

PREDICTIVE ANALYSIS OF NETWORK TRAFFIC RESOURCES UNDER CLOUD COMPUTING BASED ON NEURAL NETWORK

Ping Wang

Guangling College, Yangzhou University, Yangzhou, Jiangsu 225000, China

E-mail: pingwmail@163.com

Abstract - With the rapid development of the Internet, the network has provided us with many convenient services, and the network traffic resources have also increased greatly. The traditional traffic prediction methods have many limitations and shortcomings, which poses severe challenges to network management. The network environment is also worrying. In this study, based on the gray neural network model, the network traffic data forecasting method under cloud computing was used to study the network traffic data. By sorting and summarizing the network traffic data in a period of time, the corresponding model was established to test the data. The obtained error value of the model of the gray neural network prediction was 1.5896, its relative error value was 6.32%, and its variance value was 0.533, indicating that the error of the gray neural model was small. The experimental conclusion that this new network model was better than the traditional method was obtained, which opens a new path for network traffic resource forecasting.

Keywords: Network Traffic Resource, Neural Network, Cloud Computing, Predictive Analysis.

1. Introduction

With the rapid development of Internet technology, the network team has grown, the number of network users has doubled, the network transmission speed and capacity have also grown rapidly, and the rapid development of network applications and service types has made the Internet thrive and gradually become an integrated platform for data, language and services such as a multimedia video. However, these phenomena also make network measurement and modeling become a difficult problem [1]. In response to this problem, many experts and scholars have put forward their own opinions. Chen et al. [2] proposed a novel network architecture, HashedNets, which used the inherent redundancy in neural networks to significantly reduce the size of the model.

This structure was demonstrated on several benchmark data sets. The requirements of neural network storage were much lower, while retaining most of the generalization performance. Ganin et al. [3] proposed a new learning method of domain adaptive representation in which data for training and testing came from similar but different distributions. This method was directly inspired by the domain adaptation theory and was implemented in the context of neural network architecture. This indicated that in order to achieve effective domain transfer, prediction must be based on features that could not distinguish between training (source) and test (target) domains. Mauro et al. [4] believed that mobile devices might be exploited maliciously to invade the privacy of network users.

Therefore, the investigation for such external attackers could identify the specific operations performed by users on their mobile applications, design a system that used advanced machine learning techniques to achieve this goal and also run a complete set of experiments. These experiments showed that for most of the operations considered, the attack could achieve more than 95% accuracy and precision. Manisha and Joshi [5] proposed that the current environment was experiencing the period of users connecting to the network utilization.

Thus, clustering was needed to enhance network traffic prediction. In this study, the simulation experiment of network traffic resource prediction based on gray neural network model was used to effectively improve and optimize the network traffic data prediction model, and the accuracy and practicability of the prediction method were further illustrated, which provides a reference for network traffic resource prediction.

2. Network Traffic Resource

Data collection and monitoring subsystem of network traffic

After the data collection and monitoring subsystem of network traffic is used to obtain the relevant traffic data, the network user can be provided with the network traffic distribution status in time. Thus, the network user can control the network traffic resources of each interface in real time and timely

regulate the network congestion phenomenon. Through the system, the obtained data can be classified, summarized, counted and stored, and finally the effective network traffic data information is provided to the network user to play back network traffic and historical trend graph.

Figure 1 shows the components of the data collection and monitoring subsystem of network traffic.

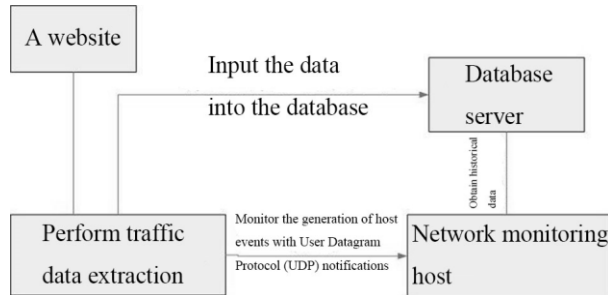


Figure 1: The components of the data collection and monitoring subsystem of network traffic

Network traffic resources and cloud computing

Cloud computing is an Internet-based computing method. At present, the storage technology of cloud computing is represented by GFS and Hadoop. Most of the data collection now uses tools such as WireShark and Winpcap, but the amount of data processed has reached PB or higher. The traditional way cannot keep up with the updated data, but with the rapid development of Internet technology, the ability of the cloud platform to process data has been more comprehensive. For Hadoop-based cloud computing technology, one of the major advantages is that it can effectively detect and process network traffic data resources [6]. It divides a large problem into multiple small problems and then solves these small problems based on the Hadoop platform. At the same time, it summarizes these scattered results to obtain the final network traffic monitoring result [7].

3. Prediction Algorithm of Gray Neural Network

Gray model prediction

The predictive model established in the gray system is called the gray model (GM), which is a differential equation based on the original data sequence. Without knowing the characteristics of the system, the system information is used to quantify the abstract model to predict the output of the system [8].

The specific modeling steps are as follows. First, through qualitative analysis, the direction and goals of the research are elaborated. Then, the relationship

between the factors in the language model and various factors is analyzed to establish a quantitative model, study quantitatively the causal relationship of each link and obtain the aggregated quantitative relationship. Second, dynamic modeling is performed to further collect input and output data for each link and the establishment of GM (1, 1).

Finally, the model is analyzed, and the solution is optimized.

Suppose that the time series have n observation data, the original sequence is $x^{(0)} = x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)$, and $x^{(1)} = x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)$ is obtained by generating a new data sequence by one accumulation. The dynamics of the new data sequence can be approximated by the differential equation in (1):

$$\frac{dx^{(1)}}{dt} + a \bullet x^{(1)} = b \tag{1}$$

Where $x^{(1)}$ is a sequence generated by an accumulation process, a is development gray number, b is an endogenous control gray number, and t is time value.

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B^T)^{-1} B^T Y \tag{2}$$

Where Y is column vector, and B is constructed data matrix, which are shown as follows:

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \dots & \dots \\ -\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{bmatrix}$$

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

$z^{(1)}$ is set as the numerical series of mean generation with consecutive neighbors of $x^{(1)}$,

$$z^{(1)}(k) = \frac{1}{2}[x^{(1)}(k-1) + x^{(1)}(k)]$$

Thus, the gray differential equation for GM (1,1) is as follows:

$$x^{(1)}(k) + a \bullet z^{(1)}(k) = b \tag{3}$$

(2) is the local equation of (3), also known as the shadow equation.

The solution to the whitening equation is also called the time response function, as shown in (4):

$$x^{(1)}(t) = [x^{(0)}(0) - \frac{b}{a}]e^{-at} + \frac{b}{a} \tag{4}$$

The series of time response functions of (3) is:

$$x^{(1)}(k+1) = [x^{(0)}(0) - \frac{b}{a}]e^{-ak} + \frac{b}{a}, k = 1, 2, \dots, n \tag{5}$$

$x^{(1)}(k+1) = x^{(1)}(1)$ is set, and there is

$$x^{(1)}(k+1) = [x^{(1)}(1) - \frac{b}{a}]e^{-ak} + \frac{b}{a} \tag{6}$$

The regressive prediction equation is:

$$\begin{aligned} x^{(0)}(k+1) &= x^{(1)}(k+1) - x^{(1)}k \\ &= (1 - e^{-a}) \times [x^{(0)}(1) - \frac{b}{a}]e^{-ak} + \frac{b}{a}, k = 1, 2, \dots, n \end{aligned} \tag{7}$$

Back-propagation (BP) neural network model

The multi-layer neural network model with BP algorithm is generally called BP neural network, and BP algorithm is the most widely used learning method [9]. The whole is roughly divided into two parts, one is the forward propagation of the signal, and the other is the back propagation of the error [10]. It consists of different parts which play their own roles, have powerful sensing functions, and can achieve multilevel network learning, giving readers a unique experience.

$$net_k = \sum_{i=1}^q w_{ki} y_i + a_k = \sum_{i=1}^q w_{ki} \phi(\sum_{j=1}^M w_{ij} x_j + \theta_i) + a_k \tag{11}$$

The output o_k of the k -th node of the output layer is as follows:

$$o_k = \psi(net_k) = \psi(\sum_{i=1}^q w_{ki} y_i + a_k) = \psi[\sum_{i=1}^q w_{ki} \phi(\sum_{j=1}^M w_{ij} x_j + \theta_i) + a_k] \tag{12}$$

• **BP processing of error**

It is assumed that the number of samples of the trained BP nerve is single, and the error value of a single data sample can be obtained from the total error criterion function:

$$E_p = \frac{1}{2} \sum_{k=1}^L (T_k - o_k)^2 \tag{13}$$

Where E_p is the error value of the actual output of a single data sample. The first layer is set as the input layer and the L -th layer is the output layer.

T is set to be the training set. If the number of training samples is P , then the error value of the P sample data can be obtained according to the total error criterion function:

$$\Delta w_{ij} = \eta \sum_{p=1}^P \sum_{k=1}^L (T_k^p - O_k^p) \cdot \psi'(net_k) \cdot w_{ki} \cdot \phi'(net_i) \cdot x_j \tag{15}$$

$$\Delta \theta_i = \eta \sum_{p=1}^P \sum_{k=1}^L (T_k^p - O_k^p) \cdot \psi'(net_k) \cdot w_{ki} \cdot \phi'(net_i) \tag{16}$$

• **Forward propagation of signals**

It is assumed that the signal size of the i -th node in the hidden end of the system is net_i , P is set as the input signal of the BP neuron, w and b are respectively the weight and the threshold, y is the result of the processing, θ_i is the threshold of the hidden neuron, q is the number of units of the output layer, M is the total number of samples, and the actual network parameter is represented by $\theta = (w, b)$.

The formula for the log sig function is:

$$a = \log sig(w + b) \tag{8}$$

$$net_i = \sum_{j=1}^M w_{ij} x_j + \theta_i \tag{9}$$

Where net_i is the input for the i -th neuron.

The output o_i of the i -th node of the hidden layer is as follows:

$$o_i = \phi(net_i) = \phi(\sum_{j=1}^M w_{ij} x_j + \theta_i) \tag{10}$$

The input net_k of the k -th node of the output layer is as follows:

$$E_p = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^L (T_k^p - O_k^p)^2 \tag{14}$$

η is the learning rate, and the system output layer connection parameters can be obtained by modifying the connection parameter values in the system based on the gradient descent method. At the same time, the change amount Δa_k of output layer threshold, the change amount Δw_{ij} of hidden layer weight, and the change amount $\Delta \theta_i$ of hidden layer threshold are obtained. Finally, the weights of the input and output layers of the BP neural network and the correction values of the thresholds are obtained:

$$\Delta w_{ki} = \eta \sum_{p=1}^P \sum_{k=1}^L (T_k^p - O_k^p) \cdot \psi'(net_k) \cdot y_i \tag{17}$$

$$\Delta a_k = \eta \sum_{p=1}^P \sum_{k=1}^L (T_k^p - O_k^p) \cdot \psi'(net_k) \tag{18}$$

Prediction of traffic by gray neural network

In order to compensate for the defects of the gray model and the neural network model, the gray model and the neural network are optimized and combined by weighting to obtain the gray neural network prediction model [11]. Through this new method, the predictive performance of network traffic data can be effectively improved and optimized [12]. Its structure is shown in Figure 2.

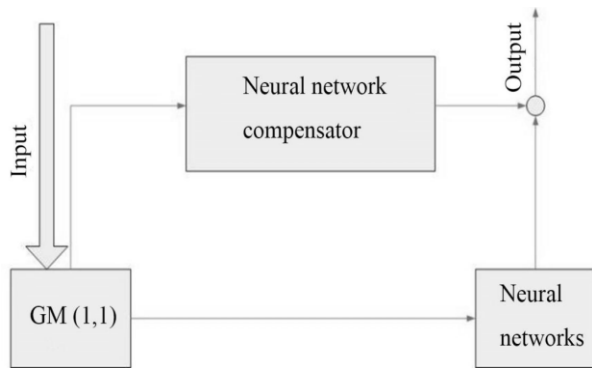


Figure 2: Structure of gray neural network combination

The specific modeling steps are as follows:

First, the original data information of the network traffic formed over a period of time is obtained, and a sequence is generated to establish a GM (1, 1) model. Then, by predicting the first to the n-th data information of the sequence of the GM (1,1) model, a new data sequence P of length n can be obtained, where n T data sequences are obtained from the first to the n-th data of the original sequence.

Second, the new sequence P is used as the input vector of the neural network, while T is used as the output vector of the neural network, and the weights and thresholds are set.

The BP neural network is trained to obtain the weights and thresholds for each node in the network.

The GM (1,1) is used to perform the prediction, the predicted value is used as the input of the network, and the relevant data is brought into the simulation experiment to obtain the corresponding output result. Under the condition of supplementing the neural network compensator, the network input is the original sequence used for modeling, and the network output is to obtain the error compensation amount.

Using the difference between two consecutive data in the original sequence as the input to the network, the goal of reducing the number of input variables and speeding up the network training speed is achieved.

Finally, the predicted output of the GM (1,1) model and the predicted output through the neural network compensator are added to obtain the output of the combined neural network.

It is assumed that x_{h1} is the predicted value obtained by the neural network prediction model in the h-th period, x_{h2} is the predicted value obtained by the gray prediction model in the h-th period, Y_h is the predicted value of the combined model in the h-th period, c_1 is the weight coefficient of neural network model, and c_2 represents the weight coefficient of the gray model. Thus,

$$Y_h = c_1 x_{h1} + c_2 x_{h2}, t = 1, 2, \dots, n \tag{19}$$

and

$$\sum_{k=1}^n c_k = 1, k = 1, 2.$$

e , e_1 and e_2 respectively represented the error values of the gray prediction model, the neural network prediction model, and the combined prediction model, $x^{(0)}(h)$ represented the original sequence, and the error values of each model can be expressed as:

$$\begin{cases} e = x_{h1} - x^{(0)}(h) \\ e_1 = x_{h2} - x^{(0)}(h) \\ e_2 = Y_h - x^{(0)}(h) \end{cases} \tag{20}$$

In order to improve the accuracy of the prediction model, the sum of the squared differences of the models needs to be minimized, which is represented by the following formula:

$$\min \sum_{h=1}^n |Y_h - x^{(0)}(h)|^2 = \min \sum_{h=1}^n |c_1 e + c_2 e_1|^2 \tag{21}$$

Each prediction value is set to have a probability of P (assuming each acquisition probability is equal), and the sum of the squared errors of the combined prediction Y_h is the minimum value, which can be converted into the minimum variance (Cov represents covariance):

$$\begin{aligned}
 \text{Var}(e_2) &= \text{Var}[Y_h - x^{(0)}(h)] \\
 &= E[e_2 - E(e_2)]^2 = \sum_{h=1}^n e_h^2 P \\
 &= P \sum_{h=1}^n [Y_h - X^{(0)}(h)]^2 \\
 &= c_1^2 \text{Var}(e) + c_2^2 \text{Var}(e_1) + 2c_1 c_2 \text{Cov}(ee_1)
 \end{aligned}
 \tag{22}$$

The partial derivative of $\text{Var}(e)$ with regard to c_1 is calculated, the minimum value is found, and the minimum value of the combined prediction Y_h can be found through this step, then

$$c_1 = \frac{\text{Var}(e_1) - \text{Cov}(e, e_1)}{\text{Var}(e) + \text{Var}(e_1) - 2\text{Cov}(e, e_1)}
 \tag{23}$$

Suppose $\text{Var}(e) = S_{11}$, $\text{Var}(e_1) = S_{22}$ and $\text{Cov}(ee_1) = S_{12}$.

The neural network model and the gray model are independent, therefore, $S_{12} = 0$,

Then

$$\begin{cases} c_1 = \frac{S_{22}}{S_{11} + S_{22}} \\ c_2 = \frac{S_{11}}{S_{11} + S_{22}} \end{cases}
 \tag{24}$$

The predicted value of the gray neural network can be obtained by substituting c_1 and c_2 into the corresponding calculation formula.

At the same time, in order to test the prediction accuracy of each model, the relative error is predicted by the relative error between the predicted value and the true value.

It is assumed that m is the relative error of the prediction model, S is the mean square error of the prediction error, Y_h is the predicted value at the h -th moment, and $x^{(0)}(h)$ is the actual value at the h -th moment, then

$$m = \frac{Y_h - x^{(0)}(h)}{x^{(0)}(h)} \quad S = \sqrt{\sum_{h=1}^n m_h^2 / (n-1)}
 \tag{25}$$

4. Simulation Experiment Prediction of Network Traffic Resources under Cloud Computing

Experimental method and parameters

- **Gray model prediction**

The gray traffic model was used to simulate the network traffic data resources. 12 sets of network traffic original data were detected by using relevant software on a website. Then, according to the network traffic data, the GM (1,1) model was built, and finally fitting analysis was performed. In this

way, 12 sets of data were predicted and compared with the actual values of 12 sets of data, and the experimental error of the gray prediction model is obtained. The resulting data was displayed in time (min) and flow (MB).

- **Prediction model of gray neural network**

The simulation experiment of network traffic data resources was carried out through the gray neural network model. Matlab8.0 was used as the programming application software. The network data traffic used by the master node router was the hourly access traffic data of the network from May 7 to July 15, 2018. For ten weeks, a total of $70 \times 24 = 1680$ data was obtained. Through the subtraction operator and the mean generator, a new sequence was generated from the original data.

The sequence of the network traffic time is $\{s(n), n = 1, 2, 3, \dots, 1680\}$, every week as 1d. In the first eight weeks, a total of $56 \times 24 = 1344$ data was obtained, and network traffic data was predicted in the last two weeks. The gray neural network model was used to predict the network traffic data for ten weeks, and then the original traffic data was compared and analyzed to obtain the experimental error value of the gray neural network prediction model. The obtained data was displayed in flow rate (MB).

Prediction analysis of gray model

Twelve sets of original data of network traffic were obtained by using relevant software on a website. The specific data is shown in Table 1.

Table 1. Network data traffic

Time (min)	Traffic (MB)
1	13.12
2	13.17
3	13.22
4	13.28
5	13.31
6	13.39
7	13.49
8	13.50
9	13.52
10	13.56
11	13.59
12	13.62

It could be seen from Table 1 that the actual values of the network traffic data were different in different time periods. The actual value of the network traffic in the first set was the lowest, and that of the network traffic in the 12th set reached the highest, and the overall trend was increasing.

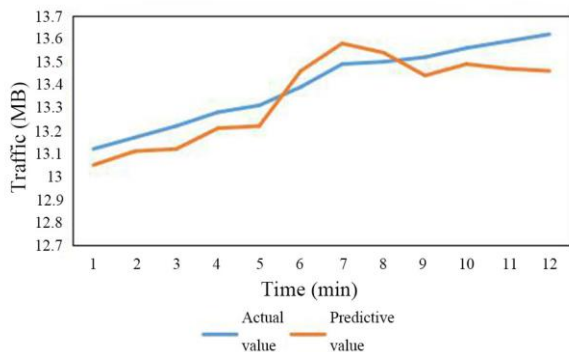


Figure 3: Fitting of predictive value of gray model and actual value

It could be seen from Figure 3 that the actual value of network traffic was generally on the rise, and the predicted value showed an upward and downward trend. The network traffic data predicted by the gray model differed greatly from the actual value. The two trends were different, and the overall fitness was relatively lower. The data predicted by the gray model was significantly different from the actual value, thus, it is necessary to strengthen and improve the prediction performance of the gray network model.

Prediction analysis of gray neural network model

The Matlab 8.0 application software obtained the hourly access traffic data of the primary node network from May 7 to July 15, 2018.

The specific values are shown in Table 2.

Table 2. Original data of network traffic

Time (h)	Traffic (MB)
50	23.322
100	42.323
150	32.963
200	34.523
250	28.689
300	41.659

It can be seen from Table 2 that the actual value of the original traffic data was different in the time interval of the master node every 50 hours. In the 50th hour, the network traffic data was the lowest. In the 100th hour, the network traffic data was the highest, the overall trend of which was on the rise.

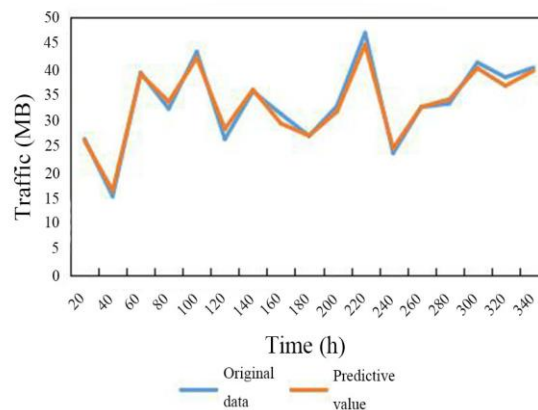


Figure 4: Simulation diagram of network traffic data

It could be seen from Figure 4 that the difference between the network traffic data predicted by the gray neural network model and the actual value was relatively small, the two trends were basically the same, and the overall fitness was relatively high.

The flow data predicted by the gray neural network model was more accurate, thus the prediction performance of the gray neural network model was relatively better.

Error comparison analysis

The data of each model was analyzed, and the error analysis results are obtained, as shown in Table 3.

Table 3. Error analysis table

	Error (e)	Relative error (m)	Variance (Var)
Gray model	1.9575	9.41%	0.752
Model of gray neural network	1.5896	6.32%	0.533

It could be seen from Table 3 that the error value of the prediction of the gray neural network model was 1.5896, the relative error was 6.32%, and the variance was 0.533. The error value of the prediction of the gray model was 1.9575, the relative error was 9.41%, and the variance was 0.752.

Compared to the predicted value of the gray model, the predicted value of the gray neural network model had small difference, a consistent trend and a high fitting degree with the actual value. Generally speaking, the data was more accurate. It could be seen that the gray neural network model had better predictive performance for network traffic data resources, and the experimental effect was more obvious.

5. Conclusion

With the rapid development of Internet technology, more and more people were engaged in networks, and various activities and services were carried out on the network, which were closely connected with the network and inseparable. It made the network traffic resources fully penetrate into the life and work of everyone. Among them, the prediction of network traffic data has become an important topic and research project. Studying the prediction of network traffic data can effectively improve the worrying network environment and have positive significance for network management and regulation [13].

In this study, the network traffic data resource prediction was researched, and a network traffic resource prediction algorithm based on gray neural network was proposed.

The simulation results showed that the error value of the gray neural network model for network traffic resource prediction was 1.5896, the relative error was 6.32%, and the variance was 0.533.

The predicted value was close to the actual value, the trend of the predicted value was similar to that of the actual value, and the fitting of the predicted value and actual value was high.

Therefore, it could be seen that the gray neural network model had better performance in predicting network traffic data compared with the gray model, which proved its feasibility and practicability to some extent. It started a new way for network traffic resource prediction and had broad development space.

References

- [1] Ali M, Khan S U, Vasilakos A V. "Security in cloud computing: Opportunities and challenges," *Information Sciences*, 2015, 305:357-383.
- [2] Chen W, Wilson J T, Tyree S, et al. "Compressing neural networks with the hashing trick," *Computer Science*, 2015:2285-2294.
- [3] Ganin Y, Ustinova E, Ajakan H, et al. "Domain-adversarial training of neural networks," *Journal of Machine Learning Research*, 2015, 17(1):2096-2030.
- [4] Conti M, Mancini L, Spolaor R, et al. "Analyzing android encrypted network traffic to identify user actions," *IEEE Transactions on Information Forensics and Security*, 2015, 11(1):1-1.
- [5] Aldhyani T H H, Joshi M R. "Clustering to enhance network traffic forecasting," 2018.
- [6] Khodabandelou G, Gauthier V, El-Yacoubi M, et al. "Population estimation from mobile network traffic metadata," *World of Wireless, Mobile & Multimedia Networks. IEEE*, 2016.
- [7] Bera S, Misra S, Rodrigues J J P C. "Cloud computing applications for smart grid: a survey," *IEEE Transactions on Parallel and Distributed Systems*, 2015, 26(5):1477-1494.
- [8] Jouini R, Lemlouma T, Maalaoui K, et al. "Employing Grey Model forecasting GM(1,1) to historical medical sensor data towards system preventive in smart home e-health for elderly person," *Wireless Communications & Mobile Computing Conference. IEEE*, 2016.
- [9] Wang S, Zhang N, Wu L, et al. "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method," *Renewable Energy*, 2016, 94:629-636.
- [10] Zhang Y, Gao X, Katayama S. "Weld appearance prediction with BP neural network improved by genetic algorithm during disk laser welding," *Journal of Manufacturing Systems*, 2015, 34:53-59.
- [11] Hu H, Fan L, Guan X. "Application on crude oil output forecasting based on gray neural network," *IEEE International Conference on Cloud Computing & Big Data Analysis. IEEE*, 2017.
- [12] Zhou G, Wang C, Wang S, et al. "Gray wavelet neural network and its application in mining waste prediction," 2016 *International Conference on Progress in Informatics and Computing (PIC)*, 2016.
- [13] Cejka T, Bartos V, Svepes M, et al. "NEMEA: A framework for network traffic analysis," *International Conference on Network & Service Management*, 2017.