

## Chicken eggshell as biosorbent: Artificial intelligence as promising approach in optimizing study

Peck Loo Kiew<sup>1</sup>, Chun Kit Ang<sup>1</sup>, Khang Wei Tan<sup>1</sup> and Shu Xin Yap<sup>1</sup>

<sup>1</sup>Faculty of Engineering, Technology & Built Environment, UCSI University, 56000 Cheras Kuala Lumpur, Malaysia

**Abstract.** Response Surface Methodology (RSM) is the most popular approach for optimization study in various biochemical processes nowadays. Artificial Neural Network (ANN) has emerged as one of the most efficient methods in empirical modeling and optimization, particularly for non-linear systems. In this study, the estimation capability of RSM and ANN models was compared in copper removal from aqueous solution. The experiments were carried out based on a 3-level and 4-variable Central Composite Design (CCD). The RSM results revealed that the relationship between the response and independent variable could be represented by the quadratic polynomial model. In the development of ANN model, the optimal configuration of the model was found to be 4-10-1. Estimated responses from both models were compared with the experimentally determined responses to determine predictive capabilities of both techniques. Comparison of two methodologies showed that the ANN model was more accurate and exhibited better generalization capability than RSM, thus indicated a clear superiority than the latter in capturing the non-linear behaviour of the adsorption process using chicken eggshell as biosorbent.

### 1 Introduction

Chicken eggshells are waste material from hatcheries, homes and fast food industries which are abundant and readily available worldwide [1]. Without proper management, its disposal contributes to environmental pollution. A recent report described that chicken eggshells contain calcium and trace amounts of other micro elements (boron, copper, iron, manganese, molybdenum, sulphur, magnesium, silicon and zinc). The presence of cellulosic structure and amino acids suggested it a good biosorbent [2]. Calcium carbonate was identified as the main element responsible for metal adsorption [3]. Several studies in recent years concentrating on the evaluation of the possibility of heavy metal removal, particularly copper, fluoride, iron and chromium from aqueous solution had shed lights on the application of discarded chicken eggshell in wastewater treatment [2, 4-6]. In Malaysia, chicken eggshells are discarded from food industry or household as municipal solid waste. They are normally disposed in landfills. Even though these waste materials are available in abundance, its potential in heavy metal removal is still less being investigated.

The efficiency of copper (heavy metal) biosorption using chicken eggshell powder can be influenced by many factors such as particle size, biosorption temperature, solution pH, ratio of sorbent to solution and duration of biosorption [2, 7, 8]. Consequently, the biosorption process or design of experiments must be conducted or designed carefully. Various experimental

design techniques had been employed for biosorption process which included the conventional One-Factor-At-A-Time (OFAT) method and stactical methods such as Central Composite Design (CCD) and Box-Behnken Design (BBD). Under circumstances where multiple process variables may influence the response variables, response surface methodology (RSM) emerges as a favourable technique in optimizing the process [3, 9]. RSM has been extensively employed in various biochemical applications for the purpose of empirical modeling and optimization. It is a collection of statistical techniques for experiment designs, model buildings, evaluation of effects of process factors, as well as identification of optimum conditions, through fitting of the experimental responses to design of experiments (DOEs) into quadratic function [10]. The applications of RSM involve three major steps that start with establishment of experimental design using CCD and BBD to set the experimental level of independent variables. Subsequently, regression analysis will be conducted for the response surface modeling. Finally, the response surface mode will be used as objective function for the process optimization [11]. RSM had been commonly employed in various bioprocesses such as extraction, fermentation and absorption processes. The ultimate objective of the application is to identify the optimum operating conditions for the respective processes.

In most of the RSM problems, the interaction between the response and process variable is not known. Hence, suitable approximation for the true functional relationship

between the respond,  $y$ , and the set of process variables,  $x_1, x_2, x_3, \dots, x_n$  must be established where  $n$  is the total number of process variables in the process or system. The function can be expressed as:

$$y = f(x_n) + \varepsilon \quad (1)$$

with  $\varepsilon$  refers to the noise or error observed in the response  $y$ . The true response function  $f$  is normally unknown, thus low-order polynomial in some region of the independent variables is preferable for the approximation of  $f$ . If the response can be defined by a linear function of the independent variables, then the approximation function is a first-order model expressed by:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots + \beta_k x_k + \varepsilon \quad (2)$$

In the case of curvature in the response surface, a higher degree polynomial can be used (e.g. the second order model) as follow:

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=2}^k \beta_{ii} x_i^2 + \sum_{1 \leq i < j}^k \beta_{ij} x_i x_j + \varepsilon \quad (3)$$

Nevertheless, neural network-based models have recently been shown by several researchers to exhibit superior performance in terms of accuracy and predictive capability [12]. Artificial Neural Network (ANN) has emerged as an attractive tool for non-linear multivariate modeling in the last two decades. It has typically been used as a "black-box" tool, that is, no prior knowledge about the process is assumed but the goal is to develop a process model based only on observations of its input-output behavior [13]. ANN is generic in structure and it exhibits the ability to learn from data gathered by users, thus widely known as a reliable modeling tool for process that is lack of information [10]. ANN is capable of handling multiple independent and dependent variables simultaneously and to do this prior knowledge on the functional relationship does not need to be known. Each neuron receives information through input connections, process the information and produce the output which is distributed through output connections. Each connection is associated with a real number quantity known as weight [14]. In other words, the neurons sum weighted inputs and then applies a linear or non-linear function to the resulting sum to determine the output. These neurons are arranged in layers and are combined through excessive connectivity [14]. A neural network in its basic form is usually composed of several layers of neurons, there being one input layer, one output layer and at least one hidden layer.

The error or bias in prediction is then propagated through the system and the inter-unit connections are changed to minimize the error in prediction [15]. This is a continuous process with multiple training sets until the minimum error is attained. Training an ANN network is therefore accomplished by adjusting these weights by minimizing a non-linear error function [16]. However, a major disadvantage of ANN is the difficulty in explaining the relation between independent and response variables resulting from the ambiguously defined weights, which as mentioned earlier, is a black box [17]. According to

Agatonovic-Kustrina et al. [15], 3 types of data sets that are inclusive of training data, testing data and validation or unseen data, are required for network training, neural network performance monitoring during the training, and measurement of the performance of trained network, respectively. The number of the layers and processing elements in layers vary from one process to another. There is no strict rule available to identify the optimal number of layers and neurons. In fact, it is generally defined by trial and error [18].

It is possible to investigate the effect of each independent process variable, alone or in combination in the process through the least number of experiments through the application RSM [18]. Yet, drawbacks associated with it have been pointed out recently. Desai et al. [16] and Lou and Nakai [17] stated that RSM was useful only for quadratic approximations and the sensitivity analysis of input variables was difficult to perform using RSM due to the presence of cross interactions. It is therefore hard to conclude that RSM is applicable to all optimization and modeling studies. Consequently, the main objective of the present work is to investigate the possibility of employing artificial intelligence approach, particularly through the application of artificial neural network to construct a nonlinear relationship among the process parameters and their corresponding responses. The ANN used in this work is the feedforward back-propagation network which is normally applied in analytical applications and is very powerful in function optimization modeling. To our best knowledge, this is the first report on the comparison of RSM and ANN methodologies in predicting the copper removal efficiency from aqueous solution using eggshell as adsorbent. Also, changes on the surface morphology of the eggshell as biosorbent for copper removal was observed and comparison was made before and after the adsorption process.

## 2 Materials and Methods

### 2.1 Materials

Chicken eggshells were collected from the restaurants or market around UCSI University Kuala Lumpur (North Wing). The sample was washed with distilled water several times to remove dirt and contaminants, followed by drying in a hot air oven at 110°C for at least 12 hours. The chicken eggshells were then grinded or crushed using pestle and mortar. The sample was stored in an airtight container for future use.

### 2.2 Chemicals

Pentahydrate copper (II) sulphate was purchased from R&M Chemicals (Malaysia).

### 2.3 Biosorption Process

The adsorption process was conducted based on the method reported by Nabil et al. [5] with slight

modifications. Chicken eggshell was mixed with copper (II) sulphate solution. Agitation was made at a constant stirring speed of 150 – 300 rpm for 60 – 180 minutes. The remaining concentration of Cu(II) after adsorption was measured using UV-Vis spectrophotometer.

## 2.4 Experimental Design

### 2.4.1 Response Surface Methodology (RSM)

RSM was employed for experimental design, data analysis, and model building with the aid of the software Design Expert (Version 6.0.6, Stat-Ease Inc., Minneapolis, Minnesota USA). Central Composite Design (CCD) with four variables was used to determine the response pattern and subsequently to establish a response model. According to CCD, the total number of experimental combinations is  $2^k + 2k + n_0$ , where k is the number of independent variables and  $n_0$  is the number of repetitions of the experiments at the centre point. Four independent variables used in this work were temperature ( $X_1$ ), agitation speed ( $X_2$ ), dosage ( $X_3$ ), and time ( $X_4$ ), while the dependent variable was the percentage of copper removal. The ranges and center point values of all independent variables were based on the results of preliminary experiments (Table 1). The copper removal percentage was analyzed by multiple regressions to fit into the following polynomial equation:

$$Y = \beta_0 + \sum_{i=1}^4 \beta_i X_i + \sum_{i=1}^4 \beta_{ii} X_i^2 + \sum_{i=1}^3 \sum_{j=i+1}^4 \beta_{ij} X_i X_j \quad (4)$$

where Y is dependent variable (copper removal percentage),  $\beta_0$  is intercept,  $\beta_i$ ,  $\beta_{ii}$ ,  $\beta_{ij}$  are regression coefficients,  $X_i$  indicates the linear terms,  $X_i^2$  for the quadratic terms for a single variable, and  $X_i X_j$  for the interaction terms. Four replicates at the central point of the designed model were used to estimate the pure error sum of squares. 30 experiments were performed in random order to cover all combinations of the factor levels in the experimental design.

**Table 1.** Experimental range and values of the independent variables used in CCD

| Independent Variables | Symbol | Coded Levels |     |     |
|-----------------------|--------|--------------|-----|-----|
|                       |        | -1           | 0   | 1   |
| Temperature (°C)      | $X_1$  | 30           | 40  | 50  |
| Agitation speed (rpm) | $X_2$  | 150          | 250 | 350 |
| Dosage (g)            | $X_3$  | 0.2          | 0.6 | 1.0 |
| Time (min)            | $X_4$  | 60           | 120 | 180 |

### 2.4.2 Development of Feedforward Artificial Neural Network (FANN)

The percentage of copper removal was evaluated as the output. Using MATLAB (version 7.8.0.347, Mathworks, Natick, MA), a feedforward artificial neural network (FANN) trained with an error-back-propagation algorithm was designed. All weights and biases were randomly initialized in the range from 0 – 1. Neural network building data were trained and Mean Squared Error (MSE) was plotted against training epoch. When

MSE started to rise, the training was terminated and the value of weight that the network had in that previous step was chosen as the result of the training run. The development of ANN (activation functions) adopted in this study was in accordance to the theory and approach suggested by Ang et al. [19].

The FANN model was developed based on the experimental data of RSM (Table 2). To cope with different magnitudes in the input and output, all the data were scaled to zero mean and unit standard deviation. Single hidden layer neural networks with different numbers of hidden neurons were trained on the training data and tested on the testing data. The network with the lowest MSE on the testing data was considered as having the best network topology. In assessing the developed models, MSE and correlation coefficient ( $R^2$ ) on the unseen validation data were used as the performance criterion.

## 2.5 Characterization Study

Scanning electron microscopy (SEM) analysis was performed to observe structural and morphology changes in eggshell surfaces both before and after adsorption process. SEM images of the eggshells were obtained using JEOL JSM 6400 LV model scanning electron microscope at 15 kV.

## 3 Results and Discussion

### 3.1 Response Surface Methodology

All 30 experimental points were evaluated and the percentages of copper removal results based on the factorial design are shown in Table 2. Response surface optimization is favorable over the conventional single parameter optimization in that it saves time, space, and raw materials [11]. The experiments were run at random in order to minimize errors from the systematic trends in the variables. The percentage of copper removal obtained ranged from 4.55 to 59.1 %.

**Table 2.** Response surface CCD and results for copper removal percentage

| Run | $X_1$ | $X_2$ | $X_3$ | $X_4$ | Copper Removal (%) |
|-----|-------|-------|-------|-------|--------------------|
| 1   | 40    | 250   | 0.6   | 120   | 50                 |
| 2   | 40    | 250   | 0.6   | 120   | 45.45              |
| 3   | 40    | 250   | 0.6   | 120   | 50                 |
| 4   | 40    | 250   | 0.6   | 60    | 50                 |
| 5   | 40    | 250   | 0.6   | 60    | 45.45              |
| 6   | 40    | 250   | 0.6   | 60    | 45.45              |
| 7   | 30    | 150   | 0.2   | 60    | 13.64              |
| 8   | 50    | 150   | 0.2   | 60    | 9.1                |
| 9   | 30    | 350   | 0.2   | 60    | 31.82              |
| 10  | 50    | 350   | 0.2   | 60    | 31.82              |
| 11  | 30    | 150   | 1.0   | 60    | 34.54              |
| 12  | 50    | 150   | 1.0   | 60    | 4.55               |
| 13  | 30    | 350   | 1.0   | 60    | 45.45              |
| 14  | 50    | 350   | 1.0   | 60    | 13.63              |
| 15  | 30    | 150   | 0.2   | 180   | 45.45              |
| 16  | 50    | 150   | 0.2   | 180   | 59.1               |
| 17  | 30    | 350   | 0.2   | 180   | 40.9               |
| 18  | 50    | 350   | 0.2   | 180   | 54.55              |

|    |    |     |     |     |       |
|----|----|-----|-----|-----|-------|
| 19 | 30 | 150 | 1.0 | 180 | 59.1  |
| 20 | 50 | 150 | 1.0 | 180 | 59.1  |
| 21 | 30 | 350 | 1.0 | 180 | 59.1  |
| 22 | 50 | 350 | 1.0 | 180 | 45.45 |
| 23 | 30 | 250 | 0.6 | 120 | 45.45 |
| 24 | 50 | 250 | 0.6 | 120 | 44.54 |
| 25 | 40 | 150 | 0.6 | 120 | 50    |
| 26 | 40 | 350 | 0.6 | 120 | 59.1  |
| 27 | 40 | 250 | 0.2 | 120 | 20    |
| 28 | 40 | 250 | 1.0 | 120 | 27.27 |
| 29 | 40 | 250 | 0.6 | 60  | 27.27 |
| 30 | 40 | 250 | 0.6 | 180 | 59.1  |

Exploration and optimization of a fitted response surface might produce poor or misleading results unless the model exhibits a good fit, which makes the checking of the model adequacy essential. Through the analysis of variance (ANOVA), the *P*-value of the model was less than 0.0001. Meanwhile, the lack of fit value of the model was not significant at 0.1417. These values confirmed that the model fitness was good.

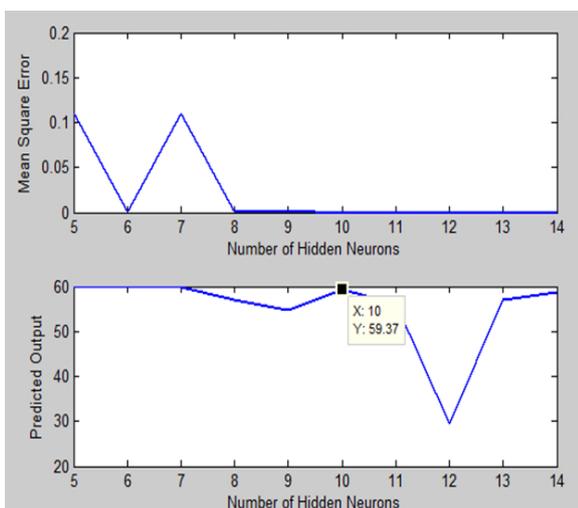
In order to check on the adequacy of the predictive model, 3 sets of experiments were repeated at the suggested optimum conditions in order to experimentally obtain the maximum copper removal using eggshell. As shown in Table 3, the percentage error differences between the experimental and predicted values were always less than 1.5 %. Since the differences between the experimental and predicted response were less than 5%, the validity of the model was verified [10].

**Table 3** Validation of data and constructed RSM model

| Run     | Experimental removal (%) | Predicted removal (%) | Error (%) |
|---------|--------------------------|-----------------------|-----------|
| Trial 1 | 59.1                     | 58.88                 | 1.31      |
| Trial 2 | 59.1                     | 58.88                 | 1.31      |
| Trial 3 | 59.1                     | 58.88                 | 1.31      |
| Average | 59.1                     | 58.88                 | 1.31      |

### 3.2 Network Topology and Training of Artificial Neural Network (ANN) Model

The copper removal adsorption process using eggshell was conducted in 30 batches contributed to 30 samples data. To create the model, these data were re-sampled

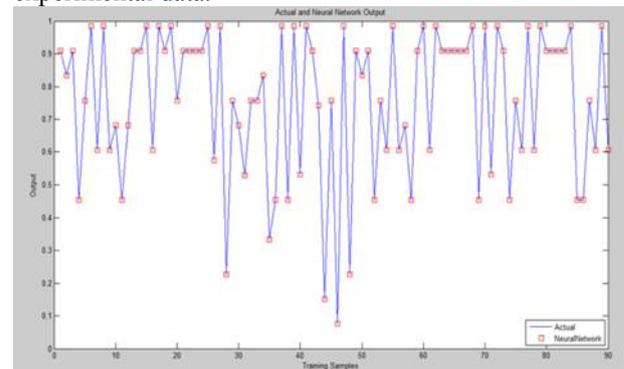


using bootstrap method into 90 data for training purposes.

**Figure 1.** Selection of number of neurons in the hidden layer.

Several structures were tested to identify the best network topology based on trial and error method. Variation in number of neurons in the hidden layer ranged from 5 – 14. Subsequently, the network was trained and tested after each addition of neuron. There was a drastic drop in the MSE obtained after number of neurons used was more than 7 (Figure 1). With the outputs predicted by ANN trained with different number of neurons, the least error between the predicted and validation result (unseen) was obtained with 10 neurons in the hidden layer. The configuration of the ANN model could therefore be defined as 4-10-1.

The predicted and actual values for the training data matched with each other as shown in Figure 2. In this study, the ANN training resulted in unity  $R^2$  value, proving that the ANN model could fit well with experimental data.



**Figure 2.** Actual and predicted output for training data.

Validation of the trained model was conducted at the suggested optimum conditions by ANN and it was predicted that 59.37 % of copper removal could be achieved. 3 sets of validation experiments were repeated. The percentage error differences between the experimental and predicted values were less than 1 % (Table 4). Therefore, the validity of the trained ANN model was verified.

**Table 4** Validation of data and constructed ANN model.

| Run     | Experimental removal (%) | Predicted removal (%) | Error (%) |
|---------|--------------------------|-----------------------|-----------|
| Trial 1 | 59.10                    | 59.37                 | 0.46      |
| Trial 2 | 59.10                    | 59.37                 | 0.46      |
| Trial 3 | 59.10                    | 59.37                 | 0.46      |
| Average | 59.10                    | 59.37                 | 0.46      |

### 3.3 Validation and Comparison of Both Predictive Models

The estimation capabilities between the statistical- and artificial intelligence-based optimization techniques namely RSM and ANN, respectively, were compared and examined. The comparison was made on the basis of various parameters such as the correlation coefficient ( $R^2$ ) for both training and validation sets, predicted copper removal percentage by both techniques and the average error of the experimental result based on the optimum

conditions suggested by RSM at temperature of 49.26 °C, agitation speed of 233.38 rpm, chicken eggshell dosage of 0.7 g for adsorption process of 178.7 minutes. Desai et al. [16] stated that the generalization ability could best be judged with the unseen (validation) dataset. Therefore, the comparative values are tabulated in Table 5.

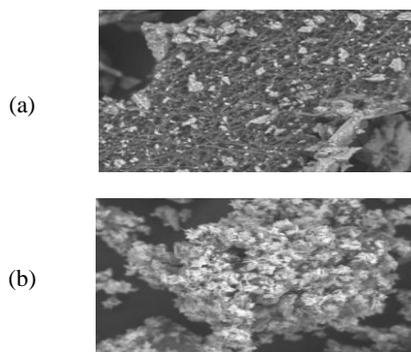
**Table 5** Comparison of predictive capability between RSM and ANN model

| Parameters                                 | RSM    | ANN   |
|--|--------|-------|
| Training $R^2$                             | 0.9734 | 1.000 |
| Predicted Copper Removal Percentage (%)    | 59.88  | 59.37 |
| Experimental Copper Removal Percentage (%) | 59.10  | 59.10 |
| Copper Removal Percentage Error (%)        | 1.31   | 0.46  |

RSM was one of the most popular and widely used techniques in optimization studies of chemical and biochemical processes. Nonetheless, it had limitations that modeling of the processes could be represented by only up to second-order equations [14]. However, it could be possible that the relationship between the process variable and response variable is more complex than that. Therefore, in the present study, the  $R^2$  values of training datasets showed that ANN model resulted in significant higher generalization and predictive capability as compared to RSM. Also, it was evidenced by the validation of the predictive results in which prediction suggested by RSM resulted in greater deviation (error) compared to ANN model prediction (Table 5).

### 3.4 Characterization Result

Scanning electron microscopy (SEM) in Figure 3 (a) shows the morphology of powdered eggshell. It confirmed the crystallinity of eggshell and the existence of pores on the surface of the eggshell. After the adsorption process, success in copper removal was evidenced by the deposition of flake-like copper crystals on the surface of the eggshell. No obvious porous structure was observed on the surface after the adsorption process (Figure 3 (b)).



**Figure 3.** SEM image of eggshell  
(a) Before adsorption and (b) After adsorption

Two models were developed to predict the percentage of copper removal using response surface methodology (RSM) and artificial neural network (ANN). A feedforward artificial neural network (FANN) was trained by the sets of input-output patterns by implementing the error-back-propagation algorithm. The present study revealed that ANN technique performed better than RSM. Thus, it could be concluded that even though RSM was widely applied in various optimization studies, ANN optimization strategy might present a better alternative to the conventional method. The results of this study also suggested that powdered eggshell, with its low cost and abundant availability has a potential for being used as an adsorbent for contaminants such as copper in wastewater.

### References

- King'or, A.M, Int. J. Poult. Sci. **10**, 11 (2010).
- Bhaumik, R., N. K. Mondal, B. Das, P. Roy, K. C. Pal, C. Das, A. Baneerjee, and J. K. Datta, E-J. Chem. **9**, 3 (2012).
- Tsai, W. T., J. M. Yang, C. W. Lai, Y. H. Cheng, C. C. Lin, and C. W. Yeh, Bioresource Technol. **97**, 3 (2006)
- Nabil, A. J., and A. A. Sameer. J. Envir. Prot. Sci. **3**, (2009).
- Aimi, R. N., A. H. Norhafizah, and C. S. Wong, Int. J. Eng. Technol. **13**, 1 (2013).
- Rubcumintara, T., J. Clean Energ. Technol. **2**, 2 (2014).
- Feng, N. C., X. Y. G. Guo, and S. Liang, Trans. Nonferrous Met. Soc. China. **20**, 1(2010).
- Liu, C. L., H. H. Ngo, W. Guo, and K. L. Tung. Bioresource Technol. **119** (2012).
- Singh, R., R. Chadetrik, R. Kumar, K. Bishnoi, D. Bhatia, A. Kumar, N. R. Bishnoi, and N. Singh. J. Hazard Mat. **174**, 1-3 (2010).
- Kiew, P. L., A. Zainal, and M. D. Mashitah. Biotechnol. Bioprocess Eng. **18** (2013).
- Zhang, L. L., M. Xu, Y. M. Wang, D. M. Wu, and J. H. Chen. Molecules. **15**, 11 (2010).
- Chou, C. J., and L. F. Chen. Int. J. Prod. Res. **50**, 7 (2012).
- Zainal, A., M. D. Mashitah, M. Siti Hatijah, and M. N. Rabiatal Adawish. Bioprocess Biosys. Eng. **5** (2010).
- Shen, C., L. Wang, and Q. Li. J. Mat. Process. Technol. **183**, 2-3 (2007).
- Agatonovic-Kustrina, S., M. Zecevic, L. J. Zivanovic, and I. G. Tucker. 1998. J. Pharm. Biomed. Anal. **17**, 1 (1998).
- Desai, K. M., S. A. Survase, S. S. Lele, and R. S. Singhal. Biochem. Eng. J. **41**, 3 (2008).
- Lou, W., and S. Nakai. Food Res. Int. **34**, 7 (2001)
- Baş, D., and İ. H. Boyacı. J. Food Eng. **78**, 3 (2007).
- Ang, C.K., S. H. Tang, S. Mashohor, M.K.A.M. Arrifin. Int. J. Comp. Comm. Contr. **9**, 3 (2014).

### 4 Conclusion