IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 3, No. 4, December 2014, pp. 145~149
ISSN: 2252-8938

Prediction of Daily Network Traffic based on Radial Basis Function Neural Network

Haviluddin*, Imam Tahyudin**
*Department of Computer Science, Faculty of Mathematic and Natural Science, Mulawarman University, Indonesia
**Department of Information System, STMIK AMIKOM Purwokerto, Indonesia

Article Info

Article history:

Received Jul 2, 2014
Revised Nov 8, 2014
Accepted Nov 20, 2014

Keyword:

Network traffic
RBFNN
SSE
MSE
MAPE
MAD

ABSTRACT

This paper presents an approach for predicting daily network traffic using artificial neural networks (ANN), namely radial basis function neural network (RBFNN) method. The data is gained from 21-24 June 2013 (192 samples series data) in ICT Unit Universitas Mulawarman, East Kalimantan, Indonesia. The results of measurement are using statistical analysis, e.g. sum of square error (SSE), mean of square error (MSE), mean of percentage error (MPE), mean of absolute percentage error (MAPE), and mean of absolute deviation (MAD). The results show that values are the same, with different goals that have been set are 0.001, 0.002, and 0.003, and spread 200. The smallest MSE value indicates a good method for accuracy. Therefore, the RBFNN model illustrates the proposed best model to predict daily network traffic.

Corresponding Author:

Haviluddin,
Department of Computer Science,
Faculty of Mathematics and Natural Science,
Mulawarman University, Indonesia.
Email: haviluddin@unmul.ac.id

1. INTRODUCTION

The management of traffic quota is an important part, especially for the organizations that use information technology. Subsequently, for the leadership, management traffic quota will help in making decisions that will benefit the efficiency and effective for organizations including universities. The predicting activities is a part of organization management. Subsequently, the daily network traffic prediction is also a process of analyzing and determining the quota of bandwidth in a network in the future, in which a technical analysis approach usage data traffic. Furthermore, the predicting techniques used in the literature can be classified into two categories: statistical and soft-computing models. The statistical models include simple regression linear (SRL), exponential smoothing, the autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH) models. Nevertheless, these models are focused around the supposition that the data of various time series linearly correlate and provide poor prediction performance [1-4].

Meanwhile, the daily network traffic data are nonlinear and non-stationary in nature. To overcome this limitation, the second model is soft-computing methods have been suggested. Furthermore, modeling using the artificial neural network (ANN) model can provide better analytical results, and it is effective for forecasting, in which this method is able to work well on the non-linear time-series data [3, 5-7]. Therefore, this paper will study one of the ANN models, namely the Radial Basis Function Neural Network (RBFNN), in order to address the issue of network traffic time series data that has non-linear characteristics. This paper consists of four sections. Introduction section is the motivation to do the writing of the article. Next, the methodology is describing of model. Third section is the analysis and discussion results, and finally conclusion section is research summaries.

2. RESEARCH METHOD

The RBFNN is the abbreviation of radial basis function neural network which is based on the function approximation theory or supervised and unsupervised manner were used together. Subsequently, it has a unique training algorithm called hybrid method that emerged as a variant of NN in late 80's. This model is a kind of feed-forward neural network (FFNN) in which includes an input layer, a hidden layer, and an output layer [8]. As shown in Figure 1.

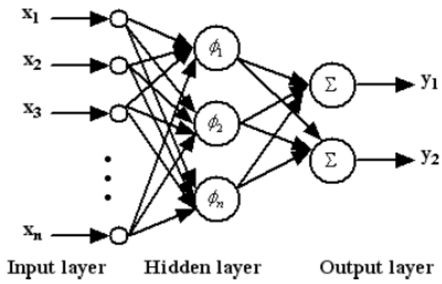


Figure 1. RBF neural network structure [9]

In general, RBFNN process the first phase is unsupervised learning between input layer and hidden layer that non-linear radial-based activation function, commonly Gaussian function. Second phase is supervised learning between hidden layer and output layer with a linear.

$$R(X^q - C_i) = exp[-(\|w1_i - X^q\| x b1_i)^2]$$

Where
$\|w1_i - X^q\|$ is the Euclidean distance, *c* is the center of Gaussian function

$X^q = \left(x_1^q, x_2^q, \dots, x_j^q, \dots, x_m^q\right)$ is the *qth* input data.

Hence, the architecture of RBFNN as shown in Figure 1, and the equation is

$$Y = \sum_{j=1}^{m} W_{jm}.\varphi$$

where:
Y = output value; φ = hidden value; W = weights (0-1)

The algorithm of RBFNN to analyze within time series data characteristics is:
1. Initialization of the network.
2. Determining the input signal to hidden layer, and find D_{ij} is a distance data i to j where $i, j = 1,2, \dots, Q$

$$D_{ij} = \sqrt{\sum_{k=1}^{R} (p_{ik} - p_{jk})^2}$$

3. Find $a1$ is a result activation from distance data multiply bias.

$$a1_{ij} = e^{-(b1*D_{ij})^2} \text{ x } b1 = \frac{\sqrt{-\ln(0.5)}}{spread}$$

4. Find weight and bias layers, $w2_k$ and $b2_k$, in each $k = 1,2, \dots, S$.

a. Determining training samples and test samples.

In this study, the data were collected from ICT server Universitas Mulawarman. Then, the data were collected from 21-24 June 2014 (192 samples series data). Then, each network traffic data was captured by the CACTI software. The daily network traffic data was analyzed using MATLAB R2013b. Since the implicit function of RBFNN is Gaussian function, in which general requires for input value between 0 and 1. The daily network traffic data need normalized using statistical data normalization, which is usually expressed as:

$$\bar{X} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Where: X is the actual value of sample; X_{max} takes a large value, and X_{min} takes a samples of data is less than the minimum value to ensure normalized value is not close to 0. Later, process inverse transform to get the actual value is obtained. Furthermore, in this experiment we used the sum of square error (SSE), mean of square error (MSE), mean of absolute percentage error (MAPE), and mean of absolute deviation (MAD) were engaged the predicted output with the desired output.

3. RESULTS AND ANALYSIS

This section presents the best achieved result by the RBFNN algorithm. In this experiment, the training data were selected from 21-24 June 2014 (192 samples series data) was captured. Based on NNs rules, the data were divided into training and testing data then, the datasets consist of 144 (90%) samples for data training and 48 (10%) samples for data testing or two neurons, P=[p(t-1), p(t)], and the number of output neurons is one, p'(t), as shown in Table 1. In order to test and validate the different error goals, four statistical; SSE, MSE, MAPE, and MAD test were carried out. The results obtained are summarized in Table 2A and 2B. From the simulations carried out, it was created a precise neural network by *newrb (P,T,error_goal,spread)* function, which is this function creates RBFNN structure, automatically selected the number of hidden layer and made the error to 0. For the error goal values were 0.01, 0.02, and 0.03. Where, *spread* is the density of basis function value of 200 was settled.

The experiment has shown values are the same, with different goals were 0.001, 0.002, and 0.003 have been set. The values in the training RBFNN are SSE value is 0986, MSE value is 0.007, MAPE value is 0.043, and MAD value is 0063 then, the values of the RBFNN testing are SSE value is 0343, MSE value is 0.007, MAPE value is 0.036, and MAD value is 0.064.

Table 1. The Daily Network Data After Normalized

Group		Input neurons P=[p(t-1),p(t)]		Output neuron T= p'(t)	Group		Input neurons P=[p(t-1),p(t)]		Output neuron T= p'(t)
		t-1	t	t+1			t-1	t	t+1
Group Train	1	0.20	0.16	0.17	Group Test	142	0.27	0.23	0.16
	2	0.16	0.17	0.12		143	0.14	0.11	0.19
	3	0.17	0.12	0.07		144	0.11	0.19	0.18
	4	0.12	0.07	0.05		145	0.19	0.18	0.14
	5	0.07	0.05	0.08		146	0.18	0.14	0.08
	…	…	…	…		…	…	…	…
	…	…	…	…		…	…	…	…
	…	…	…	…		…	…	…	…
	…	…	…	…		…	…	…	…
	…	…	…	…		…	…	…	…
	137	0.41	0.42	0.45		184	0.22	0.32	0.31
	138	0.42	0.45	0.48		185	0.32	0.31	0.27
	139	0.45	0.48	0.43		186	0.31	0.27	0.32
	140	0.48	0.43	0.27		187	0.27	0.32	0.20
	141	0.43	0.27	0.23		188	0.32	0.20	0.19

Table 2A. The RBFNN Training Results

RBFNN	Training			
	SSE	MSE	MAPE	MAD
error goal = 0.001				
SPREAD 200	0.986	0.007	0.043	0.063
error goal = 0.002				
SPREAD 200	0.986	0.007	0.043	0.063
error goal = 0.003				
SPREAD 200	0.986	0.007	0.043	0.063

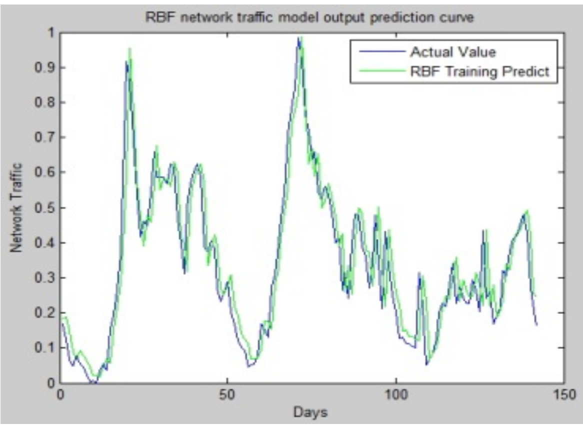
Table 2B. The RBFNN Testing Results

RBFNN	Testing			
	SSE	MSE	MAPE	MAD
error goal = 0.001 SPREAD 200	0.343	0.007	0.036	0.064
error goal = 0.002 SPREAD 200	0.343	0.007	0.036	0.064
error goal = 0.003 SPREAD 200	0.343	0.007	0.036	0.064

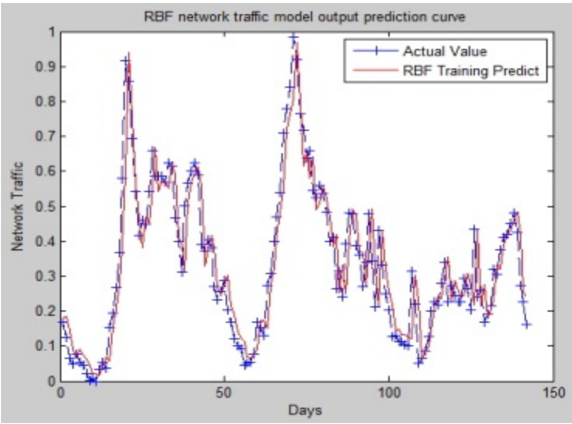
4. CONCLUSION

In this paper, the analysis using RBFNN technique to achieve the model of daily network traffic activities have been conducted in the ICT Unit, Universitas Mulawarman. According to Figure 2 and 3, the results of RBFNN training shows that for error goal value is 0.001 then SSE value is 0986, MSE value is 0.007, MAPE value is 0.043, and MAD value is 0063 Afterward, the values of the RBFNN testing are SSE value is 0343, MSE value is 0.007, MAPE value is 0.036, and MAD value is 0.064. Indicator test result of data is the smallest error value, where value indicating an error testing is the best model [9].

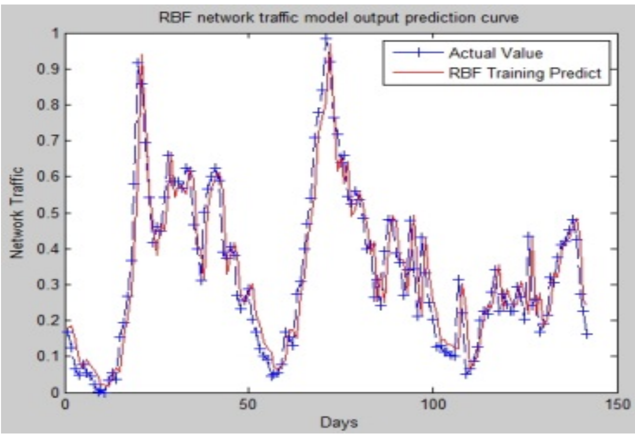
Therefore, the determination of the best model is determined by selecting the smallest value of testing error. Based on the results of RBFNN are considered closer to the actual value. In other words, the RBFNN model with different spread values illustrates the proposed best model to predict daily network traffic activities.



(a) error goal 0.001

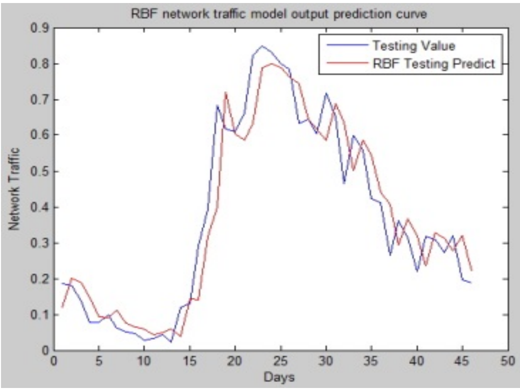


(b) error goal 0.002

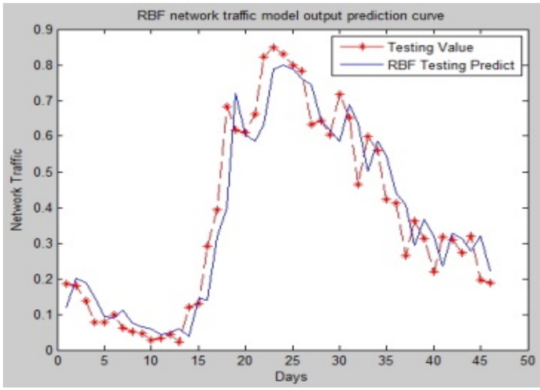


(c) error goal 0.003

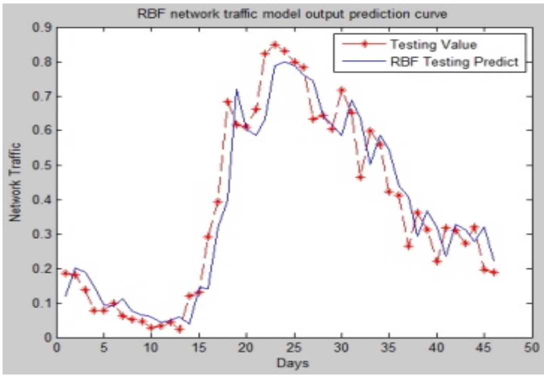
Figure 2. The RBFNN training results curves



(a) error goal 0.001



(b) error goal 0.002



(c) error goal 0.003

Figure 3. The RBFNN testing results curves

ACKNOWLEDGEMENTS

This study has been completed thanks to the help and support from various parties that cannot be mentioned one by one. Especially thank you to our family, children and wife. Researchers say a big thank you to family of Mulawarman University and STMIK AMIKOM Purwokerto who has given support to complete this study. Hopefully this research can be useful.

REFERENCES

[1] Haviluddin, and R. Alfred, "Forecasting Network Activities Using ARIMA Method," *Journal of Advances in Computer Networks,* vol. 2, no. 3, September 2014, pp. 173-179, 2014.
[2] B. Majhi, M. Rout, and V. Baghel, "On the development and performance evaluation of a multiobjective GA-based RBF adaptive model for the prediction of stock indices," *Journal of King Saud University-Computer and Information Sciences*, 2014.
[3] Y. Birdi, T. Aurora, and P. Arora, "Study of Artificial Neural Networks and Neural Implants," *International Journal on Recent and Innovation Trends in Computing and Communication,* vol. 1, no. 4, 2013.
[4] U. Yolcu, E. Egrioglu, and C. H. Aladag, "A new linear & nonlinear artificial neural network model for time series forecasting," *Decision Support Systems,* vol. 54, no. 2013, pp. 1340-1347, 2013.
[5] O. Claveria, and S. Torra, "Forecasting tourism demand to Catalonia: Neural networks vs. time series models," *Economic Modelling,* vol. 36, no. 2014, pp. 220-228, 2014.
[6] J.-Z. Wang, J.-J. Wang, Z.-G. Zhang *et al.*, "Forecasting stock indices with back propagation neural network," *Expert Systems with Applications,* vol. 38, no. 2011, pp. 14346-14355, 2011.
[7] G. Chen, K. Fu, Z. Liang *et al.*, "The genetic algorithm based back propagation neural network for MMP prediction in CO_2-EOR process," *Fuel,* vol. 126, no. 2014, pp. 202-212, 2014.
[8] W. Jia, D. Zhao, T. Shen *et al.*, "A New Optimized GA-RBF Neural Network Algorithm," *Computational Intelligence and Neuroscience*, no. 2014, pp. 1-6, 2014.
[9] J. Wu-Yu, and J. Yu, "Rainfall time series forecasting based on Modular RBF Neural Network model coupled with SSA and PLS," *Journal of Theoretical and Applied Computer Science,* vol. 6, no. 2, pp. 3-12, 2012.