different waveguide lengths and two propagation regions (slabs). The waveguide array is placed in the middle of the two free propagation regions. The difference in length between any adjacent waveguides is constant. Wavelengths that propagate in the first free propagation region (first slab) will arrive at the input of the second propagation region (second slab) with different phases due to different waveguide lengths. As a result, wavelengths will be demultiplexed over the output ports of the second slab. Based on the property of reciprocity, AWG can also be functioned as a Multiplexer. Kaneko et al [11] proposed a thermal AWG that offers high stable performance over a wide range of temperature, which is targeted at PON design. A Silicon-based AWG model with 1.26dB insertion loss has been developed by C2V Company. However, the main drawback of this model was its large size structure. Tippinit and Asawamethapant [12] spent their valuable efforts to overcome this drawback by proposing a new design through which a small size AWG with a low insertion loss (1.09dB) was obtained.

#### 2.1 AWG Characteristics

Kaneko et al [13] reviewed the two main properties provided by the AWG. The first one is Wavelength Cyclic Property (WC) while the second one is Free Spectral Range Property (FSR). Wavelength Cyclic Property is the one in which a shift to the input port of the NxN AWG is followed by an opposite shift to the output port. In other words, if two identical wavelength aggregations incident on tow different input ports of an NxN AWG; they will be distributed among its output ports with non overlapping manner. One of the main objectives in this paper is to exploit this property in order to allow several OLTs to use the same frequency band, which multiplies the utilization of the fiber. Figure 4 shows a simple illustration of this property where a 4x4 AWG and 4 OLTs with identical wavelength aggregation for each are considered. Free Spectral Range is the property that identifies the periodic operating nature of the AWG. If two wavelengths enter from the same input port of an NxN AWG, they will come out from the same output port as long as they are separated by its periodic operating range (FSR). An AWG with large FSR is desirable in optical design to allow transmission over a wide range of bandwidth. Figure 5 shows a schematic diagram of this property where a 4x4 AWG with 400GHz Free Spectral Range is considered.



#### **3 PROPOSED PON ARCHITECTURE**

The proposed PON architecture shown in figure 6 is a hybrid TDM/DWDM-based PON. It is mainly based on exploiting AWG properties. According to wavelength cyclic property (WCP), if an aggregation of wavelength ( $\lambda 1 - \lambda n$ ) transmitted by OLT1 (Service Provider 1) enters the 1st port of the NxN AWG, it will be distributed among its output ports, each wavelength is then directed to a specific 1xJ power splitter which distributes it evenly among j-number of ONUs. If an identical aggregation of wavelength ( $\lambda 1 - \lambda n$ ) transmitted by OLT2 (Service Provider 2) enters the 2nd input port of the NxN AWG, it will also be distributed among its output ports without overlapping with the 1st wavelength aggregation. Generally, if number m of identical aggregations ( $\lambda 1 - \lambda n$ ) transmitted by m numbers of OLTs enter the N input ports of the NxN AWG, where m=n=N, they will be distributed on its output ports with no overlapping manner resulting to increase the utilization of the fiber as each AWG output port carries a replica of the same aggregation albeit from different OLTs. According to Free Spectral Range Property (FSRP), each ONU can handle its traffic via the same AWG input/output ports if downstream and upstream are separated by the periodic operating range of the AWG. Moreover, FSRP can be exploited in order to double the system capacity. For example, if 1st and 2nd FSRs are used by m OLTs, 3rd and 4th FSRs can be used by the same OLTs, which leads to double the over all system capacity. C-Band has been chosen for transmission because it is the lowest attenuation window for optical transmission. Moreover, transmission at 1550nm would allow using the well-known Erbium Doped Fiber Amplifier (EDFA) if future

necessity arises. In order to verify feasibility of the proposed PON, performance has been examined using Optisystem and Matlab software packages under bit-error-rate (BER) constraints. Simulation results and discussion is provided in the next section.



#### **4 SIMULATION RESULTS AND DISCUSSIONS**

Two simulations were conducted in order to verify feasibility of our proposed PON. In the first simulation, three different bit rate OLTs were considered, where 622Mbps, 1Gbps and 2.5Gbps were transmitted simultaneously. 10e-9 BER was chosen as a reference for operational requirement (maximum allowable value).

In the second simulation, 10Gbps OLT was included. Since forward error correction (FEC) is recommended at 10Gbps, a



BER of 2.9e-4 was chosen as a reference for operational requirement because it represents the pre-FEC BER required to achieve a post-FEC BER lower than 10e-9.

#### 4.1 First simulation

In this simulation, three different bit rate OLTs were considered, where 622Mbps, 1Gbps and 2.5Gbps NRZ Pseudo Random data were transmitted simultaneously at 0dBm. 16x16 AWG and 1x16 WDM-Demux with 1.26dB insertion loss for each was considered. 1x16 Power Splitter with 14.04dB Splitting/insertion loss was assigned. A frequency range of 1.5THz with 100GHz frequency spacing (193.1THz - 194.6THz) which represents an aggregation of sixteen wavelengths was allocated for each OLT as shown in figure 7.



Variable length single-mode fiber (feeder fiber) with 0.2dB/km attenuation and 16.75ps/nm.km dispersion was considered. PIN photodiode was chosen for signal reception due to its low biasing voltage and low cost.



Figure 8 shows BER against feeder fiber at 0dBm and three different bit rates, 622Mbps, 1Gbps and 2.5Gbps, where BER increases as feeder fiber increased until it reaches 10e-9 at 24Km/2.5Gbps, 30Km/1Gbps and 33Km/622Mbps, and then continuous to increase with the increase of the feeder fiber. Based on this result, we can conclude that transmission is possible over the proposed PON until 24Km (maximum allowable distance for transmission) as it is the lowest distance obtained for the simultaneous transmission.

#### 4.2 Second simulation

In this simulation, a 10Gbps OLT was included. Since it is expected for any optical line to lose several dBs of its budget when bit rate increased to 10Gbps; alternative solutions were taken into account in order to compensate for such losses. To do so, we decided to: increase the input power, incorporate a 10GHz optical band-pass filter prior to the receiver, and use an optimized gain APD photodiode instead of PIN photodiode for signal reception. Figure 9 shows BER against input power at 10Gbps, 24Km feeder while using APD with gain (M=2), where decreasing BER values are achieved as input power increased. However, we prefered to keep away from higher values (more than 10dBm) where undesirable effects related to non linearity in optical fibers are expected as well as to comply with safety standards. Therefore we just increased the input power to 2dBm where a BER of 3.11e-3 which is higher than that is required for the pre-FEC was achieved.



band-pass filter was incorporated prior to the receiver while optimizing the APD gain. Figure 10 shows BER versus APD gain (M) at 10Gbps, 24Km feeder and 2dBm, where a BER value of 2.78e-5 which is lower than that is required for the pre-FEC was achieved at M=6.

#### **5** CONCLUSIONS

A high scalable TDM/DWDM-based passive optical network has been proposed in which a variety of service can be provided to a large number of users. Different bit rat services can be provided simultaneously to 16 passive remote terminals over 24-Km feeders, each passive remote terminal can provide a variety of service to 16 ONU group. Each ONU group can accommodate up to 16 ONU, total 256 for each passive remote terminal, resulting in over all system capacity 4096 ONU. Incorporating AWG offers the following opportunities: First, the architecture enables each ONU to handle its downstream and upstream traffic through the same input/output ports. Second, the architecture is able to allow different bit rate OLTs to use the same frequency band and handle their traffic over a common fiber (Achievable Co-Existence and Increasable Fiber Utilization). Moreover, since each received wavelength refers to an individual OLT, each ONU can move to higher bit rate service simply by using a suitable filter and simple reconnection with out affecting the legacy ones (Achievable Self Upgradeability). This would allow gradual upgrade for the whole system to 10Gbps under seamless migration scenario. i.e. 622 Mbps will be dispensed with in the near future as it is the lowest bit rate service offered, later it will be followed by 1Gbps and 2.5Gbps respectively. Finally, the over all system capacity can be multiplied by exploiting several free spectral range periods (Increasable Scalability).



#### REFERENCES

- ITU-T, Recommendation G.983.1, "Broadband optical access systems based on Passive Optical Networks (PON)," 1998.
- [2] ITU-T, Recommendation G.983.2, "ONT management and control interface specifications for ATM PON," 2000.
- [3] ITU-T, Recommendation G.983.3, "A broadband optical access system with increased service capability by wavelength allocation,"2001.
- [4] IEEE 802.3ah, "Ethernet for the First Mile website," available at http://www.ieee802.org/3/efm/
- [5] ITU-T, Recommendation G.984.1, "Giagbit-capable Passive Optical Networks (G-PON): General characteristics," 2003.
- [6] ITU-T, Recommendation G.984.2, "Gigabit-capable Passive Optical Networks (G-PON): Physical Media Dependent (PMD) layer specification," 2003.
- [7] ITU-T, Recommendation G.984.3, "Gigabit-capable Passive Optical Networks (G-PON): Transmission convergence layer specification," 2004.
- [8] IEEE Std 802.3av-2009, "Physical layer specifications and management parameters for 10 Gb/s passive optical networks," 2009.
- [9] ITU-T Recommendation G.987 series, "10-Gigabit-capable passive optical network (XG-ON) systems" 2010.
- [10] Jie Hyun Lee, Seung-Hyun Cho, Han-Hyub Lee, Eui-Suk Jung, Jea-Hoon Yu, Byoung-Whi Kim, Sang-Hyeun Lee, Jai-Sang Koh, Back-Heung Sung, Suk-Jin Kang, Jin-Hee Kim, Ki-Tae Jeong, and Sang Soo Lee, "First Commercial Deployment of a Colorless Gigabit WDM/TDM Hybrid PON System Using Remote Protocol Terminator," *Journal of Lightwave Technology*, vol. 28, pp. 344-351, 2010.
- [11] A. Kaneko, S. Kamei, Y. Inoue, H. Takahashi, and A. Sugita, "Athermal silica-based arrayed-waveguide grating (AWG) multi/demultiplexers with new low loss groove design," *Elect. Lett*, vol.36, no.4, pp318–319, Feb. 2000.

- [12] Tippinit J and Asawamethapant W, "Optical properties improvement on A WG Achieved by Adding Transmission Star Couplers into FPRI" Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2012 9th International Conference, 16-18 May 2012, IEEE Conference Publication 2012.
- [13] A. Kaneko, T. Goh, H. Yamada, T. Tanaka, and L. Ogawa, "Design and applications of silica-based planar lightwave circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol.5, no.5, pp1227–1236, Sept. /Oct. 1999.

# Animation as a Problem Solving Technique in Mechanical Engineering Education Hana Jamal Badi<sup>1</sup>, Akram M. Zeki<sup>2</sup>, Waleed F. Faris<sup>3</sup>, Roslina Bt. Othman<sup>4</sup>

Abstract— It is apparent that engineering education requires an environment that is fully functional and practical in order to achieve the best results for students. Multimedia technologies specifically animation renders the classroom more attractive and interactive. When students are accustomed to studying in such a developed environment they will not only understand faster but will also be more creative and are more likely to generate more practical and inventive ideas. Moreover, animation delivers information in a creative manner that makes it difficult for students to forget thus resulting in better knowledge retention and improved academic performance.

Keywords— animation, engineering education, learning, multimedia technologies.

#### INTRODUCTION 1

OWADAYS in many universities lecturers are using advanced technological tools in the classroom to support their teaching and to help students to better understand the subject. Reference [13] found that "the growing concern of teachers to improve their theoretical classes together with the revolution in content and methods brought about by the New Information Technologies combine to offer students a new more attractive, efficient and agreeable form of learning". Animation brings things to life, especially in sensitive and complicated topics that require machines which are either unavailable or hard to imagine by solely using speech and text. Animation also provides a clear and accurate explanation of the subject. It encourages students to be more interactive and innovative when working on projects or assignments [18].

Animation can be most effective when used in areas of education that are complicated and has practical theories such as mechanical engineering and medicine. Reference [13] found "the case of Descriptive Geometry (DG) is particularly special, since the main purpose of this subject is not only to provide students with theoretical knowledge of Geometry and Drawing, but also to enhance their spatial perception". The group in this example used Macromedia Flash software as an implementing tool to design the animation. This particular software gives the designer a chance to create animations that may help in bringing the idea of the machine's operation and inner parts closer to the students' understanding and imagining. Moreover, the group also found that animation technologies were able to authorize the learners' interactive observation in most significant subjects in Descriptive Geometry.

Another critical area of education that can benefit from animation technologies is medicine. Reference [10] research study about 'Animation-assisted CPRII program as a reminder

tool in achieving effective one-person-CPR performance' aimed to use animation in Cardiopulmonary Resuscitation (CPR) training. The researchers conducted their study on two groups of trainees, intervention group who used animation in their training and a control group who relied on their previous usual training. The results of the study were highly significant and pointed out effective uses of animation in medicine. The analysis results of the group's study that used 30-point scoring checklist showed that the AA-CPRII group performed and scored better in comparison with the control group (p < 0.001). Psychomotor skills tested with the AA-CPRII group gave positive results more than the control group in hand positioning with a score of (p = 0.025), compression depth with a score of (p = 0.035) and compression rate with a score of (p < 0.001). In conclusion, Animation-supported CPR can be a useful tool in accomplishing successful one-person-CPR operation.

#### **2 DEFINITIONS**

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

Mechanical engineering is one important area of engineering education. Students of mechanical engineering are required to run experiments on complex machines and tools. Certain mechanical machines such as, automobiles, air conditioning equipment, spacecraft, and all types of engines are particularly expensive and are seldom made available to students for learning purposes. In addition, difficult and detailed topics require detailed explanations and properly prepared introductions in order to help students gain a better understanding of the topic. Therefore, such areas need technology and multimedia to support its delivery of information.

Reference [1] found that using dynamic media in learning

<sup>&</sup>lt;sup>1</sup>Hana Jamal Badi is currently pursuing masters degree program in Information System, Kulliyyah of Information & Communication Technology, International Islamic University Malaysia (IIUM), P.O. Box 10, Kuala Lumpur 50728, Malaysia. E-mail: hanajamal88@gmail.com

<sup>2</sup>Akram M. Zeki is currently an Assistant Professor at the department of Information System, Kulliyyah of Information & Communication Technology, International Islamic University Malaysia (IIUM), P.O. Box 10, Kuala Lumpur 50728, Malaysia. His research interest including: Multimedia, Watermarking and Image Processing. E-mail: <u>akramzeki@iium.edu.my</u>

<sup>&</sup>lt;sup>3</sup>Waleed F. Faris is currently a Professor in the department of Mechanical Engineering, Kulliyyah of Engineering, International Islamic University Malaysia (IIUM), P.O. Box 10, Kuala Lumpur 50728, Malaysia. E-mail: <u>wfaris@vt.edu</u>

<sup>&</sup>lt;sup>4</sup>Roslina Bt. Othman is currently an Associate Professor at the Department of Library and Information Science, Kulliyyah of Information & Communication Technology, International Islamic University Malaysia (IIUM), P.O. Box 10, Kuala Lumpur 50728, Malaysia. E-mail: roslina@iium.edu.my

computer science is more effective than static media especially when the subject is complicated and its content rich. Their research final analyzed results approved their preliminary hypotheses, where in the post-tests dissimilar results were achieved. The group who studied using dynamic D presentations resulted in a higher average total score (10.80, SD = 1.09) than the group who used static S materials (9.39, SD = 1.56) who also scored a higher average compared to the control group C (6.95, SD = 1.76).

Multimedia technology mainly focuses on interactive, computer-based functions that give users the chance to exchange ideas and knowledge through digital and printed components [4]. Animation, graphics, audio, video, and text are multimedia elements. Reference [2] found that "when used appropriately in your application's user interface, animation can enhance the user experience while providing a more dynamic look and feel. Moving user interface elements smoothly around the screen, gradually fading them in and out, and creating new custom controls with special visual effects can combine to create a cinematic computing experience for your users".

Reference [7] agrees that interactive multimedia gives a distinctive opportunity that allows users to share and exchange ideas through suitable mediums and interactive programs. Reference [7] also found that interactive animations which are designed for learning purposes in a form that offers all multimedia elements such as, text and graphics can create an environment that is attractive and allow users to exchange ideas and thoughts easily.

#### **3** COGNITIVE THEORY OF MULTIMEDIA LEARNING

In Fig. 1, [11] presented a model that shows the learning progress with multimedia lessons. They explain it as follows: the two processing rows which are words and pictures at the beginning illustrate the dual channel principles that every individual has which contains processing visual/pictorial material and auditory/verbal material. The limited capacity principle where a person can concentrate a limited number of instructions in each channel at one time is illustrated by the large Working Memory box placed in the middle of the figure. The active processing principle where learning takes place when the appropriate material is presented properly is illustrated in the figure by five arrows, namely selecting words, selecting images, organizing words, organizing images, and integrating.



Figure 1: Cognitive theory of multimedia learning by [11]; adapted from Mayer (2005)

However, no matter what type of educational technological environment is applied, positive learning can only be achieved if the teacher is well prepared.

Reference [12] in editing (Sharpe, R. and Pawlyn, J.) research work about e-teachers found that for effective learning it is essential that teachers are well prepared in order to be ready for the tasks that they are assigned to by performing the following attributes: teachers must have mastered the subject and have a strong background on what they teach, focus on content while explaining, and most importantly teachers must be friendly and have good communication skills with students. Those attributes can be most effective and can result in students' positive responses when applied in the case of teachers in the real world.

#### **4** MULTIMEDIA IN EDUCATION

Reference [8] with the contribution of Panella, O. G., Escudero, D. F., Zaragoza, M. P., and Portero, E. V. found that "multimedia is complimentary to other disciplines, such as Telecommunications, Computer Science, Audiovisual Communication or Graphical Design". As a result, multimedia can be imbedded into several qualified areas due to its power and generality.

Reference [3] found that popularity of animations comes from its unique feature which is lasting impression on the memory of the user, and it is the single most important reason for its efficiency. In addition, animation makes the learning environment more instructive, engaging, and works well for students. Reference [3] continues by saying that all of those practical elements actually provide a vast platform for students when learning essential and valuable lessons. Reference [3] agrees that "Aside from being great to watch, animation helps in getting a message across that sometimes words, audio or even video alone cannot".

In addition, [16] found that certain media formats are effective in representing contents better than others. Therefore, it is important to choose the media correctly, namely that which is suitable for the contents type in order to attract the targeted users attention which is a basic requirement for a successful learning environment.

#### **5** ANIMATION AS A SOLUTION TO ENHANCE MECHANICAL ENGINEERING EDUCATION

Reference [9] found that the approach of mechanical engineering focuses on specific components which are; design, professional practices, forces, materials, fluids, energy, and motion. Moreover, mechanical engineers formulate machines and configurations that utilize those components to provide a helpful function and valuable product. "Mechanical Engineering has been defined as the profession in which powerproducing and power-consuming machines are researched, deigned, and manufactured" [9].

Reference [5] found that "using visualization may have strong impact on the teaching process in engineering education". Presenting information in simple clear animations provides the educational environment with attractive and positive interaction that supports student understanding. Animation has a fundamental function especially in situations when the method type is non-analytical or in cases when the analytical method is very difficult for the students to grasp [5].

classroom forms an enthusiastic and attractive feeling to the students.

#### 6 EXAMPLES OF MULTIMEDIA USED IN ENGINEERING EDUCATION

The following are some examples that focus on technology and developed ideas that show a variety of teaching methods in mechanical engineering:

Reference [17] in their study about the importance of using multimedia as an enhancing tool for training engineering students found that economic pressures on universities and the appearance of new technologies have encouraged the creation of new applications and programs for conveying engineering teaching performance. Reference [17] used 3D technologies to create an interactive practical educational system for measuring mechanical parts. The educational system introduces three sections: 3D animation of mechanisms, assembly modeling of mechanisms and modeling of each part. The principle of mechanisms operation which was introduced by using 3D animations was accurate and well organized. In addition, [17] in their study result found that mechanical engineering students had positive impact and felt more interested. They also found that the combination of both technologies 2D and 3D can be highly effective in terms of learning and gives the students the chance to be innovative.

In practical subjects such as mechanical engineering, students may face problems in understanding the information. In many cases, the traditional manner of teaching is not the only issue but also the availability of the resources and especially the tools that are related to experiments. As a new learning tool, [15] in the University of Wolverhampton Virtual Learning Environment created a multimedia learning technologybased environment using non-text animations in which students could engage. The reason of designing their animation was because they discovered that some students face difficulties in certain essential materials models which are included in the Science/Engineering subject. In addition, knowledge achievement turns to be challenging when students skip or miss a class. As a result, students understanding, recognizing, scoring and general performance decreases. The results of [15] study using multimedia to support learning showed that students managed to proceed successfully in understanding complex atomic and crystal structures.

Reference [14] found "the use of Internet technologies in this teaching tool makes it possible to conjure visualisations that cannot be achieved using traditional teaching materials such as transparencies". They also agreed that "these virtual reality simulations and animations provide the capability of training students in NC programming and operations without the need to work on actual NC machines in the laboratory".

Reference [6] introduced their topic's main idea which is 'Novel Technique to Improve Power Engineering Education Through Computer-Assisted Interactive Learning' and what the whole procedure of the study is going to be about by using "PowerPoint slides and industry-furnished videos". Those tools are useful when introducing a new topic to students. For a better result in gaining knowledge, it is essential that the

#### 7 CONCLUSION

Multimedia plays an important role in enhancing education because it is flexible and has a variety of elements such as animation, audio, video, graphics, and text which if designed appropriately can be effective for learners in any learning situation. Animation as a multimedia element contains and describes all the other elements by forming a combination that is rich in presentation and delivering information. Well studied short animations that educate students on how a machine works together with supporting text for explaining the progress and the formula will form a very useful combination of multimedia presentation that will surely be easy to understand and hard to forget. In addition, it will enrich users experience, formulate better academic performance and raise the learning process level of problem solving method. In conclusion, multimedia technologies and specifically animation provides to students an attractive, effective and dynamic presentation for the subject's contents that makes the idea more close to understanding.

#### REFERENCES

- A. Holzinger, M. Kickmeier-Rust, and D. Albert, "Dynamic media in computer science education; content complexity and learning performance: is less more?," International Forum of Educational Technology & Society, vol.11, no.1, pp. 279-290, 2008.
- [2] A. I. Mac OS X Developer Library, "What Is Animation?," Mac OS X Developer veloper Library website, http://developer.apple.com/library/mac/#documentation/GraphicsIma ging/Conceptual/Animation\_Overview/WhatisAnimation/WhatisAnim ation.html.2008
- [3] A. Miller, "Incorporating Animation in Kids Education," 1888articles website, http://www.1888articles.com/incorporating-animation-inkids-education-0106867.html.2010
- [4] DegreeDirectory.org. "What Is Multimedia Technology?", http://degreedirectory.org/articles/What\_is\_Multimedia\_Technolo gy.html.2011
- [5] E. Perge, "ANIMATION: A PROMLEM SOLVING TECHNIQUE IN ENGINEERING EDUCATION I", http://www.eng.unideb.hu/userdir/dmk/docs/20072/07\_2\_02.pdf .2007
- [6] G. G. Karady and K. E. Holbert, "Novel technique to improve power engineering education through computer-assisted interactive learning," IEEE Trans. IEEE Power & Energy Society, Power Systems, vol.19, no.1, pp. 81-87, February.2004, doi:10.1109/TPWRS.2003.821006.
- [7] Hakikur Rahman, "Interactive Multimedia Technologies for Distance Education Systems", http://encyclopedia.jrank.org/articles/pages/6638/Interactive-Multimedia-Technologies-for-Distance-Education-Systems.html.2008
- [8] I. Cheng, L. V. Safont, A. Basu, and R. Goebel, Multimedia in Education: Adaptive Learning and Testing. Singapore: World Scientific, pp. 53, September.2010.
- [9] J. Wickert and K. E. Lewis, An Introduction to Mechanical Engineer-

ing. Third edition. Stamford: Cengage Learning, pp. 2-43, January.2012.

- [10] M. Choa, J. Choa, Y. H. Choib, S. S. Kima, J. M. Sungc, and H. S. Chunga, "Animation-assisted CPRII program as a reminder tool in achieving effective one-person-CPR performance," Elsevier Journal of Simulation & Education, vol.80, no.6, pp. 680–684, June.2009.
- [11] R. C. Clark and R. E. Mayer, E-Learning and the Science of Instruction. Proven Guidelines for Consumers and Designers of Multimedia Learning. San Francisco: John Wiley & Sons, Inc., pp. 35-37, January.2011.
- [12] R. Donnelly and F. McSweeney, Applied E-Learning And E-Teaching In Higher Education. London & Pennsylvania: Idea Group Inc (IGI), 2009.
- [13] R. R. Garcia, J. S. Quiros, R. G. Santos, S. M. Gonzalez, and S. M. Fernanz, "Interactive Multimedia Animation with Macromedia Flash in Descriptive Geometry Teaching," Elsevier Journal of Computers & Education, vol.49, no.3, pp. 615-639, November.2007.
- [14] S. K. Ong and M. A. Mannan, "Virtual Reality Simulations and Animations in a Web-Based Interactive Manufacturing Engineering Module," Elsevier Journal of Computers & Education, vol.43, no.4, pp. 361-382, December.2004.
- [15] T. Laoui and J. O'Donoghue, "Development of a Support Environment for First Year Students Taking Materials Science/Engineering," Routledge Journal of Research in Science & Technological Education. Available from: Taylor & Francis, Ltd., vol.26, no.1, pp. 93-110, April.2008.
- [16] T. Murphy, "Research Based Methods for Using PowerPoint, Animation, and Video for Instruction," Williams College website, http://www.ipfw.edu/celt/technology/PDFs/p372murphy.pdf.2004
- [17] X. Chen, S. Lin, S. Yang, and Y. Liu, "The Simulation of Engineering Drawing Practice Teaching Platform Based on Assembly Modeling," Proc. First Ann. International Conf. Information Science and Engineering (ICISE), pp. 3532-3535, December.2009, doi:10.1109/ICISE.2009.1284.
- [18] Y. K. Singh, T. K. Sharma, and B. Upadhya, Educational Technology: Teaching Learning. New Delhi: APH Publishing, pp. 1-13, 2008.

## GaD-eM: An Adaptive Game Design Model for Malaysian Higher Education (HE)

Mazeyanti M Ariffin

Abstract—this paper aims to identify existing issues in designing educational game and discuss the limitation of existing GBL framework or model. The outcome of this study is an educational game design framework called GaD-eM that addresses all limitations identified earlier.

Index Terms— GaD-eM, Game-based Learning (GBL), evaluation, game design, GBL framework/model, educational game design

\_\_\_\_\_

#### **1** INTRODUCTION

GAME is a competitive activity whether physical or mental that has set of rules or constraints, with the aim to entertain or reward the players [1, 2].Game continues to gain its popularity when it is no longer limited in entertainment industry but also starts to influence the advertising, analyzing, marketing, simulating and e-learning [3] and also in military [4], education [5] and healthcare [6,7]. The usage of game in education has introduced a new term. Game that being used in education is called game-based learning (GBL). GBL refers to the innovative learning approach derived from the use of computer games that posses educational value or different kinds of software applications that use games for learning and education purposes such as learning support, teaching enhancement, assessment and evaluation of learners [8]. Ref [9] further defined that in GBL, games are utilizes as medium for conveying the learning contents. It shows that the main aim of GBL is to emphasize on learning. Unfortunately there exist imbalance between maintaining level of excitement and conveying the learning content. Many GBL imparted too much excitement on the visual effect but failed to convey the learning contents [10]. Hence, it is critical to develop a GBL model that balances between these elements.

This paper is structured as follows: Section I covers the introduction, Section II discusses the background study and a literature review, Section III discusses on the proposed framework called Gad-eM, and finally in Section IV is the conclusion.

#### **2**. LITERATURE REVIEW

#### 2.1 Existing GBL Framework or models

The literature search is carried out using various electronic databases that is relevant. The following terms were used

("computer games" OR "video games" OR "serious games" OR "simulation games" OR "simulation systems" OR "game-based learning "OR "online games" ) AND ("evaluation" OR "framework" OR "model")

Table 1 listed 16 existing evaluation frameworks or models on GBL.

Framework / Model (s)
Game Object Model v1[11]
Game Object Model v2[12]
Four Dimensional Framework [13]
Kirkpatrick Four Level Evaluation Model[14] CRESST Learning Model[14] Affective Motivation Learning Model[14]
Framework of Heuristic Evaluation in MMORPG[15]
Framework for evaluating web based learning [16]
Adaptive Digital Game-based learning[10]
Design Framework for Edutainment Environment[17]
Adopted Interaction Cycle for Games[18]
The Engaging Multimedia Design Model for Children[19]
SIG-Glue Quality Criteria Framework[20]
Evaluation framework for GBL [21]
Educational Computer Game Design Model [22]

TABLE 1 : LIST OF GBL FRAMEWORK/MODEL

Framework of Heuristic Evaluation in MMORPG [15] and Framework for evaluating web based learning [16] are extended from the study done by [23]. Ref [23] introduced a heuristics evaluation that focuses on finding interface usability problems by applying Human Computer Interaction (HCI) technique.

Ref [10] reviewed four GBL frameworks including Design Framework for Edutainment Environment, Adopted Interaction Cycle for Games framework, and The Engaging Multimedia Design Model for Children and Game Object Model and create a new framework called Adaptive Digital GBL which adds the learner characteristics as acomponent. The learner characteristics in this framework focus on the learner psychology.

Ref [17] introduces Design Framework for Edutainment Environment. In this framework, she identified that learner is an essential focus when designing educational game. She identified 4 factors that are closely related to the learners that is storytelling, feel challenges while playing, interactivity and interfaces.

Ref [18] focused on interactivity between learners and game when proposing the Adopted Interaction Cycle for Games framework. In this framework, they described how the interaction between a user and a computer game happens in term of cognitive and physical user actions.

Ref [19] proposed The Engaging Multimedia Design Model for Children framework which focuses on engagement level of the learner. In her model, she claimed that high engagement level will be achieved if these features are presents such as simulation interaction, construct interaction, immediacy, feedback and goals.

Ref[11] introduced Game Object Model that significantly addressed pedagogy and game design. The Game Object Model has been further developed using theoretical constructs and developments to become Game Object Model II. Game Object Model II is far richer model that includes social elements in designing educational game.

Ref[14] reviewed three evaluation framework related to GBL including Kirkpatrick Four Level Evaluation, CRESST Learning Model and Affective Motivation Learning Model. These three frameworks are designed to evaluate learning outcome.

Ref [22] introduced Educational Computer Game Design Model. In this framework they described the process, issues and challenges in designing educational computer game for Malaysian classrooms.

Ref [20] proposed a framework called SIG-Glue Quality Criteria Framework that covers both pedagogical and technical criteria. The outcomes of this framework are the classification of games by learning purposes and an evaluation framework for assessing games.

Ref [21] proposed a general framework for evaluating GBL called Evaluation Framework for Games-based Learning. In this framework, they identified all potential attributes that to be considered when evaluating a GBL application.

Ref [13] developed a framework that addressed pedagogy which called Four Dimensional Framework (FDF). FDF described representation, pedagogy used, context and learner specification as four vital aspects in order to assist tutor's selection and use of games in their practices.

#### 2.2 Limitation of existing GBL frameworks or models

When designing an educational game it seems logical to design it from the pedagogical perspective because the main aim of using GBL is to motivate and engage learner with the intention that effective learning may occur [21]. Designing an educational game based on pedagogy is essential because unlike COTS, the main objective of educational game is to teach and reinforce learned concept rather than entertainment [10, 24]. Ref [25] and [21] reiterates that, creating an educational game should start from formalizing the relationship between learning theory, game design, play and development. Ref [25] and [21] argued that educational game have not been designed using any coherent theory of learning. Majority of educational games were designed based on the entertainment game which depends on the effective usage of multimedia elements to create the learning experience [8]. Despite the importance of pedagogical elements, the existing frameworks or models is still has insufficient pedagogical support in educational game design.

Furthermore, as to date very few models or frameworks address the learner characteristics or learner background in designing educational game particularly their customs, languages, culture, ethnicity and learning styles. Information about learners' background helps to refine the game design so that it can provides more effective learning experiences. Ref [26] claimed that cultural context influences comprehension regardless of an individual background. Ref [27] suggests that it is important to know the characteristics of the target audience when intending to use a game in learning. Presenting the learning materials in familiar concepts and situations from the learners' own cultures allow to increase the knowledge acquisition of learners [28]. Ref [29] stress that using topic that closely related to students lives could draw them into depth and complexity of a subject therefore it is important to present to students material which they can relate to in order to engage them in the learning process. Success of educational game is due to an emotional link established between the game and learner [30].

Ref [22] failed to identify available educational game in the market that is suitable to be used in Malaysian context due to our multicultural and multiethnic background.

Therefore, the aim of this work is to propose a framework that addresses those limitations.

#### 3 Proposed game design model

Figure 1 illustrated the proposed game design model called GaD-eM. GaD-eM is derived from the Educational Computer Design Model [22]) and Game Object Model [11]



#### Figure 1 : GaD-eM

GaD-eM is basically harmonizing pedagogy dimensions and game design components. Educational games need to contain different aspects; those that promote educational objectives and those that allow for realization of these educational objectives. Hence, an educational game area consists of different *components* that contain distinct *interfaces*. Game aspects that promote the educational objectives are called as '*abstract interfaces*' and represented by black circles while game aspects that support educational objectives are called as '*concrete interfaces*' and represented by white circles. The *game area* embodies all *components* (each with their own *interfaces*) and interfaces that define the interactive learning environment. In GaD-eM model, the *components* are represented by squares.

The game area component consists of seven motivational interfaces namely play, exploration, challenges, fun, engagement, relevance and illustration. The illustration component contains story line, game genre, rules, goal, content, combination, learning outcome and instructional strategy interfaces with the aspects components embedded within it. The virtual agent, graphics, sounds, scaffold, navigation, control mechanism and language interfaces make up the story line, appearance and playability of the game. The game genre interface represents the content of the game. The goal interface represents the learning outcome of the game; while the rules interface depict the instructional strategy. The abstract interfaces therefore represent the pedagogical elements while concrete interfaces represent game elements.

#### 3.1 Theoretical constructs of GaD-eM

Game is "a physical or mental contest that has specific rules, with the aim to amuse or reward the gamers" [1,2]. Ref [2] provides the following definition of a game or computer

game: "A game is an artificially constructed, competitive activity with a specific goal, a set of rules and constraints that is located in a specific context." As a result, an educational game should present play, fun, explorative, and challenge in an engaging environment. It is also must governed by specific rules with a specific goal that is located in a specific context.

Games provide situated experiences in which players are immerse in complex problem solving task [5]. Ref [2] and [13] defined instructional games as games that have been specifically designed or modified to meet learning objectives. As a result, an educational game should allow learner to apply their prior knowledge into the current knowledge in order to achieve a goal. This type of skill is called combination

Ref [31] defines game as "an activity that is voluntary and enjoyable, separate from the real world, uncertain, unproductive (in that activity does not produce any goods of external value) and governed by rules. Ref [32] said that computer games have the following characteristics rules, goals & objectives, outcomes and feedback, conflict/competition/challenge/opposition, interaction and representation or story.

#### CONCLUSION

As a conclusion, a GBL framework is proposed to address those limitations that have been identified in the existing frameworks or models. Thorough study has identified lack of pedadogy components in game design, failure to consider learner background and game genre as components that has potential to influence learning in Higher Education. This study hopes to implement the identified components in the game development and investigate the game effectivess in imparting learning in Higher Education in Malaysia.

#### REFERENCES

- M. Zyda. "From visual simulation to virtual reality to games". IEEE Computer, 2005
- [2] R.T. Hays, "The Effectiveness of Instructional Games : A Literature Review and Discussion. Orlando" : Technical Report 2005-004 Naval Air Warfare Center Learning Systems Division.2005
- [3] A. Frank. and N. Lundblad. "The New Role Of Gaming: How games move outside Computing: Technologies entertainment". *Entertain* ment and Applications, Artificial IFIP First International Workshop on Entertainment Computing (IWEC 2002), 2002.
- [4] M. Schneider, K Carley, and I.Moon, . "Detailed Comparison of America's Army Game and Unit of Action Experiments". Technical Report CMU-ISRI-05-139, Carnegie Mellon University, School of Computer Science, Institute for Software Research International,2007
- [5] K Squire., L.Giovenetto., B Devane, & D Shree,. "From users to Designers : Building a self-organizing game-based learning environment". *Techtrends*, vol 49 no 5, p 34-42.
- [6] J.L Lennon, "Debriefings of web-based malaria games", Simulation & Gaming. Vol 37(3), pp 350-356. 2006
- [7] S. Conrad, "Effective Digital game Design for Multi-Generational Learning".

https://files.pbworks.com/download/PenGOGpS0d/immersionfall09/ 26075333/Digital\_Games\_For\_Learning.pdf 2010

[8] S Tang, M Hanneghan and A El Rhalibi . "Introduction to Games-Based Learning", Games-based Learning Advancement for Multisensory Human Computer Interfaces: Techniques and Effective Practices . T.M. Connolly, M.H. Stansfield and E. Boyle eds. Idea-Group Publishing: Hershey. ISBN: 978-1-60566-360-9. 2009

- [9] C Conati, "Probabilistic Assessment of User's Emotions in Educational Games". *Journal of Applied Intelligence*, , 2002.
- [10] P. Tan, S. Ling, and C. Ting, "Adaptive digital game-based learning framework," *The 2nd international conference on Digital*, pp. 142–146, 2007.
- [11] A Amory, "Building an educational adventure game: theory, design, and lessons." Journal of Interactive Learning Research. Vol 12 2/3, p. 249-263.2001
- [12] A Amory. and R. Seagram. "Educational Game Models: Conceptualization and Evaluation." South African Journal of Higher Education. Vol 17 no 2, p. 206-217. 2003
- [13] S de Freitas & M Oliver. "How can exploratory learning with games and simulations within the curriculum be most effectively evaluated?" *Computers & Education*, Vol 46, p 249-269.2006
- [14] H F O'Neil, R Wainess, and E L Baker, "Classification of learning outcomes: evidence from the computer games literature". *The Curriculum Journal*, vol 16 no 4, December 2005.
- [15] S Song, and J Lee,. "Key factors of heuristic evaluation for game design: Towards massively multi-player online role-playing game". International Journal of Human-Computer Studies, vol 65,p 709–723. 2007
- [16] S Ssemugabi,& R de Villiers. "A comparative study of two usability evaluation methods using a web-based e-learning application". Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries, 2007.
- [17] Z C Embi, "A Case Study on the Implementation of Framework for Edutainment Environment", PhD dissertation, Faculty of Creative Multimedia, Multimedia Univ, Malaysia, 2005.
- [18] W Barendregt and M.M. Bekker. "Towards a Framework for Design Guidelines for Young Children's Computer Games". Proceedings of the 2004 ICEC, 2004.
- [19] N S Said. "An engaging multimedia design model". Proceeding of the 2004 conference on Interaction design and children: building a community 2004.
- [20] C Dondi, and M Moretti,. "A methodological proposal for learning games selection and quality assessment". *British Journal of Educational Technology*, vol 38 no 3 p 502–512.2007
- [21] T. Hainey, "Using games-based learning to teach requirements collection and analysis at tertiary education level," PhD dissertation, Univ of West Scotland, 2010
- [22] K. Osman and N. Aini Bakar, "Educational Computer Games for Malaysian Classrooms: Issues and Challenges," Asian Social Science, vol. 8, no. 11, pp. 75–84, Aug. 2012.
- [23] J Nielsen and R Molich. "Heuristic evaluation of user interfaces". Proceedings of the ACM HI'90 pp 249-256.1990
- [24] R E Pedersen, *Game Design Foundations*. First ed., Texas: Wordware Publishing 2003.
- [25] D W Shaffer, K T Squire, R Halverson, and J P Gee, "Video Games and the Future of Learning". http://www.academiccolab.org/resources/gappspaper1.pdf. 2008
- [26] M Steffensen, C Joag-Deve., & R Anderson, "A Cross-Cultural Perspective on Reading Comprehension". *Reading Research Quarterly*, vol 15 (1), pp 10-29, 1979.
- [27] E Bae. "The Effects of Trainee Characteristic on Learning Effectiveness in Improving Organizational Performance". *Performance Improvement Quarterly*, vol 15 (2), pp 234-245. 2002
- [28] J Haynes. "SIOP: Making Content Comprehensible for ELLs". www.EverythingESL.Net .

- [29] V Perrone, "How to engage students in learning". Educational Leadership, vol 51(5), pp 4-7, 1994
- [30] H Leemkuil., T Jong, R Hoog & N Christoph. "KM QUEST: A collaborative Internet-based simulation game". Simulation and Gaming vol 34 No 1, 2003
- [31] R Caillois, "Man, Play, and Games". New York: Schocken Books.1961
- [32] I S Mayer, S. Bockstael-Blok, W. and Valentin, C. "A Building Block Approach to Simulation: An Evaluation Using Containers Adrift", *Simulation & Gaming*, vol 35, pp 29-52. 2004

Mazeyanti M Ariffin is currently pursuing doctorate degree program in Information Technology in Universiti Teknologi PETRONAS, Malaysia. E-mail: mazeyanti2012@gmail.com

## A Dynamic Self-Adaptive Music-Inspired Optimization Algorithm for the Hippocampus Localization in Histological Images: A Preliminary Study

Ali Kattan, Rosni Abdullah

**Abstract**—The hippocampus is a structure in the medial temporal lobe of the brain that is involved in episodic memory function. The texture features of the hippocampus could give better differentiation between Alzheimer's disease and normal controls. The localization of the hippocampus structure in MRI histological images is considered as a multimodal global continuous optimization problem, which is solved by means of soft computing techniques using stochastic global optimization methods. Recently, the harmony search (HS) algorithm, a music-inspired optimization method, was introduced as a new soft computing rival. However, the overall performance of this algorithm is quite sensitive to the proper settings of its parameters prior to starting the optimization process. Many have proposed HS-based variants that promote self-adaptive parameter settings. In this paper we propose a new HS-based algorithm with dynamic and self-adaptive features. Since this work represents an early step prior to considering a full implementation on actual biomedical images, the proposed algorithm is tested using a multimodal global continuous optimization benchmarking problems rather than actual hippocampus biomedical images. Results demonstrate the superiority of the proposed algorithm against many other HS-based competing methods.

Index Terms—biomedical imaging, computational intelligence, evolutionary algorithms, harmony search, soft computing, meta-heuristic.

\_\_\_\_\_

---- 🌢

#### **1** INTRODUCTION

HE hippocampus is a structure in the medial temporal L lobe of the mammal brain involved in episodic memory function. In patients with Alzheimer's disease (AD), smaller hippocampal volumes measured on magnetic resonance imaging (MRI) correlate with worse memory function [1]. The hippocampus has long been known for its crucial role in learning and memory processes where it has recently been demonstrated that the volume of the hippocampus is an early biomarker for AD [2]. AD is a brain disorder that destroys brain cells, causing problems with memory, thinking, and behavior severe enough to affect work, lifelong hobbies, or social life. However, the accurate diagnosis of AD can be challenging, in particular at the earlier stage. Early diagnosis of AD patients is important because it allows early treatment with cholinesterase inhibitors, which have been shown to delay institutionalization and improve or stabilize cognition and behavioral symptoms [3]. Most biomedical image acquisition techniques of the hippocampus region have many problematic features that will hamper tasks like localization and segmentation of structure in such images. These include fuzziness of the hippocampus boundaries and the relatively large image size in addition to many others. Object detection in general would impose some strict requirements related to accuracy and execution speed [4].

The texture features taken only from hippocampus gives better differentiation between AD and normal controls. Therefore, the textures of hippocampus are much affected by AD [3]. The localization of structures in biomedical images is considered as a multimodal global continuous optimization problem and solved by means of soft computing techniques [2, 4]. A technique used for the proper localization of the hippocampus was initially presented in [4] and then used in [2]. The hippocampus is located by detecting, as landmarks, two regions which are usually well distinguishable within the structure: the pyramidal and granule cell layers, which belong to the Ammon's Horn (CA) and Dentate Gyrus (DG) regions, respectively as shown in Figure (1) top.

In this technique a 2D deformable model of a section of the hippocampus is made to fit the corresponding region of a histological image, to accurately localize such a structure and analyze gene expression in specific sub-regions. Once the model is defined, a similarity measure must also be defined that drives the model search towards the actual configuration of the object, i.e., that reaches its maximum when the image representation of the model is perfectly superimposed to the object as it appears in the image as shown in Figure (1). The problem then becomes a global optimization problem, that is the search of the model parameters, which maximize the similarity measure. The landscape searched by the optimization algorithm is usually strongly multimodal [2]. Therefore, the next crucial step to be taken is the selection of a good heuristic, which can deal with such a rough landscape both efficiently and effectively. Stochastic global optimization (SGO) methods, such as particle swarm optimization (PSO) and deferential evolution (DE) have been successfully used to locate the hippocampal region in biomedical images [2, 4]. However, it was shown in [2] that (DE) significantly outperforms other meth-

Ali Kattan, (PhD), member of IEEE, Lecturer at the Department of IT, Ishik University, Erbil, Iraq, PH-009647505335344.
E-mail: ali.kattan@ishikuniversity.net

Rosni Abdullah, (PhD) Professor at the School of Computer Sciences, Universiti Sains Malaysia (USM), Penang, Malaysia.
E-mail: rosni@cs.usm.my

ods including genetic algorithm (GA) and PSO. In the aforementioned method, active shape models (ASM) are used based on deformable models in medical image analysis [5, 6]. The method parametrically deformed the model shape to match as closely as possible the shape of the hippocampus in the region to be located in the biomedical image used. The model is moved and deformed by altering its parametric representation using an optimization heuristic, which maximizes a function that measures the similarity between the model and the object itself [2, 4]. From another perspective, the problem could be also considered as a minimization problem whereby the algorithm tries to minimize the difference between the model and the object itself.



Figure (1) Example of a hippocampus histological image (top) and a hippocampus model (bottom). The dotted lines in the model represent the lower and upper limits for the possible deformation of the model [4]

SGO methods such as GA are commonly used in computational neuroscience [7-9]. In order to solve global optimization of continuous functions, SGO algorithms proved to be effective optimization techniques as they do not require special conditions or mathematical properties of the objective functions [10]. A recent SGO method is the harmony search (HS) algorithm [11], which is similar in concept to other SGO methods such as PSO and GA in terms of combining the rules of randomness to imitate the process that inspired it [12]. The HS algorithm is a relatively young meta-heuristic that was inspired from the improvisation process of musicians and is used successfully for many optimization problems with continuous design variables [12-14]. However, the algorithm capabilities are quite sensitive to the settings of its parameters affecting its overall performance and its ability to converge to a good solution [10, 15].

In this work, the development of a new HS-based algorithm having self-adaptive features is introduced. The proposed algorithm enables the dynamic settings of some of the important algorithm's parameters using new quality measure. Since this work represents an early step prior to considering a full implementation on actual biomedical images, the proposed algorithm is tested using multimodal global continuous optimization benchmarking problems rather than actual hippocampus biomedical images. Results demonstrate the superiority of the proposed algorithm against many other SGO methods including recent variants of HS making it a potential candidate method for a hippocampus localization and detection in MRI histological images.

The rest of this paper is organized as follows: section 2 gives a background of the HS algorithm. Section 3 presents related works covering some recent HS-based algorithms with self-adaptive features. Section 4 introduces the proposed method along with the empirical results in section 5. Section 6 is the discussion and the conclusions are given in section 7.

#### 2 BACKGROUND

The original HS algorithm, referred to as classical hereafter, was introduced as an alternative optimization technique for linear programming, non-linear programming and dynamic programming [16]. The method can handle discrete and continuous variables with similar ease [11]. HS concept is based on the improvisation process of musicians in a band where each note played by a musician represents one component of the harmony vector. The harmony vector represents a solution vector  $\mathbf{x}$ , having the size N representing all musician notes associated with a harmony quality value, an aesthetic measure, as shown in Figure (2). The harmony vector is analogous to N dimension variables and the harmony quality represents the fitness function f(x). The computational procedure for the classical HS algorithm is given in Figure (3). For a complete description of the algorithm refer to [11].

The classical HS algorithm required statically setting a certain number of parameters prior to starting the optimization process. However, the algorithm capabilities are quite sensitive to these parameters' settings affecting its overall performance and its ability to converge to a good solution where these parameters need to be skillfully assigned in order to obtain good results [17].



Figure (2) The harmony vector and its mathematical representation

1	Initialize the algorithm parameters (HMS, HMCR, PAR, BW,
	MAXIMP)
2	Initialize the harmony memory HM with random values drawn from vectors $[x^{L}, x^{U}]$
3	Iteration itr=0
4	While itr < MAXIMP Do
5	Improvise new harmony vector x'
6	Harmony Memory Consideration:
1.000	$(\dot{x}_i \in \{x_{i1}, x_{i2}, \dots, x_{iHMS}\}$ with probability HMCR
	$x_i \leftarrow \{ x_i \in X_i \text{ with probability } (1 - HMCR) \}$
7	If probability HMCR Then
	Pitch Adjustment:
	$(\dot{x}_i \pm rnd(0,1) \cdot BW$ with probability PAR
	$x_i \leftarrow \{ x_i \text{ with probability } (1 - PAR) \}$
	Bounds check:
	$\hat{x}_i \leftarrow \min\left(\max(\hat{x}_i, x_i^L), x_i^U\right)$
	EndIf
8	If x' is better than the worst harmony in HM Then Replace worst harmony in HM with x'
9	itr= itr+1
10	EndWhile
11	Best harmony vector in HM is the solution

### Figure (3) The computational procedure for classical HS algorithm

#### **3 RELATED WORKS**

Two of the HS algorithm optimization parameters, namely the pitch-adjustment rate (PAR) and the bandwidth (BW) are considered to have a great influence on the quality of the final solution [10]. Many recent works have proposed "selfadaptive" HS variants that would automatically tune and find the best settings for these two optimization parameters achieving better results than that of the classical. However, many of these variants have also introduced new additional parameters, which in some cases would also require an additional effort to set their initialization values manually prior to starting the optimization process and depending on the problem considered. In addition, many of these rivals have forced a relationship between the value of these parameters and the current iteration count imposing a linear monotonic change that is bounded by a maximum integer value (MAXIMP), as in the algorithm of Figure (3), where this value is selected subjectively by the user to signal termination. Consequently, the selection of such integer value would affect the whole optimization process.

The Improved HS (IHS) [18] and Global-best HS (GHS) [19] were the first two early attempts to auto-tune these parameters. The value of these parameters would be computed in every iteration based on the current improvisation rate value,

which is bound by MAXIMP. Both of these algorithms have reported superiority over the classical HS in a number of domains. However, each algorithm has its own limitations and many have criticized the methods' approach used in adjusting these optimization parameters [10, 17, 20].

The self-adaptive harmony search (SHS) algorithm is "an almost-parameter-free harmony search algorithm" [10]. In this method some of the parameters are automatically adjusted according to its self-consciousness. The PAR is decreased linearly with from 1.0 to 0.0 as a function of the iteration count that is bounded by MAXIMP. The BW on the other hand was replaced altogether and the new harmony is updated according to the maximal and minimal values in the harmony memory (HM). It was indicated that such technique would result in the new harmony making better utilization of its own experiences. The BW is computed for each dimension variable based on the highest and lowest value of the HM for that dimension variable. The formulas used in SHS are given in Eq. (1) through (3).

$trial(i) = HM(int(rnd() \times HMS + 1))$	(1)
$trial(i) = trial(i) + [max(HM^{i}) - trial(i)] \times rnd[0,1)$	(2)
$trial(i) = trial(i) - [trial(i) - \min(HM^{i})] \times rnd[0,1)$	(3)

The technique used in SHS would avert the need for any initial setting involving the PAR and BW values. Based on a number of full-factorial experiments involving the harmony memory consideration rate (HMCR) and the HM size (HMS), it was found that a HMS size of around 50 and a HMCR value of 0.99 are the most suitable to use for a number of continuous optimization problems. The maximum iteration count for all the problems considered was set to a value of 100,000.

The self-adaptive GHS (SGHS) algorithm [15] is based on the GHS algorithm [19] and employs a new improvisation scheme with an adaptive parameter tuning method. Three parameters are adjusted dynamically during the optimization process in SGHS, namely the HMCR, the PAR and the BW. It is assumed that the HMCR and PAR values are normally distributed in the range of [0.9,1.0] and [0.0,1.0] respectively with mean values HMCRm and PARm. It uses a standard deviations of 0.01 for the HMCR and 0.05 for the PAR. The HMCRm and the PARm are initially set prior to starting the algorithm at 0.98 and 0.9 respectively. Then SGHS starts with a HMCR & PAR values generated according to the normal distributions of these two. After a specified number of iterations, referred to as learning period (LP), which was selected at a value of 100, both the HMCRm and the PARm are recalculated by averaging all the recorded HMCR and PAR values during this period and then the procedure is repeated. It was mentioned that such approach would result in appropriate values that can be gradually learned to suit the particular problem and the particular phases of the search process. The dynamic value of the BW is set in a fashion that is similar to that used in the IHS algorithm, i.e. as function of the current iteration count and MAXIMP value and as given in Eq. (4). To sum up, the SGHS algorithm still requires the initial value setting of BWmin, BW<sub>max</sub>, HMCRm, PARm and the new parameter LP. These parameters must be determined manually before starting the optimization process.

$$BW(itr) = \begin{cases} BW_{max} - \frac{BW_{max} - BW_{min}}{MAXIMP} \times 2itr \ if \ itr < MAXIMP/2 \\ BW_{min} \ if \ itr \ge MAXIMP/2 \end{cases} \dots (4)$$

Another recent self-adaptive HS variant is the harmony search with adaptive pitch adjustment (HSASP) algorithm [20]. In this algorithm, the bandwidth value is adapted dynamically using a technique inspired from the velocity clamping in particle swarm optimization. Arguing that the simultaneous dynamic change of both the HMCR and the PAR can cause a twist of global and local search (a contradiction to the concept used in the SGHS algorithm), the HSASP algorithm uses a high fixed value of 0.995 for the HMCR parameter while the PAR, as in SHS, is decreased linearly from 1.0 to 0.0 during the optimization process. It was indicated that a large HMCR value would increase convergence rate of the algorithm in most cases and provides better performance [10, 18, 19]. It was argued that the technique used in SHS [10] using the lowest and the highest values of the *ith* dimension variable in the HM, given previously in Eq. (1) through Eq. (3), would eliminate the benefits of the exploration of search space outside the current boundary confined within a minimum and a maximum current values and would result in non-uniform distribution of position of the new harmony. In the HSASP method the improvisation process also considers using the current high and low values of each dimension variable within the HM. However, the improvisation would introduce two new terms, the *range* and the value  $\lambda$  and as given in Eq. (5) through Eq. (8). The bandwidth parameter BW is replaced with  $\lambda$  × range where it is argued that stochastic oscillation in each dimension is restricted to the current range of harmony position and is regulated by the parameter  $\lambda$  with values ranging from 0.2 to 0.8.

range(i) = max(HM(i)) - min(HM(i))	(5)
$trial(i) \pm \lambda \times range(i) \times rnd()$	(6)
If $trial(i) < x^{L}(i)$ then $trial(i) = x^{L}(i)$	(7)
If $trial(i) > x^{U}(i)$ then $trial(i) = x^{U}(i)$	(8)

The proposed technique has shown to attain good results in comparison to other methods including SHS [10] and the classical HS [11]. However, it still requires several manual setting for the newly introduced parameter  $\lambda$ . The best values were found to be 0.3, 0.4 and 0.5 considering the problems selected. As for other parameters, the HMS was set at 50 with the memory initialized using the common uniform random initialization and the maximum iteration count was set manually based on the problem dimensionality.

There are two contradicting techniques to change the PAR in the methods introduced above. The PAR either starts from a minimum value and ends with a maximum one as in IHS, GHS and SGHS [15, 18, 19] or it starts from a maximum value and ends with a minimum one as in SHS and HSASP [10, 20]. Regardless of the technique, the PAR behavior is predetermined in these methods to be monotonic and does not take

into account the quality of solutions in the current HM. A larger value of the PAR would result in further modification to the newly created dimensional variable thereby enhancing the local exploitation ability of the algorithm, whereas a smaller value of the PAR would result in the new harmony vector to select its dimensional values by perturbing the corresponding values in the HM, thus enlarging the search area and diversity of the HM [10]. The main difference among these methods is choosing which one to take place first, local exploitation or enlarging the search area and diversity? i.e. start with large PAR value or small PAR value? There is no link between the "quality" of the current solutions and the PAR value setting.

#### **4 PROPOSED METHOD**

The proposed method utilizes the Best-to-Worst (BtW) ratio of the current HM. The concept of BtW was introduced earlier by the authors for the training of artificial neural networks using the HS algorithm [12, 21]. The words "best" and "worst" are part of the HS algorithm nomenclature whereby the algorithm basically tries to find the "best" solution among a set of solutions stored in the HM by improvising new harmonies to replace those "worst" ones as in the algorithm given in Figure (3). At any time the HM would contain a number of solutions including a best solution and a worst solution in terms of their stored quality measures, i.e. fitness function values. With minimization problems in mind, the BtW value is a value in the range [0,1] and as given by the ratio of the current best harmony fitness value to the current worst harmony fitness value in HM. This is expressed in Eq. (9) where a higher BtW value in this case indicates that the quality of current HM solutions is approaching that of the current best. If maximization problems are considered however, then the problem could be treated as minimization one by using the inverse of the fitness function. In both cases a higher BtW value indicates that the quality of current HM solutions is approaching that of the current best.

$$BtW = \frac{f(\bar{x}_{best})}{f(\bar{x}_{worst})} \qquad \dots (9)$$

The PAR value in the proposed method is adjusted dynamically based on the value of the current BtW ratio rather than the value of the current iteration count. This is shown in Figure (4) where as search progresses, the BtW value would eventually decrease owing to having better quality solutions in HM. This behavior is expressed in Eq. (10) and Eq. (11) where the PAR becomes a function of the BtW value and not the iteration count. Unlike the methods introduced in the previous section, the PAR change is not monotonic and is not bound by a MAXIMP value. The PAR is to increase or decrease in response to the quality of current solutions in the HM represented by the computed BtW ratio.



Figure (4) The dynamic setting of PAR in the proposed method

$$mslope = \frac{PAR_{max} - PAR_{min}}{0} = PAR_{min} - PAR_{max} \qquad \dots (10)$$

$$PAR = mslope \cdot BtW + PAR_{max} \qquad \dots (11)$$

 $PAR_{max}$  is set at the value of 1.0, as in many of the methods presented earlier, and  $PAR_{min}$  is set at a small value that is greater than zero. Setting  $PAR_{min}$  to zero might inhibit the pitch adjustment process completely in case of HM stagnation; a condition in which BtW becomes close to unity awing to having similar solutions in HM. If the BtW value decreases, indicating a range of different quality solutions, then the PAR value increases to reflect the local exploitation ability of the algorithm and cause further modifications to the newly created harmony. On the other hand if the BtW value increases, indicating a higher quality solutions within the HM, then the PAR value decreases to enlarge search area and diversity by causing the values of the newly created harmony to be selected by perturbing the corresponding values in the HM.

The pitch-adjusting process is accomplished by using a dynamic bandwidth value (DBW) that considers each harmony vector component (dimension variable) separately and as given in Eq. (12) and Eq. (13).

$$\begin{aligned} ActiveBW(i) &= C \cdot StdDev(x_{HM}^{i}) & \dots (12) \\ DBW(i) &= rnd(-Active BW(i), Active BW(i)) & \dots (13) \end{aligned}$$

The ActiveBW for the considered dimension variable is computed by calculating the standard deviation of the respective HM column. The DBW of the dimension variable is a random value confined within the positive and negative range of the ActiveBW for that dimension variable. This is similar in concept to settings used in the SHS and the HSASP methods presented earlier in that it considers the current high and low values existing within the HM for each dimension variable. The standard deviation in this case represents how much variation exists from the average of each component vector value in the HM. A value of C>1.0 would extend the dynamic bandwidth value so that it would not eliminate the benefits of the exploration of search space outside current boundary confined within a minimum and a maximum [20]. Based on a number of experiments, using a value of C=2.0 makes the performance of our proposed algorithm comparable to those presented earlier.

Termination in all other methods presented earlier is the same as that of the classical HS; specifying a maximum number of iterations as stipulated by MAXIMP value selection [10, 15, 17, 20]. The proposed algorithm also uses this standard termination condition whereby the total number of optimization cycles is bounded by using a MAXIMP value. The proposed method however adds an additional termination conditions that is "OR" ed with the standard one. With minimization considered, if the current fitness function value of the best solution within HM is less than a very small value *delta*, then this would also signal termination.

#### **5 EMPIRICAL RESULTS**

The localization of the hippocampus in histological images is considered as a multimodal global continuous optimization problem whereby the algorithm tries to minimize the difference between the model and the object itself and as shown in Figure (1). It was stated earlier that this work is an early step prior to considering a full implementation on actual biomedical images. Tests were conducted using two multimodal benchmarking functions rather than actual biomedical images. The benchmarking optimization functions are shown in Figure (5) and Figure (6). In addition to the function's 3D graph, which gives a visual idea about the function's surface nature, the function's equation, range and optimal value are also included. The Generalized Schwefel 2.26 function is considered deceptive in that the global minimum is geometrically distant, over the parameter space, from the next best local minima. Therefore, the search algorithms are potentially prone to convergence in the wrong direction. The Rastrigin function is a non-convex function used as a performance test problem for optimization algorithms. It is a typical example of non-linear multimodal function. This function is a fairly difficult problem due to its large search space and its large number of local minima. These two multimodal functions have been commonly used as optimization problems in the evolutionary computation field [22, 23] and are characterized by having many local optima in addition to single global optima. The number of local optima increases exponentially with the dimension of the problem. This appears to be the most difficult test for optimization algorithms [22].


Figure (5) Rastrigin's multimodal benchmarking problem



Figure (6) Generalized Schwefel 2.26 multimodal benchmarking problem

The value settings for the HMS and the HMCR where investigated thoroughly in SHS and HSASP [3, 10] and common values are used for these two in the proposed method. With minimization problems in mind, fitness function values that are less than delta = 1.0E-50 are considered to be practically zero and would cause immediate termination otherwise termination is bound by the MAXIMP value. Table (1) summarizes the parameter settings used in the proposed DSHS.

#### Table (1) Parameter settings for DSHS

Ν	HMS	HMCR	PAR <sub>min</sub>	PAR <sub>max</sub>	MAXIMP
100	50	0.99	0.1	1.0	6.0E+05

In each test the proposed DSHS algorithm was run for 30 times and the bootstrap-t statistical analyses were carried out with 95% confidence interval for all test functions to verify the method. Table (2), given at the end of the paper, shows the results of the proposed method in comparison to others. The ODE results column in this table represents the results of the opposition-based differential evolution (ODE) SGO method [24] and was obtained from [20]. ODE is a recent variant of DE and represents a state-of-the-art evolutionary algorithm. In case of HSASP, the reported results are associated with differential evolution is square brackets in its respective column in the table. The best-attained value with its associated iteration count and the last iteration for which the proposed DSHS algorithm terminated are given in the last column.

## **6** DISCUSSION

As given in Table (2), the optimization results obtained using the proposed method were superior in comparison to other SGO methods. In the Rastrigin problem the proposed DSHS method was able to terminate earlier based on the opted precision for the minimum fitness value (delta <= 1.0E-50), which could be practically considered as zero value. In all of the selfadaptive rival methods considered in this work, the MAXIMP value selection must guarantee that the method would be able to converge to a good solution in a number of optimization cycles that is less than MAXIMP. Such value could be manually selected based on "experience" however it could be also obtained empirically based on actual testing for the problems considered.

In order to explain the behavior exhibited by PAR in relation with BtW, Figure (7) shows the convergence graph obtained for one of the experiments for the Rastrigin benchmarking problem. Graphs (a) shows the PAR value while graph (b) shows the acceptance rate percentage. Both graphs are drawn against accepted improvisations. Because these graphs contain condensed data due to the large number of iterations, a trendline is included to show the general behavior where it is plotted as a sixth order polynomial. As expressed earlier in Eq. (11), PAR changes dynamically in response to the current value of BtW and is inversely proportional to BtW and as shown earlier in Figure (4). The increase in PAR values at the beginning of the optimization process indicates that the HM contains a wider range of solutions having dispersed fitness values. PAR is increased to trigger the local exploitation ability of the algorithm and cause further modifications to the newly created harmony using the DBW values based on the current active bandwidth. The next decline in PAR values was in response to having higher BtW values. This indicates that the quality of the current solutions in the HM is becoming close to that of the current best solution. In response the PAR value is decreased almost to the value of PAR<sub>min</sub> in order to enlarge search area and diversity by causing the values of the newly created harmony to be selected by perturbing the corresponding values in the HM. This has resulted in a higher improvisation acceptance rate as evident in the matching response

shown in graph (b) of Figure (7).

Considering the problem of localization of the hippocampus in histological images, the more recent method discussed in [2] have indicated that using the DE SGO method gave better results than those obtained using PSO in [4]. In Table (2), the proposed DSHS algorithm performed much better than the state of the art ODE (which is based on DE). This work is an early preliminary step towards the implementation of the proposed DSHS algorithm in the localization of the hippocampus in biomedical images. The results obtained clearly indicate that the proposed method is a potential candidate for implementation of localization of the hippocampus in biomedical images.





Figure (7) The convergence graph for the Rastrigin's multimodal benchmarking problem

# 7 CONCLUSIONS

The texture features of the hippocampus could give better differentiation between Alzheimer's disease and normal controls. The localization of the hippocampus structure in MRI histological images is considered as a multimodal global continuous optimization problem that using stochastic global optimization methods. This work represents a preliminary research towards the development of an alternative dynamic and self-adaptive stochastic global optimization method that is based on the harmony search algorithm. Testing considered two commonly used multimodal global continuous optimization benchmarking problem rather than actual hippocampus biomedical images. Results indicated the superiority of the proposed method in comparison to some recent rival methods. In addition, the proposed algorithm dynamic and selfadaptive features have resulted in a competitive performance in comparison to other harmony search based methods having self-adaptive features. Future work should consider the proposed method as a potential candidate for full implementation of localization of the hippocampus in biomedical images.

#### ACKNOWLEDGMENTS

This research is supported by UNIVERSITI SAINS MALAYSIA and has been funded by the Research University Cluster (RUC) grant titled by "Reconstruction of the Neural Microcircuitry or Reward-Controlled Learning in the Rat Hippocampus" (1001/PSKBP/8630022).

#### REFERENCES

- [1] T. d. Heijer, F. v. d. Lijn, M. W. Vernooij, M. d. Groot, P. J. Koudstaal, A. v. d. Lugt, G. P. Krestin, A. Hofman, W. J. Niessen, and M. M. B. Breteler, "Structural and diffusion MRI measures of the hippocampus and memory performance", *NeuroImage*, vol. 63, pp. 1782-1789, 2012, doi:http://dx.doi.org/10.1016/j.neuroimage.2012.08.067.
- [2] P. Mesejo, R. Ugolotti, F. D. Cunto, M. Giacobini, and S. Cagnoni, "Automatic hippocampus localization in histological images using Differential Evolution-based deformable models", *Pattern Recognition Letters*, vol. 34, pp. 299-307, 2013, doi:http://dx.doi.org/10.1016/j.patrec.2012.10.012.
- [3] J. Rajeesh, R. S. Moni, S. Palanikumar, and T. Gopalakrishnan, "Discrimination of Alzheimer's disease using hippocampus texture features from MRI", *Asian Biomedicine*, vol. 6, no. 1, pp. 87-94, 2012, doi:10.5372/1905-7415.0601.131.
- [4] R. Ugolottia, Y. S. G. Nasheda, P. Mesejoa, Š. Ivekovič, L. Mussia, and S. Cagnoni, "Particle Swarm Optimization and Differential Evolution for model-based object detection", *Applied Soft Computing, preprint*, 2012, doi:http://dx.doi.org/10.1016/j.asoc.2012.11.027.
- [5] L. He, Z. Peng, B. Everding, X. Wang, C. Y. Han, K. L. Weiss, and W. G. Wee, "A comparative study of deformable contour methods on medical image segmentation", *Image and Vision Computing*, vol. 26, no. 2, pp. 141-163, 2008, doi:10.1016/j.imavis.2007.07.010.
- [6] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review.", *Medical Image Analysis*, vol. 13, no. 4, pp. 543-563, 2009, doi:10.1016/j.media.2009.05.004.
- [7] C. T. M. Choi, "Shape optimization of cochlear implant electrode array using genetic algorithms", in *Engineering in Medicine and Biology Society*, Istanbul, Turkey, pp. 1445-1448, 2001, doi:10.1109/IEMBS.2001.1020475.
- [8] T. D. Wager and T. E. Nichols, "Optimization of experimental design in fMRI: a general framework using a genetic algorithm", *NeuroImage*, vol. 18, pp. 293-309, 2003, doi:10.1016/S1053-8119(02)00046-0.
- [9] R. Ben-Shalom, A. Aviv, B. Razon, and A. Korngreen, "Optimizing ion channel models using a parallel genetic algorithm on graphical processors", *Journal of Neuroscience Methods*, vol. 206, pp. 183-194, 2012, doi:10.1016/j.jneumeth.2012.02.024.
- [10] C.-M. Wang and Y.-F. Huang, "Self-adaptive harmony search algorithm for optimization", *Expert Systems with Applications*, vol. 37, no. 4, pp. 2826-2837, 2010, doi:10.1016/j.eswa.2009.09.008.
- [11] K. S. Lee and Z. W. Geem, "A New Meta-heuristic Algorithm for Continuous Engineering Optimization: Harmony Search Theory and Practice", *Computer Methods in Applied Mechanics and Engineering*, vol. 194, pp. 3902-3933, 2005, doi:10.1016/j.cma.2004.09.007.
- [12] A. Kattan and R. Abdullah, "Training of Feed-Forward Neural Networks for Pattern-Classification Applications Using Music Inspired Algorithm", *International Journal of Computer Science and Information Security*, vol. 9, no. 11, pp. 44-57, 2011.
- [13] J.-H. Lee and Y.-S. Yoon, "Modified Harmony Search Algorithm and Neural Networks for Concrete Mix Proportion Design", *Journal of Computing in Civil Engineering*, vol. 23, no. 1, pp. 57-61, 2009, doi:10.1061/(ASCE)0887-3801(2009)23:1(57).
- [14] Z. W. Geem, Music-Inspired Harmony Search Algorithm: Theory and Applications vol. 191: Springer, 2009.

- [15] Q.-K. Pan, P. N. Suganthan, M. F. Tasgetiren, and J. J. Liang, "A selfadaptive global best harmony search algorithm for continuous optimization problems", *Applied Mathematics and Computation*, vol. 216, pp. 830-848, 2010, doi:10.1016/j.amc.2010.01.088.
- [16] Z. W. Geem, J. H. Kim, and G. V. Loganathan, "A New Heuristic Optimization Algorithm: Harmony Search", *Simulation*, vol. 72, no. 2, pp. 60-68, 2001, doi:10.1177/003754970107600201.
- [17] Z. W. Geem and K.-B. Sim, "Parameter-setting-free harmony search algorithm", *Applied Mathematics and Computation*, vol. 217, no. 8, pp. 3881-3889, 2010, doi:10.1016/j.amc.2010.09.049.
- [18] M. Mahdavi, M. Fesanghary, and E. Damangir, "An Improved Harmony Search Algorithm for Solving Optimization Problems", *Applied Mathematics and Computation*, vol. 188, pp. 1567-1579, 2007, doi:10.1016/j.amc.2006.11.033.
- [19] M. G. H. Omran and M. Mahdavi, "Globel-Best Harmony Search", Applied Mathematics and Computation, vol. 198, pp. 643-656, 2008, doi:10.1016/j.amc.2007.09.004.
- [20] C. Worasucheep, "A Harmony Search with Adaptive Pitch Adjustment for Continuous Optimization", *International Journal of Hybrid Information Technology*, vol. 4, no. 4, pp. 13-24, 2011.

- [21] A. Kattan, R. Abdullah, and R. A. Salam, "Harmony Search Based Supervised Training of Artificial Neural Networks", in *International Conference* on *Intelligent Systems, Modeling and Simulation (ISMS2010)*, Liverpool, England, pp. 105-110, 2010, doi:10.1109/ISMS.2010.31.
- [22] M. Ji and J. Klinowski, "Taboo evolutionary programming: a new method of global optimization", *Proceedings of the Royal Society A*, vol. A2006, no. 462, pp. 3613-3627, 2006, doi:10.1098/rspa.2006.1740.
- [23] X. Li, K. Liu, L. Ma, and H. Li, "A Concurrent-Hybrid Evolutionary Algorithms with Multi-child Differential Evolution and Guotao Algorithm Based on Cultural Algorithm Framework," in *LNCS: Advances in Computation and Intelligence*. vol. 6382, Z. Cai, C. Hu, Z. Kang, and Y. Liu, eds. Germany: Springer, pp. 123-133, 2010.
- [24] S. Rahnamayan, H. R. Tizhoosh, and M. M. A. Salama, "Opposition-Based Differential Evolution", *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp.64-79, Feb 2008, doi:10.1109/TEVC.2007.894200

					Proposed	l Method
Multimodal						Best value
Benchmarking	HSASP [20]	SHS [10]	SGHS [15]	ODE [24]	DSHS	@iteration,
Function					20110	Termina-
						tion
GSchwefel 2.26	1.866E+00	1.459E+01	3.57E+01	3.243E+04	8.18E-01	6.79E-01
	(1.348E+00)	(7.253E+00)	(8.60E+01)	(5.508E+02)	(7.19E-02)	586,731
	[λ=0.2]					MAXIMP
Rastrigin	1.485E+00	2.040E+01	1.24E+01	6.009E+02	0	0
	(8.498E-01)	(3.107E+00)	(2.64E+00)	(6.880E+01)	(0)	214,548
	[λ=0.3]					214,548

 Table (2). Optimization results of the proposed DSHS method against other SGO methods

# Monitoring, Tracking and Quantification of Quality of Service in Cloud Computing

Mohamed Firdhous, Suhaidi Hassan, Osman Ghazali

Abstract— Cloud computing has become a buzzword in computing circles now a days. Due to the attraction of cloud computing, it has attracted several service providers to the market in a very time. These cloud service providers have created a very competitive market for the customers to choose from. Due to the heavy competition, the cost of cloud services must be kept at a minimum in order to attract the sufficient number of customers. Also on the other hand, the service providers must assign as many customers as possible to a single physical system, so that the investment on these systems becomes profitable. When many customers are assigned to a single physical system, the Quality of services (QoS) of the cloud offerings would suffer. Hence it becomes necessary to monitor, track and quantify the QoS of the cloud systems in order to provide the right information to both customers and service providers. This information would help both customers and service providers in terms creating a match between them based on expectations and the capacity to meet them. This would increase the efficiency of the cloud systems by loading them to the optimum levels without sacrificing on the expected quality. Continuous monitoring would help to understand the behavior of the system in the short term and long term helping the service providers to take the necessary remedial actions soon. In this paper, the authors describe the research motivation, objectives, research questions, methodology adopted and significance of the PhD project carried out for developing a QoS monitoring, tracking and quantifying system. It also outlines the progress of the work so far along with the achievements.

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

Index Terms— Cloud Computing, Quality of Service, Service Level Agreement, Resource Optimization.

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

# **1** INTRODUCTION

loud computing has the changed the entire computing landscape by making the resources available over the internet as services. Similar to electricity, water, gas and telephony, computing also becomes a utility under cloud computing [1]. Under the utility computing paradigm, computing resources including hardware, development environment and user applications can be accessed remotely over the Internet and paid for only the usage. In the recent times, due to the popularity of cloud systems the market has been flooded with a large number of cloud service providers [2]. These cloud providers host their services on the Internet and make them available to any customer who would like to purchase them. In [3], Garg, Gopalaiyengar and Buyya state that at any given time, large virtualized systems may host and serve thousands of customers. Though cloud computing systems are advantageous to both customers and service providers in terms of economy and utilization of resources, if the resource providisaster [4]. Similar to any other subscription based services, prior to the commencement of the services, the service providers and customers enter into an agreement called Service Level Agreement (SLA) [5]. The SLA would contain the roles and responsibilities of the parties involved, scope of services, quality and performance requirements, charges and rates etc. Thus Quality of Services (QoS) plays an important role in making the cloud services acceptable to customers.

In this paper, the authors present a proposal of a project that has been targeted towards designing a secure reliable monitoring, tracking and quantifying system for cloud computing. The paper discusses in detail all the elements of the proposed project namely research motivation, objectives, research problem and questions, methodology adopted, and significance of the work along with the progress so far. The paper also presents a brief literature review that has been carried out as part of this project.

The rest of this paper is organized as follows. Section 2 presents the research motivation that provided impetus for further comprehensive investigation into this exciting field of cloud computing. Section 3 discusses the problem statement and research questions providing a background to the exact research area along with the problems studied. Section 4 details the proposed methodology adopted in this work. Discussion on expected contributions to be made by this PhD research is given in Section 5, while Section 6 provides a brief summary of the related work carried out in this area while Section 7 presents the preliminary research work carried out so far. Conclusion and future work is presented at last in Section 8.

# **2 RESEARCH MOTIVATION**

Though cloud computing provides many advantages to both customers and service providers in terms of cost savings and

Mohamed Firdhous is a Senior Lecturer attached to the Faculty of Information Technology, University of Moratuwa, Sri Lanka. He is currently on study leave and pursues PhD in computer networks attached to the Inter-NetWorks Research Lab, School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Malaysia, E-mail: mfirdhous@internetworks.my, Mohamed.Firdhous@uom.lk

Associate Prof. Dr. Suhaidi Hassan is an associate professor in the School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Malaysia. He presently heads the InterNetWorks Research Group and the chairman InterNetWorks Research Lab. Prior to this, he was the Assistant Vice Chancellor of the College of Arts and Sciences, Universiti Utara Malaysia, Malaysia, E-mail: suhaidi@uum.edu.my

Associate Prof. Dr. Osman Ghazali is an associate professor attached to the School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Malaysia and a member of the InterNetWorks Research Group. Previously he was the head of the Department of Computer Science, within the School of Computing and Technical Director of the Universiti Teaching Learning Center, Universiti Utara Malaysia, E-mail: osman@uum.edu.my

utilizations, it still needs to earn the confidence of the customers in certain other aspects for it to become commonly deployed successful technology. Due to the dynamic nature of cloud computing resulting from the creation and hosting of virtual systems on the fly, the performance of the system becomes unpredictable [6]. Many of the commercial applications including multi-tiered business applications, scientific data processing, multi-media applications that can benefit from cloud computing are highly sensitive to quality variations [7]. The QoS requirements to be met by the service provider along with the penalties to be imposed, in case of violation are specified in SLAs signed by both parties [5]. An example SLA to be signed between Amazon Web Services (AWS) a leading cloud provider and its customers who wish to use its Amazon Simple Storage Service (S3) can be found at [8]. As per the SLA, the AWS commits to take commercially reasonable efforts to maintain the availability of Amazon S3 at least at 99.9% during any monthly billing cycle. The compensation for failing to meet the above commitment is service credit also described in the SLA. There is no further commitment made on any other OoS expectations.

From the above discussion, it can be seen that the commitments made by leading cloud service providers at present are too simple and does not mention the complex application specific requirements. This kind of SLAs may not be strong enough to attract business customers whose applications are more sensitive to fluctuations in QoS in many dimensions. The situation has been more aggravated by news item reported in the media about the high-profile crash of Amazon EC2 cloud services [9]. This service outage affected many high profile businesses who had hosted their services at AWS. Not only the site was down for many days and but also some organizations lost their data permanently.

Hence the cloud service providers need to come up with innovative methods to provide the service quality demanded by different types of applications and also to assure them that these commitments will be maintained. Also there should be independent verification of the maintenance of the claims of meeting commitments by service providers. Only if the above can be provided, customers will have confidence on the service providers and would readily move their applications to cloud systems in order to reap the benefits of cloud computing.

The main objective of our research is to come up with an innovative model and mechanisms to monitor track and quantify the dynamically changing the QoS performance of cloud services. The proposed model would be able to track the performance in many dimensions using multiple QoS parameters and quantify them in an easily understandable form. During the request, allocation of resources and executing the required tasks, the performance of the system may face unpredictable challenges due to availability of resources, load, and throughput of hardware services. Hence it is a must to continuously monitor and track the real time QoS performance of cloud systems.

Detecting exceptions, malfunctions and degradations of service quality would help service providers to act proactively and correct them before the systems break down. This would help the service providers to maintain their service quality and confidence cultivated in the customers' minds. System degradations can be detected and handled through the development of an efficient, scalable, interoperable, easy-to-use monitoring tool. In this project our object is to conduct an in depth research in order to achieve the research goals given below.

- 1. To develop an analytical (mathematical) model that can be used to predict the QoS of cloud systems under various conditions.
- 2. To develop techniques that can dynamically monitor, track and quantify the QoS of cloud computing systems.
- 3. To develop mechanisms that can distribute the computed QoS score securely among cooperating systems.

## **3 PROBLEM STATEMENT AND RESEARCH QUESTIONS**

The main objective of this research project is to design, develop, implement and test a system that can be used to continually monitor, track and quantify the performance of cloud computing systems. The performance of a cloud computing system is very dynamic due to the very nature of the system itself. Cloud systems create virtual computers and host applications on them on the fly. Similarly they can remove these virtual computers and release the resources back, once the required work has been completed.

Cloud computing services have been divided into three main layers. They are namely, Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [10]. Fig. 1 shows the Cloud services layered model along with the underlying physical computing infrastructure and virtualized computing infrastructure as two distinct layers. The physical hardware is the real workhorse that carries out the processing. The physical hardware is generally provided in the form of computing clusters, grids or individual servers [11]. The virtualized computing infrastructure is created by installing a Virtual Machine Manager (VMM) on the physical hardware [12]. The VMM provides the necessary isolation and security between the multiple virtual machines running in parallel on a single physical computer.

Fig. 1: Cloud Computing Layered Model



IaaS is the provision of virtual hardware as a service over the Internet. These virtual machines can be brought up and removed on the fly based on customer demand. Once a virtual machine has been purchased, it can be treated as if it is real hardware and any operating system and applications can be installed on it. PaaS is the complete software development environment along with operating system, development and testing tools and application programming interface installed on virtual hardware. PaaS helps web based application developers to reduce the cost and time of bringing their applications to market from the design boards. SaaS is the new paradigm of software marketing and ownership. SaaS enables customers to access web based applications hosted in remote data centers and pay only for the usage. These applications have the capability of managing their own data and configuration information to suit individual user requirements.

From the above description of cloud computing systems, it can be seen that cloud systems can host applications and services that have drastically different requirements in terms of QoS [13]. While the transactional applications demand better response times and throughput guarantees, the noninteractive batch jobs more are concerned with job completion times and accuracy of processing [14]. Thus, it can be concluded that the QoS demands of cloud services are more complex and depends on multiple factors or parameters.

This research would be designed in such a manner to find answers to the following research questions.

- RQ1: How to design an analytical (mathematical) model that can be used to predict the QoS behavior of the cloud system under different conditions?
- RQ2: How to develop techniques that can dynamically monitor, track and quantify the QoS of cloud computing systems?
- RQ3: How to devise mechanisms that can distribute the computed QoS scores among cooperating systems?

#### **RQ1: How to Design an Analytical Model?**

In a shared distributed dynamic system like cloud computing, it is important that the systems must be loaded appropriately to achieve the competing requirements higher revenue for service providers and better services for customers. In order to achieve the higher revenue, the systems must be loaded optimally without sacrificing the quality of service performance. If the service providers have a prior knowledge of optimum load levels and the cost of the resources, they can price the services optimally that ensures both profits and maintains quality of services. The challenge in developing analytical models for predicting the behavior of complex dynamic systems is fitting the right statistical models to the different components of the systems such as request arrival pattern, service time distributions, I/O system behaviors, failure handling, resource usage etc. The challenge is further complicated due to interdependence of components with each other. Hence, there is an urgent need for a cloud service performance modeling and workload prediction technique that can ensure optimum system utilization without sacrificing the QoS requirements. RQ1 that has been formulated in order to solve handle this practical issue is: Is it possible to develop a mathematical model that can effectively predict the behavior of cloud system under various conditions specified by different QoS metric values?

#### **RQ2: How to Develop Monitoring Techniques?**

A cloud computing system is a very dynamic one compared to other distributed systems such as cluster computing, grid computing etc. The dynamic creation and hosting of virtual systems aggravates the situation due to non commitment of any resources until such a virtual system becomes active. Also cloud system can host services at different layers of abstraction and applications that have distinct QoS demands based on different metrics. Hence it is necessary to continuously monitor the performance and track the changes in them. Also, it is important to protect the system from malicious attacks and momentary fluctuations. The research question, RQ2 formulated for meeting this challenge is: Is it feasible to design a secure reliable technique that can monitor, track and quantify the performance of cloud systems?

### **RQ3: How to Devise Score Distribution Mechanisms?**

It is a common practice for service providers to host their systems to geographically distributed data centers. Distributing data centers geographically across multiple sites help the service providers to handle short term and long term disasters gracefully. Though the systems are distributed across multiple sites, it must be transparent to the end users and behave like a single system. When monitoring such a distributed system, it is not practically advisable to have a single monitoring system for all the sites. When multiple systems are employed, they should cooperate with each other sharing their information. When scores are shared among geographically distributed systems, they should consider other factors such as network dynamics when updating the scores and security of transmitted/received information. RQ3, the research question that was developed for researching into this challenging area is: Is it practical to devise a secure reliable mechanism that can cooperatively exchange QoS scores among themselves and incorporate the other information such as network dynamics into those scores?

## 4 METHODOLOGY

The main objective of this research is to develop a reliable and secure monitoring system for cloud computing performance. In order to achieve this objective, it is necessary to follow a strict scientific procedure as the results obtained as valid and repeatable under similar conditions. The methodology adopted for carrying out this research consists of five main phases. They are namely, analysis, design, testing, verification & validation and implementation. Fig. 2 shows these phases and how there are connected to each other in a graphical format.



Fig 2: Research Methodology

The project will be divided into three main tasks in finding solutions to the three research questions identified in Section 3. The tasks are briefly described below.

## Task 1: Designing an Analytical Model for Cloud Computing

Modeling of cloud computing mathematically enables researchers and other professionals to carry out in depth analyses of the system under different conditions. In solving RQ1, it expected to design and develop an analytical (mathematical) model that would correctly predict the behavior of the cloud computing system under various conditions. The model is expected to be based on computational statistical techniques as the behavior of components in a cloud system can be treated as stochastic processes for all practical purposes [15]. These models will capture the behavior of different components in terms of QoS attributes such as time, throughput, utilization, cost etc.

## Task 2: Developing Techniques for Monitoring Cloud Performance

Task 2 would mainly concentrate on how to monitor the behavior of a cloud system in terms of meeting the QoS requirements of customers and converting it to a comprehensible score. The proposed techniques must be capable of monitoring the QoS performance of the cloud system based on more than one parameter and be able to assign relative preference to these metrics depending on the user requirements. The proposed system should continuously monitor the performance and update the final score as these systems are dynamic. Also the proposed techniques must be able to identify and isolate the deviations in performance based on whether it was due to momentary fluctuations in performance or due to permanent degradations. For the proposed techniques to be successfully employed in practical systems, it should have enough resilience to attack that aims to modify the scores maliciously. Hence the ultimate mechanism that would be developed in solving RQ2 would consist of multiple components where each one would work independently towards achieving the final goal of monitoring, tracking and quantifying the service quality of cloud system effectively and securely.

#### Task 3: Devising Mechanisms for Distributing Performance Scores

Task 3 is mainly concerned with the collaboration between independent cloud monitoring systems developed in Task 2. In a cloud system that has been deployed across a wide geographical area and also services globally dispersed clients must be identify and allocate the right resources hosted at the right locations. Also it is practically not feasible to monitor and track all the systems with a single monitoring system. Hence it necessary to deploy multiple systems at various locations, so that they can monitor and track the performance of the systems independently but collaboratively exchange the information collected with each other. When the information is exchanged, they should also consider the other intervening factors that can affect the service quality such as the quality and performance of the network connecting these systems together. The other major factor that needs to be taken into account is the security of the information transmitted. The transmitted information can be attacked enroute by malicious attackers or the systems receiving the information may be fed with wrong information. Hence developing mechanism for exchange of information between collaborating monitoring units in answering RQ3, it is necessary arrive at an optimum solution that considers all the factors in order to arrive at a resilient, secure and scalable mechanism.

# **5** SIGNIFICANCE

The significance of the proposed project is manifold. The project is expected to have three main individual but related contributions. These contributions would help greatly for enhancing acceptability of cloud computing to a wider audience. Contribution 1, the analytical model would help the network designers to prepare their resources in such a manner that is most suitable to meet the requirements of the customers. This model can also be used by researchers to analyze the behavior of cloud systems under different conditions. Contributions 2 and 3 together would help the customers and service providers equally. From the customers' point of view, the monitoring, tracking and quantifying unit would help them to identify the service providers who would likely to meet their QoS requirements. From the service providers' angle, this system would help them to track the performance of their resources. The service providers would be able to identify any degradation of performance well before it becomes a disaster and irreversible.

# **6 RELATED WORK**

This section briefly discusses the related work that has been carried out by other researchers and reported as published work in conferences and journals. When selecting the literature for reviewing, special attention was paid for selecting the papers that were relevant and most recent. Instead of just listing the work, a critical analysis on these proposed mechanisms was carried out with special reference to the principles, strengths and weaknesses.

Cloud computing systems may host thousands of globally dispersed clients at any given time. These clients may access different types of services that have varying requirements depending on the type of clients, services and resources involved. In order to meet the requirements of clients and services, it is necessary to provide a certain level of QoS by the service providers. Nevertheless, providing a guaranteed QoS in such a challenging environment in a widely distributed diverse networks supporting complex hosting of services is not an easy task [16,17]. Though it is a challenging task, several researchers have undertaken to develop mechanisms, frameworks and systems which could guarantee the QoS requirements of different services. This section takes an in depth look at these mechanisms, frameworks and systems.

In [18], Liu et al have proposed a generic QoS framework for cloud workflow system. The proposed framework covers all the four stages of cloud workflow namely, QoS requirement specification, QoS-aware service selection, QoS consistency monitoring and QoS violation handling. The shortcoming of this framework is that it does not specifically identify any QoS parameters and also does not discuss how to differentiate clients requiring different QoS levels.

Chen and Zhang have proposed a workflow scheduling algorithm based on Particle Swarm Optimization (PSO) in [19]. The proposed mechanism can optimize up to seven parameters specified the users compared to traditional optimization techniques that consider only the workflow execution time. The downside of the proposed mechanism is that it lacks a monitoring scheme for catching QoS violations or punishing the violators. Buyya, Garg, and Calheiros have proposed a framework for SLA management with special reference to managing QoS requirements in [17]. The proposed architecture successfully integrates the market based resource provisioning with virtualization technologies for flexible resource allocations to user applications. But the proposed architecture does not support different cloud service offerings such as IaaS, PaaS and SaaS together in a combined manner.

In [5], Feng et al have proposed an optimal resource allocation model for revenue maximization. The model has been mathematically derived and tested using both synthetic and traced datasets. The proposed model performs better than heuristic optimization of resources in maximizing profits. But the application of this method is limited as it considers only the mean response time as the QoS attribute to be satisfied. For customers who require guaranteed performance or at least a commitment in terms of a confidence level cannot be served through this model. Hence from the customers' point of view, the model has limited application and may serve only casual users.

In [20], den Bossche, Vanmechelen and Broeckhove have proposed a set of heuristics for scheduling deadlineconstrained applications in a hybrid cloud system in a cost effective manner. The proposed system attempts to maximize the use of local resources along with minimizing the use of external resources without compromising the QoS requirements of the applications. The optimization heuristics takes the cost of both computation and data transfer along with the estimated data transfer times. The main criteria in optimization is the maximization of cost saving. The effect of different cost factors and workload characteristics on the cost savings have been analyzed along with the sensitivity of the results to the different runtime estimates. The advantages of the proposed methodology is that it can select an optimized set of resources from both in-house (private) and public cloud systems for meeting the QoS requirements. But at the same time it suffers from certain weaknesses. Though it is concerned only about the deadline concerned applications, it does not consider the failures that may occur after the scheduling has been done. The failure will increase the cost of execution and affect the application in terms of quality.

In [21], Emeakaroha et al have presented a scheduling heuristic that takes multiple SLA parameters when deploying applications in the Cloud. The attributes considered are physical requirements such as CPU time, network bandwidth and storage capacity for deploying applications. These parameters have limited application in real world systems as they need to be considered only during deployment. Once the applications have been ready for client access, the customers would be more interested in performance parameters such as response time, processing time etc. Hence this heuristic may not have much practical significance in real world business environments.

Li et al in [22] have proposed a novel customizable cloud workflow scheduling model. The authors have incorporated trust into the model in addition to the QoS targets. In order to analyze the users' requirements and design a customized schedule, the authors propose a two stage workflow model where the macro multi-workflow stage is based on trust and micro single workflow stage classifies workflows into timesensitive and cost-sensitive based on QoS demands. The classification of workflows has been carried out using fuzzy clustering technique. The proposed model restricts the QoS parameters considered to response time, bandwidth, storage, reliability and cost. Also the delivery of QoS is confined only to average values and no guarantee of service delivery is provided at least in terms of a predetermined confidence level. This is a strong limitation of the proposed technique as the users do not have the freedom to select their own QoS parameters and no guarantee of the QoS delivery at the least a statistical validation.

Alhamazani et al., in [23] have outlined the importance of dynamically monitoring the QoS of virtualized services. they further claim that the monitoring of the services would help both the cloud provider and application developer to maximize the return of their investments in terms of keeping the cloud services and hosted applications operating at peak efficiency, detecting changes in service and application performance, SLA violations, failures of cloud services and other dynamic configuration changes. The paper mainly concentrates on describing the PhD work being carried out in terms of research questions, objectives and methodology. The researchers mainly concentrate on SNMP based QoS monitoring. Since this is a concept paper describing work in progress, no concrete proposal is put forward or evaluated.

The literatures discussed above are mainly concerned with cloud workflow. The cloud workflows attempt to select the resources in such a manner that the required QoS would be satisfied. None of the literature cited above discuss continuous monitoring of cloud systems for their performance or quantifying them. Also, the reported mechanisms are unable to identify or detect system degradations as they are mainly concerned with resource selection and allocation.

# 7 PROGRESS TILL TO-DATE

The project has progressed significantly and the tasks identified in Section 4 have been carried out in parallel. The development of the cloud computing model has been almost complete and requires slight modifications in terms of performance tuning and testing. The development of monitoring and tracking system proposed to be carried out under Task 2 has progressed significantly. This work has been carried out iteratively by improving a basic model designed at the beginning until the final goal of robust mechanism has been achieved. The basic design and improvements have been presented at various international conferences and forums. These designs received very positive and encouraging comments from reviewers who reviewed these works.

# 8 CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

Cloud computing systems have become very popular in the recent times and attracted the attention of many people including researchers, service providers and customers. The cloud systems provide many advantages over the traditional computing system due to the innovative way of making the computing resources available over the internet and charging the customers. Cloud systems employ a pay-as-you-go business model similar to utilities like electricity, water, gas and telephony. Quality of Service would play an important role in making cloud computing acceptable to everyone especially the business customers.

Monitoring the cloud system is a key factor in ensuring the committed service quality is maintained. The monitoring the system will help both the customers and service providers as the customers can select the right service provider who could meet their requirements and service providers would be able design manage their systems optimally meeting the requirements of the customers.

The conclusion of the proposed study, it is expected to contribute significantly to the existing knowledge on cloud computing with special reference to enhancing the service quality. The work is also significant practically as the systems once completed can be used by both customers and service providers to obtain a better service and enhance their services and profitability respectively.

#### REFERENCES

- R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5<sup>th</sup> Utility," *Journal of Future Generation Computer Systems*, vol. 25, no. 6, pp. 599–616, June 2009.
- [2] B.P. Rimal, E. Choi, and I. Lumb, I. "A Taxonomy and Survey of Cloud Computing Systems," Proc. 5<sup>th</sup> Int. Joint Conf. INC, IMS and IDC (NCM'09), pp. 44-51, 2009.
- [3] S.K. Garg, S.K. Gopalaiyengar, and R. Buyya, "SLA-based Resource Provisioning for Heterogeneous Workloads in a Virtualized Cloud Datacenter," Proc. 11th Int. Conf. Algorithms and Architectures for Parallel Processing (ICA3PP2011), pp. 371–384, 2011.
- [4] D. Bourcier, and P.D. Filippi, "Cloud Computing: New Research Perspectives for Computers and Law," Proc. 13th Int. Conf. Artificial Intelligence & Law, pp. 73–92, 2011.
- [5] G. Feng, S. Garg, R. Buyya, and W. Li, "Revenue Maximization using Adaptive Resource Provisioning in Cloud Computing Environments," *Proc.* 13<sup>th</sup> ACM/IEEE Int. Conf. Grid Computing, pp. 192–200, 2012.
- [6] D.C. Marinescu, *Cloud Computing: Theory and Practice*. Burlington, MA: Morgan Kaufmann, 2013.
- [7] A. Iosup, N. Yigitbasi, and D. Epema, "On the Performance Variability of Production Cloud Services," *Proc. 11th IEEE/ACM Int. Symp. Cluster, Cloud and Grid Computing*, pp. 104-113, 2011.
- [8] Amazon S3 Service Level Agreement, http://aws.amazon.com/s3sla/. Accessed on February 07, 2013.
- [9] H. Blodget, "Amazon's Cloud Crash Disaster Permanently Destroyed Many Customers' Data." http://www.businessinsider.com/amazonlost-data-2011-4. Accessed on February 07, 2013.
- [10] C. Vecchiola, S. Pandey, and R. Buyya, "High-performance cloud computing: A view of scientific applications," *Proc.* 10<sup>th</sup> Int. Symp. *Pervasive Systems, Algorithms, and Networks (ISPAN)*, pp. 4–16, 2009.
- [11] R. Buyya, A. Beloglazov, and J.H. Abawajy, "Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges", Proc. Int. Conf. Parallel and Distributed Processing Techniques and Applications, pp. 6-20, 2010.
- [12] Y. Li, W. Li, and C. Jiang, "A Survey of Virtual Machine System: Current Technology and Future Trends," *Proc. Third Int. Symp. Electronic Commerce and Security*, pp. 332-336, 2010.
- [13] A. Quiroz, H. Kim, M. Parashar, N. Gnanasambandam, and N. Sharma, "Towards Autonomic Workload Provisioning for Enterprise

Grids and Clouds," Proc. 10th IEEE/ACM Int. Conf. Grid Comp., pp. 50-57, 2009.

- [14] D. Carrera, M. Steinder, I. Whalley, J. Torres, and E. Ayguade, "Enabling resource sharing between transactional and batch workloads using dynamic application placement," *Proc. 9th ACM/IFIP/USENIX Int. Conf. Middleware*, pp. 203-222, 2008.
- [15] T.M. Chen, "Network Traffic Modeling," The Handbook of Computer Networks: Distributed Networks, Network Planning, Control, Management, and New Trends and Applications, Vol. 3, H. Bidgoli, ed., Hoboken, NJ: Wiley, pp. 326-339, 2007.
- [16] J.M. Pedersen, M.T. Riaz, J.C. Júnior, B. Dubalski, D. Ledzinski, and A. Patel, "Assessing Measurements of QoS for global Cloud Computing Services," *Proc. Ninth IEEE Int. Conf. Dependable, Autonomic and Secure Computing*, pp. 682-689, 2011.
- [17] R. Buyya, S.K. Garg, and R.N. Calheiros. "SLA-Oriented Resource Provisioning for Cloud Computing: Challenges, Architecture, and Solutions," Proc. Int. Conf. Cloud and Service Computing, pp. 1-10, 2011.
- [18] X. Liu, Y. Yang, D. Yuan, G. Zhang, W. Li, and D. Cao, "A Generic QoS Framework for Cloud Workflow Systems," *Proc. Ninth IEEE Int. Conf. Dependable, Autonomic and Secure Computing*, pp. 713-720, 2011.
- [19] W.N. Chen, and J. Zhang, "A Set-Based Discrete PSO for Cloud Workflow Scheduling with User-Defined QoS Constraints," Proc. IEEE Int. Conf. Systems, Man and Cybernetics, pp. 773-778, 2012.
- [20] R.V. den Bossche, K. Vanmechelen, and J. Broeckhove, "Cost-Efficient Scheduling Heuristics for Deadline Constrained Workloads on Hybrid Clouds," *Proc.* 3<sup>rd</sup> IEEE Int. Conf. Cloud Comp. Tech. and Sc., (CloudCom), pp. 320-327, 2011.
- [21] V.C. Emeakaroha, I. Brandic, M. Maurer, and I. Breskovic, "SLA-Aware Application Deployment and Resource Allocation in Clouds," *Proc. 35th IEEE Annual Computer Software and Applications Conference Workshops (COMPSACW)*, pp. 298-303, 2011.
- [22] W. Li, Q. Zhang, J. Wu, J. Li, and H. Zhao, "Trust-based and QoS Demand Clustering Analysis Customizable Cloud Workflow Scheduling Strategies," *Proc. IEEE Int. Conf. Cluster Comp. Workshops*, pp. 111–119, 2012.
- [23] K. Alhamazani, R. Ranjan, F. Rabhi, L. Wang, and K. Mitra, "Cloud Monitoring for Optimizing the QoS of Hosted Applications," 4th IEEE Int. Conf. Cloud Comp. Tech. and Sc., pp. 765–770, 2012.

# A Comprehensive Survey on Quality of Service Implementations in Cloud Computing

Mohamed Firdhous, Suhaidi Hassan, Osman Ghazali

Abstract— Quality of service plays is an important factor in distributed computing. Cloud computing has been the paradigm in distributed computing. Under cloud computing, computing resources are hosted in the internet and delivered to customers as services. Prior to the commencement of services, the customers and cloud providers negotiate and enter into an agreement named service level agreement. The services level agreements clarify the roles, set charges and expectations and provide mechanisms for resolving service problems within a specified and agreed upon time period. Service level agreements also cover performance, reliability conditions in terms of quality of service guarantees. In this paper, the authors present a comprehensive survey on quality of service implementations in cloud computing with respect to their implementation details, strengths and weaknesses.

Index Terms— Cloud Computing, Utility Computing, QoS, SLA.

# 1\_INTRODUCTION

loud computing has the changed the entire computing landscape by making the resources available over the internet as services. Similar to electricity, water, gas and telephony, computing also becomes a utility under cloud computing [1]. Under the utility computing paradigm, computing resources including hardware, development environment and user applications can be accessed remotely over the Internet and paid for only the usage. In the recent times, due to the popularity of cloud systems the market has been flooded with a large number of cloud service providers [2]. These cloud providers host their services on the Internet and make them available to any customer who would like to purchase them. In [3], Garg, Gopalaiyengar and Buyya state that at any given time, large virtualized systems may host and serve thousands of customers. Though cloud computing systems are advantageous to both customers and service providers in terms of economy and utilization of resources, if the resource provisioning is not carried out optimally it would also become a disaster [4]. Similar to any other subscription based services, prior to the commencement of the services, the service providers and customers enter into an agreement called Service Level Agreement (SLA) [5]. The SLA would contain the roles and

ity and performance requirements, charges and rates etc. Thus Quality of Services (QoS) plays an important role in making the cloud services acceptable to customers.

Literature has reported many implementations for measuring and ensuring QoS in cloud computing systems. In this paper, the authors carry out a comprehensive survey on the mechanisms and methods proposed by various researchers with respect to their implementation principles, strengths and weaknesses.

## 2 CLOUD COMPUTING

\_\_\_\_\_

Cloud computing has been identified as the 5<sup>th</sup> utility after electricity, gas waster and telephony due to the way it transforms, how computing resources can be accessed, used and paid for [1]. Traditional distributed computing where computing resources were leased from business data centers required users to purchase fixed capacities and pay for them irrespective of actual usage. On the other hand, cloud computing only charges for the usage and the resources committed to the users is elastic and closely follow the demand patterns [6]. Hence cloud computing makes the investment on computing an asset that can immediately return the investment made on them.

Cloud computing has been divided into three main layers. They are namely, Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [7]. Fig. 1 shows the Cloud computing layers along with the underlying physical computing infrastructure and virtualized computing infrastructure as two distinct layers. The physical hardware is the real workhorse that carries out the processing. The physical hardware is generally provided in the form of computing clusters, grids or individual servers [8]. The virtualized computing infrastructure is created by installing a Virtual Machine Manager (VMM) on the physical hardware [9]. The VMM provides the necessary isolation and security between the multiple virtual machines running in parallel on a single physical computer.

Mohamed Firdhous is a Senior Lecturer attached to the Faculty of Information Technology, University of Moratuwa, Sri Lanka. He is currently on study leave and pursues PhD in computer networks attached to the Inter-NetWorks Research Lab, School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Malaysia, E-mail: mfirdhous@internetworks.my, Mohamed.Firdhous@uom.lk

Prof. Dr. Suhaidi Hassan is a professor in the School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Malaysia. He presently heads the InterNetWorks Research Group and the chairman InterNetWorks Research Lab. Prior to this, he was the Assistant Vice Chancellor of the College of Arts and Sciences, Universiti Utara Malaysia, Malaysia, E-mail: suhaidi@uum.edu.my

Associate Prof. Dr. Ösman Ghazali is an associate professor attached to the School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Malaysia and a member of the InterNetWorks Research Group. Previously he was the head of the Department of Computer Science, within the School of Computing and Technical Director of the Universiti Teaching Learning Center, Universiti Utara Malaysia, E-mail: osman@uum.edu.my



Fig. 1: Cloud Computing Layers

IaaS is the provision of virtual hardware as a service over the Internet. These virtual machines can be brought up and removed on the fly based on customer demand. Once a virtual machine has been purchased, it can be treated as if it is real hardware and any operating system and applications can be installed on it. PaaS is the complete software development environment along with operating system, development and testing tools and application programming interface installed on virtual hardware. PaaS helps web based application developers to reduce the cost and time of bringing their applications to market from the design boards. SaaS is the new paradigm of software marketing and ownership. SaaS enables customers to access web based applications hosted in remote data centers and pay only for the usage. These applications have the capability of managing their own data and configuration information to suit individual user requirements.

# **3 QUALITY OF SERVICE IN CLOUD COMPUTING**

Cloud computing systems may host thousands of globally dispersed clients at any given time. These clients may access different types of services that have varying requirements depending on the type of clients, services and resources involved. In order to meet the requirements of clients and services, it is necessary to provide a certain level of QoS by the service providers. Nevertheless, providing a guaranteed QoS in such a challenging environment in a widely distributed diverse networks supporting complex hosting of services is not an easy task [10, 11]. Though it is a challenging task, several researchers have undertaken to develop mechanisms, frameworks and systems which could guarantee the QoS requirements of different services. This section takes an in depth look at these mechanisms, frameworks and systems.

In [5], Feng et al have proposed an optimal resource allocation model for revenue maximization. The model has been mathematically derived and tested using both synthetic and traced datasets. The proposed model performs better than heuristic optimization of resources in maximizing profits. But the application of this method is limited as it considers only the mean response time as the QoS attribute to be satisfied. For customers who require guaranteed performance or at least a commitment in terms of a confidence level cannot be served through this model. Hence from the customers' point of view, the model has limited application and may serve only casual users.

Buyya, Garg, and Calheiros have proposed a framework for SLA management with special reference to managing QoS requirements in [11]. The proposed architecture successfully integrates the market based resource provisioning with virtualization technologies for flexible resource allocations to user applications. But the proposed architecture does not support different cloud service offerings such as IaaS, PaaS and SaaS together in a combined manner.

In [12], Liu et al have proposed a generic QoS framework for cloud workflow system. The proposed framework covers all the four stages of cloud workflow namely, QoS requirement specification, QoS-aware service selection, QoS consistency monitoring and QoS violation handling. The shortcoming of this framework is that it does not specifically identify any QoS parameters and also does not discuss how to differentiate clients requiring different QoS levels.

Chen and Zhang have proposed a workflow scheduling algorithm based on Particle Swarm Optimization (PSO) in [13]. The proposed mechanism can optimize up to seven parameters specified the users compared to traditional optimization techniques that consider only the workflow execution time. The downside of the proposed mechanism is that it lacks a monitoring scheme for catching QoS violations or punish the violators.

In [14], den Bossche, Vanmechelen and Broeckhove have proposed a set of heuristics for scheduling deadlineconstrained applications in a hybrid cloud system in a cost effective manner. The proposed system attempts to maximize the use of local resources along with minimizing the use of external resources without compromising the QoS requirements of the applications. The optimization heuristics takes the cost of both computation and data transfer along with the estimated data transfer times. The main criteria in optimization is the maximization of cost saving. The effect of different cost factors and workload characteristics on the cost savings have been analyzed along with the sensitivity of the results to the different runtime estimates. The advantages of the proposed methodology is that it can select an optimized set of resources from both in-house (private) and public cloud systems for meeting the QoS requirements. But at the same time it suffers from certain weaknesses. Though it is concerned only about the deadline concerned applications, it does not consider the failures that may occur after the scheduling has been done. The failure will increase the cost of execution and affect the application in terms of quality.

In [15], Emeakaroha et al., have presented a scheduling heuristic that takes multiple SLA parameters for application deployments in the Cloud. The attributes considered include CPU time, network bandwidth and storage capacity for deploying applications. These parameters have limited application in real world systems as they need to be considered only during deployment. Once the applications have been ready for client access, the customers would be more interested in performance parameters such as response time, processing time etc. Hence this heuristic may not have much practical significance in real world business environments.

Li et al in [16] have proposed a novel customizable cloud workflow scheduling model. The authors have incorporated trust into the model in addition to the QoS targets. In order to analyze the users' requirements and design a customized schedule, the authors propose a two stage workflow model where the macro multi-workflow stage is based on trust and micro single workflow stage classifies workflows into timesensitive and cost-sensitive based on QoS demands. The classification of workflows has been carried out using fuzzy clustering technique. The proposed model restricts the QoS parameters considered to response time, bandwidth, storage, reliability and cost. Also the delivery of QoS is confined only to average values and no guarantee of service delivery is provided at least in terms of a predetermined confidence level. This is a strong limitation of the proposed technique as the users do not have the freedom to select their own QoS parameters and no guarantee of the QoS delivery at the least a statistical validation.

Mushtaq, Augustin, and Mellouk have investigated the effect of different factors on the Quality of Experience (QoE) of multimedia users in a cloud computing network [17]. The authors of this paper have grouped the factors that affect the QoE into four groups. They are namely network parameters, characteristics of videos, terminal characteristics and types of users' profiles. The data collected through different methods have been classified using machine learning techniques such as Naive Bayes, Support Vector Machines, K-Nearest Neighbors, Decision Tree, Random Forest and Neural Networks. Out of these methods they have determined the best method for QoS/QoE correlation after evaluating them. Hence it can be concluded that this paper discusses more about the capabilities of machine learning techniques than about QoS or QoE. The QoS/QoE correlation is a case for evaluating the machine learning techniques.

Alhamazani et al., in [18] have outlined the importance of dynamically monitoring the QoS of virtualized services. they further claim that the monitoring of the services would help both the cloud provider and application developer to maximize the return of their investments in terms of keeping the cloud services and hosted applications operating at peak efficiency, detecting changes in service and application performance, SLA violations, failures of cloud services and other dynamic configuration changes. The paper mainly concentrates on describing the PhD work being carried out in terms of research questions, objectives and methodology. The researchers mainly concentrate on SNMP based QoS monitoring. Since this is a concept paper describing work in progress, no concrete proposal is put forward or evaluated.

In [19], Li et al., have adopted a profit model based on response time to represent the QoS requirements and shown that this model yields different analytical results to that obtained from traditional metrics. It was also shown that both under allocation and over allocation of resources affect profits. The right allocation of resources depends on many factors such as available resources, workload distribution, system configuration, and profit model. This is an innovative method of analyzing the effect of managing QoS on resource utilization. The results only discuss the effect of managing QoS, how to provide an optimal allocation is not discussed.

In [20], Adami et al., have proposed a distributed resource allocation algorithm for cloud and grid systems. The algorithm is capable of handling multiple resource requirements and the criteria for optimization is a tradeoff between the execution time and economic cost with system and network performance parameters as additional factors. The proposed algorithm successfully incorporates many system and network performance parameters but fails to consider the failures that may arise after allocating resources. The failures arising after the allocation increase the cost of computation as they would require more time for execution. Hence the cost based optimization used in the proposed algorithm may not be accurate due this shortcoming.

Gohad, Ponnalagu, and Narendra have proposed an extensible dynamic provisioning framework for multi tenant cloud system [21]. The proposed framework starts with defining a tenancy requirements model for helping map provisioned resources. The other index called the health grading model handles the QoS characteristics of tenants. Together both these indices permit dynamic re-provisioning for existing tenants based on changing tenancy requirements or health grading predictions. The proposed framework is innovative in the dynamic resource provisioning sense, but may not be suitable for applications that have bursty requirements. Also the proposed framework is based on starts small and grows large criterion. But when new tenants arrive, the allocated resources are not deallocated from the existing tenants, this would starve the new tenants of resources.

In [22], Ma, Sun and Abraham and proposed a lightweight framework for monitoring public clouds. The proposed framework is very less resource intensive but does not monitor the QoS parameters such as response time, processing time etc., which the customers may be more interested in.

Zhu and Agrawal have presented a framework for handling adaptive applications in cloud systems in [23]. The proposed framework is based on multi-input-multi-output feedback control model for resource provisioning. But the model is limited to memory and CPU performance only, hence the application may be affected by the underperformance of other resources such as network, disk drives etc.

In [24], Sharma et al., have proposed a cloud resource pricing model balancing QoS requirements and higher profits. This model uses the realistic valuation for underlying resources using the age of resources. The proposed model does not include utilization in computing the cost. Hence it may lead to inaccurate projections.

Stoicuta et al., have developed a client application for monitoring cloud QoS on iOS5 [25]. This application can be used by clients to monitor the performance of their cloud provider. But the application has been designed very narrowly focusing only on available transfer rate and one-way delay. Hence the application has limited applications.

In [26], Goyal et al., have proposed a QoS based trust management model. The proposed model claims to use multiple QoS attribute to compute the trust value, but there is no clear explanation how these parameters are combined. Also there is no prioritization between parameters is possible.

Iyer and Veeravalli have formulated a resource allocation strategy for cloud infrastructure based on bargaining [27]. They have combined the Nash Bargaining Solution and Raiffa Bargaining Solution to arrive at an optimal allocation strategy. The proposed strategy handles the dynamic nature of cloud very well during run time but the system does not permit to manage resources from multiple sources. Hence if a single service provider cannot meet all the requirements of the customer, he will be required to settle for a sub optimal allocation of resources.

In [28], Sanchez et al., have investigated the capability of Markov Arrival Processes based queuing models to predict future workload of cloud systems. The model has tested only with numerical experiments, hence the true capability of the models need to be evaluated with real data traces.

An optimization framework for cross layer cloud services has been proposed in [29] by Kouki, Ledoux, and Sharrock. The optimization across multiple layers has been carried out enforcing the SLA dependencies between them. The framework is very suitable for vendors marketing multitude of services and also takes the dynamic nature of cloud systems. The propose system currently lacks the run time management of QoS performance.

Wu, Garg and Buyya in [30] have proposed some algorithms for resource allocation for SaaS providers to balance the cost of hardware and SLA violations. This proposed algorithm takes certain QoS parameters such as response time and service initiation time for satisfying the customers while minimizing the use of hardware resources. Theses algorithms propose to reuse the already created VMs in order to minimize cost, but it may create security problems for customers as the residual information in the VMs can be used against them.

In [31], Phillips, Engen and Papay have reported the results obtained on the performance of virtualized hardware of two IaaS providers. They have used the Dwarf benchmarks for measuring the performance of these cloud providers and show that the actual performance and show that the labeling such as small, medium or large does not actually reflect the true nature of a system and also they show that certain applications may run better on certain hardware than the other ones.

Chauhan et al., have proposed a process for identifying a cloud service provider for a given set of requirements by matching SLA parameters in [32]. The proposed process finds a match by crating two models called the capability model and requirements model and then translating these models to graphs for matching them for compatibility. Based on the compatibility, each node pair is given a mark between 0 and 1 and the final score is computed by summing them all. This is a good effort for automating the process of finding a match between a customer's requirements and service provider's capability. But, it does not take the dynamic nature of the cloud service into account. It only matches the published capability of cloud providers with customers' requirements. This is a major shortcoming of this process.

Table 1 summaries all the work discussed so far with reference to their strengths and weaknesses along with the proposed model or framework. From Table 1, it can be seen that there is still a lot of scope for future work in this exciting and challenging area.

Work	Proposed Model/Framework	Strengths	Weaknesses
[5]	Optimal resource allocation	Mathematically derived and per-	Only mean performance time is considered,
	model for revenue maximiza-	forms better than heuristics	hence not suitable for QoS sensitive applica-
	tion		tions requiring guaranteed performance.
[11]	A framework for SLA man-	Successfully integrates the market	Does not integrate IaaS, PaaS and SaaS in a
	agement with special refer-	based resource provisioning with	combined manner.
	ence to managing QoS re-	virtualization technologies for flexi-	
	quirements	ble resource allocations.	
[12]	A generic QoS framework for	Covers all the four stages of cloud	QoS metrics are not identified and no mecha-
	cloud workflow	workflow.	nism for differentiating customers based on
			requirements.
[13]	A set-based PSO approach	Multiple parameter optimizations	But no monitoring mechanism is implement-
	scheduling problem in cloud	are possible.	ed for catching violations.
	computing	-	_
[14]	A set of heuristics for sched-	The optimization heuristics takes the	It does not consider the failures that may oc-
	uling deadline-constrained	cost of both computation and data	cur after the scheduling has been done. The
	applications in a hybrid cloud	transfer along with the estimated	failure will increase the cost of execution and
	system.	data transfer times and different cost	affect the application in terms of quality
		factors and workload characteristics.	
[15]	A scheduling heuristic that	Considers deployment attributes	Does not consider performance parameters
	takes multiple SLA parame-	such as CPU time, network band-	such as response time, performance time etc.,
	ters when deploying applica-	width, storage capacity etc., before	
	tions in the Cloud	installation of applications in the	
		cloud system.	
[16]	A flexible multistage work-	The proposed model is flexible due	Application is strongly limited due to strict
	flow scheduling model.	to breaking up of the workflow	restriction on the type of QoS attributed taken
		scheduling mechanism into multiple	into account and the absence of QoS delivery
		stages and grouping the requests	guarantees.
		based on the user requirements.	
[17]	The correlation between	QoS/QoE correlation has been stud-	This paper discusses more about the capabili-
	QoS/QoE has been studied	ied using a selected set of machine	ties of machine learning techniques than
		learning techniques.	about QoS or QoE. The QoS/QoE correlation
			is a case for evaluating the machine learning

Table 1: Summary of Strengths and Weaknesses of Proposed Models and Frameworks

			techniques.
[18]	Proposal for monitoring the	Only the concept and idea based	No concrete proposal or evaluation is present-
	cloud system for QoS per-	work in progress have been de-	ed.
	formance	scribed.	
[19]	Profit-Based Analysis of Re-	An innovative method for analyzing	No discussion on how to optimally allocate
	source Allocation on QoS	the impact of resource provisioning.	resources.
[20]	A distributed resource alloca-	Capable of handling multiple re-	Too simple, as it assumes perfect conditions
	tion algorithm for cloud and	source requirements	for execution. Failures after allocation of re-
	grid systems		sources are not taken into account.
[21]	Extensible dynamic provi-	The proposed framework is dynamic	May not be capable of handling bursty re-
	sioning framework for multi	and allocates resources depending on	quirements with short duration and large re-
	tenant cloud system	the tenant requirements.	source requirements. The new tenants arriv-
			ing late may suffer from resource starvation.
[22]	Lightweight framework for	Less resource intensive.	Does not monitor the real QoS parameters
	monitoring public clouds		such as response time, processing time etc.
[23]	A framework for handling	Based on multi-input-multi-output	Limited only to CPU and memory provision-
	adaptive applications in	feedback control model for resource	ing. Hence application performance may be
	cloud systems.	provisioning.	affected by other resource constraints such as
			network, storage etc.
[24]	A resource pricing model for	Uses realistic values using age as a	Utilization is not considered in computing
	QoS and profit balancing.	parameter.	cost. Hence may produce inaccurate costs.
[25]	A monitoring application for	Can be used by clients to monitor the	Very narrow application due to focusing only
	QoS parameters in iOS5.	performance of service providers.	on available transfer rate and one-way delay
[2(]		Multiple Orf generating and he	as QoS parameters.
[26]	A Qos based trust manage-	Multiple QoS parameters can be	No clear explanation on how to use the pa-
	ment model	usea.	to prioritize the peremeters
[27]	Resource allocation in a	The proposed strategy handles the	May load to sub optimal solutions from a sus
[27]	Compute Cloud through bar-	dynamic nature of cloud very well	tomer's perspective if a single provider can-
	gaining approach	during run time	not meet all the requirements
[28]	Investigation of the capability	Markov arrival processes have the	Only numerical experiments have been used
[20]	of MAP based queuing mod-	capability fir heavy tail distributions	to validate the model, hence needs further
	els for predicting workload of	that are common in web applica-	validation with real data traces.
	cloud systems.	tions.	valiaation whitefear and faces.
[29]	An optimization framework	Suitable for vendors selling products	Lacks the run time management of OoS per-
1 1	for cross layer cloud services.	across multiple layers. Dynamic na-	formance.
	5	ture of cloud has been considered.	
[30]	Algorithms for resource allo-	It helps reduce the cost of SaaS pro-	Due to reuse of already open VMs, it can cre-
	cation for SaaS providers for	viders without compromising the	ate security problems for customers.
	balancing cost and QoS.	QoS of customers.	
[31]	Results of an initial investiga-	It is shown that general labeling of	These are a set of experiments that can be in a
_	tion of using Dwarf bench-	cloud service providers for size or	laboratory by experts and may not help the
	marks to measure the per-	the number of units used is not suffi-	general set of customers who are not that tech
	formance of virtualized	cient to predict the real capabilities	savvy.
	hardware.	through real experiments.	
[32]	A process for matching pro-	A good effort for automating the	Only matches the service providers published
	viders' capability with cus-	matching process that was hitherto	capabilities with customer requirements. Thus
	tomers' requirements based	done manually by customers.	it cannot track the changes in cloud perfor-
	on SLA parameters.		mance due to dynamic nature of clouds.

# **4** CONCLUSIONS

Cloud computing has been the paradigm shift in distributed computing due to the way the resource provisioning and charging. Managing QoS is a critical task in making such an innovative technology to a larger audience. Several researchers have put forward their ideas for new and innovative solutions for handling this vital area is resource management. In this paper, the authors have carried out a critical review of the most recent work carried out in this area. The findings of the authors in terms of the strengths and weaknesses of the proposed work has been presented in a table for easy reference.

# REFERENCES

 R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5<sup>th</sup> Utility," *Journal of Future Generation Computer Systems*, vol. 25, no. 6, pp. 599–616, June 2009.

- [2] B.P. Rimal, E. Choi, and I. Lumb, I. "A Taxonomy and Survey of Cloud Computing Systems," Proc. 5<sup>th</sup> Int. Joint Conf. INC, IMS and IDC (NCM'09), pp. 44-51, 2009.
- [3] S.K. Garg, S.K. Gopalaiyengar, and R. Buyya, "SLA-based Resource Provisioning for Heterogeneous Workloads in a Virtualized Cloud Datacenter," Proc. 11th Int. Conf. Algorithms and Architectures for Parallel Processing (ICA3PP2011), pp. 371–384, 2011.
- [4] D. Bourcier, and P.D. Filippi, "Cloud Computing: New Research Perspectives for Computers and Law," Proc. 13th Int. Conf. Artificial Intelligence & Law, pp. 73–92, 2011.
- [5] G. Feng, S. Garg, R. Buyya, and W. Li, "Revenue Maximization using Adaptive Resource Provisioning in Cloud Computing Environments," *Proc.* 13th ACM/IEEE Int. Conf. Grid Computing, pp. 192–200, 2012.
- [6] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing," *Communications of the ACM*, vol. 53, no. 4. Pp. 50-58, April 2010.
- [7] C. Vecchiola, S. Pandey, and R. Buyya, "High-performance cloud computing: A view of scientific applications," *Proc.* 10<sup>th</sup> Int. Symp. *Pervasive Systems, Algorithms, and Networks (ISPAN)*, pp. 4–16, 2009.
- [8] R. Buyya, A. Beloglazov, and J.H. Abawajy, "Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges", Proc. Int. Conf. Parallel and Distributed Processing Techniques and Applications, pp.6-20, 2010.
- [9] Y. Li, W. Li, and C. Jiang, "A Survey of Virtual Machine System: Current Technology and Future Trends," *Proc. Third Int. Symp. Electronic Commerce and Security*, pp. 332-336, 2010.
- [10] J.M. Pedersen, M.T. Riaz, J.C. Júnior, B. Dubalski, D. Ledzinski, and A. Patel, "Assessing Measurements of QoS for global Cloud Computing Services," *Proc. Ninth IEEE Int. Conf. Dependable, Autonomic and Secure Computing*, pp. 682-689, 2011.
- [11] R. Buyya, S.K. Garg, and R.N. Calheiros. "SLA-Oriented Resource Provisioning for Cloud Computing: Challenges, Architecture, and Solutions," Proc. Int. Conf. Cloud and Service Computing, pp. 1-10, 2011.
- [12] X. Liu, Y. Yang, D. Yuan, G. Zhang, W. Li, and D. Cao, "A Generic QoS Framework for Cloud Workflow Systems," *Proc. Ninth IEEE Int. Conf. Dependable, Autonomic and Secure Computing*, pp. 713-720, 2011.
- [13] W.N. Chen, and J. Zhang, "A Set-Based Discrete PSO for Cloud Workflow Scheduling with User-Defined QoS Constraints," Proc. IEEE Int. Conf. Systems, Man and Cybernetics, pp. 773-778, 2012.
- [14] R.V. den Bossche, K. Vanmechelen, and J. Broeckhove, "Cost-Efficient Scheduling Heuristics for Deadline Constrained Workloads on Hybrid Clouds," Proc. 3<sup>rd</sup> IEEE Int. Conf. Cloud Comp. Tech. and Sc., (CloudCom), pp. 320-327, 2011.
- [15] V.C. Emeakaroha, I. Brandic, M. Maurer, and I. Breskovic, "SLA-Aware Application Deployment and Resource Allocation in Clouds," *Proc. 35th IEEE Annual Computer Software and Applications Conference Workshops (COMPSACW)*, pp. 298-303, 2011.
- [16] W. Li, Q. Zhang, J. Wu, J. Li, and H. Zhao, "Trust-based and QoS Demand Clustering Analysis Customizable Cloud Workflow Scheduling Strategies," *Proc. IEEE Int. Conf. Cluster Comp. Workshops*, pp. 111–119, 2012.
- [17] M.S. Mushtaq, B. Augustin, and A. Mellouk, "Empirical Study based on Machine Learning Approach to Assess the QoS/QoE Correlation", Proc. 17th European Conf. Networks and Optical Comm., pp. 1-7, 2012.
- [18] K. Alhamazani, R. Ranjan, F. Rabhi, L. Wang, and K. Mitra, "Cloud Monitoring for Optimizing the QoS of Hosted Applications," Proc. 4th IEEE Int. Conf. Cloud Comp. Tech. and Sc., pp. 765-770, 2012.

- [19] J. Li, Q. Wang, D. Jayasinghe, S. Malkowski, P. Xiong, C. Pu, Y. Kanemasa, and M. Kawaba, "Profit-Based Experimental Analysis of IaaS Cloud Performance: Impact of Software Resource Allocation," Proc. Ninth IEEE Int. Conf. Services Computing, pp. 344-351, 2012.
- [20] D. Adami, C. Callegari, S. Giordano, and M. Pagano, "A Hybrid Multidimensional Algorithm for Network-aware Resource Scheduling in Clouds and Grids," *Proc. IEEE Int. Conf. Comm.*, pp. 1297-1301, 2012.
- [21] A. Gohad, K. Ponnalagu, and N.C. Narendra, "Model Driven Provisioning in Multi-tenant Clouds." *Proc. Annual SRRI Global Conference*, pp. 11-20, 2012.
- [22] J. Ma, R. Sun, and A. Abraham, "Toward a Lightweight Framework for Monitoring Public Clouds," Proc. Fourth Int. Conf. Computational Aspects of Social Networks, pp. 361-365, 2012.
- [23] Q. Zhu and G. Agrawal, "Resource Provisioning with Budget Constraints for Adaptive Applications in Cloud Environments," *IEEE Trans. Services Computing*, vol. 5, no. 4, pp. 497-511.
- [24] B. Sharma, R.K. Thulasiram, P. Thulasiraman, S.K. Garg, and R. Buyya, "Pricing Cloud Compute Commodities: A Novel Financial Economic Model," *Proc. 12th IEEE/ACM Int. Symp. Cluster, Cloud and Grid Computing*, pp. 451-457, 2012.
- [25] F. Stoicuta, I. Ivanciu, E. Minzat, A.B. Rus, and V. Dobrota, "An OpenNetInf-Based Cloud Computing Solution for Cross-Layer QoS: Monitoring Part Using iOS Terminals," *Proc. 10th Int. Symp. Electronics and Telecomm.*, pp. 167–170, 2012.
- [26] M.K. Goyal, A. Aggarwal, P. Gupta, and P. Kumar, "QoS based Trust Management Model for Cloud IaaS," Proc. 2nd IEEE Int. Conf. Parallel, Distributed and Grid Computing, pp. 843-847, 2012.
- [27] G.N. Iyer and B, Veeravalli, "On the Resource Allocation and Pricing Strategies in Compute Clouds using Bargaining Approaches," 17th IEEE Int. Conf. Networks, pp. 147-152, 2011.
- [28] S.P. Sanchez, G. Casale, B. Scotney, S. McClean, G. Parr, and S. Dawson, "Markovian Workload Characterization for {QoS} Prediction in the Cloud," *Proc. IEEE Int. Conf. Cloud Computing*, pp. 147-154, 2011.
- [29] Y. Kouki, T. Ledoux, and R. Sharrock, "Cross-Layer SLA Selection for Cloud Services," Proc. First Int. Symp. Network Cloud Computing and Applications, pp. 143-147, 2011.
- [30] L. Wu, S.K. Garg, and R. Buyya, "SLA-based Resource Allocation for Software as a Service Provider (SaaS) in Cloud Computing Environments," *Proc. 11th IEEE/ACM Int. Symp. Cluster, Cloud and Grid Computing*, pp. 195-204, 2011.
- [31] S.C. Phillips, V. Engen, and J. Papay, "Snow White Clouds and the Seven Dwarfs," Proc. Third IEEE Int. Conf. Cloud Comp. Tech. and Sc., pp. 738-745, 2011.
- [32] T. Chauhan, S. Chaudhary, V. Kumar, and M. Bhise, "Service Level Agreement Parameter Matching in Cloud Computing," World Cong. ICT, pp. 564-570, 2011.

# Shared Services in Collaborative System Configurations

Norzilah Musa, Siti Zaleha Zainal Abiddin, Nasiroh Omar

Abstract— In network collaborative system, dispersed and co-located users work together to accomplish assigned task in order to achieve their shared goal. While working together, many work related activities are conducted such as interactive communication, sharing resources and manipulating media objects. Furthermore, these activities can happen in many ways either sequential, parallel, synchronous or asynchronous method. Hence, it is crucial to control and coordinate the shared resources, especially during the collaboration process. Currently, various kinds of effective, appropriate and sufficient shared services component used in supporting the collaborative activities. These services are important elements in the successful collaborative environment. However, the implementation involves some technical obstacle such as limited sharing of media objects, complexity in bridging the external sharing applications, lack of flexibility and ease to use sharing platform and neglect the ad-hoc sharing configuration by end-user. Thus, the shared-media are always under-utilized and services component management creates issues of object's access control and conflicts. This paper provides a general abstraction of service components and a comparative study on shared services on several collaborative system frameworks. Results show that existing shared services are not tolerably designed for end-users to design and develop their sharing methods easily.

Index Terms— Collaborative System Components, Collaborative System Framework, Media Sharing, Shared Services, Sharing Features, Usability, User Defined, Workspace.

# **1** INTRODUCTION

TODAY, rapid improvement in information and communication technology had influenced the growing of the internet availability, decrease the bandwidth cost and increase the computational power. These are some of the factors that make collaborative work become easier and accelerate various techniques for user to collaborate. Nowadays, distributed user can work under one organization that manage the shared resources in a flexible and secured manner among the collaborators. In addition, the collaborative work can occur between two or more organizations. These kind of working environment helps to solve issues in critical resources, personnel and logistics. Furthermore, naturally 30% to 50% productivity will increase by implementing this kind of working environment [1].

Dispersed users have various kinds of behavior, skills and experience, especially users from teams and group project from different organizations. Moreover, different organizations will have different information systems and communication structures that do not sufficiently support the collaborative works. Most of these applications are mainly created to be stand alone systems without considering cooperation and integrated features. Hence, simpler and ease to use technology that is well-suited with project group and organizational goals, work practices and standards should be adopted [2]. As for that, the development of information systems is evolving towards subsystem that can be easily integrated and configured [3] to make a successful collaboration process.

Most of collaboration process is about sharing activities. As

an example, users shared the communication channel to communicate to each other, important documents or artifacts are shared to make important decisions, dispersed users are sharing the same view of the whiteboard and shared same context and virtual room during the online meeting. In fact, different scenarios of collaborative works will require different sets of shared services. Hence, a successful collaborative systems should overcome any sharing service issues such as difficulties in integrating diverse software tools with collaborative environment, able to manage various sets of service capabilities provided by the subsystems and configuration service for communication channels to allow the media sharing process [4].

Currently, the choice of shared services in a collaborative system is often determined by cost, technical boundaries and the ability to incorporate with the existing system. Though, web-based shared services were designed to support collaboration of distance users, it limits the scope for the user to integrate origin data into various systems. Hence, the role of shared services in the collaborative activity works has been investigated [5][6][7][8]. This research is initiated with the focus on various aspects of the design and implementation of the shared services in the collaborative system framework. Based on the study of various design and implementation of shared services in collaborative works, future research could provide guidance for effective media sharing mechanism.

The remaining content of this paper is organized as follows: section 2 introduces the collaborative system framework. Section 3 discusses type of sharing and shared services followed by a comparative study of the shared service features in section 4 and concluding remarks in section 5.

# **2 COLLABORATIVE SYSTEM FRAMEWORK**

The framework discussed in this paper described the de-

Norzilah Musa is currently pursuing PhD program in computer science at UniversityTeknolgi MARA, Malaysia, E-mail: norzi105@salam.uitm.edu.my

Siti Zaleha Zainal Abiddin and Nasiroh Omar are currently associate professor and senior lecturer respectively at Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA. Malaysia. E-mail: sitizaleha533@salam.uitm.edu.my; nasiroh@tmsk.uitm.edu.my

sign and implementation of collaborative systems in the working environment. Software framework is defined as a common and reusable platform to develop collaborative application systems. The framework consists of compilers, supporting tools, source code libraries, application program interface(API) and service components [9]. The synergy between these elements enable the creation of effective solutions in solving the collaboration problems. It is agreeable among the researchers [9][10][5], the software framework had eased the application development process by permitting the system designer and programmer to spend more time on other important matters such as software and technical requirements. Furthermore, application programmer would also spend less time on coding and debugging to give more space for creativity as a value added features to the system.

In developing collaborative systems, it is crucial to understand the way people collaborate in order to improve the collaboration process. Furthermore, the process is bound to the collaborative environment which consist of organizational structure and policy, technical setting, working culture, communication setting and information workflow. Hence, the collaborative system framework consists of three elements; environment, process and support [11] as illustrated in figure 1.



Figure 1. The Collaborative System Framework [11]

The first element is the collaborative environment. The environment is about the setting of the organizations that took part in the collaboration process and approached used in handling the collaborative activities. In addition, the environment attributes will vary from one collaboration process to another process which also influence the value of the other two elements.Hence, this element is divided into two levels: planning and requirement level. The planning level is the initial phase in setting up the collaboration process. It defines the scope and boundary of the systems concept, risk management, costbenefits, project management plan, collaboration goals and the outcome. Meanwhile, requirement level specified the group of user requirements in order to perform their assigned task. It analyzed the requirement with respect to the task assigned, project planning, social protocol, organization and interorganization policies and group characteristics [12] . It includes the supporting requirement for the various kinds of groups due to the size of the group, locations of the users and the computer platform.

As a second element, collaboration process is about a group of users working together regardless their location but still can communicate effectively to achieve the assigned goal. During the process, workflow management between all elements in

the collaborative are important. Basically the workflow are between these elements' relationship; [1] user-to-user, user-todevice, device-to-data, application-to-application, applicationto-device, device-to-platform, platform-to-network and organization-to-organization. Therefore, client/server and peer-topeer (P2P) platforms are one of the key attributes of this element. These platforms provide user with various kinds of technology that suits with the organization background and capabilities. As an example, peer-to-peer platform support decentralization work management which communication between group of users are managed by agents and normally used JXTA technology in its implementation. While client/server that used centralized management approach rely on the server and a client network connection. Web services and extended markup language (XML) is common approaches used in developing the collaborative system on this platform. Currently, various kinds of method evolved from these two platforms to help dispersed users in the collaboration process, such as cloud computing and grid technology. On the whole, the collaboration process is about the decision making and the management of the coordination, cooperation and communication between the components of the collaboration towards the assigned task. The execution tasks are implemented in many ways. Several ways can be by integrating or incorporate external sub-system or compose new sub-process manually by the user.

The third element is the collaboration support. The collaborative systems must allow users regardless with the location and technology constraints, the user must be able to collaborate in effective ways. The involved organizations must take action and make it possible by providing users with effective and suitable tools and a range of services to support the collaborative practices. Hence, the collaboration support can be divided into three levels; capability, service and technology.

The capability level defines fundamental capabilities that support various kinds of collaborative tasks. Here, the tasks are matched up with services that contains special collaborative capabilities [12]. The matching process consists of two processes. First is the evaluation process. The available services are evaluated with its capabilities in supporting the current tasks. The second process is about the adaptation and acceptance process between the services and the group characteristics. Basically, the capability level is about the ability of the system to support the whole collaboration process. The service level identifies a right collaboration practices and tools by offering a list of services that can be used in executing the collaborative tasks. Some of the services are independent and autonomous by itself that can be directly used. Other services are created by composing different kind of services from different domain by using the platform and infrastructure services. Technology as a last level in the collaboration support element, deals with the implementation of the services that reflect the performance of overall collaborative work.

#### **3 Type of Sharing and Shared Services**

Although many systems and frameworks claim to be a collaborative system, but most of them only support certain features and criteria of collaboration [13][14]. These systems are actually closed systems which means it provides a fixed set of services, tools and functions to support certain features of collaboration. These systems lack of flexible services which require dynamic changes in a collaborative process that in line with the current trend of business process. Due to that, services play an important role in new computing technology that support independence and distributed systems in developing flexible collaborative systems [15]. These services are actually autonomous computational agents that made the integrated and distributed collaborative solutions possible [3]. As there are various kinds of services in the collaborative framework such as session management services, data distribution services, message management services and device management services; this paper focus only on software shared services in the collaborative works.

Based on the literature, in general, there are two types of sharing process; context and resource sharing processes.

### 3.1 Context Sharing

A group of users might work on a small part of a big project, but at the same time are interested of the changes made by another group of the related parts of the project. Hence, all users working on the same project shall share the same context in the same manner. Here, the awareness mechanisms work as shared services between the group of users. It will facilitate awareness of changes done to one part of the same product to the other users.

#### 3.2 Resources Sharing

As mentioned above, collaborative is about sharing process. Distance users shared the same goals and same resources in doing their work. The resources are anything that needed to support the assigned tasks and can be retrieved by the users. It can be information, document, multimedia objects and file. As depicted in figure2, the data is represents the resources and media is represents the medium of communication between the users. Each of these resources has their own method of control and coordinate to govern its activities. Furthermore, the collaborative activities also govern by its important attributes, such as policy, organization protocol, computer platform, context awareness, space and time.



Figure 2: Resources Sharing Logical Model

Dispersed users need services to assist them in managing the shared resources, as they are manipulating the same objects. The shared services in resource sharing are varied and relative to the objective of the collaborative activities. Table 1 shows the type of shared services related to the work task. These shared services can classify into four categories[16]; objects sharing services, communication sharing services, sharing collaboration space services and workflow sharing services.

Table 1: Type of shared services related to the activity

Activity	Shared Services	Type of resources
Editing	Edit,	File Document
Document	Update,	User
	Delete	
	View,	
	Storage	
	Context Awareness	
	Version Control	
	Message	
	Merge	
	Locking	
	Concurrency	
	Session Management	
On-line	Drawing	Multimedia Objects
Meeting	Presentation	File Document
	View	User
	Context Awareness	
	Chat	
	Space Control	
	Locking	
	Concurrency	
	Session Management	
Learning	View	Multimedia Objects
-	Forum	File Document
	Locking	User
	Session Management	
	Message	
	Context Awareness	
E-	Concurrency	User
Conferencing	Space Control	Multimedia Objects
	Session Management	
	Context Awareness	

#### 1) Object Sharing Services

Collaboration means sharing something important among the participant users. It is crucial to share any supporting objects such as text document and slide presentation in any collaborative works. Hence, it is important to coordinate the changes made by activity related to this sharing objects. Edit, delete, add and view are examples of common activity related to the shared object. As this object are shared by many users, shared services have to make sure the same version of objects can be viewed and interpret in the same way by all the users. The shared services also should equip with conflict management to handle any conflict that arise during the collaborative activity.

2) Communication Sharing Services

Effective communication is a key success factor in achieving activity objectives. It is important to avoid any misunderstand and discrepancy during the interaction. As users are sharing the same communication media, the shared services will make sure users will get the same message at the same time with a user friendly user-interface and easy to understand error message.

# 3) Collaboration Space

Shared services will provide users with the same context in the same style as all the others are located in the same room.

4) Workflow Sharing Services.

Users will know the chairperson or the owner of the collaborative activity, expertise and role of the existing members are examples of services provided by this category.

Most of the shared services are built-in as a functional requirement of the systems and others are stored as source code library and API that ready to serve when needed. Some of the collaborative work such as an online meeting will face difficulty when dealing with ad-hoc matters. In this situation, users probably need editing sharing application urgently that enables them to formulate new marketing strategy. As the existing systems are not supplied with functions to connect to the external standalone applications, it forces the developers to change the hard code of the existing system. As an alternative way, in these situations the developers can choose from the range of available external shared services applications and tools that map to the collaborative activity's objective [17]. This process is referred as shared service composition [18]. It allows developers to customize the shared services from dynamic searching of available standalone applications, integrate and implement the shared services in a proper order to achieve the activity objective.

These are examples of third-party or external standalone applications that can be used as shared services in the collaborative works [17]:

1) Version-control Systems

It facilities the version control systems either by peerto-peer and client/server platform. Examples of the systems are Subversion, Git, Mercurial and Darks.

2) Online Meeting

This system allows users to edit documents while others view it, whiteboard drawing, co-browsing and support synchronous and asynchronous activities. Examples of this system are WebEx, NetMeeting, conference and WorkSpace3D.

3) Document Sharing

This system allows users to review and edit concurrently with other users. Examples of this system are GoogleDoc and Google Wave.

4) Project Management

This system facilitates project information such as project milestone, planner and overview ongoing project through the web. Examples of this system are ActiveCollab, WorldView, WorkspaceActivityViewer.

5) Electronic Conferencing

The system facilitates the conference activity such as data conferencing, voice conferencing, video conferencing and discussion forums. Examples of this system are CU-SeeMe and NetMeeting [19].

These applications allow developers to customize the shared services from dynamic searching, integrate and implement the shared services in a proper.

# 4 LIMITATION OF THE AVAILABLE SHARED SERVICES IN COLLABORATIVE SYSTEM FRAMEWORK

Most of the available collaborative software framework deals with various kinds of sharing process that involve with different types of objects. The shared services offered by the frameworks are almost the same though develop on different architecture and platform. It shows that, different type of sharing activity will use almost the same set of shared services regardless the type of problem they are.

However, in comparing these frameworks, it was found that different kind of approaches are used to develop and execute the shared services. This approach mainly influenced by the availability of the experts, the organization background technology and collaborative environment. Furthermore, this service that supposedly support groups of user which engaged in the same activity are limited with flexibility and usability. This is due to most of these services are pre-defined and embedded into the system. Although some of the frameworks allow these services to be integrated with other collaboration tools, the processes are still not dynamic. The users are served with pre-defined external tools by the system. Users are not allowed to choose from the range of external tools that they prefer to use. On the other hand, some of the frameworks provide users with service composition that allow users to create new service operation. But this function is not for end-user and difficult to operate. Table 2 summarizes the software framework shared services category with it's strengths and limitations.

# **5 CONCLUSION**

Collaboration technology should develop in small stages. It starts with the design of configurations which is required at each of its components thorough consideration. As sharing is a nature of the collaborative activity, the configuration should provide modules that facilitate sharing among dispersed users. Hence, appropriate shared services are important facilities in the system to produce efficient and consistent collaborative works. In this paper, we give detail explanation about collaborative configurations. We also discuss sharing process focussing on four categories of sharing services; objects, communication, workspace and workflow. Each category has a different set of services which serves collaborative activity in its bounded territory. Currently, there are lacks of research work that are capable of providing such services. Moreover, most of the configurations have their own shared services which are being pre-defined and hard coded into the system. Therefore, the sharing platform cannot allow end-users to flexibly choose their own preferred shared services from common applications. In addition, their ability is limited to certain types of ca-

Table 2 : Shared Se	ervices Comparison	among Collaborative	System Framework
---------------------	--------------------	---------------------	------------------

Shared Services	Shared Activity	Shared Approach	Usability	User-Define	Ad-hoc	Reference
				Configuration		
Communication	Media Transmission, Communication, Data Management, Simulation, Activity Management, Communication Interoperability , Translation, Image Detector, Planning, Notebook, Communication, Remote Control,Viewing, Editing, Modeling	Embeded , Integrate	Easy to Moderate	Static , Dynamic	Allow	[13][20] - [28][36]- [38][38][40]-[42][45]
Resources	Planning, Media Transmission, Communication, Data Management, Activity Management, Planning, Notebook, Communication, Remote Control, Planning, Virtual Lab, Media Sharing, Storage, Display, Composition Models, Sketching	Incorporate, Embeded, In	Moderate to Hard	Static, Dynamic	Not Provide	[4][21]-[24][26][28][29] [31] - [45]
Workflow	Media Transmission, Communication, Data Management, communication Interoperability, Planning, Access Control, Sketching, View, Editing, Modelling, Access record	Embeded	Moderate to Hard	Static	Not Provide	[22][25][26][30][33]- [40][45]
Collaboration Spaces	Media Transmission, Communication, Data Management, Simulation, Planning, Virtual Lab, Media Sharing, Storage, Display	Embeded, Incorporate	Easy to Moderate	Static	Not Provide	[20][26][4][29][43][45]

tegory only. In general, collaborative working environment is branded by flexibility, usability and ad-hoc communication service requirements. Moreover, such services are very significant in promoting flexible controlling and coordinating mechanism in distributed media sharing activities.

## ACKNOWLEDGEMENT

The authors would like to thank Universiti Teknologi MARA and Ministry of Higher Education, Malaysia for the financial support.

#### REFERENCES

- S. Liu, B. Spencer, W. Du, and C. Chi, "Towards an open collaboration service framework," in 2011 International Conference on Collaboration Technologies and Systems (CTS), 2011, pp. 77–85.
- [2] H. Patel, M. Pettitt, and J. R. Wilson, "Factors of collaborative working: a framework for a collaboration model.," Applied ergonomics, vol. 43, no. 1, pp. 1–26, Jan. 2012.
- [3] A. L. Osório, H. Afsarmanesh, and L. M. Camarinha-matos, "A Service Integration Platform for Collaborative Networks," Studies in Informatics and Control, vol. 20, no. 1, pp. 19–30, 2011.
- [4] W. Prinz, H. Loh, M. Pallot, H. Schaffers, A. Skarmeta, and S. Decker, "ECOSPACE -- Towards an Integrated Collaboration Space for eProfessionals," in 2006 International Conference on Collaborative Computing: Networking, Applications and Worksharing, 2006, pp. 1– 7.
- [5] H. Afsarmanesh, M. Sargolzaei, and M. Shadi, "A Framework for Automated Service Composition," in Collaborative Networks in the Internet of Services, H. Camarinha-Matos, LuisM. and Xu, Lai and Afsarmanesh, Ed. Springer Berlin Heidelberg, 2012, pp. 63–73.
- [6] N. Schuster, C. Zirpins, and U. Scholten, "How to balance flexibility and coordination? Service-oriented model and architecture for document-based collaboration on the Web," in 2011 IEEE International Conference on Service-Oriented Computing and Applications (SOCA), 2011, pp. 1–9.
- [7] S. Woo, H. J. Park, and J. Kim, "Open media service architecture for advanced collaboration environments," Multimedia Tools and Applications, pp. 133–160, 2009.
- [8] R. W. H. Lau, R. Klamma, S. Chen, and B. Wah, "Advances in ubiquitous media technologies and applications," World Wide Web Internet And Web Information Systems, pp. 217–222, 2011.

- [9] "Frameworks Tridium." [Online]. Available: http://www.tridium.com/cs/products\_/\_services/frameworks. [Accessed: 07-Nov-2012].
- [10] J. Hill and C. Gutwin, "The MAUI Toolkit : Groupware Widgets for Group Awareness," Work, no. 2004, pp. 539–571, 2005.
- [11] P. E. Weiseth and B. E. Munkvold, "The Wheel of Collaboration Tools : A Typology for Analysis within a Holistic Framework," Environment, pp. 239–248, 2006.
- [12] J. Cugini, L. Damianos, L. Hirschman, R. Kozierok, J. Kurtz, S. Laskowski, and J. Scholtz, "Methodology for Evaluation of Collaboration Systems." [Online]. Available: http://zing.ncsl.nist.gov/nisticv/documents/method.html. [Accessed: 11-Feb-2013].
- [13] D. Wigdor, "WeSpace : The Design, Development, and Deployment of a Walk-Up and Share Multi-Surface Collaboration System," pp. 1237–1246, 2009.
- [14] P. Cesar, A. D. C. A. Bulterman, A. D. Geerts, B. J. Jansen, H. K. C, W. S. C, and A. Cwi, "Enhancing Social Sharing of Videos : Fragment, Annotate, Enrich, and Share," in MM '08 Proceedings of the 16th ACM international conference on Multimedia, 2008, pp. 11–20.
- [15] C. Guo, "A Service-Oriented Framework for Collaborative Working Environment," 2011 Second International Conference on Networking and Distributed Computing, pp. 209–213, Sep. 2011.
- [16] I. Jørstad, "A Service Oriented Architecture Framework for Collaborative Services," in Enabling Technologies: Infrastructure for Collaborative Enterprise, 2005. 14th IEEE International Workshops on, 2005, pp. 121–125.
- [17] F. Lanubile, C. Ebert, R. Prikladnicki, and A. Vizcaíno, "Collaboration Tools for Global Software Engineering," Software, IEEE, vol. 27, no. 2, pp. 52–55, 2010.
- [18] B. Limthanmaphon and Y. Zhang, "Web b Composition with Case-Based Reasoning," vol. 17, 2003.
- [19] M. Al-ma, "Using Electronic Collaborative Media in Knowledge Sharing Phases : Case Study in Jordan Hospitals," Education and Information Technologies, vol. 2, no. 4, 2008.
- [20] C. A. Jara, F. A. Candelas, F. Torres, S. Dormido, F. Esquembre, and O. Reinoso, "Real-time collaboration of virtual laboratories through the Internet," Comput. Educ., vol. 52, no. 1, pp. 126–140, 2009.
- [21] D. Mourtzis, "CIRP Journal of Manufacturing Science and Technology Internet based collaboration in the manufacturing supply chain," CIRP Journal of Manufacturing Science and Technology, vol. 4, no. 3, pp. 296–304, 2011.
- [22] H. Sangwoo, K. Namgon, C. Kiho, and K. Jong Won, "Design of Multiparty Meeting System for Interactive Collaboration," in Communica-

tion Systems Software and Middleware, 2007. COMSWARE 2007. 2nd International Conference on, 2007, pp. 1–8.

- [23] N. Cheaib, S. Otmane, and M. Mallem, "Collaborative Multimedia Collection for Enriching and Visualizing 3D Underwater Sites," 2009 6th IEEE Consumer Communications and Networking Conference, pp. 1–5, Jan. 2009.
- [24] K. Dube, E. Mansour, and B. Wu, "Supporting Collaboration and Information Sharing in Computer-Based Clinical Guideline Management," in 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05), 2005, pp. 232–237.
- [25] M. Ciampi, L. Gallo, A. Coronato, and G. D. Pietro, "Middleware mechanisms for interaction interoperability in Collaborative Virtual Environments," International Journal of Advanced Media and Communication, vol. 4, no. 2, p. 154, 2010.
- [26] A. Saddik, A. Rahman, S. Abdala, and B. Solomon, "PECOLE: P2P multimedia collaborative environment," Multimedia Tools and Applications, vol. 39, no. 3, pp. 353–377, Aug. 2007.
- [27] R. Ranon, L. D. Marco, A. Senerchia, S. Gabrielli, and L. Chittaro, "A Web-based Tool for Collaborative Access to Scientific Instruments in Cyberinfrastructures," in Grid Enabled Remote Instrumentation, Springer Berlin / Heidelberg, 2009, pp. 237–251.
- [28] T. Pering, R. Want, B. Rosario, S. Sud, and K. Lyons, "Enabling Pervasive Collaboration with Platform Composition," in Enabling Pervasive Collaboration with Platform Composition, Springer Berlin / Heidelberg, 2009, pp. 184–201.
- [29] I. F. Cruz, R. Gjomemo, and G. Jarzab, "An Interoperation Framework for Secure Collaboration," in Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS, 2010, pp. 4–11.
- [30] D. Sun, S. Xia, C. Sun, and D. Chen, "Operational transformation for collaborative word processing," in Proceedings of the 2004 ACM conference on Computer supported cooperative work - CSCW '04, 2004, vol. 6, no. 3, p. 437.
- [31] M. Boddy, M. Michalowski, A. Schwerdfeger, H. Shackleton, and S. Vestal, "FUSED : A Tool Integration Framework for Collaborative System Engineering."
- [32] U. B. Sangiorgi and J. Vanderdonckt, "GAMBIT : Addressing Multiplatform Collaborative Sketching with HTML5," in Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems, 2012, pp. 257–261.
- [33] W. Fagen and S. Kamin, "Developing Device-independent Applications for Active and Collaborative Learning with the SLICE Framework," in World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2012, vol. 2012, no. 1, pp. 1565–1572.
- [34] W.-H. Lee, S.-S. Tseng, and W.-Y. Shieh, "Collaborative real-time traffic information generation and sharing framework for the intelligent transportation system," Information Sciences, vol. 180, no. 1, pp. 62–70, Jan. 2010.

- [35] C. Guo, "A Service-Oriented Framework for Collaborative Working Environment," in 2011 Second International Conference on Networking and Distributed Computing, 2011, pp. 209–213.
- [36] J. Park, D. Yi, and J. Kim, "A Futuristic Service Framework for Flexible Media-oriented Service Composition in Future Internet \*," 2010.
- [37] B. Xu, H. Lin, L. Chiu, Y. Hu, J. Zhu, M. Hu, and W. Cui, "Collaborative virtual geographic environments: A case study of air pollution simulation," Information Sciences, vol. 181, no. 11, pp. 2231–2246, Jun. 2011.
- [38] D. D. He, M. Compton, and K. Taylor, "Access Control : What is Required in Business Collaboration ?," 2008.
- [39] S. W. Han and J. Kim, "A service composition oriented framework for configuring SMeet multiparty collaboration environments," Multimedia Tools and Applications, Mar. 2012.
- [40] C. Gadea, B. Soloman, B. Ionescu, D. Ionescu, and G. Prostean, "A Real-Time Browser-Based Collaboration System for Synchronized Online Multimedia Sharing." 2010.
- [41] D. R. Herrick, "Google This ! Using Google Apps for Collaboration and Productivity," 2009, pp. 55–63.
- [42] A. Saddik, D. Yang, and N. D. Georganas, "Tools for transparent synchronous collaborative environments," Multimedia Tools and Applications, vol. 33, no. 2, pp. 217–240, Nov. 2006.
- [43] T. Kawashima and J. Ma, "TOMSCOP a synchronous P2P collaboration platform over JXTA," 24th International Conference on Distributed Computing Systems Workshops, 2004. Proceedings., pp. 85–90, 2004.
- [44] M. Bourimi, F. Kühnel, and J. M. Haake, "Tailoring Collaboration According Privacy Needs in Real-Identity Collaborative Systems," L. Carriço, N. Baloian, and B. Fonseca, Eds. Springer Berlin / Heidelberg, 2009, pp. 110–125.
- [45] W. Wang, "Powermeeting," in Proceedings of the nineteenth ACM conference on Hypertext and hypermedia - HT '08, 2008, p. 251.

# Case Study: Using Spreadsheet to Deploy Chained Business Rules

Kang Mun Arturo G. Tan, Jr. DBA\*

**Abstract**— This paper presents some ideas about business rules and how they are managed. The translation of business rules into electronic format enables the firm to capitalize on its written rules. By encoding electronically these rules, a junior employee could perform the activities which were used to be performed by a senior employee. While there are better software that can be used for this purpose, the ubiquity of Excel spreadsheet makes it a choice as a vehicle for translating these techniques into its proper usage. Further, critical Excel functions were shown on how to manage and encode these business rules into spreadsheet. Eleven functions or techniques of Excel spreadsheet were identified as a key technique in transforming business rules into a viable spreadsheet application. A small case study is shown to illustrate the techniques.

Index Terms— Spreadsheet Application, Business Rules, Business Rules Management Systems, Excel Techniques, Expert Systems, Micro-lending, Knowledge Acquisition.

----- 🔶 ------

# **1** Motivation

The motivation in writing this paper is to be able to highlight the usefulness of three ideas: chained business rules, tabular interpretation of business rules and specific techniques of Excel spreadsheet. These three ideas, when used properly, bring tremendous benefits to business, especially small businesses, as they could make use of the Excel spreadsheet to delegate functions to junior, but trusted, personnel.

# 2 Review of Literature

## 2.1 Using Business Rules for Business Intelligence

Mircea et.al asserted the need for business rules in a global economy challenged by many economic difficulties. Business rules "allow automatic data interpretation, definition of performance key indicators, through important redline values and provide solutions for problem solving."[1] However, Mircea et.al. acknowledged the existence of various meanings of "business rule."

Ronald G. Ross, as quoted by Mircea et.al., provided the following definition: "From a business perspective, business rules are literally the encoded knowledge of your business practices, and from an Information Technology point of view, a business rule is an atomic piece of reusable business logic." (Ross, 2003) [1]

Mircea et.al. went on to define nine categories of proposed rule patterns. These are: 1) terminology definition; 2)relationship of terminologies; 3)classification rules; 4)enumeration of items; 5)simple constraints; 6)list constraint; 7)calculation methodology; 8) inference; and 9)reaction. [1]

From the nine categories, the inference and reaction patterns refer to an IF-THEN construct, a construct that is available in many programming languages. This implies therefore that many of the operations manual key decision elements can be captured into the IF-THEN construct. Once this is done, it will be easier to transfer them into electronic format.

# 2.2 Business Process Requirements Engineering

Roger et. al propose the formalization of business rules.[2] One very important assertion they made is that quality products could only be attained if the customer's perspectives are included in the production process. (This idea is well known in the study of quality management like 6Sigma.) Moreover, they explicitly stated that customer's perspective should be represented among the set of business rules that will drive the business processes.

Further, aside from stating quoting various business rules, Roger et.al propose various aspects and operations that can be done on business rules. For instance, business rules could be a necessity or an obligation. For instance, a "necessity" is that potatoes should be fried. On the other hand, frying the potatoes for 3 minutes plus or minus 30 seconds is an "obligation."

Operations on the rules like equivalence of rules, merging of rules and sequencing of rules have been presented. Roger et.al concluded that in the definition of rules, four perspectives needed to be included, namely: customer perspective, business perspective, employee perspective and product perspective. Customer perspective defines customer's perception of required quality levels. The business perspective defines how to supply the needed efforts to satisfy the customer perspective. The employee perspective identifies skills necessary to perform within the business perspective. Finally, the product perspective presents various measures regarding the product specifications. [2]

## 2.3 Business Rule Management System

Carlo et.al. asserted the various aspects needed for a firm to be innovative in using knowledge which are captured as rules. [3] Taylor discussed that rules could be utilized properly with the use of a Business Rule Management System (BRMS).[4] Taylor defined BRMS to be a "complete set of software components for the creation, testing, management, deployment, and on-going maintenance of business rules in a production operational environment."[4]

Among other functions, a BRMS could do the following: 1) Business Rules Repository, 2) Design Tools, 3) Rule Management Interfaces, 4) Validation and Verification Tools, 5) Testing Tools, and 6) Business Simulation Tools.

These six capabilities enable the company to transfer "rare" expertise to the rest of the organization. Further, when these rules are electronically captured, the decision maker is able to distance his judgement from his emotions. This emotional distancing enables the decision maker to have a clearer and consistent mind when making a decision. [4]

## 2.4 Synthesis and Inferences

In general, the IF-THEN construct is a predominant way to represent knowledge. The IF-THEN constructs could then be captured in a tabular format which we could search through. A successive IF-THEN could be captured using several tables, with the left column representing the IFs and the right column representing the THENs.

This observation agrees with Mr. Jan Vanthienen when he wrote that *"tabular models for decisions* is a powerful technique highly relevant for business rules, business processes, business knowledge, and related areas." [5]

Hence, we could have:

IF SignColor = Red THEN Stop\_Car. (1) IF SignColor = Green THEN Move\_Car. (1) IF Car = Stop\_Car THEN Open\_Radio. (2) IF Car = Move\_Car THEN Close\_All\_Phone. (2)

Table1:

SignColor	Action
Red	Stop_Car
Green	Move_Car

Table2:

Car	Action
Stop_Car	Open_Radio
Move_Car	Close_All_Phone

At this point, while we agree with Taylor about the need for a BRMS, this paper asserts that Excel Spreadsheet can be sufficiently used as a Business Rule Management System. Many small and medium scale enterprises are already in Excel. It will be a great service to these firms if we could show how some techniques to translate Excel into a viable BRMS.

#### **3 Lists of Important Excel Formulas and Constructs**

In using Excel as a Business Rules Management System, we are presenting here various spreadsheets. These techniques are needed in encoding business rules. Jell wrote about the improvements done on Excel by Microsoft. This improvement consists of hiring "two outside math consultancy firmst to rewrite the algorithms" and then hiring a "third firm to validate whicfh algorithm was the best for each function."[6] These activities encourage us to believe Excel is a reliable platform for decision support systems. The following identifies the eleven techniques most useful for business rules development.

#### **3.1 List of Options or Possibilities**

This is a list of all possible answer to a checklist or question, such as Red, Orange, Yellow, Green, Blue, Indigo and Blue. In other cases, prompt like "Indicate Status of Store Activity" would demand answers from any of the following choices: "Low Count, Medium Count – Staff Relax, Very Dense – Staff Very Tense."

Describing these situations in one's own words would make the response accurate, however, from the point of view of the BRMS, using a pre-defined response items would make it more amenable to analysis.

# 3.2 Drop Down

From the list of options as discussed in topic number 3.1, we indicate here options which will enable to reflect those choices in the spreadsheet.

This is reached from the Main Menu Bar ->Data ->Data Validation, Settings (List) and Source. Here, we access the List of Options or Possibilities. By doing this way, we prevent the possibility of error in spelling by being consistent in how we refer to our answer.

It has to be noted that drop-down list should not exceed nine choices. Beyond nine, the end-user of the system would make it difficult to use the system. However, there are cases when we really need to go beyond nine, for instance, when we have to identify the country of origin of an applicant. In those cases, we probably need to break them into groups using different techniques, such as use of different fonts.

#### 3.3 IF-THEN

The If-Then construct connects a certain situation to another situation. If the linkages are simple, then we can use the following:

IF(A1 = "Yes", "Yes is the Answer", "Not Yes is the Answer")

This construct is actually connecting this paper to BRMS and to expert systems as well. By chaining the responses from one rule to another, we are able to execute the business rules properly.

For example:

#### Group A:

If wearing\_a\_pant then applicant\_is\_a\_man. If wearing\_a\_trouser then applicant\_is\_a\_woman.

#### Group B:

If applicant\_is\_a\_man then could\_hire\_as\_driver. If applicant\_is\_woman then could\_hire\_as\_cook.

Group C: If could\_hire\_as\_driver then take\_garage\_test. If could\_hire\_as\_cook then take\_kitchen\_test.

By chaining these groups of rules together, we are able to reach Group C.

## 3.4 AND and OR

There are cases when we have to have all situations available. Having all situations available means we have to use AND. However, if only one needs to be available, then OR would suffice.

AND(A1 = "b", A2="c") OR(A1="b", A2= "c")

Pursuing the example earlier, let us assume that the test for male gender has gone beyond the wearing of a pant to the test of having a short hair and the test for female gender is by wearing an earing then we could modify group A to be:

Group A: If wearing\_a\_pant OR having\_short\_hair then applicant\_is\_a\_man.

If wearing\_a\_trouser OR wearing\_an\_earing then applicant\_is\_a\_woman.

Let us further assume requires that all male should be able to lift 150 kilograms of dumb bell and all female should be able to cook rice in 20 minutes, then we should modify the business rules as:

Group A: If (wearing\_a\_pant OR having\_short\_hair) AND (can\_lift\_150\_kilograms) then applicant\_is\_a\_man.

If (wearing\_a\_trouser OR wearing\_an\_earring) AND (can\_cook\_rice\_20\_minutes) then applicant\_is\_a\_woman.

These constructs of OR and AND improve our capability to capture and encode business rules.

#### 3.5 Vlookup

Since we have a tabular interpretation of the If-Then construct, we then try to get the Then part by the Vlookup construct. The Vlookup has the following syntax: Vlookup(If\_value, TabularArea, 2, True\_for\_Approximate or False for Exact).

However, there are cases when the tabular interpretation is not complete. We then add an error mechanism, like: = IF(ISERROR(VLOOKUP(F46,MYTABLE1,2,FALSE))),"NO RESPONSE",VLOOKUP(F46,MYTABLE1,2,FALSE)) The presence of the ISERROR ensures that if there is no corresponding ACTION part, then a "NO RESPONSE" will come out.

As we have asserted in the synthesis, several rules of similar structure could be grouped together into a group. For instance,

- If Traffic\_Light\_is\_Red Then Stop\_the\_Car.
- If Traffic\_Light\_is\_Green Then Move\_the\_Car.
- If Traffic\_Light\_is\_Orange Then Prepare\_to\_Stop\_Car.
- If Someone\_Does\_Not\_Pay Then Call\_Police.

While all of these can be captured in one single table of IF( first column) and THEN (second column), the fourth rule does not seem to belong in this group. We then have to set up a different table for that group, if there will be several rules of that type.

# 3.6 HyperLink and the Hyperlink function

To connect to another sheet or spreadsheet, we could use the hyperlink: Main Menu – Insert – Hyperlink or The Hyperlink function is denoted by: =HYPERLINK("[BOOK.XLSX]SHEET!COLROW")

This capability enables us to bring into the solution set the use of other Microsoft Document such as Word, Powerpoint or even Access. Quite often, many of the forms are needed as part and parcel of a solution set. By incorporating these into the spreadsheet, but located in another workbook, we able to enhance the solution we provide.

Further, there are cases when solutions or rules are needed to be quoted or analyzed. By the use of the hyperlink, we are able to bring in rules and procedures (or even checklist) into the set of business rules.

#### 3.7 Formulas

Any calculations would use the following: "= A1 + A2." We can always academic-based formulas being used extensively in the industry, for instance the computation of future value (FutureValue = PresentValue \* (1 + InterestRate^n), where n is the number of compounding.)

Another formula will be the accounting formula of Asset = Liability + Equity. Computations of Economic Order Quantity of each product can also be done instantaneously.

#### 3.8 Range

Highlighting a certain area and giving it a name would facilitate our efforts. Quite often, we would need to work on a certain range for instance B24:D50 (Cell\_B24 up to Cell\_D50). Referring to such an area of cells via a cell basis would make its usage harder. By simply identifying it as a cell\_range, referring to it in the VLookup and other statistical tool would be easy.

#### 3.9 Concatenation

Concatenation is an alternative to the situation of AND. The syntax is denoted as:

=Concatenate(A1,A2)

For instance, the rule

If having\_short\_hair AND (can\_lift\_150\_kilograms) then applicant\_is\_a\_man.

'can be represented as:

'having\_short\_hair" "can\_lift\_150\_kilograms"

Hence, we could rewrite it as:

If 'having\_short\_hair" "can\_lift\_150\_kilograms" then applicant\_is\_a\_man.

# 3.10 Cell and Sheet Protection

This technique is used when are about to distribute our work. By protecting the entire spreadsheet, except for a few items, we are enhancing the integrity of our application.

This is achieved by following the navigational sequence.

Main - Review - Protect

However we have to emphasize that before we enable the protection, we should not lock the areas of the spreadsheet where the end-user would want to enter data.

# 3.11 Cell Coloring

Another approach to enhancing the spreadsheet's utility is to color the cells. This is achived by the following navigational sequence:

Right Click On The Cell- Fill Color and Right Click - Font Color

A colored cell adds dimension to the cell. A green might require the end-user to key-in a response, while a red would mean the end-user should not touch that particular cell. In fact we can generate different colors based on the value of that cell by using conditional formatting.

# 4. Case Study: Micro-lending Enterprise

We present here a simple case for a micro-lending business. This spreadsheet is used to guide the lending agents in dealing with its clientele. For business confidentiality agreement, many of the rules were modified. However, the main essence of using it as an example has remained intact.

Rule1 If the Asset Base is more than \$1000 Then Extend Loan facility of \$100,000

Rule2 If the Asset Base is more than \$5000 Then Extend Loan Facility of \$500,000

Rule3 If the Asset Base is more than \$10,000 Then Extend Loan Facility of \$1,000,000.

Rule4 Asset Base is computed as the latest purchase price recorded \* (1-(currentYear – Latest purchase year)/100)

Rule5 If the Loan Facility is \$100,000 Then ask for a proof of continuity.

Rule6 If the Loan Facility is \$500,000 Then ask for a proof of business employing no less than 10 people.

Rule7 If the Loan Facility is \$1,000,000 Then ask for a proof of business employing no less than 50 people.

Rule8 If there is a proof of business Then approve the loan.

Rule9 If there is no proof of business Then ask for cash deposit of 50% of loan.

Rule10 If all requirements are appropriate Then Approve the Loan Otherwise Deny the Loan.

# 4.2 Key Spreadsheet Codes

We simply need three input numerical parameters, namely: Latest Purchase Price (LPP), Last Purchase Year(LPY) and the current year (NOW).

Cell\_F38 <- {space for entering Latest Purchase Price, LPP} Cell\_F40 <- {space for entering Latest Purchase Year, LPY} Cell\_F42 <- {space for entering the current year, NOW}

Cell\_F44 <- "=LPP \* (1 - (NOW - LPY)/100)" Cell\_F46 <- "=Vlookup(F44,LF,2,TRUE)" - where LF is a table consisting of AssetBase and LoanFacility.

LF	
ASSETBASE	LOANFACILITY
0	0
1000	100000
5000	500000
10000	1000000

Cell\_F48 <- "=VLOOKUP(F46,REQ,2,FALSE)" – where REQ is a table consisting of Loan Facility and Requirements.

REQ	
LOAN FAC	REQMNTS
100000	PROOF OF CONTINUITY
500000	PROOF OF BUS, EMP >= 10
1000000	PROOF OF BUS, EMP >= 50

Note that Rules 1, 2 and 3 are taken-cared of by table LF. Rules 5, 6 and 7 are taken-cared of by table REQ. Cell\_F44 contains computational rule.

Continuing, Cell\_F50 would require an entry of YES or NO regarding whether proof of business were submitted. Cell\_I50 <- "=IF(F50="YES","Y","N")"

Cell\_F52 <- "=IF(F50="YES","APPROVE LOAN","ASK FOR 50% DEPOSIT")" This captures Rule 8 and 9.

Cell\_I52 <- "=IF(I50="Y","A","X")"

Cell\_54 is a space to input whether 50% deposit were submitted.

This works with Rule 9. Cell\_I54 <- "=IF(F54="YES","Y","N")"

Cell\_F56 <- "=IF(I56="D","DENY LOAN","APPROVE LOAN")" This captures rule 10. Cell\_I56 <- "=IF(AND(I50="N",I52="X",I54="N"),"D","A")"

Cell\_F58 is an hyperlink leading towards the proper form.

# 5. Conclusion

This paper started with the motivation of being able to put together three ideas: chained business rules, tabular equivalence of business rules and specific spreadsheet techniques. These three ideas, when properly put together, can deliver a small and medium scale enterprises oriented Business Rules Management Systems (SME-Oriented BRMS) solution.

It has been noted that the Excel spreadsheet has

grown to be capable of supporting numerous types of applications. It has even been applied to mapping applications.[7] In this paper we showed Excel's viability as a tool in capturing and encoding rules into a spreadsheet solution. This paper listed eleven key constructs of Excel in delivering a usable Business Rules Management Systems. After having explored those eleven building blocks, a modified example (microlending) was shown to highlight its applicability.

We have to note that having encoded the business rules in Excel, the end-user did not feel threathen by the computer-based solution. As fear was absent, and thinking that the Excel is an ordinary tool, the process of improvement and refinement of the rules was rapid. Further, information on its improvement was spontaneous as everyone try to help correct the rules.

This validates the research result of Chanin Yoopetch[8] in stating that as far as knowledge acquisition in small firms is concerned, the role of social network and teamwork contributes significantly.

## **References:**

[1] M. Mircea and A. Andreescu, "Using Business Rules in Business Intelligence," *Journal of Applied Quantitative Methods*, Vol 4, No. 3 Fall 2009.

[2] Roger, Marcel and Lopez "Business Process Requirement Engineering," International Journal on Computer Science and Engineering, Vol 2, No 9, 2010, 2890-2899

[3] J. L. Carlo, K. Lyytinen and G. M. Rose "A Knowledge-Based Model of Radical Innovation in Small Software Firms", *MIS Quarterly*, Vol 36 No. 33 pp. 865-895/September 2012

[4] J. Taylor, Decision Management Systems – A Practical Guide to Using Business Rules and Predictive Analytics, IBM Press, 2012, ISBN-13: 978-0-13-288438-9.

[5] Jan Vanthienan, "The History of Modeling Decisions using Tables," http://www.brcommunity.com/, February 2, 2013

[6] Bill Jelen, *Microsoft Excel 2010*, Que Publishing, ISBN-13: 978-0-789-74308-4

[7] B. Whyte "On the Use of Spreadsheets such as Excel for Mapping," *Cartographica*, vol 46, issue 1, pp. 41-50

[8] C. Yoopetch "The Model of Knowledge Acquisition: The Study of SMEs in the Thai Hospitality Industry", World Academy of Science, Engineering and Technology, pp 351-35

\*Dr. Kang Mun Arturo G. Tan is an Assistant Professor at the Yanbu University College, Kingdom of Saudi Arabia. His research interests are in the area of Management Support Systems, Business Intelligence and Decision Support Systems. He primarily uses Excel, SPSS, and R to perform business analytics. He obtained his Doctor of Business Administration, Master of Business Administration and BS Electrical Engineering from the University of the Philippines, Diliman. He can be reached at arturo.tan@yuc.edu.sa

# Determination of Mathematical Model and Torque Estimation of s-EMG Signals based on Genetic Algorithm

<sup>1</sup>Ku Nurhanim, <sup>2</sup>I. Elamvazuthi, <sup>3</sup>P. Vasant, <sup>4</sup>S. Parasuraman

**Abstract**—This paper discusses the conversion of surface electromyography signals (s-EMG) to torque for robotic rehabilitation. Genetic algorithm (GA) has been applied as control algorithm for a number of selected mathematical models. s-EMG signals was treated as input to the mathematical model where pararameters of the mathematic model were optimized using GA. Hence, the estimated torque is considered as output data of mathematical model.

**Index Terms**— Electromyography, estimation torque,mathematical model,biomechanics human motion,muscle contraction,robot rehabilitation,feature extraction,genetic algorithm.

----- 🔶 ------

# **1** INTRODUCTION

In United States, about 795,000 people suffer from stroke which is the third ranking disease where one quarter of the patients are dependent on assisted activities for daily life [1]. Every year about 130,000 people in United Kingdom have stroke with 13,000 people under retirement age where over one third from 65% of the stroke survivors having disabilities of hand and upper limb. Muscular dystrophy, arthritis and regional pain syndrome are also major cause of disabilities of stroke patients [2].

Task oriented stroke rehabilitation promote the motor recovery and cerebral organisation of patient after stroke [3,4]. However the availability of such training session are limited in hospital and rehabilitation centre. There are constrained number of patients for rehabilitation program, number of facilities and required long duration time, and the ability of therapist to control and repeat the assistance for each training session [5]. Every year, with grossly higher demand for rehabilitation, the implementation of robotic rehabilitation may assist to reduce the dependency on therapist and provide a consistent training exercise. Robotic rehabilitation also provides intensive training without tiring and possible to enhance the therapy beyond the abilities of therapist [6].

The robotic rehabilitation utilizes the surface electromyography signals (s-EMG) as control input to estimate the magnitude and direction of torque for the robotic joint of exoskeleton. Researchers have done several works on mathematical model and black box model to estimate torque based on force produce of the muscle contraction. Biomechanical model has been use to convert s-EMG signals to torque to maintain the posture of arm [7]. Hill-Based model, also known as modelbased approach, acts as input, and the output depend on the muscle activation [8] level and musculoskeletal kinematics to generate torque [9]. Fleisher et al, used Hill-muscle model to predict the joint torques for controlling the leg of exoskeleton [10]. Menegaldo et al, estimate the torque of ankle by applying muscle model which summation each individual muscle and compare with torque measure [11]. Hill-based muscle model parameters matched the human ankle for torque and angle walking to control powered ankle prosthesis device [12].

Cavallaro et al, has done research on joint torques prediction for arm power exoskeleton application with and without GA optimization parameters [13].

Clancy et al, have used polynomial equation as a mathematical model for estimation joint torques in isometric and quasi-isometric contractions. Least square method has been used to to minimize the torque error [14]. S. Parasuraman et al, has used mathematical model with SA algorithm and GA optimization with least square method as cost/fitness function to get the best fit model of torque conversion [15]. Khalid et al, introduced new mathematical model using nonlinear regression to estimate joint torques. This new mathematical model has been compared with other mathematical model and neural network and proven to be the best fit model for online processing and real-time processing [16].

Neural-network also recognized as black box model does not require knowledge of physiology and joint dynamics for torque conversion. Neural network model was also used for prediction of joint moments [17], muscle force [18]. Neural network can be trained using different load and frequency to estimate the joint torques [19].Micheal E.Hahn, found that neural network model allow a greater accuracy joint torques prediction but as error goal set to lower, convergence is not achieved 100% [20]. The selection of number of hidden unit neural network model was not the main problem to allow accuracy of s-EMG to joint torques conversion [21]. This model was also applied for prediction of output mapping between s-EMG signals and joints angles and joint moments. The result from this study has demonstrated that is has good potential for biomechanical and simulation in human gait [22]. The processed data of s-EMG signals could be converted to estimation torque based on human movement. This estimation torque is important to control the motion of robots for rehabilitation. Mathematical model is used to convert the s-EMG signals to torque for estimation torque value. This estimation torque will be determined by optimizing the parameters of the mathematical model to the nearest value of actual torque. Artificial intelligence algorithms are used to optimize the internal parameters of the mathematical model to find the best fit mathematical model and minimize the error between estimated torque and actual torque. The aim of this paper is to establish mathematical models for s-EMG signals conversion to estimated torque using genetic algorithm.

# **2** THEORY AND TECHNIQUE

#### 2.1 Mathematical Model of EMG to Torque Conversion

The fitness function to evaluate the performance of each selected mathematical model using genetic algorithm is

$$SSE = \sum_{i=1}^{n} (\tau_{act} - \tau_{est(i)})^{2}$$
(1)

where,

*SSE* = sum square error as fitness function

 $\tau_{\rm act}$  = actual torque

 $\tau_{est(i)}$  = estimated torque as mathematical model

Mathematical model for estimated torque [15],[16]:

$$\tau_{est(1)} = x_1 . u_i + x_2 . u_i^{\frac{1}{2}}$$
(2)

$$\tau_{est(2)} = x_1 . u_i^{x_2} \tag{3}$$

$$\tau_{est(3)} = x_1 . u_i^{x_2} + x_3 . u_i^{x_4}$$
(4)

$$\tau_{est(4)} = x_1 + x_2 \sqrt{u_i}$$
(5)

$$\tau_{est(5)} = u_i^{x_1} \cdot e^{(x_2 - x_3 * u_i)}$$
(6)

where,

 $\tau_{est(1)}$  to  $\tau_{est(5)}$  = torque estimation mathematical model.

 $u_i = processed EMG data sample$ 

 $x_i$  = where, i=(1,2,3..) as random value parameter associated with selected model.

## 2.2 Regression analysis

Regression analysis accommodates the fitness function to minimize fitness values and obtain the optimum value of the parameter of selected mathematical model using genetic algorithm(GA).The steps to investigate the best fit mathematical model is identified as the following:

- i. Initialize population-start the population by random generation.
- Evaluation: The fitness function evaluates each chrosome. The lower fitness value indicates good solution, minimize sum square error.
- Selection: To form new population individual from previous population selected according to selection criteria. higher fitness level has higher presumption to be selected.
- iv. Genetic operator: crossover and mutation are applied to each selected individual to produce new generation.

v. Termination:

The process from Step i to iv are repeated until one of the condition as following are satisfied:

- A pre-determined number of generations or time has elapsed.
- A satisfactory solution has been achieved.
- No improvement in solution quality has taken place for a pre-determined number of generations.

Evaluate correlation between estimated torque and actual torque using coefficient of determination

$$R^{2} = \frac{\sum_{i=1}^{n} (\tau_{est(i)} - mean_{\tau_{act}})^{2}}{\sum_{i=1}^{n} (\tau_{act(i)} - \tau_{est(i)})^{2} + \sum_{i=1}^{n} (\tau_{est(i)} - mean_{\tau_{act}})^{2}}$$
(7)

where,

 $\tau_{\rm act}$  = actual torque

 $\tau_{\text{est}}$  = estimated torque

*mean*  $_{act}$  = mean of actual torque

## **3 EXPERIMENTAL METHOD**

The surface electromyography (s-EMG) signal acquisition was recorded using surface wireless-trigno sensor acquisition system from Delsys Inc. Boston, M.A, USA [23]. Two channel electrodes are placed on biceps brachii muscle with sampling frequency 1 kHz as shown in Fig.1. The experiments were carried on male healthy subjects. The experiment was conducted after obtaining approval from Ethics Committee of University and the subject. The movement task involved was elbow flexion with 90 degree.



Fig. 1: Delsys wireless sensor placed on bicep brachii muscles.

## 3.1 Estimated Torque s-EMG signal processing method

The EMG to torque conversion involved a few stages as the following:

#### A. Raw signal Acquisition

Amplitude of s-EMG signal was measured using surface electrode with built in amplifier circuit for acquiring the s-EMG signals.

#### B. Filtering

Filtering process will remove unwanted signal and leaving the information signal to be measured and recorded. For elimination of noise signal, the band pass filter are used (low pass and high pass) considered as the cut-off frequency of 20Hz-400Hz.

## C. Feature extraction:

The s-EMG signals were analysed by root mean square (RMS) based on square root calculation to obtain the s-EMG input data signal of muscle contraction.

### D. Muscle activation:

Detect the time duration of active region of muscle. This will produce post-processed signal which is the product of root mean square signal and muscle activation.

#### 3.2 Actual Torque modeling and preprocessing

#### A. Human tracking system

Modeling the elbow flexion movement of static and dynamic motion was recorded by using camera of Qualisys Tracking Markers System (QTM) [24]. Marker points were placed on the lower arm and hand as shown in Fig 2.



Fig. 2: Markers point for lower arm and hand

#### B. Modeling the biomechanical data

The static and dynamic motion data is modeling into body segment parameters using 3D-visual software of C-motion Inc. to get the signal of actual torque [25].

#### C. Processing the actual torque signal

The raw actual torques signal will be filtered to eliminate noise signal and rectified to obtain the actual torques data.

## **4** RESULT AND DISCUSSION

The results for each mathematical model for estimated torque using genetic algorithm (GA) for subject 1(S1) and subject 2(S2) are shown in Table 1.0.

Mathematical				
Model	Subjects	SSE	MSE	<b>R</b> <sup>2</sup>
$\tau_{\rm est(1)}$	S1	0.002056	3.11595E-05	0.828343
	S2	0.008582	0.000130035	0.823165
$\tau_{est(2)}$	S1	0.003208	4.86114E-05	0.818436
	S2	0.002696	4.08529E-05	0.833655
$\tau_{est(3)}$	S1	0.001622	2.45867E-05	0.830217
	S2	0.005659	8.57461E-05	0.818776
$\tau_{est(4)}$	S1	0.003121	3.85742E-05	0.841983
	S2	0.002735	4.14485E-05	0.828200
$\tau_{est(5)}$	S1	0.002909	4.40794E-05	0.862644
	S2	0.001843	2.79371E-05	0.865148

It can be seen that the sum square error (SSE) value of the fitness function decreased when the genetic algorithm met the optimum value parameters for each mathematical model. The unknown parameters were adjusted to ensure estimated torque became nearer to actual torque. Correlation between estimated torque and actual torque was found to be good when mean square error (MSE) decreased to small value of error and the coefficient of determination (R<sup>2</sup>) increased. From table 1.0, the highest coefficient of determination (R<sup>2</sup>) is 0.862644 for S1 and 0.865148 for S2 by mathemathical model  $\tau$  est(5).

Fig.3 to Fig.7, show the best correlated graph of the estimated torque that was applied to the mathematical model ( $\tau_{est(1)}$  to  $\tau_{est(5)}$ ) with actual torque for S1 and S2 based on table 1.0.



Fig 3: Estimated torque vs actual torque by  $\tau_{est(1)}$  for S1



Fig 4: Estimated torque vs actual torque by  $\tau_{est(2)}$  for S2



Fig 5: Estimated torque vs actual torque by  $\tau_{est(3)}$  for S1

Fig. 3 shows the estimated torque versus actual torque from S1 where the R<sup>2</sup> = 0.828343 is higher than S 2 with R<sup>2</sup>= 0.823165 for mathematical model  $\tau_{est(1)}$ . For mathematical model  $\tau_{est(2)}$ , S2 obtained R<sup>2</sup>=0.833655 as shown Fig.4, estimated torque versus actual torque compared to S1 with R<sup>2</sup> = 0.818436. The coefficient of determination of mathematical model  $\tau_{est(3)}$  for S1 are higher than S2, with R<sup>2</sup>=0.830217 compared to R<sup>2</sup> = 0.818776 as shown in Fig.5.



Fig 6: Estimated torque vs actual torque by  $\tau_{est(4)}$  for S1



Fig 7: Estimated torque vs actual torque by  $\tau_{est(5)}$  for S2

Fig.6 and Fig 7, show the estimated torque versus actual torque for mathematical model  $\tau_{est(4)}$  for S1 and mathematical model  $\tau_{est(5)}$  for S2. The correlation between estimated torque and actual torque of  $\tau_{est(4)}$  indicate that S1 has higher correlation R<sup>2</sup> = 0.841983 than S2 with R<sup>2</sup>= 0.828200. Subject 2 has higher correlation R<sup>2</sup>= 0.865148 than subject1 R<sup>2</sup>= 0.865148 for mathematical model  $\tau_{est(5)}$ .

Amongst all the mathematical models, mathematical model  $\tau_{est(5)}$  is the most suitable model for estimated torque for S1 and S2 which showed higher correlation compared to other models.

#### 5 CONCLUSIONS

This paper has discussed the analysis of torque estimation of mathematical model that can be applied to map EMG signal as an input and obtain the output of estimated torque. The actual torque is compared to estimated torque to verify if the estimated values are close to actual torque. It can be deduced that the best mathematical model for estimated torque is the fifth mathematical model ( $\tau$  <sub>est(5)</sub>). Estimated torque could be used to control the exoskeleton in various applications such as ergonomics, gait analysis, rehabilitation and also sports exercise.

#### REFERENCES

- "www.strokeassociation.org," American Stroke Association [ONLINE] 2011, download on 20<sup>th</sup> January 2013.
- [2] Ioannis Sarakoglou, Sophia Kousidou, Nikolaos G. Tsagarakis, Darwin G. Caldwell, "Exoskeleton-based exercisers for the disabilities of the upper arm and hand,"Ch.27 Rehabilitation Robotics,Itech Education and Publishing, pp.449-521,Aug 2007.
- [3] J.H Caraugh and J.J Summers, "Neural plasticity and bilateral movements: A rehabilitation approach for chronic stroke," Prog.Neurobiol., vol.78, pp. 309-320,2005.
- [4] R.J. Nudo, "Functional and structural plasticity in motor cortex: Implications for stroke recover," Phys.Med.Rehabil.Clin.N.America, vol. 14, pp.557-576,2003.
- [5] D.Erol and N.Sarkar," Coordinated control of assistive robotics devices for activities of daily living tasks," IEEE Trans. On Neural System and Rehab. Eng., vol.16, pp.278-285, 2008.
- [6] B.Dellon and Y. Matsuka," Presthetics, Exoskeletons and rehabilita-

tion, [Guard challenges of robotics]," IEEE Robotics \*&Automation Magazine, pp.30-34, 2007.

- [7] Duk Shin, Jaehyo Kim & Yasuhara Koike, " A Myokinetic arm model for estimating joint torque and stiffness from EMG signals during maintain posture," J. Neurophys., vol.101, pp.387-401, 2009.
- [8] C.De Luca, "The Use of Surface Electromyography in Biomechanics," J. App.Mechanics, vol.13, pp.877-878, 1997.
- [9] K. D. Pham and E. L. Morin, "Dynamic surface emg/force relation human elbow flexion," In Proc. IEEE Int. Conf. of the Eng. In Medicine and Biology Society, pp. 1203–1204, 1995.
- [10] Fleischer C, Wege.A. Kondak K, Hommel G, "Application of EMG signals for exoskeleton robot," In Biomed. Tech, vol.51, pp.314-319 Berlin, 2006.
- [11] L.L. Menegaldo, L.F. De Oliveirra, "Effect of muscle model parameter scaling for isometric plantar flexion torque prediction," J. of Biomechanics, vol.42, pp. 2597-2601, 2009.
- [12] M. F. Eilenberg, H. Geyer, H. Herr, "Control of a power ankle-foot prosthesis based on a neuromuscular model," IEEE Transaction Neural Sys. & Rehab. Eng., vol.18, pp.164-173, 2010.
- [13] E.E. Cavallaro, J. Rosen, Joel C. Perry, Stephen Burns, "Real-time myoprocessors for a neural controlled power exoskeleton arm," IEEE Transaction on Biomed.Eng., vol.53,pp.2387 - 2396, 2006.
- [14] E.Clancy, N.Hogan," Estimation of torque from the surface EMG." IEEE Annual Int. Conf. Of The Eng. In Medicine & Bio. Soc, vol.13, pp. 877- 878, 1991.
- [15] Khalil Ullah, J.H Kim, "A mathematical model for mapping EMG signal to joint torque for the human elbow joint using nonlinear regression," Proc. 4th Int. Conf.On Autonomous Robots & Agents Welington, pp.103-108, Feb. 10-12, 2009.
- [16] S.parasuraman, Arif W.O, V.L Jauw, "AI Techniques to estimate muscle force and joint torque for upperextremity," J.Key Eng.Material, vol.467-469, pp.788-793, 2011.
- [17] L.Wang, T.S Buchanan," Prediction of joint moment using of neural network of muscle activation from EMG signal," IEEE Transaction of neural systems and rehabilitations Eng., vol.10,pp.30-34, 2002.
- [18] M.M Liu, W.Herzog, H.H.C.M Savelberg, "Dynamic muscle predictions from EMG: artificial neural network approach." J. of Electromyography & Kinesiology, vol.9, pp.391-400, 1999.
- [19] R.Song, K.Y Tong, "Using recurrent artificial neural network model to estimate voluntary torque in dynamic situations," J. of Medical & Bio.Eng. & Comput, vol.43, pp.473-480, 2005.
- [20] Micheal E.Hahn, "Feasibility of estimating isokinetic knee torque using a neural network model," J. of Biomechanics, vol.40, pp.1107 -1114, 2007.
- [21] J.J LUH, G.C Ching, C,K, Ching, J.S. Lai, T,S Kuo, "Isokinetic elbow joint torques estimation from surface EMG and joint kinetic data: using an artificial neural network," J. of Electromyography Kinesiology, vol.9, pp.173-183,1999.
- [22] F.Sepulveda, D.M. Wells, C.L.Vaughan, "A neural network representation of electromyography and joints dynamics in human gait," J. of Biomechanics, vol.26, pp.101-109, 1993.

- [23] Delsys Trigno Wireless Sensor (2012), retrieved from http://www. Delsys.com/products/trignowireless.html on 20<sup>th</sup> Jan 2013.
- [24] Qualisys motion capture system retrived from http:// www.qualisys. com/products/hardware/oqus on 20<sup>th</sup> Jan 2013.
- [25] C-Motion Visual 3D software retrived from https://www.cmotion.com/prosucts/visual3d on 20th Jan 2013.

<sup>&</sup>lt;sup>1</sup> Ku Nurhanim is currently pursuing masters degree program in electrical and electronics, University Technology Petronas, Malaysia, E-mail: <u>hanim.rahim7@gmail.com</u>

<sup>&</sup>lt;sup>2</sup>I.Elamvazuthi, Dept. of Electrical & Electronics, University Tenology Petronas.

<sup>&</sup>lt;sup>3</sup>P.Vasant,Dept. of Fundamental & Applied Science,University Technology Petronas.

<sup>&</sup>lt;sup>4</sup> S.Parasuraman, School of Engineering, University of Monash, Malaysia.

<sup>&</sup>lt;sup>2</sup>Email of Corresponding Author: irraivan\_elamvazuthi@petronas.com.my

# Adoption of Iris Recognition System in Financial Sectors: A Case of Saudi Arabia Point of View

Hanan Abdullah Fatani, Nour Elhouda Tabet

Abstract—The process of creating and protecting of a unique and safe identity for citizens or customers is one of the most difficult and major challenges for organizations whether public or private sectors because of the high rate of fraud, identity theft and Plagiarism personality for illegal purposes [1]. Several studies conducted around the world have proven that identity theft and fraud is the most serious crime within the banking sector, this is the result of the new technology revolution [2],[3],[4]. Biometrics recognition proved it credibility, reliability, accuracy and validity, also, it had been implemented within many application at many countries towards maximum level of privacy and security in order of maximize accuracy identification and verification [5]. The motivation of this study to be a good basis to stimulate the decision makers to apply Iris technology in both private and public sectors towards more secure and safe environment among Saudi citizens.

Index Terms— unique, identity theft, fraud, plagiarism, biometrics, recognition, iris technology.

# **1** INTRODUCTION

When Almighty Allah (SWT) created the humans, He created them with some unique and distinctive features for identification. Those features are used to distinguish them from each other, and, as a form of identity to verify whether their physiological or behavioral and every feature of these properties are unique and enduring over time. Most organization facing difficult challenges in order to keep the business safe and secure by creating and protecting of a unique and safe identity for citizens or customers from any frauds or identity thefts [6],[7].

Several strategies and technology have been used to ensure the security and safety of the customers' [8],[9]. Some sectors used biometrics scanning technology of human body parts in order to strengthen security and curb counterfeiting of identities through the properties of behavior and physiological. Undoubtedly, the physiological characteristics considered the best and the trusted security methods in the process of identification, because, Biometrics cannot be forgotten or changed when use by the individual [1],[10],[11].

It is a technology that has a large proportion of credibility, usability and reliability. Physiological characteristic included all shapes of the person body, such as, DNA, fingerprint, iris recognition and palm print and other; meanwhile, behavioral characteristic is related to the person behavior like voice, signature and typing behavior "PIN number or password".

Traditionally every biometric device has its own way of utilization like camera or scanner in order to capture the images, and it is vary depending on the purpose and the application, it can be used generally in many areas such as, banking system, immigration system , election system and governmental system so on [8],[9].

In the past few year, there were new solutions became surpasses of all identification and verification concepts. Furthermore, it can meet all the challenges that correlating with customers' verification. Recent studies proved that Iris recognition technology play an important role towards applications that related to identity verification and security issues [12].Regrettably, it has been a little effort to investigate the intention to use Iris recognition technology for conducting banking transactions, since; Iris recognition system is quite new technology to biometric technology. It is crucial to investigate the external factors that influence the customer's acceptance to use Iris technology so that the new technology can be implemented and benefits for both banks customers and bank sectors.

#### 1.1 Costumers' identity theft

For most financial sector the main priorities is to insure safety and security of the customer's identification from frauds or identity thefts. With proximity of 27 million of Saudi Arabia population, a statistic in 2009 showed that more than 89% of Saudis working in public and private sector has bank loans [13]. Shocking fact yet it means that near of 90% of Saudi working citizen's deal with banks in Saudi Arabia.

All Banks must conform to the government- Saudi Arabian Monetary Agency (SAMA) - and application its regulation and recommendations [14].

### 1.2 ATMs pirating

Previous studies showed that, the innovation of the ATMs machines did not increase the number of customers using ATMs but also increased the fraudulent practices The studies recommended many solutions for customers in order to reduce ATMs fraud in the country one of these solutions was using biometrics technology [15],[16],[17]. Ensuring the safety and security of customer identity is a substantial nightmare to all banks and they have to prevent any frauds, swindling, or identity thefts as possible as can [18].

Hanan A. Fatani is currently pursuing Ph.D degree program in information & communication technology in International Islamic University, Malaysia, Her publications have appeared in conferences such as IEEE. Email: hanan9fa@gmail.com

Nour Elhouda Tabet is currently pursuing master's degree program in information & communication technology in International Islamic University, Malaysia. E-mail: oo.nour.oo@gmail.com

# **2 BIOMETRICS BACKGROUND**

# **2.1 Biometrics Technologies**

It would be the perfect implementation for many sectors, such as, healthcare industry, educational sector, financial sector and governmental sector. Biometrics comes in many forms and each form needs its own special portions of hardware and software, from the very simple rigorous like fingerprint to the highly rigorous like iris and retina [19]. Despite the many benefits due to the use of biometrics technology including accuracy, reliability, quality of performance and ease of use, there are some disadvantages of privacy and culture and user's anxiety issues have to be taken into consideration [20], [21].

# 2.2 Iris recognition solution

In 1936 the Ophthalmologist Frank Burch proposed that the iris of the eye can be used to identify and verify the person. By 1985 the Ophthalmologists Leonard Flom and Aran Safir were awarded for their concept that no Iris of two humans is similar. In 1994, John Dougman was awarded for his "algorithms" for the iris. By the year 1995, the Iris prototype was completed and tested for implementation [22],[23].

The Iris is considered the best and most appropriate techniques that can be used in the process of identification and to ensure the correctness and accuracy of the data given by the customer of what the Iris holds of credibility, high-speed data extraction, ease of use of technical equipment in the process of scanning, small storage space in the system and the most important reasons that the Iris cannot be repeated between people, even if they were twins, remarkably, the two irises of the same person are different [24],[25],[26],[27].

Generally, the human iris as shows in figure 1, is the only internal organ that will be remains in a stable form from the first day of a person life until death, also, it appears clear and well protected from all the factors of life time. It has been a complex and a high stable structure over a person's life and this is what makes them unique from one person to another. Nor the use of glasses or contact lenses has a little effect on the clarification of the iris and thus it does not conflict with the recognition technology [28].



Fig. 1. Image of a human Iris showed the structure of the Iris [29].

Figure 2 shows a real life case issued in 2002 by National Geographic channel proved how a girl photographed in an Afghanistan in 1984 was found 18 years later and identified by her iris. In another word, that is why iris considered the best and most appropriate techniques that can be used in identification.

# 2.3 Limitations



The nature of Iris technology makes adopting it limited by certain measures. There are a set of factors that are defined as the main factors influencing the adoption of new technologies. These factors are generally categorized by factors influencing users, developers, and technology providers [31].

On the users' side, there are two users that the system influences; the customers and the staff. Both users should have a level of acceptance to the technology in order to ensure the success of the adoption. The main factors influencing user acceptance are:

- Desiring the change:

This factor is derived by the human nature of desiring comfort. Unless users realize the importance of change and the benefit of adopting the new technology, else it will be difficult to adapt Iris technology into the market.

- Self-disciplined and a high frustration tolerance:

Even after technology is adopted, the user should have a level of discipline in order to maintain the changes, and tolerate the difficulties that might occur after the technology is adapted.

Developers on the other hand play an important role of allowing the adoption of the Iris technology. If developers were not qualified with the required skills to build an Iris-based system that aligns with the firms' regulations and strategies, there will be a barrier that limits the adoption. Furthermore, the firm will need to outsource in order to reach the desired system specifications. This will lead to other concerns such as increasing development cost.

On the other hand, technology providers also affect the adoption of Iris into the systems. In the case of financial sector, banks and government are the core service providers both should have a high level of technology understanding, and awareness of the benefits that it will bring to the financial sector, including increasing security level.

Moving from the main players of Iris-based systems, financial barrier could form a level of limitation to the adoption of the system. The financial sector should design a solid financial plan that does not only include development costs, but as well and maintenance costs. In order to ensure continues operation of the Iris system, these limitations should be studies and taken in consideration. All the limitations could be overcome by having a strong plan and strategies that takes all the different aspects into consideration.

# 2.4 Expectations

The main expected outcome of adopting Iris technology into

Fig. 2. Afghan girl was identified 18 year later by her eye [30].

the financial sector is increasing the level of security, minimizing the cost of failure, preventing threats, and overcoming the current system vulnerabilities. These advantages can be seen in the sectors that have adopted such as the Cairo Amman Bank in Jordan. It is reported in The New York Times Magazine, that in 2010, Cairo Amman Bank in Jordan has won awards for its outstanding contribution to financial services in the Middle East and in recognition as the best bank in Jordan. However, the advantages can be seen, and the expectation could be based on the experience of financial sectors that already developed the Iris technology into their system.

#### **3** RELATED WORK

The rapid evolution of the technology makes many ways to the massive use of it; unfortunately it can be misused for illegal process, such as, swindle or identity theft. In 2011 a study conducted in Nigeria by Adeoti illustrates that the growth of ATMs in Nigerian banks has not only increased the number of customer using ATMs, it has also increased the tendency of fraudulent practices by the fraud perpetrators. Therefore, the study recommended many solutions for customers in order to reduce ATMs fraud in the country one of these solutions using biometrics technology [15], [32].

In 2008 Cairo Amman Bank, Jordan had become the pioneer in the Arab world to implemented Iris system among their branches in Jordan and Palestine, which assist the bank to eliminated ATMs fraud and ID thefts. The system improved the customer service in banks branches and was ease of use by customers and tellers, also, help to protect the customer identification through a strong and secure authentication procedure, enhance customer confidence and security and reduce the overcoats of cards and pins [16].

In the fall of 2003 at the nationwide building society in Wiltshire, Swindon they implemented iris system to eliminate the need of a card for ATMs, and to reduce the effect of credit card theft [17].

#### **3** CONCLUSION

The purpose of this study is to prove the important needs of adopting iris recognition technology among all sector, especially financial sector towards conducting banking transactions in Saudi Arabia, by enhancing the safety and security of our identification. The researcher wishes to confirm the important of adopting iris recognition technology in banks for customers' verification to reduce frauds and improve security to individual identification for better environment for Saudi citizens.

Iris recognition technology can be used to enhance business processes since it proved it credibility, reliability, highly accurate performance, stability, ease to use and the comparisons are made in few seconds so on [33]. The technology method had been already used in identification and authentication process in many banks around the global like Jordan and Swindon [16], [17].

## REFERENCES

 Ernst and Young, "A more secure future, Enabling security through biometrics, http://www.ey.com/Publication/vwLUAssets/A\_more\_secure\_future\_

-\_enabling\_security\_through\_biometrics/\$FILE/EY\_A\_more\_secure\_ future.pdf. 2009.

- [1] J. Lewis, "Biometrics for secure Identity Verification: Trends and Developments"," 2002.
- [2] A. Cavoukian and A. Stoianov, "Biometric Encryption: A Positive-Sum Technology that Achieves Strong Authentication, Security AND Privacy," Information and Privacy Commissioner of Ontario, Canada, 2007.
- [3] L. Coventry, "Usable Biometrics," 2005.
- [4] J. Daugman and I. Malhas, "Iris recognition border-crossing system in the UAE," 2004.
- [5] R. H. Abiyev and K. I. Kilic, "Robust Feature Extraction and Iris Recognition for Biometric Personal Identification," 1998.
- [6] M. Breward, M. Head and K. Hassanein, "Consumer Acceptance of Biometrics for Identity Verification in Financial Transactions," http://is2.lse.ac.uk/asp/aspecis/20090083.pdf. 2012
- [7] E. Spinella, "Biometric Scanning Technologies: Finger, Facial and Retinal Scanning," SANS Institute, 2003.
- [8] P. Li and R. Zhang, "The Evolution of Biometrics," in *IEEE*, Guiyang, China, 2012.
- [9] P. Backman and C. Kennedy, "Biometric Identification and Privacy Concerns: A Canadian Perspective," Aird & Berlis LLP.
- [10] M. Dunker, "Don't link: Iris recognition for biometric identification," SANS Institute, 2003.
- [12] S. H. Moi, N. Abdul Rahim and S. Ibrahim, "Iris Biometric Cryptography for Identity Document," in *International Conference of Soft Computing and Pattern Recognition*, 2009.
- [13] Central Department of Statistics and Information, Consumer And Credit Card Loans, http://www.cdsi.gov.sa/english/index.php?option=com\_docman&Ite mid=158, 2012.
- [14] K. Hawsey, "Doing business in Saudi Arabia," Saudi Arabia, 2010.
- [15] J. O. Adeoti, "Automated Teller Machine (ATM) Frauds in Nigeria: The Way Out," 2011.
- [16] S. Ken, "IrisGuard Makes Inroads, from Mideast Banks to U.S. Prisons," http://www.bloomberg.com/news/2011-05-04/irisguard-makesinroad s-from-mideast-banks-to-u-s-prisons.html. 2011.
- [17] N. Hawkes, "Machines will pay up in blink of an eye," Swindon, 2003.
- [18] J. Nyasulu and T. Fomene, "Report on the Literature Study of Iris Biometric Recognition," Sweden, 2008.
- [19] J. Napua, "Growth of Biometirc Technology in Self-Service Situations," in FUJITSU Sci, USA, 2011.
- [20] M. El-Abed, R. Giot, B. Hemery and C. Rosenberger, "A study of users' acceptance and satisfaction of biometric Systems," in *IEEE*, France, 2010.
- [21] Y. Hussien, S. B. Sarapol, H. Singh and A. H. MatNoor, "Biometrics," Malaysia, 2009.
- [22] L. Flom and A. Safir, "Iris Recognition," 1994.
- [23] R. P. Wildes, "Iris Recognition : An Emerging Biometric Technology," 1997.
- [24] J. Daugman, "The importance of being random: statistical principles of iris recognition," *Pattern Recognition*, 2003.
- [25] M. Pereira and A. Veiga, "A Method For Improving The Reliability Of An Iris Recognition System," 2005.
- [26] K. S. Kyaw, "Iris Recognition System using statistical features for Biometric Identification," in *International Conference on Electronic Computer Technology*, 2009.

- [27] X. Liu, K. Bowyer and P. Flynn, "Experimental Evaluation of Iris Recognition," in *Computer Vision and Pattern Recognition*, 2005.
- [28] P. Khaw, "Iris Recognition Technology for Improved Authentication," SANS Institute, 2002.
- [29] Wikipedia, "iris recognition, http://en.wikipedia.org/wiki/Iris\_recognition. 2012.
- [30] C. Newman, "A Life Revealed," April 2002. [Online]. Available: http://ngm.nationalgeographic.com/2002/04/afghan-girl/index-text?fs =photography.nationalgeographic.com. 2013.
- [31] A. Kintsch, R. DePaula, "A Framework for the Adoption of Assistive Technology"
- [32] S. Batunanggar, "Comparison of Problem Bank Identification, Intervention and Resolution In The Seacen Countries," 2008.
- [33] C. D. Patel, S. Trivedi and S. Patel, "Biometrics in IRIS Technology: A Survey," in International Journal of Scientific and Research Publication (IJSRP), 2012.

# Mobile learning in the sphere of physical problems solving using the semantic structures method

Tatyana Gnitetskaya, Elena Karnaukhova, Elena Ivanova

**Abstract**— The article proposes an original way to train skills of solving problems in physics using the method of semantic structures developed by T.N. Gnitetskaya withing the framework of the informational model of intra-subject connections. The analysis of a physical task is performed in the course of building the structure, which semantic elements are physical concepts, values, laws located on subordinate levels and logically related to each other by internal connections. Using this method allows three levels of correction and self-monitoring of students' work in mobile learning, the only one of which is full-time. The proposed approach to mobile learning in the sphere of solving physical problems promotes improvement of analytical skills and students' intellectual level.

Index Terms— Mobile learning, semantic structures method, information model of intra-subject connections, distant learning, solution of problems in physics, analytical thinking, mobile technologies.

----- ♦ ------

# **1** INTRODUCTION

TODAY'S information technologies have become an integral part of learning process. Recently appeared wireless technologies have reflected in education as mobile technologies that are typically implemented in alternative courses. Of course, it is easy for students to use wireless devices such as Ipads and Iphones, so it started active discussions on their use for educational purposes.

At first glance, the possibilities of mobile and distant learning are not so different. Indeed, both cases suppose that students' independent activity is possible at any convenient place and time involving on-line educational resources and background information. However, it is important that mobile learning gives a possibility to correct students' independent activity. It involves text-messages - correspondence in the "question-answer" mode, and organization of students' free access to self-testing. Certainly, the advantages of mobile devices are quality and quick playback of graphic, audio and video files.

Despite the positive aspects of mobile learning, there are several disadvantages, the main of which are: 1) perception difficulties of the content of lengthy texts and volume formulas using mobile devices with small screens; 2) mobile devices fragility that becomes a serious obstacle when combining selflearning and everyday activities; 3) mobile devices dependence on power supply devices is significantly perceived when playing graphic files, many of which, such as video files, have megabit sizes.

If the second two shortcomings identified above can be removed with the assistance of the relevant technical equipment, the first drawback can be leveled only by involving another method of educational material representation.

As it is shown by comparative study conducted in Russia, the U.S. and Japan, the most accessible form of educational material representation for students from the countries listed above is the content of the educational material structured by portions, each of which is represented as a diagram [1]. We should admit that the scheme is fully reproduced on the screen of a mobile device, what is an additional argument in favor of this method.

# 2. THE STUDY – THE PROBLEMS IN THE FORM OF STRUCTURES

In previous studies we have developed the semantic structures method based on informational models of an intra-and inter-subject connnections [2]. This method is applied to the process of independent activity while solving physical problems.

Solving physical problems is known to cause difficulties for students that study physics. It has logic, because not every student has analytical skills that are used when solving physical problems, and no study of problem solving techniques can help here. Physical problems solving, in fact, is a creative process of hooking up disparate facts stated in an experimental or theoretical problem. There is an idle opinion that distance or wireless technologies oriented to students' individual work are easy to be applied in liberal arts, because they require less explanation than natural sciences that may not be comprehended by self-studying. This is not quite true. It is known fact that the content of the training material in natural sciences, such as physics, is built step by step according to the principle of causality. Fundamental physical laws reflect a cause-andeffect relation, where the cause expressed by a physical value multiplied by the coefficient of proportionality is placed to the right side of an equation and the physical quantity associated with a conclusion - to the left. Therefore, the content of physics, as a rule, is built by both deductive (from general to particular) and inductive (from particular to general) methods. It is important that the content has intra-subject connections. From this point of view, stating a portion of the physical content in a physical problem with the help of semantic structure of intra-subject connections doesn't contradict to the principle of the unity of theoretical and experimental skills and abilities to solve physical problems. The performability of this principle can be measured by two criteria [2]:

1) criterion of skills well-formedness and ability to solve physical problems by experiment. This criterion represents an evaluation of knowledge level of measurement technology and instruments, measurement techniques, and skills of using
measurement technology and instruments and ability to choose a physically grounded method of measuring and conduct physical experiments;

2) criterion of skills well-formedness and ability to solve physical problems in theory. This criterion represents an evaluation of skills and ability to analyze in theory physical phenomena, select the basic physical laws, to use mathematical methods of solving physical problems.

In the context of the second criterion there is an example of solving a theoretical physical problem, where the analysis of the condition is performed in the process of building a deductive structure, and the problem is solved by building an inductive structure. We chose a rather simple problem from I.E. Irodov's collection № 3.202 "Problems in General Physics" from "Electrodynamics" section. These problems are given to second-year students claiming for the physics major degree [3]. Condition of the problem: there is a metal plate placed in a parallel way between a plate capacitor's plates, the thickness of the plate is  $\eta$ =0,60 of the distance between the capacitor's plates. The capacitor's capacity in the absence of the plate is C=20 nF. The capacitor is connected to a dc source U=100V. The plate was slowly removed from the capacitor. Find: a) capacitor's energy gain; b) mechanical work expended in extracting the plate. As usual, problem solution is carried out in two stages: Stage 1 - analysis of the problem's condition; Stage 2 - direct solution of the physical problem.

The first step is the analysis of the problem's conditions, when a deductive structure of the conditions is being built. All the concepts involved in the problem's condition are written out and numbered, the complex concepts are divided into their simple components. Then, the written out concepts are sequentially drawn up in the form of a semantic structure (Figure 1). While the concepts involved in the problem's condition are drawn up on the first (the highest) level of the structure conditions, the second (the lower) level is filled up with the concepts of the second order, which belong to the content of the concepts of the first row. And so on, until the content of every concept involved in the problem's condition is divided into elements.



Fig.1. Semantic structure of the condition of the problem  $N\!\!\!\!0$  3.202

Agreed notations: 1 – capacitor's plates; 2 - electric field; 3 – conductor; 4 – charge; 5 - plate 1; 6 - plate 2; 7 - plate capacitor; 8 – capacitor; 9 - flat plates; 10 – capacitor's plate 1; 11 - capacitor's plate 2; 12 – dielectric; 13 - dielectric conductivity; 14 - distance between the plates; 15 - electric field inside the dielectric; 16 - metal plate; 17 – plate; 18 -  $\eta$  = 0,60 of the distance

between the plates; 19 - capacitor  $C_1$ ; 20 - total capacitor's capacity; 21 – voltage; 22 – plate capacitor C capacity; 23 – vacuum's dielectric conductivity; 24 - area of the plate; 25 - distance between the plates; 26 - capacitor  $C_2$ ; 27 – IP device.

In our example the content of the first concept of capacitor's plates includes the following concepts: plate 1, plate 2, conductor, charge, electric field. A list of the components of the concept of capacitor's plates may be increased, for example, you can expand the concept of the electric field, but in the framework of this task no further specification is required.

Once the condition structure is built and analysis is finished we may proceed to the second stage, namely, to the solution of the problem. The solution of the problem is also performed with the help of the structure (Figure 2) that is built downright. At the top of the structure there is the required in problem physical value whose connection with the other values is determined by the corresponding relation. The correlating physical values take the next row below and so on, until a logical chain between the unknown value and the data in the problem is built. In general, this structure represents a way of problem solution.



Fig.2. Semantic structure of problem solution № 3.202

Agreed notations: 1 – work; 2 - energy difference; 3 - energy  $W_1$ ; 4 - energy  $W_2$ ; 5 – voltage, 6 - capacitor  $C_1$  electrocapacity; 7 – capacitor  $C_2$  electrocapacity; 8 – dielectric conductivity; 9 - vacuum's dielectric conductivity; 10 – area; 11 – distance.

The purpose of this method of solving physics problems is to teach students to determine by themselves internal connections of physical values, concepts and laws that promote formation of critical thinking skills and skills of analyzing physical situations.

The undeniable advantage of the method is the possibility to control and correct student's independent activity in mobile learning, the organization of which is a serious problem when preforming distance tasks.

Assignment for independent work is sent to students after being prepared by a teacher. It contains: 1) the condition of the problem, 2) a list of the concepts included in the content of the problem, 3) an unfilled structure of the condition, 4) an unfilled structure of the solution. The first three are available at the website for independent work and students can download the assignment to their mobile devices at any time. In this case one should analyze the condition of the problem by filling in the structure blanks with physical concepts or values chosen from the list. For the majority of students this stage will be enough to build the structure of the solution by themselves and thus to solve the problem.

Students who didn't manage to solve the problem may compare their condition structure with the correct one on the website - so that is how the first level of students' selfcorrection is carried out.

If there are any difficulties after this stage, the student may request the structure of the problem solution with blank cells from a teacher. That is the second stage of students' selfcorrection.

Finally, the third stage of students' control and selfcorrection is performed during the defence of the problems in the classroom. The defence requires the structure of solving the problem drawn on a paper and the solution itself with the necessary mathematical transformations. If necessary, a teacher may ask for the structure analysis of the problem's conditions, but, as a rule, it is not required.

Hence, the proposed method is implemented in three stages of correction and knowledge control, two of which are carried out distantly without teacher's consultation, and this organization is preferable.

The small size of the image of the structure makes it possible to work with it using the limited screen of a mobile device, and this fact removes one of the main drawbacks of mobile devices usage.

The method allows to focus on the student's individual intelligence level, it excludes algorithmization in performing distance work. In circumstances where mobile learning is limited it promotes development of student's creative approach to problem solving and skill of taking non-standard decisions. What is very important today to organize while preparing students studying natural sciences.

In the age of searching for new types of fuel, new technologies of environment refinement and other strategic problems, we face the need of new specialists with a high knowledge level in natural sciences who will be able to analyze the problem, set goals and develop an optimal strategy for achieving it. So any new directions in natural sciences including mobile technology are possible, but they must match the current needs of the society and be science-based.

# REFERENCES

- [1] Gnitetskaya Tatyana, Gnitetskiy Pavel V. FENU, Russia;Rodriguez Marizella, Rodriguez Joe A., SFASU, USA Information Technologies and Students' Computer Skills In C. Crawford et al. (Eds.), Proceedings of Society for Information Technology and Teacher Education International Conference 2006 (pp. 1955-1962). Chesapeake, VA: AACE.
- [2] Afremov, L.L, Gnitetskaya, T.N. The theory of Inter- and Intra-subject connections. Vladivostok: Far Eastern University Press, 2005. -176.
- [3] Gnitetskaya, T.N. Modern educational technologies. -Vladivostok: Far Eastern University Press, 2004. - 256.
- [4] Irodov, I.E. Problems in General Physics. M.: Nauka, 1988.-416 p. (PrePrint)
- [5] Theng, Y.L. (2009). Mobile Learning for Tertiary Students: An Exploratory Study of Acceptance of Use. In G. Siemens & C. Fulford (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009 (pp. 937-946). Chesapeake, VA: AACE.
- [6] Chen, T. (2010). A Case Study of the Development and Diffusion of Mobile

Learning. In D. Gibson & B. Dodge (Eds.), Proceedings of Society for Information Technology & Teacher Education International Conference 2010 (pp. 2184-2189). Chesapeake, VA: AACE.

[7] MacLean, G. & Elwood, J. (2010). Mobile Phone Usage and Perceptions in Japanese University EFL Classrooms. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2010 (pp. 1774-1783). Chesapeake, VA: AACE.

# The effect of the seeding power on RSOA-based colorless ONU performance in FTTH-PON

Mohammad Syuhaimi Ab-Rahman, Farhat M. Shaltami

Abstract – In this paper, the limitations of the RSOA-based colorless ONU in FTTH-PON are investigated and observed. First of all, the architecture of the colorless receiver used is described. Then, the relationship between the input power and the output power of RSOA is explained, while the RSOA is an important element in colorless schemes as a light source. After that, this paper will explain the effects of the relationship between input power and output power of RSOA on the upstream performance. It is clear from the simulation results that the seeding power must be high to saturate the RSOA and to make upstream performance better.

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

Index Terms— colorless ONU, Fiber to the home, passive optical network, semiconductor optical amplifier.

# **1** INTRODUCTION

During the last three decades, colorless optical network units (ONU) have been developed and studied. Optical network unit (ONU) does not depend on a particular wavelength in the network. This property gives the network flexibility in operation and maintenance. Besides flexibility, colorless is a cost solution, hence every ONU has the same equipments and mass production is possible. The mass production decreases the cost of the unit which makes it cost effective. There are many studies on colorless schemes to make it more efficient [1].

In the most common architecture in colorless technique, the OLT sends the downstream wavelength, in the same time sends the upstream wavelength supply to the ONU, which called seeding light. The ONU receives these wavelengths and divides them by WDM to demodulate the downstream data and modulate the upstream data. In the ONU side, there are two stages which the seeding light goes through: (i) amplification and (ii) modulation. Reflective semiconductor optical amplifiers (RSOA) can be the amplifier, which can be considered as the light source of the ONU, and the modulator is electroabsorption modulator (EAM) [2-3].

In this paper, we study the effect of the seeding light power on the upstream performance. The rest of the paper will be organized as follows: The second section in this paper will investigate and report some characteristics of RSOA, while the RSOA is an important element in colorless schemes as a light source. After that, this paper explains the effects of these characteristics of RSOA on the upstream performance

# 2 COLORLESS FTTH-PON

Usually, RSOA is also used as a modulator. RSOA can transmit up to 2.5 Gbps, This data rate is suitable for 1Gbps channels. However, 10G EPON requires a higher bit rate, which RSOA

(This information is optional; change it according to your need.)

cannot transmit [1]. Therefore, an external modulator is added to the upstream transmitter, to fulfill the requirements of IEEE802.3av standard. Figure 1. shows the Tree topology of EPON.



Figure 1. EPON Tree topology

When this technique is used in power-splitter-based-FTTH-PON, the upstream will be very dependent on the downstream situation, because the seeding light will be divided in term of power, which makes the received seeding light to the RSOA too low. In this case, RSOA will not be saturated. This will cause two main issues; the first is the increasing of the noise power generated by the RSOA, which makes the OSNR very low.



Figure 2 Colorless ONU

The second issue is the dependency of the output power of the RSOA on the downstream status. The goal of this work is to these issues and study FTTH-PON performance (BER) when

Author name is currently pursuing masters degree program in electric power engineering in University, Country, PH-01123456789. E-mail: author\_name@mail.com

Co-Author name is currently pursuing masters degree program in electric power engineering in University, Country, PH-01123456789. E-mail: author\_name@mail.com

colorless ONU is implemented.

#### **3** SIMULATION RESULTS AND DISCUSSION

The effect of the input power on the performance of the RSOA is explained in this section. Figure 3 shows the relationship between input power and output power of RSOA when the bias current is 130mA, 100mA and 80mA. The RSOA becomes saturated when the input power is more than -15dBm. Less than -15dBm input power, the relationship between the input power and the output power behaves almost as a linear relationship.



Figure 3 relationship between input power and output power of RSOA

In RSOA-based colorless ONU, upstream performance is dependent on the seeding power comes from OLT.



Figure 4. The effect of the seeding power on upstream performance.

Figure 4. shows the effect of the seeding power on the BER performance for the upstream data. It is clear that the performance is better when the seeding power in higher. That is a result of the saturation of the RSOA. When the RSOA become saturated, the amplified spontaneous emitting power (ASE) will decrease [4]. This improves the optical signal to noise ratio (OSNR). In case of low seeding power, the ASE increases and the signal power is not in a sufficient level, which makes the OSNR very low and reduce the performance of the upstream.

To guarantee good performance, the seeding light power should be high enough to saturate the RSOA. In case of increasing splitting ratio to increase customers, the seeding power sent from OLT must be pre-amplified.

#### **4** CONCLUSION

As a conclusion, the limitations of the RSOA-based colorless ONU in FTTH-PON are caused by the effect of the low seeding light power. When the seeding power is low, the RSOA will not be saturated. In this case, ASE power will increase and makes the OSNR worse. To guarantee good performance, the seeding light power should be high enough to saturate the RSOA.

## REFERENCES

- Chang-Hee, Lee, "Colorless optical sources for WDM-PON." In Lasers & Electro Optics & The Pacific Rim Conference on Lasers and Electro-Optics, 2009. CLEO/PACIFIC RIM '09, pp: 1-2, 2009.
- [2]. Sengda Tu, Ho-Chun Lin, Ching-Sheu Wang, Hsin-Han Liao, Hai-Lin Wang, Gong-Cheng Lin, Ting-Chung Chang, Jy-Wang Liaw, Shian-Ming Chen. "The pilot trial of colorless WDM PON system in Taiwan". In OptoeElectronics and Communications Conference (OECC), 16th, pp: 13-14, 2011.
- [3]. Iwatsuki, K. and Kani, J. I,. "Applications and Technical Issues of Wavelength-Division Multiplexing Passive Optical Networks With Colorless Optical Network Units [Invited]." Journal of Optical Communications and Networking, IEEE/OSA 1(4): C17-C24, 2009
- [4]. Ziad A. El-Sahn, Walid Mathlouthi, Habib Fathallah, Sophie LaRochelle, "Dense SS-WDM Over Legacy PONs: Smooth Upgrade of Existing FTTH Networks" *Journal of lightwave technology*, vol. 28, No. 10, MAY 15, 2010

# The Transformation of Saudi Arabia through Tablet-Based Technology

Ebtisam Alqhtani, Zarina Samsudin & Hesham Fakrany

Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, Penang, Malaysia

ebtisam.alqhtani@hotmail.com, ina@usm.my, hesham.fakrany@hotmail.com

Abstract— The aim of this paper is to explore how tablet technology will eventually transform our education system by heightening both student and teacher engagement in a variety of learning contexts. We expect to discern from our research the positive impact that Tablet PCs will have on learning, although the use of this technology is not yet widespread. Preliminary findings clearly demonstrate that exploration by students and teachers alike will be greatly facilitated by using this newest technological innovation. Functionality, cost-effectiveness, and ease of operation of tablet-based technology are also additional factors to be considered when considering the incorporation of tablets into the classroom or workplace.

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

Index Terms— Tablet, m-learning, learning technology, computer in education

# **1** INTRODUCTION

RINTED textbooks are some of the conventional learning tools used by both students and teachers in many educational institutions in the world. However, the era for textbooks may slowly be ending thanks to technology innovations and advancements. According to Gimbert and Cristol in [1], technology is quickly becoming an integral learning tool for promoting, linguistic, social and cognitive development of learners. Today, the question is not whether or to what extent we should use technology in classrooms, but rather how should we use it. With innovations being borne everyday each having greater potential to improve the learning process than its predecessor, the application of technology in classrooms presents educators with many ongoing challenges. Nevertheless, the benefits of technology to both teachers and learners are apparent and the only thing to do is carefully plan out its role in the curriculum [2]. Our current education system is dealing with like most agencies in a down economy crippling budget constraints. Like these other organizations our schools are finding themselves having to do more with less people, less supplies, less help for our students. With tablet computers, we will not only be delivering the most up-to-date information for our students, we will be doing it at a reduced cost.

There are many uses for these efficient cost-effective devices. For example, tablets may be utilized in the classroom (along with a projector) as an electronic blackboard. As such, they have numerous functions including the display of supply and demand curves, drawings of the human body, musical scores, electronic circuit diagrams, and even mathematical expressions (e.g., summation and integral signs) [3-4]. Also text (in any language's alphabet) may be captured and displayed. The tablet computer exhibits a wide range of colors and widths of "ink", and allows for an unlimited number of pages on which to write,

without the need for erasing [5-6]. Students benefit greatly by creating class notes on the tablet, which may then be saved and posted to the Web, where other students can easily retrieve and view them using freely available software (tablets not required). There are instructional opportunities outside of the classroom as well. For example, tablets have the potential to be used for electronic grading; assignments may be emailed to faculty or dropped off in digital drop boxes whereupon faculty members may write their comments on the work and return it to the students the same way. Not surprisingly, there have been some initial reports of faculty use of tablets in the classroom. For instance, Wassgren and Krousgrill in [7] used a tablet-based system to record their pen-annotated, PowerPoint lectures in class for future storage and dissemination via the Web. Anderson in [8] also describes the use of a similar system. In this application, one faculty member, responding to the increased practicality said, "Being able to diagram and spontaneously work examples instead of having to use a pre-scripted PowerPoint slide deckfelt like teaching a real class."

# **2** THE TABLET'S IMPACT ON LEARNING

Bilen, et al. in [9] reports their findings on the evaluation of Tablet PC technology use and its' impact on learning across a variety of undergraduate and graduate classes from four technology and engineering disciplines at Pennsylvania State University. They measured the process using an assessment model measures to carry the task to the technology used and incorporates social learning theories. This dual approach allows us to evaluate a wide range of variables such as how the tablet's characteristics influence users, its' versatility in the classroom and the role social influence (e.g., classmates, team members) plays in the extent to which the Tablet PC is used. Perhaps most important is the tablet's impact on students' learning. Early results from one undergraduate engineering course and one graduate technology course show that approximately 65% of students in a design-oriented course found that Tablet PC use enhanced their learning experience, while only 35% of students in a non-design-oriented course thought so. In both courses, nearly 50% of students found their classmates were helpful in instructing them on how to use the Tablet PC.

<sup>•</sup> Ebtisam Alghtani is currently pursuing PhD degree program at Universiti Sains Malaysia, Malysia. E-mail: ebtisam.alghtani@hotmail.com.

Preliminary regression results show that the level of tasktechnology fit seems to positively influence students' use of the Tablet PC and that social influence positively affects students' learning gain as the result of Tablet PC use.

Koile and Singer in [10] describes a strictly controlled study of the impact that Classroom Learning Partner, a Tablet-PCbased classroom interaction system, had on student performance. The goal of their study was to test the hypothesis that the use of such a system improves student learning, especially among the poorest performing students. They also describe the validation of that hypothesis, and the controls, performance metric, and assessment methodology that were developed in the course of their needs.

El-Gayar, Moran and Hawkes in [11] develops and empirically tests a factor model for understanding college students' acceptance of Tablet PC (TPC) as a means to forecast, explain, and improve their usage pattern in education. The analysis involved more than 230 students from a regional Midwestern institution. Overall, the proposed model exhibited a good fit with the data and provided a satisfactory explanation for students' acceptance of TPC in an educational setting. Analysis of the results suggests a number of implications for educational institutions. Most notably among these is the need for programs aimed at influencing students' attitudes toward and perceptions of tablet-based technology. It is highly probable that the creation of such programs will not only enhance the image of the tablet on campus, but also promote the widespread use of TPC.

# **3 METHOD**

To gather comprehensive insights about how tablets are used, we employed a variety of research methods including written diaries and contextual inquiry observations. We selected research methods to help us understand different aspects of tablet usage and triangulate varying insights, while the diaries permitted us to collect a large amount of self-reported data over a longer period of time.

#### 3.1 Sample

A total of 30 participants were recruited from two locations in Saudi Arabia: Jeddah and Abha. This allowed us to draw participants from urban, suburban, and rural areas. Participants were recruited according to a number of criteria including gender, age, education, occupation, family situation, commuting habits, length of tablet ownership, and characteristics of their tablet. These criteria were guided by previous research by Fischer [12], to target a sample of participants that was as representative of the current Malaysian population of tablet users as possible. Participants comprised 30 individuals who took part in the diary study of which 80% (24) were women and 20% (6) were men. They ranged in age from 18 to 70, with the following distribution across the different age ranges: 18-23 (13%), 24-30 (27%), 31-40 (40%), 41-70 (20%). All participants owned a tablet: 84% owned an Apple iPad and 16% owned an Android-based tablet. For 39% of those users, the tablet they used during the time of the study was the first tablet they had ever owned. At the time of the study, while 43% of all those surveyed were still new

to using their tablet PCs (ownership of less than three months), 18% had possessed their tablet for more than three months. To gain further insights into tablet usage, we conducted contextual information sessions with four of the study participants.

#### 3.2 Diary study methodology

To uncover all of the activities carried out on tablets, participants were asked to complete a written diary entry whenever they used their tablet over the two-week study period. We used Google Forms to collect written diary entries, accessible to the participants through the browser of their desktop computers. The diary questions were designed to capture details about each activity, including an explanation of the activity itself, the time and length of the activity, the context in which that activity was conducted (e.g. location, secondary activity, other people), as well as information about transitions to and from other connected devices or the real world. The diary questions included a set of open-ended as well as several closed-ended questions. Researchers periodically called on participants to verify their involvement, reminding them to complete the question set every time they used their tablets. The latter were not able to review their previous diary submissions. A major limitation of this diary study research method is the self-reported nature of the data gathering technique. For example, participants may forget to record some activities, resulting in under-reported activity frequencies. Overall, though, the diary study methodology allowed us to collect in-depth information about each instance of tablet usage, when it occurred naturally, without us having to be present for observation.

#### **3.3 Personal Interviews**

To broaden our understanding of the function of the tablet in the participants' lives, we also interviewed several individuals. We wanted to understand how it is used alongside other devices they own, learn about and observe the context in which they are using the tablet, and follow up with questions about specific use cases described in their diary entries. Each interview lasted 10 minutes, and we used a semi-structured guide in the interview process. Contextual Inquiry was essential both when we reviewed the diary submissions and when we conducted the follow-up interviews. As researchers, we identified areas in which we would be able to collect more in-depth data, thus allowing for unique insights.

After the initial interviews, researchers followed up with four participants previously interviewed. There were three contextual inquiries lasting 10 minutes each. During each inquiry, we observed our subjects' activities, recorded their interactions with the tablet and other devices, and periodically asked follow-up questions for clarification. The activities that we observed included looking up recipes and cooking, family use before, during and after dinner, and productivity tasks in a workplace environment.

#### 3.4 Most Frequent Tablet Activities

Across all reports of tablet use, the most frequent activities were found to be: Checking emails (with light responding), playing games, social networking, searching for information, listening to music, shopping (browsing and purchasing), reading a book, checking the weather, and watching movies. These results are similar to those in a Nielsen report which found that email, social networking, and watching videos were the top three activities, although we found that playing games, informationseeking, and listening to music were more common than watching videos. It is important to discern whether the reported tablet activities are common among all participants, or if they can be accounted for by just a few participants' more frequent engagement. To determine this, we divided the activities into four distinct groups: Activities reported 1) by few participants with low frequency, 2) by few participants with high frequency, 3) by a high number of participants with low frequency, and 4) by a high number of participants with high frequency. Clearly, email checking, playing games, social networking, looking up information, and watching videos are distinct among activities, as they are common to all participants and represent activities that are done frequently.

# 4. OVERALL RESULTS OF OUR RESEARCH

Once the research was completed, we had collected 120 written diary entries, 11 detailed participant profile write-ups from the field visits, and observations from five contextual inquiries. We conducted the analysis in two stages: 1) quantitative analysis of the written diaries; and 2) triangulation of insights to develop a set of conclusions.

Over a two-week period, our participants reported 120 incidences of tablet use, with an average of 4 incidences per participant.

These numbers clearly indicate a range in the number of tablet uses reported by each participant since 80% of all diary entries were submitted by only12 out of the 30 participants. It is important to note that due to limitations of the diary study approach, it is unlikely that these numbers reflect the exact volume of actual tablet use, but rather a lower-range estimate. We coded all open-ended responses to diary questions using a bottom-up approach; similar activities were grouped and then turned into frequency distributions. Finally, we reviewed the diary study analysis to triangulate its' findings. This analysis resulted in a set of insights about the overall distribution of tablet activities and correlations of tablet activities to contextual factors.

# 5. CONCLUSION

During the course of this paper, we have explored the diverse uses of tablet technology in both the classroom and the workplace. We have seen that the tablet's versatility, functionality and ease of usage not only enhance learner engagement, but also greatly facilitate the learning process itself. Research methods such as diary studies and personal interviews also proved essential in collecting in-depth information about tablet usage, and personal interviews we conducted allowed us to gather unique insights into the potential benefits of Tablet PCs. Delivering up-to-date information in an interesting way at a reduced cost is the implicit mandate of tablet-based technology. Unfortunately while it is true that Tablet PCs are beginning to revolutionize learning for students and learners worldwide, it is evident from our research that their usage is still very limited among Saudi Arabian students. This paper, then, should serve as the impetus for exploring the reasons behind this limited use of such a life-changing technology.

#### References

- Gimbert, B. and D. Cristol, *Teaching curriculum with* technology: Enhancing children's technological competence during early childhood. Early Childhood Education Journal, 2004. **31**(3): p. 207-216.
- [2] Swaminathan, S. and J. Wright, *Education technology in the early an primary years (pp. 136-149).* 2003, Major Trends and Issues in Early Childhood Education: Challenges, Controversies and Insights (2nd ed.). New York: Teachers College Press.
- [3] Cromack, J. Technology and learning-centered education: Research-based support for how the tablet PC embodies the Seven Principles of Good Practice in Undergraduate Education. in Frontiers in Education Conference, 2008. FIE 2008. 38th Annual. 2008: IEEE.
- [4] Alvarez, C., C. Brown, and M. Nussbaum, Comparative study of netbooks and tablet PCs for fostering face-to-face collaborative learning. Computers in Human Behavior, 2011. 27(2): p. 834-844.
- [5] Couse, L.J. and D.W. Chen, A tablet computer for young children? Exploring its viability for early childhood education. Journal of Research on Technology in Education, 2010. 43(1): p. 75.
- [6] Roschelle, J., et al., Ink, improvisation, and interactive engagement: Learning with tablets. Computer, 2007. 40(9): p. 42-48.
- [7] Weitz, R., B. Wachsmuth, and D. Mirliss. The Tablet PC Goes to College: Evaluating the Utility of Tablet PCs for Faculty in the Classroom. in World Conference on Educational Multimedia, Hypermedia and Telecommunications. 2005.
- [8] Anderson, R., et al. Experiences with a tablet PC based lecture presentation system in computer science courses. in ACM SIGCSE Bulletin. 2004: ACM.
- [9] Bilén, S.G., et al., Tablet PC Use and Impact on Learning in Technology and Engineering Classrooms: A Preliminary Study. The Impact of Tablet PCs and Pen-Based Technology on Education: Evidence and Outcomes, 2008, 2008: p. 11.
- [10] Koile, K. and D. Singer. Assessing the impact of a tablet-pcbased classroom interaction system. in Monograph of the 3rd Workshop on the Impact of Tablet PCs and Pen-based Technology on Education. Evidence and Outcomes, Purdue Univ., West Lafayette, Indiana, USA. 2008.
- [11] El-Gayar, O., M. Moran, and M. Hawkes, Students' Acceptance of Tablet PCs and Implications for Educational Institutions. Journal of Educational Technology & Society, 2011. 14(2): p. 58-70.
- [12] Fischer, J., Sweet Fruit or Poisoned Apple-The iPad's Effect on Newspapers. J. on Telecomm. & High Tech. L., 2012. 10: p. 173.

# Towards The Use of Advance Technology in E-Learning: A Review

Hesham Fakrany, Merza Abbas & Ebtisam Alqhtani

Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, Penang, Malaysia

hesham.fakrany@hotmail.com, merza@usm.my, ebtisam.alqhtani@hotmail.com

**Abstract**— E-learning is the use of technology to enable people to learn anytime and anywhere. The use of latest learning technology in e-learning comprises all forms of electronically supported learning and teaching. The information and communication systems, whether networked learning or not, serve as specific media to implement the learning process. In this paper we introduce the current status of the technology use in e-learning by looking at different e-learning technologies and briefly discuss its effects in learning. It then explains five related studies that addressed the impact from using a definite technology in e-learning. We aim to acknowledge the current status o using technology in e-learning.

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_

Index Terms—ICT, e-learning, technology, LMS

## **1** INTRODUCTION

THE current issues in teaching programming language among university students have discussed by different researchers as the need to provide an alternative support to help those students interact with learning content. Such aspect comes under the classification of blended learning which refers to the various teaching strategies, which focus on understanding how the user interacts with the computer during a particular task, and how the computer fits into the scheme of the user's total job or environment (the system) [1].

From the other hands, the wide perspective of presenting learning contents in e-learning among learners has addressed and specified the main problem of these presentation techniques over the Web that e-learners may be operating during the teaching process [2]. Online learning applications designed and developed by following a certain strategy for processing and displaying user queries are accessible on the Internet. Nevertheless, most of learners find their self not familiar with the unlimited pedagogical representations, and this may affect their way of understanding and interacting with the content, especially in teaching programming language which demands different learning and design requirements. In addition, instructional designers of e-learning environments need to take into account the various patterns for providing users with the desired representation to teach programming language. Furthermore, the query process also depends on the way knowledge is represented.

This leads to the need to carry out an evaluation and adoption of teaching associated methods among learners. The importance of adaptation in teaching plays a significant role in the utilization of technology during learning. In e-learning environments, there are primarily two types of adaptations: a) to client device, and b) to client activity [3].

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

An alteration to learner behaviour means that the learning environment should fit the learner's assumptions. For that, learners usually face difficulties in understanding the learning materials, which cause them to search for alternate ways to access and represent knowledge so that they can browse through content that fits their needs. In other words, learning environment-based-technology should fit different learners and learning styles [4].

### **2.** LITERATURE

Abdullah, Sufian and Nantha in [5] investigated the problems of learning programming subjects. They also explorerd the various attempts made to solve teaching and learning problems. They propsoed the use of virtual pair programming (VPP), which was used to provide a synchronous mode of collaboration between learners. They studied the effectiveness of asynchronous VPP in the learning of object-oriented programming among students at Open University Malaysia (OUM). They found that most of the learners have given positive feedback, indicating that they are happy with the use of asynchronous VPP.

Hadjerrouit in [6] introduced the effects of blended learning tools in helping learners to understand the new programming language. He addressed the current issues faced by learners from using the face-to-face learning with information technologies that found to be lacked in providing an effective teaching and efficient solutions for learning. Thus, he referred in his study different learning theory and pedagogical strategies. After reviewing these theories and strategies, he found that there is still a need for design-based research approach to blending learning through successive cycles of experimentations, where the shortcomings of each cycle are identified, redesigned, and reevaluated. He also used blended learning model in Java programming at the introductory level. It presents the design, application, and evaluation of the approach and its implications for the learning of introductory computer programming.

Hesham Fakrany is currently pursuing PhD degree program at Universiti Sains Malaysia, Malysia. E-mail: hesham.fakrany@hotmail.com.

Zin, Othman and Yue in [7] focused on content for elearning based on embedding pedagogical aspect in learning activities and processing. Zin and others presented features of learning activity in an e-learning system called Learning Activities E-Learning Management System (LAE-LMS) based on level of processing in Instructional Management Systems Global Learning Consortium Learning Design (IMS LD) Conceptual Model. This system allows teacher to plan, manage and monitor learning activities, while students carry out the activities. LAE-LMS consists of five modules: registration module, synopsis module, authoring module, monitoring module and learner module. Acceptance testing showed that teachers and learners have high positive perception on security, suitability, accurateness and satisfaction of the features of LAE-LMS.

Zhong in [8] intended to determine how a definite strutual representation of learning contents in e-learning based pedagogical strategies helps students develop better understanding of contetn classification. The students were asked to examine the pedagogical strategies employed by different teachers in 5 primary classes. The findings indicated that integrated use of multimedia resources, teacher scaffolding and provocative questions focusing on classification criteria, as well as students' collaborative learning have impacted on students' learning that could enhance students' understanding of animal classification.

Srimathi and Srivatsa in [9] described the benefits of elearning systems among university students as 1) delivering and managing educational elements through Web-based applications, creating a useful learning environment; and 2) customising information in an educational environment while allowing control over learning. Srimathi and Srivatsa aimed to design an interactive learning application with simple context objects using the power of elearning principles. The feasibility study used the derived Instructional Design Template for further incorporation with the Sharable Content Object Reference *Model* (SCORM) model. The usage of SCORM principles helped to enhance the quality of representation of e-learning content.

Stary and Totter in [10] assessed the examination process to provide learners with the capability to control the selection and presentation of content, as well as the learning process, according to their needs, learning styles and preferences. They employed two e-learning design theories, both strengthening learner self-control, but developed on different grounds and following different patterns. It was reported that both approaches revealed several types of learner control to be helpful in self-managed learning.

Another study highlighted the representation learning contents in a number of virtual learning environments (VLEs) that were used specifically in educational fields as demonstrated by [11] the importance of VLEs as a learning method; such VLEs depend on the integration of ICT components into a formal representation of learning contents. Meanwhile, the study examined the suitability of presenting the textbook contents over VLE in higher education, along with traditional lectures, in a professional safety engineering course. Figure 2.7 shows the proposed VLE Virtu learning structure.

A study conducted by [12] deployed agents through the elearning environment. They aimed to enable numerous users to coordinate collectively and intelligently with the proposed learning environment. The development of the e-learning system involved the interaction of users with the system, adopting an artificial agent combination. The system structure was geared towards working within actual Internet/intranet settings. The conceptual structure was classified into three parts: characteristic learning space, collaborative space and cooperative space. The system was combined with the agent-based approach as a suitable solution to assess these elements. Finally, Harbouche and Djoudi used a certain methodology to design the agent-based approach.

Latif, Hassan and Hasan in [13] developed an effective interaction protocol requirement for automating document downloading and notification in an e-learning system. They presented three main agents in system development: the e-learning agent, lecturer agent and student agent. The protocol was characterised using ISLANDER editor as the graphical authentication, which aims to monitor the correctness of the fundamental interaction protocol using the Electronic Institution (EI) approach. However, the protocol focused on one-toone agent interaction through identifying learners' relations with the content. The e-learning performance structure is shown below.



## 3. CONCLUSION

This study showed the different implication of various learning tools in e-learning. The study emphasized the ability of conducting online learning on the current technological status. Finally we distinguished between these implications by comparing different studies conducted based on the use of technology in learning.

# REFERENCES

- Graham, C.R., *Blended learning systems*. Handbook of blended learning: Global Perspectives, local designs. Pfeiffer Publishing, San Francisco, <u>http://www</u>. publicationshare. com/graham\_intro. pdf, 2006.
- [2] Mercer, N., The guided construction of knowledge: Talk amongst teachers and learners. 1995: Multilingual Matters Ltd.
- [3] Bonk, C., R. Wisher, and J. Lee, *Moderating learnercentered e-learning: Problems and solutions, benefits and implications.* Online collaborative learning: Theory and practice, 2004: p. 54-85.
- [4] Chen, C., H. Lee, and Y. Chen, Personalized e-learning system using Item Response Theory. Computers & Education, 2005. 44(3): p. 237-255.
- [5] Abdullah, M.Z., I. Sufian, and S. Nantha Kumar, Improving Learning of Programming Through E-Learning by Using Asynchronous Virtual Pair Programming. Turkish Online Journal of Distance Education (TOJDE), 2006. 7(3).
- [6] Hadjerrouit, S. A blended learning model in java programming: A design-based research approach. in Proceedings of the 2007 Computer Science and IT Education Conference. 2007. Norway.
- [7] Zin, N.A.M., Z.A.H. Othman, and W.S. Yue, LAE-LMS: Applying Pedagogical Aspect of Learning Activity in E-Learning System. Asian Journal of Information Technology, 2009. 8(3): p. 74-83.
- [8] Zhong, Y., A Study of Autonomy English Learning on the Internet. English Language Teaching, 2009. 1(2): p. P147.
- [9] Srimathi, H. and S. Srivatsa, Knowledge representation in personalized elearning. Academic Open Internet Journal, 2008. 23.
- [10]Stary, C. and A. Totter. *On learner control in e-learning*. in 13th Eurpoean conference on Cognitive ergonomics: trust and control in complex socio-technical systems 2006: ACM.
- [11]Koskela, M., et al., *Suitability of a virtual learning environment for higher education*. The Electronic Journal of e-Learning, 2005. **3**(1): p. 21-30.
- [12]Harbouche, K. and M. Djoudi, Agent-Based Design for E-learning Environment. Journal of Computer Science, 2007. 3(6): p. 383-389.
- [13]Latif, N., M. Hassan, and M. Hasan. Interaction protocol for agent-based e-learning system. in 11th International Conference on Information Integration and Web-Based Applications & Services (iiWAS 2009). 2009. Kuala Lumpur: ACM.
- [14] Marivate, V., G. Ssali, and T. Marwala, An Intelligent Multi-Agent Recommender System for Human Capacity Building. Arxiv preprint arXiv:0806.2216, 2008.
- [15]Otsuka, J., V. Bernardes, and H. Da Rocha, A multiagent system for formative assessment support in Learning Management Systems. TIDIA First Workshop–E-Learning Project. 2004.

# Synthesis and Characterization of Silver Oxide Nanoparticles by a Novel Method

Ng Law Yong, Akil Ahmad, Abdul Wahab Mohammad

**Abstract**— Silver oxide (Ag<sub>2</sub>O) nanoparticles were synthesized through a simple chemical method using silver nitrate (AgNO<sub>3</sub>) in the presence of polyethylene glycol (PEG) as reducing agent. In current study, synthesis process by maintaining the solution pH using sodium hydroxide (NaOH) solution was investigated and compared. Zetasizer Nano series, X-ray diffraction (XRD) analysis, and Fourier transform infrared spectroscopy (FT-IR) studies were performed to confirm the size distribution measurement, chemical compound composition, crystallographic structure and functional group corresponding to Ag-O. Silver nanoparticles synthesized in the form of silver oxide were confirmed by the XRD study.

Index Terms— Fourier transform infrared spectroscopy, Nanoparticles, Polyethylene glycol, Silver oxide, Size distribution, Crystallographic study, X-ray diffraction analysis

----- 🔶 ------

#### **1** INTRODUCTION

IN recent years, nano-sized metal particles have got much attention due to its unique optical, electrical and magnetic properties, which depends on the size and shape of the particle. Silver oxide (Ag<sub>2</sub>O) nanoparticles is a well-known material having vast applications in the field of oxidation catalysis [1, 2], sensors [3], fuel cells [4], photovoltaic cells [5], all-optical switching devices, optical data storage systems [6] and as a diagnostic biological probes [7].

Many physical and chemical properties including luminescence, conductivity, and catalytic activity depend upon the size of nanoscale materials. The most common and simplest bulk-solution synthetic method that has been used for the preparation of metal nanoparticles is chemical reduction of metal salts. For the synthesis of nanoparticles we generally use the soluble metal salt, a reducing agent, and a stabilizing agent (caps the particle and prevents further growth of the particle or aggregation).

It was, therefore, thought worthwhile to prepare Ag<sub>2</sub>O nanoparticles for its various applications through simple chemical modification using polyethylene glycol (PEG), sodium hydroxide (NaOH) and silver nitrate (AgNO<sub>3</sub>).

#### 2 MATERIALS AND METHODS

The silver oxide (Ag<sub>2</sub>O) particles were produced using starting materials such as polyethylene glycol (PEG) of molecular weight 6000 (supplied by Sigma-Aldrich), silver nitrate (AgNO<sub>3</sub>) salt (supplied by Sigma-Aldrich) and sodium hy-

droxide (NaOH) pellets (supplied by Merck). All of the chemicals were of analytical grades and used without further purification. Reverse osmosis (RO) water used throughout the experiment for solution preparation, cleaning purposes or other usages were produced using Millipore RiOs 3 water purification system.

#### 2.1 Synthesis of Silver Oxide Particles

In current study, wet chemical route was employed for the silver oxide particle synthesis. This study had conducted and compared two different routes of silver oxide particle synthesis. In the Route 1, the pH of the solution mixture was not controlled or adjusted. The pH of the solution mixture was allowed to change accordingly after the reaction took place. In the Route 2, the pH of the solution mixture was fixed at the range of pH 9.8 to pH 10. Addition of the NaOH solution was employed to maintain the solution pH until the completion of the reaction. Some of the details of the Route 1 and Route 2 were further discussed in the following sections.

#### 2.1.1 Route 1

20 g of the PEG was dissolved in 1 liter of the RO water before being heated up to 50 degree Celsius. The solution was allowed to be stirred for another 1 hour to ensure all the PEG was completely dissolved to form a homogeneous solution. Aqueous PEG solution obtained was then filtered to remove impurities, if any. Silver nitrate solution prepared using 0.5 g of silver nitrate salt was added into the PEG solution prepared under a constant stirring rate and at constant temperature of 50 degree Celsius. pH of the solution was not controlled. The solution was continuously stirred for 1 hour to complete the chemical reactions. After the formation of the particles, the solution was filtered through the filter paper to separate the particles from the mother solution. Particles obtained were rinsed with RO water several times before it was rinsed again using ethanol. The particles were dried in oven at 60 degree Celsius for overnight.

#### 2.1.2 Route 2

All the steps employed in the Route 2 were similar to that of Route 1, except where the addition of the silver nitrate solu-

Ng Law Yong is currently pursuing doctorate degree program in Chemical and Process Engineering, Universiti Kebangsaan Malaysia, Malaysia, PH-+60389216825. E-mail: nglawyong@gmail.com

Akil Ahmad and Abdul Wahab Mohammad are postdoc and professor, respectively, in Chemical and Process Engineering, Universiti Kebangsaan Malaysia, Malaysia, PH-+60389216825. E-mail: akilchem@yahoo.com (Corresponding author for this paper is Ng Law Yong)

tion was followed by the pH adjustment. The solution mixture pH was set at pH 9.8 to 10 throughout the reaction process. During the pH adjustment, NaOH solution of 0.1 M was slow-ly added into the solution mixture when the solution pH was reducing.

# 2.2 X-ray Diffraction Analysis

The particles formed from both Route 1 and Route 2 was characterized using X-ray diffraction analysis. BrukerD8 Advance X-ray diffractometer, Germany, was employed to provide Xray diffraction patterns of the particles.

# 2.3 Size Distribution Analysis

Zetasizer Nano series was used for the size distribution measurement. Water has been used as dilution medium and polar dispersant in the measurement of particle size distribution. Refractive index of water was set as 1.330 with the viscosity 0.8872 cP. Temperature of the measurement was conducted at 25 °C.

#### 2.4 Fourier Transform Infrared (FTIR) Spectroscopy

Nicolet 6700 FTIR spectrometer (Thermo Fisher Scientific Inc., MA) with a DTGS KBr detector and a KBr beam splitter was used for FTIR analysis. All spectra were obtained from 32 scans at a 4.00 cm<sup>-1</sup> resolution with wave numbers ranging from 400 to 4000 cm<sup>-1</sup> and with an optical velocity of 0.6329.

# **3 RESULTS AND DISCUSSION**

During the preparation of the silver oxide particles, the pH of the solution was monitored continuously. However, in the Route 1 method, no adjustment of the solution pH will be conducted. In the Route 2 method, NaOH solution will be introduced into the solution mixture to keep the solution pH at the range of pH 9.8 to pH 10. In both methods, the solution will change from yellowish to brownish colour to indicate that the chemical reaction took place in the solution mixture (Figure 1 (a) and Figure 1 (b)). After the completion of the chemical reaction which took place slowly, brownish black precipitates of silver oxide particles were observed to be formed in the solution mixture (Figure 1 (c)). The proposed chemical reaction according to reported study [8] was represented by the chemical equation below:

$$2Ag^{+} + 2OH^{-} \rightarrow Ag_2O + H_2O \tag{1}$$



(a)



(b)



(c)

Figure 1. (a) Yellowish solution of the solution mixture at initial stage (b) brownish solution of the solution mixture at intermediate stage and (c) brownish-black silver oxide particles formed at the final stage of the chemical reaction.

#### 3.1 X-ray Diffraction Analysis

According to the Figure 2, Route 1 method was less efficient to produce high purity of silver oxide. It consisted of other compound such as silver carbonate. This could be explained by the low concentration of hydroxide ions in the solution. The conversion of the silver ions into silver oxide was less effective in comparison to Route 2 method. However, Route 2 method was successful to produce high purity of silver oxide particles with high crystallinity. By referring to the reported case study [9], formation of metallic silver particles crystallized in the face centered cubic (fcc) structure would produce the XRD pattern which is similar to the XRD pattern produced through Route 2. Maintaining the solution pH at high alkalinity such as pH 9.8 to pH 10, high concentration of hydroxide ions can contribute to more silver oxide particles formation. Using the Scherrer formula D =  $n\lambda/\beta\cos\theta$ , where D is the crystallite size, n is a constant (=0.9 assuming that the particles are spherical),  $\lambda$  is the wavelength of the X-ray radiation,  $\beta$  is the line width (obtained after correction for the instrumental broadening) and  $\theta$  is the angle of diffraction we have calculated the crystallite size of the Ag<sub>2</sub>O particles. The average particle size obtained from XRD data is found to be about 37.90 nm.



Figure 2. XRD patterns obtained for both Route 1 and Route 2 method.

#### 3.2 Size Distribution Analysis

Size distribution analysis using silver oxide particles verified that the averaged size produced using Route 1 were 1411 nm. However, particles with smaller averaged size distribution were produced using Route 2, which was 319.6 nm. Particle sizes are normally affected by various factors such as solution pH, stirring speed, and the present of stabilizer. In Route 1, large particle sizes were obtained and these could be reasoned with its low pH after the addition of the reactant. Route 2 produced much smaller-sized particles as the pH was maintained around pH 9.8 to 10. However, the occurrence of particle agglomeration which can hardly avoided using mechanical stirring caused the final obtained particles to have sizes at around 319.6 nm. This issue has to be overcome in the near future by controlling the stirring speed, temperature of the mixture, and the presence of stabilizer.



Figure 3. Particle size distribution for both particles produced using Route 1 and Route 2

#### 3.3 Fourier Transform Infrared Spectroscopy (FT-IR) Study

As shown in Fig. 4, the intense peak appeared in the range of 513 cm<sup>-1</sup> which correspond to the stretching vibration of Ag-O group. This further verified the compound as silver oxide particles in addition to the XRD analysis conducted.



Figure 4. FT-IR spectra of Ag<sub>2</sub>O nanoparticles.

# **4** CONCLUSION

In this article, we have successfully synthesized the silver oxide nanoparticles through a simple chemical route. X-ray diffraction analysis reveals that the crystallite size of the Ag<sub>2</sub>O particles was found to be 37.90 nm. From the FT-IR Study the stretching vibration of Ag-O group was found to be in range of 513 cm<sup>-1</sup> which confirms the formation of Ag<sub>2</sub>O particles.

# REFERENCES

[1] F. Derikvand, F. Bigi, R. Maggi, C.G. Piscopo, G. Sartori, Journal of Catalysis 271 (2010) 99.

[2] W. Wang, Q. Zhao, J. Dong, J. Li, International Journal of Hydrogen Energy 36 (2011) 7374.

[3] V.V. Petrov, T.N. Nazarova, A.N. Korolev, N.F. Kopilova, Sensors and Actuators B: Chemical 133 (2008) 291.

[4] E. Sanli, B.Z. Uysal, M.L. Aksu, International Journal of Hydrogen Energy 33 (2008) 2097.

[5] Y. Ida, S. Watase, T. Shinagawa, M. Watanabe, M. Chigane, M. Inaba, A. Tasaka, M. Izaki, Chemistry of Materials 20 (2008) 1254.

[6] W.-X. Li, C. Stampfl, M. Scheffler, Physical Review B 68 (2003) p.165412.

[7] Y.-H.W.a.H.-Y. Gu, Microchimica Acta 164 (2009) 41.

[8] P.K. Sahoo, S.S. Kalyan Kamal, T. Jagadeesh Kumar, B. Sreedhar, A.K. Singh, S.K. Srivastava, Def. Sci. J. 59 (2009) 447.

[9] M. Popa, T. Pradell, D. Crespo, J.M. Calderón-Moreno, Colloids and Surfaces A: Physicochemical and Engineering Aspects 303 (2007) 184.

# CMOS BASED THERMAL ENERGY GENERATOR FOR LOW POWER DEVICES

Z. H. Abdul Rahman, M. H. Md Khir, Z. A. Burhanudin, K. Ashraf, A. A. Abd. Rahman, S. Sulaiman

**Abstract**— This paper presents a thermal energy generator (TEG) designed using complementary metal oxide semiconductor (CMOS) process which converts thermal energy into electrical power. Energy harvesting techniques provide viable option to improve battery performance of low power devices. TEGs are of special interest due to its energy efficient, have no moving part and free maintenance. Secondly, thermal energy or heat gradient is an unremitting energy that is abundantly found from various sources such as the sun, industrial machines, automotives, mobile systems and human body. The proposed energy harvesting device takes advantage of the temperature differences between the hot to cold parts to produce an electrical power and provides a solution for micro-scale electronic systems. A 220 pair of thermopiles made of *n*-doped and *p*-doped polysilicon materials are electrically connected in series and thermally connected in parallel. Simulation shows that at a temperature gradient of 3 K, the proposed device can produce an output voltage and power of 0.29 V and 0.04 mW, respectively. Compatibility of the proposed TEG design with the standard CMOS processes enables to realize a novel on-chip power supply capable of powering many low-power wireless sensor networks and devices.

Index Terms— CMOS based TEG, Thermal energy harvesting, Polysilicon, Seebeck effect, Thermocouple, Thermoelectric generator.

\_\_\_\_\_

# **1** INTRODUCTION

**RENDs** in current electronic technologies are reduction in L terms of device sizes but increasing the performances. Wireless sensor networks (WSNs) have been utilized in various applications ranging from structural monitoring for buildings and bridges, homeland security, aircraft engine monitoring, agriculture to biomedical applications [1-3]. Furthermore, over the past several years, low power electronic technology has become more imperative [4] and are widely used in consumer applications such as smartphones, hearing aid and cardiac pacemaker. Batteries have so far been employed as the power source for the devices. However, batteries exhibit several disadvantages. The use of batteries limit the lifetime of these devices. Power can only be provided over finite period of time [5] due to power drain. Replacement and recharging of batteries will significantly increase the cost and lead to serious environmental pollution [6]. Thus, other possibilities need to be lookout to reduce dependency on batteries. A promising solution to this problem is to use energy harvesting technologies. Energy is found in the target environment of these devices in several forms such as kinetic, thermal and radiation energy. The process of extracting these energies into electrical energy is known as energy harvesting, or energy scavenging. Negligible maintenance effort [7] is require and unlimited power source can be provided using energy harvesting technique.

There are two types of energy harvesters which are macro and micro types. Macro-energy harvesters mainly focus to reduce carbon emission and for oil dependency. The sun, wind, tides and waves are the most suitable energy sources for the macro- energy harvesters. Typical power produced by macro-energy harvester ranges from kilowatts to megawatts. For micro-energy harvester, the ultimate goal is to power up wireless sensor networks and wearable devices [8]. Among the energy harvesting technologies, TEG is of special interest as a micro-energy harvester. TEGs are prevailing, as they are robust, environmental friendly, have no moving part, compact and provide virtually limitless lifetimes [9]. Abundant waste heat available in the ambience can be converted into electrical energy by using Seebeck effect. With the recent rapid advancement, the CMOS technology has become the predominant fabrication technology. With the aid of CMOS technology, extreme miniaturization of various sensors and actuators has been achieved [10,11]. Furthermore, capabilities of monolithic circuit integration are enhance through CMOS process. Thus, realizing a micro-scale TEG device has becomes a significant research subject in order to scavenge thermal energy that is suitable for low-power applications and wearable devices.

In this paper, the design, modeling and simulation of a micro-scale TEG, compatible with the CMOS fabrication technology is presented. The *p*-doped and *n*-doped polysilicon is chosen as the material of the TEG thermopile, as it is available and can be monolithically integrated in the normal CMOS manufacturing line. The proposed design is simulated to foresee the temperature distribution in the structural parts of the TEG. The temperature difference predicted in the simulation step is used to obtain the output voltage and power.

# **2 THEORY OF TEG**

The term thermoelectric is literally associated with thermal and electrical phenomena. TEG is a solid state device that can convert thermal energy from a temperature gradient into electrical energy [12]. Typically, there are three main thermoelectric effects which are Seebeck effect, Peltier effect and Thomson effect [13].

Seebeck power generation or Seebeck effect is a phenomenon by which an electromotive force or a potential difference is produced by a circuit made of two wires of dissimilar materials when the junctions of the two wires are maintained at different temperatures. This phenomenon was discovered by T. J. Seebeck in 1821. In 1834, 13 years after Seebeck made his discovery, Jean Peltier discovered the reverse process of Seebeck effect -the Peltier effect. He discovered that the passage of an electric current through a thermocouple produces heating and cooling effects depending on the direction of the current. The relation between Seebeck effect and Peltier effect was later on discovered by W. Thomson in 1855 and is known as Thomson effect. This law relates to the rate of generation of reversible heat which results from the passage of a current along a portion of a single conductor along which there is a temperature difference [14].



Fig. 1. Schematic of TEG.

A simple electrical circuit of a TEG is shown in Fig. 1. TEG device consists of an array of thermoelectric couple (thermocouple) composed of two elements: *p*-type and *n*-type semiconductors that are connected thermally in parallel and electrically in series to form a thermoelectric module. Heat is transferred from the hot temperature junction,  $T_h$  at the rate of

 $Q_{\scriptscriptstyle h}$  to the cold temperature junction,  $T_{\scriptscriptstyle c}$  at the rate of  $Q_{\scriptscriptstyle c}$  . To simplify the theoretical analysis, it is assumed that the contact resistance is neglected and all the thermoelectric elements have the same length, L and area, A. Seebeck coefficient, thermal conductivity and electrical resistivity are assumed to be temperature independent. Furthermore, it also assumed that heat transfer only happens due to conduction by thermoelectric elements. The assumptions made above are reasonable when the temperature difference between the two junctions is small. Heat generations in a TEG are resulted from the Joule heating, heat associated with the Seebeck effect and thermal conduction. The heat flow at the hot junction can be described as

$$Q_{h} = \alpha_{pn} T_{h} I - \frac{1}{2} I^{2} R + K (T_{h} - T_{c})$$
<sup>(1)</sup>

In a similar way, we can obtain the heat balance equation at the junction at temperature,  $T_c$  as

$$Q_{c} = \alpha_{pn} T_{h} I + \frac{1}{2} I^{2} R + K (T_{h} - T_{c})$$
<sup>(2)</sup>

where I is the current and K is thermal conductance. As heat flows from the hot to the cold junction, free charge carriers (electrons from *n*-type and holes from *p*-type) are also driven to the cold end. Thus, when an electron conducting (ntype) leg and hole conducting (p-type) leg are connected, an output voltage is produced.

Output voltage related to Seebeck effect is best described with correlation in Eq. (1)

$$V = \alpha_{pn} \Delta T \tag{3}$$

where  $\alpha_{pn}$  is the relative Seebeck coefficients of the p-type

and n-type semiconductor elements and  $\Delta T$  is the temperature gradient between the hot junction,  $T_h$  and the cold junction,  $T_c$ . Output voltage is proportional to the temperature gradient and the relative Seebeck coefficients. Seebeck coefficient of a material is also known as thermopower. It is a measure of the magnitude of an induced thermoelectric voltage in response to a temperature difference across the two junctions and it is dependent on the molecular structures of the materials and the absolute temperature.

The Seebeck coefficient of *p*-type and *n*-type semiconductor elements can be obtained from Eq. (4) and Eq. (5) given below

$$\alpha_{p} = \frac{k_{B}}{q} \left[ \frac{5}{2} - \left[ \left( \frac{E_{i} - E_{f}}{k_{B}T} \right) \right] \right]$$
(4)

$$\alpha_n = -\frac{k_B}{q} \left[ \frac{5}{2} - \left[ \left( \frac{E_f - E_i}{k_B T} \right) \right] \right]$$
(5)

where  $\alpha_p$  and  $\alpha_n$  are the Seebeck coefficient of the p-type and the n-type semiconductor respectively,  $k_{B}$  is the Boltzmann constant, q is the elementary charge,  $E_i$  and  $E_F$  is the intrinsic Fermi energy and Fermi energy of the materials used and T is the temperature of dopants.

The performance or the efficiency of a TEG is governed by the properties of the thermoelectric material which is represented by the figure of merit (Z). The figure of merit of a material is given as

$$Z = \frac{\alpha^2}{\rho k} \tag{6}$$

where  $\alpha$  is the Seebeck coefficient ,  $\rho$  is the electrical resistivity and k is the thermal conductivity. To achieve an adequate value of Z, a high Seebeck coefficient with low electrical resistivity and low thermal conductivity are indispensable.

If we considered, the output power is obtained under matched load resistance, the maximum output power can be expressed as

$$P_{\max} = \frac{V^2}{4R} \tag{7}$$

where R is the internal resistance of the TEG. Substituting Eq. (3) into Eq. (7), the maximum power can be expressed as

$$P_{\rm max} = \frac{\left(\alpha_{pn}\Delta T\right)^2}{4R} \tag{8}$$

If we considered, n- number of thermocouples are connected thermally in parallel and electrically in series, Eq. (6) can be written as

$$P_{\rm max} = \frac{n^2 \left(\alpha_{pn} \Delta T\right)^2}{4R} \tag{9}$$

#### **3 LITERATURE REVIEW**

Many studies have employed Seebeck effect to generate electricity from heat gradient. For a macro-scale TEG, Ahmad Nazri *et al.* [15] have designed an energy harvesting device based on Seebeck effect. The dimension of this device is 40 mm x 40 mm x 3.2 mm with a relative Seebeck coefficent of of 42.36 mV/K and an internal electrical resistance of 0.03  $\Omega$ . The device is capable to produce 5.5 V to 6 V output voltage at  $\Delta T$  = 80 °F and the output current is measured at 400 to 500 mA. Bavel *et al.* [16] fabricated EEG system using thermoelectric fundamental in headband with a total hot plate area of 64 cm<sup>2</sup>. It was designed in ten sections of 1.6 x 4 cm<sup>2</sup> each. The TEG produced 2 to 2.5 mW of power. Hongyun *et al.* [17] developed a hybrid power source consisted of solar cells and TEG. Results showed that the open circuit voltage was 0.3 V for a device with dimension of 150 mm x 80 mm x 24 mm.

Micro-scale TEGs have also been reported in literature. Ziyang Wang et al. [18] fabricated a TEG chip consisted of 4700 thermocouples by using micromachining technology. The authors were able to produce 0.25 V open circuit output voltage for a 1 K of temperature difference. Open circuit output voltage per unit temperature difference per unit area was measured at 12.5 VK<sup>-1</sup>cm<sup>-2</sup> and the output power was measured at 0.026 µWK-2cm-2. Till Huesgen et al. [19] fabricated a TEG device by using combined surface and bulk micromachining processes. An output voltage of 9.51 mVK<sup>-1</sup> was obtained from this device under a temperature gradient of 1 K. Kockmann et al, [20] presented a 1 cm<sup>2</sup> microstructured TEG with 7500 thermocouples. Wang et al. [21] modeled a thermoelectric micro generator based on p-type and n-type Bismuth Telluride (Bi<sub>2</sub>Te<sub>3</sub>) material by using MEMS technology and achieved a Seebeck coefficient of about 260 and -188 µV/K.

Commercially used materials for TEG is Bi<sub>2</sub>Te<sub>3</sub>. Bi<sub>2</sub>Te<sub>3</sub> is a promising semiconductor compound, as it provides high thermoelectric figure of merit (FOM). However, it is less compatible with the standard micromachining techniques and extra efforts are needed to make it compatible with CMOS processes. Thus, it is vital to find the best material, companionable with normal CMOS process line. It may be noted that a low thermal conductivity, large Seebeck coefficients and low electrical resistivity make a material an excellent choice for a TEG device. Doped polysilicon has been utilized as thermoelectric material in many reported research works. Polysilicon is able to solve the compatibility issue and possesses the aforementioned characteristics of a suitable material for a TEG.

Several works related to CMOS TEG based polysilicon materials have been reported. Jin Xie *et al.* [22] presented the design, modeling, fabrication and characterization of a TEG. Materials used by the authors are phosphorus and boron heavily doped polysilicon thin films. The device area was 1 cm<sup>2</sup>. At a temperature gradient of 5 K, an open circuit voltage of 16.7 V and an output power of 1.3  $\mu$ W, under matched load resistance, were reported. Hsu Kao *et al.* [23] presented a thermoelectric micro generator fabricated by using 0.35  $\mu$ m CMOS process. Experimental results showed an output voltage of 67  $\mu$ V at a temperature gradient of 1 K. Yang *et al.* developed a TEG using standard CMOS process [24]. This device was able to produce a power factor of 0.0427  $\mu$ W/cm<sup>2</sup> K<sup>2</sup> and voltage factor of 3.417 V/cm<sup>2</sup> K.

# 4 DESIGN OF CMOS BASED TEG

Fig. 2 illustrates the schematic of CMOS TEG design. The harvester consists of 220 thermocouples in series. Phosphorus and boron doped polysilicon is utilized to create *p*-type and *n*-type semiconductor elements. Thermocouple is arranged on the top of the substrate. One junction of the *p*-type and the *n*-type polysilicon legs is coupled to the hot part of the TEG, and the other junction of the *p*-type and the *n*-type polysilicon legs is connected to the cold part. Materials used for the hot plate is made from aluminium. Aluminium is utilized to act as heat receiving area to conduct heat from the hot part to the cold part of TEG. In order to increase the temperature difference between the hot and the cold junctions, heat must flow through the thermopile from the hot to the cold part by using trenches.



Fig. 2. CMOS TEG design.

The working principle of the proposed device is based on Seebeck effect. All the thermocouples are electrically connected in series in order to obtain an array of thermocouple elements.

TEG is designed to meet the requirements given in Table 1. These requirements are decided based on the previous research works and current technology that is available and suitable for CMOS harvesters.

Table 1. Requirements for CMOS TEG			
Temperature gradient ( $\Delta T$ )	3 K		
Open Circuit Voltage (V <sub>open</sub> )	> 0.5 V		
Load Resistance (R <sub>load</sub> )	$< 1000 \Omega$		
Output Power (P)	~1 mW		
Current (I)	~ 1 mA		

To meet the given requirements, the optimized values of the design parameters are given in Table 2. The theoretical model presented in section 2 is used to optimize the parameters of the TEG design.

Table 1. Structure parameter			
Properties	<i>l</i> (µm)	w (µm)	t (µm)

Substrate	3000	1300	100
Thin Film	3000	1300	5
Thermocouple	300	10	0.5
Hot plate	2000	500	1

# **5 RESULTS AND DISCUSSIONS**

CoventorWare, finite element analysis software is engaged to simulate the temperature difference between the hot and cold plates of the TEG. First, a 3D model in accordance with the design presented in Fig. 2 is constructed. The device consists of a thick silicon layer followed by a polysilicon layer, oxide layers and metal layers. Metal layers are connected with the polysilicon layer by using vias. The via is made of tungsten. Boundary conditions are defined by setting a temperature of 303 K for the hot plate, 301 K for the cold part, and 300 K for the substrate.

Fig. 3 shows the simulated temperature distribution for the CMOS TEG.



Fig. 1. Temperature distribution of CMOS TEG.

The result showed that for the hot and cold parts, temperature is uniformly distributed. Trenches has successfully isolated the hot part from the cold part. This thermal isolation of the two parts is required in order to achieve a high temperature differences them. Fig. 4 shows the cross sectional temperature distribution in the CMOS TEG





Probing is done at the hot part, and temperature is measured at 303 K. At the top surface of the cold part, temperature was set at 301 K and at the bottom part it was set at 300 K. Therefore, when probing was done at the polysilicon layer, temperature was measured at 300 K. It showed a temperature drop of 1 K from 301 K to 300 K across from the top of cold part to polysilicon layer. As the hot plate is heated at 303 K, we can conclude that temperature difference between the hot junction and cold junction is 3K.

After done with the thermal analysis of the CMOS TEG, output voltage and output power can be calculated using the theoretical model presented in section 2 and prediction of the behavior and performance of the TEG can be made. On the basis of these predictions, the device can be further be optimized for enhancement performance before transferring the layout in 2D for fabrication process. The simulation and formulation results attained are tabulated in Table 3.

Parameters	Value
Seebeck coefficient of p-type, $lpha_p$	132 µV/K
Seebeck coefficient of n-type, $\alpha_n$	-299 μV/K
Relative Seebeck coefficients, $\alpha_{_{pn}}$	431 µV/K
Load resistance, $R$	500 Ω
Open Circuit Voltage, (Vopen)	0.2845 V
Output Power, $P_{\max}$	0.04 mW
Output current, I	0.140 mA

Table 3 Simulated performance of CMOS TEG based on simulation and formula-

Fig. 5 shows the output voltage versus temperature difference. It is shown that at a temperature gradient of 3 K, this device can deliver an output voltage of 0.2845 V. Graph of the output power versus temperature difference is given in Fig. 6. The output power is measured at 0.04 mW and the output current through a matched load resistance is predicted at 0.140 mA.



Fig. 3. Simulated result of the output voltage for CMOS TEG.



Fig. 4. Simulated result of the output power for CMOS TEG.

# 6 CONCLUSIONS

A thermal energy generator using CMOS technology has been presented. The proposed CMOS TEG device consists of 220 thermocouples connected in series. Materials used for the thermocouple is *p*-type and *n*-type polysilicon. Doped polysilicon gives high Seebeck coefficient and more importantly, it is more compatible with CMOS process as compared to Bismuth Telluride. The TEG can be extended into an array in order to achieve higher power. An output voltage of 0.2845 V and output power of 0.04 mW is achieved based on the simulation and analytical analysis, proving the feasibility of the concept. Voltage is proportional to the Seebeck coefficients and the temperature gradient between the two junctions of the TEG. Therefore, the performance of the CMOS TEG can be improved if higher temperature gradient can be obtained. Further investigation on the effect of contact resistance is indeed a good way to evaluate the real behavior of the TEG performance. Optimization can also be done towards improving the performance for even larger and complex designs of the TEG.

# REFERENCES

- P. D. Mitcheson, E. M. Yeatman, G. K. Rao, A. S. Holmes and T. C. Green, "Energy Harvesting From Human and Machine Motion for Wireless Electronic Devices," Proceedings of the IEEE, vol. 96, pp. 1457-1486, 2008.
- [2] Haowei Bai, M. Atiquzzaman and D. Lilja, "Wireless sensor network for aircraft health monitoring," in Broadband Networks, 2004. BroadNets 2004. Proceedings. First International Conference on, 2004, pp. 748-750.
- [3] L. M. Miller, P. K. Wright, C. C. Ho, J. W. Evans, P. C. Shafer and R. Ramesh, "Integration of a low frequency, tunable MEMS piezoelectric energy harvester and a thick film micro capacitor as a power supply system for wireless sensor nodes," in Energy Conversion Congress and Exposition, 2009. ECCE 2009. IEEE, 2009, pp. 2627-2634.
- [4] A. Ibragimov, H. Pleteit, C. Pille and W. Lang, "A Thermoelectric Energy Harvester Directly Embedded Into Casted Aluminum," Electron Device Letters, IEEE, vol. PP, pp. 1-3, 2011.
- [5] M. Marzencki, M. Defosseux and S. Basrour, "MEMS Vibration Energy Harvesting Devices With Passive Resonance Frequency Adaptation Capability," Microelectromechanical Systems, Journal of, vol. 18, pp. 1444-1453, 2009.
- [6] Xin Lu and Shuang-Hua Yang, "Thermal energy harvesting for WSNs," in Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, 2010, pp. 3045-3052.
- [7] T. Kawahara, T. Fujii, M. Emoto, M. Hamabe, H. Watanabe, J. Sun, Y.

Ivanov and S. Yamaguchi, "Double Peltier Current Lead for Heat Leak Reduction at the Terminals for Superconducting Direct Current Applications," Applied Superconductivity, IEEE Transactions on, vol. 21, pp. 1070-1073, 2011.

- [8] J. P. Carmo, L. M. Goncalves and J. H. Correia, "Thermoelectric Microconverter for Energy Harvesting Systems," Industrial Electronics, IEEE Transactions on, vol. 57, pp. 861-867, 2010.
- [9] M. Strasser, R. Aigner, C. Lauterbach, T. F. Sturm, M. Franosch and G. Wachutka, "Micromachined CMOS thermoelectric generators as on-chip power supply," Sensors and Actuators A: Physical, vol. 114, pp. 362-370, 9/1, 2004.
- [10] I. F. Akyildiz, Weilian Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks," Communications Magazine, IEEE, vol. 40, pp. 102-114, 2002.
- [11] M. Aleksandar, O. Chris and J. Emil, "Wireless sensor networks for personal health monitoring: Issues and an implementation," Computer Communications, vol. 29, pp. 2521-2533, August 2006.
- [12] S. B. Riffat and X. Ma, "Thermoelectrics: a review of present and potential applications," Appl. Therm. Eng., vol. 23, pp. 913-935, 6, 2003.
- [13] S. Lineykin and S. Ben-Yaakov, "Modeling and Analysis of Thermoelectric Modules," Industry Applications, IEEE Transactions on, vol. 43, pp. 505-512, 2007.
- [14] C. Ionescu, N. Codreanu and P. Svasta, "Performance evaluation of a thermoelectric cooler using finite element analysis," in Electronics Technology (ISSE), 2011 34th International Spring Seminar on, 2011, pp. 436-441.
- [15] A. N. A. Razak, N. M. Nor and T. Ibrahim, "Heat energy harvesting for portable power supply," in Power Engineering and Optimization Conference (PEOCO), 2011 5th International, 2011, pp. 436-439.
- [16] M. Van Bavel, V. Leonov, R. F. Yazicioglu, T. Torfs, C. Van Hoof, N. E. Posthuma and R. J. M. Vullers, "Wearable battery-free wirelss 2-channel EEG systems powered by energy scavengers.," Sensors&Transducers Journal, vol. 94, pp. 103-115, 2008.
- [17] Hongyun Yu, Yanqiu Li, Yonghong Shang and Bo Su, "Design and investigation of photovoltaic and thermoelectric hybrid power source for wireless sensor networks," in Nano/Micro Engineered and Molecular Systems, 2008. NEMS 2008. 3rd IEEE International Conference on, 2008, pp. 196-201.
- [18] Z. Wang, V. Leonov, P. Fiorini and C. Van Hoof, "Realization of a wearable miniaturized thermoelectric generator for human body applications," Sensors and Actuators A: Physical, vol. 156, pp. 95-102, 11, 2009.
- [19] T. Huesgen, P. Woias and N. Kockmann, "Design and fabrication of MEMS thermoelectric generators with high temperature efficiency," Sensors and Actuators A: Physical, vol. 145–146, pp. 423-429, 0, 2008.
- [20] N. Kockmann, T. Huesgen and P. Woias, "Microstructured in-plane thermoelectric generators with optimized heat path," in Solid-State Sensors, Actuators and Microsystems Conference, 2007. TRANSDUCERS 2007. International, 2007, pp. 133-136.
- [21] Jin Xie, Chengkuo Lee and Hanhua Feng, "Design, Fabrication, and Characterization of CMOS MEMS-Based Thermoelectric Power Generators," Microelectromechanical Systems, Journal of, vol. 19, pp. 317-324, 2010.
- [22] Pin Hsu Kao, Po Jen Shih, Ching Liang Dai and Mao Chen Liu, "Fabrication and Characterization of CMOS-MEMS Thermoelectric Generator," *Sensors*, vol. 10, pp. 1315-1325, 2010.
- [23] S. M. Yang, T. Lee and M. Cong, "Design and verification of a thermoelectric energy harvester with stacked polysilicon thermocouples by CMOS process," Sensors and Actuators A: Physical, vol. 157, pp. 258-266, 2, 2010.
- [24] Xin Lu and Shuang-Hua Yang, "Thermal energy harvesting for WSNs," in Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, 2010, pp. 3045-3052.

- Zati Hanani Abdul Rahman is currently pursuing masters degree program in Electrical and Electronics Engineering in Universiti Teknologi Petronas , Malaysia, PH-019 5258503. E-mail: zatihananirahman@gmail.com
- Mohd Haris Md Khir is currently working as Senior Lecturer in Electrical and Electronics Engineering department in Universiti Teknologi Petronas,
- Zainal Arif Burhanudin is currently working as Senior Lecturer in Electrical and Electronics Engineering in Universiti Teknologi PETRONAS, Malaysia, PH-0125211773, Email: zainabh@petronas.com.my
- Khalid Ashraf is currently pursuing doctorate degree program in Electrical and Electronics Engineering in Universiti Teknologi Petronas, Malaysia, PH-0192914917. Email: ashraf.balghari@gmail.com
- Airul Azha Abd. Rahman is currently working as staff researcher at MIMOS BERHAD, PH-0192193013. Email:airul@mimos.my
   Suraya Sulaiman is currently working as researcher at MIMOS
- Berhad.PH-0172380389. Email:suraya.sulaiman@mimos.my

# Before and After: User's Knowledge Maturity within Personal Information Management

Azrina Kamaruddin, Novia Admodisastro, Alan Dix

Abstract— Personal Information Management (PIM) is an important discipline that evolves with the technology advancements in their operating systems. The understanding of its four main activities in PIM: organizing, retrieving, deleting and archiving are crucial because the information growth in our desktop increased overtime. The paper presents the study conducted to enhance understanding of one of the causes for core phenomena – user's knowledge maturity which was identified in prior studies. The qualitative approach is taken to deepen our understanding on this cause. As a result, we found that several issues pertaining to time and users' task have increased their knowledge on organizing-keeping information on their desktop.

Index Terms— Users' knowledge maturity, Personal Information System, Keeping behavior, Qualitative analysis, Longititidal approach, Grounded theory

#### **1** INTRODUCTION

Personal Information Management (PIM) is an activity that involves information acquisition, retrieving, maintaining and deleting. It has grown over time [1] due to the large increase in digital information that we encounter in our working and personal lives. Boardman [2] classifies it as an "umbrella" that managed digital information such as email, bookmark, files, meetings and appointments in a personal computing environment [3,4]. PIM is an important aspect in our life, where studies have reported that people always consider PIM as a chore [5,6,7,8,9] just like we possessively take control of our belongings. In order, to design tool that asissts users with their information on the desktop, we first need to acquire user's specification and to understand the issues regarding the PIM activities before a suitable design tool is proposed. The application and combination of Human Computer Interaction (HCI) discipline is important for Software Engineer to provide a user-centred design PIM tool for the intended users [9].

HCI defined as a set of "processes and resources that the users employ to interact with the computer" [10]. It is crucial to understand the users' experiences and behavior (a set of processes) when they interact with their computers/machines. In our study scope, we would like to investigate and understand the users' behavior exhibit within the PIM important activities. Several studies reported Malone [5], Lansdale [6], Barreau and Nardi [7], Whittaker and Sidner [8] and Jones et al. [9] discover that people are still struggling to manage their information on their desktop, file, email and bookmark even though they have been supported by latest technology. The challenges in PIM give impact to work productivity [7;11;12] and user experience [13].

Early research works on PIM focus at particular domain such as hierarchy [14], email [15], and bookmark [16]. Gradually, several studies focused on interaction across our Personal Space of Information (PSI) have been conducted and reported such as [14;17;18;19]. However, there is little study on the keeping activities within PIM frameworks.

A study has been conducted as an extension of further investigating one (i.e. user's knowledge maturity) of the causes resulted from prior studies conducted. Table 1 shows a summary of studies that we had conducted to understand the core phenomena identified in re-analysed Study I. We have developed an understanding of keeping activities specifically the core phenomena that we named as the notion of dumping. The notion of dumping is defined as "a procrastination decision [20]". Both decision making and 'dumping' behavior are inter-related [20].

The objectives of this study are:

- 1. To observe user's maturity in knowledge related to their task.
- 2. To use the user's maturity knowledge resulted with their frequent task creates better understanding of their action especially managing their PIM.

Study	Method/	Analysis	Outcome
	Techniques	Technique	
Study I	<ul> <li>Interview</li> <li>Observation</li> <li>Capture images</li> </ul>	Informal analysis	Build preliminary model of PIM user's behaviour
Study II	<ul> <li>Diary study</li> <li>Simple questionnaire</li> <li>Interview</li> </ul>	Informal analysis	Understanding the issues and methodological
Re- analysed Study I	<ul><li>Interview</li><li>Observation</li><li>Sketches</li></ul>	Grounded theory	Build core phenomena and its relationship model

Table 1. Summary of studies conducted

Azrina Kamaruddin is pursuing a PhD degree program in Human Computer Interaction in Lancaster University, United Kingdom, PH-60389471763. Email: azrina@fsktm.upm.edu.my

Novia Admodisastro is currently a lecturer in Information System department, Universiti Putra Malaysia, Malaysia, PH-60389471755. E-mail: novia@fsktm.upm.edu.my

Alan Dix is a professor at Talis Birmingham, UK and University of Birmingham, UK, E-mail: alan@hcibook.com

Study III	٠	Interview	Grounded	Detail under-
	٠	Observation	theory –	standing
	٠	Partial longi-	until axial	Elaborate on the
		tudinal study	coding	causes model
	٠	Capture im-	_	from main model
		ages		

The paper presents the work that has been conducted in Study III only, which starts with introduction, literature review, methodology, analysis, results and conclusion.

# **2 LITERATURE REVIEW**

There were many studies about information seeking related with domain knowledge such as domain knowledge on search tactic formulation [21], distinguished subject familiarity [22], topic knowledge of the search topic [23]. Recently, a study about finding and refinding based on the effect of task domain by Toms et al. [24] and factors and evaluation of refinding behaviours by Capra and Quinones [5;25]. However, there is yet little study about user's knowledge maturity in the information organization scope.

According to Capra [5;25] domain knowledge is part of knowledge where the user are familiar with the topic. He stated that the user only search information based on their domain in general. Another category of knowledge is stated by Capra [5;25] as task knowledge that measures knowledge about how to do a specific task. The factors that contribute to the task knowledge are familiarity and frequency. Capra [5;25] also reported in his thesis the effects of frequency and familiarity in refinding information.

Our study is about how the user's domain knowledge affects their information management understanding when their domain knowledge increase. We defined the users' knowledge maturity as the users' knowledge about information hierarchy on their desktop. Level of the user knowledge is changed before their knowledge of their task increased and after their knowledge of their task reached certain maturity level. The maturity supports their decision making upon organisation or retrieval of their information. Increasing in user knowledge also help them to decide about their information dumping either in certain location is only for temporary or information is suitable to categorise in specific category.

# 3 METHODOLOGY

This study was conducted nearly a year after the first and second studies. Throughout Study I and Study II, we identified few users clearly mentioned about their experiences and maturity that evolved when handling their new task. This indicates that the participants have increased understanding about their PIM. For example, Participant P002 from Study I reported that her first involvement in the study is as a first year PhD student which is new to the research domain.

The following section 3.1 describes in details the implementation of Study III.

# 3.1 Longitudinal study

Data collection for a given period of time can be either retrospective that is asking participant to reflect back upon their experiences and attitudes, or contemporary that is collecting data at different times about the current situation. The contemporary longitudinal study involves repeat follow-ups of a single sample, panel or cohort. This is more favoured approach in conducting Study III. Although retrospective data collection is an important part of any study, longitudinal research usually focuses on short-term reflection to avoid deterioration of reliability and validity when asking respondents to reflect back over long time times.

In addition a case study is designed as follow:

- We only select two participants from Study I.
- The two participants that we select are: one uses spring cleaner information management strategy and the other uses no filer information management strategy.
- These two participants used intensive dumping behaviour based on their information management strategy.

This case study consists of some aspects of longitudinal study in terms of time frame and changes over time in people behaviour [30]. The third study is conducted few years after the first and second studies. In order, to observe user knowledge maturity has an impact on user core phenomena, the notion of dumping, we purposely selected two similar users from different information management strategy to be Study III participants. The two participants were selected in Study III because in prior studies they were very co-operative and easy to communicate. We refer the two participants in Study III throughout this paper as S3-01 and S3-02. Further information about the participants is presented in section 3.2.

# 3.2 Naturalistic Approaches

In this case study, again we gathered information using naturalistic observation. We made the observations at their office desktop where they always do their jobs [31]. The observation on their desktop focused on their dumping files (e.g. folders, single information) that they mentioned earlier in the first study. However, the observations and interview were about their understanding and description of their files when they were more experiences with the task that they performed during study I. When interviewed, we could see the increased of knowledge about their desktop. Hence, they could provide reasons why they acted certain way when it they were at task versus time or urgency.

The interview has been conducted. The questions we asked were as follows:

- What is happening on their desktop (email and bookmark)?
- Why is the transition of desktop interface (email and bookmark)?
- When is the transition happen?
- When so they feel that they are getting matured in understanding their desktop?

• What happen to their information management before, during and after they achieved certain knowledge maturity.

Interview technique was selected when we wish to obtain more detailed and thorough information on the issue of user's maturity that might be gleaned from our GT analysis. We often prepared several questions, but also deviated where necessary in order to maximise the information obtained [26].We recorded our interview and transcribed the interview sheet. The study was conducted at the participants' workspace. Interview the participants was still our preferred approach because we wanted to acquire qualitative data [27] to describe changes in user's information management behaviour. The interview duration is between 1 – 1.5 hours. The images of users desktop screen was captured and compared with the screens captured in Study I.

We audio-recorded the interviewed and captured few images on the participant's computer screen. Then, we transcribed the interview. S3-01 is a non-native English speaker while S3-02 is a native English speaker. There was no translation made for the non-native English speaker transcription.

# **Participants Selection**

The participants in Study III were taken from Study I. We decided to select two participants that we observed and identified from the first study as active participants. In Study I, the participants were referred as P002<sup>1</sup> and P019<sup>2</sup>, whilst in Study III we referred them as S3-01 and S3-02 respectively. The participant S3-01 used spring cleaner, while the participant S3-02 used no filer information management strategies when managing information on his desktop.

## a) Participants background

Study III was only conducted by interviewing two participants that we purposely selected from Study I. We decided to choose them as the participants because they were very responsive during Study I. Even though, we only have two participants, it would not give problem for us to understand the issues. In fact, we could study them properly and deeply [28]. It is important to know the participants' chronologies personal backgrounds before we proceed with data analysis process. Table 2 shows summary description of the participants.

Table 2. Study I	III partici	pant descri	ptions
------------------	-------------	-------------	--------

	Participants Id		
	S3-01	S3-02	
Level of	High compotency	High compe-	
Competency	Fligh competency	tency	
Information			
management	Spring cleaner	No filer	
strategy			
Native	Non-English	English native	
Sex	Female	Male	
At time of Study	Writing up PhD	Postdoc (data	

<sup>1</sup> P002 referring to participant who used Window OS as her operating system. <sup>2</sup> P019 referring to participant who used Mac as his operating system

III		analysis task)
At time of Study	2nd woor PhD	Postdoc (litera-
Ι	2 <sup>nd</sup> year PhD	ture review)

S3-01 was a female PhD student. She is also a mother and currently in writing stage of her PhD. She is an international postgraduate student and her native language is Bahasa Melayu. Her husband and children were all in UK. She is a valuable and main contributor to the Study I and Study II. She was in second year PhD when we interviewed her during Study I.

While, the participant S3-02 was a postdoc researcher. He was also our participant that involved in Study I. He had an experienced being a PhD student and this is an advantage for him. He knew that experiences and skills that he gained during PhD was not only about the research area but also skills and experiences managing information to complete the task. He was an international post-doc and his native language is English.

#### b) Participants Workspace

In this study, the important entity that we observed and analysed was their desktop. Both of our participants were at different stage when delivering their main tasks. S3-01 was nearly completed her PhD thesis whereas, S3-02 was at the analysis stage of his main task. While doing the interview session with them, we also observed their desk. They talked about their desk management especially participant S3-01. During the first interview, we saw her desk was occupied and scattered with papers, articles and books. It was a partially unsorted desk. However, when we observed her desk again in Study III, paper, articles and books were sorted and stacked nicely on her desk. She said that "Ohm...that is where all information has been sorted and cleared because she had used and referred to that stack few months back, now, might not used it anymore".

While, participant S3-02 desktop was quite tidy in both studies, however his desktop interface was dynamically changed due to his task. His desktop changed from a cluttered interface to a tidier and categorised interface. When questioned about it? He reported that, "*The reading and finding article stage has stopped and completed at this time, now is the time where you start to get your hands dirty with da-ta. Once a while, you might need to refer to your paper and article so you just find it in the category. I already categorised them under appropriate label.*"

We interviewed both the participants at their workspace. The interview was recorded and transcribed. They were still at their workspace provided by their department. However, participant S3-02 moved to a new bigger office room, in which the room contains one experimental lab with glass window. Nevertheless, the things in the room were not much different from the one we observed in Study I.

While another participant S3-01 was at the same office room but the condition was much neater. This participant

desk was tidier, because she is nearly finished her thesis writing. Papers were sorted and stacked on her desk, she said that they were done. She reported that she categorised all her printed papers in the folder. The physical folder labels were similar with the digital folder version. The folders were nicely sorted in the folder. However, at the moment folder that was currently used was unordered.

# **4 STUDY III ANALYSIS PROCESS**

The data analysis technique selected was GT. Given that, we have experiences with the technique, we decided to applied few stages in GT. We only interested with categories that related with the participants maturity in their knowledge. We started with open coding in GT or finding concepts and categories [29] that related with users' knowledge maturity. The process is about users' knowledge maturity with time when we integrate using axial code. In this analysis, we aim to:

- Understand users' knowledge maturity in information organization especially in the context of core phenomena of notion of dumping.
- The understanding is arranged towards time factor e.g. before maturity, during maturity and matured.

#### 4.1 Results

The issues pertaining to time and users' task are important properties for users' knowledge maturity that cause the core phenomena the notion of dumping. In this section, we showed the in-depth understanding of users'knowledge maturity on how they behave before and after they achieved knowledge maturity.

#### Achieving Users' Knowledge Maturity

Users' knowledge maturity is very much related with time and users' task. Most of the participants in studies conducted in our research were PhD students varied in their intakes. Some were in their first year while some were finishing. Besides PhD students, the participants were academicians or researcher. Clearly, in our observation, almost all PhD students are having problems in managing their information on the desktop and adapting with their new role as a PhD student. According to both of Study III participants, they reported that when they were reaching their final stage in PhD they realized that they had discovered many skills such as increased their knowledge vocabulary within their new task and dynamic information management strategies. These are due to the participants increase in knowledge of understanding in their research area.

As a first year PhD student, there are many skills need to be learnt and a lot of adjustment to be made in the research area. Being a research student that dealt with many reading, assessing much information affected their information management. However, when they have reached maturity in their research area or task, their knowledge on information management has also increased. There is no course taught on how to manage information in PSI. The skills is known through participants experiencing themselves and as a result based on what they feel comfortable and applicable about it. Next sub-section, we presented our analysis based on the users' knowledge maturity evolve with time.

#### a) Pending Categorisation

Categorizing information is one of the techniques primarily to keep information. However, pending Categorisation is when users are busy in delivering their primary task, categorising is the last option to do. Users' only focus in completing the task/sub-task rather than involving their effort and cognitive to categorise their information.

#### b) Early categorisation issues

When users are new to the task and knowledge domain, they have vague idea of how their hierarchy tree on their desktop will look like or how they will structure their files and folder due to their shallow knowledge about the new task. They admitted that, they have problems if they quickly categorised information but after they gained understanding of their new task and knowledge they allow retrospective categorisation. They have increased their vocabulary of the knowledge and also their folder labels.

At times, they have problems in retrieving information because they thought that they had moved the files into new folder. This uncertainty of being more knowledgeable is an issue when user is becoming dynamic in their personality pertaining to their categorisation of information. However, both of the participants are aware about their changes in making decisions on categorising their information at early stage or new to their task. When they became matured, they have the desire and tendency to change information in their old folder into new folder based on their maturity.

c) Maturity in Research Knowledge Change Decision Criteria

When users achieved certain level of knowledge maturity, they also have reached maturity in making decision. In addition, earlier technique that was used to make decision also changed or extended. For example, in Study I, one of the participants read the abstract of the paper to decide whether to save the paper or not. When, she becoming more knowledgeable in her research area, she decided that in order to save information, reading an abstract was not enough. She extended her reading technique from abstract section to discussion and conclusion sections in the paper. When, she decided the paper is relevant, then she continue reading the rest of the paper sections and finally save the paper.

### d) Better Picture of Personal Space of Information

Getting matured in their research area or task has made the participants to give a better picture and understanding on what they are doing at that time (i.e. achieving maturity). At the beginning of their task and managing information, they do not have a clear picture of how the hierarchy structure will look like, and what information that they need to search, save and etc. When, they nearly achieved understanding of their task and knowledge they know what they are doing. Even though, they dumped information in a partial category and in 'My Document' or desktop, they know what these information are and why they do that and later they know where the information should go.

One of the participants reported that, he was able to manage the way he did now because of the experiences he had during PhD study. Based on this experience, he applied the information management skill until now.

#### e) Early Categorisation Advantages

Categorisation makes our PSI looks tidy and organize. User with instant filer strategy definitely categorised their information once they acquired information. However, our participants are from no filer and spring cleaner strategies said that quickly categorisation would cause few problems when they reached maturity. So, some of the participants prefer to dump information and later make a decision of where it should be located/categorised. Their PSI looks cleaned and tidy.

# f) Understanding Dumping Action

Before participants becoming matured in their knowledge or research area, their information management especially their information organisation may have been structured or unorganized as in dumping. At the early point of time, participants did not get a clear idea where exactly the information gathered need to be categorised. One of the participants uncertain about the location and did not have a clear picture of where it should be located.

However, when the maturity increased, they understood better their information and where it should be located. They also reported that they knew their reasons why they dumped information. Reasons are compliment with Study I [20] because they are running out of time, they are focusing on their primary task and by thinking of where and what label will distract their main focus on delivering their main task.

When they became knowledgeable in their research area, they knew exactly what is happening in their information management. As a result, they were becoming more knowledgeable on what are going on their PSI. Some of the behaviour that matured is: their judgement on keeping information increased and knew what to find, how to find and where to read in the content. Furthermore, they know that for example this information is only temporarily and later it will be properly categorised and kept. If they came back looking at their scattered information unfiled, they knew why and when it will get filed. In addition, they also know why they dumped the information and how they will organise the information later. As evidence, through our observation for the second time with the participants they can easily said without feeling uncertain about their information.

Excerpt: ...Katakan pasal friendship. Saya dulu tak ada folder pasal Friendship. Actually, I just can just create right away tapi

is like I feel like tak ada masa jadi dump saja sini. Katakan saya ada 1 chapter cakap pasal Friendship. Jadi tengah-tengah tulis ke, tengah-tengah baca ke and then I need to get that information. Jadi saya tahu I dumped somewhere here that information [laughing]. So saya cari ajela kat sini. Saya tengok macam ni kalau benda tu tak susun lagi ke dalam folder maknanya saya dump kat sini semuanya dump kat sini, saya pun dump aje la kat sini. Tapi kalau misalnya kalau benda tu katakana dah ada folder. Saya masuk jelah tapi itu yang normal condition tapi kalau saya nak rush ke walaupun ada folder tapi dump juga dekat sini sebabnya nantilah baru buat housekeeping sekarang ni tak buat lagi. Apa tu yang saya dump aje la tapi macam kita tahu macam mana nak cari kat sini, aaa apa tu bila kita dump kat mana tapi mungkin kita tak tahu exactly laa... apa benda yang itu kena manually search la pulak..... (S3-1 nonnative English speaker)

# The translation:

...If it is about friendship. I don't have any folder about friendship. Actually, I can just create the folder Friendship straight away, but I just don't have time, I dumped the information regarding Friendship somewhere. When I need the information, if by browsing the hierarchy and if it is not there, I know I dumped somewhere outside that particular folder...

# **5** CONCLUSION

The study reveals that user's knowledge maturity affects user's PIM especially in the dumping and organisation scope. When users are becoming matured they not only understand more about their task and knowledge but also their information organisation. At this point, even though they dumped their information, they dumped in partially category where they think the information should reside. Later, they will recategorise the information in a proper category. They also know the information dumped is for temporary purposes. They are uncertain to categorise at the early stage because they are not sure on how the information structure looks in their desktop.

# REFERENCES

- C. C. Marshall and W. Jones, Keeping Encountered Information. Communication ACM, 49(1): 66-67, 2006.
- [2] R. Boardman, Improving Tool Support for Personal Information Management, PhD Thesis, Dept of Electrical and Electronic Engineering, Imperial College London, 2004.
- [3] T. W. Malone, How Do People Organize Their Desks?: Implications for the Design of Office Information Systems, ACM Transactions on Information Systems (TOIS), 1(1):99–112, 1983.
- [4] O. Bergman, R. Beyth-Marom, and R. Nachmias, The User-Subjective Approach to Personal Information Management Systems, *Journal of* the American Society for Information Science and Technology, 54(9), 2003.
- [5] R.G. Capra and M.A. Perez-Quinones, *Factors and Evaluation of Refinding Behaviors*, in Proceedings of the Workshop on Personal Information Management (SIGIR'06), 2006.
- [6] M. Lansdale, The Psychology of Personal Information Management, Applied Ergonomics, 19(1):55–66, 1988.

- [7] D. Barreau and B. A. Nardi. Finding and Reminding: File Organization from the Desktop, ACM SIGCHI Bulletin, 27(3):39–43, 1995.
- [8] S. Whittaker and C. Sidner. *Email Overload: Exploring Personal Information Management of Email*, in Proceedings of the SIGCHI Conference On Human Factors in Computing Systems, pp. 276–283. ACM Press, 1996.
- [9] W. Jones, H. Bruce, and S. Dumais, *Keeping Found Things Found on the Web*, in Proceedings of the 10<sup>th</sup> International Conference on Information and Knowledge Management, pp. 119–126. ACM Press, 2001.
- [10] D. Te'eni, J.M. Carey and P. Zhang, *Developing Effective Organizational Information Systems*, Jones Wiley n Sons Publisher, 2006.
- [11] A. Sellen and R. Harper, *The Myth of the Paperless Office*, MIT Press, 2001.
- [12] W. Jones, Finders, Keepers? The Present and Future Perfect in Support of Personal Information Management, First Monday, 9(3), 2004, retrieved March 10, 2004 from

http://www.firstmonday.org/issues/issue9\_3/jones/index.html.

- [13] V. Bellotti and I. Smith, Informing the Design of an Information Management System with Iterative Fieldwork, in Proceeding of the 3<sup>rd</sup> Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, 2000.
- [14] R. Boardman, Category Overlap between Hierarchies in User Workspace, in Proceeding of the 8th IFIP TC 13 International Conference on Human-Computer Interaction (Interact), 2001.
- [15] N. Ducheneaut and V. Bellotti, E-Mail as Habitat: An Exploration of Embedded Personal Information Management, *Interactions*, 8(5):30– 38, 2001.
- [16] R. Boardman, *Multiple Hierarchies in User Workspace*, in Proceeding of the CHI '01 extended abstracts on Human Factors in Computer Systems, pp. 403–404. ACM Press, 2001.
- [17] D. Abrams, R. Baecker, and M. Chignell, *Information Archiving with Bookmarks: Personal Web Space Construction and Organization*, in Proceedings of the SIGCHI conference on Human factors in computing systems, pages 41–48. ACM Press, 1998.
- [18] R. Boardman, Workspaces that Work: The Need for Integrated Personal Information Management, in Proceedings of the 6<sup>th</sup> Human Centred Technology Postgraduate Workshop, 2002.
- [19] R. Boardman, Workspaces that Work: Towards Unified Personal Information Management, in Proceeding of the HCI2002, People and Computers XVI - Memorable yet Invisible, Volume 2, 2002.
- [20] A.Kamaruddin and A. Dix, To 'Dump' or Not To 'Dump': Changing, Supporting or Distracting Behavior, in Proceedings of the International Confrence on iUser, pp. 49-54, 2010.
- [21] M., Wildemuth. The Effects of Domain Knowledge on Search Tactic Formulation, *Journal of the American Society for Information Science and Technology*, vol. 55(3): 246-258, February, 2004.
- [22] M.J. Bates, Factors Affecting Subject Catalog Search Success, Journal of the American Society for Information Science, vol. 28: 161-169. May, 1977.
- [23] B.L. Allen, Cognitive Research in Information Science: Implications for Design. Annual Review of Information Science and Technology, 26:3– 37, 1991.
- [24] E.G., Toms et al., *The Effect of Task Domain on Search*, in Proceedings of the Conference of the Centre for Advanced Studies on Collaborative Research (CASCON'03), IBM Press, pp. 303-312, 2003.
- [25] R.G., Capra, An Investigation of Finding and Refinding Information on the Web, PhD Thesis, Computer Science and Applications, Virginia Polytechnic Institute and State University Blacksburg, Virginia, February, 2006.
- [26] A. Adams and A.L. Cox, "Questionnaires, In-Depth Interviews and Focus Groups", in C. Paul and A.L. Cox (eds.), Research Methods for Human Computer Interaction, Cambridge University Press, pp. 17–34, 2008.
- [27] M.D, Myers, Qualitative Research in Business & Management, SAGE Publication Ltd., 2009.

- [28] F.H. Abd Razak, Single Person Study: Methodological Issues, PhD Dissertation, Computing Department, Lancaster University, Lancaster, United Kingdom, 2008.
- [29] B.G. Glaser and A.L. Strauss, The Discovery of Grounded Theory: Strategies for Qualitative Research, Chicago: Aldine Publishing Company, 1967.

# Overhead Analysis as One Factor Scalibility of Private Cloud Computing for IAAS Service

# Teguh Bharata Adji, Faisal Suryadi Nggilu, Sujoko Sumaryono

Abstract— The rapid development of technology leads to improvement in a variety of computing concepts, information and communication technology(ICT). The ICT environment is become more complex and costly. Cloud Computing is one of phenomenon int the new ICT services, Cloud Computing model has three categories of services : Infrastructure As A Service (IAAS), Platform As A Service (PAAS), Software As A Service (SAAS). It also has four deployment models which are private cloud computing, community cloud, public cloud and hybrid Cloud.

This research aims to determine the overhead of the virtualization environment. It is expected the private cloud with virtual technology, that utilizes the maximum resources does not degrade server scalability. The implementation of private cloud is using OpenStack with configuration multiple interfaces multiple servers.

The results of the thesis indicate that the overhead of a single virtual machine is 114 ms (database server) and 212 ms (web server), to ten virtual machines are active 615 ms (database server) and 786 ms (web server). Overhead that occurs on a single VM can still be neglected, despite the performance degradation with increasing number of active virtual machines the execution time of servers application tends to be linear close to shape belongs to the physical servers.

Index Terms— Overhead, private cloud computing, IAAS, OpenStack, multiple interfaces multiple servers.

---- 🔶

# **1** INTRODUCTION

As the rapid development of technology leads to improvement in a variety of computing concepts, the information and communication technology (ICT) environment becomes more complex and costly. Provision of reliable infrastructure have consequences for the ICT users to be able to use it effectively and efficiently.

Cloud computing is one of development phenomenon in the new ICT services. Cloud Computing is a model which provides computing resources or information technology (e.g., software, processing power, storage.) that allows customers to "take and use" resources on-demand, via the internet (NIST, 2011).

With the cloud computing, it will change the paradigm of the company or organization in ICT investment. Cloud computing can be promoted to consolidate ICT infrastructure, simplify management, maximizing the utilization of computing resources and at the end it would be able to save the operational costs.

Infrastructure As a Service (IAAS) is a service that "rent" the basic technology information resources, which includes media storage, processing power, memory, operating system, network capacity, etc. which can be used by users to run applications virtually.

The implementation of virtual technology that utilizes the maximum resources of server, is expected to reduce its scalability to the virtualization environment such as overhead.

# **2 LITERATURE REVIEW**

Cloud computing is one of development phenomenon in the new ICT services. Cloud Computing is a model which provides computing resources (e.g., software, processing power, storage.) that allows customers to "take and use" resources ondemand via the internet (NIST, 2011). There are four proposed ontology which is expected to be a framework and guidance in conducting research in cloud computing. This ontology consist of cloud application (software as a service), cloud software environment (platform as a service), cloud software infrastructure (infrastructure as a service, data as a service, communication as a service), software kernel, firmware / hardware ( hardware as a service) (Youseff, 2008).

Infrastructure As a Service (IAAS) is a service that "rent" the basic information technology resources, which include media storage, processing power, memory, operating system, network capacity, etc. which can be used by users to run applications virtually. In a virtualized environment, (Nurhaida, 2009) performed overhead measurement, linearity, performance isolation and the use of hardware resources in a virtualized environment. The research used VMWare ESX 3.5.0 as virtualization application.

#### 2.1 Virtual Server scalibility

\_\_\_\_\_

Scalability is the ability to improve system performance gradually in accordance with the work load by increasing the number of resources. To evaluate overhead due to the virtualization mechanisms, we compare the execution time of an application running on a non-virtualized operating system (Ta) with another execution time of the same application run within a single virtual machine (Tav). The overhead may be negligible for a virtual machine and becomes significant when several virtual machines running at the same time (context switch overhead may exist even if no application is executed on the other virtual machines) and the formula is :

Meanwhile, Ta is also compared with Tanv when some virtual machines n running at the same time. In this scenario, only a virtual machine actually runs the application. The other n - 1 VMs are not running any application.

where Ovn = Virtualization Overhead and Tanv = Executing time for n virtual machines.

# **3 RESEARCH METHODOLOGY**

Two scenarios used in this study has been proposed. Previously, resource analysis study has been conducted to run the scenario.

• For the first scenario, private cloud computing was implemented with Infrastructure As A Service (IAAS) in MTI network as shown in Figure 3.1. The design of private cloud consisted of four computers, which consist of one master computer and three slave computers. The configuration is using multiple interfaces of multiple servers.



Figure 3.1 The design of Private Cloud Computing

- The specification of master computer is 8 core processor, 8 GB memory, Hardisk 1 Terabyte and two Network Card. The master were installed and configured with OpenStack Compute and Glance.
- The specification of slave computer is 4 core processor, 8 GB memory, 1 Terabyte Hardisk and two Network Card. The slave were installed and configured with Nova Compute and Nova Volume.
- In the second scenario, overhead will be measured. Network topology for testing overhead can be seen in Figure 3.2. In this scenario, the execution time between physical servers and virtual servers will be compared. The application will be installed in both servers (Database Server i.e. mysql server and webserver apache2). The server were installed with ubuntu server 10.04 (Lucid Lynx) 64 bit operating system, and the specifications of both servers are 8 processors, 8 GB memory and 1 Terabyte hardisk. Every server will be given different loads as follows:
- Load for database server by backuping data of 146 MB taken from mediawiki database downloaded

from http://opensourcetelkomspee dy.com / wiki / index.php / Main\_Page.

- Load for web server by http request as many as 5000 requests
- Request process will be done by using apache benchmark tool which is available in apache2 applications.





#### **4 RESULTS AND DISCUSSION**

Based from the testing process, the data was then collected for further analysis, as shown on Table 4.1.

<b>Table 4.1</b> Evaluation of the overhead with loading scenarios
on the Server

Skenario pembebanan database server	Jumlah proses	Native (server fisik)	Virt	tual <u>Mesin</u>
Backup database 146	1	0m3.903s	1	0m4.017s
MB (Databaseserver)			2	0m4.085s
		1	3	0m4.271s
			4	0m4.299s
			5	0m4.439s
			6	0m4.335s
	3.	32 X	7	0m4.530s
		3	8	0m4.554s
			9	0m4.529s
			10	0m4.518s
5000 http request	1	0m8.734s	1	0m8.946s
(webserver)			2	0m9.467s
			3	0m9.375s
			4	0m9.478s
			5	0m9.400s
			6	0m9.528s
		2	7	0m9.544s
		1	8	0m9.446s
			9	0m9.473s
			10	0m9.520s

Based on data in Table 4.1, a graph of database and webserver measurement results is obtained as shown in Figure 4.1 and Figure 4.2



Figure 4.1 Graph of overhead measurement on mysql server



Figure 4.2 Graph of overhead measurement on the webserver

The execution time of the computing process in one physical server is correlated with that of in one overhead server. It can be seen from Figure 4.1 and Figure 4.2 that if one virtual server is running then the overhead is almost zero and can thus be neglected. However, when the number of virtual machines increases, the execution time of database and web server applications in more than one virtual servers tend to be linier and the shape of the curve close to shape belongs to the physical servers. From the measurements which have been done, the overhead values are obtained as follows:

- The overhead of the database server process are as follows. Ov = 114 ms and Ovn value = 615 ms. Database performance degradation after ten virtual machines run = 18.54%
- The overhead of the web server process are as follow. Ov = 212 ms and Ovn value = 786 ms. Webserver performance degradation after ten virtual machine run= In addition, the use of virtualization server with OpenStack has increased the optimalization of processor usage, from the level of 0.1% in the presence of an active virtual machine to 99.8% with 4 actives virtual machines. The memory usage also has been increased from 714 MB to 6731 MB as can be seen in Table 4.2. These values are obtained from the resource monitoring tool namely Top software in one of the slave. Hence, the use of virtual technology is a promising infrastructure that is reliable and optimizes all resources belong to servers.

# Table 4.2 The optimization of the processor and memory usage

Number of active VM	Prosesor usage (%)	Memori usage (MB)
No VM is active	0,1	714
1 VM is active	25,1	6678
2 VM are active	49,9	6694
3 VM are active	74,9	6711
4 VM are active	99,8	6731

Another OpenStack ability is its fast system recovery during the system error by using the snapshot feature (i.e. backup and recovery process). Application of the snapshot can reduce almost half of the stages of recovery steps and saves time significantly. The snapshot result can be seen in Figure 4.3.

root@av	van1:~# euca-desc	ribe-images				
IMAGE	ami-0000000a	snapshot (mas	teradit) m	ycloud avail	able	private
IMAGE	ami-00000009	snapshot (nat	ivebaru) creemy	ycloud avail	able of 36.phg	private
IMAGE	ami-00000008	snapshot (eri	ckpovray) m	ycloud avail	able	private
IMAGE	ami-00000007	snapshot (fai	salnative) m	ycloud avail	able	private
IMAGE	ami-00000005	snapshot (mas	terpovray) m	ycloud avail	able	private
IMAGE	ami-00000004	bucket/lucid-	server-cloud:	img-amd64.in	ng.manifest	.xml
nstance	e-store Screenshot-40.					

Figure 4.3 Snapshot results in OpenStack

# **5 CONCLUSION AND FUTURE WORKS**

# 5.1 Conclusion

- Private cloud computing with infrastructure as a service (IAAS) with multiple interfaces multiple servers has been successfully designed and built by using 4 servers and OpenStack in Master of Information Technology (MTI) Lab, Faculty of Engineering, Gadjah Mada University, Yogyakarta.
- The overhead values for one virtual machine are 114 ms (database server) and 212 ms (webserver). Mean while for ten active virtual machines are 615 ms (database server) and 786 ms (the webserver). Overhead that occurs in one virtual machine can be ignored and system perfomance degrades as the number of active virtual machines increase.
- Overheads of the webserver and database server applications in more than one virtual machine tend to be liniear and the shape of the curve close to the shape belongs to the physical servers.
- OpenStack server virtualization has increased the optimization of the processors from 0.1% in the absence of virtual machine to 99.8% with four active virtual machines. The memory usage has also been increased from 714 MB to 6731 MB.
- The use of snapshot provides a significant improvement of the recovery system. If there is error in the system or server error then snapshot technology can restores quickly.

# 5.2 Future works

• One of the drawback of virtualization is that all services are centralized in a machine. If one machine is damage or there is

error then it will fail to run all systems. This can be overcome by creating a Fail Over Server mechanism as a backup, using a tool called Heartbeat.

- The use of the GPU (Graphics Processor Unit) in the cloud system for parallel processing such as rendering.
- Development of the portal to software services (i.e. rendering -http://www.blendercloud.net/)
- Performance analysis of network communications in the cloud computing environment.

#### REFERENCES

- Johnson., Jha, A., Murari, K., Raju, M., Cherian, V., Girikumar, Y (2011). OpenS tack Beginner's Guide – UEC Edition (Ubuntu Server 11.10-Oneiric Ocelot). CSS Crop.
- [2] National Institute of Standard and Technology NIST (2011). The NIST Definition of Cloud Computing (Draft)
- [3] Nurhaida, I. (2009). Pengukuran overhead, Lineeritas, Isolasi Kinerja dan Penggu naan Sumber Daya Perangkat Keras pada Server Virtual Fakultas Teknik Universitas Indonesia.
- [4] Nurmi, D., Wolski, R., Grzegorczyk, C., Obertelli, G., Soman, S., Youseff, L., Zagorodnov, D. The Eucalyptus Open-source Cloudcomputing System, Computer Science Department University of California,Santa Barbara, California 93106.
- [5] Openstack (2011), OpenStack Compute Administration Manual http://docs.OpenS tack.org/diablo/OpenStackcompute/admin/content/ (access: 10 October 2011).
- [6] Peng, J., Zhang, X., Lei, Z. (2009). Comparison of Several Cloud Computing Platforms Second International Symposium on Information Science and Engineering.
- [7] Pepple, K. (2011). Deploying OpenStack, O'Reilly Publishing Services United States of America.
- [8] Purnoma. Membangun Virtual PC dengan VirtualBox. Penerbit Andi Yogyakarta. Yogyakarta 2010.
- [9] Sugianto, M. (2011) Mengenal partisi LVM http://vavai.com/2011/09/22/ berkenalan-dengan-lvm (diakses tanggal 16 januari 2012)
- [10] Sugianto, M. (2011). Panduan Virtualisasi & Cloud Computing pada Sistem Linux, Bandung revisi2
- [11] Wang, Lizhe. (2008). Virtual Environments for Grid Computing Dissertation, Universität Karlsruhe (TH) Fakultät für Maschinenbau,
- [12] Yousef, L., Butrico, M., Dilma S. (2008) Toward a Unified Ontology of Cloud Computing Grid Computing Environments Workshop

Teguh Bharata Adji is currently a lecturer of Gadjah Mada University. He was graduated from doctorate degree program in Information Technology Universiti Teknologi Petronas in 2010. E-mail: adji@mti.ugm.ac.com

Sujoko Sumaryono is currently a lecturer of Gadjah Mada University. He was graduated from doctorate degree program in Information Technology Universiti Teknologi Petronas in 2010. E-mail: adji@mti.ugm.ac.com

Faisal Nggilu is previously a student of Master Program in IT Gadjah Mada University, Indonesia. Email: faisal.sn@gmail.com

# The Impact of Intergration HCI in Software Processing Models

Saad Masood Butt

Computer and Information Sciences Department, Universiti Teknologi PETRONAS, Tronoh, Perak, Malaysia saadmasoodbutt668@yahoo.com

Wan Fatimah Wan Ahmad

Computer and Information Sciences Department, Universiti Teknologi PETRONAS, Tronoh, Perak, Malaysia

**Abstract** – Users' dissatisfaction with the software used will impact the efficiency. Moreover, the lack of knowledge of users' involvement in the development of the software will cause issues to the user's later on. In the case of human-computer Interaction (HCI), it has been suggested that a user's participation and HCI concern in the application growth life-cycle (SDLC) as an important procedure for a successful program execution. However, it is still not sure to what extend user participation is important and HCI problem has been settled by system professionals. The result which is mentioned in this paper and the review opinions from the experts' point of view are taken from analysis on the value of HCI in SDLC. The objectives of the analysis are to identify the condition of the users' contribution in SDLC and to identify the HCI elements that have been settled. Results show that many of the experts have engaged the customers in SDLC, but the majority only during the need research stage. The conclusions have also shown that HCI components on performance are well resolved. However, the non-functional components such as social, environmental issues have not been highlighted by experts. This paper indicates with recommendations to further analyze the users' interest on the value of the users' contribution in the program development.

•

Index Terms— Human Computer Interaction, Software Process Models, Human Computer Interaction, SDLC

# **1** INTRODUCTION

The disappointment of the application software used to L affect efficiency. The performance of application software based work and the face of the program. Research shows that there is a debate on the use of software that allows users frustrated [4, 8, 9] frustrated user is suggested due to the decline in the value of the consideration HCI in technical growth. system In some cases, a person is not included. But only the chosen of the process of growing commercial software of HCI in software development can not be ignored, and the lack of knowledge about the involvement of users in software development continues to be a problem with customers in the [12] program, especially the disappointment of the software affects the performance of the Office of the Public and the overall well-being [13] Among the techniques developed many software development cycle. life (SDLC) is a methodology that has been approved by the General to explain the process and issues involved in the development of [1] on developing applications with data (IS), SDLC strategy focusing on the application [. 2], but recently concerns have been adjusted to the user experience [3] The dispute related to HCI in SDLC is essential. However, it does not get enough attention to the experts, it was announced that the SDLC new strategic focus of human-centered, it is necessary to confront the compliment ', they recommended SDLC called new circuits of humancentered development. [4].n 7 concludes the paper. modify the header or footer on subsequent pages. Prior to HCI research, targeting the needs of the user. HCI is seen to be constructed from the user's perspective and not from the point of view of Research on the value of HCI concerns from the perspective of the operator is still common in the software industry [5,6,7] The purpose of this paper is to analyze the effects of HCI in SDLC

from analysis expertise. identify the situation of users in SDLC and evaluating the HCI elements that have been made in the SDLC.

# **2 RELATED WORK**

#### 2.1 HCI APPROACHES

The strategy focuses on HCI and human relations and partnerships, describes what the program should do it from the perspective of the users of the restrictions, such as physical, intellectual and behavioral success. In addition, the results of which have contributed towards the growth of the program and use must also be considered. Growth between HCI 'obligations and commitments between the use of the interface and how users can interact with the program.

By Hoffer et al, [1], the SDLC current strategy focuses more information about the features and specifications of the system to meet business needs. Incredible HCI is concerned with methods and techniques to fit the needs of active lifestyles and well-being. To build information systems that meet business and personal needs and concerns HCI should be included in the strategy, especially for the development of this reason, Zhang et al [4] introduced a strategy perspective concerns HCI. The particular case of the evaluation results.

HCI containing information about the four elements which are physical, intellectual, emotional and behavior together with examples of their listing. These issues are not HCI research specifications of software development.

# 2.2 ROLE OF HCI IN SDLC

Zhang et al, [4] disputes related to HCI in SDLC is essential. However, this has not been emphasized by application experts. They argue that the SDLC new focus on humancentered strategies need to be concerned about the user experience compliments. So that tactical human-centered they recommend SDLC new system called the human-centered lifestyle (HCSDLC) Set 1 features the popular HCSDLC four steps have been regarded as an alternative investment and. Preparing the research design and the implementation is proven in Figure 1, this structure will focus on the HCI concerns, unlike the SDLC is considered. Collaboration between perspectives SDLC and HCI programs to ensure growth is achieved. However, the current SDLC too many business needs. (Functionalist) rather than individual needs. (The physical and the intellectual, the psychological character and situational factors) to each regarded much in the SDLC so often happens in the gap between the needs of business and the needs of each

# **2.3 USER PREVENTATIONS**

As defined by Lawson [4], Disappoint the software is used. "The emergence of barriers that would prevent the need to satisfy the" recent reports about the disappointment of the problem behind the screen [4] and the problem of using a business website [8]. These problems arise when the software has been successfully delivered to customers and [17] Recent reports on the disappointment of a problem with the screen, and the problem of the business website. These problems arise when the software has achieved and delivered to customers.

Research more disappointed by the Besserie et al, [10] described the frustrations of using the computer to have knowledge of the client during the execution of their daily lives. The results of their research showed that one-third to one-half of the time invested prior to the application of the system is due to the use of a disappointment. Frustration can affect the performance level of compliance and the public well-being.

# **3 METHODS**

Analysis of the primary means to select a set of questions, and chosen. A series of questions were designed based on the structure of the HCI HCSDLC [4] mentioned in the previous Participation in the study was the use of the College Student Information System (ISIS) and the performance of UTP, Malaysia.

Interviews were among the end-user program in ISIS analysis of the disappointment in the program. Which included five professors and five students who are definitely used. Study on the HCI in SDLC is performed among participants 32 to a developer application and professional growth of applications such as IT, professional design, web development application, professional IT. support and systems analysts.

#### 4 Analysis

### 4.1 VIEWS OF THE SYSTEM

The conclusions from the appointment show that the participants are not fulfilled with the use of the program. Participant decided that the writing and number is readable. However, the results showed that 67% of participants were puzzled about how the data is structured. Application performance is not enough, and 89% of the participants said that the mistake was not on the program. Overall response of respondents to the program is awesome and exciting. The experts suggest that to improve the design and efficiency. Aware of this and rectifies the involvement of users as to be considered during the whole application life cycle was introduced by Bryant [17].

#### **4.2 USER INVOVLMENT**

Show that 96.7% of experts decide whether a client should be involved in the SDLC and 90.4% decided that the participation of users is very important. However, most of the reactions that they engage with customers in SDLC, results have indicated that some level of participation of users is very low, only 16.13% of a customer between growth on other hand, 77.42% of the experts were engaged customers during the research necessary to (77.42%) Further research to understand user participation has proven to be 80.7% of the participants decide that. use 'Deficit of knowledge is limited due to their relationship with customers, especially in the design and growth. These findings show that there is no need for customers to have sufficient understanding. This research has exposed serious problem than the deficit on the possibility of the involvement and the participation of users in the software life cycle.

# **4.2 HCI CONSIDERATIONS**

Two were found to be significantly related to the level of what has been fixed in the SDLC component among HCI professionals. In this study, three groups of the HCI components which are suggested by Zhang et al [4] used these groups with the purpose of the purchase and focus on the details of each. The results proved that the HCI with regards to the objectives have been resolved except for "Safe for use" Table 2 shows the specific HCI concerns, depending on the target.

Table 2. HCI Deliberation Based on Usability Goal

		-
HCI elements	%	Sum
Effective to use	74.19	23
Efficient to use	74.19	23
Safe to use	32.36	10
Easy to learn	64.52	20
Easy to remember	51.61	16
how to use		

The results show that significant performance and operational performance. But very little protection. Results of the concerns HCI, depending on the user experience are given in Table 3 is proved in Table 3 was (77.4%), response (67.7%) and friendly (51.6%) were resolved by Internet experts. But the fun does not have a profile (0%) and emotional responses (0%), indicating

that this component in the joint success principles are not being addressed well.

Table 5. Descriptive Statistics of Osers Experience							
HCI elements	%	Sum					
Satisfying	67.7	21					
Helpful	77.4	24					
Fun	0	0					
Friendly	51.6	16					
Emotionally fulfilling	0	0					
Entertaining	3.2	1					

Table 3. Descriptive Statistics of Users' Experience

That the HCI is based on the important details of each style are shown in Table 4, most of the experts prefer to identify their focus on each profile style of each individual in the system depends on the skill (. 67.7%), knowledge (61.3%) and to work (51.6%) percent of the very few to have a look at the background of social (9.7%) and sexual problems (3.2%), it is clear that the aim focus on the details of each style is highlighted in terms of the two groups.

HCI elements	%	Sum		
Gender	3.2	1		
Computer training	61.3	19		
Experience with similar	67.7	21		
systems				
Occupation	51.6	16		
Cultural background	9.7	3		

Table 4. Target User Model Profile

The analysis shows that in the world of real problem is the focus of HCI in SDLC needs efficiently and very few non functional requirements such as security, performance, and public issues.

#### 4 Conclusion

From the research, it reveals that the customer's disappointment with the system has impacted users' performance and group well-being. The outcome indicates that there are problems in the style such as information company and unhelpful mistake concept. Such style problems may have took place due to the unquestionable confirmed truth that the customers were not being engaged in the whole content management procedure. This can be confirmed by the outcome from the research on users' participation which reveals that the customers were only being engaged at the beginning on to be able to collect system need and at the later level of program development procedure as to confirm and confirm their need. The outcomes have confirmed that all associates have had engaged customers in their SDLC. However, the users' participation is mainly targeted on the need research level, and only a few have had engaged customers in the style and development level. The outcome also indicates that due to the users' lackof particular understanding in SDLC may have described why users' participation is still little particularly during the style and development level. The outcomes on HCI problem have exposed that the non-functional specifications such as group and efficient problems have not been given enough concentrate by experts. From these outcomes it is suggested to further evaluate the users' participation and its importance in the content management.

#### REFERENCES

- P. V. Jani," Security within Adhoc Networks", Position Paper, PAMPAS Workshop, Sept. 16/17 2002, London.
- [2] Hoffer, J.A., George, J.F., and Valacich, J.S. "Modern Systems Analysis and Design", (4th ed.), Upper Saddle River, NJ: Prentice Hall, 2005.
- [3] Hirschheim, R., and Klein, H.K. "Four Paradigms of Information Systems Development". Communication of the ACM, vol. 10, no. 32, pp. 1199-1216, 1989Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.
- [4] Zhang, P., Benbasat, I., Carey, J., Davis, F., Galletta, D., and Strong, D. "Human Computer Interaction Research in the MIS Discipline". Communication of the AIS, vol. 20, no. 9, pp. 334- 355, 2002
- [5] Zhang, P., Carey, J., Te'eni, D., and Tremaine, M. "Integrating Human-Computer Interaction Development into the Systems Development Life Cycle: A Methodology". Communications of the Association for Information Systems, vol. 15, pp. 512-543, 2005.
- [6] Thiam, K.C., and Siti, S.S. "Webuse: Website Usability Evaluation Tool". Malaysian Journal of Computer Science, vol. 16, no. 1, pp. 47-57, 2003.
- [7] Hisham, S., and Edwards, A.D. "Incorporating Culture in User-interface: A Case Study of Older Adults in Malaysia". Proceedings of the Eighteenth Conference on Hypertext and Hypermedia, pp.145-146, 2005.
- [8] Balakrishnan, V., and Paul, H.P. 2008. "A Study of the Effect of Thumb Sizes on Mobile Phone Texting Satisfaction". Journal of Usability Studies, vol.3, no. 3, pp. 118-128, 2008.
- [9] Patrick, J. R. "Future of the Internet. Keynote Speech". Americas Conference on Information Systems, 2003.
- [10] Tetard, F., Patokorpi. E & Kadyte. V. "User-Centered Design of Mobile. Services for Tourist: A Case Study on Student Work on Mobile Design". Institute for Advanced Management Systems Research, TUCS, Abo Akademi University, Datacity, Finland, 2005.
- [11] Bessiere, L. J, Ceaparu, I., Robinson, J. and Shneiderman, B. "Help! I'm lost: user frustration in web Navigation". Journal of IT & Society, vol.1, no. 3, pp18-26, 2003.
- [12] Vansderdonckt, J. and Harning, M.B. "Closing the Gaps: Software Engineering and Human- Computer Interaction". Interact 2003, Workshop. http://www.interact2003.org/workshops/ws9- description.html, 2003.
- [13] Singh, S. and Kotzé, P. "An Overview of Systems Design and Development Methodologies with Regard to the Involvement of Users and Other Stakeholders". Proceedings of SAICSITConference, Pages 37 – 47, 2003.
- [14] Klein j., Moon Y. and Picard R. W. "This computer respond to user frustration: Theory, Design and Results", Journal of Interacting with Computer, vol. 14, pp. 119-140, 2002.
- [15] Lawson, R. "Frustration: The development of a scientific concept". New York: MacMillan, 1965.
- [16] Scheirer, J., Fernand, R., Klein, J. and Picard, R. W. "Frustrating the User on Purpose: A Step Toward Building An Affective Computer". Journal of Interacting with Computers, vol. 14, pp. 93-118, 2002.
- [17] Lazar, J. Jones, A. and Shneiderman, B. "Workplace User Frustration with Computers: An Exploratory Investigation of the Causes and Severity". Journal of Behaviour and Information Technology, vol. 25, no.3, pp. 239-251, 2006.
- [18] Bryant, M. "Introduction to user involvement", The Sainsbury Centre for Mental Health, 2001.

# Comparative Analysis of Ecological Footprint of Two Different Neighbourhoods in Minna, Nigeria

#### ABD'RAZACK, Nelson T.A and Ahmad Nazri bin MUHAMAD LUDIN

**Abstract** – This paper discusses the comparative analysis of Ecological Footprint of two different neighbourhoods in Minna: M.I Wushisi and Tunga Low-Cost estate. It considers the building design, types, consumption pattern, lifestyle, and land-use in these estates. One prominent way of curbing sprawl is the design of eco-village that encourages human-scale settlement that encompasses social interaction, environmental sustainability and low-impact lifestyle. Globalization has been seen as a factor that determine environmental and social problem in the society. The method employed in this study makes use of both primary and secondary data to analyse and present the analysis in tabular form to show the relationship of different consumption that contributes to Ecological Footprint of these estates. There was the use of questionnaire for data collection, and a total of 360 and 370 questionnaired administered in both estates. The result indicated that the Ecological Footprint of Tunga Low-Cost (0.94gha) is lower than M.I Wushishi (0.98gha) due to building type, household size and lifestyle. Thus, this implies that the urban planners and designer has to be abreast of necessary information that will allow them to design a city that will be sustainable and consider consumption and lifestyle of the inhabitants of such city or neighbourhood. In addition, the study illustrates that Ecological Footprint could play a useful role in conducting such assessments, by documenting some of the behaviours that are most crucial to a person's total environmental impact and how they are related to design and building form.

Keywords— Consumption, Ecological Footprint, Land-use, Lifestyle, Neighbourhood, Resources, Sustainability

\_\_\_\_\_

---- - - - - - •

#### **1** INTRODUCTION

here is significant relationship between building design and community development in terms of social, economic and environmental impact on the inhabitants of such communities[19]. There is a lot of criticism against urban sprawl that engulf Rural Urban Fringe (RUF) in its development [4, 2]. There are a lot of problems associated with urban sprawl such as high cost of urban services due to distance, high level of segregation and lack of social cohesion among residents and poor environmental management in the fringe. These have diverse effects on Ecological Footprint of the city as distance causes a lot of energy to be used in the process of transportation and building materials. The majority of the buildings in the sprawl area are single family dwelling. Different methods have been proposed to stem the development of urban sprawl [26, 1]. The process of curtailing sprawl has not been successful especially in the third world countries [16, 9, 26, 3, and 14]. Poor planning and individualism caused by economic growth have not been met by physical planning process.

One of the prominent ways of curbing sprawl is the design

of eco-village that encourages human-scale settlement that encompasses social interaction, environmental sustainability and low-impact lifestyle [20]. Globalization has been seen as one of the factors that determine environmental and social problem in the society [29]. The use of ecovillage has been successful in many nations of North America, Europe and South East Asia [14, 29, and 16].

This paper considers Ecological Footprint as useful tool for sustainable development in planning practice. The most important way to determine sustainability is through local planning and urban design. Urban sprawl causes Ecological Footprint to be higher due to consumption of energy for transportation that is critical to planning issues. The most challenging in Ecological Footprint is the consumption by people that form the factors aggravating the environmental impact to the world [27] leading to global warming and climate change. In other to achieve sustainability therefore, it requires critical analysis of both land use and consumption pattern of people in the face of dwindling resources of cities. Planners have little effect on the behaviour of people but could use design of different form to improve the consumption habit of the people.

Basically there are four different types of development approaches in literature for sustainable development. One of the approaches is the extent of protection of resources and ecological function of development; this compares how well two different designs conserve resources in the development of land. This is known as ecosystem planning [12].

This system have not been able to solve the problem of urban sprawl especially energy used for transportation. Other

ABD'RAZACK, Nelson .T A is currently pursuing PhD degree program in Urban and Regional Planning in Universiti Teknologi Malaysia. PH=+60146379331. E-mail: isbamat\_2@yahoo.co.uk

Ahmad Nazri bin MUHAMAD LUDIN is a Prof. in Urbn and Regional Planning Department, Universiti Teknologi Malaysia. PH=+60137798449 E-mail: b-anazri@utm.my

method (ecosystem indicator) measures human impact on the environment. This requires collection of data over time and comparison of such data on different types of development. Good as it may look; it does not present useful information itself well for planners to use in the area of providing useful information for sustainable development. Third is the measurement of sustainability of established project against laydown criteria for development such as Green Building Council's LEED certification process [18]. The LEED sustainability criteria establish benchmark for building and process of neighbourhood certification in USA. Though it is a criterion, it is only used for architectural and building design sustainability process. The forth approach combine various environmental impact in the determination of its usefulness over time. This process compares the building design and consumption pattern of different land use [27].

[25] defined Ecological Footprint as "total area of productive land and water ecosystems required for producing the resources that the population consumes and assimilating the waste that the population produces, wherever on earth that land and water may be located". The overall impact of man on the environment in relation to its carrying capacity is the important aspect that Ecological Footprint considerations [5]. The calculation of Ecological Footprint of nation or city requires the difference between biocapacity of the area compared to the rate of consumption of such resources to determine whether the population is living sustainably. [10], indicated that the average global Ecological Footprint is 2.7gha while the bio-capacity is 1.8gha which leaves deficits 0.4gha. This implies that there is overshoot in the consumption pattern and if care is not taken, there will be depletion in the natural resources and the world will become unsustainable. [11] reports that the world has overshoot by 25% as at 2003. The process of calculating Ecological Footprint incorporates all bio-productive tendencies of all land using equivalent factors. The unit of Ecological Footprint is standardized and expressed as global hectare (gha) of worldaverage productivity.

Ecological Footprint has come to be one of the clear measures of sustainability due to its calculation of impact of human behaviour on the natural ecosystem [25]. Majority of the Ecological Footprint at present are done at national and international level with little on city level [19]. There is need to study Ecological Footprint at local level especially of individual or neighbourhood as case might be. This research focuses on Ecological Footprint of two housing estates in Minna.

# 2.0 STUDY AREA

This paper focused on the lifestyle of the people of Minna and the consumption habit of the people to measure whether the city is sustainable or not within the global ecological limits. It has in its objective the examination of the sociodemographic and economic survey of inhabitant of Minna, assessment of household Ecological footprints in relation to food, housing, transportation, goods and services and the waste generation in the city. It also evaluates the consequential effect on consumption pattern and the waste management style in Minna.

This section examines Minna in relation to location, popula-

tion, people, land use, economic and other activities in the city. Minna lies at latitude 9°37' North and longitude 6°33' East on a geological base of undifferentiated basement complex of mainly gneiss and magmatite. To the Northeast of the city a more or less continuous steep outcrop of granite occurs limiting any urban development in that direction [34]. But the event of urbanisation has led to the encroachment of the base of the outcrop for urban development. The city of Minna has grown from a mere settlement to a city that now has a dual function of Niger State capital and the headquarters of Chanchaga Local government. Due to expansion of the city Minna has now have part of Bosso Local Government as part of it. This implies that Minna has two different local governments (Chanchaga and Part of Bosso LGA).

The present city today is widely spread along the main spine dual carriage road from Chanchaga in the south to Maikunkele in the North- a distance of about 20Kilometres. There is uneven development of the city (Low development at the North-East) due to steep slope, erosion, and flooding and soil type. The other constraints are drainage valley at the centre of the city which flow South-Westward with many minor drainage channels, this has cause flooding in the city in recent past. The land that lies beyond the present built p are suitable for urban development, especially at the fringes, but this requires careful planning to keep the engineering costs of culverts, bridges, embankments and drainage work as reasonable as possible. Minna is approximately 170 Kilometres from Abuja the Federal Capital. It covers an approximately 70,000 Ha of land at the present development. The map of Study area is shown thus in figure 1 and 2 thus:



Figure 1: Map Nigeria showing Niger State



Figure 2: Map of Niger State showing study area **3.0 METHODOLOGY** 

The use of both primary and secondary data was used in the analysis of the research. The two estates have been fully developed which make it possible to collect necessary information from the residents of the estate. There is the use of structured questionnaire to conduct the survey from the residents and information from the ministry of land which is the main implementer of the project. The use of Ecological Footprint spreadsheet for calculation was formulated by [31] and made available by redefine organization was use (Household Ecological Footprint 2.0 by redefine.org). This sheet estimates the Ecological Footprint of households within the neighbourhood. The Ecological Footprint of each estate was calculated using the physical and consumption variables so as to determine the estimate of household and the per capita Ecological Footprint. The physical parameters used are the land for development of each neighbourhood and the material consumed for the construction of the buildings, the lifestyle and consumption habit of the inhabitants of the two neighbourhoods. The lands for building material include space for road, building, parking space, walkways, courtyards etc.

Minna is assumed to have a bio-capacity that can be used to determine the bio-capacity of the two estates. The main focus here is comparative value of environmental effect on the city. The building used are based on design model of the estates (that it its size). There was also inclusion of parking, building and courtyards to determine the per capita consumption. The total Ecological Footprint was then calculated and divided by the number of building in each estate (500 in M.I Wushishi and 400 in Tunga Low cost estate respectively) to estimate the Ecological Footprint of individuals. The consumption data used was average Footprinting of Nigeria due to dearth of data on consumption pattern of the study area.

The consumption focuses on food, automobile, and utilities (electricity and water consumptions) because it is the major consumption variables that show higher energy used on daily basis and which can easily be estimated. A total of 360 and 370 households were sampled using systematic random sampling. The sample size was arrived to from [17] estimate. The data on food consumption was obtained from household expenditure on food. Transport energy, water and electricity usage was also obtained. These include the number of cars owned by each household and alternative usage to public electricity supply and alternative water supply. Information on total number of employed people per household, total households working, distance to place of work, the type of cars. The fuel used was determined using average of 10 Km per litre [6, 7, nd 8]. The amount of water used, the electricity and fuel for generators were also estimated on yearly basis as obtained from the household consumption.

# 4.0 RESULT AND DISCUSSION

#### 4.1 Ecological Footprint Analysis

The population of the two (M.I Wushishi and Tunga Lowcost) estates according to 2006 National and Housing census of Nigeria put them as 5,000 and 9,856 people respectively. This is then segregated to 1,110 and 1,816 households respectively, this implies that the average per household vary between 5 and 7 people per household [20].

The analysis of the total number of building and types of development indicated that there is difference in the design of both estates as shown in table 1 below.

	Tal	ble	1:	De	sign	Ty	ypes	of	Buil	lding	s in	the	Estate	es
--	-----	-----	----	----	------	----	------	----	------	-------	------	-----	--------	----

Design	M.I Wushishi	Tunga Low-Cost
Total Land Size	68.84 Ha	44.65 Ha
Gross Density	8.9	3.6
Dwelling per Hectare	2.4	4.2
Type of Dwelling	3-4 Bedroom Flat	1-2 Bedroom Flats
Future Development	24.8%	Nil
Road Construction	14.0%	10.0%

**Source:** Authors Field Survey, 2012

The design density shows that the gross density per acre of land is 8.9 in the Tunga Low-cost due to age of the estate and the number of people per building, while M.I Wushishi has density of 3.6. The space provides for parking and road construction is higher in Tunga estate compared to M.I Wushishi estate (14%). The space p provided for road construction has been encroached and change into the physical appearances of the building as alter the plan of the estate. The future development can only take place in Wushishi estate (about 25%) while the whole land in the Tunga estate has been developed due to its situation at the centre of the city.

#### 4.2 Land-use Analysis

The land use of the two estates shows contrast due to the age of the estates and location of each estate. The Tunga estate was developed in the 1980s while Wushishi was developed in 2009. The land use pattern shows that the Tunga Low-cost was developed using typical suburban development which covers completely the estate with houses, roads, parking and corner shops. This is contrary to Wushishi which have space for recreation and open spaces and area for future development. There was no fence in Tunga estate while Wushishi estate was fenced. The majority of the inhabitants of the Tunga estate are civil servants while the inhabitants of the Wushishi estate are politicians, technocrats and high income earners.

#### 4.3 Ecological Footprint Analysis

The principle of Ecological Footprint indicated that a compact neighbourhood have lower Ecological Footprint and vice
versa [19]. The analysis of the total land used shows that Tunga Low-Cost has 44.65 Ha of land and M.I Wushishi estate has 68.84 Ha of land. This translates to 3.6 acre per person in Tunga Low-cost and 8.9 acre in Wushishi estate. This implies that Tunga Low-cost have 0.94 gha as Ecological Footprint while Wushishi have 1.15 gha as its Ecological Footprint. This is shown in table 3:

	Table 2: Ecological	Footprint Analy	vsis of the estates
--	---------------------	-----------------	---------------------

Variables	M.I W	Wushishi Tunga		a Low-Cost	
	Area (ha)	Percentage	e Area (ha)	Percentage	
Road	9.64	14.0	4.47	10.0	
Buildings	37.17	54.0	37.51	84.0	
Parking	0.34	0.5	0.45	1.0	
Open space	0.69	1.0	0.89	2.0	
Foot path	0.14	0.2	0.45	1.0	
Recreation	2.75	4.0	0.89	2.0	
Courtyard	0.35	0.5	0.00	0.0	
Future Dev.	. 17.76	25.8	0.00	0.0	
Total	68.84	100.0	44.65	100.0	
Average EF	0.98ha	(2.49 acre)	0.94 (2	0.94 (2.39 acre)	

Source: Author's Field Survey, 2012

The variable that contributes greatly to Ecological Footprint is the building which account for 68.2% in Tunga Low-cost and 83.5% in M.I Wushishi estate due to building material and total material used. But there is difference in Ecological Footprint (0.98 Ha in Wushishis and 0.94 Ha in Tunga Low-cost). There is large courtyard in Wushishi while it is minimal in Tunga Low-cost due to illegal conversion of land for building purposes. The percentage is 4.0% and 2.0% respectively in the land use analysis.

## 4.4 Per capital comparison of Consumption and Building Types

The calculation of Ecological Footprint of three variables of building form which is important to households is calculated. They are the building, parking and courtyards and the behavioural pattern of households consumption for food, utilities and transportation energy. Table 3 shows the analysis of the Ecological Footprint

Table 3: Per capita Ecological Footprint for built form

Types o	of	Bld. size	EF	Ave. EF
Buildin	g	(m²)	(gha)	(gha)
Wushis	i 4 bedroom	700	1.01	0.98
	3 bedroom	600	0.95	
Tunga	2 bedroom	350	1.00	0.94
	1 bedroom	550	0.88	

Source: Authors Field Survey, 2012

The Ecological Footprint of building indicated that per capita index for housing range between 1.01 and 0.95 in Wushishi Estate and 0.88 and 1.00 in Tunga Low-cost. The implication of this corroborates the[4] that compact cities reduces environmental impact and also put forward by [32] that compact cities support Ecological Footprint. The total land developed in Tunga Low-cost is lower compared to the Wushishi estate. This is does to higher density of land 80% in Tunga and 54% in Wushishi.

Electricity consumption in Tunga Low-cost is lower than that of Wushishi estate. This was obtained from the electricity bill payable by households in both estates and the record of [23]. This is in line with the fact that compact buildings uses less electricity compared to bogus buildings as seen in the Wushishi estate. Also the number of rooms and other appliances used by the households in both estates lso affect the energy consuption. Other factors that affect electricity consumption in both estates include income level, lower occupancy ratio. Table 4 indicated the average electricity and water consumption by type of buildings in both estates.

Table 4: Average Othity Consumption by Households					
Estate Ave	e. Electricity (KWh)	Ave. water (litre)			
Tunga 1 bedroom	112	1,200			
2 bedroom	180	1,800			
Wushishi 3 bedroor	m 230	2,500			
4 bedroom	270	4,400			

Table 4: Average Utility Consumption by Households

Source: Authors Field Survey 2012

There is inadequate calculation for water consumption in both estates due to epileptic nature of urban water supply and residents depends mainly on alternative source of water which therefore limits its usage and accurate estimate. Therefore, the Ecological Footprint of these utilities is about 1.4ha and 1.8ha respectively.

## 4.5 Ecological Footprint of Food, Transport and Utilities

There is difference in the food consumption in both estates due to income level and household size. The variables that constitute the food vary and all the variables are combined together for the analysis. The people of Tunga Low-cost eat less meat compared to the Wushishi estate and consume more of vegetables. There is also high consumption of milk, yoghurt, eggs etc. in Wushishi estate. The average consumption of food constituted in Ecological Footprint need some assumptions because of local content. The assumptions are

- (i) Low organic/local content
- (ii) Average organic/local content
- (iii) High organic/local content

Comparing these assumptions indicated that people of Wushishi use high organic content, which are packaged food items than people in Tunga Low-cost. The average Ecological Footprint shows that Tunga estate is 0.94gha while Wushishi have 1.15gha.

The assumed fuel usage per household in the transportation and alternative power supply during outage of public electricity supply shows that it is higher in the Wushishi estate because it is far from workplace of the residents thereby making them to have high number of vehicles compared to Tunga estate. The average number of vehicle in Wushishi is about 3 cars per household whereas it is about 1 car per household in Tunga estate. Also the people of Tunga estate enjoy more electricity supply than Wushishi estate which makes them to depend less on generators. Almost all the households in Wushishi have generating set which make them to buy more fuel to power the machine whereas only few households in Tunga estate have the generating set. The analysis of average fuel consumption per household during electricity outage is about 400 litres per month while it is about 180-200 litres per month in Tunga Low-cost. Table below shows the average fuel consumption per household

Table J. Average Fuel Consumption ber mousenoid	Table 5: Average	<b>Fuel Consum</b>	ption per Household
-------------------------------------------------	------------------	--------------------	---------------------

Estate	Wushishi Estate	Tunga Low-cost
Ave. number of Cars	3	1
per Households		
Ave. number of	1	0.4
Generating sets		
Ave. litre of Fuel consur	ned 380-400	180-200
per households		
Ave. cost of Fuel N2	24,700 –N 26,000	N 9,750 - N 13,000
per month		

Source: Authors Field Survey, 2011

## 4.6 Per-capita Comparison Based on Selected Residential and Consumption Components

The analysis of per-capita Ecological Footprint when calculated from the consumption patter of people in these estates depend on households income and types of residents they occupy and the consumption of food, travels by different mode of transportation, electricity and water. The analysis is shown in table 6.

Table 6: Per-Capita EF of building Form

	<u>1</u>		0			
Estate		Type of	Dwelling	Yards	EF per	
		Dwellings	Size(m <sup>2</sup> )	Size(m <sup>2</sup> )	Acre	
M.I Wus	hishi	4 bedroom Flat	700	200	1.15	
		3 bedroom Flat	600	150	1.02	
		2 bedroom Flat	380	70	0.98	
Tunga	3 Bedı	oom Detarched	350	50	0.94	
	2 Bed	room semi-D	550	100	0.92	
	1 Bed	lroom Semi-D	250	100	0.88	
C	NT.	C	(т 1	1 T D	1 ·	

**Source:** Niger State Ministry of Land and Town Planning, 2010.

The Ecological Footprint as shown in table 6 indicate the different construction of dwellings in the estates and shows the per-capital Ecological Footprint that ranges between 1.15gha and 0.98gha in M.I Wushishi estate for building measuring between 700m<sup>2</sup> and 380m<sup>2</sup> respectively. Also the Ecological Footprint of dwellings in Tunga Low-Cost implies that it ranges between 0.94 and 0.88 from building of sizes between 350m<sup>2</sup> and 250m<sup>2</sup> respectively. This indicated that a compact Tunga Low-Cost estate is more sustainable and corroborated the findings of [34 and 4] that cities have an opportunity to reduce their environmental impact by encouraging higher densities.

Another factor considered in the sustainability measure of the two estates is the consumption of amenities in the dwellings, the average water consumption; electricity consumption and petrol consumption by inhabitants of these estates are shown in figure 3 thus:



Figure 3: Consumption of Resources by Households per-Capita

Source: Author's Field Survey, 2012

This implies that the average consumption of these amenities in Tunga Low-Cost is lower than the M.I Wushishi estate due to factors such as the household size, lifestyle, income level and ability of households to spend on these utilities. All these activities constitute what make the Ecological Footprint to differ in these estates despite the fact that they are situated in the same city of Minna.

The survey conducted has its own limitation which range from Ecological Footprint as an assessment tool for urban development because it is not easy to combine the dwelling form with consumption habit of the inhabitants from the onset. This though have effect on the environment, there is no adequate data to relate these variable together. The Ecological Footprint calculation for individual, household and neighbourhood differs due to resources consumed the ability to afford the resources and the lifestyle adopted by households. Many of the resources consumed cannot be measure directly as it requires modification to suit Ecological Footprint calculation. Another dimension to this problem is the estimation of food consumed by households from different sources. Indeed, most of the data for food would necessarily come from self-reports, which are well known to have limitations in terms of accuracy and completeness [22, 32].

Data for Origin-Destination is not available to determine accuracy of the distance covered by households annually which are estimated based on the availability of cars to households and ability of the households to afford the cost of transportation.

## 5.0 CONCLUSION

The analysis shown above indicated that there is different estimates of Ecological Footprint of different neighbourhood of a city rather than estimating the city's Ecological Footprint. There is more analysis to be able to generalize the EF of a particular city in developing countries as dearth of data is a problem in the process of calculating the Ecological Footprint. The Ecological Footprint per-capital in Tunga Low-Cost is lower than the M.I Wushishi estate, but the difference is negligible and could be due to several factors such as consumption habit, lifestyle, household size, availability of private cars, parking etc. and building form. The most contributing factor to EF is what the household consume; therefore, there is critical need to review the building design and lifestyle to determine the overall EF of a particular neighbourhood. Due to compactiness of Tunga Estate it may reduce the EF and consideration must be giving to how the way we build our cities influences consumptive behaviour.

What this implies that the urban planners and designer has to be abreast of necessary information that will allow them to design a city that will be sustainable and consider consumption and lifestyle of the inhabitants of such city or neighbourhood.

In addition, the study illustrates that Ecological Footprint could play a useful role in conducting such assessments, by documenting some of the behaviours that are most crucial to a person's total environmental impact and how they are related to design and built form. The Ecological Footprint tool was deemed to have considerable promise as a neighbourhood planning tool, despite challenges associated with data assembly and conversion and limitations in its ability to deal with cause-and-effect processes.

#### REFERENCES

- Arendt, R. (1999). Growing Greener: Putting Conservation into Local Plans and Ordinances (Washington, DC: Island Press).
- Barnett, J. (2003). Redesigning Cities (Chicago, IL: American Planning Association Press).
- [3] Burchell, R. W. and Mukherji, S. (2003). Conventional Development versus Managed Growth: The Costs of Sprawl, American Journal of Public Health, 93(9), 1534–1540.
- [4] Calthorpe, P. and Fulton, W. (2001). The Regional City: Planning for the End of Sprawl (Washington, DC: Island Press).
- [5] Chambers, N., Simmons, C. and Wackernagel, M. (2000). Sharing Nature's Interest: Ecological Footprints as an Indicator of Sustainability (London: Earthscan).
- [6] Department of Energy (2000) Chapter 4: Fuel Economy, Consumption and Expenditures, in: Household Vehicles Energy Consumption 1994. Available at <u>http://www.eia.doe.gov/emeu/rtecs/chapter3.html</u>) (accessed August 2009).
- [7] Department of Energy (2000). Chapter 3: Vehicle Miles Traveled, in: Household Vehicles Energy Consumption 1994. Available at <u>http://www.eia.doe.gov/emeu/rtecs/chapter3.html</u> (accessed August 2009).
- [8] Department of Petroleum Resources (2004). Fuel Consumption per Litre of Different types of Vehicles in Nigeria. NNPC/Joint Venture Publications. NNPC Tower, Abuja Nigeria.
- [9] Duany, A., Plater-Zyberk, E. and Speck, J. (2000). Suburban Nation: The Rise of Sprawl and the Decline of the American Dream (New York: North Point Press).
- [10] Global Footprint Network (2008): Ecological Footprint Standards 2008. Oakland: Global Footprint Network. Retrieved from <u>http://www.footprintstandards.org in April, 2011</u>
- [11] Global Footprint Network (2010): Ecological Footprint Atlas, 2010. GFN Publishers, Oakland, California. USA. Retrieved from www.footprintnetwork.org on 2<sup>nd</sup> April, 2011.
- [12] Gordon, D. and Tamminga, K. (2002). Large-Scale Traditional Neighbourhood Development and Pre-Emptive Ecosystem Planning: The Markham Experience, 1989–2001, *Journal of Urban Design*, 7(3), 321–340.
- [13] Holden, E. (2004). Ecological Footprints and Sustainable Urban Form, Journal of Housing and the Built Environment, 19(1), 91–109.
- [14] Hu, D. and Wang, R. (1998). Exploring Eco-Construction for Local Sustainability: An Eco-Village Case Study in China, *Ecological Eco-*

nomics, 11, 167-176.

- [15] Kelbaugh, D. (1997). Common Place: Toward Neighborhood and Regional Design (Seattle, WA: University of Washington Press).
- [16] Kirby, A. (2003). Redefining Social and Environmental Relations at the Eco-Village at Ithaca: A case study, *Journal of Environmental Psychology*, 23, 323–332.
- [17] Krejcie, R and Morgan, D.W. (1970). Determining the Sample Size for Research Activities. *Educational and Psychological Measurement*, 30, 607-610
- [18] Leadership in Energy and Environmental Design (2005). LEED: Leadership in Energy and Environmental Design. Available at <u>http://www.usgbc.org/DisplayPage.aspx?CategoryID=19</u> (accessed October 2011).
- [19] Moos, M., Whitfield, J., Johnson, J.C., and Andrey, J. (2006). Does Design Matter? The Ecological Footprint as a Planning Tool at the Local Level. *Journal of Urban Design*, 11(2), 195–224.
- [20] National Population Commission (2010). The 2006 National Population and Housing Census Data: Priority Tables. NPC Publication, Presidency, Abuja, Nigeria.
- [21] Norberg-Hodge, H. (2002). Why Eco-villages? The Ecologist, 32(1), 38.
- [22] Parslow, R., Hepworth, S. and McKinney, P. (2003). Recall of Past use of Mobile Phone Handsets, *Radiation Protection Dosimetry*, **106**(3), 233–240.
- [23] Power Holding Company of Nigeria (2010). Metre Reading for Households Energy Consumption in Niger State, Nigeria. QuarterlyBulletin, PHCN, Minna, Nigeria
- [24] Pollard, T. (2001). Greening the American Dream? *Planning*, 67(10), 10–16.
- [25] Rees, W. (2000) Eco-Footprint Analysis: Merits and Brickbats, *Ecological Economics*, 32(3), 371–374.
- [26] Roseland, M. (1992). Toward Sustainable Communities (Ottawa: National Round Table on the Environment and Economy).
- [27] Spangenberg, J. H. and Lorek, S. (2002). Environmentally Sustainable Household Consumption: from Aggregate Environmental Pressures to Priority Fields of Action, *Ecological Economics*, 43, 127–140.
- [28] Takeuchi, K., Namiki, Y. and Tanaka, H. (1998). Designing Eco-Villages for Revitalizing Japanese Rural Areas, *Ecological Economics*, 11, 177–197.
- [29] Trainer, T. (2000). Where are we, Where do we want to be, How do we get there? *Democracy and Nature*, 6(2), 267–283.
- [30] Wackernagel, M. and Yount, D. (1998). The Ecological Footprint: An Indicator of Progress Toward Regional Sustainability, *Environmental Monitoring and Assessment*, **51**, 511–529.
- [31] Wackernagel, M., Deumling, D., Manfreda, C. and Dholakia, R. (2003). Assessing Your Household's Ecological Footprint, Version 3.2, Redefining Progress. Available at <u>http://www.redefiningprogess.org</u> (accessed January 2010).
- [32] Wackernagel, M., Schulz, N., Deumling, D., Linares, A., Jenkins, M., Kapos, V., Monfreda, C., Loh, J., Myers, N., Norgaard, R. and Randers, J. (2002). Tracking the Ecological Overshoot of the Human Economy, *Proceedings of the National Academy of Sciences of the USA*, 99(14), 9266–9271.
- [33] Walker, L. and Rees, W. (1997) Urban Density and Ecological Footprints: An Analysis of Canadian Households, in: M. Roseland (Ed.) Eco-City Dimensions, 96–112 (Gabriola Island, BC: New Society Publishers).
- [34] Maclox Limited (1980). Minna Master Plan. Maxclox Publication for Niger State Ministry of Land and Town Planning. Maidiguri, Nigeria.

# Impact of public space utilization frequencies on tacit knowledge sharing

KAYODE OLORUNTOBA, MOHD HISYAM RASIDI, ISMAIL SAID

Abstract – Public space influence on knowledge sharing acquisition has become an interesting issue among knowledge management, urban landscaper, and knowledge based development researchers. Thus, this study aim to investigate the contributing influences of the degree of public space usages on tacit knowledge sharing. Cyberjaya Malaysia was considered as the study area having been adjudged as a high technology park that foothold in knowledge sharing for its knowledge based development. A total of 384 survey questionnaires were administered by the residents of the study area in various public spaces within the city to collate the respondent's perception on their degree of public space utilization in relation to tacit knowledge sharing behaviour. Validated variables were adapted to measure knowledge sharing while the frequencies of public space utilization were measure with the user's degree of visits to public spaces. Data collated were analysed with statistical packages for social science "SPSS" to access the differences and similarities in respondents perception. Our findings revealed that the frequencies at which people utilized public spaces exhibited similarity differences in their tacit knowledge sharing tendencies. Human attitudes towards sharing their knowledge indicated to require much social interactions and public space utilization.

Index Terms— Cyberjaya, Human attitudes, Knowledge based development, Knowledge sharing, Public space, social interactions Tacit knowledge

\_\_\_\_ **♦** 

# **1** INTRODUCTION

Dublic space can be considered as all social spaces that is easily accessible to users on non-conditional basis [1]. Public spaces are not just open spaces. Public space is characterised by the facilities, and infrastructures that geared towards providing comfortable natural environment for the users. The forms and settings of public space provide a defined demarcation from other surrounding of the city. Public space consists of numerous social spaces within the realm of a city. In the context of science or technological city, public space encompass majorly the communal spaces, public squares and urban courtyards, and the extended public parks that characterised in high tech amenities and facilities. It's served as a natural area that accord users the common sense of environmental appreciation [2]. In a landscape perspective, Public space is a tool that capable of attracting human mind and patronages. It provides avenue for human contact and togetherness. Human social discussion and interaction is foothold in the availability of natural environment [3]. When people visited public spaces, they meet others public space users which set a good platform for human contacts. However, human contact encourages interaction while social cohesion is an offshoot of the integration [4]. Thus, Informal study does occur when there's a group interaction. The meaningfulness of an area can be link with the public space standard. When a city has well maintained and high quality public spaces, users continue to increase and more people continue to visit the city and develop good idea about the area. Therefore such users psychologically developed good sense of association and images for the city. The potential of public space to facilitate social cohesion necessitates that it has an association with knowledge sharing. It has been argued that experiences sharing occur through human interactions [5] and public spaces provide an effective avenue for human social interactions [3], [5]. Therefore public

space can be considered as an important tool for knowledge sharing. Human frequents contact develop social interaction that gives rise to social cohesion and social capital needed for sharing.

#### 1.1 Study area

\_\_\_\_\_

Cyberjaya is a modern science city that was designed to establish and sustained the Malaysia multimedia super corridor center. The conception of Cyberjaya originates from a study by management consultancy McKinsey towards Malaysia multimedia super corridor. The city was commissioned by the Federal Government of Malaysia in 1995 [24]. Cyberjaya is situated in Sepang, Selangor and about 50 km south of Kuala Lumpur Malaysia. It occupied about 28.94 square kilometers of land with population of around 45,000 that comprises of 19,000 of workforces, 16,000 of students, and 10,000 that are residents [25].

# 2 TACIT KNOWLEDGE

Knowledge involved the act of exchanging the know-how among people for the purpose of rendering assistance or by way of solving problem [6). Its involve getting to know about untold fact and skills. Validated information becomes knowledge [7]. Thus, knowledge can either be explicit or tacit. Explicit knowledge is obtained through analytical or scientific reasoning via the course of formal training. Tacit knowledge encompasses the act if gaining knowledge through experiences and interaction that occur among group of people. It's often informal and subjective [8] as it hinged on the social cognitive of people. However, tacit knowledge has held the power for innovation and technology productivity. Tangible knowledge can be obtained through group discussion and experience exchange [9]. It's strongly rooted in human action as in difficult to be coded or store in explicit form. Human social environment is required for effective tacit knowledge development since individual or group of people can acquire tacit knowledge through social interactions [9]. Therefore, social interactions constitute a major factor that facilitates tacit knowledge.

## **3 KNOWLEDGE SHARING**

Knowledge sharing encompasses the transfer of relevant ideas among people. It's a process that involve individual or groups of people sharing there know- how, experiences and skills among others that were in need [10]. In recent time, numerous researches have emphases on the significant of knowledge sharing in innovation [11], [12]. Valuable part of knowledge resides in the tacit knowledge [13]. Knowledge without sharing retards development and limit the development of such knowledge itself. It's through sharing that knowledge expanded and becomes productive. One of the easier channels for firms and organizational innovation and technological development is through knowledge sharing [14]. Knowledge sharing provides conducive avenue for competitive advantage [15] and refined professionalism. In the context of science city, knowledge sharing can be defined as exchanges or transfer on individual or group know - how among others that exhibits social bond and cohesiveness within a defined environment.

# **4 METHODOLOGY AND MEASURES**

In this study, 384 questionnaires were administered to collate the regularity of respondent frequencies of visit to public space. The regularity of respondent utilization of public space were categorized on everyday visit; twice a week visit; weekly visit; monthly visit; and occasional visit graded with Likertscale ratings of 5 point to 1 point for everyday visit to the occasional visits to public space respectively. Knowledgesharing was measured by three constructs (subjective norms to share knowledge, attitude to share knowledge, and intention to share knowledge), adopted from [16], and [17]. The collated data were analyzed using "ANOVA" analysis of variance to obtain the data average mean, and the groups mean discrepancies. Attitude toward sharing encompasses transferring knowledge and experience that arising from individual desires to share. Subjective norms to share are as influence by others friends or community associates to share. Intention to share can be define as the degree of human belief that will be engaged in the knowledge sharing behavior. Hence, the test of data reliability was done using Cronbarch's Alpha while data collated consistency was tested using factor analysis.

## **5 ANALYSIS AND DISCUSSIONS**

#### 5.1 Demographical survey

Demographical influences of resident's status of age, and gender were used to explore their effects on public space utilization. The finding consisted with literature that hypothesized knowledge city as a community of advanced literacy residents [18], [19]. Higher fractions of respondent were university degree and post graduate degree holders having 66.1% and 17.7% respectively as showed in Table 1. The population percentages of the male and female respondents recorded 65.3% and 34.7% respectively thereby reflecting a gender justification.

Measure	Items	Per
		cent
		(%)
Gender	Male	66.4
	Female	33.6
<b>Residents Status</b>	Yes	75.5
	No	24.5
Duration of	0-3yrs	28.9
Residents	4 – 6yrs	18.2
	7 – 9yrs	41.9
	10 yrs. and above	10.9
Educational	High School or equivalent	2.1
status	undergraduate	14.1
	graduate	66.1
	postgraduate degree	17.7
	neighbourhood/communal	
Types of public	spaces-	56.3
space visited	public parks-	10.9
-	public square/urban	
	cluster court-	16.7
	yards-	
	other(canopies, entrance	16.1
	porch, etc.)	
N-294		

Table 1. Demographical valuation

N = 384

## 6 ANALYSIS AND RESULTS

Cronbarch's Aphal of the variables exceeded 0.700 which demonstrating reliable value [20]. The confirmatory factor analysis (CFA) applied to check the measurement model variables as recommended [21]. The indicators factor loadings were all significant at 0.01 which is considered as good model [22]. One-Way ANOVA (Analysis of variance) applied to measure collated data. Thus, statistically significant different was establish between the five levels of public space visitations indicators in relation to knowledge sharing measuring constructs (subjective norm to share knowledge, intention to share knowledge, and attitude to share knowledge).

The analysis of variance showed that; subjective norm to share knowledge exhibited F (4, 379)= 636.864, p = .000. Intention to share knowledge exhibited F (4, 379)=13.412, p = .000. Attitude to share knowledge exhibited F (4, 379) =459.009, p = .000 as showed in Table 2.

Source		Sum of Squares	df	Mean Square	F	р
Subjective norm to share	Between Groups	245.787	4	61.447	636.864	.000
knowledge Within Groups	Within Groups	36.567	379	.096		
	Total	282.354	383			
Intention to I share knowledge	Between Groups	8.189	4	2.047	13.412	.000
	Within Groups	57.851	379	.153		
	Total	66.040	383			
Attitude to share	Between Groups	461.309	4	115.327	459.009	.000
knowledge With	Within Groups	95.225	379	.251		
	Total	556.533	383			

**Table 2.** One-Way Analysis of Variance Summary Table Comparing respondents

 Public space visitation on Knowledge sharing

In Table 3, all the public space degree of usages indicators exhibited statically significant on knowledge sharing constructs except the subjective norm to share knowledge that exhibited non-significant mean of 1.9 to respondents that visits public space on occasional basis. Also, attitudes to share knowledge exhibited non-significant to those respondents that visit public spaces on monthly and occasional basis (See Figure 3). People's intention to share knowledge exhibited average mean of 4.2 to 4.5 that indicates high significant to occasionally visit, monthly visit, weekly visit, twice a week visit, and daily visits to public.

Table 3. Means and Standard Deviat	tions Comparing public space visitation indicators
	actione expansion opage rightation indicatore

Community Attachment Variables	Degree of Public Space Visits	Number of Re- spondents	Mean	SD
Subjective norm to share	Occasionally	59	1.9379	.35274
knowledge	Monthly	120	4.1222	.27983
	Weekly	58	4.0230	.24868
	Twice a week	100	4.0733	.31633
	Daily	47	4.4255	.37880
	Total	384	3.7960	.85861
Intention to share knowledge	Occasionally	59	4.4271	.36852
	Monthly	120	4.2400	.38944
	weekly	58	4.1517	.41263
	Twice a week	100	4.1760	.40129
	Daily	47	4.5957	.36946
	Total	384	4.2823	.41524
Attitude to share knowledge	Occasionally	59	2.0305	.75275
	Monthly	120	2.1200	.47697
	weekly	58	4.1862	.31313

Twice a week	100	4.2160	.46509
Daily	47	4.5234	.42437
Total	384	3.2583	1.20544

The Post hoc HSD Test was used to compare mean difference of verified variables. It was shown that the significant mean differences occur among respondents that visited public space on daily basis and those respondents that do visits public space on weekly, twice a week, monthly, and occasionally basis (See Table 4). Table 4 reflect the Post Hoc HSD summary of the respondent's knowledge sharing behaviour. It indicates that the respondents perceptions defers on their understanding and responses to knowledge sharing in the context of the degree on public space visitation. On the basis of occasional visits to public space, the respondents demonstrated no significant differences with those that do visit public space monthly, weekly, twice a week, and on daily basis.

	How often do you visit the	Mean Difference
	public space?	(I-J)
occasionally	Monthly	.18712*
	Weekly	.27539*
	Twice in a week	.25112*
	Daily	16863
Monthly	vocationally	18712*
	Weekly	.08828
	Twice in a week	.06400
	Daily	35574*
Weekly	Occasionally	27539*
	Monthly	08828
	Twice in a week	02428
	Daily	44402*
Twice in a week	Occasionally	25112*
	Monthly	06400
	Weekly	.02428
	Daily	41974*
Daily	Occasionally	.16863
	Monthly	.35574*
	Weekly	.44402*
	Twice in a week	.41974*

Table 4. Po	st Hoc table	of group	differences
-------------	--------------	----------	-------------

Table 5: Summary of significant relationships among tested variables

Community Attach-	Degree Of Public Space Visitation				
ment Variables	Daily	Twice a	Weekly	Monthly	Occasionally
	Visit	Week Visit	Visit	Visit	Visit
Attitude to share	•	•	•		
knowledge					
Subjective norm to	•	•	•	•	
share knowledge					
Intention to share	•	•	•	•	•
knowledge					

# **7 DISCUSSIONS AND CONCLUSION**

This study indicates that knowledge sharing is obtainable in public spaces as primary social place. Human intention to share knowledge as part of the knowledge sharing

determinants can be obtained through public space usage. The respondents in the study area demonstrated that occasional usage of public space is capable of trigger their intention towards sharing of knowledge among friends and co- workers. However, intention to share knowledge cannot independently determined people actual knowledge sharing behaviour. As such, attitudes towards sharing of knowledge and the subjective norm towards sharing knowledge should be considered as well. In this study, people's attitudes to share knowledge reflect no significant relationship with monthly visit, and occasional visit to public space. Daily visit, weekly visit, and twice a week visit are the minimum requirement exhibited by human attitude towards sharing of knowledge in the study area. Subjective norm to share knowledge exhibited significant relationship with daily visit, twice a week visit, weekly visits, and monthly visits to public space but demonstrated poor significant relationship with occasional visits to public space. Therefore, it's of importance to argue that the respondents in this study area needed to be utilizing public spaces at the minimum visits of monthly basis to develop subjective norms among others towards sharing of their knowledge.

Visiting social places and engaging in social activities in public space provides opportunity to meet others and facilitates trust and confidence among neighbours. Constant relationship can facilitates social bonds and close familiarity [26]. Familiarity refines human minds and ideas of others about life and taught. Therefore, it implies that much of time is needed for people to develop habit that capable of shaping their subjective norms. The result of this study supported the aforementioned as it indicates that subjective norms prove unattainable in the situation where the people choose to visit social places (public space) only on occasional basis. More so, individual attitudes to share knowledge demonstrated to be more time demanding to acquire among the respondents. Attitudes associated with human personal ways of reacting to issues, it's has much to do with the state of human minds. Therefore, social contacts and mutual relationship are required in reshaping human inborn characters and attitudes. The frequency in human interaction and friendship possess the potential to influence their attitudes. Monthly and occasional visits to public spaces proved not sufficient to facilitate human attitudes towards sharing their know-how among others (See Figure 5). Therefore, our findings revealed that human attitudes is needed more social engagement. Thus, public space remains a freely accessible social arena in every community and urban centres. Regular visits to public space can develop sharing habit among neighbourhood and groups of users. The study area has been adjudged as a high technology park that foothold in knowledge sharing for its knowledge based development [23]. Its advised that knowledge community and technological oriented cities that foothold in knowledge sharing should imbibes in public space frequent utilization culture as a functioning tools towards knowledge sharing actualisation

while public space should be accord outstanding consideration in the urban and rural dwellers community.

## REFERENCES

797

- J. Gehl, "Three types of outdoor Activities and quality of outdoor space in Gehl(1996).life between Buildings using public space, ArkitektensForlag, skive, 2001. pp. 11-40
- [2] Y C.Jim, and W.S Chen 'Comprehensive green space planning based on landscape ecology principles in compact Nanjing city, China. Landscape and Urban Plan ning 998 (2003) 1–22
- M. armona, C. Magalhães, and L. Ham mond. 'Public Space: The manage ment dimension'' Firstpublished 2008 by Routledge, 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN. Simul taneously published in the USA and Canada by Routledge 270 Madison Avenue, 2008. New York, NY 10016, USA. ISBN 0- 203- 92722-2 Master e-book ISBN
- [5] R. Putnam. 'Bowling along. America's de clining social capital. Journal of Democ racy, 1993. 6(1),65-78

[6]K. Peters, B. Elands, and A. Buijs, 'InteraCtions in urbanparks: Stimulating socialcohesion?Urban Forestry & Urbanban Greening 9 (2010)pp. 93–100

- K. E. Kelloway, 'Labor unions and safety: Conflits and cooperation" in J. Barling and M Frone (eds). Psycho logu of occupation and safety. Washington DC. APA Books. 2003. Pp. 249 - 264
- [8] M. Alavi, and E.D Leidner. 'Knowledge maagement and knowledge managementsystems'': Conceptual foundations and research issues", MIS Quarterly, 2001.Vol. 25 No. 1, pp. 107-36.
- [9] I. Nonaka. 'The knowledge-creating com pany. Har vard Business Review, 1991. 69, 96-104.
- [10] B. Choi and H.Lee. 'An empirical investigation of KM styles and their effect on corporate performance. *Information & Management*, 2003. 40, 403–417.
- [11] D. Hong, E. Suh, and C. Koo. 'Developing strategies for overcoming barriers to knowledge sharing based on conversa tional knowledge management: A case study of a financial company. Expert Systems with Applica tions 38 (2011) 14417–14427
- [12] C.Huang. "Knowledgesharingandgroup cohesivenessonperfor mance:An empirical study of technol gyR&DteamsinTaiwan. Technovation 29 (2009) 786–
- [13] N. Mooradian, "Tacit knowledge: Philosophical roots and role in KM", Journal of Knowledge Management, 2005. Vol. 9 No. 6, pp. 104-13.
- [14] I. Nonaka and H. Takeuchi. "The Knowledge Creating Company": How Jap anese Companies Create the Dynamics of In novation, Oxford University Press, New York, NY. 1995.
- F.H. Lin, "Knowledge sharing and firm innovative capability:An empirical study", *International Journal of Manpower*, 2007. Vol. 28 No. 3/4, pp. 315-32.

- [16] F. Requena, "Social capital, satisfaction and quality of life in the workplace", *Social Indicators Research*, 2003. Vol. 61No. 3, pp. 331-60.
- [17] W.S. Chow and L.S. Chan. "Socialnetwork, social trust and sharedgoals inorganizational knowledge sharing", *Information* and Management, Vol. 45 (2008)pp.458-465
- [18] A.Anttiroiko, Global Competition of high-Tech centre. Jour nal of technology Management. 2004. Vol 28 number 3/4/5/6. Pp259-294. 259-294.
- [19] K. Ergazakis, K. Metaxiotis, and J Psar ras, 'An Emerging Pat tern of Succesful Knowledge Cities: Main Fea tures' in Fran cisco Javier Carillo (ed.), Knowledge Cit ies: Approaches, Experiences and Perspectives, Oxford: Butterworth-Heine mann,2006. pp. 3-16.
- [20] H.J Nunnally and H.I Bernstein, '*Psychometric Theory*, McGraw-Hill, New York, NY.1994
- [21] C.J. Anderson and W.D Gerbing, "Assumptions of the twostep approach to latentvariable modeling", *Sociological Methods* and Research, 1992. Vol. 20 No. 3, pp. 321-33.
- [22] R.P. Bagozzi and Y. Yi, "On the evaluation of structural equationmodel", *Journal of Academy of Marketing Science*, 1998.
   Vol. 16 No. 1,pp. 74-94.
- [23]. SetiaHaruman. 'Cyberjaya: Over view.2011.' Retrieved 17 Oct, 2012, <u>http://www.cyberjaya-msc.com/cyberjaya-</u>town.asp.
- [24] Federal Department of Town Planning. The physical plannibg guidelines for the multimedia super corridor. Kuala lumpur: ministry of housing and local government Malaysia. 2000.
- [25] SetiaHaruman. 'Cyberjaya: Over view.2011.' Retrieved 17 Oct, 2012,<u>http://www.cyberjaya-msc.com/cyberjaya-</u>town.asp.
- [26] J. Macinko and B. Starfield. "The utility of social capital in re search on health determinants. Milbank Quarterly, 2000. 79(3), pp. 387–427.

Kayode OLORUNTOBA received Masters Degrees in Architecture from Federal University of Technology, Akure, Nigeria. He is currently a Doctoral Candidate in the Depart. of Architecture, Universiti Teknologi

Malaysia. +60102117988 kayodetoba@yahoo.com

Mohd Hisyam RASIDI received PhD in Regional Environment Systems, Shibaura Institute of Technology .He is a Senior Lecturer, Depart. Landscape Architecture, Universiti Teknologi Malaysia b-hisham@utm.my

Ismail Said received PhD in Architecture from Universiti Teknologi Malaysia. He is an Associate Professor at Depart. Landscape Architecture, Universiti. Teknologi Malaysia. b-ismail@utm.my

# E-Learning Acceptance among Tertiary Education Students

Lee Nyuk Ling, Lee Yoke, Nik Kamariah Nik Mat

**Abstract**— In current year, e-learning received special attention from higher education in implementing distance learning courses. Therefore, in this study, we examined the factors influence the acceptance of e-learning among tertiary education students. The factors tested were results demonstrability, performance expectancy, effort expectancy, and social influence. A total of 213 respondents received from several private higher education learners in Malaysia who are currently pursuing Certificate, Foundation, Diploma and Degree. The findings of this study revealed that results demonstrability and performance expectancy is significant positively influence the acceptance of e-learning. Whereas, effort expectancy and social influence is insignificantly influence the acceptance of e-learning.

-----

Index Terms—E-Learning, effort expectancy, performance expectancy, results demonstrability, social influence

# 1. Introduction

The term e-learning referred to the learning methodology using any electronic as a delivery medium. E-learning can be synchronous in which participants have to attend online sessions on particular time [1]. There are so many synonymous of e-learning like web-based learning, e-education, open-learning, open courseware, and virtual education. The other type is asynchronous in which participants do not need to attend any classes and follow restricted time for attending class. In November 2005, the Sloan Consortium published a report on elearning and defined "online learning" or "e-learning" as a learning in which 80% to 100% of the content is using Internet as a medium of delivery [2].

In current year, due to the rapid growth of webbased technologies and high usage of Internet have made learning and teaching via the e-learning more feasible. Elearning (on-line learning, web-based learning or internet learning) is a method which developed from distance education. It has received special attention from higher education in implementing distance learning courses. According to Chai and Poh [3], e-learning is the most recent evolution of distance learning that creates, delivers, fosters, and facilitates learning, anytime and anywhere, with the use of electronic technologies as a medium of instruction in distance learning. However, distance participants need to be ready for this type of learning tools as past experiences showed that new technologies do not necessarily lead to major enhancement in education [4]. According to Raja Hussain [5], due to the increase in demand for higher education, many institutions in Malaysia have planned for e-learning. A

• Lee Yoke is currently pursuing PhD in Banking & Finance in Universiti Utara Malaysia, Malaysia. E-mail: leetingtong@hotmail.com number of distance learning higher education institutions such as Wawasan Open University and the Open University of Malaysia had emerged in the growth of distance learning in Malaysia [6]. Meanwhile, O'Malley [7] reported e-learning as method distance learning is being promoted as the educational medium of the future.

E-learning allows a new way for many adults who have been tied up with many commitments in life and facilitate them to learn anytime and anywhere they want at their flexibility and convenience. Access to learning via the Internet has made physical or geographical limitations no longer a critical issue for adults to enroll in any course with any university where e-learning opportunities are available. Many past studies showed positive results on the acceptance of e-learning by working adults. However, not much has been found on the acceptance of e-learning by young students.

## 2. Literature Review

According to a several studies on information technology systems [8], [9], [10] and the technology acceptance model (TAM) proposed by Davis [11], can efficiently explain and predict learners' intention and behaviour. Chang et al. [12] investigated on perceived convenience in the extended TAM for examining the technology acceptance model of the mobile learning activities. Besides that the antecedent factors (perceived ease of use and perceived usefulness) that affected acceptance of English mobile learning were also examined. The results showed that perceived convenience positively affected attitude towards using but indirectly affect continuance intention to use through perceived usefulness and attitude toward using.

Chai and Poh [3] studied on criteria such as program content, web page accessibility, learner's participation and involvement, web site security and support, institution commitment, interactive learning environment, instructor competency, and presentation and design to identify successful criteria in implementing an e-learning program in Malaysia. All the factors were deemed important for the successful implementation of e-learning

Lee Nyuk Ling is currently pursuing PhD in Banking & Finance in Universiti Utara Malaysia, Malaysia. E-mail: leenyukling@hotmail.com

NikKamariahNik Mat is a Professor in Universiti Utara Malaysia, Malaysia. E-mail: drnik@uum.edu.my

program. Meanwhile, Selim [13] studied on four main critical factors (student, instructor, information technology, and university support) and reported all the factors deemed as critical determinants of e-learning acceptance. Apart from that, there were some researcher found that a number of factors such as students' and instructors' characteristics [14], [15], [16], technology support and system [16], [17], institutional support [18], [19], course content and knowledge management [13], [20] and online tasks and discussion groups [21], [22] could influence learners' towards acceptance of e-learning.

Hong et al. [15] studied on students' satisfaction and perceived learning with a web-based course by undertaken among postgraduate students in University Malaysia Sarawak University. The findings revealed that students had high level of acceptance with their Webbased courses. The students who had high level of acceptance found that the web-based course was flexible and convenient [15], [23]. Furthermore, majority of the studies reported that there was no difference in learning achievement between the students taking web-based courses and those students enrolled in traditional face-toface educational environment [15], [24], [25], [26], [27]. Hong et al. [15] found some students faced problems with the web learning environment and they needed more time and guidance to adapt to the web-learning environment.

On the other hand, Poon et al. [16] investigated on the main five factors (students' behaviour, characteristics of lecturers, interactive application, technology or system, and the institutions.) that affect the effectiveness of webbased learning environment in various courses at eight universities in Malaysia. The study reported that the students were not fully comfortable with web-based learning environment. However, according to Oliver and Omari [28], Collin [24], Swan et al. [29], Motiwalla and Tello [23] and Fredericksen et al. [30] the study on asynchronous web-based learning in general reported high levels of students' satisfaction with the courses

While, Malik [31] investigated the factors that influence learner satisfaction towards e-learning and reported that learner's satisfaction is positively influenced by student and instructor attitude towards technology, their computer efficiency, and instructor response, on-line course flexibility and proper facilitation of technical matters. Kuldip and Zoraini [32] studied on a group of Open University Malaysia learners and tutors to determine the E-learning readiness. However the study showed that learners and tutors are moderately ready for e-learning.

## 3.0 Research methodology 3.1 Research Framework





## 3.2Research participants

A total of 250 questionnaires were distributed to learners who are currently undertaking their tertiary education using the traditional classroom in private higher education in Malaysia to express their views and interest in doing their studies the "e" way. Several private higher education providers in Malaysia were selected to provide feedback on their acceptance of e-learning approach.The questionnaires were distributed to learners who are currently pursuing Certificate, Foundation, Diploma and Degree. The collections of data were carried out between 14<sup>th</sup> January 2013 to 21<sup>st</sup> February and we managed to collect 213 responses.

## 3.3Research instruments

Data were collected from learners using questionnaires. The questionnaires were designed using a five-point Likert scale (5-strongly agree, 4-agree, 3neutral, 2-disagree and 1-strongly disagree). The first part of the questionnairegathered information pertaining to learners' profile such as gender, nationality, age, highest education, and currentlevel of programmes pursuing. In addition to demographic characteristics section, the survey instrument consists of five factors. The five factors measured were e-learning readiness (ELREADY), results demonstrability (RD), performance expectancy (PE), effort expectancy (EE), and social influence (SE). (Keller, Christina et al. [33])

### 3.4Data analysis

The data obtained from the questionnaires were analysed using SPSS and the results of descriptive analysis and regression were obtained.

## 4. Findings

From the sample, the male and female respondents are of 46.5% and 53.5% respectively. The distribution on the nationality of respondents; 81.2% are Malaysian and only 18.8% are foreign respondents. The highest education obtained by respondents is from others qualifications (foundation) and followed by SPM. Meanwhile, majority of the respondents are currently pursuing Degree.

Table 1		
Item Reliability		
Analysis		
Independent	No of	Cronbach's
Variable	Items	alpha
RD	2	0.775
PE	8	0.879
EE	6	0.780
SI	3	0.719

A reliability analysis using the measure of Cronbach's alpha was used to estimate the reliability of the Independent Variables. From Table 1 above, the reliability result of independent variables generated in this study is ranged from 0.719 to 0.879. According to Hair, Anderson, Tatham and Black [34], an academic research with Alpha value above 0.7 is generally accepted and this study showed the reliability of the Independent Variables is above the accepted threshold.

In Regression output via the Enter Method, the adjusted R Square showed that the model with the inclusion of the independent variables (RD, PE, EE, SI) collectively explains for 45.5% of the variance in acceptance of e-learning. R<sup>2</sup>ranges from the value of 0 and 1. It showed that there is a linear relationship between acceptance of e-learning and the independent variables (RD, PE, EE, SI). It is possible to predict the influence of independent variables to acceptance of e-learning.

Whereas in the ANOVA (analysis of variance) which showed the significant relationship between the dependent variables and the independent variables in the regression equation. From the results generated, the p value is 0.000 (p<0.05, F=45.175) and this showed that the linear relationship is significant.

Table 2

Results From Regression Analysis

Independent Variable	t	Sig
RD	2.264	0.025
PE	5.865	0
EE	1.491	0.137
SI	1.109	0.269

*RD* - *Results demonstrability, PE* - *Performance Expectancy* 

*EE - Effort Expectancy, SI - Social Influence* 

Table 2 above showed the results of the regression testing. The regression coefficient suggest that RD and PE is positively significant influence acceptance of e-learning with p value generated 0.025 (p<0.05, t=2.264) and 0 (p<0.05, t=5.865), respectively.

Whereas, the independent variables for EE and SI is insignificant influence theacceptance of e-learning with p value generated 0.137 (p>0.05,t=1.491) and 0.269 (p>0.05, t=1.109), respectively.

## 5. Conclusions

In this study, four factorswere tested to study the influence towards acceptance of e-learning among tertiary education learners. These factors are results demonstrability (RD), performance expectancy (PE), effort expectancy (EE) and social influence (SI). The findings of this study revealed that results demonstrability and performance expectancy is significant positively influence the acceptance of elearning. Whereas, effort expectancy and social influence is insignificantly influence the acceptance of e-learning. However, these findings have limitations in terms to generalise to the whole country because the data were only collected from a few private higher education providers.

## REFERENCES

- [1] Davoud, M. (2006). Critical factors for effective e-learning.
- [2] Charmonman, S. (2006, June). Level Elearning in ASEAN and Thailand Paper presented at the B3 – E-Learning, Euro Southeast Asia 2006, Singapore, Thailand
- [3] Chai Lee Goi & Poh Yen Ng (2009). E-learning in Malaysia: Success factors in implementing e-learning program, International Journal of teaching and Learning in Higher Education, 20(2), 237-246.
- [4] Stephenson, J. (2001). Teaching and learning online, pedagogies for new technologies. London, U.K.: Kogan Page Limited.
- [5] Raja Hussain, R. M. (2004).eLearning in higher education institutions in Malaysia. E-mentor, 5(7).
- [6] Jefferies, P., & Hussain, F. (1998).Using the Internet as a teaching resource. Education& Training, 40 (8), 359-365.
- [7] O'Malley, J. (1999). Student perceptions of distance learning, online learning and the traditional classroom. Online Journal of Distance Learning Administration, 2 (4), 1-13.
- [8] Yoon, C. & Kim, S. (2007). Convenience and TAM in a ubiquitous computing environment: The case of wireless LAN. Electronic Commerce Research & Applications, 6(1), 102-112.
- [9] Shin, D.-H. (2007). User acceptance of mobile Internet: Implication for convergence technologies. Interacting with Computers, 19(4), 472-483.
- [10] Moon, J.-W.& Kim, Y.-G. (2001). Extending the TAM for a World-Wide-Web context. Information & Management, 38(4), 217-230.

- [11] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance information technology.MIS Quarterly, 13(3), 319-340.
- [12] Chang, C.C., Yan, C.S., Tseng, J.S., (2012). Perceived convenience in an extended technology acceptance model: Mobile technology and English learning for college students. Australasian Journal of Education Technology, 28(5), 809-826.
- [13] Selim, H.M. (2005). Critical success factors for e-learning acceptance: Confirmatory factor models. Computers and Education.
- [14] Ndubisi, N.O., & Chukwunonso, N.C. (2004).On-line learning adoption intention: Comparing the predictive power of two competing models. Paper presented at the HERDSA 2004.
- [15] Hong, K.S., Lai, K.W., & Holton, D. (2003). Students' satisfaction and perceived learning with a Webbased course. Journal of Educational Technology & Society 6(1).
- [16] Poon, W.C., Low, L.T., & Yong, G. F. (2004). A study of Webbased learning (WBL) environment in Malaysia. The International Journal of Educational Management, 18(6), 374-385.
- [17] Rafaeli, S., & Sudweeks, F. (1997). Networked interactivity. Journal of Computer-Mediated Communications, 2(4).
- [18] Latifah, A.L., &Ramli, B. (2005).Priority-satisfaction survey: A tool in developing effective retention strategies. Paper presented at the Conference on Research in Distance and Adult Learning in Asia.(Open University of Hong Kong, P.R. China, June 20-22, 2005).
- [19] Passmore, D.L. (2000). Impediments to adoption of webbased course delivery among university faculty.ALN Magazine, 4 (2).
- [20] Rosenberg, M.J. (2001). E-learning: Strategies for delivering knowledge in the digital age. New York: McGraw Hill.
- [21] MacDonald, J. (2001). Exploiting online interactivity to enhance assignment development and feedback in distance education. Open Learning, 16(2), 179-189.
- [22] Webb, N.M., Nemer, K., Chizhik, A., & Surgue, B. (1998). Equity issues in collaborative group assessment: Group composition and performance. American Educational Research Journal, 35(4), 607-651.
- [23] Motiwalla, L., & Tello, S. (2000). Distance learning on the Internet: An exploratory study. The Internet and Higher Education, 2 (4), 253-264.
- [24] Collins, M. (2000).Comparing Web, correspondence and lecture versions of a second-year non-major biology course.British Journal of Educational Technology, 31 (1), 21-27.
- [25] Carswell, L. (2000). Distance education via the Internet: The student experience. British Journal of Educational Technology, 31 (1), 29-46.
- [26] Kearsley, G. (2000). Online education: Learning and teaching in cyberspace, Belmont, California: Wadsworth.

- [27] Wegner, S. B., Holloway, K. C., & Gordon, E. M. (1999).The effects of Internet-based instruction on student learning. Journal of Asynchronous Learning Networks, 3 (2).
- [28] Oliver, R., & Omari, A. (2001). Student responses to collaborating and learning in a Web-based environment. Journal of Computer-Assisted Learning, 17 (1), 34-47.
- [29] Swan, K., Shea, P., Fredericksen, E., Pickett, A., Pelz, W., & Maher, G. (2000). Building knowledge building communities: Consistencies, contact and communication in the virtual classroom. Journal of Educational Computing Research, 23 (4), 359-383.
- [30] Fredericksen, E., Pickett, A., Pelz, W., Shea, P., & Swan, K. (2000). Student satisfaction and perceived learning with online courses: Principles and examples from the SUNY Learning Network. Journal of Asynchronous Learning Network, 14 (2).
- [31] Malik, Mahwish W. (2010). Factor Effecting Learner's Satisfaction Towards E-Learning: A Conceptual Framework. OIDA International Journal of Sustainable Development, Vol. 2, No. 3, pp. 77-82.
- [32] Kaur, Kuldip and Zoraini Wati ,Abas (2004) An assessment of e-learning readiness at Open University Malaysia. In: International Conference on Computers in Education 2004.
- [33] Keller, C., Hrastinski,S., and Carlsson, S. (2007). Students' acceptance of e-learning environments: a comparative study in Sweden and Lithuania. Proceedings of the Fifteen European Conference on Information Systems. University of St. Gallen.
- [34] Conference (Miri, Proceedings ascilite Melbourne 2008: Full paper: Lim, Hong & Tan 550 Sarawak, July, 4-7 2004).